The INSTITUTE AND FACULTY OF ACTUARIES

SA0 Project

Estimation and projection of cancer mortality in England

Author:
Dr. Ayşe Arik

A project submitted in fulfilment of the requirements to become a fellow actuary

The Institute and Faculty of Actuaries, UK

March 2025

Summary

This study investigates cancer mortality rates based on the population data of England between 2001 and 2018 using a Bayesian hierarchical modelling setting.

Major cancer types, to be specific breast cancer (BC) and lung cancer (LC) mortality, are examined with a focus on projection of cancer mortality in the future. We investigate patterns in type-specific cancer mortality by year of death and various risk factors: age, gender, regions of England, income deprivation quintile, average age-at-diagnosis (AAD), and non-smoker (NS) prevalence rates. We analyse patterns in 2001–2018 (as baseline) for each cancer type and project these in subsequent years.

We then assess the impact of diagnosis delays on cancer mortality, associated with national lockdowns that were introduced as a result of the COVID-19 pandemic.

The dataset

The cause of deaths data underlying the analysis is provided by the Office for National Statistics (ONS) in the UK. The data is anonymised and stratified by five-year age groups (from 20 to 89), gender, single calendar year (2001–2018), nine regions of England, and income deprivation deciles (1–10). The raw data used in this study cannot be released for the purposes of maintaining data protection and confidentiality based on our agreement with the data provider. Yet, the data can be accessible by contacting the ONS.

We also utilise publicly available data, including age-specific NS prevalence between 1993–2019 from the Health Survey of England, and age-specific cancer mortality by regions of England up to 2022 from the ONS.

Principal findings

We assess BC and LC mortality by various risk factors over time, and we project mortality from these cancer types by deprivation and region (where relevant). Our results show:

- BC and LC mortality vary by age, year, gender, and region.
- Both AAD and deprivation are significant variables for explaining changes in LC mortality over time.
- NS prevalence, used as a proxy for smoking, is found to be a significant factor for both BC and LC mortality.
- There are significant socio-economic differences in LC mortality, and these differences persist over time.

- There are marginally significant regional differences in BC mortality, and this remains relevant to future years.
- There are notable differences in BC mortality across different age groups, including those associated with screening ages, in future years.
- The AAD variable can be used to construct COVID-19 pandemic-related scenarios.
- For a 1-month diagnosis delay, our models have estimated 2,340 (95% CI 1,743 to 2,869) excess deaths for women and 5,164 (4,353 to 6,066) for men. When a 6-month delay is considered, our models suggest 10,180 (7,944 to 12,340) and 28,660 (23,040 to 35,090) excess deaths for women and men respectively.
- The excess LC deaths, as a result of delays in cancer diagnosis, significantly differ by age, region and deprivation. Particularly, the excess mortality for both men and women is found to be significantly
  - higher at older age groups including 60–64 years old;
  - higher in the northern regions of England compared to the southern regions; and
  - higher for those living in the most deprived quintiles compared to those in the least deprived quintiles.

Contents

Abbreviations

APS	**A**nnual **P**opulation **S**urvey
AAD	**A**verage **A**ge-at-**D**iagnosis
BC	**B**reast **C**ancer
CRUK	**C**ancer **R**esearck UK
CII	**C**ritical **I**llness Insurance
ESP	**E**uropean **S**tandard **P**opulation
IMD	**I**ndex of **M**ultiple **D**eprivation
LC	**L**ung **C**ancer
NHS	**N**ational **H**ealth **S**ystem UK
NS	**N**on-**S**moker
ONS	**O**ffice for **N**ational **S**tatistics
PHS	**P**ublic **H**ealth **S**cotland
STZ	**S**um-**T**o-**Z**ero
UK	**U**nited **K**ingdom

If everything around seems dark, look again, you may be the light (Rumi).

Chapter 1

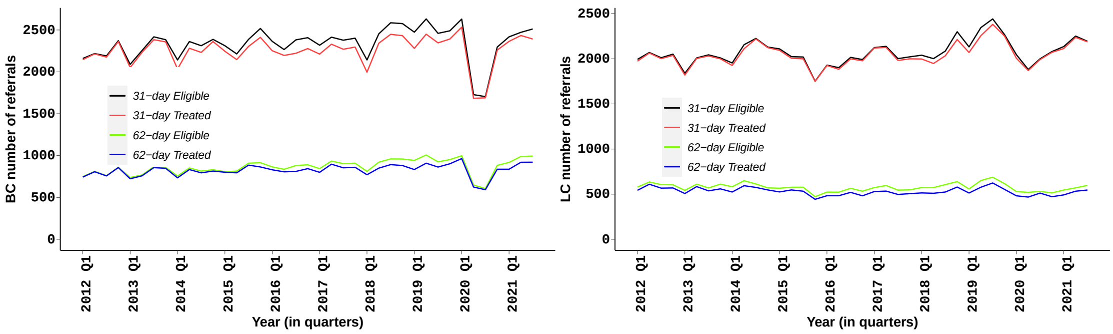
Introduction

1.1 Background

Cancer remains one of the major causes of mortality and morbidity in England, representing 27–28% of all deaths per year [ONS, 2020b]. Cancer is also considered to be the largest driving cause of 'avoidable' mortality, while the rate of avoidable mortality in the United Kingdom (UK) has significantly increased in 2020 as compared to 2010 [ONS, 2022a]. Furthermore, cancer is one of the core causes in critical illness insurance (CII), which is an insurance policy paying a benefit on the occurrence of a serious illness, and often including an accelerated death benefit [Macdonald et al., 2003]. Importantly, cancer in CII contracts, along with heart attack and stroke, account for the largest percentage of claims. The proportion of cancer claims has been reported to be as high as 54% in 2002 in the UK suggesting higher risk for females. See Kimball [2002] for a detailed discussion.

Cancer has attracted more attention due to the global COVID-19 pandemic. The pandemic was first identified in Wuhan, China in December 2019 and then rapidly spread to other parts of world in 2020, by claiming more than 6.5 million lives worldwide as of November 2022 [WHO, 2022]. As reported by Baker [2021], a large number of people, around 12.3% of all deaths, died in England between 2020 and 2021 because of the COVID-19 pandemic. Thus, cancer accounted for a lower percentage of all deaths in those years, to be specific 24% of all deaths [ONS, 2022b]. As a response to the pandemic, the UK entered three national lockdowns, with the first being introduced on 23 March, 2020. Meanwhile, cancer pathways have been seriously affected by changes in health practices due to a halt in cancer screening (from late March 2020 till June 2020), significant increases in the number of patients waiting for key diagnostic tests for more than 6 weeks, and significant reductions in the number of patients starting cancer treatment. Cancer Research UK (CRUK) reported that 3 million fewer people were screened for cancer in the UK between March and September 2020 compared to the same period in 2019 [CRUK, 2021a]. The effect on cancer pathways goes beyond 2020 due to re-occurring national lockdowns. For instance, 590,000 fewer people participated in breast cancer (BC) screening between April 2020 and March 2021, 33% lower in comparison to the pre-pandemic levels, 2018–2019. Moreover, the number of cancer patients starting a cancer treatment decreased by 12% between April 2020 and March 2021 compared to the pre-pandemic levels, whereas the number of people waiting for more than 6 weeks for key diagnostic tests soared to 215,000 in March 2021 from 67,000 in March 2020. See CRUK [2022b] for further statistics.

Additionally, Public Health Scotland (PHS) reported a sharp decline in type-specific cancer referrals, which demonstrate national standards with respect to how long cancer patients wait for their first cancer treatment in Scotland [PHS, 2021]. Figure 1.1 shows both 31- and 62-day breast and lung cancer (LC) referrals until the period of July to September 2021 along with the previous quarters to allow comparison to the period before the COVID-19 pandemic. In Figure 1.1, cancer referrals manifest a decrease after the first quarter of 2021, particularly in 31-day referrals. However, we see a sharper decline in BC referrals, including 62-day referrals. As pointed by PHS [2022], this could be linked to the availability of national BC screening programme where women aged 50–70 are eligible for screening once every three years in Scotland. Note that the BC screening programme in England currently targets all women between 47 to 73 years old [CRUK, 2022a].



(a) Breast cancer referrals

(b) Lung cancer referrals

Figure 1.1: 31- and 62-day eligible and treated breast and lung cancer referrals in Scotland between 2012 and 2021. Source: Public Health Scotland.

All these changes in cancer pathways sparked the fear of a shift to later diagnosis for people having the disease but not diagnosed yet. This is considered to be a serious concern since a late cancer diagnosis could restrict the opportunities for feasible treatment and worsen cancer survival.

Meanwhile, early empirical studies suggested that the COVID-19 pandemic has disproportionately affected certain groups, such as the elderly, people with comorbidities or people who are more deprived. For a broader discussion, see Chen et al. [2020], CRUK [2020], Grasselli et al. [2020], Richardson et al. [2020], Zhou et al. [2020]. One potential implication of this is to exacerbate socio-economic inequalities in cancer risk, which has been a staggering issue, mostly getting worse rather than better in several countries including the UK. For a more in-depth discussion, see Arık et al. [2021], Bennett et al. [2018], Brown et al. [2007], Mouw et al. [2008], Riaz et al. [2011], Singh et al. [2011].

Most of the recent published studies have focussed on identifying the impact of various changes in availability of cancer treatment and services in addition to health-seeking behaviour, as a result of national lockdowns, on cancer survival in England based on the National Health System (NHS) UK cancer registration and hospital administrative dataset. Lai et al. [2020] point out dramatic reductions in the demand for, and supply of, cancer services in response to the COVID-19 pandemic by showing that these reductions could largely contribute to excess mortality among cancer patients. Sud et al. [2020] indicate a significant reduction in cancer survival as a result of treatment delay, mostly disruption in cancer surgery, in England. Maringe et al. [2020] also note substantial increases in avoidable cancer deaths in England as a result of diagnostic delays over a year on. Arık et al. [2021] report significant increases in type-specific cancer mortality as a result of diagnostic delays based on a population-based study in England. Arık et al. [2023a] further point out medium to large size increases in BC mortality from aged 65 and above based on a modelling study calibrated with respect to available population data of England and medical literature.

1.2 Aim of this study

In this study we are particularly interested in malignant neoplasm of trachea, bronchus, and lung, and malignant neoplasm of BC. The reason why we choose these two cancer types is because they still represent the largest percentage of overall cancer deaths, and they are among the leading causes of death in the UK. It is also worth noting that LC became the main leading cause of death for women aged 50 to 64 years old in 2008 by replacing BC, that accounted for 10.1% of overall deaths among these age groups in 2018 [ONS, 2020c].

To be more precise, we have two main interests: (a) providing a deeper insight into LC and BC mortality; and (b) understanding the impact of diagnosis delays on future cancer mortality. Part of the contribution of this study is providing a modelling framework in order to project LC and BC mortality on regional and deprivation level, where appropriate. We develop type- and gender-specific Bayesian hierarchical models to project cancer mortality, together with 95% credible intervals, where we use a Poisson distribution assumption for cancer deaths. For similar applications of the Poisson distribution, see Arık et al. [2021], Czado et al. [2005], Wong et al. [2018]. We carry out the considered models based on cancer deaths and mid-year population estimates, stratified by five-year age groups, single year, deprivation quintiles in regions of England between 2001 and 2018, provided by the Office for National Statistics (ONS). Our approach presents a detailed modelling structure for each cancer type accounting for various risk factors by avoiding coalition between different variables, such as region and deprivation level. We also incorporate smoking information into our modelling as an additional risk factor. Moreover, we consider an estimate of age-at-diagnosis as a risk factor for a given cancer mortality modelling. The estimate of age-at-diagnosis is based on the related cancer morbidity, as developed and discussed in Arık et al. [2021], and it is used to provide a proxy to delays in diagnosis while examining the impact of COVID-19 related health disruptions.

This study is organised as follows. In Chapter 2 we introduce the available data and important concepts used in different parts of the report. In Chapter 3 we explain the modelling framework for LC and BC mortality rates. In Chapter 4 we carry out the developed models with respect to the population LC and BC mortality in England and present main findings, accordingly. In Chapter 5 we discuss the main implications of our findings and conclude.

Chapter 2

Data and Concepts

This chapter summarises all data used in this study. Some relevant metrics supporting this research are also introduced. In general, we rely on a set of cancer data provided by the ONS along with some publicly available data. More details are provided below.

2.1 Number of cancer deaths and population exposures

We have type- and gender-specific cancer deaths and mid-year population estimates exposed to mortality risk, aka 'exposures', at five-year age bands in different regions (specifically the north east, the north west, Yorkshire and the Humber, the East Midlands, the West Midlands, the east, London, the south east, and the south west according to Nomenclature of Territorial Units for Statistics described by Eurostat [2007]) and deprivation deciles (1 to 10) of England in single years from 2001 to 2018. The data were provided by the ONS. Although the data are anonymised, it cannot be released openly due to the level of granularity in order to maintain data protection and confidentiality. Currently, cancer death numbers and population estimates can be obtained with a data request from the ONS.

We also note that causes of death data is accessible at a lower granularity up to 2022 through a service, namely 'NOMIS', provided by the ONS [ONS, 2022d]. Specifically, we use 'NOMIS' as an additional source to obtain BC deaths in women by five-year age bands in the regions of England from 2019 to 2022.

For each region r and deprivation level d of England, whenever applicable, we have death and exposure counts as follows:

- $D_{a,g,d,r,t}$ is number of deaths from cancer at age-at-death a in year t for gender g, where the youngest age group is 20–24 and the oldest age group is 85–89;
- $E_{a,g,d,r,t}$ is number of exposures, i.e. the corresponding mid-year population estimates;
- $\theta^c_{a,g,d,r,t}$ is crude mortality rates, that can be calculated as $\theta^c_{a,g,d,r,t} = \frac{D_{a,g,d,r,t}}{E_{a,g,d,r,t}}$.

2.2 Index of Multiple Deprivation

The index of multiple deprivation (IMD) is a national deprivation index, dividing small areas in England into 10 equal groups. This index ranks these areas from the most deprived 10% of small areas nationally to the least deprived 10%, where 1 is labelled to be the most deprived and 10 to be the least deprived. The IMD can be used to make direct comparisons in relation to the relative deprivation across different areas in England. For further discussion, see DCLG [2015].

According to the technical report by Gill [2015], the IMD is derived as a weighted average of 7 different sub-indices as follows:

- Income deprivation (weight 22.5%)
- Employment deprivation (22.5%)
- Education, skills and training deprivation (13.5%)
- Health deprivation and disability (13.5%)
- Crime (9.3%)
- Barriers to housing and services (9.3%)
- Living environment deprivation (9.3%)

The seven IMD sub-indices listed above measure different aspects of deprivation, and also these indices are related to mortality. Hereby, the IMD can be used as a predictive variable for estimating and projecting mortality in a given area r.

This study is based on the 'income deprivation' groups under the IMD published in 2015, namely 'IMD 2015'. Importantly, we group income deprivation (deprivation from now on) deciles into quintiles for modelling purposes, where each quintile refers 20% of the corresponding sub-population in a given region.

For completeness of information, we note that there are earlier and later versions of IMD over time. See Smith et al. [2015a] and Smith et al. [2015b] for further details, including the methodology used for constructing the IMD, the type of data used, and so on.

2.3 Data on smoking

There are two sets of publicly available smoking data. First smoking data is collected on the Labour Force Survey, as part of the Annual Population Survey (APS), based on an annual sample size around 320,000. In this data, prevalence numbers (%) are available for current smokers, ex-smokers, and non-smokers by age groups, specifically 18–24, 25–34, 35–44, ..., 65+, in the nine regions of England between 2011 and 2021. For further details, see ONS [2020a], and to access the data, refer to ONS [2022c]. As part of the APS, smoking prevalence by the IMD started being reported since 2012, as well [Archbold et al., 2023]. Figure 2.1 displays that smoking prevalence has declined in each deprivation decile since 2012 with statistically higher smoking prevalence in the most deprived neighbourhoods as compared to the least deprived ones in England.

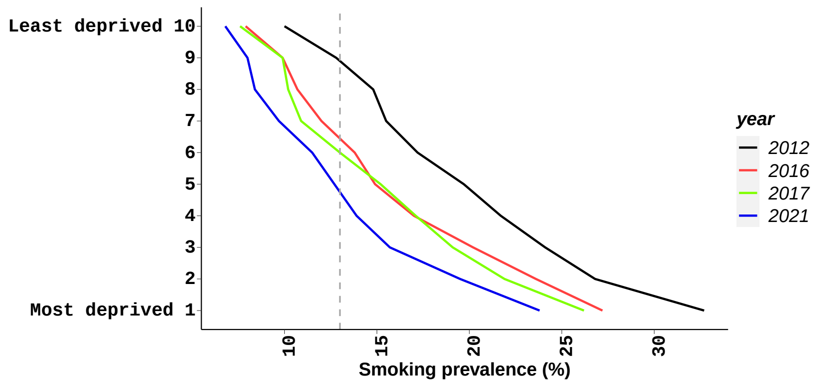


Figure 2.1: Prevalence rates of current smokers by deprivation deciles, all persons aged 18 years and over, England, 2012 to 2021, where the dashed line is the national average, 13%, in 2021. Source: Annual Population Survey from the Office for National Statistics.

Second smoking data is collected as part of the Health Survey for England. This data, available from the NHS Digital, shows age-specific prevalence rates for wider age groups, including ages 65–74, and 75+, between 1993 and 2019. However, unlike the first dataset, it does not include regional breakdowns within England. For more details, see Digital [2020].

Since the two datasets are not compatible, we use a single source in this study. In particular, we prefer the Health Survey for England dataset because it covers a longer time period. See Section 2.3.1 for further discussion.

2.3.1 Non-smoking prevalence rates as a proxy for smoking

Cigarette smoking remains to be the greatest cause of preventable death and disease globally [Archbold et al., 2023]. Smoking is also considered to be a significant risk factor for certain diseases, including LC, as noted in NHS [2022]. Nonetheless, researchers are cautioned using smoking data most appropriate for specific research questions [Ryan et al., 2012]. This is mainly because differences in smoking definitions across different surveys could have a notable impact on smoking estimates. An immediate example to this is a change in survey questions in the APS impacting the calculation of ex-smoker estimates from 2016. For a broader discussion, see Windsor-Shellard et al. [2018].

In this study non-smoker (NS) prevalence in the Health Survey of England is considered as a proxy for smoking in LC models. This decision serves two purposes: (a) avoiding changes across different definitions of smoking information, and (b) having a simple and clear interpretation of smoking in the projection models.

Note that we attempted to distribute NS prevalence by deprivation using Figure 2.1 and by regions of England based on the APS data. Yet, these assumptions caused correlation issues in the existence of region and deprivation variables in the projection models. Hereby, NS prevalence is assumed to be the same in all regions and deprivation quintiles of England.

Meanwhile, age-specific fitted NS prevalence rates have been obtained based on the data between 1993 and 2019 by using a simple linear model as follows:

$$\mathrm{NS}_{a,t} = \beta_0 + \beta_{1,a} + \beta_2 t + \beta_3 t^2 + \beta_{4,a} t. \tag{2.1}$$

Afterwards, the parameter estimates in (2.1) are employed to re-construct NS prevalence backwards to 1981. These rates are involved in LC and BC modelling in a way to accommodate a lag of 20 years by following the study of Luo et al. [2022]. This means, for example, we use NS prevalence in 1981 as an input to estimate LC mortality in 2001.

Figure 2.2 demonstrates fitted and crude age-specific NS prevalence rates from 1981 to 2019. The figure reveals an increasing trend in NS prevalence for both genders, notably showing a more homogeneous and faster increase among males over this period.

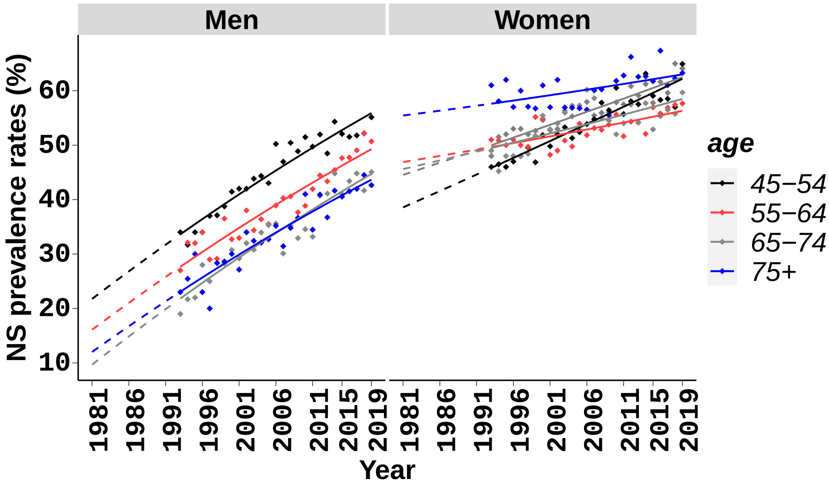


Figure 2.2: Non-smoker prevalence rates at selected age groups between 1981 and 2019: observed rates (dots), fitted rates (lines), and re-constructed rates (dashed lines).

2.4 Age-standardised mortality rate

Age-standardisation is a useful tool to compare different sub-populations by eliminating variations raised by different ages and corresponding population sizes. We obtain age-standardised mortality rates by gender to make comparisons across different deprivation levels and regions of England. This calculation is done by choosing distribution of a population over ages as the reference population. For this sake, we use the European Standard Population (ESP) 2013 [Eurostat, 2013].

Age-standardised mortality rate for gender g in a given year t, region r and deprivation quintile d in England, denoted by $\text{ASR}_{g,d,r,t}$, is defined based on the weighted average of crude mortality rates, $\theta^c_{a,g,d,r,t}$, over a specific age range as follows:

$$\text{ASR}_{g,d,r,t} = \frac{\sum_a \theta^c_{a,g,d,r,t} E_a^{\text{std}}}{\sum_a E_a^{\text{std}}}, \tag{2.2}$$

where E_a^{std} denotes the 'standard population' at age a. In this study, the relevant age range is determined depending on the modelling age group for each cancer type under consideration where the underlying age group is 45–89 years old for LC mortality and 35–89 years old for BC mortality.

2.4.1 Mortality differences over sub-populations

Inequalities in cancer mortality need to be addressed while considering future projections. Age-standardised rates by region and deprivation can be a useful tool to understand trends in mortality inequalities. Figure 2.3 and Figure 2.4 respectively demonstrate how LC and BC mortality rates have realised across different deprivation levels and regions of England based on crude age-standardised rates from 2001 onwards.

Specifically, Figure 2.3 shows that there are distinct differences in age-standardised observed LC mortality rates, for instance, in women across different deprivation quintiles in the regions of England. The figure clearly demonstrates higher mortality rates in more deprived quintiles over the years. We also see a widening deprivation gap in certain regions, such as the north east of England, as compared to others, such as London.

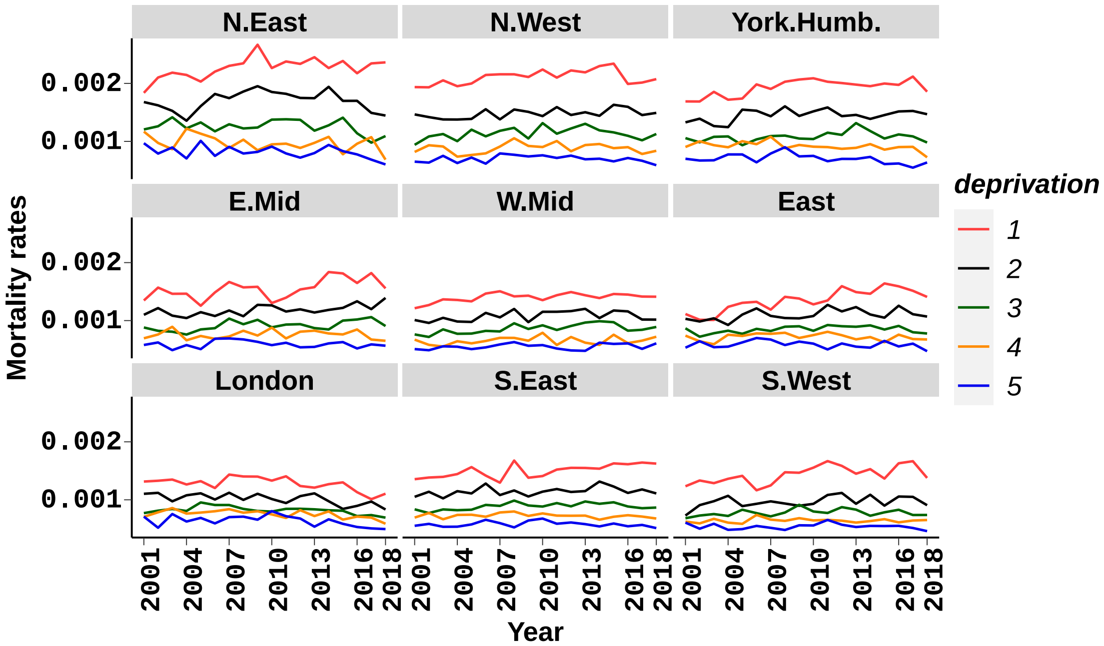


Figure 2.3: Age-standardised observed lung cancer mortality rates in women in deprivation quintiles 1 (most deprived) to 5 (least deprived) of England: the standardisation is done over ages 45–89.

Figure 2.4 demonstrates regional variability in age-standardised observed BC mortality rates in women between 2001 and 2022. We observe much smaller variation in BC mortality across the regions of England as compared to LC mortality. We note that different than LC mortality, deprivation is not found to be a significant variable to explain variations in BC mortality. See, for example, Arık et al. [2021] for further discussion. Hereby, in this study, deprivation is not considered as a separate risk factor while predicting future BC mortality.

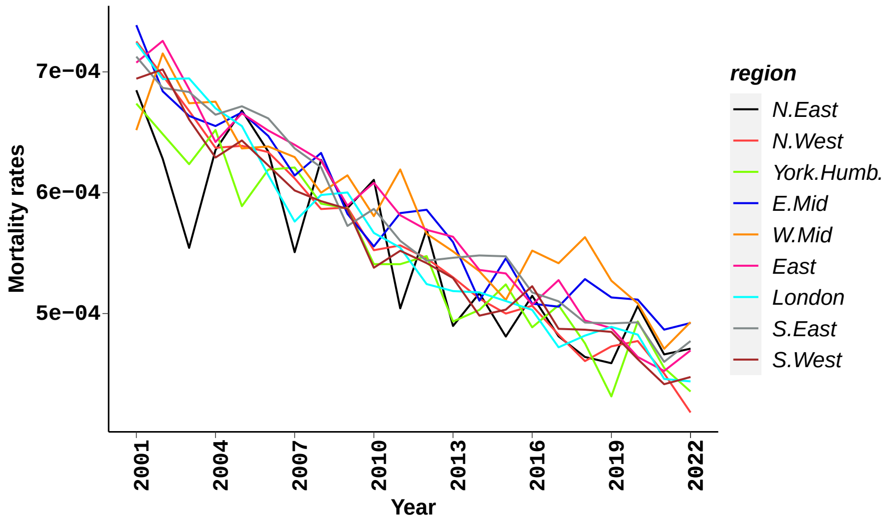


Figure 2.4: Age-standardised observed breast cancer mortality rates in regions of England: the standardisation is done over ages 35–89.

2.5 A useful metric: relative deprivation gap

It is important to quantify socio-economic differences in cancer mortality over time, with the aim of understanding how these differences have changed. We quantify relative deprivation gap in a given cancer type, denoted by $\mathrm{RD}_{g,r,t}$, for gender g in region r of England at time t such that

$$\mathrm{RD}_{g,r,t} = \frac{\mathrm{A\hat{S}R}_{g,\text{quintile } 1,r,t} - \mathrm{A\hat{S}R}_{g,\text{quintile } 5,r,t}}{\mathrm{A\hat{S}R}_{g,\text{quintile } 1,r,t}}, \tag{2.3}$$

where fitted age-standardised mortality rates in the most (1) and least (5) deprived quintiles are used as an input. Note that the numerator in this formula, which is

$$\mathrm{AD}_{g,r,t} = \mathrm{A\hat{S}R}_{g,\text{quintile } 1,r,t} - \mathrm{A\hat{S}R}_{g,\text{quintile } 5,r,t},$$

shows the absolute deprivation gap, $\mathrm{AD}_{g,r,t}$, in region r at time t. In (2.3), the change in the deprivation gap in a given region at time t is expressed with respect to the mortality rate in the most deprived quintile of the same region.

2.6 A useful variable: average age-at-diagnosis

Age-at-diagnosis is known to be a crucial factor for cancer survival. In the study of Arık et al. [2021], a link has been established between cancer morbidity and mortality through a variable, namely average age-at-diagnosis (AAD). In that study, it is shown that AAD is a statistically significant variable to explain type-specific cancer mortality. Hereby, this variable is also used as an input in our projection models after verifying its significance through appropriate variable selection procedures for the data under inspection (see Appendix A for details).

AAD for gender g in deprivation quintile d of region r at the time of diagnosis t, denoted by $\mathrm{AAD}^{\mathrm{morbidity}}_{g,d,r,t}$, is estimated as follows:

$$\mathrm{AAD}^{\mathrm{morbidity}}_{g,d,r,t} = \frac{\sum_a a\hat{\lambda}_{a,g,d,r,t} E^{\mathrm{std}}_a}{\sum_a \hat{\lambda}_{a,g,d,r,t} E^{\mathrm{std}}_a}, \tag{2.4}$$

where E^{std}_a shows population numbers at age-at-diagnosis a according to the ESP 2013, and $\hat{\lambda}_{a,g,d,r,t}$ is the relevant type-specific fitted incidence rate obtained based on the best fitted models in Arık et al. [2021]. For modelling purposes, AAD is then weighted over years as described below:

$$\mathrm{AAD}^{\mathrm{morbidity}}_{g,d,r} = \frac{\sum_t \mathrm{AAD}^{\mathrm{morbidity}}_{g,d,r,t} E_{g,d,r,t}}{\sum_t E_{g,d,r,t}}, \tag{2.5}$$

by using the relevant mid-year population estimates $E_{g,d,r,t}$ in deprivation quintile d of region r. Note that, if deprivation is not a significant variable in the model under inspection, AAD could also be averaged over deprivation quintiles so that it would be by region only.

In this study, AAD variable is found to be significant to explain differences in LC mortality but not in BC mortality (refer to Appendix A for further details). Figure 2.5 shows estimated AAD values in LC for women according to 2.4 from 2001 to 2017. We note that 2017 is the latest available calendar year in the LC morbidity data. We estimate an increasing trend in AAD values over the calendar years, with comparable results in the regions of England. However, the estimates across deprivation quintiles in a given region are notably different, where lower AAD values are calculated in more deprived quintiles (as opposed to higher mortality rates in the same quintiles, see Figure 2.3).

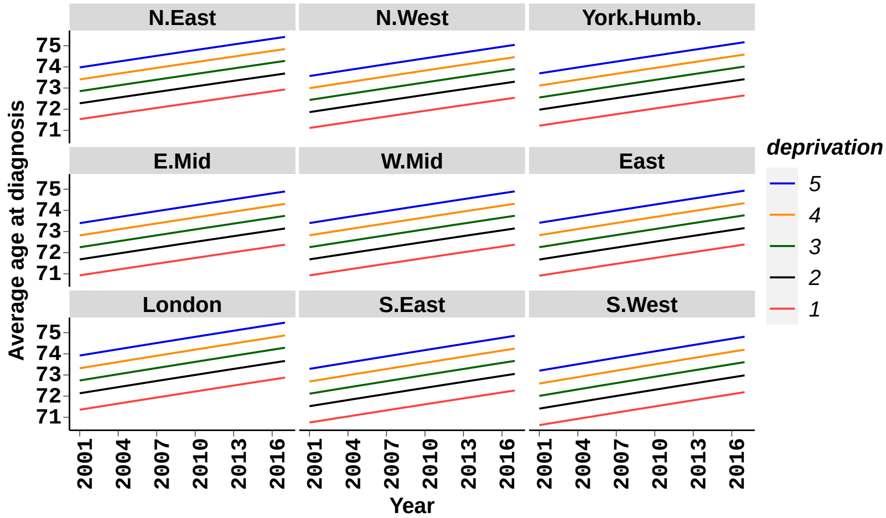


Figure 2.5: Average age-at-diagnosis in lung cancer mortality, women, in deprivation quintiles 1 (most deprived) to 5 (least deprived), of regions of England between 2001 and 2017.

Figure 2.6 demonstrates AAD estimates in LC for men between 2001 and 2017. Similar to the female counterparts, there is an increasing trend in calculated AAD values over the time, with lower AAD values estimated in more deprived quintiles of a given region. We note higher AAD estimates for men as opposed to women.

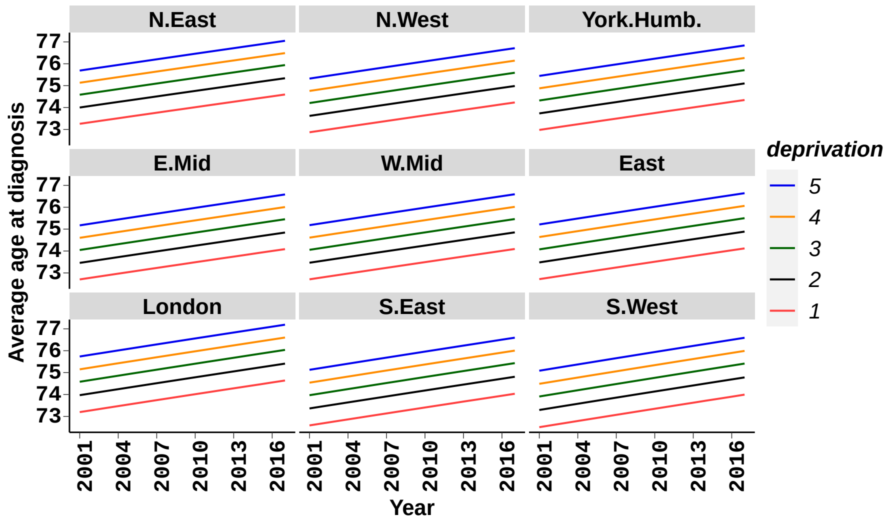


Figure 2.6: Average age-at-diagnosis in lung cancer mortality, men, in deprivation quintiles 1 (most deprived) to 5 (least deprived), of regions of England between 2001 and 2017.

Chapter 3

Statistical Modelling

We apply separate projection models to each type- and gender-specific cancer data described in Section 2.1. In each model, we consider different groups in the population of England, stratified by deprivation quintiles in the nine regions of England. Although age-period-cohort type models are mostly applied for projection purposes (see, for instance, Antonio et al. [2015], Czado et al. [2005], Smittenaar et al. [2016]), we implement more elaborate models by taking the advantage of our granular data with the aim of identifying regional and deprivation differences in future years.

We note that some results in this part have been presented in some international conferences including the Actuarial Research Conference 2022 in Chicago, US and the International Congress of Insurance: Mathematics and Economics 2023, Edinburgh, UK.

3.1 Projection models for cancer mortality

We assume that the number of gender- and type-specific cancer deaths $D_{a,d,r,t}$ at age a and year t in deprivation quintile d of region r in England follows a Poisson distribution with mean and variance shown by $\hat{\mathbb{E}}(D_{a,d,r,t}) = \hat{\theta}_{a,d,r,t}\ E_{a,d,r,t}$. Although this is a common assumption in the literature since the study of Brouhns et al. [2002], there are issues with this assumption as discussed later in this chapter.

Thus, in order to account the heterogeneous structure in the sub-populations of England, we construct a baseline model for a given gender- and type-specific cancer mortality using a Poisson-lognormal Bayesian hierarchical model. The general structure of our model is:

$$
\begin{aligned}
D_{a,d,r,t} &\sim \text{Poisson}(\theta_{a,d,r,t}\ E_{a,d,r,t}) \\
\theta_{a,d,r,t} &\sim \text{Lognormal}(\mu_{a,d,r,t}, \sigma^2) \\
\mu_{a,d,r,t} &= \boldsymbol{\beta X} + f(\boldsymbol{X}, \boldsymbol{\kappa}) \\
\sigma^2 &\sim \text{Inv.Gamma}(1, 0.1) \\
\boldsymbol{\beta} &\sim \text{Normal}(0, 10^4),
\end{aligned}
\tag{3.1}
$$

where:

- $\theta_{a,d,r,t}$ is gender-specific mortality rates at age-at-death a in year t in deprivation quintile d of region r (where applicable).
- $\mu_{a,d,r,t}$ is the location parameter of lognormal distribution for a given cancer type. It is defined based on various covariates - including *age-at-death a, deprivation d, region r, AAD, and NS prevalence rates* - which are collectively denoted by $\boldsymbol{X}$, with associated model parameters $\boldsymbol{\beta}$.

  Meanwhile, the function $f(\boldsymbol{X}, \boldsymbol{\kappa})$ explicitly captures the period effect t and potential two-way interactions between time and other covariates. The structure of $\boldsymbol{\kappa}$ is explicitly defined in (3.3).
- Non-informative prior distributions are assumed for model parameters $\boldsymbol{\beta}$ and σ^2, where appropriate, to reflect relative prior ignorance on their values (see, for example, Gelman et al. [2013], Ntzoufras [2009], for discussions on non-informative and informative priors).

The structure of the location parameter, $\mu_{a,d,r,t}$, differs for each gender- and type-specific cancer. It mostly depends on main variables, namely age, year, deprivation quintile, region, and AAD, along with two-way interaction terms between main variables, where appropriate. For example, $\mu_{a,d,r,t}$ in (3.1) might have the following form:

$$
\mu_{a,d,r,t} = \beta_0 + \beta_{1,a} + \beta_{2,r} + \beta_{3,d} + \beta_4 \text{AAD}_{r,d} + \beta_5 \text{NS}_{a,t-20} + \kappa_t + \text{interaction terms},
\tag{3.2}
$$

where age-at-death a, region r, and deprivation quintile d are considered as categorical variables, and NS prevalence rates and AAD variable are assumed to be numerical variables, standardised to have zero mean and unit variance to facilitate the calculations. We note that NS prevalence rates are involved with a lag of 20 years based on the study of

Luo et al. [2022]. Meanwhile, period effect t is defined differently from other covariates, with associated coefficient denoted by k_t.

In our model, we describe the period effect in (3.1) using a random walk with drift. Depending on the model specification, there can be more than one coefficient related to period effects, denoted by $\boldsymbol{\kappa} = (\kappa_{1,t}, \kappa_{2,t})^T$, between 2001 and 2018. Specifically, we assume that each $\kappa_{i,t}$ for $i = 1, 2$ follows the process below:

$$
\begin{aligned}
\kappa_{i,t} &= \psi_{\kappa_i} + \kappa_{i,t-1} + \epsilon_{\kappa_i,t} \\
\psi_{\kappa_i} &\sim \text{Normal}(0, {\sigma^2}_{\psi_{\kappa_i}}) \\
\epsilon_{\kappa_i,t} &\sim \text{Normal}(0, \sigma^2_{\kappa_i}) \\
\sigma^2_{\kappa_i} &\sim \text{Inv.Gamma}(1, 0.001).
\end{aligned} \tag{3.3}
$$

Here, the variance of the drift term is estimated as $\hat{\sigma}^2_{\psi_{\kappa_i}} = \frac{1}{2018-2001}\,\hat{\sigma}^2_{\kappa_i}$.

Note that, for model identifiability and interpretability, in general, sum-to-zero (STZ) constraints are imposed to all categorical variables. STZ constraint allows us to make comparisons between a given level of a categorical variable with respect to the corresponding average effect as reference level. However, we adopt a different constraint for the period effect $\boldsymbol{\kappa}$ such that $\boldsymbol{\kappa}_1 = (\kappa_{1,0}, \kappa_{2,0})^T = (0, 0)^T$. This changes the interpretation of the period-related coefficients by setting the first year as the baseline year. To be specific, the values related to the following years would be estimated with respect to $\boldsymbol{\kappa}_1$ and thus should be interpreted accordingly. See, for example, Wong et al. [2018] for a similar approach.

Several models are implemented to fit the historical mortality rates under inspection and then compared systematically. A variable selection procedure is carried out to determine the structure of parameter μ for a given cancer type, as explained in Appendix A.

3.2 Model for lung cancer mortality

LC mortality exhibits distinct differences by gender in its historical trends as presented and discussed in Section 4.1. This is considered to be associated with changing smoking patterns as noted by ONS [2017]. Although average daily smoking has stayed higher among men in the last years, the decline in average tobacco consumption in time, along with other things, e.g. healthy diet, seems to be reflected on LC mortality. Thus, men have enjoyed LC mortality improvements in more deprived areas whereas women experienced a deterioration in the most recent years. For further discussion, see Arık et al. [2021], ONS [2017].

In order to capture gender-specific differences in LC, we implement separate models for each gender. We rely on a set of data, *split by five-year age groups, region, deprivation decile, and single years* between 2001-2018 (Section 2.1). The LC mortality rates are examined from ages 45–49 to 85–89, where the two youngest age groups, ages 45–49 and 50–54, are combined to represent ages 45–54, and deprivation deciles to be quintiles in order to avoid small numbers. Note that LC is considered to be a bigger issue starting from age 45 according to the cancer statistics in the UK and available literature [Collin et al., 2008, ONS, 2015, 2017, 2018a, Wakelee et al., 2007].

Importantly, for each gender, we first consider a model consisting of available main variables, *that are age-at-death, year, region, deprivation quintile, and AAD*, along with related two-way interaction terms. Second, NS prevalence rates (Section 2.3.1) are integrated into the modelling approach in order to improve the developed models further. This is because smoking is indicated to be the biggest risk factor for developing LC and considered to be responsible of 72% of LC cases [DHSC, 2023, NHS, 2022].

Female lung cancer mortality

Female LC mortality is more complicated than male LC mortality, for instance, suggesting slowdown in mortality improvement at different age groups in the recent years. Although this could be related to various factors, one reason is considered to be changes in smoking patters, where women started smoking more and men less after the Second World War [ONS, 2017] (see Appendices C.1–C.2 for modelling results).

We first establish a model, where the location parameter of lognormal distribution in (3.1) is defined based on available main variables, as

$$\mu^{\text{lung}}_{a,d,r,t} = \beta_0 + \beta_{1,a} + \beta_{2,r} + \beta_{3,d} + \beta_4 \text{AAD}^{\text{morbidity}}_{r,d} + \beta_{5,d,a} + \beta_{6,r,a} + \kappa_{1,t} + \kappa_{2,t} \text{AAD}^{\text{morbidity}}_{r,d}. \tag{3.4}$$

Note that (3.4) has been determined with respect to a variable selection procedure, where the details can be found in Appendix A.1. Here, $\beta_{1,a}$ is the age coefficient for age group a with levels $a = 1, 2, \ldots, 8$, where a maps to $\{45-54, 55-59, 60-64, \ldots, 85-89\}$; $\kappa_{1,t}$ is the coefficient associated with period t with levels $t = 1, 2, \ldots, 18$, where t maps to $\{2001, 2002, \ldots, 2018\}$; $\beta_{2,r}$ is the coefficient of the region component for region r with levels $r = 1, 2, \ldots, 9$, where r maps to {North East, North West, Yorkshire and the

Humber, East Midlands, West Midlands, East, London, South East and South West}; $\beta_{3,d}$ is the deprivation component for quintile d with levels $d = 1, 2, \ldots, 5$, respectively; $\kappa_{2,t}$ is the coefficient of interaction between period effect and AAD component; $\beta_{5,d,a}$ is the coefficient of interaction between age-at-death and deprivation quintile; and $\beta_{6,r,a}$, for the interaction between age-at-death and region.

We further improve this model by introducing NS prevalence rates in women and an interaction between region and AAD, as additional risk factors, such that

$$\mu^{\text{lung}}_{a,d,r,t} = \beta_0 + \beta_{1,a} + \beta_{2,r} + \beta_{3,d} + \beta_4 \text{AAD}^{\text{morbidity}}_{r,d} + \beta_{5,d,a} + \beta_{6,r,a} + \kappa_{1,t} + \big(\kappa_{2,t} + \beta_{7,r}\big)\text{AAD}^{\text{morbidity}}_{r,d} + \beta_8 \text{NS}^{\text{women}}_{a,t-20}. \tag{3.5}$$

Specifically, $\beta_{7,r}$ is the coefficient of interaction between region effect and AAD component, and β_8 is the coefficient for the NS prevalence rates (Table A.2). We note that NS prevalence rates and the AAD variable are introduced as numerical variables, *standardised to have zero mean and unit variance to facilitate the calculations.*

We further note that the models in (3.4) and (3.5) have been implemented to quantify the impact of diagnosis delays on LC mortality in women, with the help of AAD under different scenarios.

Male lung cancer mortality

Male LC mortality demonstrates a generally decreasing trend at different age groups over the years, with distinct differences between the most and least deprived quintiles (see Appendices C.3–C.4 for modelling results).

Once more, we first establish a model considering the main variables in the available data as follows:

$$\mu^{\text{lung}}_{a,d,r,t} = \beta_0 + \beta_{1,a} + \beta_{2,r} + \beta_{3,d} + \beta_4 \text{AAD}^{\text{morbidity}}_{r,d} + \beta_{5,d,a} + \kappa_{1,t} + \big(\kappa_{2,t} + \beta_{6,r}\big)\text{AAD}^{\text{morbidity}}_{r,d}, \tag{3.6}$$

where there are two differences as compared to the female model in (3.4): (a) no interaction term between age and region, and (b) the interaction between region effect and AAD component ($\beta_{6,r}$) involved following the variable selection procedure (see Table A.4 in Appendix A.2).

We then improve this model by using male NS prevalence rates as an additional risk factor such that

$$\mu^{\text{lung}}_{a,d,r,t} = \beta_0 + \beta_{1,a} + \beta_{2,r} + \beta_{3,d} + \beta_4 \text{AAD}^{\text{morbidity}}_{r,d} + \beta_{5,d,a} + \kappa_{1,t} + \\ \big(\kappa_{2,t} + \beta_{6,r}\big)\text{AAD}^{\text{morbidity}}_{r,d} + \beta_7 \text{NS}^{\text{men}}_{a,t-20}, \tag{3.7}$$

where smoking information is incorporated by adding β_7 to represent the coefficient for the NS prevalence rates (Table A.5).

Similar to female LC models in (3.4) and (3.5), the models in (3.6) and (3.7) have been used to establish pre- and post-pandemic scenarios by introducing diagnosis delays in AAD component and estimating related increases in LC mortality in men accordingly.

We further note that different best fitted models can be identified by changing the description of null model in the variable selection process, as demonstrated in Appendix A. We implemented several model specifications before determining the final modelling structure(s). The overall decision has been made with the aim of finding a compromise between model complexity, data fitting, and potential correlations across different variables. We also acknowledge that the interaction term between the period effect and AAD component in (3.4) is added to have a consistent modelling structure for LC in both men and women. Note that this term is found to be significant in the existence of smoking information and also in the alternative variable selection procedure.

3.3 Model for breast cancer mortality

We focus on female BC mortality as there are very few records regarding male BC. In particular, we use the population data of England from 2001 to 2018, *stratified by five-year age groups and region,* pointing out a generally decreasing trend in BC mortality over the time (see Appendices C.7–C.8 for modelling results). It is important to note that deprivation is not found to be a significant variable to explain the changes in BC mortality (see Appendix A.3 and Arık et al. [2021] for wider discussion). Hereby, BC mortality projection is considered on regional level. Furthermore, the AAD variable by the regions of England has not been found to be statistically important to explain BC mortality either. This is considered because we have more ‘equality’ in BC mortality, leading to comparable AAD estimates in different regions.

Following the variable selection procedure, see Appendix A.3, we have first considered a much simpler projection model as compared to LC models such that

$$\mu_{a,r,t}^{\text{breast}} = \beta_0 + \beta_{1,a} + \beta_{2,r} + \kappa_{1,t}. \tag{3.8}$$

Here, $\beta_{1,a}$ is the age coefficient for age group a with levels $a = 1, 2, \ldots, 11$, where a maps to $\{35-39, 40-44, 45-49, \ldots, 85-89\}$; $\kappa_{1,t}$ is the coefficient for the period component for period t with levels $t = 1, 2, \ldots, 18$, where t maps to $\{2001, 2002, \ldots, 2018\}$; $\beta_{2,r}$ is the coefficient of the region component for region r with levels $r = 1, 2, \ldots, 9$, where r maps to {North East, North West, Yorkshire and the Humber, East Midlands, West Midlands, East, London, South East and South West}, respectively.

Afterwards, we account for female NS prevalence rates as an additional risk factor that contributes to explain BC mortality. The association between BC risk and smoking has been extensively studied by considering the amount of cigarette consumption [Hunter et al., 1997], duration of smoking [Reynolds et al., 2004], and smoking initiation at different ages [Al-Delaimy et al., 2004], sometimes leading to conflicting results. However, there is more evidence suggesting a potential causality between smoking and BC, especially in the case of long-term heavy smoking and smoking initiation at a young age [Reynolds, 2013, Xue et al., 2011].

Hereby, we focus on an improved version of (3.8) as follows:

$$\mu_{a,r,t}^{\text{breast}} = \beta_0 + \beta_{1,a} + \beta_{2,r} + \beta_3 \text{NS}_{a,t-20}^{\text{women}} + \kappa_{1,t}, \tag{3.9}$$

where β_3 is the smoking coefficient.

Provided regional-level cancer mortality is available up to 2022, see Section 2.1, we have utilised observations from 2019 until 2022 in order to make comparisons between observed and projected BC mortality.

3.4 Projection method for deaths

We develop a Bayesian methodology as described in (3.1) for the population data of England. Different covariates, i.e. region, deprivation quintile, and AAD, are regressed on the location parameter of lognormal distribution in addition to usual variables, such as age and period, in (3.4) and (3.6) for female and male LC mortality, respectively, and (3.9) for BC mortality.

In order to project a given type- and gender-specific cancer mortality beyond the observed calender period, we assume that the age-at-death-, region-, and deprivation-related effects remain unchanged over time. Specifically, the future mortality rates can be derived as

$$\theta^*_{a,d,r,t} \sim \text{Lognormal}(\mu^*_{a,d,r,t}, \sigma^2),$$

where a new location parameter, $\mu^*_{a,d,r,t}$, is defined considering changes in time-related terms.

To be precise, the period-related effects for LC could be extrapolated from 2019 to 2036 by setting the baseline year as the last year of the observed calendar year such that $\boldsymbol{\kappa}^*_1 = (\kappa_{1,19}, \kappa_{2,19})^T = (\hat{\kappa}_{1,18}, \hat{\kappa}_{2,18})^T$ in (3.3). BC data is available for a longer period, up to 2022. However, the last four calendar years, 2019 to 2022, are used to check with model validation in BC mortality instead of involving these years in our modelling. Hereby, the extrapolation could go until 2036, where the baseline year would be $\boldsymbol{\kappa}^*_1 = \kappa_{1,19} = \hat{\kappa}_{1,18}$ in that case.

3.5 Overdispersion of the Poisson-lognormal model

Poisson distribution imposes mean-variance equality, that is

$$\hat{\mathbb{E}}(D_{a,d,r,t}) = \hat{\text{var}}(D_{a,d,r,t}) = \hat{\theta}_{a,d,r,t}\ E_{a,d,r,t},$$

where $\hat{\theta}_{a,d,r,t}$ denotes the expected mean of gender- and type-specific fitted mortality rates at a given age a in deprivation quintile d of region r in England.

This is a strong assumption implying that individuals born in the same year would have the same mortality experience despite several different factors, such as smoking, income, and education, impacting mortality. See, for example, Arık et al. [2020], Brown [2003], Redondo Loures and Cairns [2019] for relevant discussion. This causes an additional variation across individuals, also known to be 'overdispersion'. In order to deal with overdispersion, we establish a hierarchical modelling structure using lognormal distribution in (3.1).

Furthermore, we check with the model fit of a given gender-specific model using Pearson residuals across different ages and years for a given region and deprivation quintile. As explained in Wong et al. [2018], the residuals can be derived as follows:

$$r_{a,d,r,t} = \frac{D_{a,d,r,t} - \hat{\mathbb{E}}(D_{a,d,r,t})}{\sqrt{\hat{\text{var}}(D_{a,d,r,t})}},$$

where $\hat{\mathbb{E}}(D_{a,d,r,t}) = \hat{\theta}_{a,d,r,t} E_{a,d,r,t}$ and
$\hat{\text{var}}(D_{a,d,r,t}) = \hat{\mathbb{E}}(D_{a,d,r,t}) \times (1 + \hat{\mathbb{E}}(D_{a,d,r,t}) \exp(\sigma^2 - 1))$. The corresponding fitted mortality rate, $\hat{\theta}_{a,d,r,t}$, is derived using the mean of lognormal distribution as $\hat{\theta}_{a,d,r,t} = \exp(\mu_{a,d,r,t} + \sigma^2/2)$.

3.6 Pandemic scenarios and excess deaths

We develop pandemic scenarios by considering an increase in the AAD covariate. This is motivated by delays in cancer diagnosis due to initial health disruptions caused by the COVID-19 pandemic. *Excess cancer deaths* for a given gender at various age groups a in deprivation level d and region r of England for cancer type c in the projection years, $\text{ED}^c_{a,d,r,t}$, are calculated by subtracting the estimated number of deaths in pre-pandemic scenarios from those in post-pandemic scenarios as follows:

$$\text{ED}^c_{a,d,r,t} = \hat{\mathbb{E}}(D^{\text{pandemic}}_{a,d,r,t}) - \hat{\mathbb{E}}(D^{\text{baseline}}_{a,d,r,t}),$$

where $\hat{\mathbb{E}}(D^{\text{baseline}}_{a,d,r,t})$ refers to the pre-pandemic estimates with no COVID-impact.

Meanwhile, *excess type-specific cancer mortality* in the projection years is obtained by dividing excess type-specific cancer deaths by the corresponding mid-year population estimates. Thus, in order to calculate age-specific excess cancer mortality for a given gender in projected year t, $\text{EAM}^c_{a,t}$, we use

$$\text{EAM}^c_{a,t} = \frac{\hat{\mathbb{E}}(D^{\text{pandemic}}_{a,t}) - \hat{\mathbb{E}}(D^{\text{baseline}}_{a,t})}{E_{a,t}},$$

whereas in order to calculate region-specific excess cancer mortality in year t, $\text{ERM}^c_{r,t}$, we apply

$$\text{ERM}^c_{r,t} = \frac{\hat{\mathbb{E}}(D^{\text{pandemic}}_{r,t}) - \hat{\mathbb{E}}(D^{\text{baseline}}_{r,t})}{E_{r,t}}.$$

Last, deprivation-specific cancer mortality for a given gender in year t, $\text{EDM}^c_{d,t}$, is

$$\text{EDM}^c_{d,t} = \frac{\hat{\mathbb{E}}(D^{\text{pandemic}}_{d,t}) - \hat{\mathbb{E}}(D^{\text{baseline}}_{d,t})}{E_{d,t}}.$$

Provided that AAD is not found to be statistically significant to explain BC mortality, a different model has been developed for BC mortality based on a Markov approach in separate work by the author [Arık et al., 2023a,b].

Hereby, in this study, we examine three pandemic scenarios for LC, not for BC. Specifically, Scenario 1 (S1) introduces a 1-month delay in AAD. In Scenario 2 (S2) we assume a 3-month delay in AAD. Last, in Scenario 3 (S3) we assume a 6-month delay in AAD.

Assumption 1: cancer survival

We take into account for net cancer survival to distribute an overall increase (1- to 6-month) in AAD over time. Any increase in AAD in a given year would lead to an increase in type-specific cancer mortality under inspection in the same year. Hereby, the aim is to allow a gradual increase in the related cancer mortality in the future years. Particularly, we assume that the pandemic-related health disruptions could cause a bigger increase in AAD in the first year of the pandemic, e.g. COVID-19, by gradually declining later on.

LC survival would be gradually declined over time such that 40% of people with LC would survive from this disease for one year or more, 15% for 5 years or more, and 10% for 10 years or more [CRUK, 2021b]. Hereby, we assume that a 60% increase of a particular delay, e.g. 1-month, in the AAD variable would realise in the second year of the projection period, *the first year of the COVID-19 pandemic*, 2020. This would be followed by a 25% increase up to 5 years, 10% from 6 to 10 years, and 5% in the rest of the projection period.

Assumption 2: population estimates in future years

Estimated cause-specific number of deaths in a particular year involves multiplying relevant mortality rates by the corresponding mid-year population estimates. The general modelling structure given in (3.1) provides the framework to project mortality rates. Nevertheless, for the calculation of the related number of deaths, the relevant mid-year population estimates must be provided as well.

We rely on the national population projections provided by the ONS in the projection years. The data are split by five-year age groups and gender in the regions of England from 2019 to 2043 [Nash, 2020]. To facilitate our calculations, we specifically require the mid-year future population estimates by deprivation quintiles. Consequently, we assume that the distribution of population estimates across deprivation quintiles within a specific region in 2018 will remain unchanged throughout the projected years.

Chapter 4

Numerical Results

This chapter mainly discusses estimated historical and future cancer mortality trends based on the methodology introduced in Chapter 3. The numerical results are obtained by calibrating the developed models with respect to the population LC and BC mortality in England.

In each model, we check with the model adequacy using heat maps of Pearson residuals as diagnostic tools across age and year for a given region and deprivation quintile of England. These maps allow us to detect whether or not a regular pattern appears over ages or years. If such pattern is observed, this would mean that residual mortality is not explained well by the underlying model. We also examine age-specific fitted and crude type-specific mortality rates across different deprivation quintiles in each region of England over the observed calendar years 2001–2018, where appropriate, along with future mortality rates up to 2036. Furthermore, for LC mortality, we investigate age-, region-, and deprivation-specific excess mortality in England over the projected period.

4.1 Lung cancer mortality

LC has been the leading cause of death for ages 50 to 79 before the COVID-19 pandemic, where COVID-19 replaced LC as the leading cause of death in women aged 65 to 79 years old in 2020 and 2021 [ONS, 2018b, 2021]. ONS [2023] further reported that the leading causes in women were BC and LC in 2022, in a similar manner with the pre-pandemic period. Age-specific LC mortality rates for both men and women in different deprivation groups and regions of England are presented in Appendices C.1 to C.4 based on two separate models, where one consists of smoking data and the other does not.

In this section, we mainly focus on examining variations in LC mortality across different deprivation levels and regions of England under different modelling assumptions. We then examine trends in deprivation gap over time across the regions of England. Last, we investigate the impact of diagnosis delays on future LC mortality by age, region, and deprivation.

4.1.1 Parameter estimates

We have investigated LC mortality using two different models, each developed separately for each gender, as described in Section 3. The key difference between the models arises from variable selection: one incorporates NS prevalence rates as a proxy for smoking, in addition to existing variables such as age-at-death and region.

Appendices B.2 and B.1 present the parameter estimates for female LC models with and without NS prevalence rates, respectively. Estimates for male LC models are in Appendices B.4 and B.3.

Our findings confirm that individuals over 65 and those in more deprived quintiles (specifically quintiles 1 and 2) face higher LC mortality risk, as indicated by positive estimates associated with related levels of age and deprivation variables. Regional effects show no clear pattern, likely due to the inclusion of deprivation and interaction terms.

Furthermore, NS prevalence rates negatively correlate with LC mortality risk, while increasing age contributes positively. The common period effect for women shows a positive contribution to LC mortality until around 2015 (when NS prevalence is included), followed by a decline. For men, without NS prevalence rates, the period effect follows a negative trend, indicating LC mortality improvement in recent decades. However, once NS prevalence is incorporated, the main period effect for men becomes less interpretable, likely because NS prevalence rates absorb much of the temporal effect.

4.1.2 Age-standardised mortality rates

Age-standardised fitted and projected LC mortality for men and women in deprivation quintiles 1, 3, and 5 across regions of England from 2001 to 2036 are shown in Figure 4.1 and Figure 4.2. Figure 4.1 presents results from LC models without smoking data, while Figure 4.2 shows results from models that include smoking data, as described in Section 3.2.

LC mortality in women has mostly deteriorated in the past decade(s), with a generally widening deprivation gap in the regions of England, whereas there has been a decreasing trend in male LC mortality between 2001 and 2018. Historical trends in Figure 4.1 and Figure 4.2 point out comparable outcomes with the earlier literature, such as Arık et al. [2021]. Particularly, there has been generally a non-decreasing trend in LC mortality for women in the most deprived quintile in the regions of England. Meanwhile, the least deprived quintiles present mostly levelled rates for women. Although LC mortality in men is higher than their female counterparts, there has been mortality improvement across all quintiles and regions over the years.

Deprivation gap, *which is the difference between the highest (in quintile 1) and lowest (in quintile 5) mortality rates in a given region at a particular time*, has widened for women from 2001 to 2018, with a bigger gap in the northern regions of England in comparison to the southern regions of England (see Section 2.5 for formal definition(s) of 'deprivation gap'). The change in deprivation gap for men is less clear than women, requiring a more in depth analysis. Importantly, we predict that the differences between the most and least deprived quintiles for both men and women persist in the future, where projected mortality rates in deprivation quintiles 1 and 5 remain to be significantly different from each other in each region.

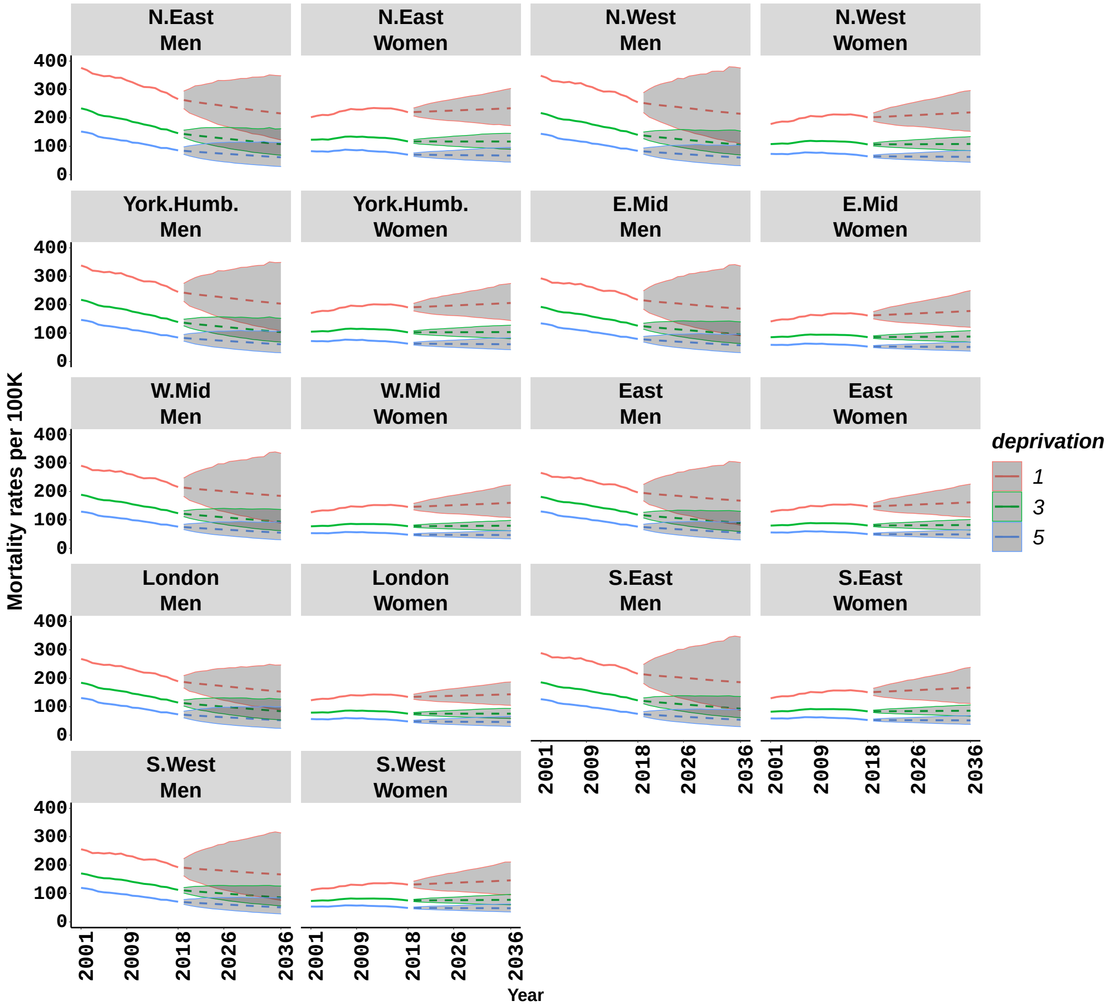


Figure 4.1: Age-standardised fitted (lines) and projected (dashed lines) lung cancer mortality, with 95% credible intervals, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) and regions of England based on **full models without smoking data**.

Figure 4.1 and Figure 4.2 demonstrate that the key difference between different modelling approaches arises from the predicted rates in the future. For example, in Figure 4.2, we estimate mortality improvement for women at a higher degree in different deprivation quintiles, especially the most deprived neighbourhoods in each region. This implicitly indicates a marginal decrease in deprivation gap based on the model including the smoking data over the future years.

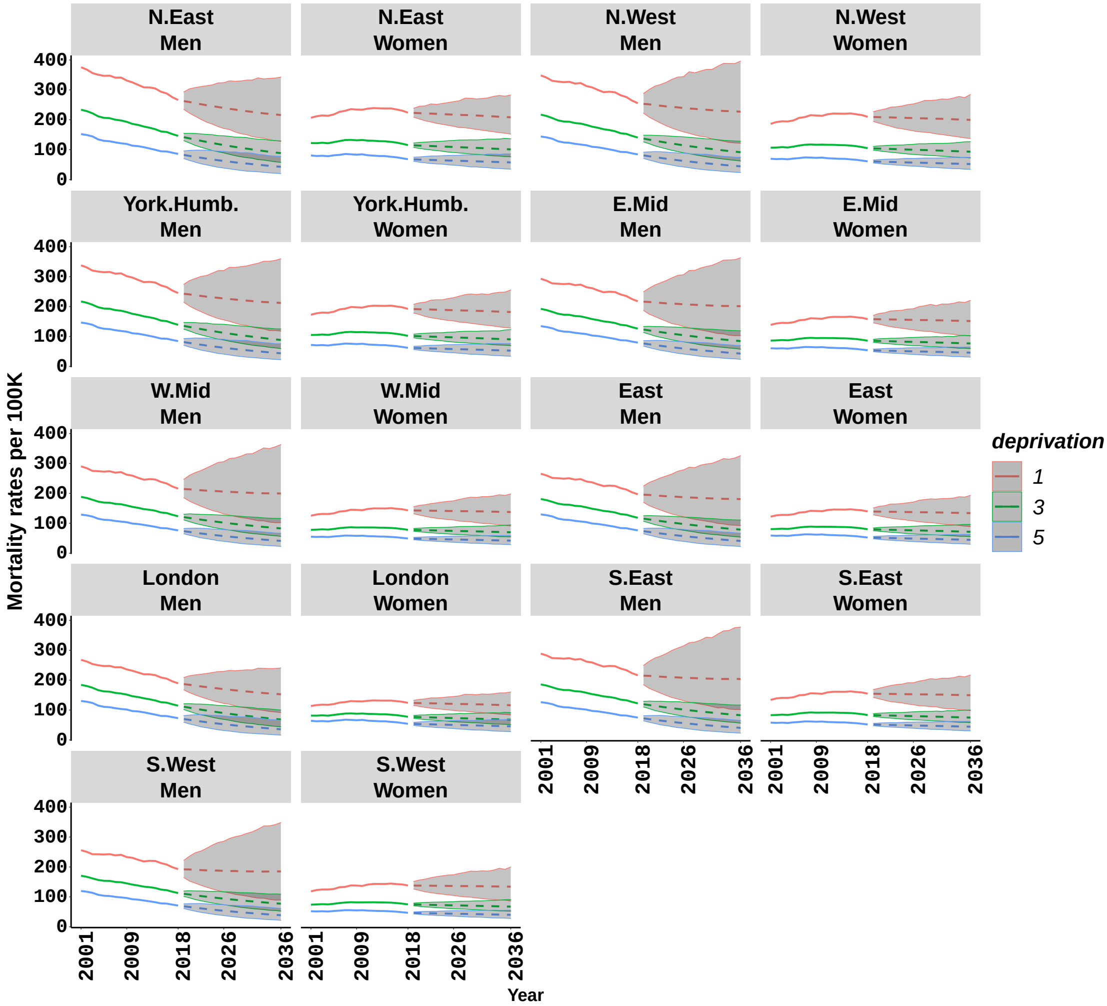


Figure 4.2: Age-standardised fitted (lines) and projected (dashed lines) lung cancer mortality, with 95% credible intervals, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) and regions of England based on **full models with smoking data**.

4.1.3 Disparities in lung cancer

Socio-economic and regional differences are relevant to LC mortality as discussed in Section 4.1.2. In order to tackle these inequalities, it is crucial to understand how these differences have changed throughout the years and are expected to change in the future. Therefore, we quantify socio-economic inequalities in LC by using a relative deprivation metric, defined in Section 2.5.

Figure 4.3 displays how relative deprivation gap in LC mortality has changed from 2001 to 2018 and is expected to further change in 2036 in England based on the model specifications with and without smoking data. The figure points out a widening deprivation gap for women, with an increasing trend for both genders in each region from 2001 to

2018, alongside marginal increases in 2036. Figure 4.3 also indicates a persistent deprivation gap, still associated with worse outcomes in the most deprived quintiles as compared to the least deprived quintiles.

Although both models show some differences - more pronounced in LC for women - they agree on worsening LC outcomes in the northern regions of England in comparison to the southern regions. We note significant differences across the north and south of England in 2018, as evidenced by non-overlapping 95% confidence intervals.

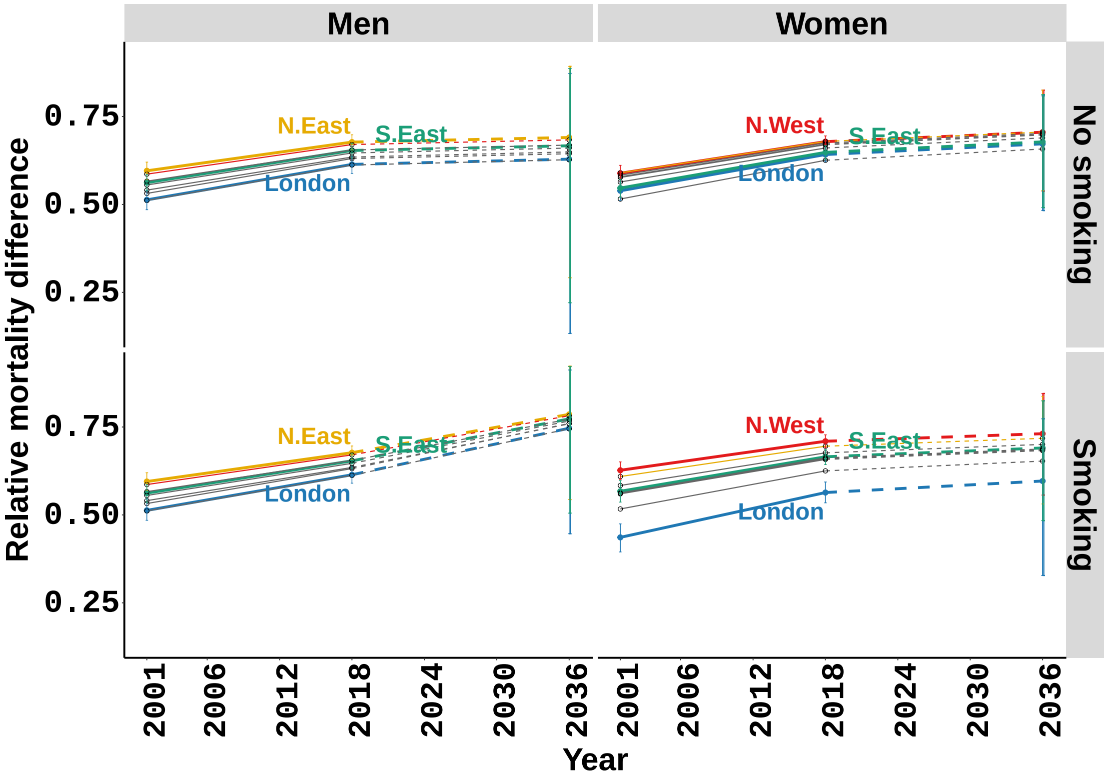


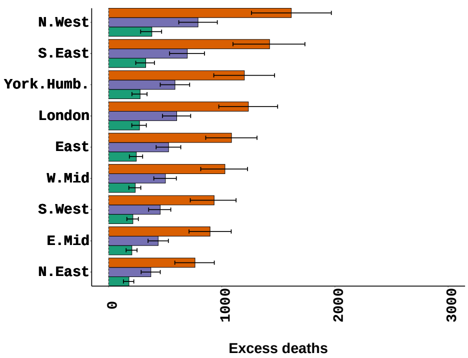
Figure 4.3: Relative deprivation gap, $RD_{\text{female},r,t}$, in comparison to the most deprived quintile, in lung cancer mortality in 2001, 2018, and 2036 for the regions of England, with 95% credible intervals.

4.1.4 Excess number of deaths up to 2036

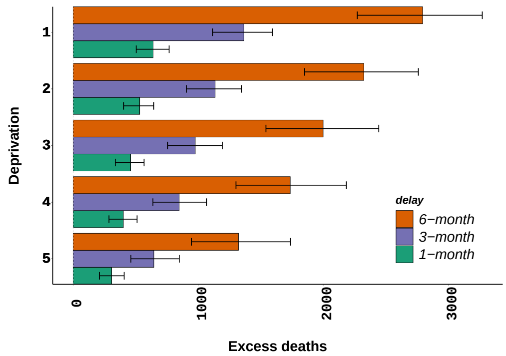
In this part of our study, we quantify the impact of diagnosis delays caused by the COVID-19 pandemic on LC mortality under three pandemic scenarios. Specifically, we present estimates of LC 'excess deaths' and 'excess mortality' for men and women. Technical specifications of excess deaths and excess mortality, along with our modelling assumptions, can be found in Section 3.6. Our findings are based on the models that incorporate NS prevalence rates as a proxy for smoking (see (3.5) for women and (3.6) for men). These models are more refined than alternative models that do not include NS prevalence rates, the results of which are provided in Appendix C.5.

Figure 4.4 illustrates overall excess deaths over a 17-year period (2020–2036) for women and men in England. Specifically, we estimate 2,340 (1,743; 2,869) and 10,180 (7,944; 12,340) excess deaths for women due to 1-month and 6-month diagnosis delays, respectively. Meanwhile, for men, 1-month and 6-month diagnosis delays are predicted to result in 5,164 (4,353; 6,066) and 28,660 (23,040; 35,090) additional deaths, respectively.

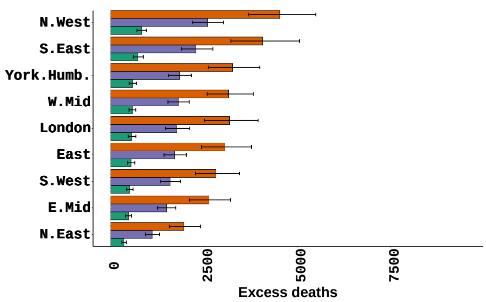
Importantly, we present total excess deaths up to 2036 in different regions and deprivation quintiles of England, with 95% credible intervals, as a result of varying diagnosis delays. We see significant differences in excess deaths, with higher numbers for men than for women, across different regions (such as *the East Midlands vs. the north west of England*), and deprivation quintiles (such as *the most vs. the least deprived quintile*). Furthermore, varying diagnosis delays lead to substantial variations in excess deaths in each region or deprivation quintile.



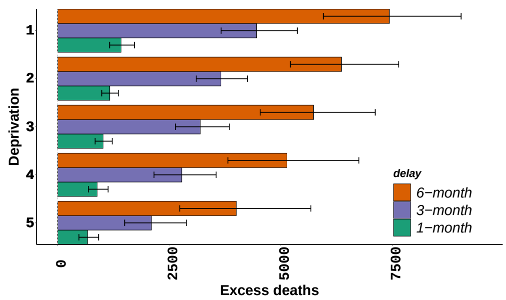
(a) Excess deaths, **women**, region



(b) Excess deaths, **women**, deprivation quintiles



(c) Excess deaths, **men**, region

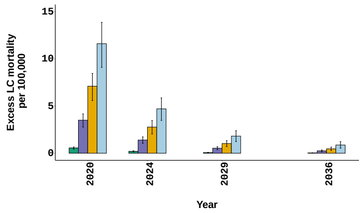


(d) Excess deaths, **men**, deprivation quintiles

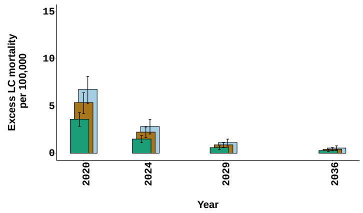
Figure 4.4: Cumulative lung cancer excess deaths from 2020 to 2036 based on the full models **with smoking data**. Total excess deaths (over 17 years) in different deprivation quintiles and regions of England, with 95% credible intervals.

Figure 4.5 presents annual changes in age-, region-, and deprivation-specific excess LC mortality per 100,000 men and women, based on the models with NS prevalence rates, for 1- and 6-month diagnosis delays. We predict significantly higher excess mortality with increasing age, such as between ages 60–64 and 70–74, for both genders. Furthermore, the highest excess deaths (in absolute terms) are predicted to occur between ages 70–79 in 2020, gradually decreasing over the projection period. See Appendix C.5 for findings based on the model excluding NS prevalence rates, and Appendix C.6 for results from the model incorporating NS prevalence rates. This is aligned with empirical evidence suggesting LC to be the leading cause of death between ages 65 and 79 before the COVID-19 pandemic and after 2021 [ONS, 2023].

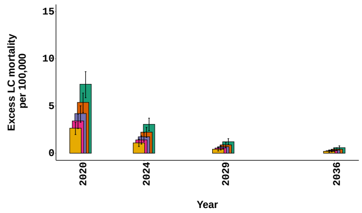
Our results also show marked regional differences in excess mortality, with northern regions (e.g. the north east of England) experiencing higher excess mortality than southern regions (e.g. the south east of England), especially in the first year of the COVID-19 pandemic. However, these regional disparities diminish over time. Similarly, we estimate significant differences in excess mortality between the most and least deprived quintiles in future years. A longer diagnosis delay results in significantly higher excess deaths across all deprivation quintiles, though the overall impact on mortality declines over time.



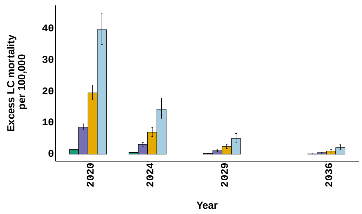
(a) 1-month delay, **women**, age-specific



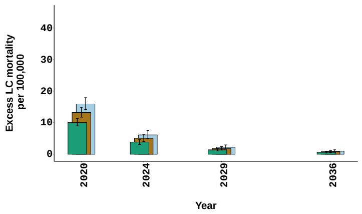
(b) 1-month delay, **women**, region-specific



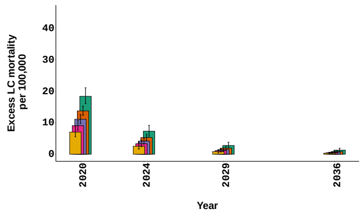
(c) 1-month delay, **women**, deprivation-specific



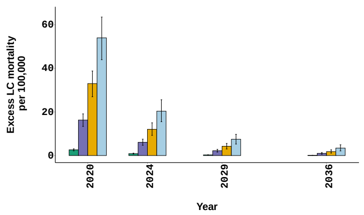
(d) 1-month delay, **men**, age-specific



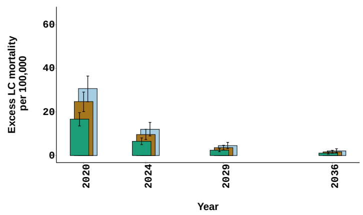
(e) 1-month delay, **men**, region-specific



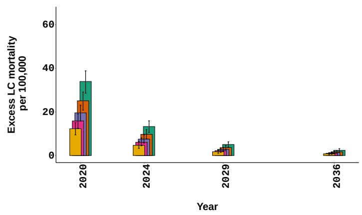
(f) 1-month delay, **men**, deprivation-specific



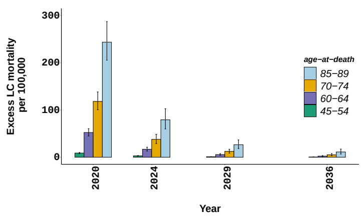
(g) 6-month delay, **women**, age-specific



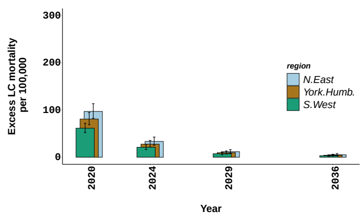
(h) 6-month delay, **women**, region-specific



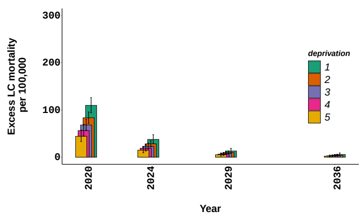
(i) 6-month delay, **women**, deprivation-specific



(j) 6-month delay, **men**, age-specific



(k) 6-month delay, **men**, region-specific



(l) 6-month delay, **men**, deprivation-specific

Figure 4.5: Lung cancer excess mortality, **per 100,000 people**, by age-at-death, selected regions and deprivation quintiles in England based on **full models with smoking data**. Annual excess deaths from 2020 to 2036, with 95% credible intervals. Note that differences in lung cancer excess mortality at other ages, in other regions or deprivation quintiles in intermediate years, are comparable to the presented years.

Significant differences are observed in the LC models for women when comparing models with and without NS prevalence rates. For example, at ages 85–89, Figure 4.5 shows noticeably lower excess mortality compared to Figure C.43, where the latter estimates nearly twice as many excess deaths. Notably, this difference arises due to two additional components in the model, (3.5), with NS prevalence rates: (i) NS prevalence rates and (ii) an interaction term between AAD and region. The inclusion of both terms enable further refinement of model estimates across different regions and deprivation quintiles in England. For details on the variable selection procedure and related model metrics, see Appendix A.

4.2 Breast cancer mortality

This section presents the main findings on BC mortality based on two model specifications. We compare results from (3.8) and (3.9), where the latter refines the earlier model by incorporating NS prevalence rates. Details on the variable selection procedure can be found in Appendix A.3.

Deprivation and AAD are not found to be significant explaining differences in historical BC mortality trends. Therefore, these variables have not been used to estimate future rates either. In the absence of AAD, our main focus has been on projection of future BC mortality rates in the regions of England.

4.2.1 Parameter estimates

Appendices B.6 and B.5 provide parameter estimates for BC models with and without NS prevalence rates, respectively.

Our findings indicate that individuals over 60 and those living in southern regions have higher BC mortality risk, as reflected in the positive parameter estimates of related levels of age and region variables. Similar to LC mortality, NS prevalence rates show a negative association with BC mortality. Additionally, estimates for the common period effect suggest a negative trend, indicating improvements in BC mortality over time. When NS prevalence rates are incorporated, part of the temporal effect is absorbed by this variable.

4.2.2 Age-standardised mortality rates

Figure 4.6 exhibits age-standardised fitted and projected BC mortality rates, with 95% credible intervals, based on the models described in (3.8) and (3.9), without and with NS prevalence rates, respectively, across the regions of England up to 2036. There is a declining trend in all regions over the calendar years, that is also expected to continue in the future years.

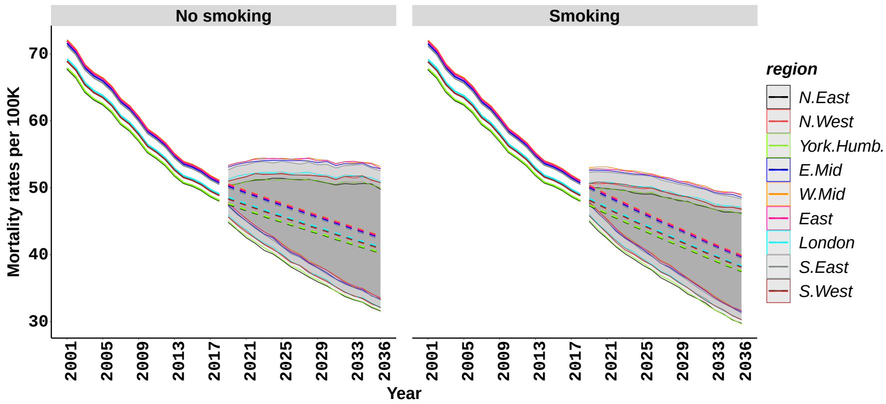


Figure 4.6: Age-standardised fitted and projected breast cancer mortality rates, females, in regions of England: fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

Figure 4.6 also highlights that the main differences between the models in (3.8) and (3.9) arise due to the time trend's slope in the future years and the associated 95% credible intervals. To be precise, we estimate more substantial mortality improvements in (3.9), with increased certainty, when we factor in smoking information.

4.2.3 Disparities in breast cancer for women

Figure 4.7 presents age-standardised BC fitted mortality rates based on the models in (3.8) and (3.9), without and with smoking information, in different regions of England in three time points: 2001, 2018, and 2036.

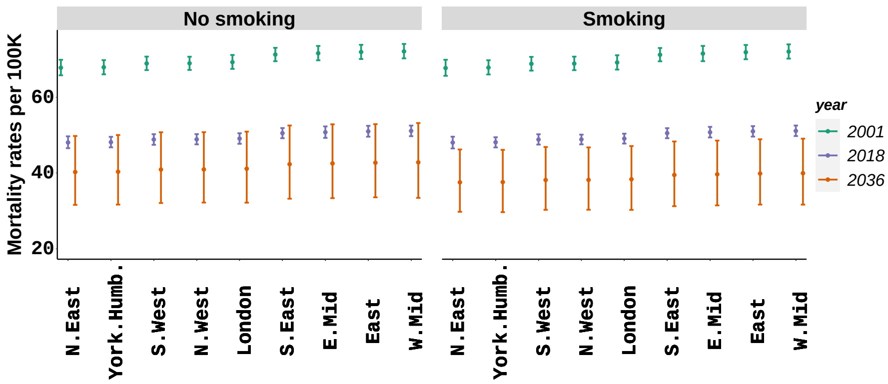


Figure 4.7: Age-standardised breast cancer mortality rates, females, in selected years in the regions of England with 95% credible intervals.

Figure 4.7 conveys the following main messages:

- A significant mortality improvement has been estimated in BC mortality from 2001 to 2018, and a further improvement is expected to happen in the future.
- Mortality improvement in the future is expected to happen at a different degree in each model, where the model with smoking information, (3.9), predicts slightly lower rates. This leads to different interpretation in terms of significance of the improvement in the projected rates. Specifically, the estimates in 2036 in all regions under (3.9) suggest a statistically significant change from 2018 as opposed to the estimates under (3.8).
- Despite the fact that region is a statistically significant variable to explain BC mortality, there are only marginal differences in mortality rates across different regions. This can be deduced, for instance, by looking at overlapped confidence intervals of each estimate across regions of England in the future years.

Chapter 5

Discussion and Conclusions

5.1 Statement of principal findings

In this project we have delved into two cancer types: LC and BC. We have constructed type- and gender-specific cancer mortality models, implemented using a Bayesian hierarchical modelling framework (Section 3). First, our analysis focused on estimating the historical trends in LC and BC mortality rates in various deprivation quintiles within the nine regions of England between 2001 and 2018. Following that we projected future cancer mortality rates on regional and deprivation level, where appropriate, up to 2036. Last, part of our analysis focused on quantifying the impact of diagnosis delays on cancer mortality (Section 4).

The main findings are summarised as follows:

- NS prevalence rates, used as a proxy for smoking, have been found to be significant for explaining both LC and BC mortality.
- The AAD variable, estimated based on type-specific cancer morbidity, has been shown to be significant in explaining LC mortality.
- Incorporating NS prevalence rates into the projection models has resulted in estimating higher cancer mortality improvements in future years.
- We found a widening deprivation gap in LC mortality from 2001 to 2018, with a persistent gap in future years.
- Our analysis revealed marginal differences in BC mortality across the nine regions of England between 2001 and 2018, with comparable estimates in the projected

years. Our estimates also indicated substantial differences in BC mortality between the youngest and oldest screening age groups in 2036.

- The AAD variable has been instrumental in constructing various COVID-19 related scenarios. Specifically, we introduced 1-, 3-, and 6-month delays in LC diagnosis with the aid of AAD variable.
- We have quantified excess deaths from LC from 2020 until 2036 based on the differences in the projected mortality rates between the pre- and post-pandemic scenarios. We have also calculated excess mortality from LC in the same years by dividing the excess deaths by the corresponding mid-year population estimates.
- Our findings have shown higher excess mortality from LC in men as compared to women. Furthermore, estimates for excess deaths from LC varied notably across the regions and deprivation quintiles of England, as a result of a given delay in cancer diagnosis.
- The impact of delays in LC diagnosis has resulted in varied excess mortality estimates for both men and women by
  - age groups, with the highest deaths between ages 65–85;
  - regions of England, with the highest in the north east of England; and
  - deprivation quintiles, with the highest in the most deprived quintile.

5.2 Strengths and limitations of this study

We provide a detailed modelling framework for LC and BC mortality, integrating model uncertainty through the structure of linear predictors. We also account for parameter uncertainty in a natural manner. Smoking is a significant risk factor for various diseases. We demonstrate how to incorporate this risk factor, along with the AAD variable, into cancer mortality modelling. The models are carried out to estimate LC and BC mortality between 2001 and 2018, and then project these mortality rates until 2036, together with 95% uncertainty intervals, on different regions and deprivation levels of England. We further show how one of the risk factors, the AAD variable, could be used to create separate scenarios associated with an extreme event, such as the COVID-19 pandemic.

The modelling results are broadly consistent with the existing literature, such as the reported numbers in Digital [2023], Luo et al. [2022]. Specifically, recent data from NHS Digital reveals that the age-standardised cancer mortality in England was highest for those living in the most deprived areas in 2020, in addition to an increasing deprivation gap compared to 2019 [Digital, 2023]. Specifically, the age-standardised all-cause cancer mortality in the most deprived quintiles was 53% higher for men and 55% higher for women, in contrast to those in the least deprived areas. This marks an increase from the 2019 cancer mortality, which showed a 49% disparity for both genders. Besides, LC is noted to be impacted the most in 2020, with age-standardised mortality of 103, per 100,000 people, for men in the most deprived areas, compared to 37 for those in the least deprived areas, and 78 for women as opposed to 26 for the ones in the least deprived areas. Aligned with these observations, our modelling approach has provided estimates of excess deaths from LC with 95% credible intervals, pointing out highest excess mortality in the most deprived quintiles or in the northern regions of England in the first year of the COVID-19 pandemic.

Hereby, our study can be significant to inform decision makers by increasing awareness about the continuing impact of the COVID-19 pandemic. The estimated results can also be helpful while implementing evidence-based health interventions.

We note that access to cancer mortality data at both deprivation and regional levels has been unavailable for recent years, specifically from 2019 onwards. Besides, smoking data has not been available in the same granularity as for cancer data. Similar challenges appeared when mid-year population estimates have been required by region and deprivation quintiles for obtaining excess deaths. This limited our ability to make data-driven inferences. Nonetheless, suitable adjustments were used when data, e.g. NS prevalence rates, are not provided to suitable resolution. Thus, our analysis is mostly based on data and estimation. Furthermore, we conducted a comparison between the observed and projected BC mortality rates between 2019 and 2022 in order to evaluate the performance of our models to some extent. Despite the disruptions caused by the COVID-19 pandemic in 2020 and 2022, the observed age-specific BC mortality rates were largely consistent with the corresponding 95% credible intervals established by our projections.

5.3 Implications for actuaries

Our findings can help life insurers understand the impact of late diagnosis on cancer mortality and survival rates. The modelling framework developed here can be useful for assessing different scenarios, not only those related to extreme events (as exemplified here with the COVID-19 related scenarios) but also those linked to regional or nationwide government initiatives. For example, it can be used to evaluate the potential impact of earlier diagnoses on cancer mortality, which could be facilitated by expanding existing screening programs to broader age groups or introducing pilot screening programmes in specific regions.

This study can add value while considering insurance pricing and valuation assumptions related to different sub-populations. For example, examining variations in cancer mortality among the most and least deprived population groups can provide valuable insights into potential differences between insured and general populations. Particularly, our findings can help actuaries to obtain informed estimates of future trends of cancer mortality in heterogeneous populations.

Understanding trends in cancer mortality is highly relevant to long-term insurance policies as this is an important underlying assumption for price guarantees. Our models can be relevant to cancer life insurance and CII, as well, where cancer is a major disease leading to large insurance payouts. Furthermore, accounting for the impact of diagnosis delays on cancer rates can allow actuaries to adjust, for example, current cancer life insurance or CII premiums, and reserve calculations.

5.4 Further research

A useful and detailed modelling framework is developed for two major cancer types as part of this project. Our approach provides a valuable model by leading to region- and deprivation-specific cancer mortality estimates in future years, apart from the usual variables such as age and gender. Yet, there are important areas for further research. Particularly, the modelling framework can be extended in the following ways:

- expanding our methodology to other common cancer types, such as bowel and prostate cancer, to obtain further insights in cancer risk;
- incorporating availability of national cancer screening with the aim of quantifying the impact of a halt or delay in cancer screening on cancer mortality;
- incorporating competing risk methods for other-cause mortality while examining the impact of the COVID-19 pandemic; and
- developing scenarios to evaluate the potential effects of earlier cancer diagnosis on mortality, particularly in the context of government initiatives related to existing or newly introduced screening programmes.

Appendix A

Variable Selection

In this study, we implement a Bayesian hierarchical modelling structure for a given gender- and cause-specific cancer, as described in (3.1). The key difference between models arises from the definition of location parameter μ in the log-normal distribution. Specifically, the models presented in Section 3.2 and Section 3.3 are determined using a Bayesian variable selection procedure in the R-INLA software [Lindgren and Rue, 2015] based on two criteria: Deviance Information Criterion (DIC) and Bayes factors.

In these models, age and year are treated as categorical variables to facilitate the development of an appropriate projection model. In contrast, NS prevalence rates and AAD are treated as numerical variables for two reasons: (a) they are not part of the original dataset, and given this, we aim to estimate their average impact on historical data, and (b) we seek a parsimonious model by minimising the number of parameters.

We employ a forward variable selection approach, starting with the simplest (null) model and adding variables iteratively, provided they improve model fit. Specifically, variables are included in the best-fitting models if they result in a lower DIC or a Bayes factor greater than 3, following the studies of Kass and Raftery [1995], Spiegelhalter et al. [2002].

For completeness, the DIC can be viewed as a Bayesian counterpart to the Akaike Information Criterion (AIC), with two key modifications: (a) replacing the maximum likelihood estimator of given parameter vector $\boldsymbol{\beta}$ with its posterior mean, mode, or median, and (b) incorporating a data-driven bias correction in place of the number of parameters. Specifically, the DIC is defined as

$$\text{DIC} = -4E_{\boldsymbol{\beta}|D}(\log f(D|\boldsymbol{\beta})) + 2\log f(D|\hat{\boldsymbol{\beta}}),$$

where $\hat{\boldsymbol{\beta}}$ is the posterior mean, mode or median of parameter vector $\boldsymbol{\beta}$, f is the likelihood function, and D is the related data. See, for example, Chapter 7 in Gelman et al. [2013] for further information.

The Bayes factor compares marginal likelihoods and is defined as a ratio of posterior odds of two competing models, H_j and H_k, based on data D, provided that the same prior distributions are assumed for both models, see Lindgren and Rue [2015], such that

$$B_{jk} = \frac{P(D|H_j)}{P(D|H_k)};\ j \neq k.$$

See Spiegelhalter et al. [2002] for further details on Bayesian measures of complexity and fit.

Section A.1 and Section A.3 summarise the variable selection process for LC and BC mortality, considering different null models. Our primary focus is on full models derived from a variable selection process starting with the simplest null model, which includes only the offset variable. However, we also explore alternative full models based on the availability of different variable choices, splitting the overall process into two separate variable selection stages to understand how trends in historical data can be further improved.

A.1 Variable selection for female lung cancer mortality

Table A.1: Variable selection procedure in the R-INLA software to determine the best fitted model for female lung cancer mortality, without considering 'smoking proxy'.

variable added	Bayes factor	marginal likelihood	diff. in marginal likelihood	DIC
null		-30356.43		45717.28
age	Inf	-25067.03	5289.39	44475.35
income	Inf	-22314.66	2752.36	42788.21
region	Inf	-21187.68	1126.98	41748.87
year	126951622561447328	-21148.30	39.38	41652.67
AAD	28463406452108.4	-21117.31	30.97	41614.40
income:age	2.26320928808427e+125	-20828.67	288.63	41114.45
region:age	312056207.62	-20809.11	19.55	40943.17

Note: The null model only includes the related offset variable, and age and year are defined to be categorical variables.

Table A.2: Variable selection procedure in the R-INLA software to determine the best fitted model for female lung cancer mortality, with ‘smoking proxy’.

variable added	Bayes factor	marginal likelihood	diff. in marginal likelihood	DIC
null		-20809.12		40943.37
smoking	1.53e+43	-20709.68	99.43	40757.45
AAD:region	4.00e+20	-20662.25	47.44	40644.90
AAD:year	36.24	-20658.66	3.59	40564.79

Note: The null model includes ALL variables in Table A.1; age and year are defined to be categorical variables; and non-smoker prevalence rates are used with 20-year time lag.

Table A.3: An alternative variable selection procedure in the R-INLA software to determine the best fitted model for female lung cancer mortality, with ‘smoking proxy’.

variable added	Bayes factor	marginal likelihood	diff. in marginal likelihood	DIC
null		-30356.45		45718.20
age	Inf	-25069.12	5287.33	44475.38
income	Inf	-22318.36	2750.76	42788.29
region	Inf	-21193.58	1124.77	41748.84
year	129115643505238544	-21154.18	39.39	41652.72
smoking	5.59e+41	-21058.05	96.12	41491.94
AAD	52164756508979.5	-21026.46	31.58	41452.27
age:income	1.05e+117	-20757.01	269.45	40937.04
age:smoking	3.80e+73	-20587.58	169.42	40649.96
region:AAD	8.93e+19	-20541.64	45.93	40535.07
income:smoking	3500548.57	-20526.58	15.06	40490.82
AAD:smoking	985210.84	-20512.78	13.80	40462.12
year:AAD	9706822596297934848	-20469.06	43.71	40310.30

Note: This is a two-stage variable selection procedure. The first stage begins with a null model that includes only the related offset variable and performs variable selection among all main variables. The second stage starts with a null model that includes the selected main variables and then applies variable selection among potential two-way interaction terms. Note that non-smoker prevalence rates are used with a 20-year time lag.

A.2 Variable selection for male lung cancer mortality

Table A.4: Variable selection procedure in the R-INLA software to determine the best fitted model for male lung cancer mortality, without considering 'smoking proxy'.

variable added	Bayes factor	marginal likelihood	diff. in marginal likelihood	DIC
null		-33057.97		47607.99
age	Inf	-26665.01	6392.96	46598.55
income	Inf	-23252.59	3412.41	44710.56
year	Inf	-22404.99	847.59	43789.62
region	1.62e+286	-21745.97	659.02	43003.53
AAD	8634573.64	-21730.00	15.97	42981.00
income:age	6.67e+228	-21203.12	526.88	42012.54
AAD:year	60480674313.05	-21178.29	24.82	41888.20
AAD:region	35.60	-21174.72	3.57	41843.68

Note: The null model only includes the related offset variable, and age and year are defined to be categorical variables.

Table A.5: Variable selection procedure in the R-INLA software to determine the best fitted model for male lung cancer mortality, with 'smoking proxy'.

variable added	Bayes factor	marginal likelihood	diff. in marginal likelihood	DIC
null		-21174.72		41843.63
smoking	1399391.13	-21160.57	14.15	41814.23

Note: The null model includes ALL variables in Table A.4; age and year are defined to be categorical variables; and non-smoker prevalence rates are used with 20-year time lag.

Table A.6: An alternative variable selection procedure in the R-INLA software to determine the best fitted model for male lung cancer mortality, with ‘smoking proxy’.

variable added	Bayes factor	marginal likelihood	diff. in marginal likelihood	DIC
null		-33057.98		47608.01
age	Inf	-26665.02	6392.96	46598.59
income	Inf	-23252.59	3412.42	44710.58
smoking	Inf	-22361.73	890.86	43784.52
region	4.63e+285	-21703.96	657.76	43002.49
AAD	8245920.92	-21688.03	15.92	42980.08
income:age	6.26e+226	-21165.81	522.21	42019.89
AAD:smoking	2.73e+38	-21077.30	88.50	41855.82
age:smoking	2.97e+42	-20979.50	97.79	41638.24
income:smoking	322067.75	-20966.82	12.68	41593.15
AAD:region	28.06	-20963.49	3.33	41545.50

Note: This is a two-stage variable selection procedure. The first stage begins with a null model that includes only the related offset variable and performs variable selection among all main variables. The second stage starts with a null model that includes the selected main variables and then applies variable selection among potential two-way interaction terms. Note that non-smoker prevalence rates are used with a 20-year time lag.

A.3 Variable selection for breast cancer mortality

Table A.7: Variable selection procedure in the R-INLA software to determine the best fitted model for breast cancer mortality, without considering 'smoking proxy'.

variable added	Bayes factor	marginal likelihood	diff. in marginal likelihood	DIC
null		-10133.46		14460.45
age	Inf	-7120.84	3012.62	13686.88
year	3.31e+204	-6649.91	470.92	13093.15
region	13.28	-6647.33	2.58	13053.13

Note: The null model only includes the related offset variable; and age and year are defined to be categorical variables.

Table A.8: Variable selection procedure in the R-INLA software to determine the best fitted model for breast cancer mortality, with 'smoking proxy'.

variable added	Bayes factor	marginal likelihood	diff. in marginal likelihood	DIC
null		-10133.46		14460.46
age	Inf	-7120.84	3012.62	13686.90
year	3.32e+204	-6649.915	470.92	13093.17
smoking	1822.77	-6642.40	7.50	13075.32
region	25.68	-6639.16	3.24	13039.60

Note: The null model only includes the related offset variable; age and year are defined to be categorical variables; and non-smoker prevalence rates are used with 20-year time lag.

Appendix B

Parameter Estimates

B.1 Parameter estimates for female lung cancer based on the model in (3.4), without smoking data

Table B.1: Estimated coefficients for lung cancer mortality in women based on (3.4).

Covariate	Parameter	Mean	SD	%2.5	%97.5	Covariate	Parameter	Mean	SD	%2.5	%97.5
	β_0	-6.9190	0.0103	-6.9390	-6.8990		$\beta_{5,\text{deprivation}_5,\text{age}_6}$	0.0286	0.0124	0.0042	0.0535
Age	β_{1,age_1}	-1.9320	0.0101	-1.9510	-1.9110		$\beta_{5,\text{deprivation}_1,\text{age}_7}$	0.0623	0.0125	0.0368	0.0864
	β_{1,age_2}	-0.8800	0.0089	-0.8975	-0.8625		$\beta_{5,\text{deprivation}_2,\text{age}_7}$	0.0840	0.0143	0.0567	0.1129
	β_{1,age_3}	-0.3391	0.0075	-0.3536	-0.3242		$\beta_{5,\text{deprivation}_3,\text{age}_7}$	-0.1278	0.0227	-0.1715	-0.0829
	β_{1,age_4}	0.1138	0.0069	0.1002	0.1273		$\beta_{5,\text{deprivation}_4,\text{age}_7}$	-0.0995	0.0208	-0.1421	-0.0613
	β_{1,age_5}	0.4920	0.0063	0.4797	0.5044		$\beta_{5,\text{deprivation}_5,\text{age}_7}$	-0.0669	0.0172	-0.1002	-0.0327
	β_{1,age_6}	0.7546	0.0061	0.7426	0.7667		$\beta_{5,\text{deprivation}_1,\text{age}_8}$	-0.0867	0.0150	-0.1160	-0.0569
	β_{1,age_7}	0.8899	0.0064	0.8774	0.9023		$\beta_{5,\text{deprivation}_2,\text{age}_8}$	-0.0333	0.0151	-0.0635	-0.0035
	β_{1,age_8}	0.9004	0.0073	0.8859	0.9148		$\beta_{5,\text{deprivation}_3,\text{age}_8}$	0.0362	0.0134	0.0093	0.0619
	$\beta_{2,\text{region}_1}$	-0.4669	0.0502	-0.5902	-0.4065		$\beta_{5,\text{deprivation}_4,\text{age}_8}$	0.1367	0.0138	0.1089	0.1635
	$\beta_{2,\text{region}_2}$	0.1369	0.0074	0.1221	0.1518		$\beta_{5,\text{deprivation}_5,\text{age}_8}$	0.2413	0.0158	0.2100	0.2714
	$\beta_{2,\text{region}_3}$	-0.0989	0.0178	-0.1429	-0.0742		$\beta_{6,\text{region}_1,\text{age}_1}$	-0.1005	0.0315	-0.1621	-0.0386
	$\beta_{2,\text{region}_4}$	0.2099	0.0180	0.1828	0.2525		$\beta_{6,\text{region}_2,\text{age}_1}$	-0.0654	0.0278	-0.1208	-0.0110
	$\beta_{2,\text{region}_5}$	0.1144	0.0181	0.0885	0.1586		$\beta_{6,\text{region}_3,\text{age}_1}$	-0.0345	0.0243	-0.0834	0.0127
	$\beta_{2,\text{region}_6}$	0.1236	0.0162	0.0990	0.1632		$\beta_{6,\text{region}_4,\text{age}_1}$	-0.0336	0.0213	-0.0747	0.0091
	$\beta_{2,\text{region}_7}$	-0.8271	0.0434	-0.9339	-0.7753		$\beta_{6,\text{region}_5,\text{age}_1}$	0.0292	0.0194	-0.0089	0.0668
	$\beta_{2,\text{region}_8}$	0.3776	0.0290	0.3417	0.4493		$\beta_{6,\text{region}_6,\text{age}_1}$	0.0732	0.0186	0.0365	0.1091
	$\beta_{2,\text{region}_9}$	0.4305	0.0385	0.3851	0.5240		$\beta_{6,\text{region}_7,\text{age}_1}$	0.0606	0.0196	0.0217	0.0993
	$\beta_{3,\text{deprivation}_1}$	2.8620	0.1527	2.6850	3.2360		$\beta_{6,\text{region}_8,\text{age}_1}$	0.0709	0.0229	0.0272	0.1175
	$\beta_{3,\text{deprivation}_2}$	1.2020	0.0640	1.1270	1.3580		$\beta_{6,\text{region}_9,\text{age}_1}$	-0.0639	0.0224	-0.1086	-0.0190
	$\beta_{3,\text{deprivation}_3}$	-0.1165	0.0077	-0.1334	-0.1034		$\beta_{6,\text{region}_1,\text{age}_2}$	-0.0498	0.0212	-0.0924	-0.0088
	$\beta_{3,\text{deprivation}_4}$	-1.3390	0.0717	-1.5140	-1.2570		$\beta_{6,\text{region}_2,\text{age}_2}$	0.0332	0.0175	-0.0011	0.0672
	$\beta_{3,\text{deprivation}_5}$	-2.6090	0.1398	-2.9490	-2.4480		$\beta_{6,\text{region}_3,\text{age}_2}$	0.0256	0.0152	-0.0044	0.0549
	β_4	1.6740	0.1083	1.5460	1.9380		$\beta_{6,\text{region}_4,\text{age}_2}$	0.0319	0.0147	0.0037	0.0600
	$\beta_{5,\text{deprivation}_1,\text{age}_1}$	0.1262	0.0167	0.0924	0.1581		$\beta_{6,\text{region}_5,\text{age}_2}$	0.0313	0.0151	0.0017	0.0605
	$\beta_{5,\text{deprivation}_2,\text{age}_1}$	0.1162	0.0153	0.0873	0.1462		$\beta_{6,\text{region}_6,\text{age}_2}$	0.0070	0.0154	-0.0237	0.0376
	$\beta_{5,\text{deprivation}_3,\text{age}_1}$	0.0838	0.0136	0.0575	0.1099		$\beta_{6,\text{region}_7,\text{age}_2}$	-0.0152	0.0171	-0.0478	0.0180
	$\beta_{5,\text{deprivation}_4,\text{age}_1}$	0.0814	0.0122	0.0575	0.1053		$\beta_{6,\text{region}_8,\text{age}_2}$	-0.0935	0.0260	-0.1435	-0.0430
	$\beta_{5,\text{deprivation}_5,\text{age}_1}$	0.0125	0.0115	-0.0108	0.0343		$\beta_{6,\text{region}_9,\text{age}_2}$	0.0331	0.0222	-0.0098	0.0769
	$\beta_{5,\text{deprivation}_1,\text{age}_2}$	-0.0597	0.0114	-0.0821	-0.0370		$\beta_{6,\text{region}_1,\text{age}_3}$	0.0280	0.0195	-0.0119	0.0663
	$\beta_{5,\text{deprivation}_2,\text{age}_2}$	-0.1401	0.0114	-0.1621	-0.1166		$\beta_{6,\text{region}_2,\text{age}_3}$	0.0056	0.0180	-0.0310	0.0403
	$\beta_{5,\text{deprivation}_3,\text{age}_2}$	-0.2203	0.0137	-0.2471	-0.1938		$\beta_{6,\text{region}_3,\text{age}_3}$	0.0006	0.0161	-0.0302	0.0331
	$\beta_{5,\text{deprivation}_4,\text{age}_2}$	0.0615	0.0175	0.0261	0.0938		$\beta_{6,\text{region}_4,\text{age}_3}$	0.0136	0.0163	-0.0174	0.0457
	$\beta_{5,\text{deprivation}_5,\text{age}_2}$	0.0306	0.0166	-0.0032	0.0631		$\beta_{6,\text{region}_5,\text{age}_3}$	-0.0050	0.0170	-0.0378	0.0283
	$\beta_{5,\text{deprivation}_1,\text{age}_3}$	0.0422	0.0138	0.0152	0.0693		$\beta_{6,\text{region}_6,\text{age}_3}$	0.0177	0.0191	-0.0209	0.0551
	$\beta_{5,\text{deprivation}_2,\text{age}_3}$	0.0379	0.0125	0.0136	0.0628		$\beta_{6,\text{region}_7,\text{age}_3}$	0.0434	0.0296	-0.0160	0.1009
	$\beta_{5,\text{deprivation}_3,\text{age}_3}$	0.0276	0.0118	0.0049	0.0517		$\beta_{6,\text{region}_8,\text{age}_3}$	0.0382	0.0272	-0.0172	0.0882
	$\beta_{5,\text{deprivation}_4,\text{age}_3}$	-0.0130	0.0111	-0.0352	0.0082		$\beta_{6,\text{region}_9,\text{age}_3}$	0.0112	0.0231	-0.0343	0.0560
	$\beta_{5,\text{deprivation}_5,\text{age}_3}$	-0.0698	0.0113	-0.0917	-0.0473		$\beta_{6,\text{region}_1,\text{age}_4}$	0.0295	0.0194	-0.0090	0.0666
	$\beta_{5,\text{deprivation}_1,\text{age}_4}$	-0.1171	0.0131	-0.1428	-0.0915		$\beta_{6,\text{region}_2,\text{age}_4}$	0.0251	0.0188	-0.0117	0.0626
	$\beta_{5,\text{deprivation}_2,\text{age}_4}$	-0.0232	0.0193	-0.0610	0.0150		$\beta_{6,\text{region}_3,\text{age}_4}$	-0.0136	0.0181	-0.0490	0.0219
	$\beta_{5,\text{deprivation}_3,\text{age}_4}$	0.0052	0.0171	-0.0274	0.0404		$\beta_{6,\text{region}_4,\text{age}_4}$	-0.0385	0.0188	-0.0742	-0.0009
	$\beta_{5,\text{deprivation}_4,\text{age}_4}$	0.0132	0.0152	-0.0158	0.0437		$\beta_{6,\text{region}_5,\text{age}_4}$	-0.0954	0.0217	-0.1378	-0.0534
	$\beta_{5,\text{deprivation}_5,\text{age}_4}$	-0.0119	0.0131	-0.0378	0.0130		$\beta_{6,\text{region}_6,\text{age}_4}$	0.0034	0.0284	-0.0516	0.0572
	$\beta_{5,\text{deprivation}_1,\text{age}_5}$	-0.0141	0.0124	-0.0399	0.0103		$\beta_{6,\text{region}_7,\text{age}_4}$	0.0274	0.0267	-0.0257	0.0774
	$\beta_{5,\text{deprivation}_2,\text{age}_5}$	0.0079	0.0123	-0.0180	0.0310		$\beta_{6,\text{region}_8,\text{age}_4}$	-0.0294	0.0219	-0.0718	0.0138
	$\beta_{5,\text{deprivation}_3,\text{age}_5}$	0.0109	0.0126	-0.0151	0.0352		$\beta_{6,\text{region}_9,\text{age}_4}$	0.0091	0.0197	-0.0291	0.0490
	$\beta_{5,\text{deprivation}_4,\text{age}_5}$	0.0121	0.0142	-0.0157	0.0403		$\beta_{6,\text{region}_1,\text{age}_5}$	0.0006	0.0172	-0.0325	0.0332
	$\beta_{5,\text{deprivation}_5,\text{age}_5}$	-0.0367	0.0212	-0.0794	0.0037		$\beta_{6,\text{region}_2,\text{age}_5}$	0.0120	0.0177	-0.0230	0.0464
	$\beta_{5,\text{deprivation}_1,\text{age}_6}$	-0.0524	0.0179	-0.0870	-0.0164		$\beta_{6,\text{region}_3,\text{age}_5}$	0.0077	0.0181	-0.0277	0.0432
	$\beta_{5,\text{deprivation}_2,\text{age}_6}$	-0.0723	0.0159	-0.1034	-0.0416		$\beta_{6,\text{region}_4,\text{age}_5}$	-0.0308	0.0208	-0.0720	0.0091
	$\beta_{5,\text{deprivation}_3,\text{age}_6}$	-0.0208	0.0135	-0.0474	0.0067		$\beta_{6,\text{region}_5,\text{age}_5}$	0.0277	0.0282	-0.0273	0.0854
	$\beta_{5,\text{deprivation}_4,\text{age}_6}$	0.0073	0.0128	-0.0180	0.0321		$\beta_{6,\text{region}_6,\text{age}_5}$	0.0332	0.0233	-0.0126	0.0790

Covariate	Parameter	Mean	SD	%2.5	%97.5	Covariate	Parameter	Mean	SD	%2.5	%97.5
	$\beta_{6,\text{region}_7,\text{age}_5}$	0.0027	0.0210	-0.0393	0.0426		$\kappa^*_{1,\text{year}_7}$	0.0064	0.0655	-0.1247	0.1326
	$\beta_{6,\text{region}_8,\text{age}_5}$	0.0042	0.0191	-0.0332	0.0414		$\kappa^*_{1,\text{year}_8}$	0.0084	0.0701	-0.1349	0.1472
	$\beta_{6,\text{region}_9,\text{age}_5}$	-0.0048	0.0182	-0.0420	0.0305		$\kappa^*_{1,\text{year}_9}$	0.0108	0.0746	-0.1405	0.1622
	$\beta_{6,\text{region}_1,\text{age}_6}$	-0.0336	0.0173	-0.0669	0.0011		$\kappa^*_{1,\text{year}_{10}}$	0.0120	0.0785	-0.1472	0.1680
	$\beta_{6,\text{region}_2,\text{age}_6}$	-0.0023	0.0180	-0.0376	0.0337		$\kappa^*_{1,\text{year}_{11}}$	0.0130	0.0834	-0.1547	0.1847
	$\beta_{6,\text{region}_3,\text{age}_6}$	-0.0270	0.0197	-0.0638	0.0129		$\kappa^*_{1,\text{year}_{12}}$	0.0144	0.0891	-0.1605	0.1928
	$\beta_{6,\text{region}_4,\text{age}_6}$	-0.1037	0.0260	-0.1547	-0.0534		$\kappa^*_{1,\text{year}_{13}}$	0.0159	0.0937	-0.1641	0.2049
	$\beta_{6,\text{region}_5,\text{age}_6}$	-0.1159	0.0243	-0.1636	-0.0692		$\kappa^*_{1,\text{year}_{14}}$	0.0172	0.0982	-0.1779	0.2129
	$\beta_{6,\text{region}_6,\text{age}_6}$	-0.0607	0.0214	-0.1028	-0.0188		$\kappa^*_{1,\text{year}_{15}}$	0.0192	0.1028	-0.1802	0.2266
	$\beta_{6,\text{region}_7,\text{age}_6}$	-0.0198	0.0191	-0.0573	0.0181		$\kappa^*_{1,\text{year}_{16}}$	0.0201	0.1059	-0.1895	0.2401
	$\beta_{6,\text{region}_8,\text{age}_6}$	-0.0118	0.0183	-0.0470	0.0242		$\kappa^*_{1,\text{year}_{17}}$	0.0203	0.1099	-0.2009	0.2420
	$\beta_{6,\text{region}_9,\text{age}_6}$	0.0333	0.0172	0.0004	0.0685		$\kappa^*_{1,\text{year}_{18}}$	0.0228	0.1139	-0.2062	0.2535
	$\beta_{6,\text{region}_1,\text{age}_7}$	0.1129	0.0177	0.0774	0.1486		κ_{2,year_2}	-0.0172	0.0113	-0.0399	0.0054
	$\beta_{6,\text{region}_2,\text{age}_7}$	0.1658	0.0202	0.1262	0.2060		κ_{2,year_3}	-0.0205	0.0126	-0.0454	0.0056
	$\beta_{6,\text{region}_3,\text{age}_7}$	0.0848	0.0248	0.0353	0.1345		κ_{2,year_4}	-0.0245	0.0141	-0.0515	0.0047
	$\beta_{6,\text{region}_4,\text{age}_7}$	0.0410	0.0231	-0.0052	0.0840		κ_{2,year_5}	-0.0177	0.0140	-0.0437	0.0113
	$\beta_{6,\text{region}_5,\text{age}_7}$	0.0339	0.0191	-0.0029	0.0711		κ_{2,year_6}	-0.0275	0.0141	-0.0534	-0.0001
	$\beta_{6,\text{region}_6,\text{age}_7}$	-0.0037	0.0172	-0.0373	0.0304		κ_{2,year_7}	-0.0192	0.0138	-0.0458	0.0072
	$\beta_{6,\text{region}_7,\text{age}_7}$	-0.0120	0.0162	-0.0435	0.0200		κ_{2,year_8}	-0.0335	0.0132	-0.0573	-0.0058
	$\beta_{6,\text{region}_8,\text{age}_7}$	-0.0431	0.0160	-0.0745	-0.0116		κ_{2,year_9}	-0.0362	0.0130	-0.0600	-0.0094
	$\beta_{6,\text{region}_9,\text{age}_7}$	-0.0655	0.0167	-0.0984	-0.0324		$\kappa_{2,\text{year}_{10}}$	-0.0295	0.0135	-0.0545	-0.0009
	$\beta_{6,\text{region}_1,\text{age}_8}$	-0.0354	0.0181	-0.0712	0.0000		$\kappa_{2,\text{year}_{11}}$	-0.0508	0.0136	-0.0768	-0.0244
	$\beta_{6,\text{region}_2,\text{age}_8}$	0.2024	0.0282	0.1475	0.2567		$\kappa_{2,\text{year}_{12}}$	-0.0580	0.0136	-0.0840	-0.0312
	$\beta_{6,\text{region}_3,\text{age}_8}$	0.0582	0.0260	0.0051	0.1067		$\kappa_{2,\text{year}_{13}}$	-0.0643	0.0135	-0.0922	-0.0392
	$\beta_{6,\text{region}_4,\text{age}_8}$	0.0157	0.0226	-0.0283	0.0610		$\kappa_{2,\text{year}_{14}}$	-0.0649	0.0131	-0.0907	-0.0385
	$\beta_{6,\text{region}_5,\text{age}_8}$	-0.0169	0.0192	-0.0539	0.0215		$\kappa_{2,\text{year}_{15}}$	-0.0751	0.0133	-0.1007	-0.0478
	$\beta_{6,\text{region}_6,\text{age}_8}$	-0.0588	0.0186	-0.0949	-0.0233		$\kappa_{2,\text{year}_{16}}$	-0.0779	0.0135	-0.1041	-0.0505
	$\beta_{6,\text{region}_7,\text{age}_8}$	-0.0730	0.0180	-0.1083	-0.0363		$\kappa_{2,\text{year}_{17}}$	-0.0872	0.0138	-0.1144	-0.0600
	$\beta_{6,\text{region}_8,\text{age}_8}$	-0.0768	0.0189	-0.1140	-0.0406		$\kappa_{2,\text{year}_{18}}$	-0.0924	0.0143	-0.1203	-0.0644
	$\beta_{6,\text{region}_9,\text{age}_8}$	-0.0506	0.0206	-0.0917	-0.0095		$\kappa^*_{2,\text{year}_1}$	-0.0941	0.0224	-0.1385	-0.0484
	κ_{1,year_2}	0.0085	0.0124	-0.0134	0.0354		$\kappa^*_{2,\text{year}_2}$	-0.0960	0.0287	-0.1517	-0.0405
	κ_{1,year_3}	0.0197	0.0129	-0.0060	0.0447		$\kappa^*_{2,\text{year}_3}$	-0.0976	0.0340	-0.1639	-0.0292
	κ_{1,year_4}	0.0113	0.0132	-0.0130	0.0372		$\kappa^*_{2,\text{year}_4}$	-0.0993	0.0388	-0.1738	-0.0230
	κ_{1,year_5}	0.0376	0.0138	0.0101	0.0644		$\kappa^*_{2,\text{year}_5}$	-0.1021	0.0442	-0.1858	-0.0112
	κ_{1,year_6}	0.0676	0.0144	0.0390	0.0964		$\kappa^*_{2,\text{year}_6}$	-0.1039	0.0489	-0.1982	-0.0075
	κ_{1,year_7}	0.0887	0.0139	0.0601	0.1142		$\kappa^*_{2,\text{year}_7}$	-0.1054	0.0523	-0.2065	0.0041
	κ_{1,year_8}	0.1019	0.0137	0.0753	0.1267		$\kappa^*_{2,\text{year}_8}$	-0.1079	0.0560	-0.2155	0.0072
	κ_{1,year_9}	0.0922	0.0129	0.0664	0.1179		$\kappa^*_{2,\text{year}_9}$	-0.1098	0.0603	-0.2273	0.0163
	$\kappa_{1,\text{year}_{10}}$	0.0969	0.0140	0.0702	0.1243		$\kappa^*_{2,\text{year}_{10}}$	-0.1114	0.0644	-0.2328	0.0265
	$\kappa_{1,\text{year}_{11}}$	0.0932	0.0141	0.0650	0.1204		$\kappa^*_{2,\text{year}_{11}}$	-0.1134	0.0675	-0.2399	0.0259
	$\kappa_{1,\text{year}_{12}}$	0.0917	0.0138	0.0651	0.1203		$\kappa^*_{2,\text{year}_{12}}$	-0.1151	0.0704	-0.2503	0.0311
	$\kappa_{1,\text{year}_{13}}$	0.0832	0.0142	0.0567	0.1102		$\kappa^*_{2,\text{year}_{13}}$	-0.1172	0.0738	-0.2591	0.0344
	$\kappa_{1,\text{year}_{14}}$	0.0801	0.0137	0.0531	0.1078		$\kappa^*_{2,\text{year}_{14}}$	-0.1191	0.0767	-0.2682	0.0377
	$\kappa_{1,\text{year}_{15}}$	0.0706	0.0136	0.0434	0.0964		$\kappa^*_{2,\text{year}_{15}}$	-0.1217	0.0805	-0.2796	0.0396
	$\kappa_{1,\text{year}_{16}}$	0.0522	0.0136	0.0251	0.0786		$\kappa^*_{2,\text{year}_{16}}$	-0.1231	0.0834	-0.2797	0.0435
	$\kappa_{1,\text{year}_{17}}$	0.0257	0.0137	-0.0016	0.0527		$\kappa^*_{2,\text{year}_{17}}$	-0.1248	0.0875	-0.2911	0.0511
	$\kappa_{1,\text{year}_{18}}$	-0.0061	0.0142	-0.0332	0.0219		$\kappa^*_{2,\text{year}_{18}}$	-0.1265	0.0915	-0.2967	0.0594
	$\kappa^*_{1,\text{year}_1}$	-0.0044	0.0271	-0.0605	0.0485		ψ_{κ_1}	0.0025	0.0054	-0.0048	0.0167
	$\kappa^*_{1,\text{year}_2}$	-0.0031	0.0365	-0.0771	0.0682		σ^2	0.0078	0.0007	0.0066	0.0092
	$\kappa^*_{1,\text{year}_3}$	-0.0007	0.0438	-0.0871	0.0848		$\sigma^2_{\psi_{\kappa_2}}$	0.0000	0.0000	0.0000	0.0000
	$\kappa^*_{1,\text{year}_4}$	0.0014	0.0495	-0.0969	0.0993		$\sigma^2_{\psi_{\kappa_1}}$	0.0000	0.0000	0.0000	0.0001
	$\kappa^*_{1,\text{year}_5}$	0.0023	0.0555	-0.1137	0.1105		$\sigma^2_{\kappa_1}$	0.0005	0.0002	0.0003	0.0011
	$\kappa^*_{1,\text{year}_6}$	0.0043	0.0609	-0.1221	0.1280		$\sigma^2_{\kappa_2}$	0.0003	0.0001	0.0001	0.0007

B.2 Parameter estimates for female lung cancer based on the model in (3.5), with smoking data

Table B.2: Estimated coefficients for lung cancer mortality in women based on (3.5).

Covariate Parameter	Mean	SD	%2.5	%97.5	Covariate Parameter	Mean	SD	%2.5	%97.5
β_0	-7.2760	0.0264	-7.3240	-7.2260	$\beta_{5,\text{deprivation}_2,\text{age}_7}$	0.0859	0.0142	0.0573	0.1146
β_{1,age_1}	-2.4300	0.0339	-2.4960	-2.3670	$\beta_{5,\text{deprivation}_3,\text{age}_7}$	-0.1328	0.0239	-0.1826	-0.0895
β_{1,age_2}	-1.0340	0.0132	-1.0590	-1.0080	$\beta_{5,\text{deprivation}_4,\text{age}_7}$	-0.0941	0.0185	-0.1306	-0.0564
β_{1,age_3}	-0.4921	0.0121	-0.5149	-0.4687	$\beta_{5,\text{deprivation}_5,\text{age}_7}$	-0.0649	0.0165	-0.0955	-0.0302
β_{1,age_4}	-0.0717	0.0132	-0.0973	-0.0458	$\beta_{5,\text{deprivation}_1,\text{age}_8}$	-0.0852	0.0145	-0.1137	-0.0549
β_{1,age_5}	0.3056	0.0133	0.2810	0.3316	$\beta_{5,\text{deprivation}_2,\text{age}_8}$	-0.0311	0.0137	-0.0575	-0.0029
β_{1,age_6}	1.1480	0.0253	1.1010	1.1970	$\beta_{5,\text{deprivation}_3,\text{age}_8}$	0.0351	0.0140	0.0089	0.0631
β_{1,age_7}	1.2830	0.0257	1.2340	1.3330	$\beta_{5,\text{deprivation}_4,\text{age}_8}$	0.1340	0.0141	0.1071	0.1613
β_{1,age_8}	1.2920	0.0265	1.2410	1.3430	$\beta_{5,\text{deprivation}_5,\text{age}_8}$	0.2390	0.0160	0.2081	0.2696
$\beta_{2,\text{region}_1}$	-0.1256	0.0386	-0.1850	-0.0552	$\beta_{6,\text{region}_1,\text{age}_1}$	-0.1074	0.0322	-0.1710	-0.0481
$\beta_{2,\text{region}_2}$	0.1451	0.0073	0.1310	0.1592	$\beta_{6,\text{region}_2,\text{age}_1}$	-0.0668	0.0279	-0.1201	-0.0084
$\beta_{2,\text{region}_3}$	0.0152	0.0145	-0.0096	0.0437	$\beta_{6,\text{region}_3,\text{age}_1}$	-0.0325	0.0243	-0.0784	0.0160
$\beta_{2,\text{region}_4}$	0.1079	0.0154	0.0780	0.1358	$\beta_{6,\text{region}_4,\text{age}_1}$	-0.0331	0.0206	-0.0747	0.0063
$\beta_{2,\text{region}_5}$	0.0113	0.0153	-0.0175	0.0378	$\beta_{6,\text{region}_5,\text{age}_1}$	0.0299	0.0192	-0.0067	0.0664
$\beta_{2,\text{region}_6}$	0.0305	0.0136	0.0020	0.0534	$\beta_{6,\text{region}_6,\text{age}_1}$	0.0750	0.0182	0.0382	0.1086
$\beta_{2,\text{region}_7}$	-0.5302	0.0344	-0.5821	-0.4674	$\beta_{6,\text{region}_7,\text{age}_1}$	0.0618	0.0192	0.0264	0.0992
$\beta_{2,\text{region}_8}$	0.1846	0.0228	0.1423	0.2207	$\beta_{6,\text{region}_8,\text{age}_1}$	0.0731	0.0239	0.0264	0.1169
$\beta_{2,\text{region}_9}$	0.1611	0.0293	0.1088	0.2087	$\beta_{6,\text{region}_9,\text{age}_1}$	-0.0718	0.0229	-0.1184	-0.0293
$\beta_{3,\text{deprivation}_1}$	1.8050	0.1168	1.5980	1.9890	$\beta_{6,\text{region}_1,\text{age}_2}$	-0.0535	0.0200	-0.0918	-0.0136
$\beta_{3,\text{deprivation}_2}$	0.7603	0.0488	0.6740	0.8345	$\beta_{6,\text{region}_2,\text{age}_2}$	0.0328	0.0175	-0.0020	0.0657
$\beta_{3,\text{deprivation}_3}$	-0.0806	0.0070	-0.0937	-0.0668	$\beta_{6,\text{region}_3,\text{age}_2}$	0.0255	0.0165	-0.0068	0.0573
$\beta_{3,\text{deprivation}_4}$	-0.8448	0.0547	-0.9253	-0.7480	$\beta_{6,\text{region}_4,\text{age}_2}$	0.0339	0.0147	0.0049	0.0623
$\beta_{3,\text{deprivation}_5}$	-1.6400	0.1067	-1.8030	-1.4500	$\beta_{6,\text{region}_5,\text{age}_2}$	0.0355	0.0146	0.0063	0.0641
β_4	0.9437	0.0803	0.8022	1.0700	$\beta_{6,\text{region}_6,\text{age}_2}$	0.0102	0.0152	-0.0185	0.0399
$\beta_{5,\text{deprivation}_1,\text{age}_1}$	0.1318	0.0167	0.1005	0.1634	$\beta_{6,\text{region}_7,\text{age}_2}$	-0.0127	0.0173	-0.0457	0.0228
$\beta_{5,\text{deprivation}_2,\text{age}_1}$	0.1150	0.0150	0.0869	0.1455	$\beta_{6,\text{region}_8,\text{age}_2}$	-0.0956	0.0264	-0.1475	-0.0420
$\beta_{5,\text{deprivation}_3,\text{age}_1}$	0.0810	0.0132	0.0551	0.1069	$\beta_{6,\text{region}_9,\text{age}_2}$	0.0341	0.0216	-0.0098	0.0775
$\beta_{5,\text{deprivation}_4,\text{age}_1}$	0.0797	0.0118	0.0563	0.1028	$\beta_{6,\text{region}_1,\text{age}_3}$	0.0281	0.0194	-0.0122	0.0655
$\beta_{5,\text{deprivation}_5,\text{age}_1}$	0.0109	0.0116	-0.0109	0.0347	$\beta_{6,\text{region}_2,\text{age}_3}$	0.0046	0.0174	-0.0298	0.0394
$\beta_{5,\text{deprivation}_1,\text{age}_2}$	-0.0592	0.0111	-0.0810	-0.0365	$\beta_{6,\text{region}_3,\text{age}_3}$	-0.0002	0.0160	-0.0323	0.0312
$\beta_{5,\text{deprivation}_2,\text{age}_2}$	-0.1391	0.0114	-0.1616	-0.1171	$\beta_{6,\text{region}_4,\text{age}_3}$	0.0144	0.0153	-0.0162	0.0441
$\beta_{5,\text{deprivation}_3,\text{age}_2}$	-0.2201	0.0135	-0.2458	-0.1924	$\beta_{6,\text{region}_5,\text{age}_3}$	-0.0042	0.0164	-0.0378	0.0269
$\beta_{5,\text{deprivation}_4,\text{age}_2}$	0.0635	0.0177	0.0296	0.0973	$\beta_{6,\text{region}_6,\text{age}_3}$	0.0186	0.0188	-0.0167	0.0575
$\beta_{5,\text{deprivation}_5,\text{age}_2}$	0.0298	0.0159	-0.0020	0.0608	$\beta_{6,\text{region}_7,\text{age}_3}$	0.0476	0.0299	-0.0083	0.1060
$\beta_{5,\text{deprivation}_1,\text{age}_3}$	0.0403	0.0140	0.0127	0.0674	$\beta_{6,\text{region}_8,\text{age}_3}$	0.0372	0.0281	-0.0215	0.0929
$\beta_{5,\text{deprivation}_2,\text{age}_3}$	0.0371	0.0119	0.0135	0.0591	$\beta_{6,\text{region}_9,\text{age}_3}$	0.0095	0.0221	-0.0307	0.0539
$\beta_{5,\text{deprivation}_3,\text{age}_3}$	0.0263	0.0118	0.0026	0.0499	$\beta_{6,\text{region}_1,\text{age}_4}$	0.0305	0.0200	-0.0092	0.0698
$\beta_{5,\text{deprivation}_4,\text{age}_3}$	-0.0126	0.0115	-0.0353	0.0081	$\beta_{6,\text{region}_2,\text{age}_4}$	0.0261	0.0180	-0.0077	0.0588
$\beta_{5,\text{deprivation}_5,\text{age}_3}$	-0.0684	0.0121	-0.0930	-0.0440	$\beta_{6,\text{region}_3,\text{age}_4}$	-0.0153	0.0169	-0.0481	0.0172
$\beta_{5,\text{deprivation}_1,\text{age}_4}$	-0.1159	0.0135	-0.1423	-0.0889	$\beta_{6,\text{region}_4,\text{age}_4}$	-0.0403	0.0190	-0.0764	-0.0021
$\beta_{5,\text{deprivation}_2,\text{age}_4}$	-0.0237	0.0199	-0.0624	0.0146	$\beta_{6,\text{region}_5,\text{age}_4}$	-0.0953	0.0221	-0.1360	-0.0514
$\beta_{5,\text{deprivation}_3,\text{age}_4}$	0.0030	0.0180	-0.0340	0.0378	$\beta_{6,\text{region}_6,\text{age}_4}$	0.0052	0.0272	-0.0475	0.0546
$\beta_{5,\text{deprivation}_4,\text{age}_4}$	0.0151	0.0147	-0.0141	0.0434	$\beta_{6,\text{region}_7,\text{age}_4}$	0.0301	0.0285	-0.0256	0.0828
$\beta_{5,\text{deprivation}_5,\text{age}_4}$	-0.0107	0.0137	-0.0367	0.0181	$\beta_{6,\text{region}_8,\text{age}_4}$	-0.0267	0.0217	-0.0687	0.0176
$\beta_{5,\text{deprivation}_1,\text{age}_5}$	-0.0143	0.0125	-0.0393	0.0098	$\beta_{6,\text{region}_9,\text{age}_4}$	0.0088	0.0192	-0.0282	0.0471
$\beta_{5,\text{deprivation}_2,\text{age}_5}$	0.0087	0.0116	-0.0139	0.0308	$\beta_{6,\text{region}_1,\text{age}_5}$	-0.0006	0.0168	-0.0320	0.0316
$\beta_{5,\text{deprivation}_3,\text{age}_5}$	0.0108	0.0124	-0.0123	0.0351	$\beta_{6,\text{region}_2,\text{age}_5}$	0.0109	0.0170	-0.0204	0.0443
$\beta_{5,\text{deprivation}_4,\text{age}_5}$	0.0111	0.0139	-0.0155	0.0368	$\beta_{6,\text{region}_3,\text{age}_5}$	0.0088	0.0179	-0.0261	0.0433
$\beta_{5,\text{deprivation}_5,\text{age}_5}$	-0.0388	0.0207	-0.0820	0.0012	$\beta_{6,\text{region}_4,\text{age}_5}$	-0.0365	0.0215	-0.0781	0.0056
$\beta_{5,\text{deprivation}_1,\text{age}_6}$	-0.0537	0.0183	-0.0893	-0.0169	$\beta_{6,\text{region}_5,\text{age}_5}$	0.0282	0.0304	-0.0336	0.0866
$\beta_{5,\text{deprivation}_2,\text{age}_6}$	-0.0714	0.0151	-0.1018	-0.0433	$\beta_{6,\text{region}_6,\text{age}_5}$	0.0347	0.0250	-0.0139	0.0845
$\beta_{5,\text{deprivation}_3,\text{age}_6}$	-0.0209	0.0133	-0.0459	0.0046	$\beta_{6,\text{region}_7,\text{age}_5}$	0.0055	0.0211	-0.0382	0.0485
$\beta_{5,\text{deprivation}_4,\text{age}_6}$	0.0082	0.0130	-0.0176	0.0325	$\beta_{6,\text{region}_8,\text{age}_5}$	0.0028	0.0187	-0.0339	0.0393
$\beta_{5,\text{deprivation}_5,\text{age}_6}$	0.0280	0.0125	0.0026	0.0522	$\beta_{6,\text{region}_9,\text{age}_5}$	-0.0058	0.0183	-0.0423	0.0295
$\beta_{5,\text{deprivation}_1,\text{age}_7}$	0.0626	0.0122	0.0392	0.0874	$\beta_{6,\text{region}_1,\text{age}_6}$	-0.0362	0.0174	-0.0700	-0.0036

Covariate	Parameter	Mean	SD	%2.5	%97.5	Covariate	Parameter	Mean	SD	%2.5	%97.5
	$\beta_{6,\text{region}_2,\text{age}_6}$	-0.0016	0.0167	-0.0339	0.0316		$\kappa^*_{1,\text{year}_3}$	0.3224	0.0532	0.2238	0.4296
	$\beta_{6,\text{region}_3,\text{age}_6}$	-0.0274	0.0212	-0.0721	0.0120		$\kappa^*_{1,\text{year}_4}$	0.3353	0.0606	0.2211	0.4599
	$\beta_{6,\text{region}_4,\text{age}_6}$	-0.0904	0.0253	-0.1400	-0.0406		$\kappa^*_{1,\text{year}_5}$	0.3472	0.0655	0.2217	0.4758
	$\beta_{6,\text{region}_5,\text{age}_6}$	-0.1121	0.0240	-0.1575	-0.0633		$\kappa^*_{1,\text{year}_6}$	0.3596	0.0721	0.2189	0.5008
	$\beta_{6,\text{region}_6,\text{age}_6}$	-0.0584	0.0221	-0.1020	-0.0153		$\kappa^*_{1,\text{year}_7}$	0.3695	0.0796	0.2139	0.5235
	$\beta_{6,\text{region}_7,\text{age}_6}$	-0.0204	0.0191	-0.0569	0.0158		$\kappa^*_{1,\text{year}_8}$	0.3822	0.0854	0.2199	0.5491
	$\beta_{6,\text{region}_8,\text{age}_6}$	-0.0135	0.0167	-0.0482	0.0185		$\kappa^*_{1,\text{year}_9}$	0.3936	0.0915	0.2164	0.5855
	$\beta_{6,\text{region}_9,\text{age}_6}$	0.0310	0.0164	-0.0001	0.0641		$\kappa^*_{1,\text{year}_{10}}$	0.4060	0.0970	0.2287	0.6055
	$\beta_{6,\text{region}_1,\text{age}_7}$	0.1050	0.0175	0.0692	0.1369		$\kappa^*_{1,\text{year}_{11}}$	0.4174	0.1028	0.2307	0.6234
	$\beta_{6,\text{region}_2,\text{age}_7}$	0.1588	0.0204	0.1195	0.1991		$\kappa^*_{1,\text{year}_{12}}$	0.4292	0.1059	0.2389	0.6426
	$\beta_{6,\text{region}_3,\text{age}_7}$	0.0845	0.0238	0.0360	0.1307		$\kappa^*_{1,\text{year}_{13}}$	0.4405	0.1112	0.2390	0.6685
	$\beta_{6,\text{region}_4,\text{age}_7}$	0.0423	0.0221	0.0005	0.0867		$\kappa^*_{1,\text{year}_{14}}$	0.4509	0.1160	0.2306	0.6871
	$\beta_{6,\text{region}_5,\text{age}_7}$	0.0304	0.0188	-0.0089	0.0665		$\kappa^*_{1,\text{year}_{15}}$	0.4623	0.1203	0.2380	0.7114
	$\beta_{6,\text{region}_6,\text{age}_7}$	-0.0044	0.0165	-0.0358	0.0283		$\kappa^*_{1,\text{year}_{16}}$	0.4728	0.1240	0.2419	0.7243
	$\beta_{6,\text{region}_7,\text{age}_7}$	-0.0120	0.0162	-0.0433	0.0199		$\kappa^*_{1,\text{year}_{17}}$	0.4856	0.1271	0.2570	0.7562
	$\beta_{6,\text{region}_8,\text{age}_7}$	-0.0449	0.0157	-0.0749	-0.0154		$\kappa^*_{1,\text{year}_{18}}$	0.4956	0.1324	0.2544	0.7788
	$\beta_{6,\text{region}_9,\text{age}_7}$	-0.0635	0.0158	-0.0944	-0.0326		κ_{2,year_2}	-0.0178	0.0123	-0.0431	0.0049
	$\beta_{6,\text{region}_1,\text{age}_8}$	-0.0325	0.0175	-0.0665	0.0015		κ_{2,year_3}	-0.0189	0.0138	-0.0481	0.0090
	$\beta_{6,\text{region}_2,\text{age}_8}$	0.1998	0.0278	0.1482	0.2558		κ_{2,year_4}	-0.0251	0.0123	-0.0512	-0.0031
	$\beta_{6,\text{region}_3,\text{age}_8}$	0.0540	0.0254	0.0023	0.1022		κ_{2,year_5}	-0.0185	0.0129	-0.0425	0.0064
	$\beta_{6,\text{region}_4,\text{age}_8}$	0.0112	0.0220	-0.0292	0.0559		κ_{2,year_6}	-0.0268	0.0136	-0.0527	0.0010
	$\beta_{6,\text{region}_5,\text{age}_8}$	-0.0144	0.0193	-0.0520	0.0243		κ_{2,year_7}	-0.0200	0.0136	-0.0466	0.0057
	$\beta_{6,\text{region}_6,\text{age}_8}$	-0.0580	0.0181	-0.0923	-0.0219		κ_{2,year_8}	-0.0339	0.0136	-0.0593	-0.0076
	$\beta_{6,\text{region}_7,\text{age}_8}$	-0.0703	0.0182	-0.1059	-0.0338		κ_{2,year_9}	-0.0373	0.0143	-0.0672	-0.0100
	$\beta_{6,\text{region}_8,\text{age}_8}$	-0.0763	0.0191	-0.1114	-0.0373		$\kappa_{2,\text{year}_{10}}$	-0.0313	0.0138	-0.0589	-0.0062
	$\beta_{6,\text{region}_9,\text{age}_8}$	-0.0460	0.0210	-0.0862	-0.0023		$\kappa_{2,\text{year}_{11}}$	-0.0521	0.0134	-0.0801	-0.0268
	$\beta_{7,\text{region}_1}$	-0.0294	0.0098	-0.0480	-0.0103		$\kappa_{2,\text{year}_{12}}$	-0.0604	0.0131	-0.0861	-0.0358
	$\beta_{7,\text{region}_2}$	-0.0461	0.0068	-0.0598	-0.0326		$\kappa_{2,\text{year}_{13}}$	-0.0655	0.0141	-0.0946	-0.0384
	$\beta_{7,\text{region}_3}$	-0.0128	0.0070	-0.0267	0.0008		$\kappa_{2,\text{year}_{14}}$	-0.0664	0.0141	-0.0965	-0.0394
	$\beta_{7,\text{region}_4}$	0.0103	0.0085	-0.0065	0.0271		$\kappa_{2,\text{year}_{15}}$	-0.0774	0.0135	-0.1044	-0.0511
	$\beta_{7,\text{region}_5}$	0.0091	0.0076	-0.0058	0.0242		$\kappa_{2,\text{year}_{16}}$	-0.0811	0.0137	-0.1090	-0.0542
	$\beta_{7,\text{region}_6}$	0.0344	0.0085	0.0180	0.0518		$\kappa_{2,\text{year}_{17}}$	-0.0912	0.0140	-0.1199	-0.0643
	$\beta_{7,\text{region}_7}$	0.0741	0.0081	0.0584	0.0893		$\kappa_{2,\text{year}_{18}}$	-0.0975	0.0144	-0.1257	-0.0708
	$\beta_{7,\text{region}_8}$	-0.0145	0.0076	-0.0289	-0.0000		$\kappa^*_{2,\text{year}_1}$	-0.0997	0.0226	-0.1451	-0.0578
	$\beta_{7,\text{region}_9}$	-0.0252	0.0090	-0.0427	-0.0080		$\kappa^*_{2,\text{year}_2}$	-0.1022	0.0293	-0.1604	-0.0465
	β_8	-0.3542	0.0223	-0.3958	-0.3101		$\kappa^*_{2,\text{year}_3}$	-0.1053	0.0356	-0.1818	-0.0373
	κ_{1,year_2}	0.0235	0.0146	-0.0104	0.0471		$\kappa^*_{2,\text{year}_4}$	-0.1077	0.0411	-0.1975	-0.0262
	κ_{1,year_3}	0.0533	0.0162	0.0204	0.0830		$\kappa^*_{2,\text{year}_5}$	-0.1099	0.0459	-0.2049	-0.0220
	κ_{1,year_4}	0.0593	0.0159	0.0264	0.0871		$\kappa^*_{2,\text{year}_6}$	-0.1126	0.0510	-0.2145	-0.0076
	κ_{1,year_5}	0.1041	0.0163	0.0709	0.1347		$\kappa^*_{2,\text{year}_7}$	-0.1155	0.0550	-0.2309	-0.0057
	κ_{1,year_6}	0.1529	0.0147	0.1207	0.1791		$\kappa^*_{2,\text{year}_8}$	-0.1175	0.0590	-0.2329	0.0047
	κ_{1,year_7}	0.1924	0.0160	0.1578	0.2218		$\kappa^*_{2,\text{year}_9}$	-0.1198	0.0638	-0.2485	0.0106
	κ_{1,year_8}	0.2208	0.0165	0.1849	0.2482		$\kappa^*_{2,\text{year}_{10}}$	-0.1224	0.0678	-0.2559	0.0152
	κ_{1,year_9}	0.2292	0.0178	0.1940	0.2606		$\kappa^*_{2,\text{year}_{11}}$	-0.1263	0.0721	-0.2719	0.0126
	$\kappa_{1,\text{year}_{10}}$	0.2505	0.0181	0.2156	0.2848		$\kappa^*_{2,\text{year}_{12}}$	-0.1278	0.0760	-0.2891	0.0153
	$\kappa_{1,\text{year}_{11}}$	0.2641	0.0193	0.2258	0.3002		$\kappa^*_{2,\text{year}_{13}}$	-0.1312	0.0802	-0.2892	0.0197
	$\kappa_{1,\text{year}_{12}}$	0.2813	0.0200	0.2412	0.3181		$\kappa^*_{2,\text{year}_{14}}$	-0.1341	0.0845	-0.3026	0.0290
	$\kappa_{1,\text{year}_{13}}$	0.2906	0.0202	0.2498	0.3309		$\kappa^*_{2,\text{year}_{15}}$	-0.1363	0.0881	-0.3160	0.0328
	$\kappa_{1,\text{year}_{14}}$	0.3061	0.0208	0.2665	0.3448		$\kappa^*_{2,\text{year}_{16}}$	-0.1388	0.0923	-0.3199	0.0360
	$\kappa_{1,\text{year}_{15}}$	0.3145	0.0221	0.2719	0.3557		$\kappa^*_{2,\text{year}_{17}}$	-0.1407	0.0956	-0.3304	0.0452
	$\kappa_{1,\text{year}_{16}}$	0.3138	0.0234	0.2706	0.3566		$\kappa^*_{2,\text{year}_{18}}$	-0.1437	0.0998	-0.3472	0.0584
	$\kappa_{1,\text{year}_{17}}$	0.3049	0.0238	0.2587	0.3473		ψ_{κ_1}	0.0112	0.0040	0.0048	0.0204
	$\kappa_{1,\text{year}_{18}}$	0.2881	0.0240	0.2400	0.3329		ψ_{κ_2}	-0.0026	0.0032	-0.0095	0.0031
	$\kappa^*_{1,\text{year}_1}$	0.2999	0.0349	0.2328	0.3711		σ^2	0.0068	0.0007	0.0056	0.0082
	$\kappa^*_{1,\text{year}_2}$	0.3113	0.0455	0.2230	0.4041		$\sigma^2_{\psi_{\kappa_2}}$	0.0000	0.0000	0.0000	0.0000

B.3 Parameter estimates for male lung cancer based on the model in (3.6), without smoking data

Table B.3: Estimated coefficients for lung cancer mortality in men based on (3.6).

Covariate	Parameter	Mean	SD	%2.5	%97.5	Covariate	Parameter	Mean	SD	%2.5	%97.5
	β_0	-6.1940	0.0084	-6.2090	-6.1780		κ_{1,year_3}	-0.0539	0.0119	-0.0753	-0.0300
	β_{1,age_1}	-2.2110	0.0091	-2.2270	-2.1930		κ_{1,year_4}	-0.0892	0.0119	-0.1124	-0.0678
	β_{1,age_2}	-1.0710	0.0073	-1.0850	-1.0550		κ_{1,year_5}	-0.1100	0.0118	-0.1330	-0.0878
	β_{1,age_3}	-0.4075	0.0059	-0.4192	-0.3959		κ_{1,year_6}	-0.1150	0.0118	-0.1382	-0.0932
	β_{1,age_4}	0.0908	0.0053	0.0801	0.1016		κ_{1,year_7}	-0.1357	0.0122	-0.1610	-0.1116
	β_{1,age_5}	0.5010	0.0049	0.4915	0.5108		κ_{1,year_8}	-0.1478	0.0115	-0.1718	-0.1271
	β_{1,age_6}	0.8494	0.0047	0.8401	0.8586		κ_{1,year_9}	-0.1691	0.0126	-0.1948	-0.1453
	β_{1,age_7}	1.0490	0.0050	1.0390	1.0580		$\kappa_{1,\text{year}_{10}}$	-0.2025	0.0127	-0.2262	-0.1783
	β_{1,age_8}	1.1990	0.0062	1.1870	1.2110		$\kappa_{1,\text{year}_{11}}$	-0.2241	0.0118	-0.2484	-0.2027
	$\beta_{2,\text{region}_1}$	-0.4923	0.0361	-0.5506	-0.4345		$\kappa_{1,\text{year}_{12}}$	-0.2528	0.0118	-0.2765	-0.2309
	$\beta_{2,\text{region}_2}$	0.1290	0.0055	0.1182	0.1396		$\kappa_{1,\text{year}_{13}}$	-0.2684	0.0119	-0.2919	-0.2461
	$\beta_{2,\text{region}_3}$	-0.0991	0.0122	-0.1214	-0.0776		$\kappa_{1,\text{year}_{14}}$	-0.2915	0.0114	-0.3159	-0.2697
	$\beta_{2,\text{region}_4}$	0.2805	0.0168	0.2508	0.3092		$\kappa_{1,\text{year}_{15}}$	-0.3354	0.0121	-0.3601	-0.3128
	$\beta_{2,\text{region}_5}$	0.2531	0.0161	0.2253	0.2815		$\kappa_{1,\text{year}_{16}}$	-0.3479	0.0115	-0.3707	-0.3256
	$\beta_{2,\text{region}_6}$	0.1472	0.0129	0.1236	0.1697		$\kappa_{1,\text{year}_{17}}$	-0.3944	0.0120	-0.4192	-0.3715
	$\beta_{2,\text{region}_7}$	-0.8548	0.0406	-0.9173	-0.7886		$\kappa_{1,\text{year}_{18}}$	-0.4319	0.0122	-0.4556	-0.4082
	$\beta_{2,\text{region}_8}$	0.3206	0.0202	0.2869	0.3538		$\kappa^*_{1,\text{year}_1}$	-0.4499	0.0445	-0.5337	-0.3637
	$\beta_{2,\text{region}_9}$	0.3158	0.0245	0.2757	0.3566		$\kappa^*_{1,\text{year}_2}$	-0.4685	0.0611	-0.5888	-0.3461
	$\beta_{3,\text{deprivation}_1}$	2.9590	0.1314	2.7500	3.1650		$\kappa^*_{1,\text{year}_3}$	-0.4847	0.0753	-0.6343	-0.3312
	$\beta_{3,\text{deprivation}_2}$	1.2520	0.0547	1.1660	1.3390		$\kappa^*_{1,\text{year}_4}$	-0.5016	0.0875	-0.6804	-0.3218
	$\beta_{3,\text{deprivation}_3}$	-0.1192	0.0074	-0.1328	-0.1050		$\kappa^*_{1,\text{year}_5}$	-0.5162	0.0982	-0.7054	-0.3190
	$\beta_{3,\text{deprivation}_4}$	-1.3950	0.0617	-1.4920	-1.2980		$\kappa^*_{1,\text{year}_6}$	-0.5327	0.1097	-0.7525	-0.3137
	$\beta_{3,\text{deprivation}_5}$	-2.6980	0.1187	-2.8810	-2.5060		$\kappa^*_{1,\text{year}_7}$	-0.5494	0.1199	-0.7867	-0.3058
	β_4	1.7290	0.0910	1.5840	1.8720		$\kappa^*_{1,\text{year}_8}$	-0.5657	0.1292	-0.8218	-0.3021
	$\beta_{5,\text{deprivation}_1,\text{age}_1}$	0.1289	0.0144	0.1006	0.1563		$\kappa^*_{1,\text{year}_9}$	-0.5813	0.1370	-0.8640	-0.3136
	$\beta_{5,\text{deprivation}_2,\text{age}_1}$	0.1460	0.0125	0.1213	0.1700		$\kappa^*_{1,\text{year}_{10}}$	-0.5979	0.1453	-0.8924	-0.3161
	$\beta_{5,\text{deprivation}_3,\text{age}_1}$	0.1107	0.0105	0.0902	0.1314		$\kappa^*_{1,\text{year}_{11}}$	-0.6149	0.1529	-0.9231	-0.3152
	$\beta_{5,\text{deprivation}_4,\text{age}_1}$	0.0659	0.0089	0.0486	0.0835		$\kappa^*_{1,\text{year}_{12}}$	-0.6324	0.1592	-0.9602	-0.3203
	$\beta_{5,\text{deprivation}_5,\text{age}_1}$	-0.0054	0.0086	-0.0219	0.0117		$\kappa^*_{1,\text{year}_{13}}$	-0.6506	0.1664	-0.9945	-0.3153
	$\beta_{5,\text{deprivation}_1,\text{age}_2}$	-0.0711	0.0086	-0.0881	-0.0543		$\kappa^*_{1,\text{year}_{14}}$	-0.6677	0.1727	-1.0230	-0.3207
	$\beta_{5,\text{deprivation}_2,\text{age}_2}$	-0.1484	0.0095	-0.1654	-0.1289		$\kappa^*_{1,\text{year}_{15}}$	-0.6841	0.1803	-1.0560	-0.3183
	$\beta_{5,\text{deprivation}_3,\text{age}_2}$	-0.2267	0.0120	-0.2501	-0.2026		$\kappa^*_{1,\text{year}_{16}}$	-0.7014	0.1885	-1.0730	-0.3128
	$\beta_{5,\text{deprivation}_4,\text{age}_2}$	0.0645	0.0163	0.0316	0.0942		$\kappa^*_{1,\text{year}_{17}}$	-0.7187	0.1938	-1.1090	-0.3260
	$\beta_{5,\text{deprivation}_5,\text{age}_2}$	0.0458	0.0138	0.0204	0.0745		$\kappa^*_{1,\text{year}_{18}}$	-0.7347	0.2008	-1.1270	-0.3421
	$\beta_{5,\text{deprivation}_1,\text{age}_3}$	0.0272	0.0107	0.0058	0.0483		κ_{2,year_2}	-0.0006	0.0110	-0.0220	0.0207
	$\beta_{5,\text{deprivation}_2,\text{age}_3}$	0.0381	0.0096	0.0193	0.0568		κ_{2,year_3}	0.0008	0.0118	-0.0250	0.0220
	$\beta_{5,\text{deprivation}_3,\text{age}_3}$	0.0050	0.0092	-0.0126	0.0239		κ_{2,year_4}	-0.0215	0.0124	-0.0476	0.0033
	$\beta_{5,\text{deprivation}_4,\text{age}_3}$	-0.0184	0.0086	-0.0356	-0.0015		κ_{2,year_5}	-0.0278	0.0117	-0.0502	-0.0036
	$\beta_{5,\text{deprivation}_5,\text{age}_3}$	-0.0712	0.0097	-0.0905	-0.0515		κ_{2,year_6}	-0.0352	0.0119	-0.0591	-0.0122
	$\beta_{5,\text{deprivation}_1,\text{age}_4}$	-0.0908	0.0117	-0.1146	-0.0686		κ_{2,year_7}	-0.0369	0.0119	-0.0598	-0.0133
	$\beta_{5,\text{deprivation}_2,\text{age}_4}$	0.0050	0.0176	-0.0288	0.0377		κ_{2,year_8}	-0.0489	0.0107	-0.0700	-0.0272
	$\beta_{5,\text{deprivation}_3,\text{age}_4}$	0.0032	0.0145	-0.0259	0.0301		κ_{2,year_9}	-0.0431	0.0116	-0.0657	-0.0203
	$\beta_{5,\text{deprivation}_4,\text{age}_4}$	-0.0234	0.0121	-0.0475	0.0005		$\kappa_{2,\text{year}_{10}}$	-0.0556	0.0120	-0.0791	-0.0318
	$\beta_{5,\text{deprivation}_5,\text{age}_4}$	0.0037	0.0099	-0.0150	0.0240		$\kappa_{2,\text{year}_{11}}$	-0.0494	0.0125	-0.0738	-0.0234
	$\beta_{5,\text{deprivation}_1,\text{age}_5}$	-0.0090	0.0098	-0.0281	0.0107		$\kappa_{2,\text{year}_{12}}$	-0.0542	0.0116	-0.0754	-0.0314
	$\beta_{5,\text{deprivation}_2,\text{age}_5}$	0.0106	0.0090	-0.0069	0.0279		$\kappa_{2,\text{year}_{13}}$	-0.0664	0.0118	-0.0909	-0.0444
	$\beta_{5,\text{deprivation}_3,\text{age}_5}$	0.0043	0.0101	-0.0150	0.0245		$\kappa_{2,\text{year}_{14}}$	-0.0784	0.0120	-0.1027	-0.0542
	$\beta_{5,\text{deprivation}_4,\text{age}_5}$	0.0058	0.0116	-0.0166	0.0291		$\kappa_{2,\text{year}_{15}}$	-0.0856	0.0124	-0.1107	-0.0603
	$\beta_{5,\text{deprivation}_5,\text{age}_5}$	-0.0823	0.0194	-0.1208	-0.0435		$\kappa_{2,\text{year}_{16}}$	-0.0786	0.0115	-0.1014	-0.0564
	$\beta_{5,\text{deprivation}_1,\text{age}_6}$	-0.0707	0.0168	-0.1033	-0.0373		$\kappa_{2,\text{year}_{17}}$	-0.0803	0.0120	-0.1037	-0.0560
	$\beta_{5,\text{deprivation}_2,\text{age}_6}$	-0.0474	0.0118	-0.0706	-0.0240		$\kappa_{2,\text{year}_{18}}$	-0.0832	0.0122	-0.1072	-0.0596
	$\beta_{5,\text{deprivation}_3,\text{age}_6}$	-0.0426	0.0104	-0.0632	-0.0214		$\kappa^*_{2,\text{year}_1}$	-0.0868	0.0386	-0.1626	-0.0082
	$\beta_{5,\text{deprivation}_4,\text{age}_6}$	0.0057	0.0102	-0.0146	0.0258		$\kappa^*_{2,\text{year}_2}$	-0.0892	0.0547	-0.1946	0.0251
	$\beta_{5,\text{deprivation}_5,\text{age}_6}$	0.0339	0.0098	0.0145	0.0532		$\kappa^*_{2,\text{year}_3}$	-0.0926	0.0659	-0.2246	0.0376
	$\beta_{5,\text{deprivation}_1,\text{age}_7}$	0.0770	0.0100	0.0577	0.0966		$\kappa^*_{2,\text{year}_4}$	-0.0954	0.0761	-0.2516	0.0552
	$\beta_{5,\text{deprivation}_2,\text{age}_7}$	0.1263	0.0122	0.1025	0.1502		$\kappa^*_{2,\text{year}_5}$	-0.0980	0.0852	-0.2660	0.0696
	$\beta_{5,\text{deprivation}_3,\text{age}_7}$	-0.1161	0.0215	-0.1586	-0.0721		$\kappa^*_{2,\text{year}_6}$	-0.1015	0.0941	-0.2827	0.0850
	$\beta_{5,\text{deprivation}_4,\text{age}_7}$	-0.1242	0.0174	-0.1591	-0.0906		$\kappa^*_{2,\text{year}_7}$	-0.1055	0.1032	-0.3129	0.0982
	$\beta_{5,\text{deprivation}_5,\text{age}_7}$	-0.0671	0.0134	-0.0933	-0.0410		$\kappa^*_{2,\text{year}_8}$	-0.1065	0.1101	-0.3269	0.1152
	$\beta_{5,\text{deprivation}_1,\text{age}_8}$	-0.0651	0.0118	-0.0878	-0.0409		$\kappa^*_{2,\text{year}_9}$	-0.1103	0.1170	-0.3489	0.1251
	$\beta_{5,\text{deprivation}_2,\text{age}_8}$	0.0037	0.0112	-0.0184	0.0260		$\kappa^*_{2,\text{year}_{10}}$	-0.1131	0.1243	-0.3624	0.1421
	$\beta_{5,\text{deprivation}_3,\text{age}_8}$	0.0451	0.0104	0.0257	0.0665		$\kappa^*_{2,\text{year}_{11}}$	-0.1159	0.1298	-0.3696	0.1472
	$\beta_{5,\text{deprivation}_4,\text{age}_8}$	0.1383	0.0113	0.1165	0.1602		$\kappa^*_{2,\text{year}_{12}}$	-0.1192	0.1361	-0.3878	0.1568
	$\beta_{5,\text{deprivation}_5,\text{age}_8}$	0.1854	0.0135	0.1603	0.2138		$\kappa^*_{2,\text{year}_{13}}$	-0.1228	0.1429	-0.4013	0.1735
	$\beta_{6,\text{region}_1}$	-0.0004	0.0080	-0.0156	0.0157		$\kappa^*_{2,\text{year}_{14}}$	-0.1248	0.1468	-0.4009	0.1808
	$\beta_{6,\text{region}_2}$	-0.0043	0.0055	-0.0146	0.0065		$\kappa^*_{2,\text{year}_{15}}$	-0.1295	0.1522	-0.4243	0.1866
	$\beta_{6,\text{region}_3}$	0.0027	0.0058	-0.0085	0.0142		$\kappa^*_{2,\text{year}_{16}}$	-0.1325	0.1586	-0.4474	0.2000
	$\beta_{6,\text{region}_4}$	0.0198	0.0068	0.0062	0.0323		$\kappa^*_{2,\text{year}_{17}}$	-0.1356	0.1643	-0.4615	0.1977
	$\beta_{6,\text{region}_5}$	0.0072	0.0059	-0.0043	0.0191		$\kappa^*_{2,\text{year}_{18}}$	-0.1382	0.1705	-0.4767	0.2011
	$\beta_{6,\text{region}_6}$	0.0286	0.0069	0.0153	0.0425		σ^2	0.0021	0.0003	0.0014	0.0028
	$\beta_{6,\text{region}_7}$	0.0187	0.0065	0.0060	0.0312		$\sigma^2_{\psi_{\kappa_2}}$	0.0001	0.0000	0.0000	0.0001
	$\beta_{6,\text{region}_8}$	-0.0335	0.0059	-0.0449	-0.0220		$\sigma^2_{\psi_{\kappa_1}}$	0.0001	0.0000	0.0001	0.0002
	$\beta_{6,\text{region}_9}$	-0.0387	0.0076	-0.0536	-0.0237		$\sigma^2_{\kappa_1}$	0.0018	0.0006	0.0009	0.0034
	κ_{1,year_2}	-0.0208	0.0118	-0.0420	0.0047		$\sigma^2_{\kappa_2}$	0.0013	0.0005	0.0007	0.0025

B.4 Parameter estimates for male lung cancer based on the model in (3.7), with smoking data

Table B.4: Estimated coefficients for lung cancer mortality in men based on (3.7).

Covariate	Parameter	Mean	SD	%2.5	%97.5	Covariate	Parameter	Mean	SD	%2.5	%97.5
	β_0	-6.8200	0.0326	-6.8960	-6.7620		κ_{1,year_3}	0.0186	0.0105	-0.0019	0.0380
	β_{1,age_1}	-1.9120	0.0185	-1.9480	-1.8740		κ_{1,year_4}	0.0168	0.0119	-0.0041	0.0414
	β_{1,age_2}	-0.9831	0.0088	-1.0010	-0.9657		κ_{1,year_5}	0.0319	0.0122	0.0084	0.0568
	β_{1,age_3}	-0.3204	0.0074	-0.3352	-0.3056		κ_{1,year_6}	0.0637	0.0146	0.0364	0.0917
	β_{1,age_4}	-0.0378	0.0088	-0.0553	-0.0207		κ_{1,year_7}	0.0748	0.0142	0.0485	0.1053
	β_{1,age_5}	0.3721	0.0088	0.3539	0.3890		κ_{1,year_8}	0.1005	0.0164	0.0699	0.1361
	β_{1,age_6}	0.7774	0.0059	0.7657	0.7893		κ_{1,year_9}	0.1094	0.0168	0.0775	0.1448
	β_{1,age_7}	0.9769	0.0065	0.9640	0.9889		$\kappa_{1,\text{year}_{10}}$	0.1122	0.0198	0.0746	0.1508
	β_{1,age_8}	1.1270	0.0069	1.1140	1.1400		$\kappa_{1,\text{year}_{11}}$	0.1249	0.0196	0.0935	0.1699
	$\beta_{2,\text{region}_1}$	-0.5047	0.0110	-0.5255	-0.4826		$\kappa_{1,\text{year}_{12}}$	0.1298	0.0213	0.0920	0.1771
	$\beta_{2,\text{region}_2}$	0.1296	0.0052	0.1195	0.1400		$\kappa_{1,\text{year}_{13}}$	0.1469	0.0230	0.1062	0.1950
	$\beta_{2,\text{region}_3}$	-0.1027	0.0062	-0.1150	-0.0901		$\kappa_{1,\text{year}_{14}}$	0.1562	0.0252	0.1121	0.2073
	$\beta_{2,\text{region}_4}$	0.2865	0.0075	0.2714	0.3011		$\kappa_{1,\text{year}_{15}}$	0.1463	0.0263	0.0988	0.2029
	$\beta_{2,\text{region}_5}$	0.2578	0.0075	0.2427	0.2722		$\kappa_{1,\text{year}_{16}}$	0.1647	0.0285	0.1161	0.2254
	$\beta_{2,\text{region}_6}$	0.1515	0.0066	0.1383	0.1636		$\kappa_{1,\text{year}_{17}}$	0.1515	0.0295	0.0982	0.2175
	$\beta_{2,\text{region}_7}$	-0.8694	0.0108	-0.8904	-0.8468		$\kappa_{1,\text{year}_{18}}$	0.1456	0.0309	0.0915	0.2124
	$\beta_{2,\text{region}_8}$	0.3276	0.0070	0.3142	0.3420		$\kappa^*_{1,\text{year}_1}$	0.1509	0.0498	0.0554	0.2516
	$\beta_{2,\text{region}_9}$	0.3238	0.0084	0.3077	0.3407		$\kappa^*_{1,\text{year}_2}$	0.1573	0.0645	0.0349	0.2834
	$\beta_{3,\text{deprivation}_1}$	3.0050	0.0314	2.9370	3.0660		$\kappa^*_{1,\text{year}_3}$	0.1619	0.0760	0.0097	0.3099
	$\beta_{3,\text{deprivation}_2}$	1.2720	0.0134	1.2420	1.3000		$\kappa^*_{1,\text{year}_4}$	0.1665	0.0862	-0.0043	0.3388
	$\beta_{3,\text{deprivation}_3}$	-0.1210	0.0051	-0.1306	-0.1108		$\kappa^*_{1,\text{year}_5}$	0.1715	0.0960	-0.0191	0.3537
	$\beta_{3,\text{deprivation}_4}$	-1.4170	0.0151	-1.4450	-1.3850		$\kappa^*_{1,\text{year}_6}$	0.1760	0.1051	-0.0347	0.3808
	$\beta_{3,\text{deprivation}_5}$	-2.7390	0.0281	-2.7960	-2.6770		$\kappa^*_{1,\text{year}_7}$	0.1813	0.1155	-0.0516	0.4082
	β_4	1.7610	0.0214	1.7170	1.8040		$\kappa^*_{1,\text{year}_8}$	0.1860	0.1237	-0.0570	0.4335
	$\beta_{5,\text{deprivation}_1,\text{age}_1}$	0.1282	0.0135	0.1013	0.1543		$\kappa^*_{1,\text{year}_9}$	0.1911	0.1310	-0.0692	0.4480
	$\beta_{5,\text{deprivation}_2,\text{age}_1}$	0.1466	0.0127	0.1216	0.1720		$\kappa^*_{1,\text{year}_{10}}$	0.1961	0.1403	-0.0899	0.4747
	$\beta_{5,\text{deprivation}_3,\text{age}_1}$	0.1108	0.0105	0.0901	0.1320		$\kappa^*_{1,\text{year}_{11}}$	0.2001	0.1494	-0.1095	0.4897
	$\beta_{5,\text{deprivation}_4,\text{age}_1}$	0.0657	0.0090	0.0483	0.0828		$\kappa^*_{1,\text{year}_{12}}$	0.2075	0.1567	-0.1180	0.5154
	$\beta_{5,\text{deprivation}_5,\text{age}_1}$	-0.0059	0.0086	-0.0226	0.0115		$\kappa^*_{1,\text{year}_{13}}$	0.2133	0.1635	-0.1248	0.5335
	$\beta_{5,\text{deprivation}_1,\text{age}_2}$	-0.0708	0.0083	-0.0867	-0.0545		$\kappa^*_{1,\text{year}_{14}}$	0.2185	0.1692	-0.1257	0.5457
	$\beta_{5,\text{deprivation}_2,\text{age}_2}$	-0.1486	0.0093	-0.1670	-0.1302		$\kappa^*_{1,\text{year}_{15}}$	0.2232	0.1751	-0.1326	0.5643
	$\beta_{5,\text{deprivation}_3,\text{age}_2}$	-0.2260	0.0118	-0.2487	-0.2026		$\kappa^*_{1,\text{year}_{16}}$	0.2287	0.1814	-0.1341	0.5802
	$\beta_{5,\text{deprivation}_4,\text{age}_2}$	0.0638	0.0159	0.0316	0.0938		$\kappa^*_{1,\text{year}_{17}}$	0.2351	0.1881	-0.1330	0.5934
	$\beta_{5,\text{deprivation}_5,\text{age}_2}$	0.0445	0.0135	0.0164	0.0705		$\kappa^*_{1,\text{year}_{18}}$	0.2402	0.1940	-0.1444	0.6009
	$\beta_{5,\text{deprivation}_1,\text{age}_3}$	0.0278	0.0117	0.0051	0.0503		κ_{2,year_2}	0.0003	0.0138	-0.0270	0.0277
	$\beta_{5,\text{deprivation}_2,\text{age}_3}$	0.0391	0.0099	0.0191	0.0594		κ_{2,year_3}	0.0003	0.0137	-0.0256	0.0280
	$\beta_{5,\text{deprivation}_3,\text{age}_3}$	0.0052	0.0096	-0.0134	0.0240		κ_{2,year_4}	-0.0205	0.0126	-0.0448	0.0063
	$\beta_{5,\text{deprivation}_4,\text{age}_3}$	-0.0179	0.0091	-0.0359	-0.0008		κ_{2,year_5}	-0.0299	0.0119	-0.0521	-0.0057
	$\beta_{5,\text{deprivation}_5,\text{age}_3}$	-0.0710	0.0099	-0.0904	-0.0517		κ_{2,year_6}	-0.0348	0.0119	-0.0563	-0.0086
	$\beta_{5,\text{deprivation}_1,\text{age}_4}$	-0.0916	0.0121	-0.1157	-0.0680		κ_{2,year_7}	-0.0387	0.0140	-0.0686	-0.0129
	$\beta_{5,\text{deprivation}_2,\text{age}_4}$	0.0056	0.0176	-0.0276	0.0434		κ_{2,year_8}	-0.0480	0.0118	-0.0716	-0.0270
	$\beta_{5,\text{deprivation}_3,\text{age}_4}$	0.0044	0.0146	-0.0230	0.0339		κ_{2,year_9}	-0.0440	0.0123	-0.0695	-0.0218
	$\beta_{5,\text{deprivation}_4,\text{age}_4}$	-0.0242	0.0126	-0.0484	0.0015		$\kappa_{2,\text{year}_{10}}$	-0.0556	0.0118	-0.0773	-0.0290
	$\beta_{5,\text{deprivation}_5,\text{age}_4}$	0.0020	0.0100	-0.0178	0.0212		$\kappa_{2,\text{year}_{11}}$	-0.0511	0.0119	-0.0750	-0.0251
	$\beta_{5,\text{deprivation}_1,\text{age}_5}$	-0.0077	0.0098	-0.0275	0.0116		$\kappa_{2,\text{year}_{12}}$	-0.0532	0.0129	-0.0774	-0.0275
	$\beta_{5,\text{deprivation}_2,\text{age}_5}$	0.0100	0.0092	-0.0082	0.0279		$\kappa_{2,\text{year}_{13}}$	-0.0680	0.0125	-0.0938	-0.0437
	$\beta_{5,\text{deprivation}_3,\text{age}_5}$	0.0041	0.0096	-0.0150	0.0230		$\kappa_{2,\text{year}_{14}}$	-0.0798	0.0125	-0.1044	-0.0552
	$\beta_{5,\text{deprivation}_4,\text{age}_5}$	0.0057	0.0116	-0.0168	0.0295		$\kappa_{2,\text{year}_{15}}$	-0.0855	0.0123	-0.1082	-0.0605
	$\beta_{5,\text{deprivation}_5,\text{age}_5}$	-0.0834	0.0183	-0.1206	-0.0491		$\kappa_{2,\text{year}_{16}}$	-0.0804	0.0126	-0.1049	-0.0542
	$\beta_{5,\text{deprivation}_1,\text{age}_6}$	-0.0724	0.0165	-0.1048	-0.0397		$\kappa_{2,\text{year}_{17}}$	-0.0810	0.0122	-0.1039	-0.0552
	$\beta_{5,\text{deprivation}_2,\text{age}_6}$	-0.0474	0.0127	-0.0739	-0.0223		$\kappa_{2,\text{year}_{18}}$	-0.0847	0.0126	-0.1094	-0.0588
	$\beta_{5,\text{deprivation}_3,\text{age}_6}$	-0.0418	0.0103	-0.0617	-0.0216		$\kappa^*_{2,\text{year}_1}$	-0.0952	0.0395	-0.1749	-0.0175
	$\beta_{5,\text{deprivation}_4,\text{age}_6}$	0.0048	0.0096	-0.0143	0.0236		$\kappa^*_{2,\text{year}_2}$	-0.1065	0.0551	-0.2142	0.0033
	$\beta_{5,\text{deprivation}_5,\text{age}_6}$	0.0338	0.0095	0.0158	0.0525		$\kappa^*_{2,\text{year}_3}$	-0.1165	0.0656	-0.2447	0.0140
	$\beta_{5,\text{deprivation}_1,\text{age}_7}$	0.0784	0.0099	0.0592	0.0985		$\kappa^*_{2,\text{year}_4}$	-0.1271	0.0781	-0.2823	0.0249
	$\beta_{5,\text{deprivation}_2,\text{age}_7}$	0.1280	0.0129	0.1028	0.1531		$\kappa^*_{2,\text{year}_5}$	-0.1371	0.0882	-0.3114	0.0427
	$\beta_{5,\text{deprivation}_3,\text{age}_7}$	-0.1142	0.0209	-0.1592	-0.0742		$\kappa^*_{2,\text{year}_6}$	-0.1489	0.0966	-0.3392	0.0406
	$\beta_{5,\text{deprivation}_4,\text{age}_7}$	-0.1231	0.0189	-0.1605	-0.0868		$\kappa^*_{2,\text{year}_7}$	-0.1596	0.1049	-0.3703	0.0389
	$\beta_{5,\text{deprivation}_5,\text{age}_7}$	-0.0670	0.0145	-0.0934	-0.0366		$\kappa^*_{2,\text{year}_8}$	-0.1691	0.1131	-0.4038	0.0427
	$\beta_{5,\text{deprivation}_1,\text{age}_8}$	-0.0651	0.0120	-0.0878	-0.0422		$\kappa^*_{2,\text{year}_9}$	-0.1796	0.1216	-0.4359	0.0547
	$\beta_{5,\text{deprivation}_2,\text{age}_8}$	0.0036	0.0104	-0.0173	0.0228		$\kappa^*_{2,\text{year}_{10}}$	-0.1911	0.1281	-0.4500	0.0554
	$\beta_{5,\text{deprivation}_3,\text{age}_8}$	0.0448	0.0109	0.0238	0.0667		$\kappa^*_{2,\text{year}_{11}}$	-0.2015	0.1339	-0.4754	0.0511
	$\beta_{5,\text{deprivation}_4,\text{age}_8}$	0.1371	0.0114	0.1163	0.1594		$\kappa^*_{2,\text{year}_{12}}$	-0.2113	0.1400	-0.5031	0.0546
	$\beta_{5,\text{deprivation}_5,\text{age}_8}$	0.1839	0.0134	0.1570	0.2094		$\kappa^*_{2,\text{year}_{13}}$	-0.2212	0.1449	-0.5205	0.0590
	$\beta_{6,\text{region}_1}$	0.0010	0.0075	-0.0133	0.0164		$\kappa^*_{2,\text{year}_{14}}$	-0.2310	0.1517	-0.5447	0.0617
	$\beta_{6,\text{region}_2}$	-0.0040	0.0053	-0.0140	0.0061		$\kappa^*_{2,\text{year}_{15}}$	-0.2423	0.1571	-0.5715	0.0563
	$\beta_{6,\text{region}_3}$	0.0030	0.0058	-0.0085	0.0143		$\kappa^*_{2,\text{year}_{16}}$	-0.2542	0.1613	-0.5719	0.0402
	$\beta_{6,\text{region}_4}$	0.0201	0.0068	0.0069	0.0332		$\kappa^*_{2,\text{year}_{17}}$	-0.2656	0.1661	-0.6038	0.0482
	$\beta_{6,\text{region}_5}$	0.0074	0.0058	-0.0041	0.0190		$\kappa^*_{2,\text{year}_{18}}$	-0.2760	0.1699	-0.6196	0.0458
	$\beta_{6,\text{region}_6}$	0.0285	0.0063	0.0161	0.0406		σ^2	0.0020	0.0004	0.0014	0.0028
	$\beta_{6,\text{region}_7}$	0.0185	0.0062	0.0062	0.0309		$\sigma^2_{\psi_{\kappa_2}}$	0.0001	0.0000	0.0000	0.0002
	$\beta_{6,\text{region}_8}$	-0.0340	0.0058	-0.0454	-0.0228		$\sigma^2_{\psi_{\kappa_1}}$	0.0001	0.0000	0.0000	0.0002
	$\beta_{6,\text{region}_9}$	-0.0404	0.0066	-0.0533	-0.0275		$\sigma^2_{\kappa_1}$	0.0014	0.0005	0.0008	0.0028
	β_7	-0.3810	0.0207	-0.4265	-0.3425		$\sigma^2_{\kappa_2}$	0.0014	0.0005	0.0007	0.0027
	κ_{1,year_2}	0.0152	0.0101	-0.0037	0.0341						

B.5 Parameter estimates for female breast cancer based on the model in (3.8), without smoking data

Table B.5: Estimated coefficients for breast cancer mortality based on (3.8).

Covariate Parameter	Mean	SD	%2.5	%97.5	Covariate Parameter	Mean	SD	%2.5	%97.5
β_0	-7.3260	0.0105	-7.3470	-7.3050	$\kappa_{1,\text{year}_{11}}$	-0.2203	0.0153	-0.2502	-0.1909
β_{1,age_1}	-1.8560	0.0172	-1.8890	-1.8230	$\kappa_{1,\text{year}_{12}}$	-0.2403	0.0141	-0.2674	-0.2120
β_{1,age_2}	-1.2210	0.0136	-1.2480	-1.1930	$\kappa_{1,\text{year}_{13}}$	-0.2690	0.0149	-0.2979	-0.2394
β_{1,age_3}	-0.7361	0.0110	-0.7578	-0.7146	$\kappa_{1,\text{year}_{14}}$	-0.2903	0.0149	-0.3191	-0.2615
β_{1,age_4}	-0.3474	0.0099	-0.3664	-0.3285	$\kappa_{1,\text{year}_{15}}$	-0.2988	0.0148	-0.3281	-0.2707
β_{1,age_5}	-0.0900	0.0096	-0.1086	-0.0715	$\kappa_{1,\text{year}_{16}}$	-0.3118	0.0145	-0.3382	-0.2825
β_{1,age_6}	0.1072	0.0090	0.0894	0.1247	$\kappa_{1,\text{year}_{17}}$	-0.3305	0.0149	-0.3594	-0.3006
β_{1,age_7}	0.2801	0.0089	0.2633	0.2979	$\kappa_{1,\text{year}_{18}}$	-0.3441	0.0161	-0.3753	-0.3120
β_{1,age_8}	0.5136	0.0087	0.4968	0.5306	$\kappa^*_{1,\text{year}_1}$	-0.3549	0.0293	-0.4144	-0.2994
β_{1,age_9}	0.8116	0.0084	0.7953	0.8279	$\kappa^*_{1,\text{year}_2}$	-0.3652	0.0380	-0.4420	-0.2935
$\beta_{1,\text{age}_{10}}$	1.1040	0.0084	1.0880	1.1210	$\kappa^*_{1,\text{year}_3}$	-0.3751	0.0451	-0.4638	-0.2870
$\beta_{1,\text{age}_{11}}$	1.4340	0.0086	1.4180	1.4500	$\kappa^*_{1,\text{year}_4}$	-0.3849	0.0517	-0.4864	-0.2793
$\beta_{2,\text{region}_1}$	-0.0314	0.0108	-0.0525	-0.0099	$\kappa^*_{1,\text{year}_5}$	-0.3953	0.0574	-0.5072	-0.2816
$\beta_{2,\text{region}_2}$	-0.0143	0.0080	-0.0298	0.0017	$\kappa^*_{1,\text{year}_6}$	-0.4052	0.0628	-0.5269	-0.2797
$\beta_{2,\text{region}_3}$	-0.0296	0.0088	-0.0468	-0.0124	$\kappa^*_{1,\text{year}_7}$	-0.4146	0.0683	-0.5499	-0.2768
$\beta_{2,\text{region}_4}$	0.0234	0.0090	0.0059	0.0413	$\kappa^*_{1,\text{year}_8}$	-0.4250	0.0735	-0.5741	-0.2833
$\beta_{2,\text{region}_5}$	0.0302	0.0085	0.0129	0.0471	$\kappa^*_{1,\text{year}_9}$	-0.4340	0.0771	-0.5875	-0.2819
$\beta_{2,\text{region}_6}$	0.0279	0.0084	0.0114	0.0447	$\kappa^*_{1,\text{year}_{10}}$	-0.4449	0.0826	-0.6089	-0.2827
$\beta_{2,\text{region}_7}$	-0.0100	0.0083	-0.0263	0.0065	$\kappa^*_{1,\text{year}_{11}}$	-0.4544	0.0868	-0.6338	-0.2837
$\beta_{2,\text{region}_8}$	0.0187	0.0075	0.0034	0.0337	$\kappa^*_{1,\text{year}_{12}}$	-0.4645	0.0903	-0.6536	-0.2863
$\beta_{2,\text{region}_9}$	-0.0150	0.0085	-0.0316	0.0010	$\kappa^*_{1,\text{year}_{13}}$	-0.4755	0.0951	-0.6726	-0.2957
κ_{1,year_2}	-0.0199	0.0121	-0.0433	0.0041	$\kappa^*_{1,\text{year}_{14}}$	-0.4855	0.1014	-0.6960	-0.2883
κ_{1,year_3}	-0.0526	0.0138	-0.0815	-0.0258	$\kappa^*_{1,\text{year}_{15}}$	-0.4964	0.1070	-0.7137	-0.2859
κ_{1,year_4}	-0.0707	0.0149	-0.1004	-0.0405	$\kappa^*_{1,\text{year}_{16}}$	-0.5065	0.1108	-0.7248	-0.2841
κ_{1,year_5}	-0.0822	0.0138	-0.1096	-0.0556	$\kappa^*_{1,\text{year}_{17}}$	-0.5170	0.1150	-0.7428	-0.2932
κ_{1,year_6}	-0.1014	0.0143	-0.1296	-0.0748	$\kappa^*_{1,\text{year}_{18}}$	-0.5280	0.1193	-0.7659	-0.3007
κ_{1,year_7}	-0.1302	0.0149	-0.1606	-0.1030	ψ_{κ_1}	-0.0100	0.0036	-0.0181	-0.0038
κ_{1,year_8}	-0.1473	0.0145	-0.1756	-0.1185	σ^2	0.0040	0.0004	0.0032	0.0047
κ_{1,year_9}	-0.1731	0.0141	-0.2009	-0.1457	$\sigma^2_{\kappa_1}$	0.0005	0.0002	0.0003	0.0011
$\kappa_{1,\text{year}_{10}}$	-0.2041	0.0153	-0.2338	-0.1746					

B.6 Parameter estimates for female breast cancer based on the model in (3.9), with smoking data

Table B.6: Estimated coefficients for breast cancer mortality based on (3.9).

Covariate	Parameter	Mean	SD	%2.5	%97.5	Covariate	Parameter	Mean	SD	%2.5	%97.5
	β_0	-7.4490	0.0231	-7.4920	-7.3980		$\kappa_{1,\text{year}_{10}}$	-0.1507	0.0171	-0.1860	-0.1184
	β_{1,age_1}	-1.8890	0.0184	-1.9250	-1.8530		$\kappa_{1,\text{year}_{11}}$	-0.1618	0.0178	-0.1985	-0.1304
	β_{1,age_2}	-1.2540	0.0141	-1.2810	-1.2260		$\kappa_{1,\text{year}_{12}}$	-0.1756	0.0182	-0.2114	-0.1395
	β_{1,age_3}	-0.8577	0.0234	-0.9029	-0.8071		$\kappa_{1,\text{year}_{13}}$	-0.1973	0.0189	-0.2351	-0.1589
	β_{1,age_4}	-0.4689	0.0230	-0.5098	-0.4189		$\kappa_{1,\text{year}_{14}}$	-0.2132	0.0194	-0.2525	-0.1760
	β_{1,age_5}	-0.1101	0.0100	-0.1297	-0.0896		$\kappa_{1,\text{year}_{15}}$	-0.2164	0.0202	-0.2582	-0.1770
	β_{1,age_6}	0.0867	0.0098	0.0679	0.1056		$\kappa_{1,\text{year}_{16}}$	-0.2238	0.0206	-0.2666	-0.1841
	β_{1,age_7}	0.2495	0.0102	0.2295	0.2691		$\kappa_{1,\text{year}_{17}}$	-0.2360	0.0211	-0.2807	-0.1962
	β_{1,age_8}	0.4826	0.0104	0.4631	0.5027		$\kappa_{1,\text{year}_{18}}$	-0.2448	0.0225	-0.2911	-0.2019
	β_{1,age_9}	0.9480	0.0244	0.8951	0.9916		$\kappa^*_{1,\text{year}_1}$	-0.2530	0.0310	-0.3150	-0.1953
	$\beta_{1,\text{age}_{10}}$	1.2420	0.0240	1.1890	1.2840		$\kappa^*_{1,\text{year}_2}$	-0.2608	0.0373	-0.3341	-0.1875
	$\beta_{1,\text{age}_{11}}$	1.5710	0.0242	1.5170	1.6170		$\kappa^*_{1,\text{year}_3}$	-0.2688	0.0440	-0.3582	-0.1805
	$\beta_{2,\text{region}_1}$	-0.0316	0.0112	-0.0540	-0.0099		$\kappa^*_{1,\text{year}_4}$	-0.2758	0.0502	-0.3781	-0.1785
	$\beta_{2,\text{region}_2}$	-0.0144	0.0081	-0.0302	0.0013		$\kappa^*_{1,\text{year}_5}$	-0.2839	0.0558	-0.4041	-0.1789
	$\beta_{2,\text{region}_3}$	-0.0298	0.0091	-0.0479	-0.0117		$\kappa^*_{1,\text{year}_6}$	-0.2923	0.0597	-0.4121	-0.1759
	$\beta_{2,\text{region}_4}$	0.0234	0.0092	0.0053	0.0413		$\kappa^*_{1,\text{year}_7}$	-0.2997	0.0645	-0.4307	-0.1720
	$\beta_{2,\text{region}_5}$	0.0305	0.0084	0.0143	0.0470		$\kappa^*_{1,\text{year}_8}$	-0.3074	0.0693	-0.4450	-0.1677
	$\beta_{2,\text{region}_6}$	0.0283	0.0081	0.0123	0.0441		$\kappa^*_{1,\text{year}_9}$	-0.3154	0.0740	-0.4592	-0.1708
	$\beta_{2,\text{region}_7}$	-0.0101	0.0085	-0.0267	0.0066		$\kappa^*_{1,\text{year}_{10}}$	-0.3230	0.0776	-0.4727	-0.1717
	$\beta_{2,\text{region}_8}$	0.0189	0.0074	0.0045	0.0330		$\kappa^*_{1,\text{year}_{11}}$	-0.3302	0.0823	-0.4918	-0.1688
	$\beta_{2,\text{region}_9}$	-0.0152	0.0083	-0.0320	0.0006		$\kappa^*_{1,\text{year}_{12}}$	-0.3384	0.0870	-0.5060	-0.1723
	β_3	-0.1079	0.0181	-0.1409	-0.0673		$\kappa^*_{1,\text{year}_{13}}$	-0.3466	0.0910	-0.5243	-0.1722
	κ_{1,year_2}	-0.0137	0.0129	-0.0391	0.0131		$\kappa^*_{1,\text{year}_{14}}$	-0.3550	0.0949	-0.5470	-0.1732
	κ_{1,year_3}	-0.0396	0.0135	-0.0673	-0.0135		$\kappa^*_{1,\text{year}_{15}}$	-0.3631	0.0989	-0.5630	-0.1711
	κ_{1,year_4}	-0.0540	0.0146	-0.0835	-0.0265		$\kappa^*_{1,\text{year}_{16}}$	-0.3716	0.1029	-0.5788	-0.1732
	κ_{1,year_5}	-0.0585	0.0151	-0.0882	-0.0286		$\kappa^*_{1,\text{year}_{17}}$	-0.3793	0.1073	-0.6001	-0.1722
	κ_{1,year_6}	-0.0723	0.0150	-0.1019	-0.0426		$\kappa^*_{1,\text{year}_{18}}$	-0.3870	0.1116	-0.6129	-0.1692
	κ_{1,year_7}	-0.0954	0.0156	-0.1256	-0.0632		ψ_{κ_1}	-0.0081	0.0037	-0.0156	-0.0013
	κ_{1,year_8}	-0.1065	0.0159	-0.1373	-0.0743		σ^2	0.0038	0.0004	0.0031	0.0046
	κ_{1,year_9}	-0.1269	0.0158	-0.1578	-0.0961		$\sigma^2_{\kappa_1}$	0.0004	0.0002	0.0002	0.0009

Appendix C

Age-specific Fitted and Projected Mortality Rates

C.1 Findings on female lung cancer mortality based on the model in (3.4), without smoking data

C.1.1 Pearson residuals

Figure C.1 and Figure C.2 respectively show Pearson residuals based on the female LC model, without smoking information, depending on (3.4), fitted in the most (1) and least (5) deprived quintiles of the regions of England from 2001 to 2018. The figure does not point out any significant non-random clustering over ages or years.

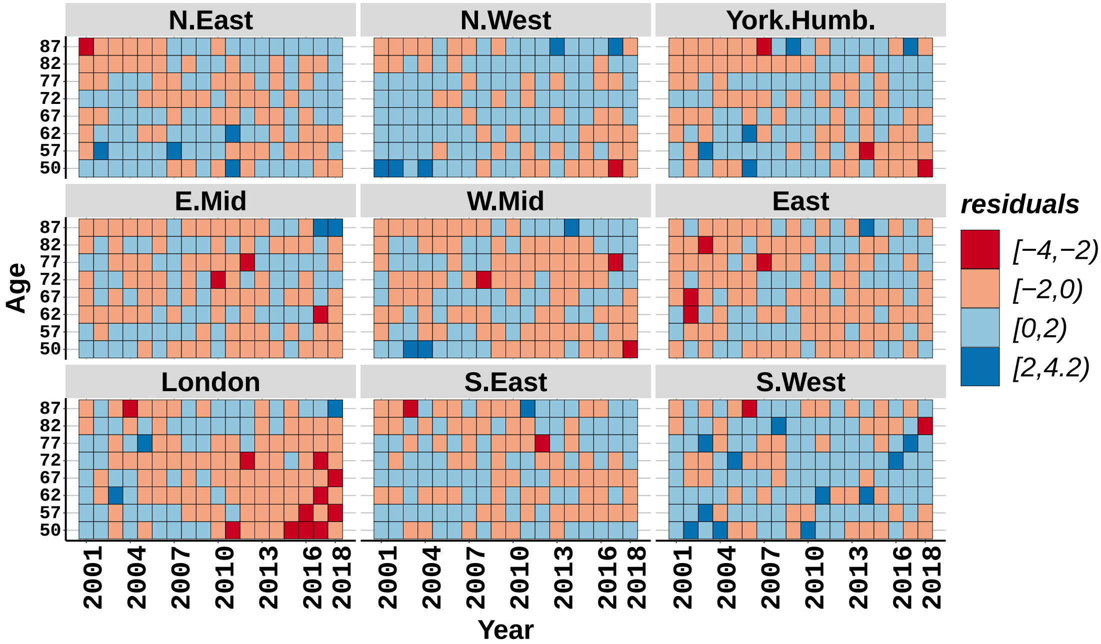


Figure C.1: Heat map of Pearson residuals for **female** lung cancer mortality in regions of England, **deprivation quintile 1 (most deprived), excluding smoking data**, based on (3.4): orange/light blue cells indicate areas with good fit, while red/dark blue cells indicate areas with poor fit. Note that there is a small number of residuals greater than 4, and these are included in the last category.

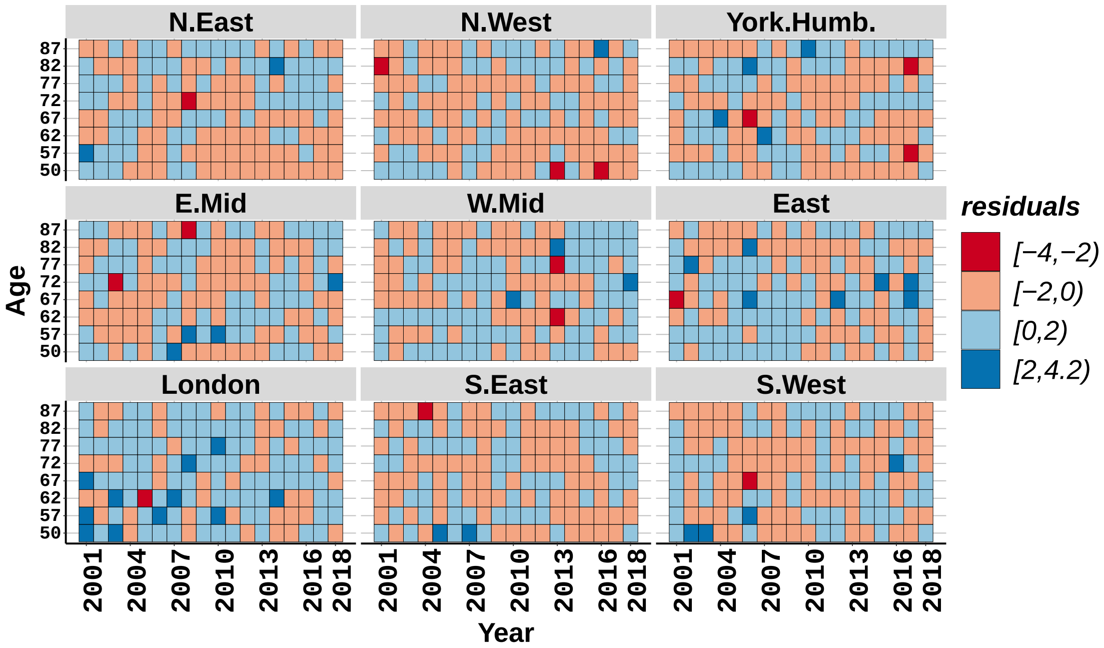


Figure C.2: Heat map of Pearson residuals for **female** lung cancer mortality in regions of England, **deprivation quintile 5 (least deprived), excluding smoking data**, based on (3.4): orange/light blue cells indicate areas with good fit, while red/dark blue cells indicate areas with poor fit. Note that there is a small number of residuals greater than 4, and these are included in the last category.

C.1.2 Age-specific rates from 2001 to 2036

We present fitted and projected LC mortality in women across various age groups, using the mid-age of each group as a reference, for selected deprivation quintiles (1, 3, and 5) in the regions of England from 2001 to 2036. These projections are based on (3.4), which excludes NS prevalence rates.

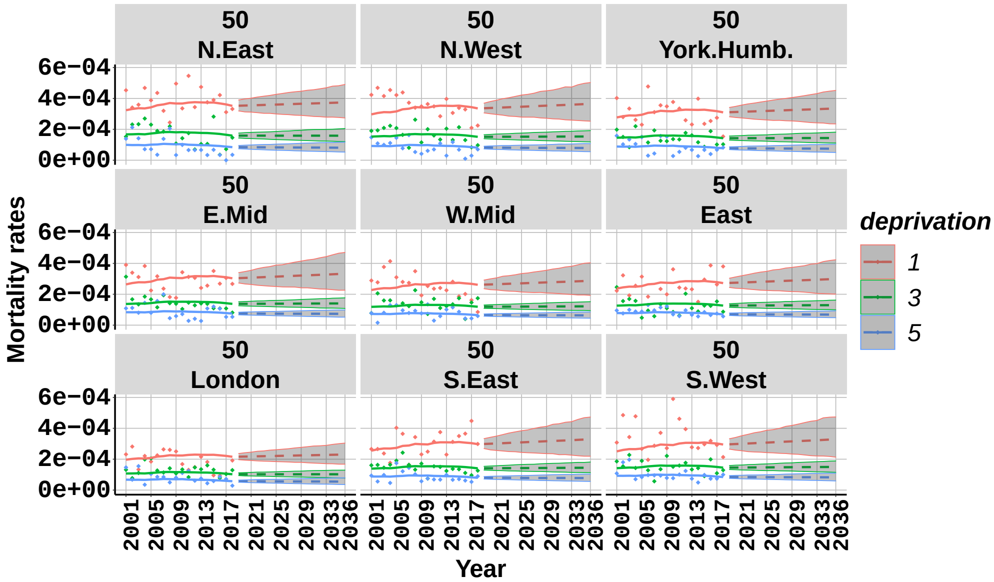


Figure C.3: Lung cancer mortality, **females**, **ages at death 50**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.4), **excluding smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

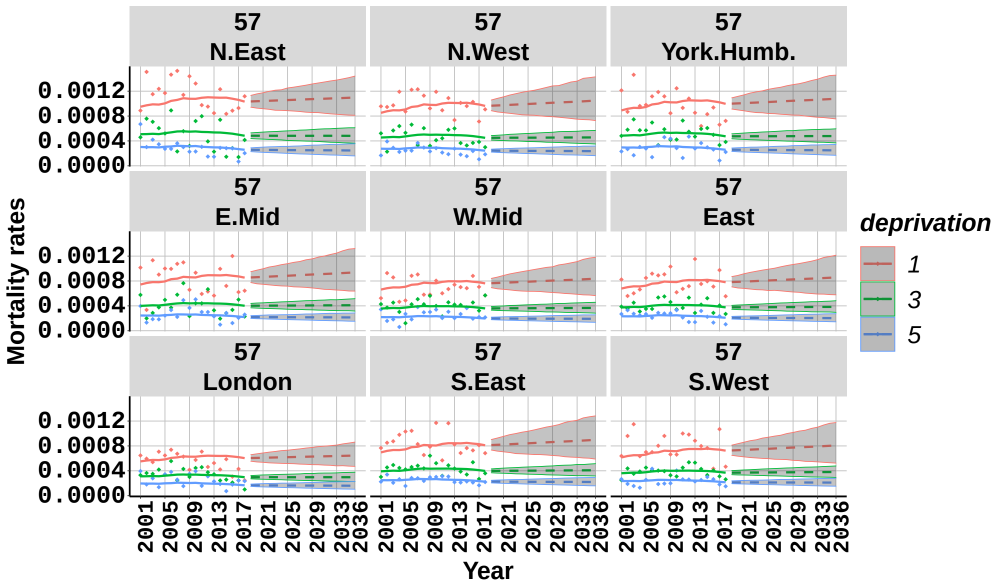


Figure C.4: Lung cancer mortality, **females**, **ages at death 57**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.4), **excluding smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

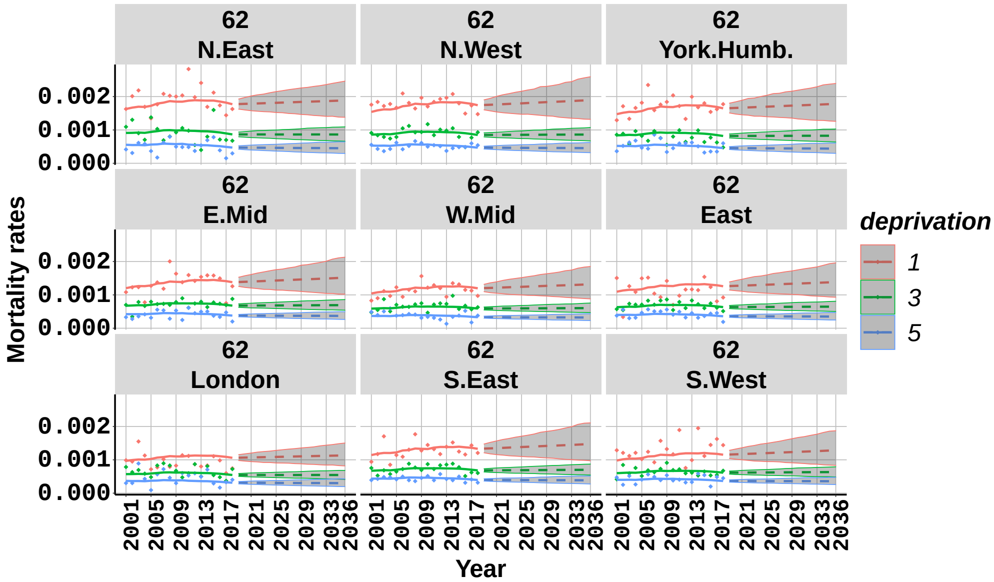


Figure C.5: Lung cancer mortality, **females**, **ages at death 62**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.4), **excluding smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

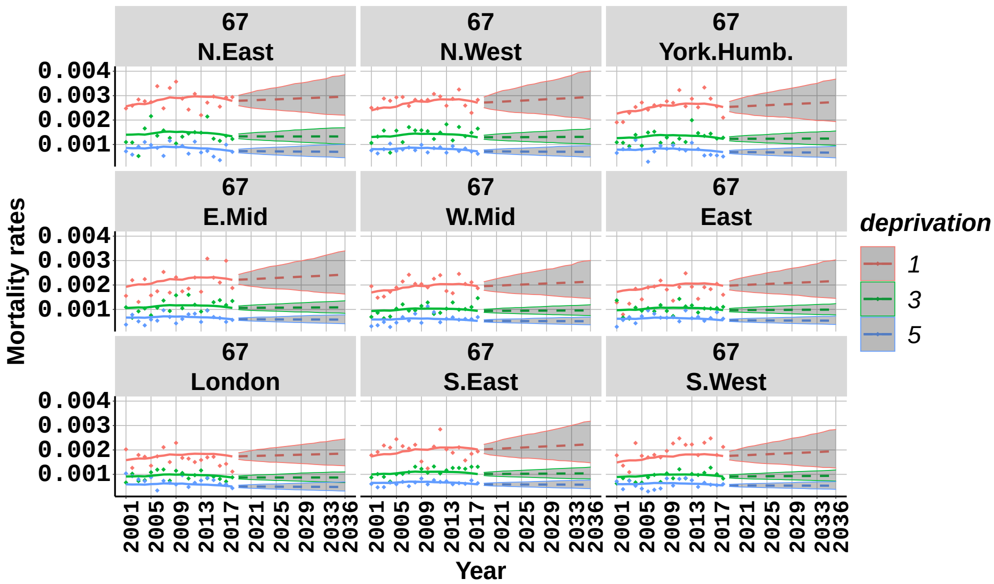


Figure C.6: Lung cancer mortality, **females**, **ages at death 67**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.4), **excluding smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

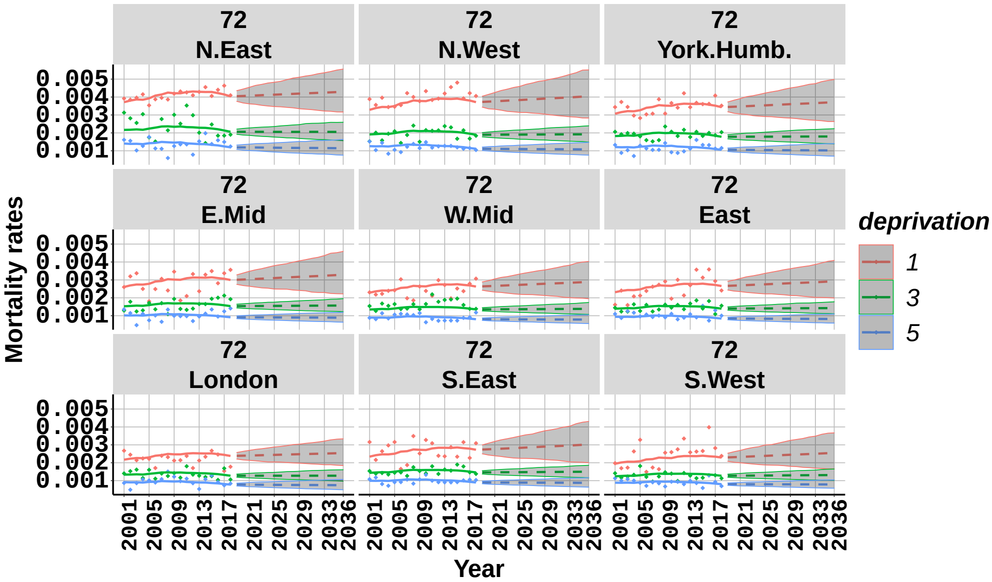


Figure C.7: Lung cancer mortality, **females**, **ages at death 72**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.4), **excluding smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

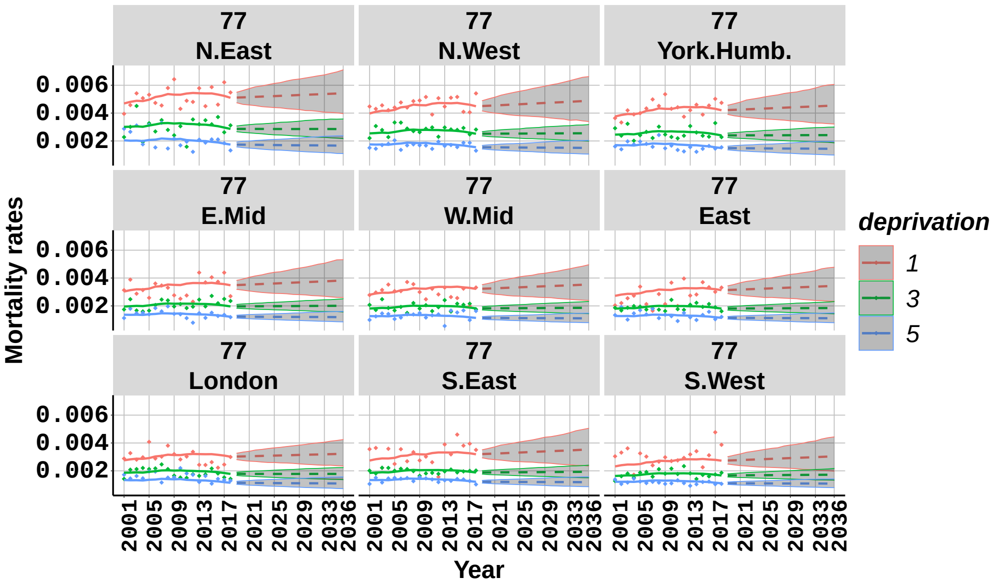


Figure C.8: Lung cancer mortality, **females**, **ages at death 77**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.4), **excluding smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

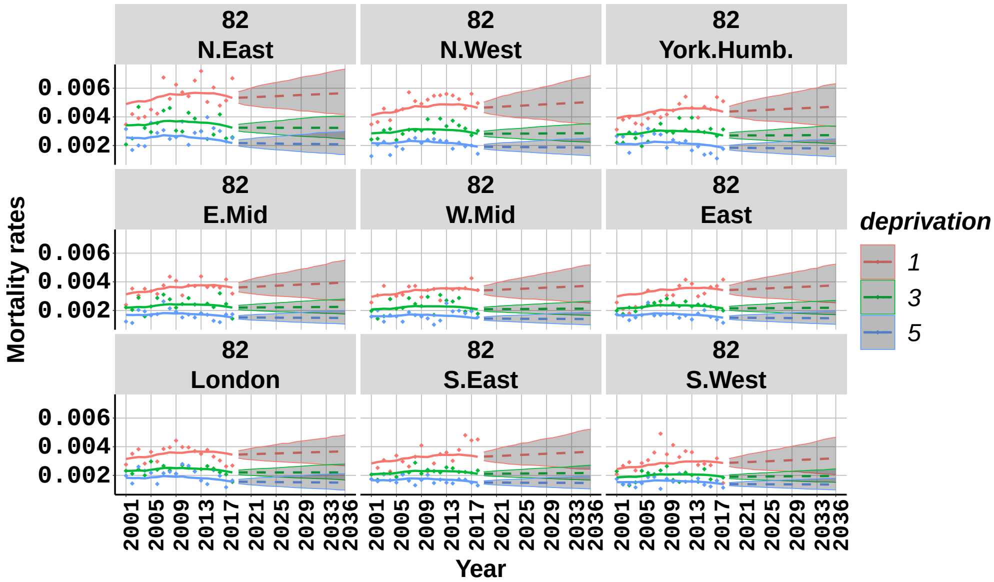


Figure C.9: Lung cancer mortality, **females**, **ages at death 82**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.4), **excluding smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

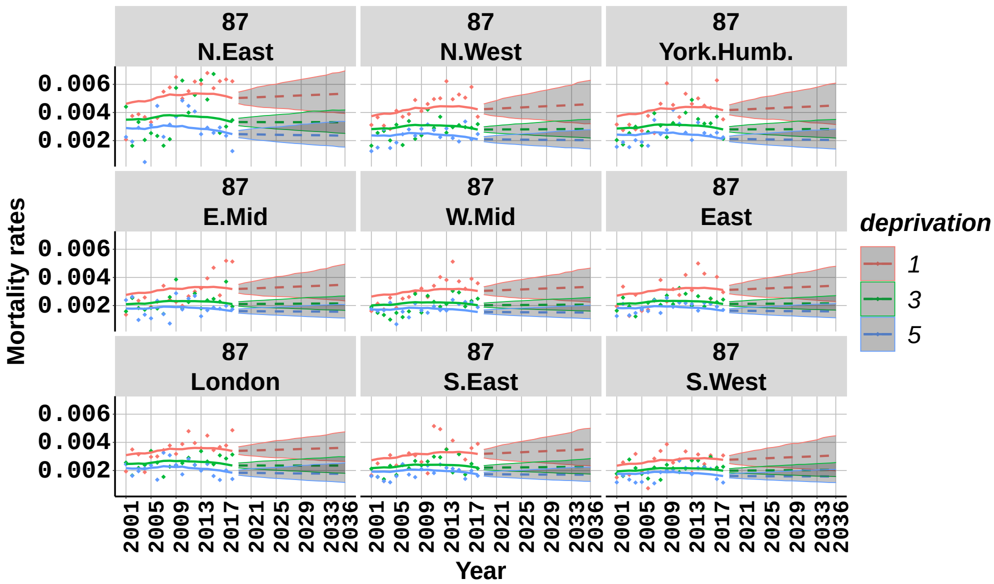


Figure C.10: Lung cancer mortality, **females**, **ages at death 87**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.4), **excluding smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

C.2 Findings on female lung cancer mortality based on the model in (3.5), with smoking data

C.2.1 Pearson residuals

Figure C.11 and Figure C.12 demonstrate the Pearson residuals in the same granularity with Figure C.1 and Figure C.2, depending on (3.5), where we additionally introduce NS prevalence rates and an interaction term between AAD and region into our modelling framework (suggested by variable selection in Appendix A.1). This model indicates a better fit, as compared to the model in (3.4), as demonstrated with, for example, fewer red/dark blue cells in Figure C.11 compared to Figure C.1. Besides, there is no significant non-random clustering over ages or years.

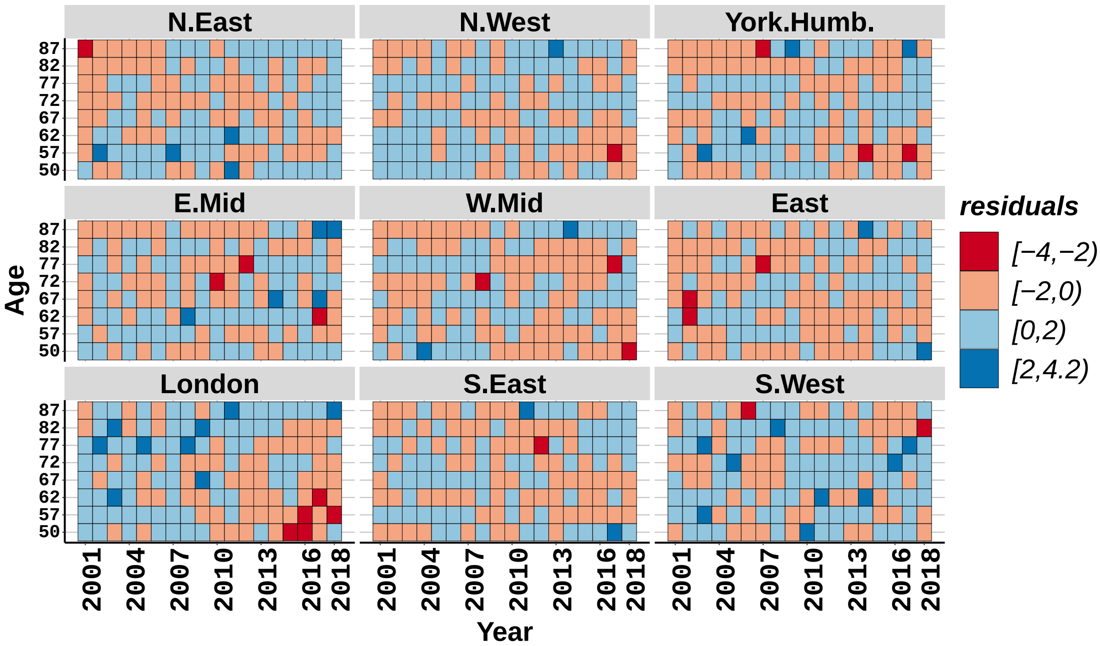


Figure C.11: Heat map of Pearson residuals for **female** lung cancer mortality in regions of England, **deprivation quintile 1 (most deprived), including smoking data**, based on (3.5): orange/light blue cells indicate areas with good fit, while red/dark blue cells indicate areas with poor fit. Note that there is a small number of residuals greater than 4, and these are included in the last category.

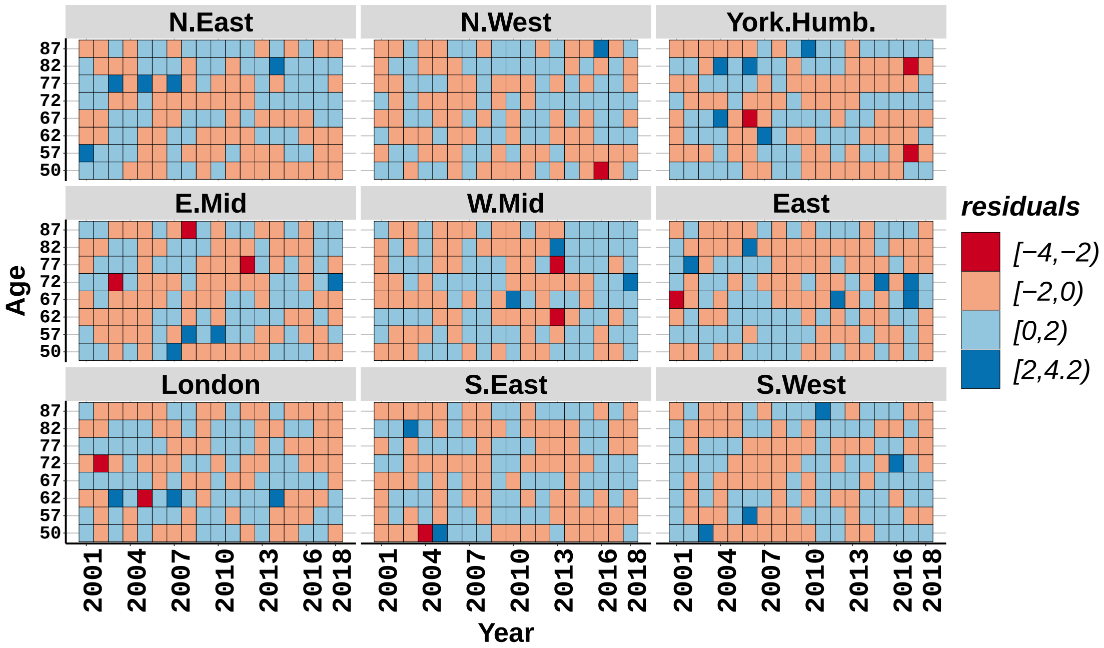


Figure C.12: Heat map of Pearson residuals for **female** lung cancer mortality in regions of England, **deprivation quintile 5 (least deprived), including smoking data**, based on (3.5): orange/light blue cells indicate areas with good fit, while red/dark blue cells indicate areas with poor fit. Note that there is a small number of residuals greater than 4, and these are included in the last category.

C.2.2 Age-specific rates from 2001 to 2036

We present fitted and projected LC mortality in women across various age groups, using the mid-age of each group as a reference, for selected deprivation quintiles (1, 3, and 5) in the regions of England from 2001 to 2036. These projections are based on (3.5), which includes NS prevalence rates.

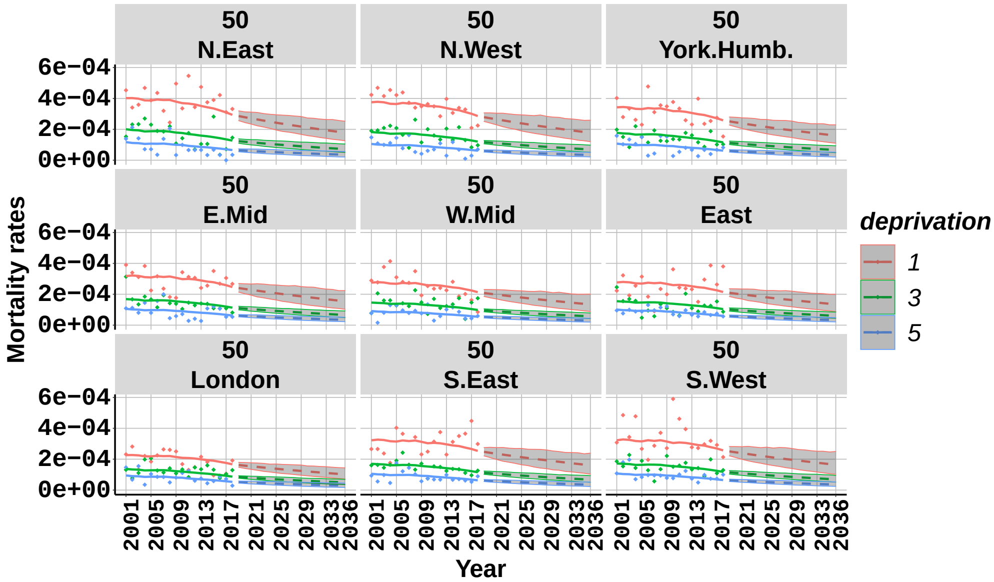


Figure C.13: Lung cancer mortality, **females**, **ages at death 50**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.5), **including smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

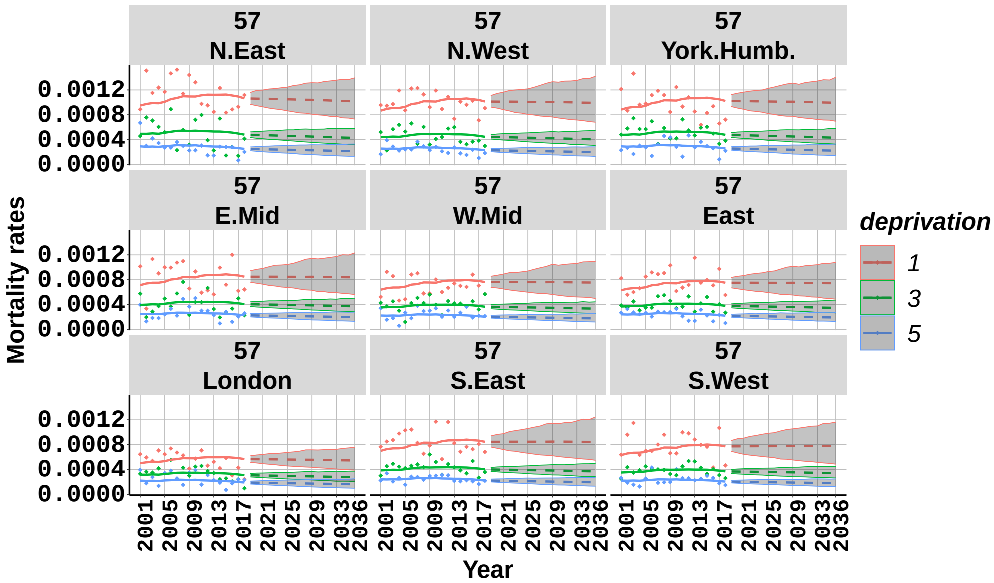


Figure C.14: Lung cancer mortality, **females**, **ages at death 57**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.5), **including smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

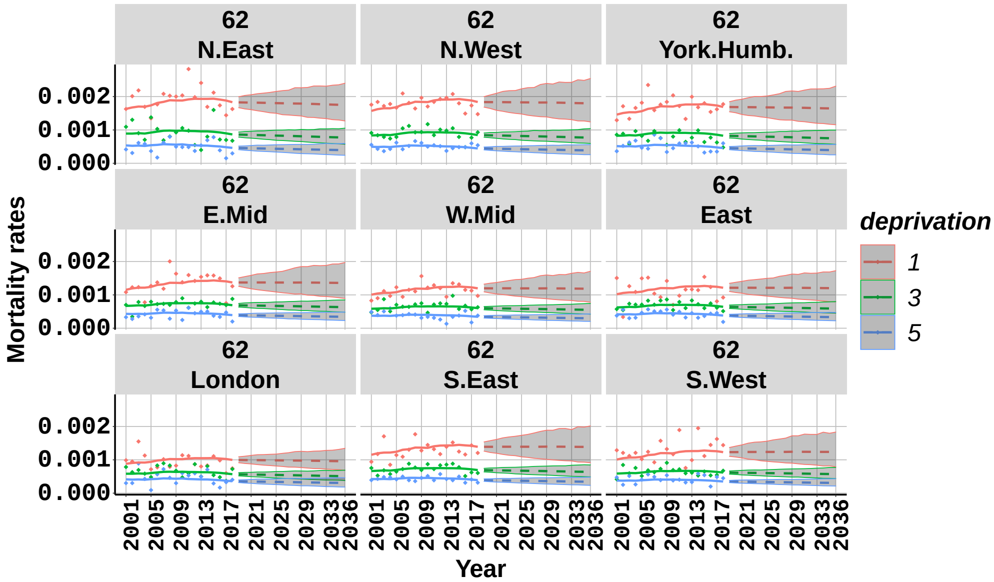


Figure C.15: Lung cancer mortality, **females**, **ages at death 62**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.5), **including smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

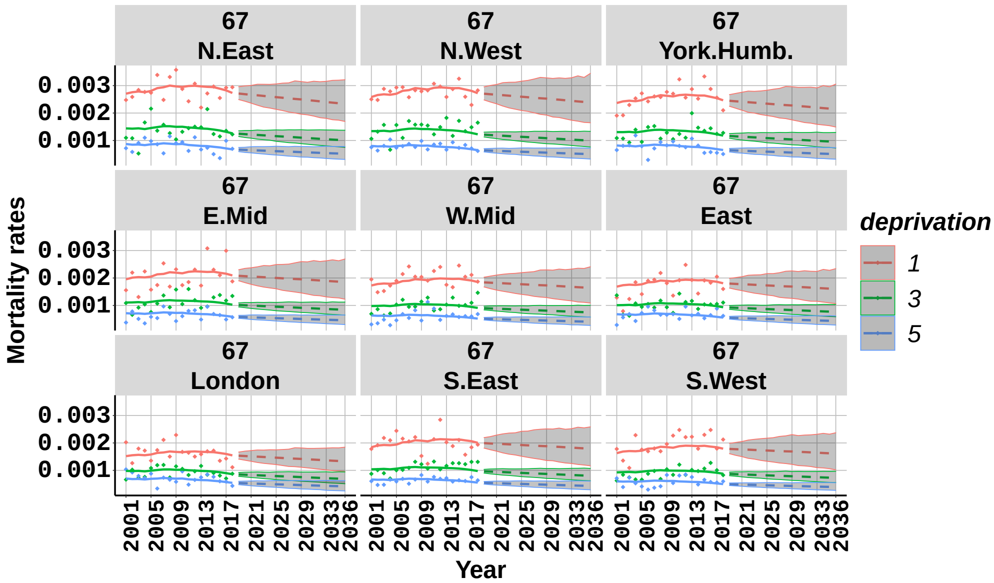


Figure C.16: Lung cancer mortality, **females**, **ages at death 67**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.5), **including smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

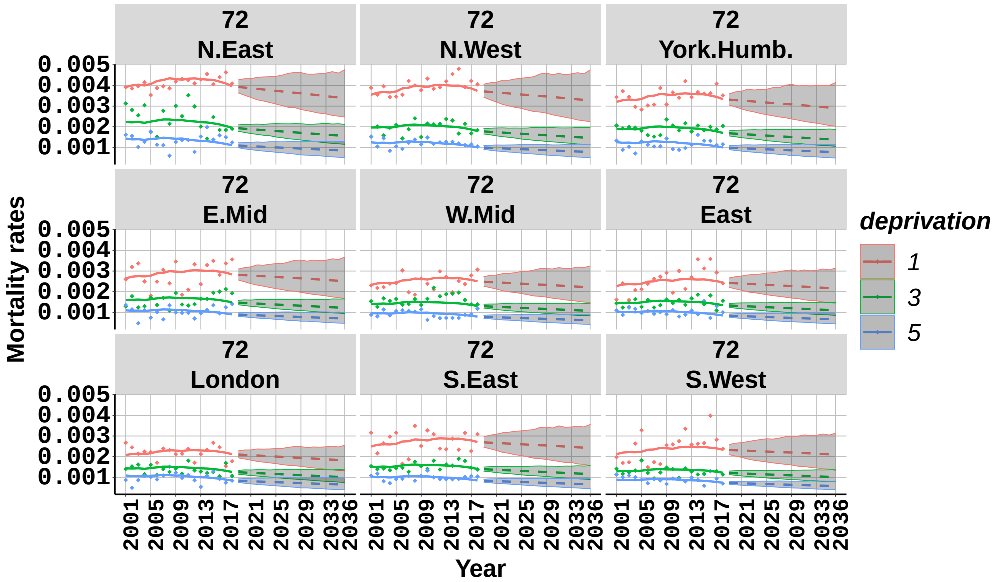


Figure C.17: Lung cancer mortality, **females**, **ages at death 72**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.5), **including smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

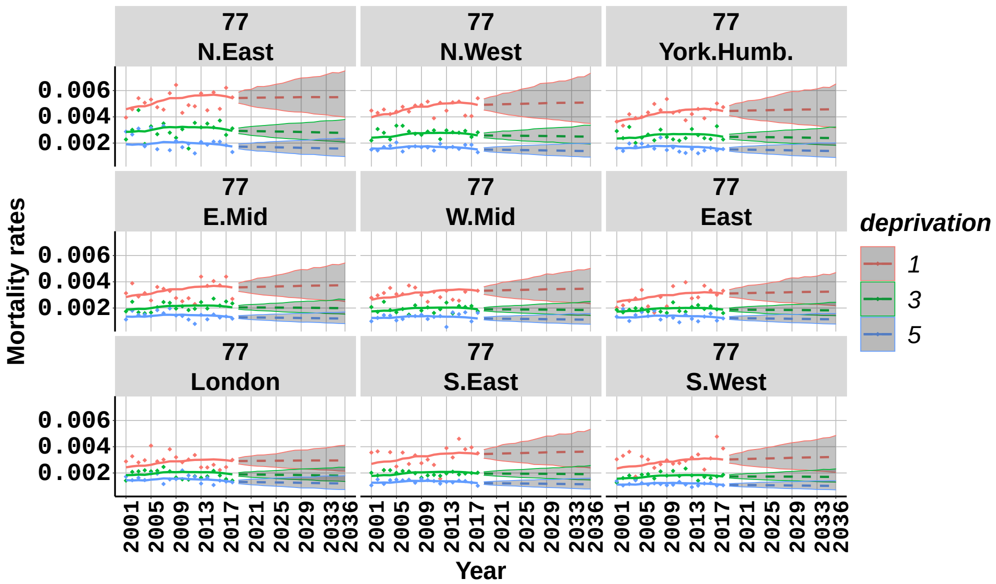


Figure C.18: Lung cancer mortality, **females**, **ages at death 77,** in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.5), **including smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

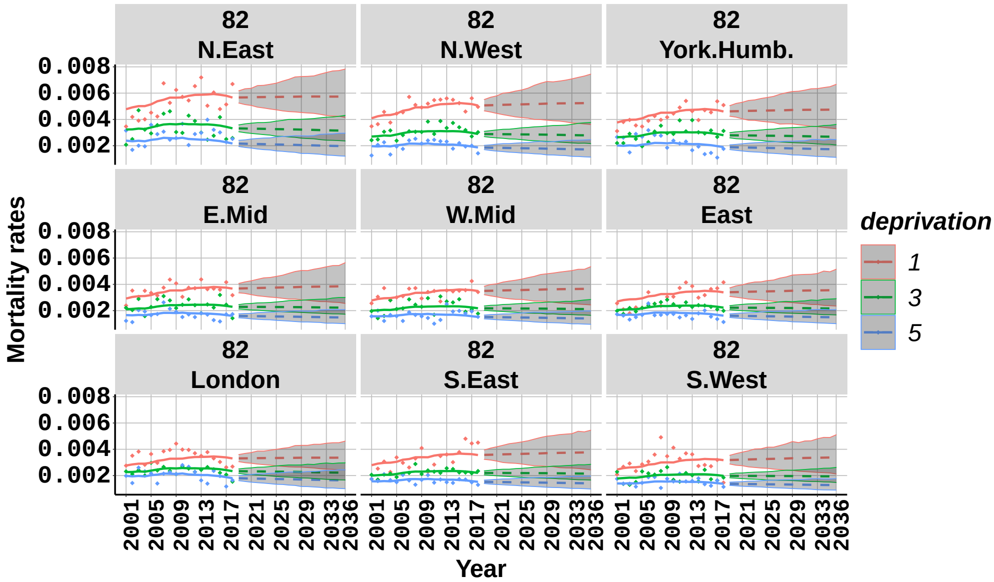


Figure C.19: Lung cancer mortality, **females**, **ages at death 82**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived)in regions of England based on (3.5), **including smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

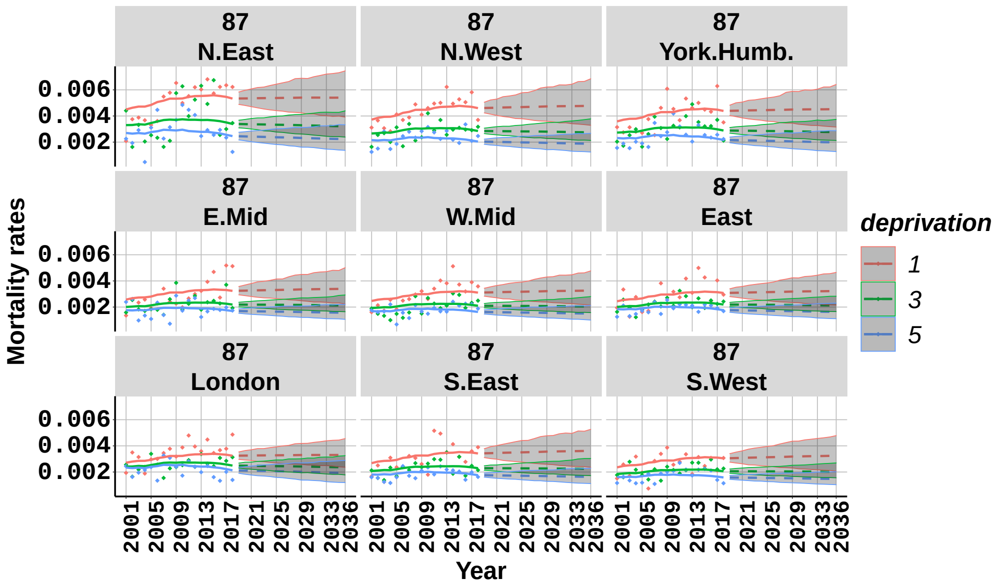


Figure C.20: Lung cancer mortality, **females**, **ages at death 87**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.5), **including smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

C.3 Findings on male lung cancer mortality based on the model in (3.6), without smoking data

C.3.1 Pearson residuals

Figure C.21 and Figure C.22 display Pearson residuals across different age groups between 2001 and 2018 on the male LC model in (3.6). Similar to the ones reported for female LC mortality, the residuals are shown in deprivation quintiles 1 (most deprived) and 5 (least deprived) in the regions of England. Similarly, we do not see any non-random clustering across ages or years that points out a good fit.

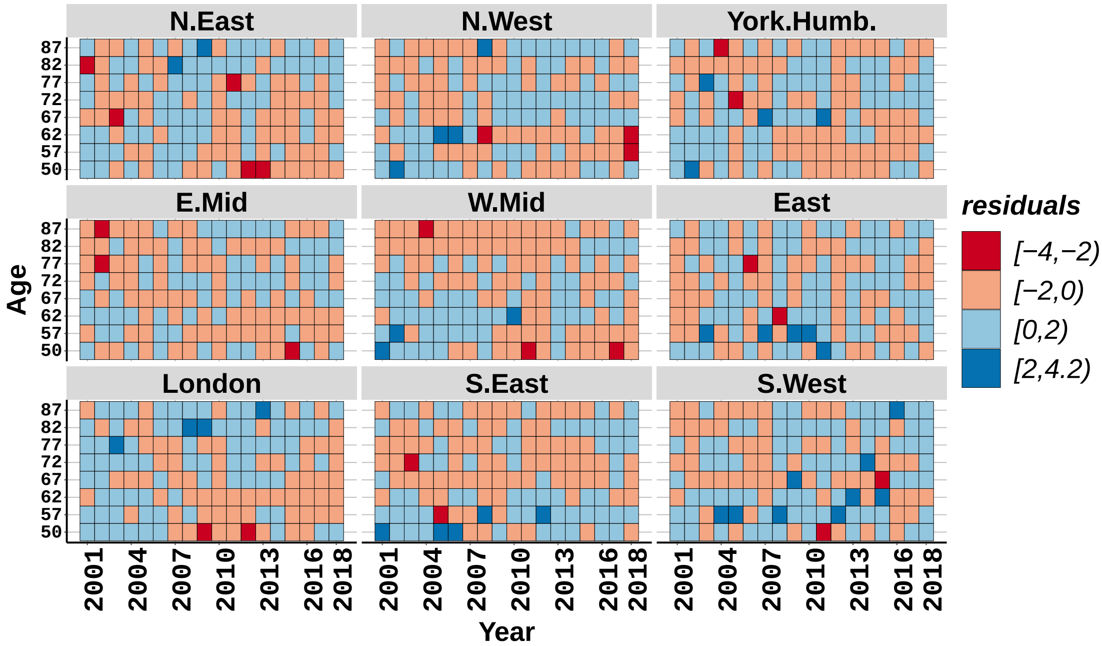


Figure C.21: Heat map of Pearson residuals for **male** lung cancer mortality in regions of England, **deprivation quintile 1 (most deprived), excluding smoking data**, based on (3.6): orange/light blue cells indicate areas with good fit, while red/dark blue cells indicate areas with poor fit. Note that there is a small number of residuals greater than 4, and these are included in the last category.

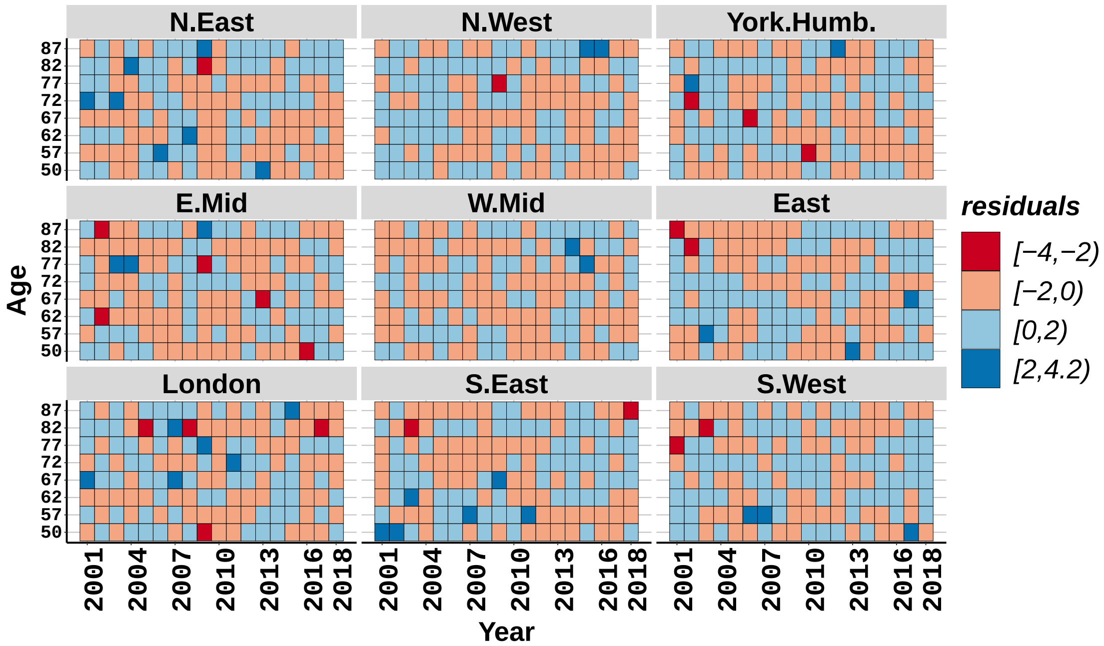


Figure C.22: Heat map of Pearson residuals for **male** lung cancer mortality in regions of England, **deprivation quintile 5 (least deprived), excluding smoking data**, based on (3.6): orange/light blue cells indicate areas with good fit, while red/dark blue cells indicate areas with poor fit. Note that there is a small number of residuals greater than 4, and these are included in the last category.

C.3.2 Age-specific rates from 2001 to 2036

We present fitted and projected LC mortality in men across various age groups, using the mid-age of each group as a reference, for selected deprivation quintiles (1, 3, and 5) in the regions of England from 2001 to 2036. These projections are based on (3.6), which excludes NS prevalence rates.

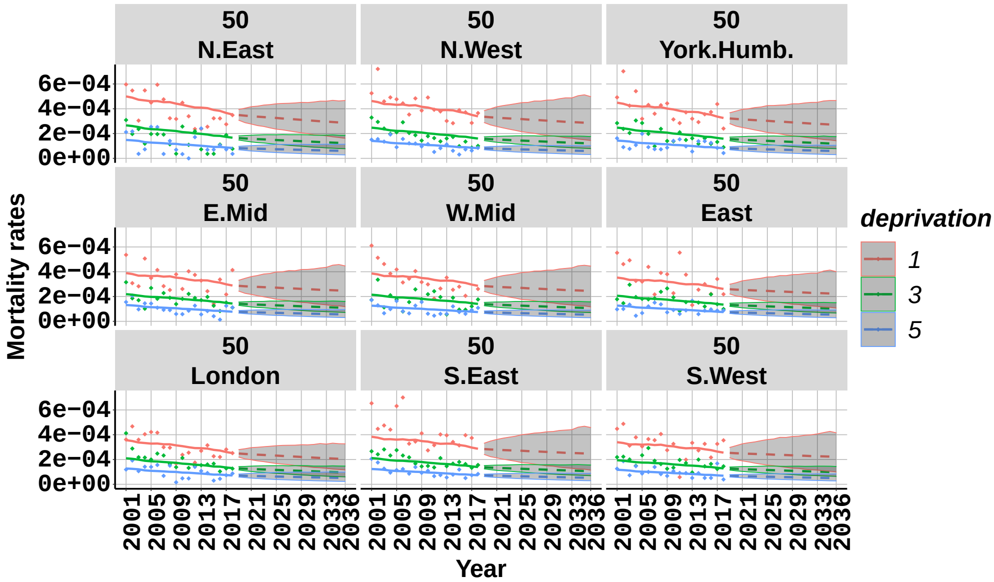


Figure C.23: Lung cancer mortality, **males, ages at death 50**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.6), **excluding smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

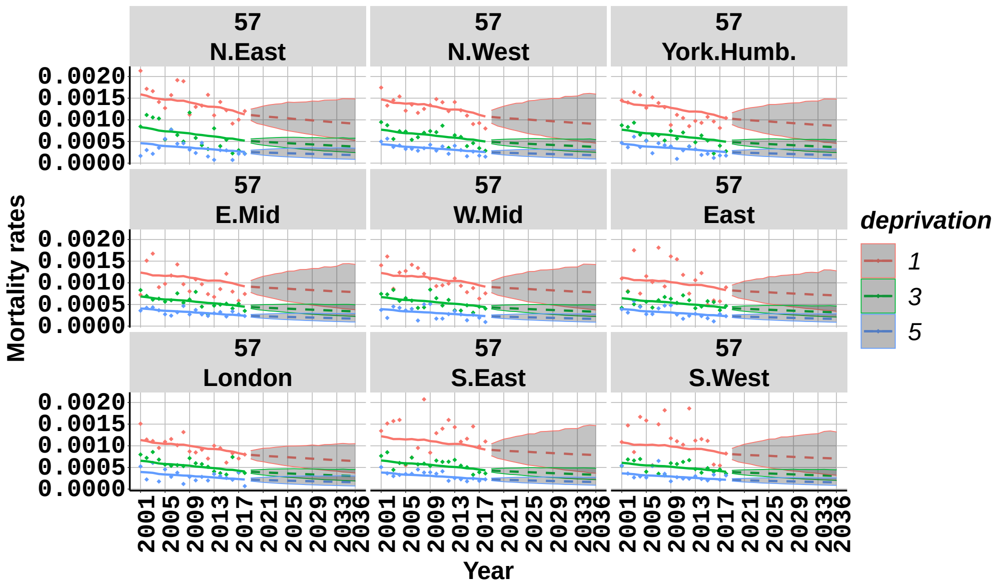


Figure C.24: Lung cancer mortality, **males, ages at death 57**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.6), **excluding smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

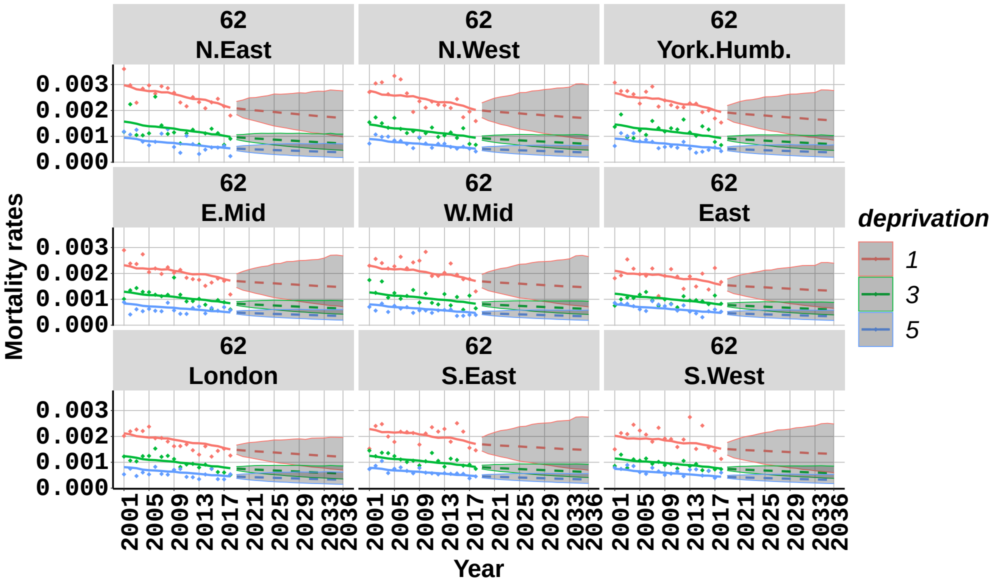


Figure C.25: Lung cancer mortality, **males, ages at death 62**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.6), **excluding smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

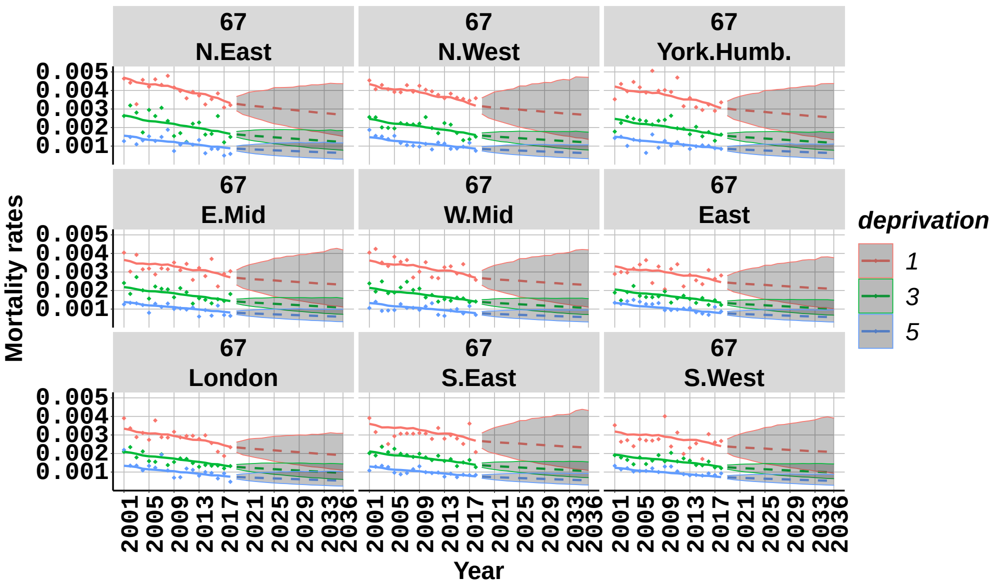


Figure C.26: Lung cancer mortality, **males, ages at death 67**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived)in regions of England based on (3.6), **excluding smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

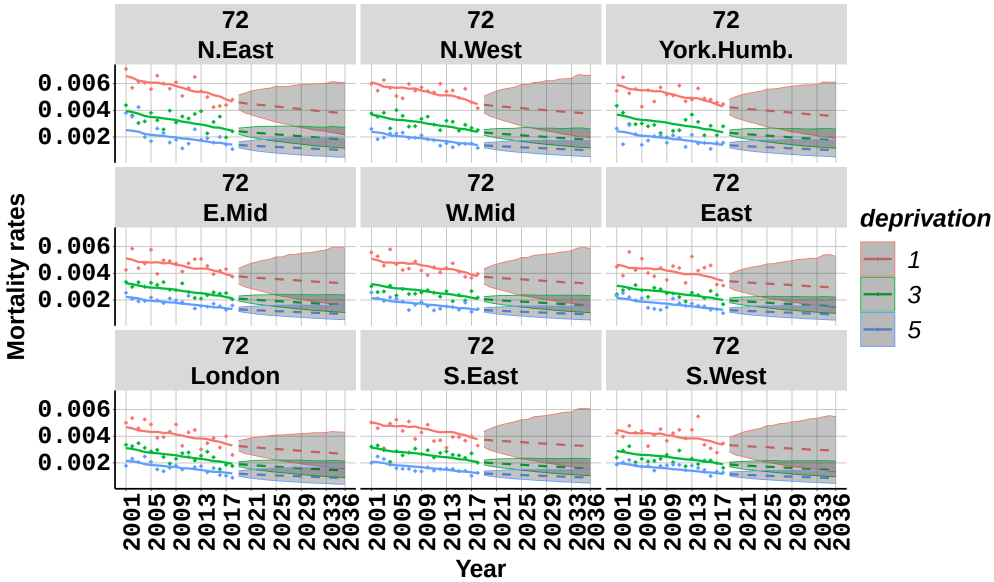


Figure C.27: Lung cancer mortality, **males, ages at death 72**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.6), **excluding smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

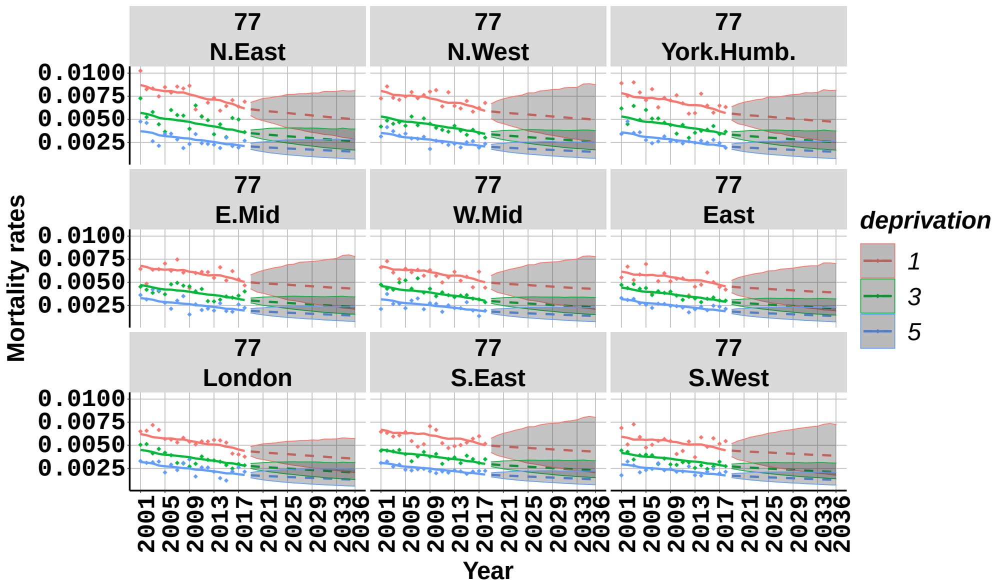


Figure C.28: Lung cancer mortality, **males, ages at death 77**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.6), **excluding smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

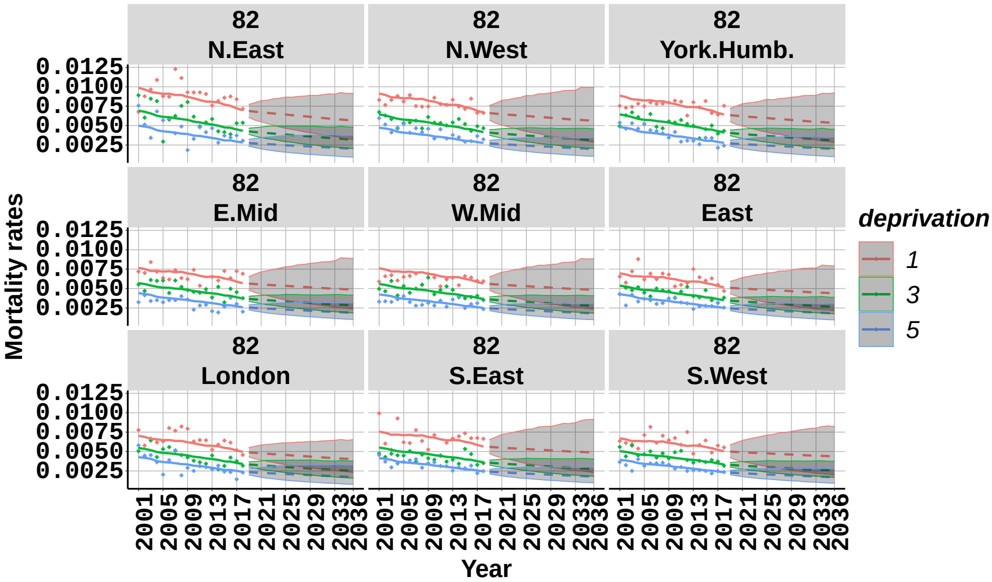


Figure C.29: Lung cancer mortality, **males, ages at death 82**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.6), **excluding smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

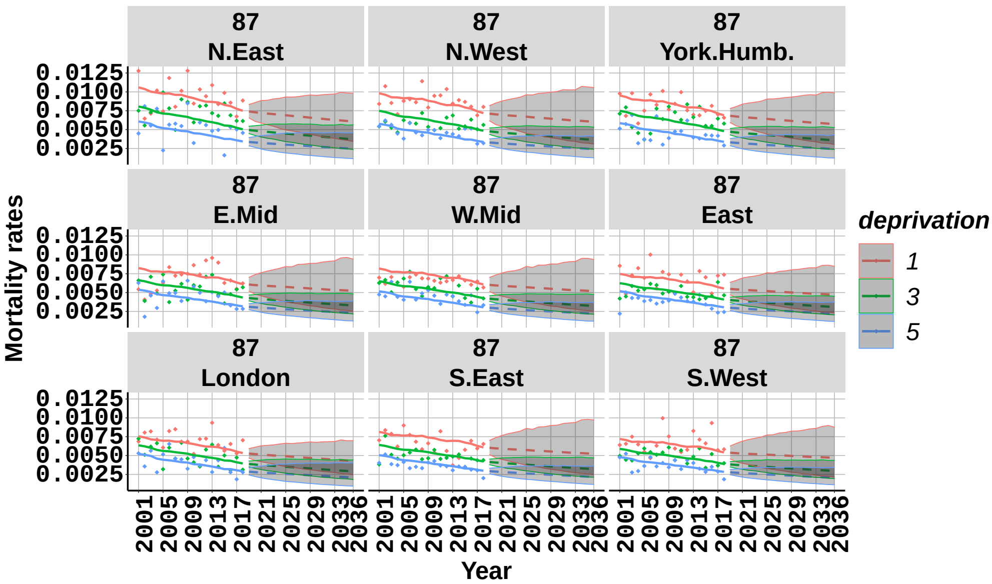


Figure C.30: Lung cancer mortality, **males, ages at death 87**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.6), **excluding smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

C.4 Findings on male lung cancer mortality based on the model in (3.7), with smoking data

C.4.1 Pearson residuals

Figure C.31 and Figure C.32 exhibit the residuals depending on (3.7). These figures present quite similar results shown in Figure C.21 and Figure C.22. Yet, the model in (3.7) suggests a better fit as evidenced by a lower DIC (41,814) and a higher marginal likelihood (-21,160) as compared to the earlier model results with 41,843 for the DIC and -21,174 for the marginal likelihood (see Table A.4 and Table A.5).

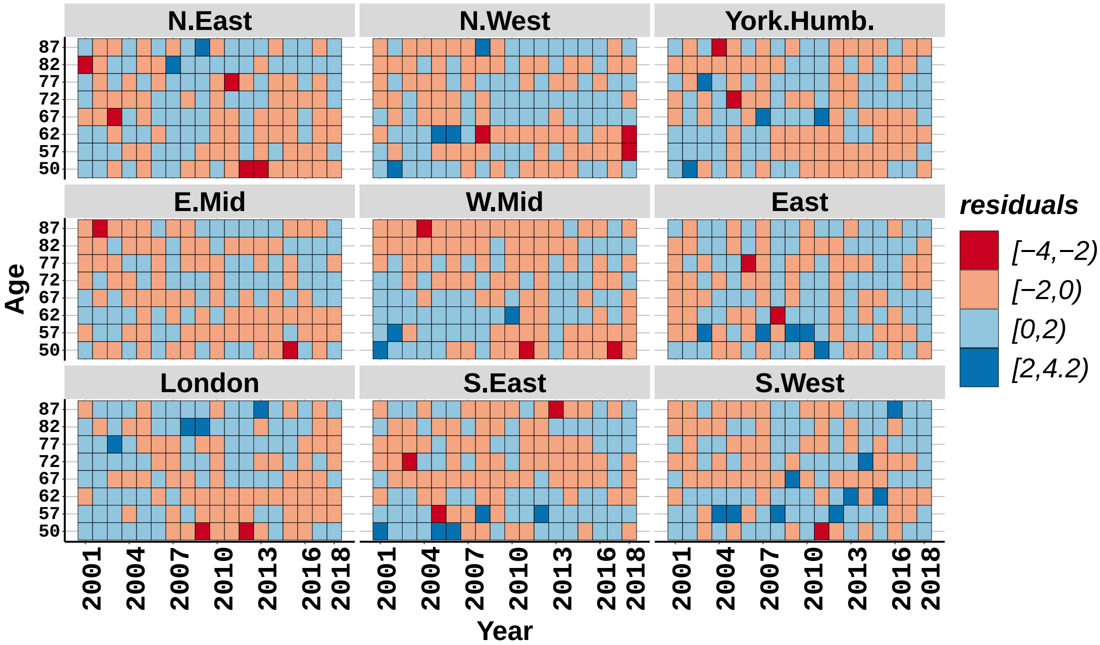


Figure C.31: Heat map of Pearson residuals for **male** lung cancer mortality in regions of England, **deprivation quintile 1 (most deprived), including smoking data**, based on (3.7): orange/light blue cells indicate areas with good fit, while red/dark blue cells indicate areas with poor fit. Note that there is a small number of residuals greater than 4, and these are included in the last category.

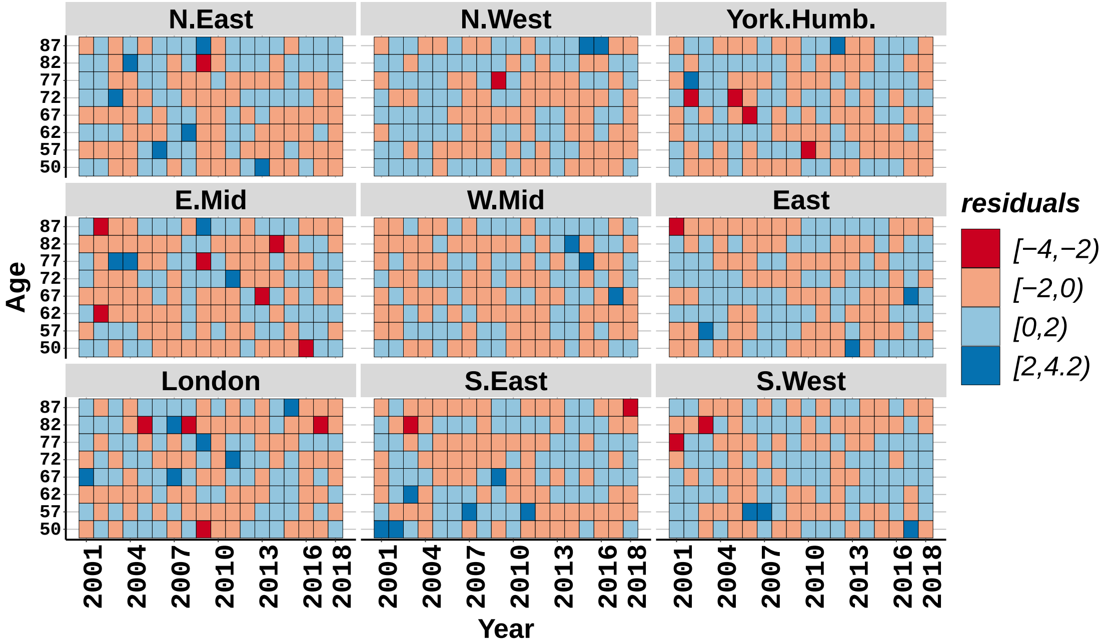


Figure C.32: Heat map of Pearson residuals for **male** lung cancer mortality in regions of England, **deprivation quintile 5 (least deprived), including smoking data**, based on (3.7): orange/light blue cells indicate areas with good fit, while red/dark blue cells indicate areas with poor fit. Note that there is a small number of residuals greater than 4, and these are included in the last category.

C.4.2 Age-specific rates from 2001 to 2036

We present fitted and projected LC mortality in men across various age groups, using the mid-age of each group as a reference, for selected deprivation quintiles (1, 3, and 5) in the regions of England from 2001 to 2036. These projections are based on (3.7), which includes NS prevalence rates.

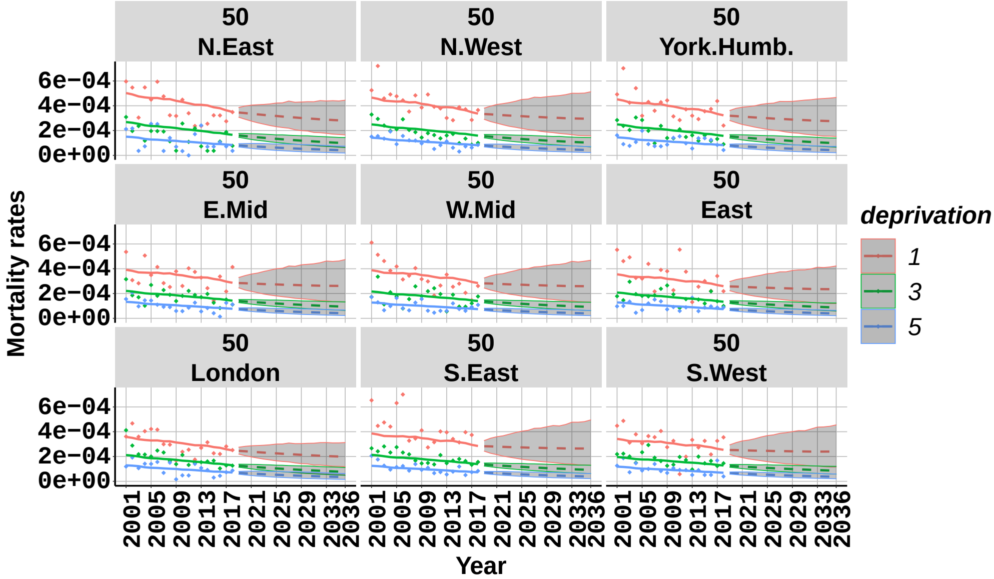


Figure C.33: Lung cancer mortality, **males, ages at death 50**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.7), **including smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

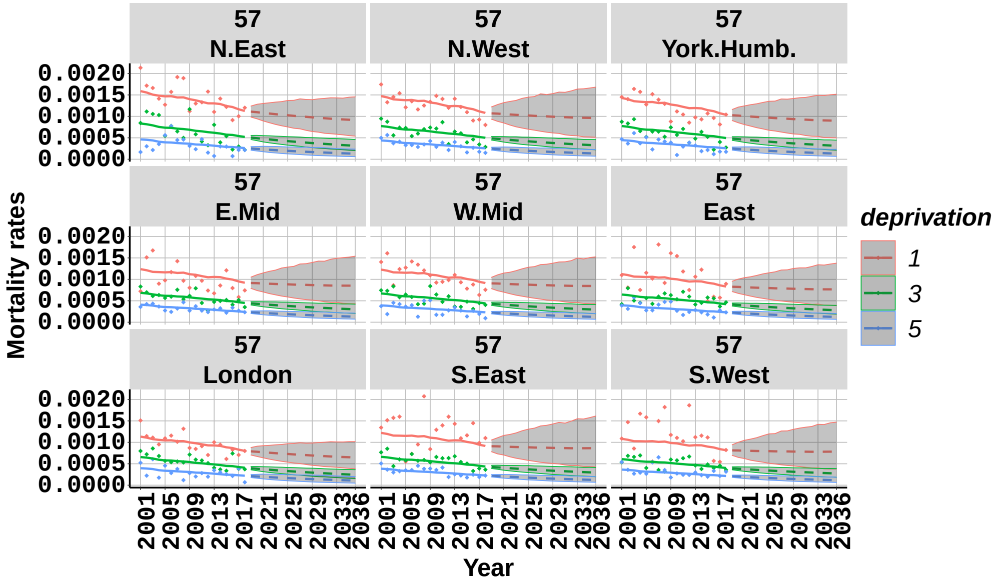


Figure C.34: Lung cancer mortality, **males, ages at death 57**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.7), **including smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

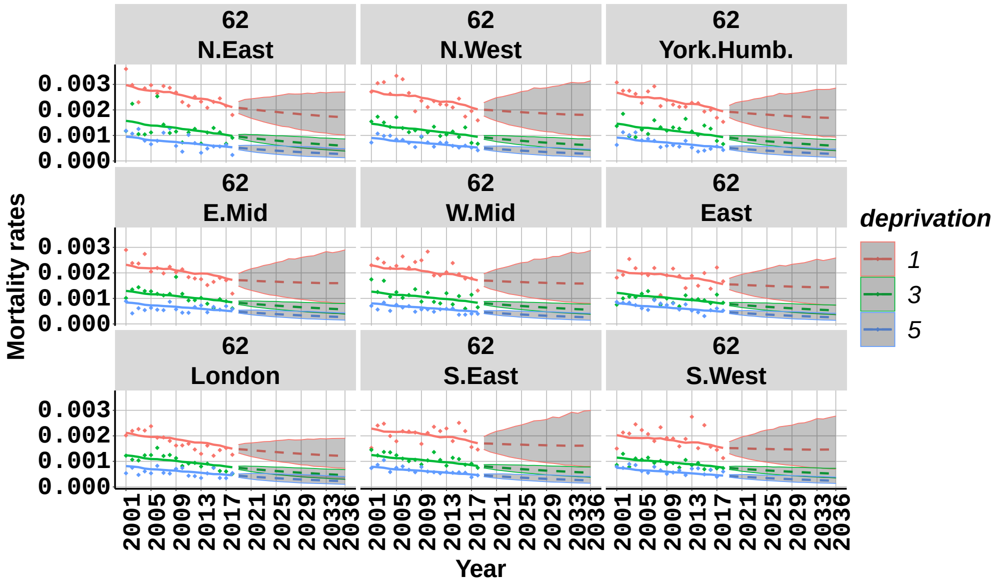


Figure C.35: Lung cancer mortality, **males, ages at death 62**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.7), **including smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

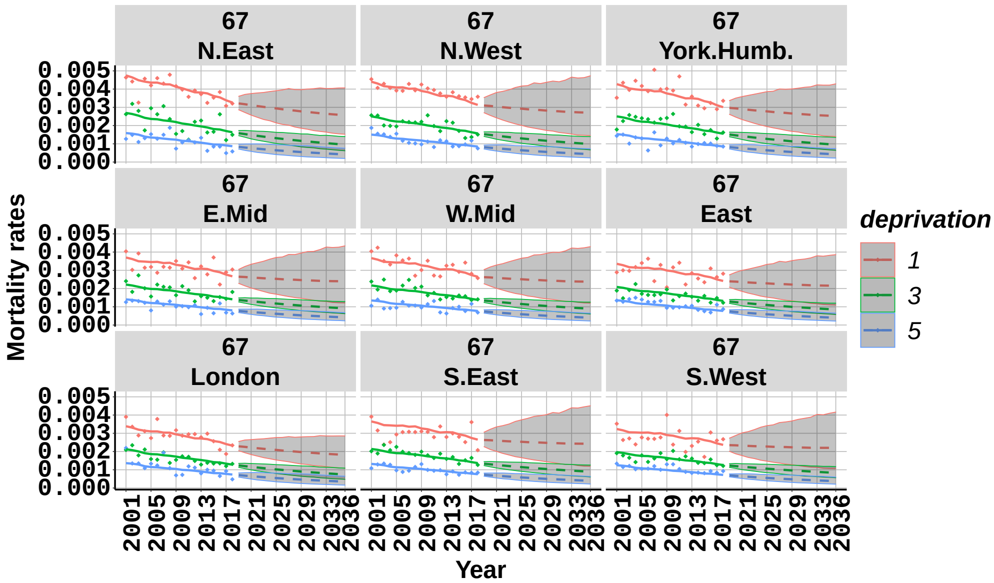


Figure C.36: Lung cancer mortality, **males, ages at death 67**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived)in regions of England based on (3.7), **including smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

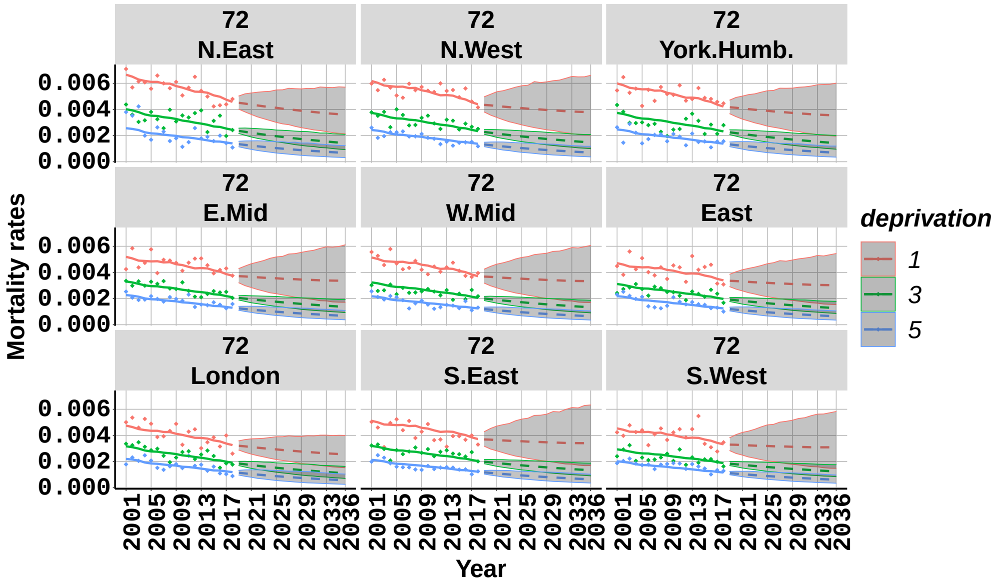


Figure C.37: Lung cancer mortality, **males, ages at death 72**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.7), **including smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

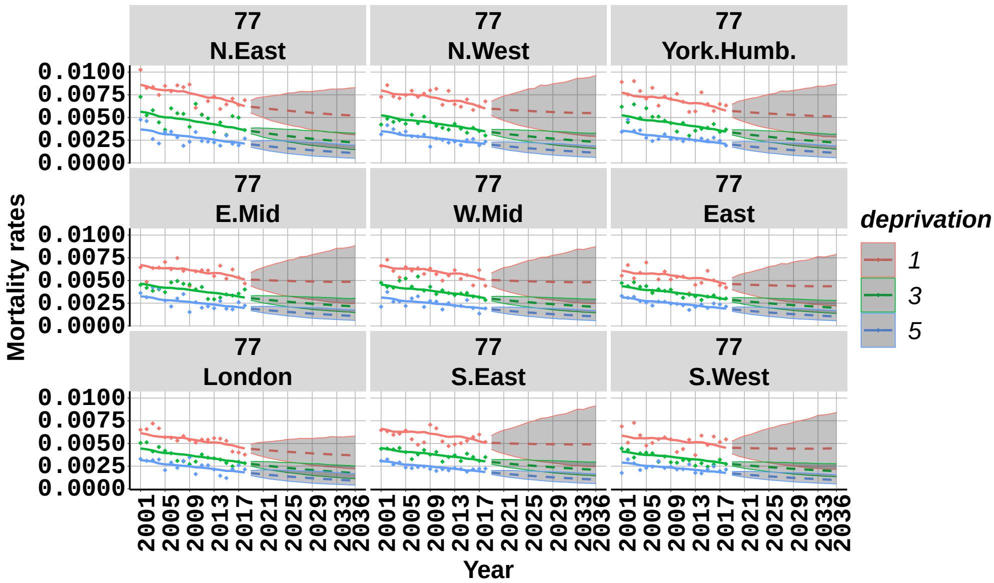


Figure C.38: Lung cancer mortality, **males, ages at death 77**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.7), **including smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

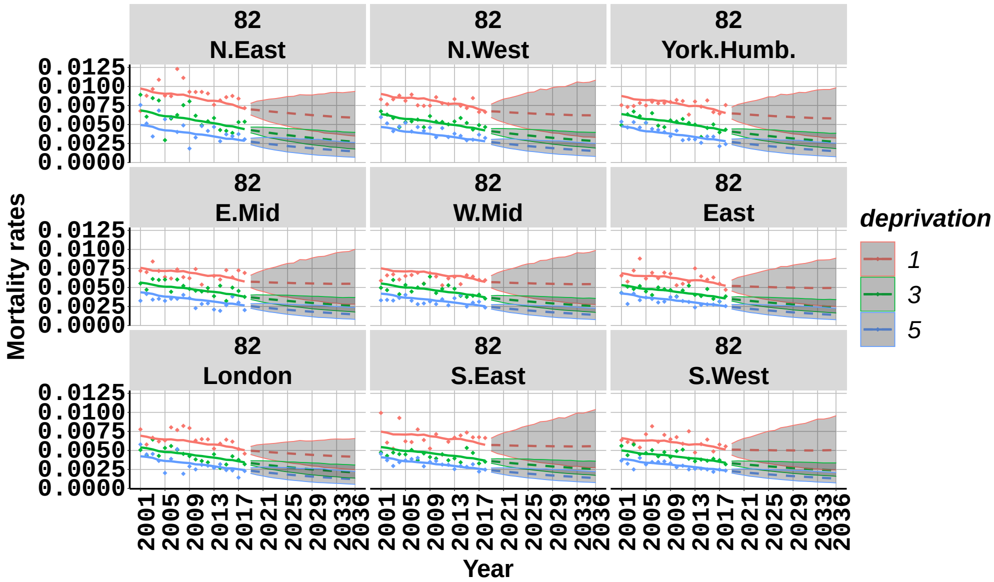


Figure C.39: Lung cancer mortality, **males, ages at death 82**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.7), **including smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

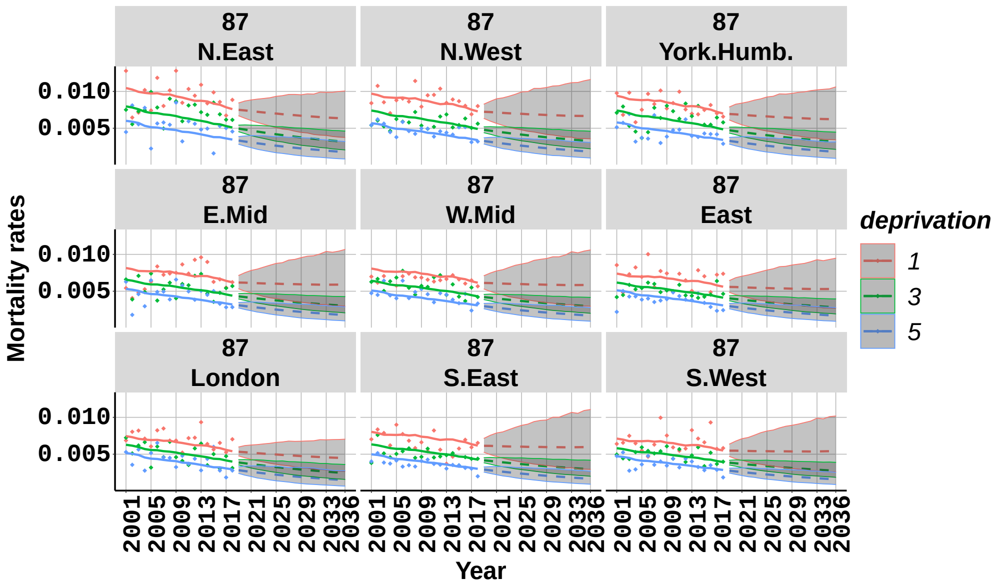
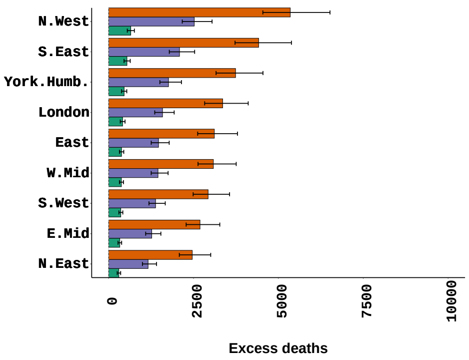
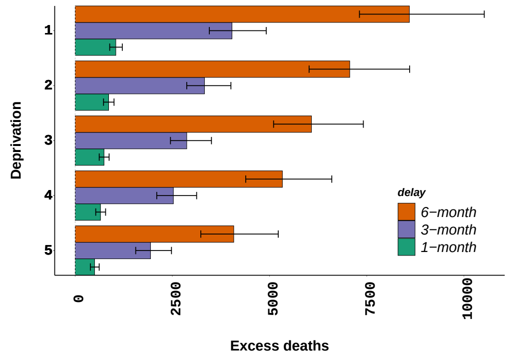


Figure C.40: Lung cancer mortality, **males, ages at death 87**, in selected deprivation quintiles 1 (most deprived), 3, and 5 (least deprived) in regions of England based on (3.7), **including smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

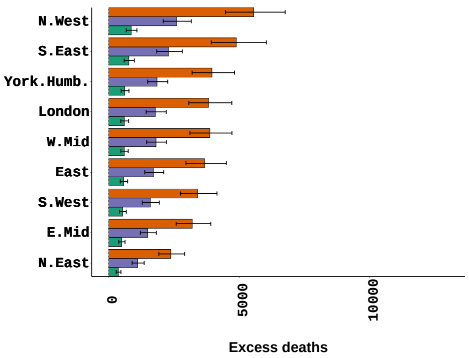
C.5 Excess lung cancer deaths and mortality based on the models in (3.4) for women and (3.6) for men, without smoking data



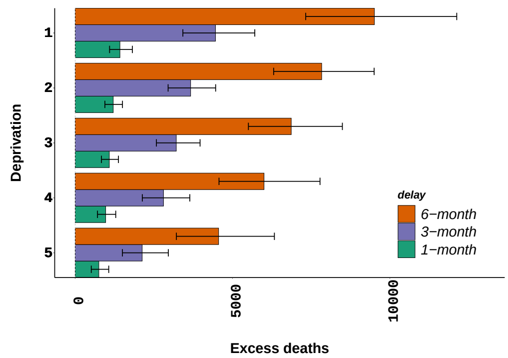
(a) Excess deaths, **women**, region



(b) Excess deaths, **women**, deprivation quintiles

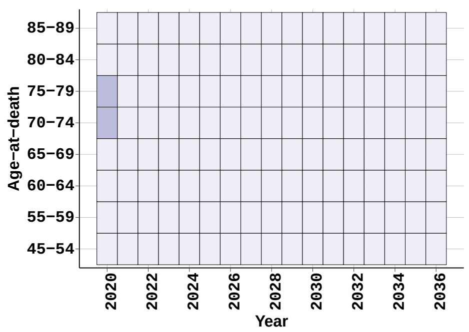


(c) Excess deaths, **men**, region

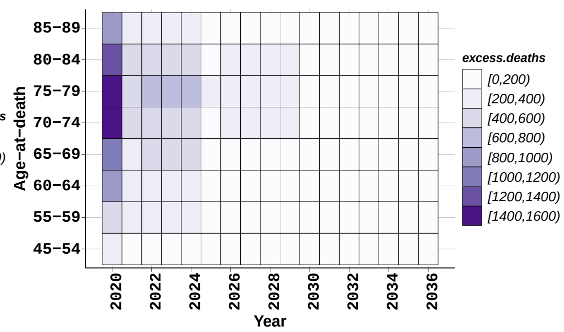


(d) Excess deaths, **men**, deprivation quintiles

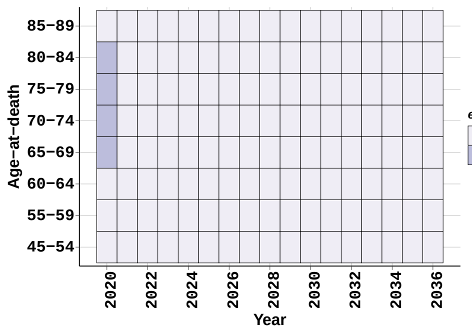
Figure C.41: Cumulative lung cancer excess deaths from 2020 to 2036 based on the full models **without smoking data**. Total excess deaths (over 17 years) in different deprivation quintiles and regions of England, with 95% credible intervals.



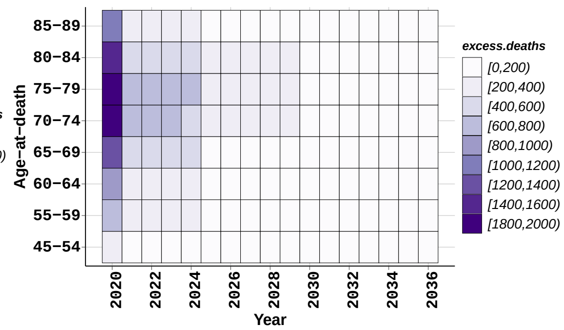
(a) Excess deaths, 1-month delay, **women**



(b) Excess deaths, 6-month delay, **women**

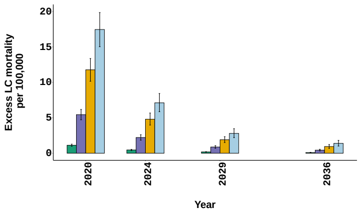


(c) Excess deaths, 1-month delay, **men**

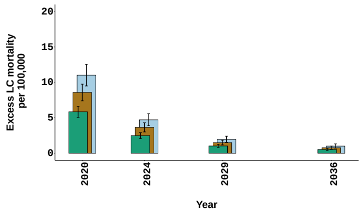


(d) Excess deaths, 6-month delay, **men**

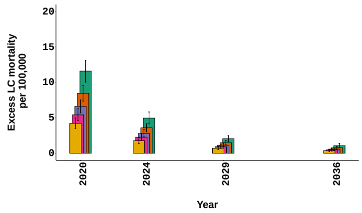
Figure C.42: Lung cancer excess deaths (absolute numbers) by age-at-death in England from 2020 to 2036, based on **full models without smoking data**.



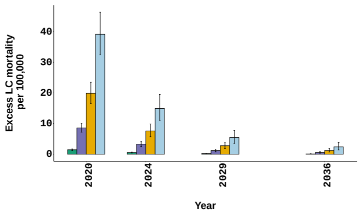
(a) 1-month delay, **women**, age-specific



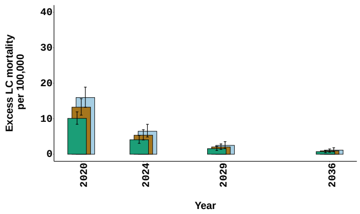
(b) 1-month delay, **women**, region-specific



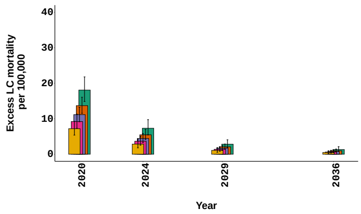
(c) 1-month delay, **women**, deprivation-specific



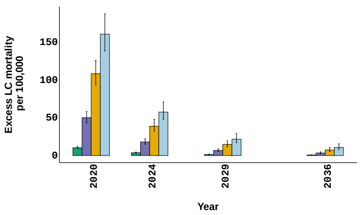
(d) 1-month delay, **men**, age-specific



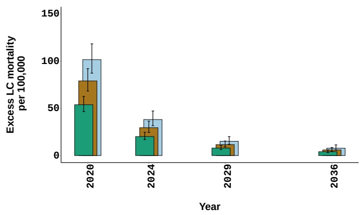
(e) 1-month delay, **men**, region-specific



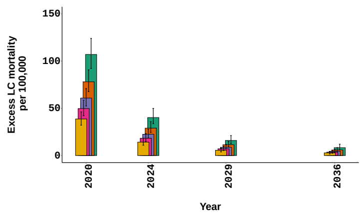
(f) 1-month delay, **men**, deprivation-specific



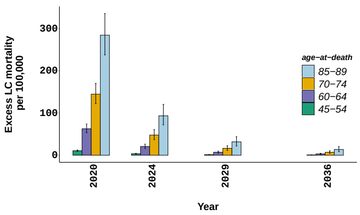
(g) 6-month delay, **women**, age-specific



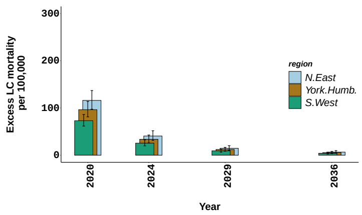
(h) 6-month delay, **women**, region-specific



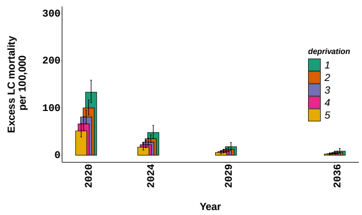
(i) 6-month delay, **women**, deprivation-specific



(j) 6-month delay, **men**, age-specific



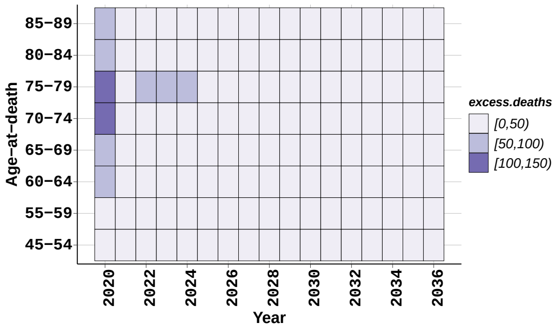
(k) 6-month delay, **men**, region-specific



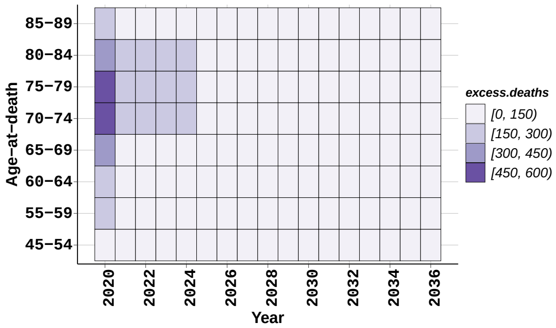
(l) 6-month delay, **men**, deprivation-specific

Figure C.43: Lung cancer excess mortality, **per 100,000 people**, by age-at-death, selected regions and deprivation quintiles in England based on **full models without smoking data**. Annual excess deaths from 2020 to 2036, with 95% credible intervals. Note that differences in lung cancer excess mortality at other ages, in other regions or deprivation quintiles in intermediate years, are comparable to the presented years.

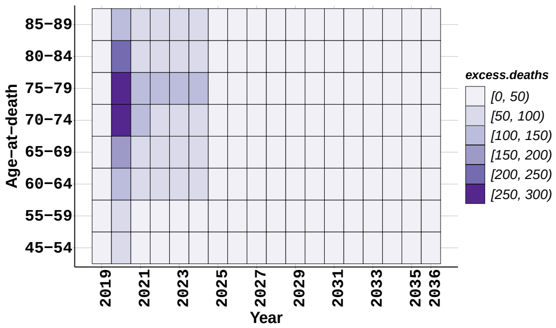
C.6 Excess lung cancer deaths based on the models in (3.5) for women and (3.7) for men, with smoking data



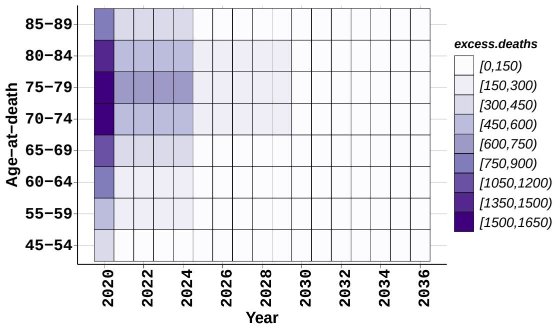
(a) Excess deaths, 1-month delay, **women**



(b) Excess deaths, 6-month delay,**women**



(c) Excess deaths, 1-month delay, **men**



(d) Excess deaths, 6-month delay, **men**

Figure C.44: Lung cancer excess deaths (absolute numbers) by age-at-death in England from 2020 to 2036, based on **full models with smoking data**.

C.7 Findings on breast cancer mortality based on the model in (3.8), without smoking data

C.7.1 Pearson residuals

Figure C.45 exhibits the distribution of Pearson residuals in selected regions of England across modelling ages 35–39 to 85–89 based on the model specification in (3.8) between 2001 and 2018. The figure does not suggest a particular pattern or clustering over ages or years, pointing towards a good fit of the model to the BC data.

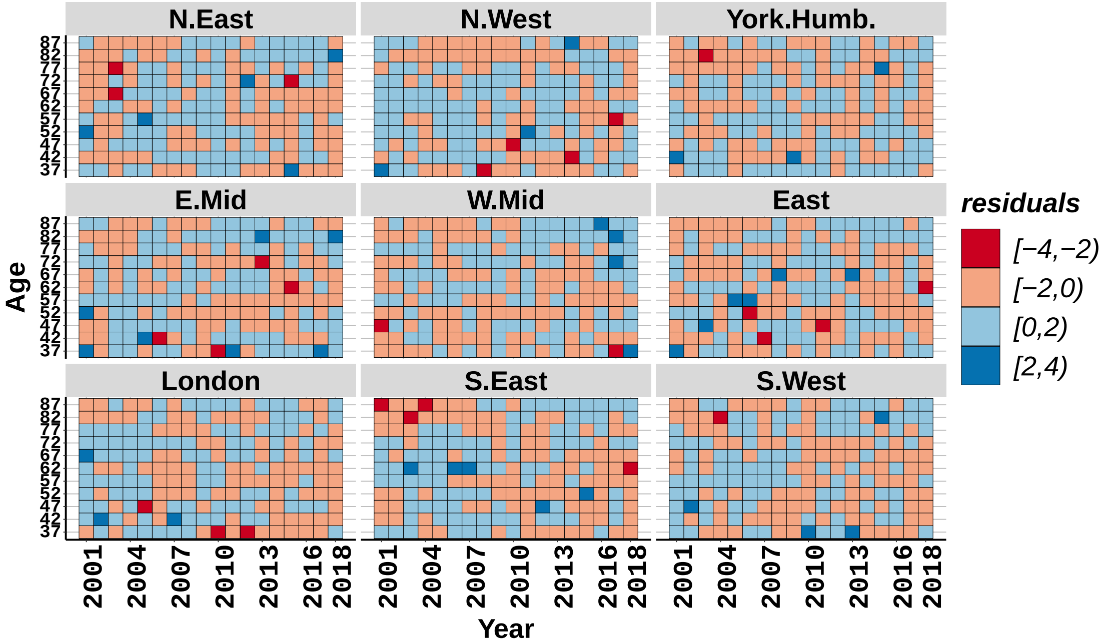


Figure C.45: Heat map of Pearson residuals for breast cancer mortality in regions of England based on (3.8), **excluding smoking data**: orange/light blue cells indicate areas with good fit, while red/dark blue cells indicate areas with poor fit. Note that there is a small number of residuals greater than 4, and these are included in the last category.

C.7.2 Age-specific rates from 2001 to 2036

We present fitted and projected BC mortality in women across various age groups, using the mid-age of each group as a reference, in the regions of England from 2001 to 2036. These projections are based on (3.8), which excludes NS prevalence rates.

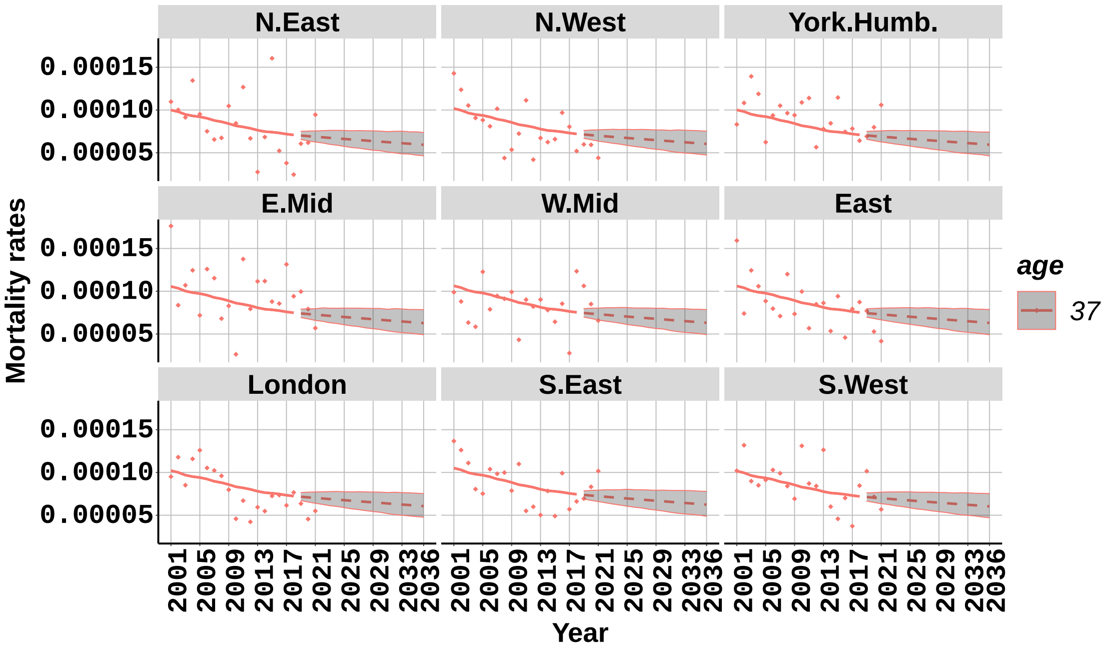


Figure C.46: Breast cancer mortality, females, **age at death 37**, in regions of England based on (3.8), **excluding smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

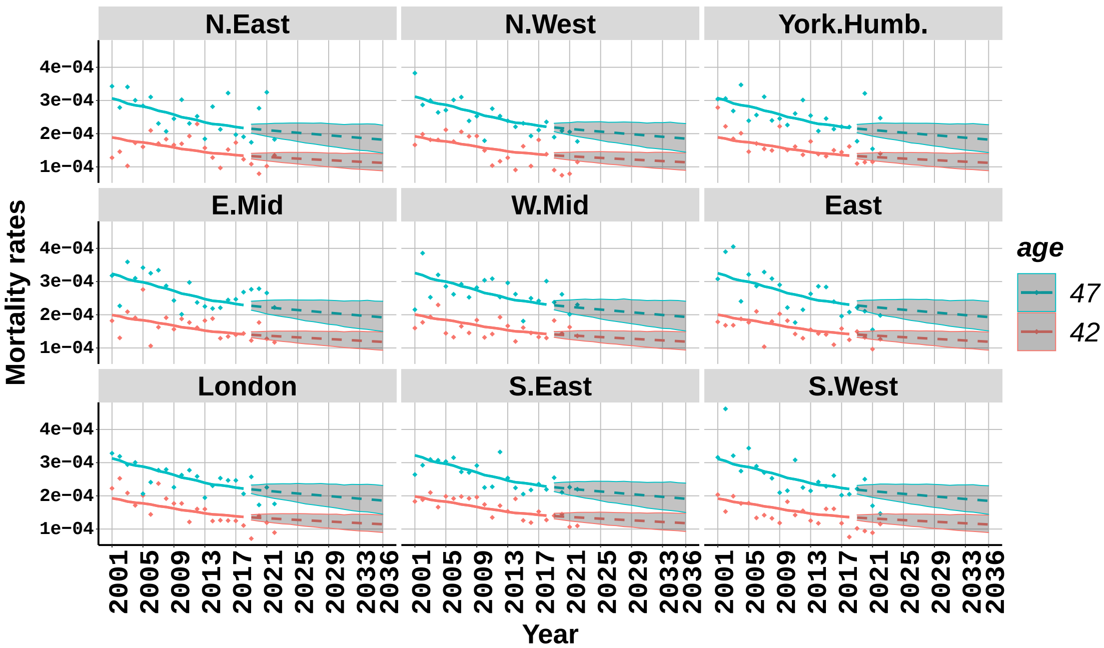


Figure C.47: Breast cancer mortality, females, **ages at death 42 and 47**, in regions of England based on (3.8), **excluding smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

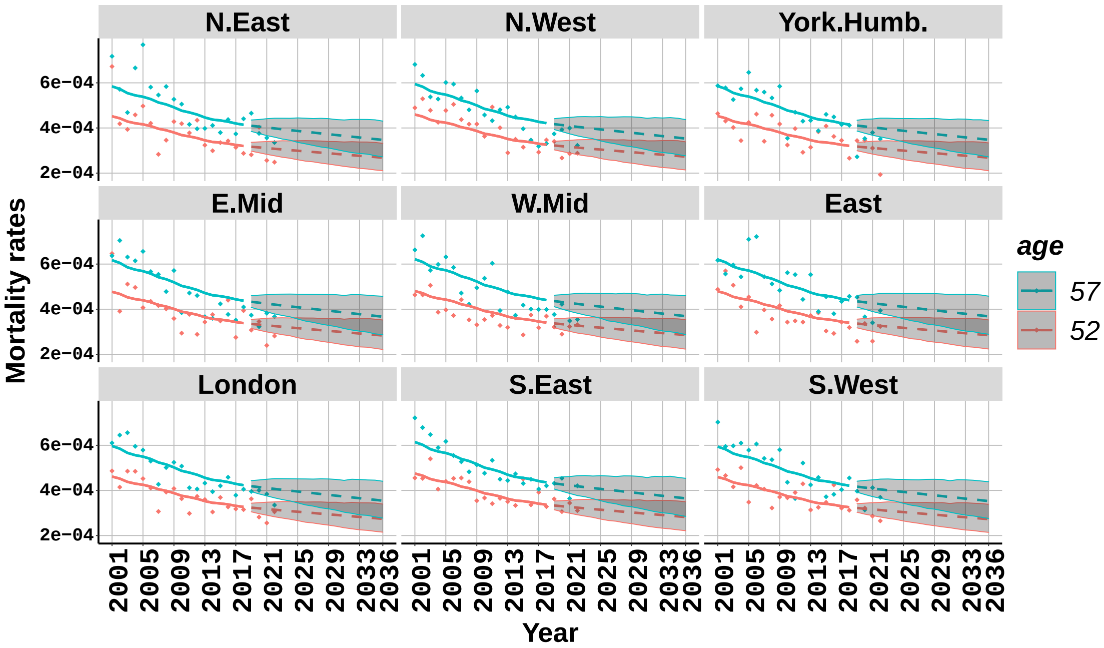


Figure C.48: Breast cancer mortality, females, **ages at death 52 and 57**, in regions of England based on (3.8), **excluding smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

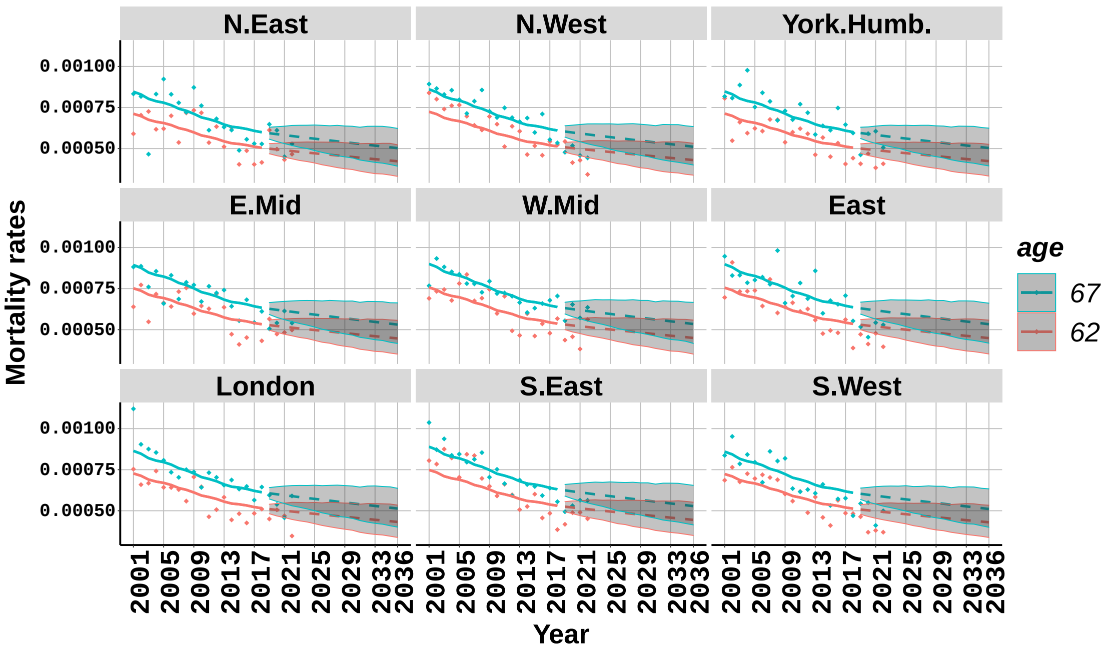


Figure C.49: Breast cancer mortality, females, **ages at death 62 and 67**, in regions of England based on (3.8), **excluding smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

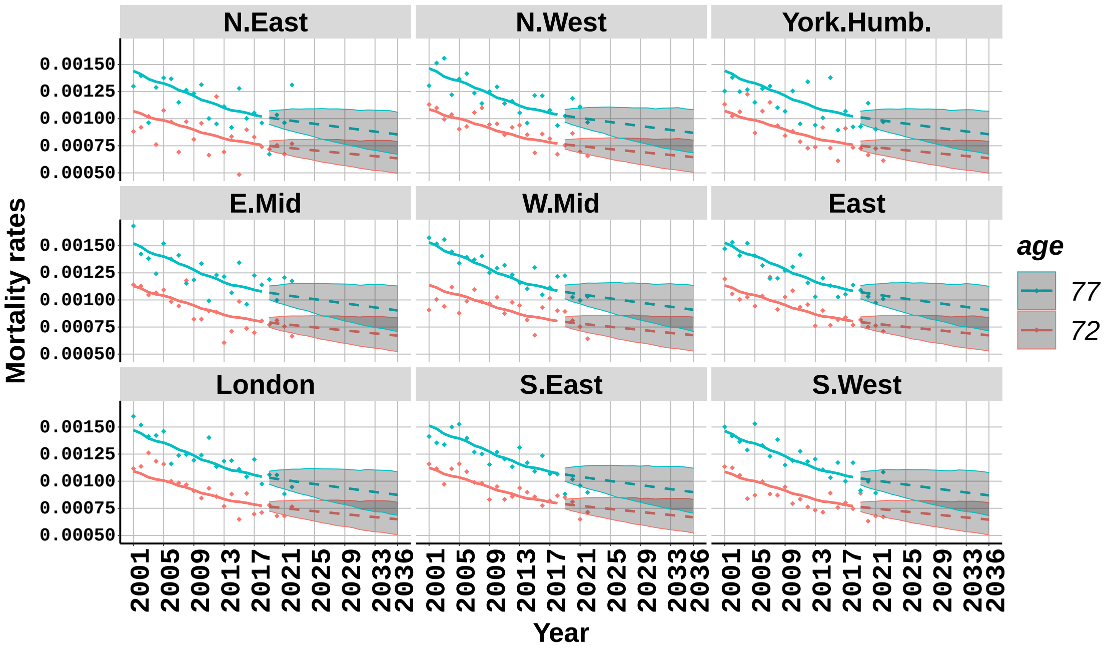


Figure C.50: Breast cancer mortality, females, **ages at death 72 and 77**, in regions of England based on (3.8), **excluding smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

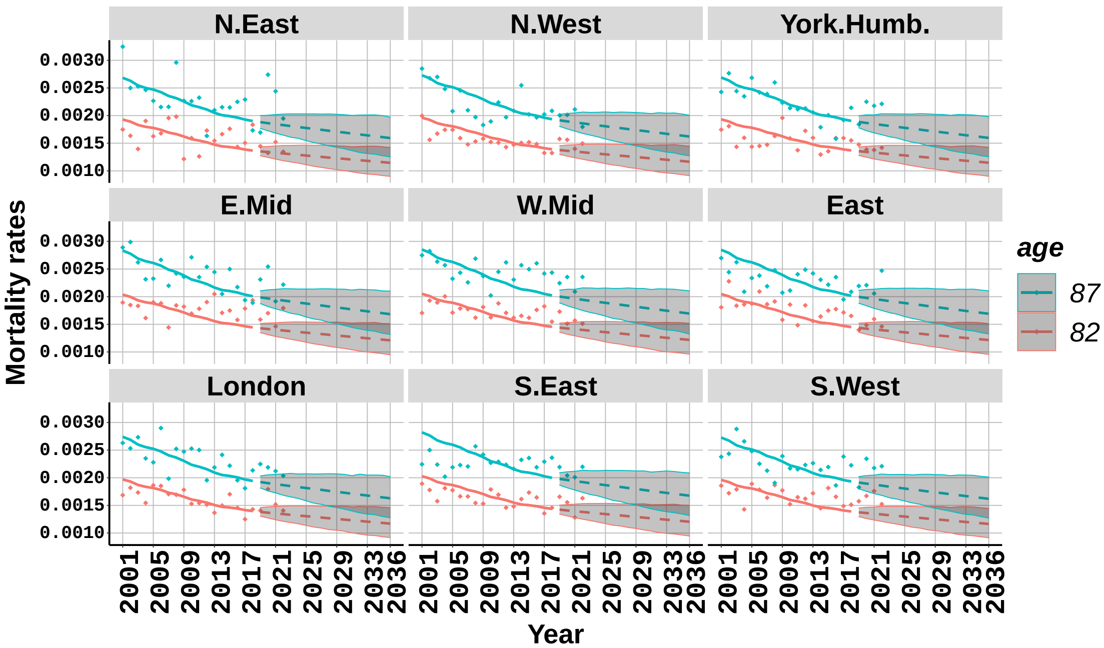


Figure C.51: Breast cancer mortality, females, **ages at death 82 and 87**, in regions of England based on (3.8), **excluding smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

C.8 Findings on breast cancer mortality based on the model in (3.9), with smoking data

C.8.1 Pearson residuals

Figure C.52 displays the distribution of Pearson residuals across selected regions of England for modelling ages 35–39 to 85–89, based on the model specification in (3.9), for the period 2001–2018. The residual patterns in Figure C.52 closely resemble those in Figure C.45. However, the model incorporating NS prevalence rates, as specified in (3.9), demonstrates improved marginal likelihood and DIC results (Appendix A.3).

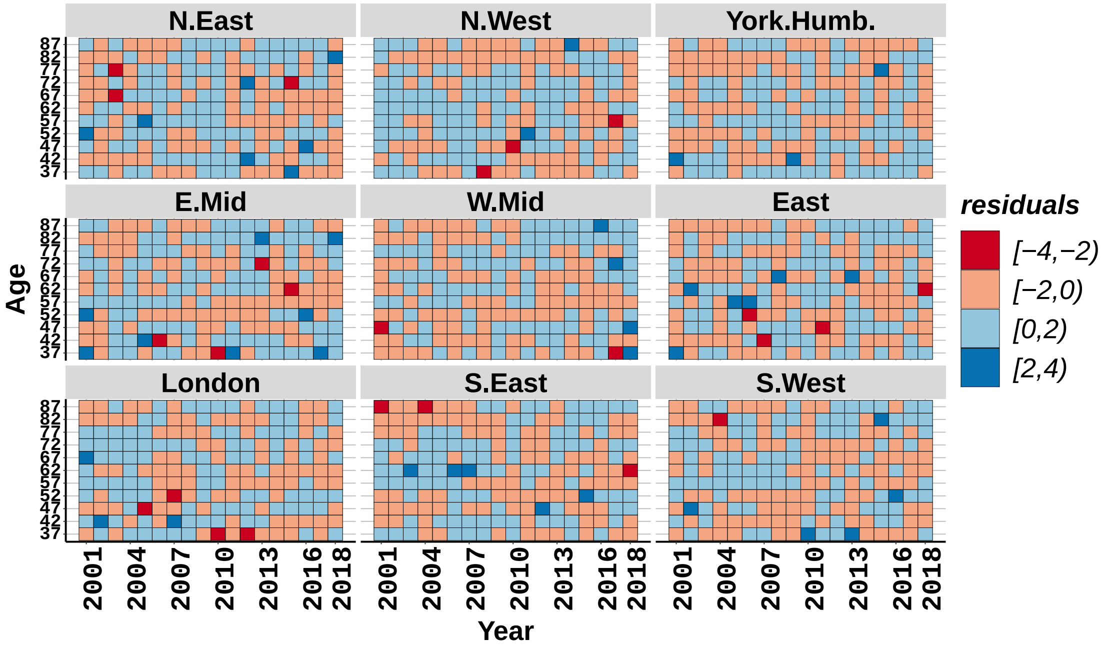


Figure C.52: Heat map of Pearson residuals for breast cancer mortality in regions of England based on (3.9), **including smoking data**: orange/light blue cells indicate areas with good fit, while red/dark blue cells indicate areas with poor fit. Note that there is a small number of residuals greater than 4, and these are included in the last category.

C.8.2 Age-specific rates from 2001 to 2036

We present fitted and projected BC mortality in women across various age groups, using the mid-age of each group as a reference, in the regions of England from 2001 to 2036. These projections are based on (3.9), which includes NS prevalence rates.

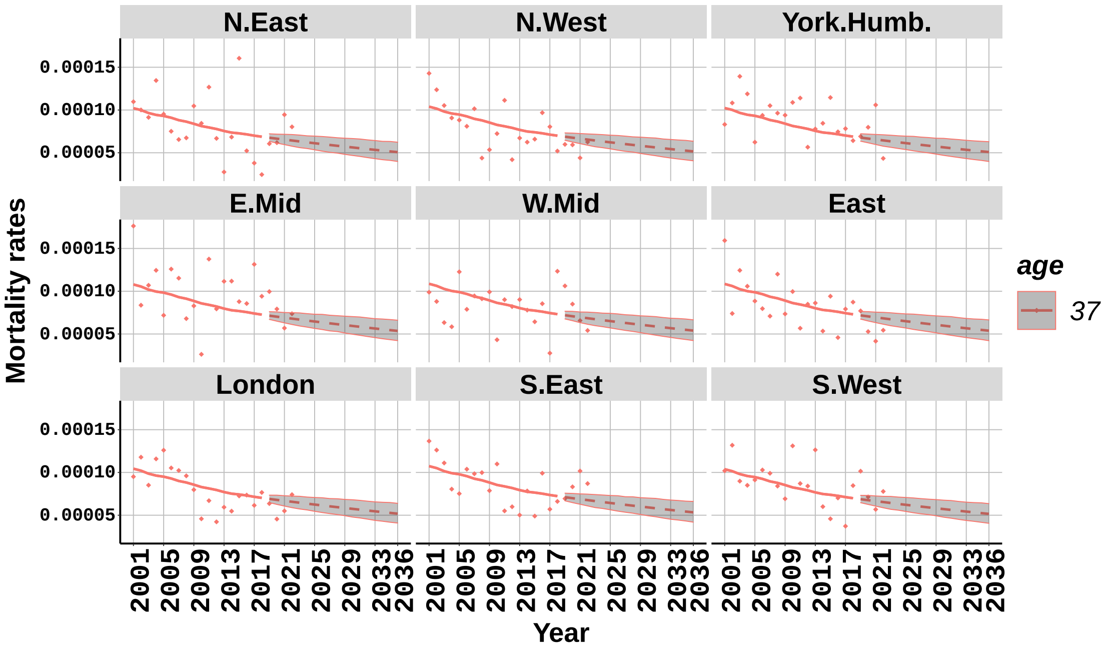


Figure C.53: Breast cancer mortality, females, **age at death 37**, in regions of England based on (3.9), **including smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

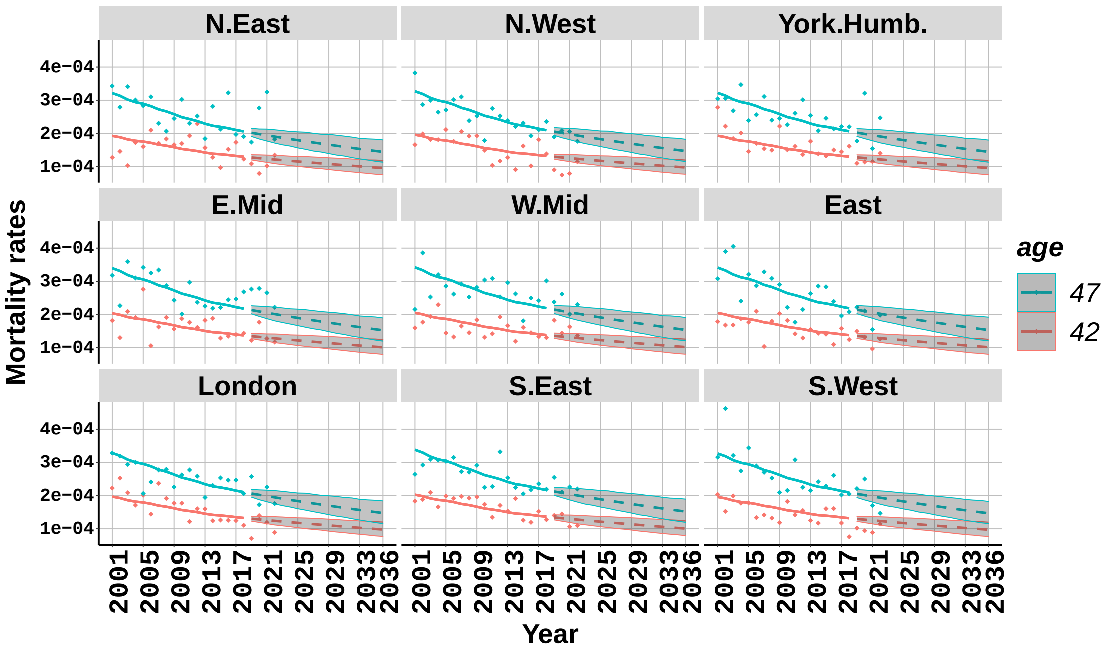


Figure C.54: Breast cancer mortality, females, **ages at death 42 and 47**, in regions of England based on (3.9), **including smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

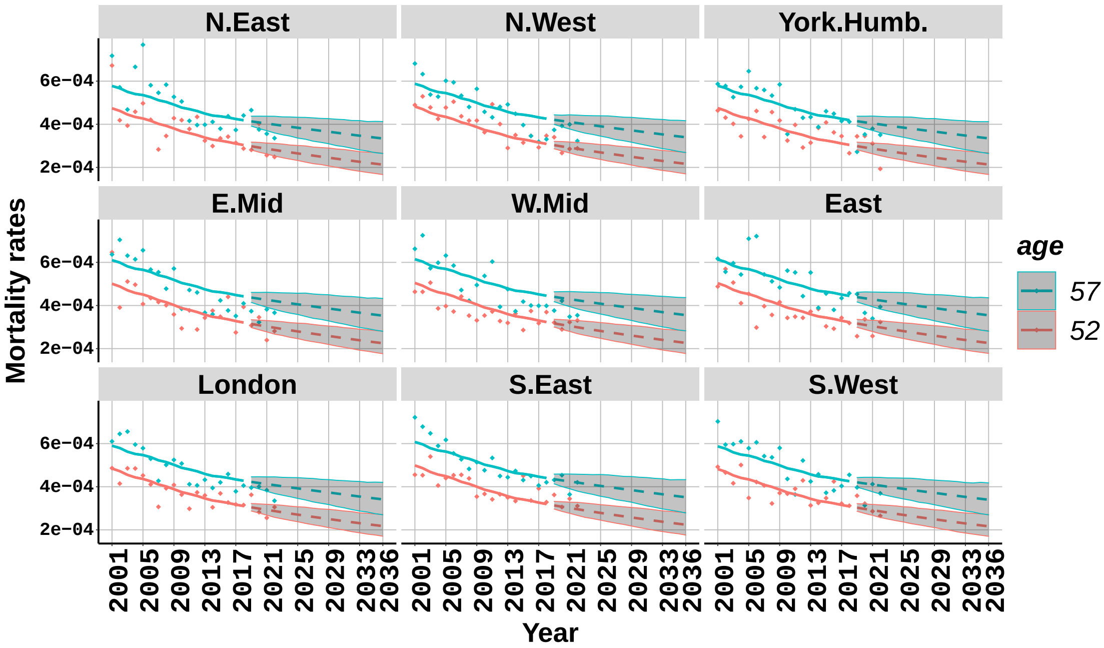


Figure C.55: Breast cancer mortality, females, **ages at death 52 and 57**, in regions of England based on (3.9), **including smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

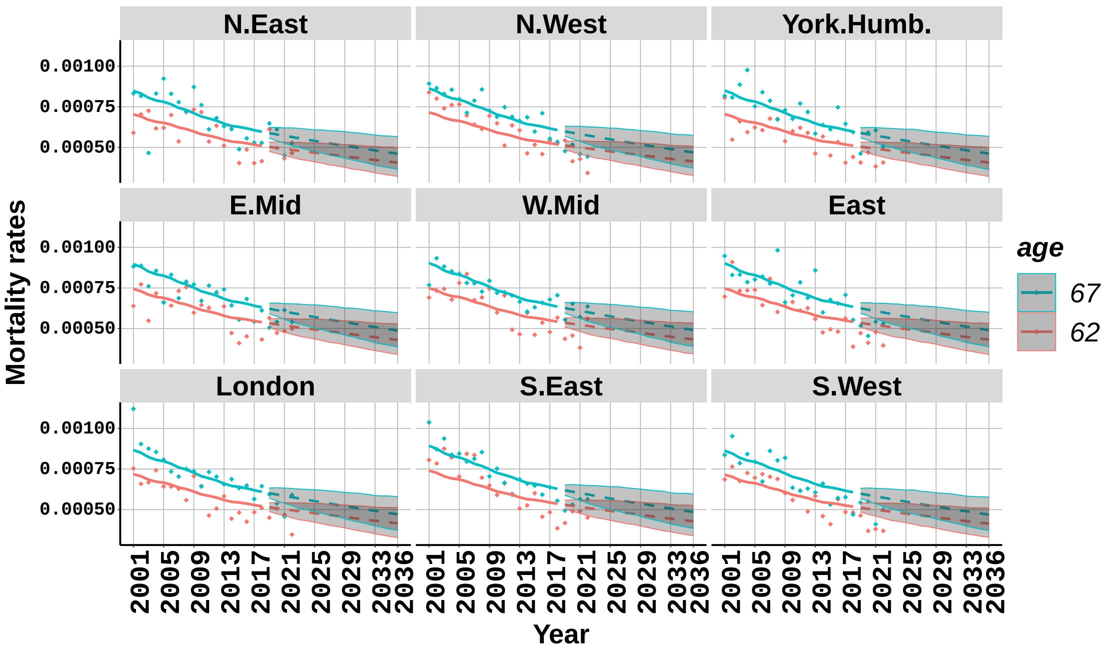


Figure C.56: Breast cancer mortality, females, **ages at death 62 and 67**, in regions of England based on (3.9), **including smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

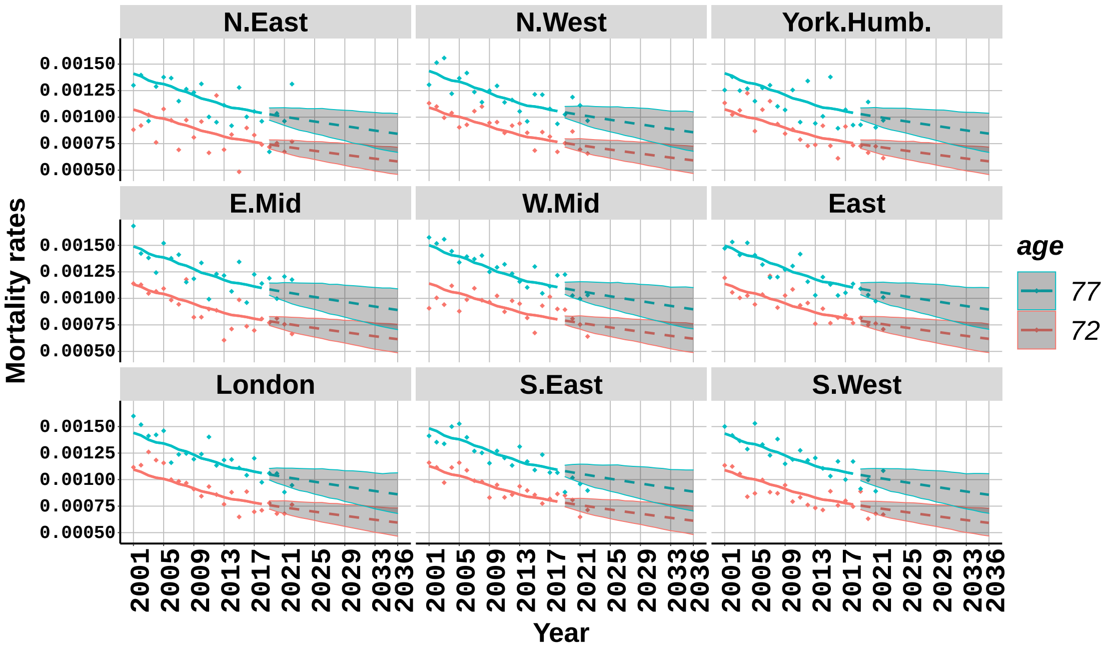


Figure C.57: Breast cancer mortality, females, **ages at death 72 and 77**, in regions of England based on (3.9), **including smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

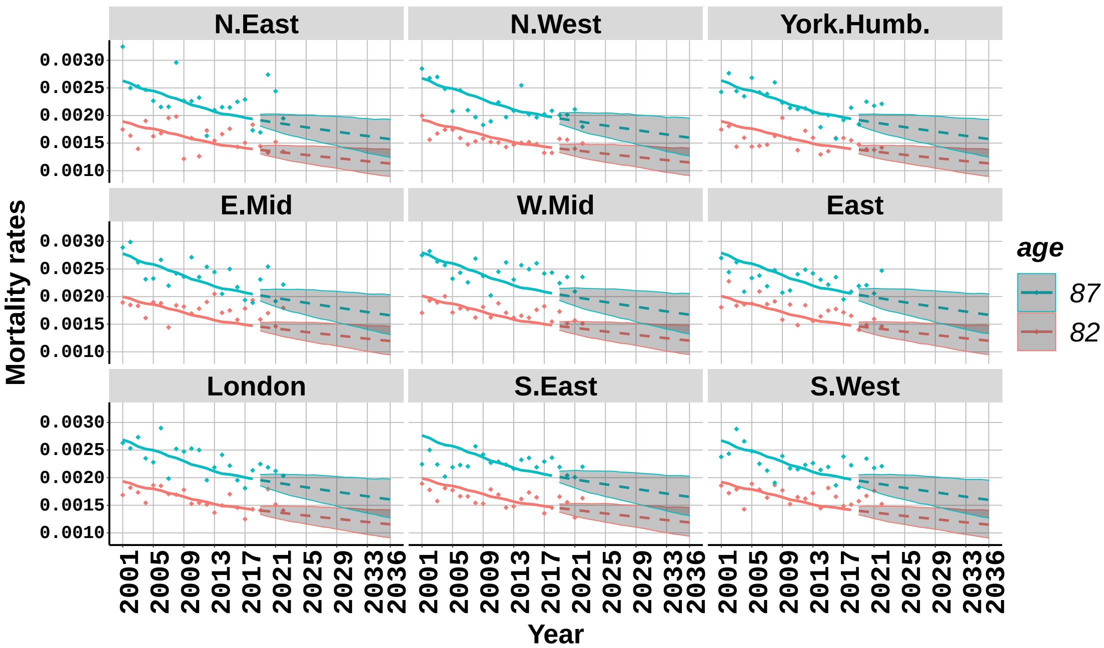


Figure C.58: Breast cancer mortality, females, **ages at death 82 and 87**, in regions of England based on (3.9), **including smoking data**: observed rates (dots), fitted rates (lines), projected rates (dashed lines) with 95% credible intervals for the projected rates.

Bibliography

Al-Delaimy, W., Cho, E., Chen, W., Colditz, G., and Willet, W. (2004). A prospective study of smoking and risk of breast cancer in young adult women. *Cancer Epidemiol Biomarkers Prev.*, 13(3):398–404.

Antonio, K., Bardoutsos, A., and Ouburg, W. (2015). Bayesian poisson log-bilinear models for mortality projections with multiple populations. *European Actuarial Journal*, 5:245–281.

Archbold, M., Davies, B., and Mais, D. (2023). Deprivation and the impact on smoking prevalence, England and Wales: 2017 to 2021. Technical report, Office for National Statistics.

Arık, A., Cairns, A., Dodd, E., Macdonald, A., and Streftaris, G. (2023a). The effect of the covid-19 health disruptions on breast cancer mortality for older women: a semi-markov modelling approach. https://arxiv.org/abs/2303.16573.

Arık, A., Cairns, A., Dodd, E., Macdonald, A., and Streftaris, G. (2023b). Estimating the impact of the COVID-19 pandemic on breast cancer deaths among older women. Technical report, Society of Actuaries.

Arık, A., Dodd, E., Cairns, A., and Streftaris, G. (2021). Socioeconomic disparities in cancer incidence and mortality in England and the impact of age-at-diagnosis on cancer mortality. *PLoS One*, 16(7).

Arık, A., Dodd, E., and Streftaris, G. (2020). Cancer morbidity trends and regional differences in England - a bayesian analysis. *PLoS One*, 15(5).

Baker, C. (2021). Cancer: summary of statistics (England). https://researchbriefings.files.parliament.uk/documents/SN06887/SN06887.pdf.

Bennett, J., Pearson-Stuttard, J., Kontis, V., Capewell, S., Wolfe, I., and Ezzati, M. (2018). Contributions of diseases and injuries to widening life expectancy inequalities in England from 2001 to 2016: a population-based analysis of vital registration data. *LANCET*, 3:e586–97.

Brouhns, N., Denuit, M., and Vermunt, J. (2002). A poisson log-bilinear regression approach to the construction of projected life tables. *Insurance: Mathematics and Economics*, 31:373–393.

Brown, J. (2003). Redistribution and insurance: Mandatory annuitization with mortality heterogenity. *Journal of Risk and Insurance*, 70(1):17–41.

Brown, S., Hole, D., and Cooke, T. (2007). Breast cancer incidence trends in deprived and affluent Scottish women. *Breast Cancer Res Treat*, 103:233–238.

Chen, N., Zhou, M., Dong, X., and et al. (2020). Epidemiological and clinical characteristics of 99 cases of 2019 coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet*, 395(10223):507–13.

Collin, S., Martin, R., Metcalfe, M., Gunnell, D., Albertsen, P., Neal, D., Hamdy, F., Stephens, P., Lane, J., Moore, R., and Donovan, J. (2008). Lung cancer incidence in never-smokers. *Lancet Oncology*, 9:445–52.

CRUK (2020). Cancer in the UK 2020: Socio-economic deprivation. Technical report, Cancer Research UK.

CRUK (2021a). Evidence of the impact of covid-19 across the cancer pathway: Key stats. Technical report, Cancer Research UK.

CRUK (2021b). Survival for all stages of lung cancer. https://www.cancerresearchuk.org/about-cancer/lung-cancer/survival.

CRUK (2022a). 2012 review of the UK breast screening programme. https://www.cancerresearchuk.org/about-cancer/breast-cancer/getting-diagnosed/screening/screening-2012-review.

CRUK (2022b). Performance measures across the cancer pathway: Key stats. Technical report, Cancer Research UK.

Czado, C., Delwarde, A., and Denuit, M. (2005). Bayesian Poisson log-bilinear mortality projections. *Insurance: Mathematics and Economics*, 36:260–284.

DCLG (2015). The english index of multiple deprivation (imd) 2015 - guidance. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/464430/English_Index_of_Multiple_Deprivation_2015_-_Guidance.pdf.

DHSC (2023). New lung cancer screening roll out to detect cancer sooner. https://www.gov.uk/government/news/new-lung-cancer-screening-roll-out-to-detect-cancer-sooner.

Digital, N. (2020). Health survey for England. https://digital.nhs.uk/data-and-information/publications/statistical/health-survey-for-england/2019.

Digital, N. (2023). Deaths from cancer increased with deprivation. https://digital.nhs.uk/data-and-information/publications/statistical/cancer-registration-statistics/england-2020/deaths-from-cancer-increased-with-deprivation.

Eurostat (2007). Regions in the European Union: Nomenclature of territorial units for statistics. *Eurostat:Methodologies and Working Papers*.

Eurostat (2013). Revision of the european standard population - report of eurostat's task force. https://ec.europa.eu/eurostat/documents/3859598/5926869/KS-%20RA-13-028-%20EN.PDF/e713fa79-1add-44e8-b23d-5e8fa09b3f8f.

Gelman, A., Carlin, J., Stern, H., and Rubin, D. (2013). *Bayesian Data Analysis*. Chapman&HALL/CRC.

Gill, B. (2015). The english indices of deprivation 2015. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/465791/English_Indices_of_Deprivation_2015_-_Statistical_Release.pdf.

Grasselli, G., Zangrillo, A., Zanella, A., and et al. (2020). Baseline characteristics and outcomes of 1591 patients infected with SARS-CoV-2 admitted to ICUs of the Lombardy Region, Italy. *JAMA*, 323(16):1574–81.

Hunter, D., Hankinson, S., Hough, H., Gertig, D., and et al. (1997). A prospective study of nat2 acetylation genotype, cigarette smoking, and risk of breast cancer. *Carcinogenesis*, 18(11):2127–32.

Kass, R. and Raftery, A. (1995). Bayes factors. *Journal of the American Statistical Association*, 90:773–795.

Kimball, S. (2002). Product matters. Technical report, Society of Actuaries.

Lai, A., Pasea, L., Banerjee, A., and et al. (2020). Estimated impact of the covid-19 pandemic on cancer services and excess 1-year mortality in people with cancer and multimorbidity: near real-time data on cancer care, cancer deaths and a population-based cohort study. *BMJ Open*.

Lindgren, F. and Rue, H. (2015). Bayesian spatial modelling with r-inla. *Journal of Statistical Software*, 63(19).

Luo, Q., O'Connell, D., Yu, X., Kahn, C., Caruana, M., and et al. (2022). Cancer incidence and mortality in Australia from 2020 to 2044 and an exploratory analysis of the potential effect of treatment delays during the COVID-19 pandemic: a statistical modelling study. *Lancet Public Health*, 7:537–48.

Macdonald, A., Waters, H., and Wekwete, C. (2003). The genetics of breast and ovarian cancer ii: A model of critical illness insurance. *Scandinavian Actuarial Journal*, 1:28–50.

Maringe, C., Spicer, J., Morris, M., Purushotham, A., Nolte, E., and Sullivan, R. e. a. (2020). The impact of the COVID-19 pandemic on cancer deaths due to delays in diagnosis in England, UK: a national, population-based, modelling study. *The LANCET Oncology*, 21(8):1023–1034.

Mouw, T., Koster, A., Wright, M., Blank, M., Moore, S., Hollenbeck, A., and Schatzkin, A. (2008). Education and risk of cancer in a large cohort of men and women in the United States. *PLOS ONE*.

Nash, A. (2020). Subnational population projections for england: 2018-based. Technical report, Office for National Statistics.

NHS (2022). Causes: Lung cancer. https://www.nhs.uk/conditions/lung-cancer/causes/.

Ntzoufras, I. (2009). *Bayesian Modelling using WinBUGS*. WILEY Series in Computational Statistics.

ONS (2015). Cancer registration statistics, England, 2013. Technical report, Office for National Statistics.

ONS (2017). Cancer registration statistics, England, 2015. Technical report, Office for National Statistics.

ONS (2018a). Cancer registration statistics, England, 2016. Technical report, Office for National Statistics.

ONS (2018b). Deaths registered in England and Wales (series DR): 2018. Technical report, Office for National Statistics.

ONS (2020a). Adult smoking habits in the UK: 2019. Technical report, Office for National Statistics.

ONS (2020b). Coronavirus (COVID-19): 2020 in charts. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid192020incharts/2020-12-18.

ONS (2020c). Leading causes of death, UK: 2001 to 2018. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/causesofdeath/articles/leadingcausesofdeathuk/2001to2018.

ONS (2021). Deaths registered in England and Wales: 2020. Technical report, Office for National Statistics.

ONS (2022a). Avoidable mortality in Great Britain: 2020. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/causesofdeath/bulletins/avoidablemortalityinenglandandwales/2020#avoidable-mortality-by-cause.

ONS (2022b). Cancer mortality statistics 2021 to 2022. https://www.ons.gov.uk/aboutus/transparencyandgovernance/freedomofinformationfoi/cancermortalitystatistics2021to2022.

ONS (2022c). Dataset: Smoking habits in the UK and its constituent countries. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandlifeexpectancies/datasets/smokinghabitsintheukanditsconstituentcountries.

ONS (2022d). Nomis - offical census and labour statistics. https://www.nomisweb.co.uk/query/construct/summary.asp?mode=construct&version=0&dataset=161.

ONS (2023). Death registration summary statistics, England and Wales: 2022. Technical report, Office for National Statistics.

PHS (2021). Cancer waiting times in NHS Scotland. Technical report, Public Health Scotland. https://www.publichealthscotland.scot/media/9358/2021-09-28-cwt-report.pdf.

PHS (2022). Breast screening. https://www.nhsinform.scot/healthy-living/screening/breast/breast-screening.

Redondo Loures, C. and Cairns, A. (2019). Mortality in the US by education level. *Annals of Actuarial Science*, 14(2):384–419.

Reynolds, P. (2013). Smoking and breast cancer. *Journal of Mammary Gland Biology and Neoplasia*, 18:15–23.

Reynolds, P., Hurley, S., Goldberg, D., Anton-Culver, H., and et al. (2004). Active smoking, household passive smoking, and breast cancer: evidence from the California Teachers Study. *Journal of the National Cancer Institute*, 96(1):29–37.

Riaz, S., Horton, M., Kang, J., Mak, V., Luchtenborg, M., and Møller, H. (2011). Lung cancer incidence and survival in England: an analysis by socioeconomic deprivation and urbanization. *Journal of Thoracic Oncology*, 6(12).

Richardson, S., Hirsch, J., Narasimhan, M., and et al. (2020). Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York city area. *JAMA*, 323(20):2052–9.

Ryan, H., Trosclair, A., and Gfroerer, J. (2012). Adult current smoking: Differences in definitions and prevalence estimates – NHIS and NSDUH, 2008. *Journal of Environmental and Public Health*, 2012:918368.

Singh, G., Williams, S., Siahpush, M., and Mulhollen, A. (2011). Socioeconomic, rural-urban, and racial inequalities in US cancer mortality: Part i-all cancers and lung cancer and Part ii-colorectal, prostate, breast, and cervical cancers. *Journal of Cancer Epidemiology*, 2011.

Smith, T., Noble, M., Noble, S., Wright, G., McLennan, D., and Plunkett, E. (2015a). The english indices of deprivation 2015 - research report.

Smith, T., Noble, M., Noble, S., Wright, G., McLennan, D., and Plunkett, E. (2015b). The english indices of deprivation 2015 - technical report.

Smittenaar, C., Petersen, K., Stewart, K., and Moitt, N. (2016). Cancer incidence and mortality projections in the UK until 2035. *British Journal of Cancer*, 115:1147–1155.

Spiegelhalter, D., Best, N., Carlin, B., and Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society*, B(64):583–640.

Sud, A., Torr, B., Jones, M., and et al. (2020). Effect of delays in the 2-week-wait cancer referral pathway during the COVID-19 pandemic on cancer survival in the UK: a modelling study. *The LANCET: Oncology*.

Wakelee, H., Chang, E., Gomez, S., Keegan, T., Feskanich, D., Clarke, C., Holmberg, L., Yong, L., Kolonel, L., Gould, M., and West, D. (2007). Lung cancer incidence in never-smokers. *Journal of Clinical Oncology*, 25(5):472–478.

WHO (2022). WHO coronavirus (COVID-19) dashboard. https://covid19.who.int.

Windsor-Shellard, B., Pullin, L., and Horton, M. (2018). Adult smoking habits in the UK:2017. Technical report, Office for National Statistics.

Wong, J., Forster, J., and Smith, P. (2018). Bayesian mortality forecasting with overdispersion. *Insurance: Mathematics and Economics*, 83:206–221.

Xue, F., Willett, W., Rosner, B., and et al. (2011). Cigarette smoking and the incidence of breast cancer. *JAMA Internal Medicine*, 171(2):125–133.

Zhou, F., Yu, T., Du, R., and et al. (2020). Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet*, 395(10229):1054–62.