

Imperial College
London

Spatiotemporal modelling of all-cause and cause-specific mortality in England

Theo Alexander Omar Rashid

Supervisors: Majid Ezzati, James E Bennett, Seth Flaxman

A thesis submitted for the degree of Doctor of Philosophy (PhD)
Department of Epidemiology and Biostatistics, School of Public Health
Imperial College London

September 2023

Copyright

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution Non-Commercial 4.0 International Licence (CC BY-NC).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Declaration of Authorship

I, Theo Alexander Omar Rashid, hereby declare that the work in this thesis is my own original research, and that I have appropriately cited any work within that is not my own.

Abstract

High-resolution data for changes in mortality and longevity are scarce. Estimating mortality for specific combinations of spatial units, time periods, age groups, and causes of death poses statistical and computational challenges which typically result in compromises in the granularity of the results. I applied Bayesian hierarchical models based on patterns of mortality over age, space, and time, to obtain robust yearly estimates of life expectancy and cause-specific mortality for small areas. Using vital registration data held by the UK Small Area Health Statistics Unit, I investigated trends in mortality for subnational units in England from 2002 to 2019.

I examined trends in life expectancy in England's 6791 middle-layer super output areas (MSOAs). In the years 2014-19, 1270 (18.7%) MSOAs for women and 784 (11.5%) MSOAs for men saw declines in life expectancy. The same analysis was performed for the 4835 lower-layer super output areas which comprise London. At this smaller level, issues with the population data in the older age groups affected the reliability of the life expectancy estimates.

I modelled cause-specific mortality for 314 districts in England. The inequality in life expectancy increase since 2010 was driven largely by that of deaths from dementias and the residual group of non-communicable diseases, as well as ischaemic heart disease in men. The analysis was extended to look specifically at the top ten leading cancer causes of death. Preventable cancers showed the greatest spatial inequality in 2019. Unlike areas in the rest of the country, mortality in London from several cancers did not increase in poorer districts, suggesting that some features of London weaken the relationship between poverty and mortality.

England has seen increasing inequalities in all-cause mortality over the past two decades, driven by rises in mortality for a large number of communities and from several preventable causes of death.

Acknowledgements

Firstly, I would like to thank my supervisor Majid Ezzati for his guidance and mentorship. I have learnt so much from you over the past years, and it's inspiring to see the level of care and attention you pay towards your team. I am incredibly grateful for the energy you have put into our work. It's an honour to join the long and distinguished list of your former PhD students.

The stalwart of the group, James Bennett, has been far more than just a supervisor, but also a patient sounding board, a guide to the light when no model will converge, and a great friend. I hope I didn't bother you too much, but to work so closely with you over the past years has been a pleasure. My supervisor, Seth Flaxman, has been a fountain of ideas and energy, and every conversation has stretched my thinking. Robbie Parks and Kyle Foreman have looked out for me throughout the past years and I'd like to thank them for pushing me towards new opportunities whenever they can.

I am part of an incredibly talented and welcoming group of researchers at Imperial. Ricky Nathvani, Sierra Clark, Honor Bixby, Perviz Asaria, Bin Zhou, Jonny Pearson-Stuttard, Vasilis Kontis, and the wider group of NCD-RisC collaborators have provided a constant stream of insights and entertainment. My fellow PhD students, Barbara Metzler, Emily Muller, and Victor Lhoste, have survived the journey with me and made the experience so much fun. Beyond the group, I'd like to thank Adam Howes for helpful, technical conversations and Sam Acors for unhelpful, non-technical conversations. There would be no results in this thesis without the dedicated IT support from Eric Johnson and the invaluable probabilistic modelling advice from the oracle-like Chris Paciorek.

Thank you to my mother, Claire, and to my father, Aly, for my incredible upbringing and the foundation you have laid for me. Thanks to Rosalind as well. Of everyone, my biggest thanks go to Solange for keeping me sane, happy and loved throughout. I am very lucky to have you.

This thesis is dedicated to my grandfather, Geoff Hardern, who sadly passed during my PhD studies. It was his dream for all of his grandchildren to go to university, and I hope he'd be proud of how long I have spent there.

Table of contents

Copyright	3
Declaration of Authorship	5
Abstract	7
Acknowledgements	9
1 Rationale, aims and objectives	19
1.1 Rationale	19
1.2 Objectives	19
1.3 Structure of the thesis	20
2 Background	21
2.1 Overview	21
2.2 Mapping mortality and disease at small areas	21
2.2.1 Disease mapping methods	22
2.2.2 Applications of disease mapping methods	26
2.3 Mortality from specific causes of death	29
2.3.1 The Epidemiologic Transition	29
2.3.2 Modelling cause-specific mortality	31
2.4 Health inequalities in the UK	33
2.5 Summary	37
3 Data	39
3.1 Overview	39
3.2 Geographies of England	39

3.3	Data sources	40
3.3.1	Deaths	40
3.3.2	Population	41
3.3.3	Community deprivation	41
3.3.4	Migration	42
3.4	Exploratory data analysis	42
3.5	Summary	44
4	Modelling death rates	47
4.1	Overview	47
4.2	Why we need to smooth death rates	47
4.3	A model for smoothing death rates	49
4.4	Inference	52
4.5	Clean code and open source	53
4.6	Summary	53
5	Life expectancy trends in England at the MSOA level	55
5.1	Methods	55
5.1.1	Mortality and population data	55
5.1.2	Statistical analysis	56
5.2	Results	57
5.2.1	Inequalities in life expectancy	57
5.2.2	Change in life expectancy	60
5.2.3	Life expectancy and deprivation	64
5.2.4	Inequalities in probability of survival	64
5.3	Discussion	65
5.3.1	Strengths and limitations	66
5.3.2	Comparison with previous literature	68
5.3.3	Analogies with trends in the USA	68
5.4	Summary	70
6	Life expectancy inequality in London at the LSOA level	71
6.1	Methods	72

6.1.1	Mortality and population data	72
6.1.2	Statistical analysis	72
6.2	Results	73
6.2.1	Inequalities in life expectancy	73
6.2.2	Change in life expectancy and inequality	75
6.3	Discussion	76
6.3.1	Strengths and limitations	79
6.3.2	Comparison with previous literature	81
6.3.3	Urban policy and public health implications	82
6.4	Summary	82
7	Cause-specific mortality at the district level	85
7.1	Methods	85
7.1.1	Mortality and population data	85
7.1.2	Grouping causes of death	86
7.1.3	Statistical analysis	88
7.2	Results	90
7.2.1	Variation in space	90
7.2.2	Variation over time	92
7.2.3	Contribution to life expectancy inequality	94
7.2.4	Contribution to life expectancy change	96
7.3	Discussion	98
7.3.1	Strengths and limitations	98
7.3.2	Comparison with previous literature	102
7.3.3	Explaining the variation	103
7.4	Summary	106
8	Trends in cancer mortality at the district level	107
8.1	Methods	108
8.2	Results	110
8.2.1	All cancers	110
8.2.2	Site-specific cancers	110

8.2.3	Change from 2002 to 2019	114
8.3	Discussion	116
8.3.1	Strengths and limitations	116
8.3.2	Comparison with previous work	118
8.3.3	Explaining the variation and implications	119
8.4	Summary	121
9	Discussion	123
9.1	Public health and policy implications of findings	123
9.2	Future work	125
	References	129
	Appendices	145
A	Life table methods	145
A.1	Period life tables	145
A.1.1	The very young ages and the very old ages	147
A.2	Probability of dying	148
A.3	Cause-specific decomposition of differences in life expectancy	148
A.4	Mean age at death	149
B	Model parameters and priors	151
C	Hex-cartogram of life expectancy in 2019 at the MSOA level	157
D	Groups of causes of death	159
E	Maps of cause-specific mortality	163

List of Figures

3.1	Age-specific death rates in England from 2002 to 2019.	43
3.2	Age-specific death rates for broad age groups and life expectancy in England from 2002 to 2019.	43
3.3	District-level life expectancy for total death rates from 2002 to 2019.	44
4.1	Simulated and expected deaths from 2002 to 2019 for a young age group and an old age group.	48
5.1	Ranked MSOA life expectancies in 2002 and 2019.	58
5.2	Map of life expectancy and the distribution of life expectancy in 2019.	59
5.3	Median estimates of MSOA-level life expectancy in England in 2019 against the 95% credible intervals.	60
5.4	Comparison of female and male life expectancy in 2019 and change from 2002 to 2019.	61
5.5	Geography of change in life expectancy from 2002 to 2019.	61
5.6	Map of posterior probability that the estimated change represents an increase in life expectancy from 2002 to 2019.	62
5.7	Distribution of MSOA life expectancies in each year from 2002 to 2019.	63
5.8	Maximum to minimum and 99 th to first percentile differences in life expectancy across 6791 MSOAs, 2002–19.	63
5.9	Change in MSOA life expectancy in different time periods, 2002–19.	64
5.10	MSOA life expectancy in relation to measures of socioeconomic deprivation in the MSOA in 2002 and 2019.	65
5.11	Probability of dying in specific ages in 6791 MSOAs in England in 2002 and 2019.	66
6.1	Map of life expectancy and the distribution of life expectancy in London in 2019.	74

6.2	Median estimates of LSOA-level life expectancy in London in 2019 against the 95% credible intervals.	74
6.3	Geography of change in life expectancy in London from 2002 to 2019.	75
6.4	Change in LSOA life expectancy in different time periods, 2002–19.	76
6.5	Life expectancy at birth and inequality in life expectancy within London's 33 local authority districts from 2002 to 2019.	77
6.6	Distribution of median estimates of LSOA life expectancy at birth for 2002 and 2019, and of the change from 2002 to 2019 in 33 London districts.	78
7.1	Total number of deaths for the twelve leading causes of death in England from 2002 to 2019 and residual groups.	87
7.2	Ranked probability of dying between birth and 80 years of age in 314 local authority districts in England in 2002 and 2019 for the twelve leading causes of death and residual groups.	91
7.3	Ranked change in probability of dying between birth and 80 years of age in 314 local authority districts in England in 2002 and 2019 for the twelve leading causes of death and residual groups.	93
7.4	Contributions of deaths from specific causes to the life expectancy difference in 2019 between the district with the highest life expectancy and all other local authority districts in England.	95
7.5	Contributions of deaths from specific causes to the difference between the district life expectancy in 2002 and 2019 in 314 local authority districts in England. . . .	97
7.6	Contributions of deaths from specific causes to difference between the district life expectancy for women in 2002 and 2010, and in 2010 and 2019, in 314 districts in England.	99
7.7	Contributions of deaths from specific causes to difference between the district life expectancy for men in 2002 and 2010, and in 2010 and 2019, in 314 districts in England.	100
8.1	Total number of deaths for the ten leading cancers in England from 2002 to 2019.	109
8.2	Probability of dying from cancer between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.	111

8.3	Ranked probability of dying between birth and 80 years of age in 314 local authority districts in England in 2002 and 2019 for the ten leading cancers.	112
8.4	Pairwise correlations between the leading cancers for each sex in 2019 women and men.	113
8.5	Local authority district probability of death in 2019 in relation to poverty for each of the leading cancers.	115
8.6	Ranked change in probability of dying between birth and 80 years of age in 314 local authority districts in England in 2002 and 2019 for the ten leading cancers.	117
C.1	Cartogram of life expectancy and the distribution of life expectancy in 2019. . .	157
E.1	Probability of dying from ischaemic heart disease between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.	164
E.2	Probability of dying from a stroke between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.	165
E.3	Probability of dying from all other CVDs between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019. . . .	166
E.4	Probability of dying from Alzheimer's and other dementias between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.	167
E.5	Probability of dying from COPD between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.	168
E.6	Probability of dying from diabetes mellitus, nephritis and nephrosis between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.	169
E.7	Probability of dying for men from liver cirrhosis between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.	170
E.8	Probability of dying from all other NCDs between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019. .	171
E.9	Probability of dying from trachea, bronchus and lung cancers between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.	172

E.10 Probability of dying from colorectal cancer between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019. .	173
E.11 Probability of dying from oesophageal cancer between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.	174
E.12 Probability of dying from pancreatic cancer between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019. .	175
E.13 Probability of dying from stomach cancer between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019. .	176
E.14 Probability of dying from leukaemia between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.	177
E.15 Probability of dying from lymphoma and multiple myeloma between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.	178
E.16 Probability of dying for women from breast cancer between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.	179
E.17 Probability of dying for women from ovarian cancer between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.	179
E.18 Probability of dying for women from corpus uteri cancer between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.	180
E.19 Probability of dying for men from prostate cancer between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.	180
E.20 Probability of dying for men from liver cancer between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.	181
E.21 Probability of dying for men from bladder cancer between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.	181

E.22 Probability of dying from all other cancers (cancers only study) between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.	182
E.23 Probability of dying from all other cancers (cause-specific mortality study) between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.	183
E.24 Probability of dying from lower respiratory infections between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.	184
E.25 Probability of dying from all other infections, maternal, perinatal and nutritional conditions between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.	185
E.26 Probability of dying from injuries between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.	186

List of Tables

3.1	The numbers of each spatial unit of England used in the thesis and their populations in 2019.	39
4.1	Summary of model posterior checks.	51
B.1	Specification of the Bayesian statistical model in Equation 4.4.	151
B.2	Specification of the Bayesian statistical model in Equation 6.1.	152
B.3	Specification of the Bayesian statistical model in Equation 7.1.	154
D.1	Groups of causes of death used in the district-level cause-specific mortality analysis in Chapter 7 with ICD-10 codes.	159
D.2	Groups of causes of death used in the district-level cancer mortality analysis in Chapter 8 with ICD-10 codes.	161

List of Abbreviations

BUGS	B ayesian inference u sing G ibbs sampling
CrI	C redible interval
CAR	C onditional a utoregressive
COPD	C hronic O bstuctive P ulmonary D isease
CVD	C ardiovascular D isease
GBD	G lobal B urden of D isease
GHE	G lobal H ealth E stimates
GMRF	G aussian M arkov R andom F ield
ICAR	I ntrinsic conditional a utoregressive
ICD	I nternational C lassification of D iseases
IMD	I ndex of M ultiple D eprivation
IMPN	I nfections, M aternal, P erinatal and N utritional
LSOA	L ower-layer S uper O utput A rea
MCMC	M arkov chain M onte C arlo
MSOA	M iddle-layer S uper O utput A rea
NCD	N on-communicable D isease
NHS	N ational H ealth S ervice
NUTS	N o U -turn S ampler
OA	O utput A rea
ONS	O ffice for N ational S tatistics
SAHSU	S mall A rea H ealth S tatistics U nit

Chapter 1

Rationale, aims and objectives

1.1 Rationale

Although the UK has, by global standards, a relatively high life expectancy, the country still suffers vast inequalities in mortality. While there have been subnational studies on trends in life expectancy in England, these have been limited in resolution to the district level (Bennett et al., 2015). There is vast heterogeneity in health outcomes and their determinants within districts. As a notable example of mortality from recent years, the district of Kensington and Chelsea, which contains both the most affluent areas in the country in its south and some of the poorest in its north, saw a tragic fire in a highrise residential building in a council housing complex, Grenfell Tower, in 2017. There is a clear need to uncover further variation in mortality below the district level.

Since around 2010, although national life expectancy has continued to increase, there is evidence that the rise in female life expectancy has reversed in the most deprived deciles of the population (Bennett et al., 2018; Marmot et al., 2020). There is a need to identify the specific places where these declines are happening. There are likely to be important changes in mortality from specific causes of death in these areas that have driven this reversal.

1.2 Objectives

The aim of my thesis is to estimate trends in mortality for small areas over the past 20 years in England. There are two specific objectives which will help achieve this aim:

1. To estimate trends in life expectancy for very small areas in England from 2002 to 2019.
2. To investigate trends in cause-specific mortality for districts in England from 2002 to 2019.

1.3 Structure of the thesis

Chapter 2 reviews the literature on spatial methods for mapping disease and mortality for small subnational regions, followed by the literature of separating total mortality into different causes of death. I will then explore inequalities in the UK over the past few decades through to the present. Chapter 3 presents the data sources, and Chapter 4 the statistical modelling choices common to all objectives of this thesis. Chapter 5 concerns the first objective of the thesis - estimating trends in life expectancy for very small areas in England. Chapter 6 extends the first objective by focussing on London at a finer scale than the previous chapter in an attempt to gauge whether higher-resolution analyses are possible. Chapter 7 addresses objective two of this thesis, modelling specific causes of deaths at a coarser scale, and looking at potential drivers of the observed trends in life expectancy. Chapter 8 follows the methods of Chapter 7, but focussing solely on deaths from cancers. Chapter 9 concludes with a discussion on the public health implications of the findings and areas for future research building on the work in this thesis.

Chapter 2

Background

2.1 Overview

The section begins with an overview of the spatial methods for mapping mortality for small areas and describes some examples of applications from the UK and beyond looking at mortality for small subnational regions. There is a brief history of the Small Area Health Statistics Unit (SAHSU), who manage the mortality data used in the thesis and have developed and applied many of the spatial methods discussed.

This is followed by a history of separating total mortality into different causes of death and the epidemiologic transition theory.

The chapter finishes by exploring the picture of inequalities in the UK over the past few decades through to the present, focussing on class, income, geography, and deprivation.

2.2 Mapping mortality and disease at small areas

Many studies compare the prevalence of diseases or mortality in different subgroups of the population by dividing the population geographically into small areas. The number of cases, or number of deaths, in an area is likely to be small. This sparseness issue is even more pertinent when the population is further stratified by age group. When calculating rates of incidence from the observed data, there is an apparent variability between spatial units, which is often larger than the true differences in risk due to the noise in the data. To overcome these issues, we can

use statistical smoothing techniques to obtain robust estimates of rates by sharing information between strata.

2.2.1 Disease mapping methods

In small-area studies, it is common to smooth data using models with explicit spatial dependence, which are designed to give more weight to nearby areas than those further away. There are three main categories for modelling spatial effects. First, we can treat space as a continuous surface using Gaussian processes or splines. Second, we can use hierarchical models for areal data, which make use of the spatial neighbourhood structure of the units. Third, we can again use hierarchical models for areal data but instead we can exploit a nested hierarchy of geographical units, for example between state, county and census tract in the US. Each of these methods, which can be used separately or in combination if the context of the problem allows, rely on assumptions which may make them more or less appropriate in different applications.

Space as a continuous process

In the context of disease mapping, events are usually aggregated to areas rather than assigned specific geographical coordinates. Wakefield and Elliott (1999) model aggregated counts as realisations of a Poisson process, in which the expected number of cases is calculated by integrating a continuous surface that generates the cases over the area of the spatial unit. The surface was a function of spatially-referenced covariates. Kelsall and Wakefield (2002) describe an alternative model, where the log-transformed risk surface is modelled by a Gaussian process, whose correlation (or “kernel”) function depends on distance.

Best et al. (2005) provide a review of the use of hierarchical models with spatial dependence for disease mapping. In particular, the authors focus on Bayesian estimation, and different classes of spatial prior distributions.

The first prior proposed for spatial effects $\mathbf{S} = S_1, \dots, S_n$ is the multivariate normal

$$\mathbf{S} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}), \quad (2.1)$$

where $\boldsymbol{\mu}$ is the mean effect vector, $\boldsymbol{\Sigma} = \sigma^2 \boldsymbol{\Omega}$ and $\boldsymbol{\Omega}$ is a symmetric, positive semi-definite matrix defining the correlation between spatial units. A common choice when specifying the structure of the correlation matrix is to assume a kernel function that decays with the distance between the centroids of the areas, so that places nearby in space share similar disease profiles. Note, this is mathematically equivalent to the practical implementation of a Gaussian process for a specified set of spatial locations, which uses a finite set of points. An example in Elliott et al. (2001b) chooses an exponential decay function to map cancer risk in northwest England. Kernel functions based on distance, however, do not allow the variability of the spatial surface to change with location. Paciorek and Schervish (2006) describe nonstationary extensions to common kernel choices in spatial settings with this property.

Space as discrete units

A more popular prior is the conditional autoregressive (CAR) prior, also known as a Gaussian Markov random field (GMRF), which was first introduced by Besag et al. (1991). These form a joint distribution as in Equation 2.1, but the covariance is usually defined instead in terms of the precision matrix

$$\mathbf{P} = \boldsymbol{\Sigma}^{-1} = \tau(\mathbf{D} - \rho \mathbf{A}), \quad (2.2)$$

where τ controls the overall precision of the effects, $\mathbf{A}$ is the spatial adjacency matrix formed by the small areas, $\mathbf{D}$ is a diagonal matrix with entries equal to the number of neighbours for each spatial unit, and the autocorrelation parameter ρ describes the amount of correlation. This can be seen as tuning the degree of spatial dependence, where $\rho = 0$ implies independence between areas, and $\rho = 1$ full dependence. The case with $\rho = 1$ is called the intrinsic conditional autoregressive (ICAR) model. There sometimes exists further overdispersion in the residuals that cannot be modelled by purely spatially-structured random effects. Besag et al. (1991) proposed the model (hereafter called BYM)

$$S_i = U_i + V_i, \quad (2.3)$$

where U_i follow an ICAR distribution, and V_i are independent and identically distributed random effects. The addition of the spatially-unstructured component V accounts for any non-spatial heterogeneity.

Space as a nested hierarchy of geographies

The relationships between different levels of a hierarchy of geographical units are often incorporated into models as a nested hierarchy of random effects. These models account for when spatial units lie within common administrative boundaries. This is often a desirable property of the model for certain geographies, like states in the US, which are administrative. Policy is decided at these geographies, so there is reason to believe these boundaries may have a greater effect on health outcomes than spatial structure. Finucane et al. (2014) demonstrate how country-level blood pressure can be modelled by exploiting the hierarchy of global, super-region, region and country. Note, although these models group by geographical region, these models are not *spatial* as they do not contain any information on the relative position of the areas.

Of the two specifications that are spatial, either as a continuous process or discrete units, the Markov random field priors are often preferred for computational reasons, as we can exploit the sparseness of the adjacency matrix in our inference algorithms rather than computing the covariance between each pair of spatial units as in the general case of Equation 2.1. There are concerns, however, that the GMRF representation of space as an adjacency matrix, which was originally proposed for a regular lattice of pixels in image analysis (Besag et al., 1991), is reductive for more complicated spatial problems. Despite this, in an epidemiological context, Duncan et al. (2017) found the standard ICAR model with binary, first-order neighbour weights outperformed models with a variety of different weighting schemes, including matrix weights based on higher-order degrees of neighbours, distance between neighbours, and distance between covariate values.

In applications to disease mapping, spatial models are the natural choice when the disease exhibits a spatial pattern. This is the case for mortality from infectious diseases, particularly on short timescales like Covid-19 (Konstantinoudis et al., 2022). Nested hierarchies are a more suitable choice when administrative areas are meaningful and have an effect on the health outcomes of the population. For example, state-specific abortion laws in the USA could affect maternal mortality, and so a model should include an effect for each state.

Modelling variation beyond space

As computational power has improved, it has become feasible to model patterns over other features of the population, such as time period and age group. Trends over time can be modelled as linear through slopes, or using nonlinear effects which allow neighbouring time points to be alike, the simplest of which is a first-order Gaussian random walk process. All-cause mortality varies smoothly over ages, following a characteristic J-shape with higher mortality in the infant and older age groups (Preston et al., 2001), and therefore can be modelled using a nonlinear process such as a random walk.

Difficulties arise when considering interactions between the space, age, and time variables. One can imagine situations in which different spatial units will have different age patterns in disease rates. For example, if certain age groups were vaccinated against disease in that spatial unit before others. There are also social or behavioural risk factors, such as alcohol consumption or smoking rates, that are likely to exhibit different age patterns over space. After implementing a base model with the main effects, the question is how to model additional terms which account for the interactions between the variables. Space-time interactions could range from fully independent, to each spatial unit having independent temporal patterns, to inseparable space-time variation where interactions borrow strength across neighbouring spatial units and neighbouring time periods (Knorr-Held, 2000).

However, it should be considered that by breaking the population down into smaller and smaller subgroups through space, age and time period, the counts of cases become more sparse and there is a need for stronger smoothing to produce robust estimates, particularly for data that are already at the small-area level. Although interaction effects are plausible, modellers should consider whether there is evidence for the interaction in the data or whether they can simplify the model if the interaction effect turns out to be negligible.

It should be noted that there are situations where statistical smoothing would not be appropriate. There might be true variability in the data which a smoothing model would conceal. For example, certain spatial units might contain isolated populations with high mortality over a sustained period, such as counties with Native American reservations in the USA (Dwyer-Lindgren et al., 2017a). There can also be spatially- and temporally-specific events that cause a spike in mortality such as the Grenfell Tower fire in 2017. Without accounting for these events, the models described above

would either attenuate their effect on mortality, or a spike in deaths would cause estimates of mortality in nearby spatial units or years to be erroneously high. Beyond the use of subject matter experts, posterior predictive checks and plots of modelled death rates against the observed data can help to identify outlier spikes in mortality which are specific to a particular time or place, and which we do not want our model to smooth.

2.2.2 Applications of disease mapping methods

Small-area analyses of mortality

In order to compare the health status between areas, health authorities require a measure of mortality that collapses age-specific information into a single number. Indirectly standardised measures such as the standardised mortality ratio – the ratio between total deaths and expected deaths in an area – are easy to calculate, but are not easily understood by laypeople. Directly standardised methods, in contrast, require knowledge of the full age structure of death rates rather than just the total number of deaths. Age-standardised death rates, however, suffer the same interpretability issue as the standardised mortality ratio, and are only comparable between studies if the same reference population is used. An alternative choice is *life expectancy*. Silcocks et al. (2001) explain that life expectancy is a “more intuitive and immediate measure of the mortality experience of a population, [and] is likely to have greater impact... than other measures that are incomprehensible to most people.” However, although the metric appears more interpretable, life expectancy at birth constructed from a period life table is often misinterpreted as the mean length of life of the cohort into which the newborn is born. In fact, it measures the expectation of life assuming that the newborn will be exposed to age-specific mortality conditions throughout their life that are exactly the same as the current population.

The estimation of death rates requires two data sources: deaths counts and populations. Modern death registration systems, such as that of the UK, are almost entirely complete and accurate. On the other hand, although usually treated as a known quantity, the population denominator is often problematic. Populations for small geographies are only recorded during a decennial census, and estimates are generated for the years in-between using limited survey data on births, deaths and migration. And although the census is considered the “gold standard”, it is subject to enumeration errors, particularly for areas with special populations such as students or armed forces (Elliott et al., 2001b).

Beyond the population issue, finer scale studies are restricted by data availability. Where data are available, there is still the need to overcome small number issues before feeding death rates through the life table to calculate life expectancy. Eayres and Williams (2004) recommend a minimum population size of 5000 when using traditional life table methods, below which the calculation of life expectancy is unstable¹, or the error estimates become so large that any comparison between subgroups becomes meaningless. One approach, often taken by statistical agencies, is to build larger populations by either aggregating multiple years of data (Bahk et al., 2020; Office for National Statistics, 2015; Public Health England, 2021a) or combining spatial units (Ezzati et al., 2008). Here, we focus on studies using Bayesian hierarchical models to generate robust estimates of age-specific death rates by recognising the correlations between spatial units and age groups, which produce more accurate estimates for small population studies of life expectancy (Congdon, 2009; Jonker et al., 2012).

Jonker et al. (2012) demonstrated the advantages of the Bayesian approach for 89 small areas in Rotterdam using a joint model for sex, space and age effects, finding a 8.2 year and 9.2 year gap between the neighbourhoods with the highest and lowest life expectancies for women and men. Stephens et al. (2013) employed the same model for 153 administrative areas in New South Wales, Australia.

Bayesian spatial models for mortality have been scaled to small areas for entire countries, and also consider trends in these regions over time. Bennett et al. (2015) forecasted life expectancy for 375 districts in England and Wales using a spatiotemporal model trained over a 31 year period, and Dwyer-Lindgren et al. (2017a) explored mortality trends in 3110 US counties from 1980 to 2014.

There have also been studies on specific cities at a finer resolution. In order to improve estimates for disability-free life expectancy, Congdon (2014) considered both ill-health and mortality in a joint likelihood with spatial effects for 625 wards in London, finding more than a two-fold variation in the percent of life spent in disability for men. Bilal et al. (2019) looked at 266 subcity units for six large cities in Latin America. As there is no contiguous boundary in this case, a random

¹Or, if the open-ended age group contains zero deaths, the calculation of life expectancy is impossible. This is because the probability of dying in the final age group will be zero, so the life table cannot be closed and the life expectancy will be infinite.

effects model for each city was used instead of a spatial model. The largest difference between the top and bottom decile of life expectancy at birth was 17.7 years for women in Santiago, Chile.

Two studies in North America have looked below the county level, at census tracts, with wide-ranging population sizes as small as 40. Dwyer-Lindgren et al. (2017b), using a model that relied heavily on sociodemographic covariates, studied trends for life expectancy and many causes of death for 397 tracts in King County, Washington, uncovering an 18.3 year gap in life expectancy for men. Using the same model for Vancouver, Canada, Yu et al. (2021) found widening inequalities over time and a difference of 9.5 years for men.

Small Area Health Statistics Unit

In 1983, a documentary on the radioactive fallout from a fire at the Sellafield nuclear site in Cumbria claimed that there was a ten-fold increase in cases of childhood leukaemia in the surrounding community. This anomaly had gone undetected by public health authorities, raising concern that routinely collected data were not able to identify local clusters of disease. The subsequent enquiry confirmed the excess, and recommended that a research unit was set up to monitor small-area statistics and respond quickly to *ad hoc* queries on local health hazards. The Small Area Health Statistics Unit (SAHSU) was established in 1987 (Elliott et al., 1992).

Beyond producing substantive research on environment and health, a core aim of SAHSU is to develop small-area statistical methodology (Wakefield and Elliott, 1999) for:

- *Point source type studies.* Is there an increased risk close to an environmental hazard? SAHSU has investigated increased mortality from mesothelioma and asbestosis near Plymouth docks (Elliott et al., 1992); excess respiratory disease mortality near two factories in Barking and Havering (Aylin et al., 1999); kidney disease mortality near chemical plants in Runcorn (Hodgson et al., 2004); and possible excess of several morbidities near landfill sites (Elliott et al., 2001a; Jarup et al., 2007, 2002b).
- *Geographic correlation studies.* Is there a correlation between disease risk and spatially-varying environmental variables? SAHSU have looked at several exposures, including a plume of mercury pollution (Hodgson et al., 2007); mobile phone base stations during pregnancy (Elliott et al., 2010); noise from aircraft near Heathrow (Hansell et al., 2013);

road traffic noise in London (Halonen et al., 2015); and particulate matter from incinerators during pregnancy (Parkes et al., 2020).

- *Clustering*. Does a disease produce non-random spatial patterns of incidence? If the aetiology is unknown, this could suggest that the disease is infectious.
- *Disease mapping*. Summarising the spatial variation in risk.

SAHSU has been at the forefront of both methodology and applications in disease mapping. Aylin et al. (1999) mapped diseases for wards in Kensington, Chelsea and Westminster using a model that smoothed rates towards the mean risk across the region. Thereafter, SAHSU published a plethora of studies for disease mapping models with explicit spatial dependence, including using the BYM model (Equation 2.3) to map spatial variation in the relative risk of testicular (Toledano et al., 2001) and prostate (Jarup et al., 2002a) cancers for small areas in regions of England. In a landmark piece bringing together work on disease mapping and environmental exposures, SAHSU published an environment and health atlas for England and Wales, showing the spatial patterns of 14 health conditions at census ward level over an aggregated 25 year period alongside five environmental exposure surfaces (Hansell et al., 2014).

Further disease mapping studies at SAHSU using spatially structured effects have also extended the methodology to look at age patterns and trends over time. Asaria et al. (2012) analysed cardiovascular disease death rates by fitting a spatial model for all wards in England separately for each age group and time period. Bennett et al. (2015) designed a model to jointly forecast all-cause mortality for districts in England by age group and year. The model used BYM spatial effects and random walk effects over age and time to capture nonlinear relationships.

2.3 Mortality from specific causes of death

2.3.1 The Epidemiologic Transition

In the mid-twentieth century, a team in the US Public Health Service, led by Iwao Moriyama, began investigating the cause-specific composition of mortality into all diseases and injuries for the first half of the century. Moriyama and Gover (1948) grouped vital registration data into primary causes. Notably, they found, as the US saw an overall downward trend in mortality, the leading causes of death changed from communicable diseases, such as tuberculosis and diphtheria, toward non-communicable, “chronic diseases of older ages”, such as heart diseases and cancers.

The success of the reduction – and in the case of typhoid fever, near-elimination – of infectious diseases was attributed to the strategy of the health officer in the early 1900s, who was focussed on improving water and sanitation, and public health interventions such as immunisation and quarantines.

By comparing vital registration data over several centuries, Abdel Omran observed this shift of mortality from communicable to non-communicable diseases (NCDs) in many countries (Omran, 1977, 1971). Although the pace and determinants of the transition varied between countries, Omran was able to formalise three common successive stages of the shift in mortality:

1. *The Age of Pestilence and Famine*. Mortality is high and largely governed by Malthusian “positive checks” – epidemics, famines, and wars.
2. *The Age of Receding Pandemics*. Mortality decreases as epidemics become less frequent, but infectious diseases remain the leading causes of death.
3. *The Age of Degenerative and Man-made diseases*. Mortality declines further along with fertility, increasing the average age of population and NCDs take over as the leading causes of death.

He termed this the *Epidemiologic Transition theory*. Omran (1971) explained that England and Wales took the classic transition path followed by western societies, whereby socioeconomic factors such as improvements to living standards are crucial in causing easily preventable diseases to subside and shifting towards the third phase of the transition, whilst medical and other public health technology only help society much later in the final stage. Later, Olshansky and Ault (1986) would propose a fourth stage to the theory, *the Age of Delayed Degenerative Diseases*, in which the structure of causes of death is stable, but the age at which degenerative diseases kill is postponed, thus decreasing older age mortality. There are, however, questions around the universality and unidirectionality of the theory, with many examples in which age-specific death rates for population subgroups have risen over time, most notably the HIV/AIDS pandemic (Gaylin and Kates, 1997). Gersten and Wilmoth (2002) also criticise the lack of attention Omran's theory pays towards the role of infection in chronic and degenerative diseases, in particular certain cancers.

Around the same time as Omran, Preston collated cause-specific mortality data for a large number

of populations, spanning 48 nations and nearly a century (Preston, 1970; Preston and Nelson, 1974). This would enable international comparisons of groups of causes of death over different time periods, and a deeper understanding of the upward trends in life expectancy. In particular, by plotting cause-specific disease rates against overall mortality, Preston and Nelson (1974) saw that, over time, the contribution of infectious diseases to a particular *level* of mortality had become ever smaller. That is to say, as mortality declined, the contribution from infectious diseases also declined. Preston attributed this to an accelerating rate of medical progress guided by the “germ theory of disease”, which public health and science were not able to replicate for NCDs. Preston also traced the excess deaths in older males observed in western societies to cardiovascular diseases, cancer and bronchitis – a direct result of dramatic increases in cigarette smoking (Preston, 1970).

Since its first edition in 1990, the subject of international comparisons of the cause-specific composition of mortality has been the remit of the Global Burden of Disease (GBD) studies (Murray and Lopez, 1996). The studies aim to quantify and compare the burden of diseases, injuries, and risk factors, usually through cross-sectional methods but occasionally by examining trends and subnational populations (Dwyer-Lindgren et al., 2017a; Ezzati et al., 2008). An important innovation of the GBD study was the introduction of a hierarchical classification of groups of causes, with the broadest level divided into three groups: communicable, maternal, perinatal, and nutritional diseases (Group 1), NCDs (Group 2), and injuries (Group 3). Salomon and Murray (2002) made use of the wide-ranging dataset and grouping from the GBD to revisit the epidemiologic transition for the second half of the twentieth century. They found the majority of the change in cause structure occurred among children, with a shift from Group 1 to Groups 2 and 3, and in young adults, where the role of injuries is more dominant for men.

2.3.2 Modelling cause-specific mortality

In studies looking at multiple causes of death it may be desirable to extend disease mapping models to capture the interdependence between causes of death. Cause-specific mortality can exhibit complex correlation structures, with correlations for diseases with common risk factors but anti-correlations for competing causes of death. For studies that already look at mortality for small areas and narrow age groups, breaking the population down further by cause of death

increases the sparseness in the data, and it may be necessary to introduce terms to the smoothing model which share information between causes in order to stabilise the estimation of death rates.

The simplest and most computationally-scalable method is to ignore any correlations between diseases and run separate regression models for each cause of death. This is the approach taken by GBD studies for a vast array of causes of death. When looking at both total mortality and a mutually exclusive, collectively exhaustive list of causes of death, studies typically constrain the cause-specific death rate estimates from separate regressions to sum to the all-cause mortality value. This is because estimates of all-cause mortality are more robust, as the death counts are larger and therefore have lower variance, and the data do not suffer from errors in assignment of cause of death. The GBD studies scale death rates so that the proportions of each cause of death sum to unity. In the context of mortality projections, Wilmoth (1995) also points out that aggregating forecasts from multiple causes of death often leads to upward bias when compared to the total mortality forecast.

A joint modelling approach would allow the borrowing of strength across causes to express correlations between diseases that share common risk factors and similar aetiologies. There is also some redundancy in separate regression models as a unique spatial surface is specified for each cause. In the disease mapping literature, studies have built spatial models to look at a small number of diseases which share spatial components between diseases. Firstly, Knorr-Held and Best (2001) considered a model (also described in Best et al. (2005)) for two diseases, with disease-specific spatial components and one shared component. Held et al. (2005) generalised this approach to model any number of diseases as a weighted sum of shared components. Rather than a general approach, both Downing et al. (2008) and Mahaki et al. (2018) allocated the spatial components to the diseases *a priori* based on knowledge of the common risk factors between cancers such as smoking, obesity and alcohol consumption. However, unless the components are pre-specified based on prior knowledge, the required number of shared components to capture variation in the data needs to be determined through trial and error. This is especially problematic when the number of diseases in the study, and hence the possible number of combinations of shared components, increases.

Foreman et al. (2017) modelled a larger number of causes of death to jointly forecast cause-specific mortality for states in the US. The model was similar to the spatiotemporal models

described earlier, with random walk effects for temporal nonlinearities and a CAR prior for spatial effects, but with the introduction of a multivariate normal prior for causes of death whereby the covariance matrix describes the correlation structure between the 15 cause groups. The model did not, however, share information between age groups.

In studies of the cause composition of total mortality, rather than estimating the absolute death rate for each cause of death, it is possible to reframe the problem using a compositional model which considers the fraction of each cause of death composing total mortality. This was the approach taken by Salomon and Murray (2002) to investigate the dynamics of the proportions of mortality from GBD Groups 1, 2, and 3. The benefit of a compositional model is that the proportions are constrained to sum to unity, and the model can capture covariance between the component causes of death. However, it is not possible to recover absolute cause-specific death rates using the compositional approach without estimating the overall death rate.

2.4 Health inequalities in the UK

While the UK is, by global standards, a wealthy nation with relatively high life expectancy, the nation still suffers vast, preventable inequalities in mortality and morbidity. Health inequalities can be reduced through, amongst other initiatives, progressive social and economic policies, better nutrition programmes, and improved health care. It is important to estimate and understand differences in health outcomes between population subgroups to aid the design of such policies. There are several ways to stratify the UK population and compare inequalities between subgroups. Here, I focus on class, income, geography, and deprivation.

Class and income inequality

The notion of class is prominent in UK society, but health outcomes between classes are difficult to separate from other risk factors such as hazards in manual labour or smoking rates. The Whitehall study of 1967 followed 17,530 men working in the civil service and recorded their mortality over a 10-year period. Marmot et al. (1984) found, by classifying the civil servants into social class according to their employment grade, there was a three-fold difference in mortality between the highest class, administrators, and men in the lowest class, mainly messengers and manual workers. They found, in general, a strong inverse association between grade and mortality,

which Marmot described as a “social gradient”. The men were working stable, sedentary jobs in the same office building in London, so the gradient could not be explained by industrial exposure alone, and the gradient remained even after controlling for smoking. The authors concluded there must be other factors inherent to social class (defined here by employment grade), which explain the mortality differences. A second cohort of Whitehall employees from 1985 to 1988, this time including women as well as men, were screened and asked to answer questions on self-reported ill-health. Marmot et al. (1991) found the social gradient in health had persisted in the 20 years separating the studies. In 2008, Marmot was asked by the Secretary of State to conduct a review into the state of health inequalities in the UK and to use the evidence to design policy for reducing these inequalities. A key plot in the first Marmot Review, released in 2010, depicted the social gradient in mortality for regions in England by socio-economic classification of employment (Marmot et al., 2010).

Income is not a routinely collected statistic in the UK. Nevertheless, using a small survey of 7000 people on three measures of morbidity, Wilkinson (1992) showed health improved sharply from the lowest to the middle of the income range.

Spatial inequality

In 2015, the GBD study released its first subnational estimates of mortality, starting with the UK and Japan. Steel et al. (2018) assessed these data, which divided the UK into 150 regions, finding mortality from all-causes varied twofold across the country, with the highest years of life lost in Blackpool and the lowest in Wokingham. In a study on forecasting subnational life expectancy in England and Wales, Bennett et al. (2015) estimated a 8.2 year range in life expectancy for men and 7.1 year range for women in 2012 between 375 districts. The lowest life expectancies were seen in urban northern England, and the highest in the south and London’s affluent districts. Within London itself, Cheshire (2012) visualised the heterogeneity of mortality in London by assigning tube stops the life expectancy of the nearest ward, revealing that 10 years are lost between two consecutive stops, Canary Wharf and North Greenwich, on the Jubilee line.

Deprivation

There have been substantial efforts in the UK to measure the deprivation of an area. Since 2004, the standard deprivation indicator in England has been the Index of Multiple Deprivation (IMD)

– a composite indicator for each Lower-layer Super Output Area (LSOA²) covering income, unemployment, health, crime and environmental data sources (Ministry of Housing, Communities & Local Government, 2019). The Marmot Review presented life expectancy and disability-free life expectancy against IMD at the Middle-layer Super Output Area, which exhibit strong social gradients (Marmot et al., 2010). The GBD study found the 15 most deprived areas had consistently raised mortality, especially for all causes, lung cancer and chronic obstructive pulmonary disease. Deprived areas in London, such as Tower Hamlets, Hackney, Barking and Dagenham had lower rates of premature mortality than expected for that level of deprivation (Steel et al., 2018). Bennett et al. (2018) jointly estimated death rates by age, year and deprivation decile. They found that since 2011, although national life expectancy has continued to increase, the rise in female life expectancy has reversed in the two most deprived deciles. Using data from Public Health England, the second Marmot Review in 2020 also reported that female life expectancy declined in the most deprived decile between the periods 2010-12 and 2016-18 (Marmot et al., 2020). Digging further into these trends by region, the report found this trend was seen in all regions except London, the West Midlands and the North West, and that male life expectancy in the bottom decile also decreased in the North East, Yorkshire and the Humber, and the East of England.

Recent trends in health outcomes

Since the turn of the millennium, there have been two periods of contrasting public health policy in the UK. The early 2000s saw the implementation of the English health inequalities strategy under New Labour, with explicit goals of reducing geographical inequalities in life expectancy. The strategy saw a large increase in public spending targeting the social determinants of health, with policies on supporting families, tackling deprivation, and preventative healthcare.

Following the change in government in 2010, the strategy came to an end. The Conservative government implemented a widespread series of cuts to public services, collectively known as austerity. This included both tight restrictions on the healthcare budget as well as a sweeping reorganisation of the NHS in hope of improving the organisation's efficiency, and cuts to the wider determinants of social health, such as housing and education (Ham, 2023).

²See Chapter 3 for descriptions of spatial units.

Although it is difficult to isolate the causal effect of the period of austerity on health outcomes, there is some evidence of differences in health outcomes in these periods. Barr et al. (2017) analysed the trends in life expectancy for different quantiles of deprivation and provided evidence that the English health inequalities strategy achieved its aim of reducing the gap in life expectancy between the 20% most-deprived areas and the rest of the English population, and that the trends were reversing since 2012. These trends have been found across the life course: with rising infant mortality associated with childhood poverty (Taylor-Robinson et al., 2019); increases in “deaths of despair” (drug overdose, suicide, alcoholic liver disease) for those in middle ages (Angus et al., 2023; Hiam et al., 2020); and falls in female life expectancy at 65 and 85 (Hiam et al., 2018). Alexiou et al. (2021) found strong associations between cuts to local government and the change in district-level life expectancy from 2013 to 2017. As written in the *The New York Times*, “after eight years of budget cutting, Britain is looking less like the rest of Europe and more like the United States, with a shrinking welfare state and spreading poverty” (Goodman, 2018).

Although fiscal policies of austerity were adopted by many countries in response to the global financial crash of 2008, there is evidence that public health has deteriorated more in the UK than in other countries. In an international study comparing mortality trends in England and Wales to 22 industrialised countries, Leon et al. (2019) showed that although there was a general slowdown in improvement of life expectancy across many nations, the slowdown in the most recent period of the study, 2011-16, was more pronounced in England and Wales. More recently, *The Economist* found the same evidence, comparing the long-run trend from 1980-2011 through to 2022 for 12 European countries: “longer-run slowdowns in life expectancy are observable in other European countries... but none has stalled quite as much as Britain” (*The Economist*, 2023). The UK has also performed worse as measured by cancer survival rates and infant mortality compared to other industrialised countries (OECD, 2016).

After a decade of cuts, the UK entered the 2020s facing the greatest public health challenge for a generation: the Covid-19 pandemic. Unsurprisingly, England and Wales suffered one of the highest excess deaths tolls relative to other high-income countries (Kontis et al., 2020).

It is important to estimate how health inequalities have changed in different areas of the country through this period of substantial change in economic, social, and healthcare policy. Small-area health statistics, and in particular those at high resolutions, not only reveal the extent of the

mortality differences between neighbourhoods, but can also identify the areas at highest risk, allowing public health interventions to target the most disadvantaged groups.

2.5 Summary

Death rates vary by sex, age group and across time and space. Studying mortality variations at the small-area level introduces sparsity in the data, and statistical smoothing models must be used to obtain robust estimates of death rates. These models should be flexible to allow for variation across age, space and time, and should consider interactions between each of the dimensions. Modelling mortality from specific causes of death introduces further challenges, both through increased sparsity in the death counts and through correlations between diseases with common risk factors.

Following a change in government in 2010, the UK abandoned its strategy to reduce health inequalities. Since then, there has been a decline in female life expectancy in the most deprived areas.

Chapter 3

Data

3.1 Overview

This chapter presents the datasets and data cleaning processes that are common between the proceeding analysis chapters.

3.2 Geographies of England

This thesis is concerned only with England, as Scotland, Wales, and Northern Ireland each have their own separate deprivation data which are not comparable. The geographies in England form a nested hierarchy of spatial units from regions to districts to Middle-layer Super Output Areas (MSOAs) to Lower-layer Super Output Areas (LSOAs). The number of units for each geography are summarised in Table 3.1.

Table 3.1: The numbers of each spatial unit of England used in the thesis and their populations in 2019.

Geography	Number of units	Median (5 th -95 th percentile) population in 2019
region	9	5,934,037 (3,536,336-9,092,877)
district	314	140,271 (68,238-380,483)
MSOA	6791	7985 (5760–11,917)
LSOA	32,844	1620 (1235-2468)

England is divided into nine regions (North East, North West, Yorkshire and the Humber, East Midlands, West Midlands, East of England, South East, South West, London). Within these

regions, there are 314 local authority districts. Districts are administrative geographies formed from a mixture of London boroughs, metropolitan and non-metropolitan districts, and unitary authorities. They are responsible for local policies, and are therefore subject to local government restructuring and boundary changes. To ensure geographic consistency, all data were mapped to the district boundaries from 2020.

Output Areas (OAs) are the smallest building block for spatial census statistics, with between 40 and 250 households and typically 100 to 625 people, and are designed to have some socioeconomic homogeneity. LSOAs are a type of census geography made up of around four or five OAs. MSOAs are then composed of around four or five LSOAs, and these MSOAs fit within district boundaries. OAs, LSOAs, and MSOAs are all statistical units which are designed by the Office for National Statistics (ONS) purely for analysis purposes, so researchers can use spatial units with similar, but small, population sizes. No policies are created using these boundaries (Office for National Statistics, 2022a). Again, for geographic consistency, all data were mapped to the output area hierarchy from the 2011 census.

3.3 Data sources

3.3.1 Deaths

This thesis is primarily concerned with modelling death rates for small areas in England. This requires two data sources: counts of deaths, and population counts. The counts of deaths come from de-identified civil registration data for all deaths in England from 2002 to 2019. In other words, every death in England from 2002 to 2019.

The data is extracted from the ONS database and held by SAHSU in a secure environment as individual death records are identifiable data. The data are updated every year and are mostly complete for previous years, but a handful of deaths are registered in later extracts if the ONS have been waiting on a coroner's report to identify the underlying cause of death.

Each record comes with information on postcode of residence, allowing us to assign each death into a spatial unit for analysis. For each analysis, deaths were stratified into the following age groups: 0, 1–4, 5–9, 10–14, then 5-year age groups up to 80–84, and 85 years and older. There are also a series of ICD-10 (International Classification of Diseases, Tenth Revision) codes from

the death certificate associated with the underlying and contributory causes leading to the death. Here, I focus only on the underlying cause of death, which has been assigned using selection algorithms to improve consistency between doctors (Office for National Statistics, 2022b).

3.3.2 Population

The second data source we require are population counts. These are taken from mid-year population estimates of the usual resident population by the ONS (Office for National Statistics, 2021a, 2021b). The ONS estimates inter-censal populations on a rolling basis, updating the previous year's value using the change in the population in GP patient registration data as an indicator of the true population change. The LSOA populations are fully consistent with estimates for higher levels in the nested geographical hierarchical including MSOAs, districts, regions and the national total for England (Office for National Statistics, 2021c).

3.3.3 Community deprivation

I used data for the following measures of socioeconomic deprivation from the English Indices of Deprivation:

- Income deprivation (also referred to as *poverty*). The proportion of the geographical population claiming income-related benefits due to being out of work or having low earnings.
- Employment deprivation (also referred to as *unemployment*). The proportion of the relevant population of the geography involuntarily excluded from the labour market due to unemployment, sickness or disability, or caring responsibilities.
- Education, skills and training deprivation (also referred to as *low education*). Lack of attainment and skills, including education attainment levels, school attendance, and language proficiency indicators in the geographical population.

The above measures are the three largest contributors to the Index of Multiple Deprivation (IMD), excluding a domain on health that also uses mortality data. The data are produced at the LSOA level (Ministry of Housing, Communities & Local Government, 2019).

IMD data are not available for every year. The analysis period for the thesis is 2002 to 2019, so I used data for these measures for 2004, as data for 2002 were not available, and 2019. The 2004 data on deprivation domains were reported for LSOA boundaries from the 2001 census. I

mapped these data to the 2011 census LSOA boundaries by assigning the 2001 LSOA score to all postcodes contained within it, then overlaying the 2011 LSOA boundaries, and averaging the score for all constituent postcodes of each LSOA, to obtain the corresponding score for each 2011 LSOA.

The definition of the indicators can change over time. Further, the indicator used for measuring education, skills and training deprivation (low education) is not directly interpretable because it combines multiple concepts and cannot be simply expressed as a proportion of the population. Therefore, I used ranking rather than scores so that comparisons can be made not only across spatial units in a single year, but also across the different years.

The deprivation data for geographies larger than LSOAs in Table 3.1 were created by ranking the population-weighted average of scores for all constituent LSOAs, as done previously for districts (Ministry of Housing, Communities & Local Government, 2019).

3.3.4 Migration

I also used estimates of population turnover, defined as the proportion of households in each LSOA in 2019 who were different from those who had lived there in 2002, from the Consumer Data Research Centre. The Consumer Data Research Centre estimates these proportions by using the names of households members, individually and in combination, and addresses and dates of records from electoral and consumer registers and land registry sales data (van Dijk et al., 2021). Estimates of population turnover for MSOAs were created by taking the mean across all constituent LSOAs (Ministry of Housing, Communities & Local Government, 2019).

3.4 Exploratory data analysis

Before fitting any models, I explored the data visually, in this case looking at how total mortality varies over different cross sections: sex, age, space, time.

The general age pattern, after aggregating all years and separating into age groups, follows a J-shape curve, with raised infant and older age mortality (Figure 3.1). Male mortality is higher at all ages, but particularly in young adulthood (15-29 years) due to injuries resulting from risky behaviours.

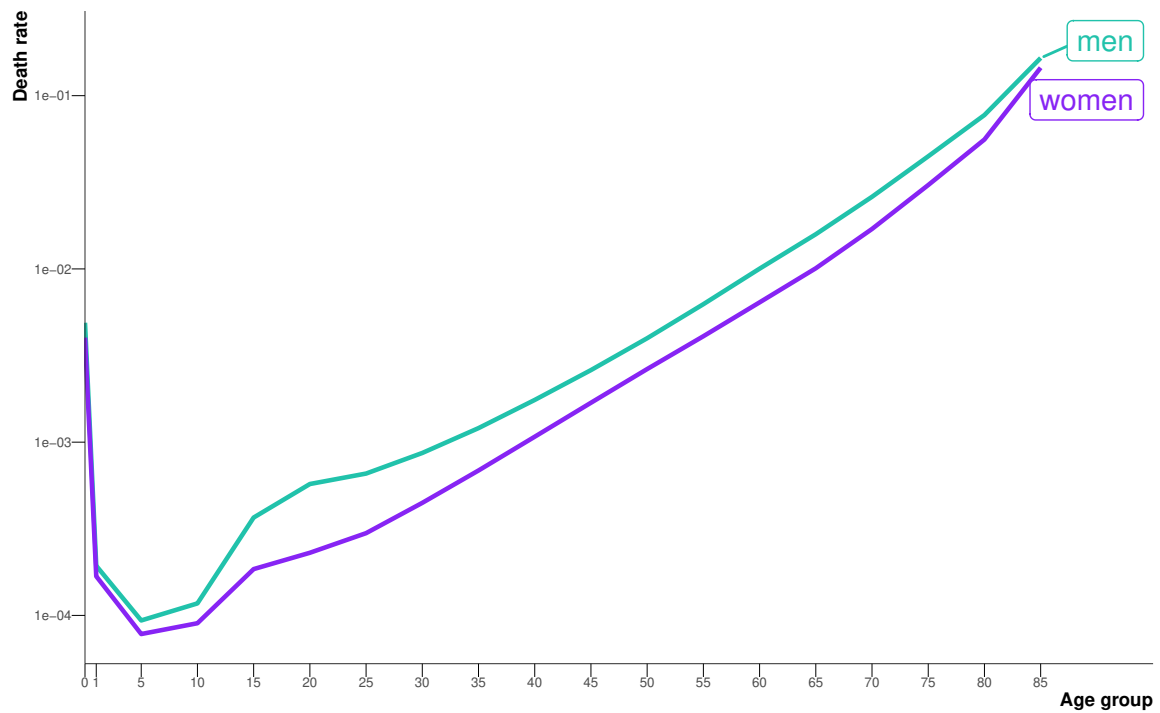


Figure 3.1: Age-specific death rates in England from 2002 to 2019.

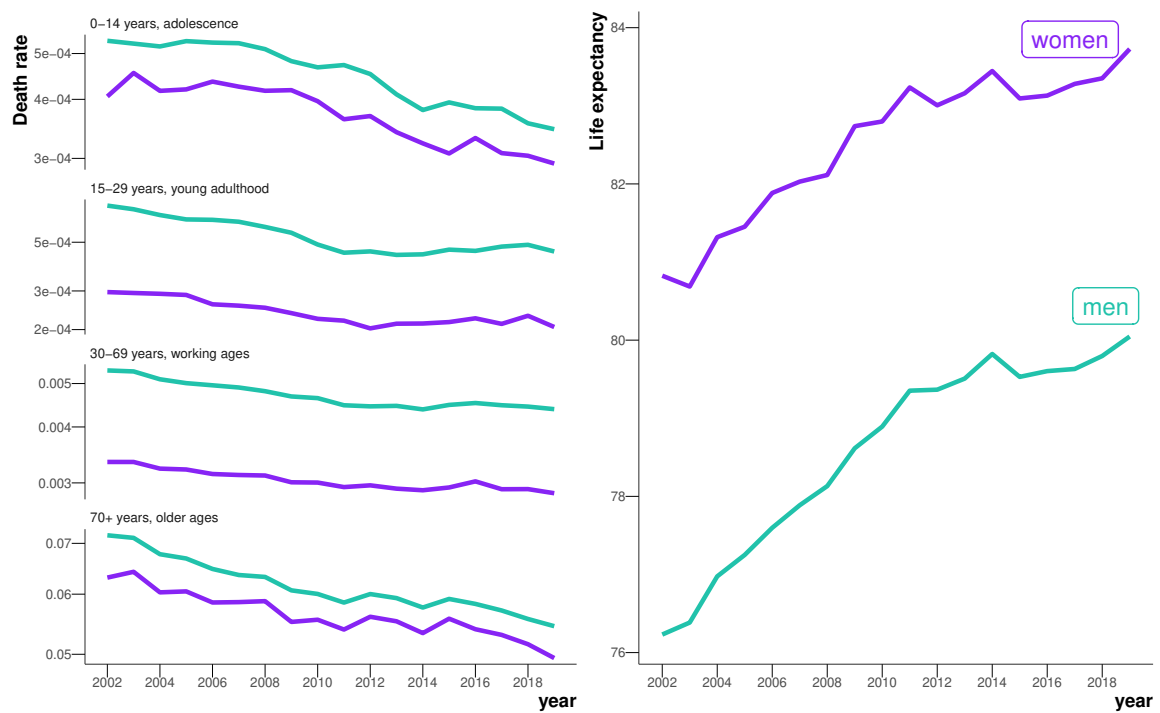


Figure 3.2: Age-specific death rates for broad age groups and life expectancy in England from 2002 to 2019.

Figure 3.2 looks at the trends over time for wider age groups. I did not use the 5-year age groups because the number of deaths for certain age groups at single years were small, and any data presented should be non-disclosive in accordance with SAHSU's data sharing agreement. In general, death rates have decreased from 2002 to 2019 in all age groups, but with slowing progress in young adulthood and working ages (30-69 years). Likewise, life expectancy has improved throughout the study period, but has stalled since around 2010 for both sexes.

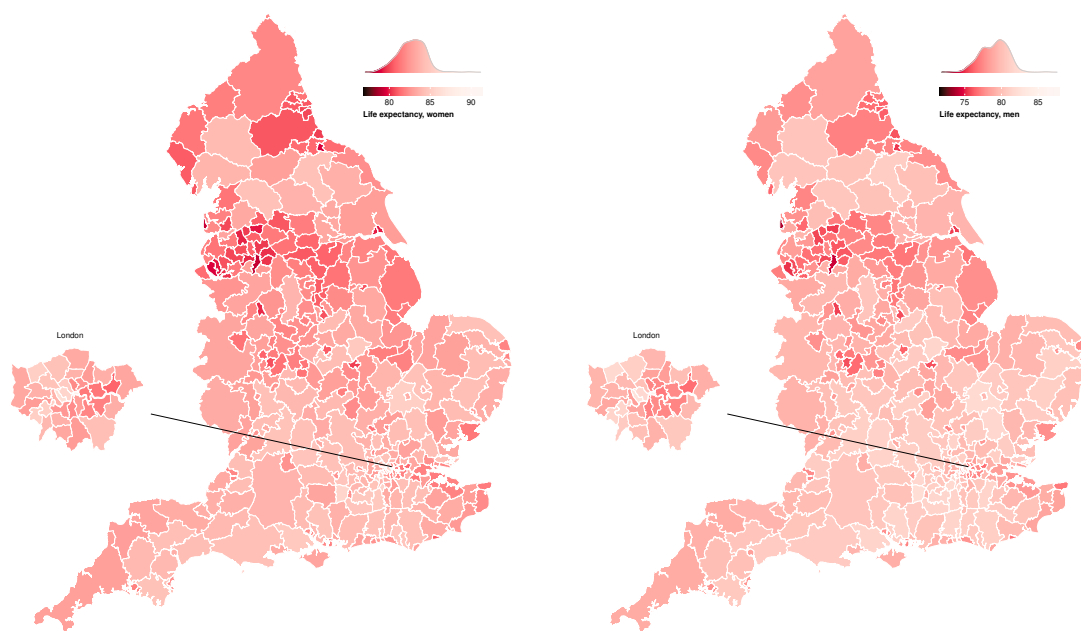


Figure 3.3: District-level life expectancy for total death rates from 2002 to 2019.

Figure 3.3 shows the geography of life expectancy after aggregating deaths over the entire study period by district. For both sexes, the picture is similar: pockets of low life expectancy in the urban North West, North East, and West Midlands.

In this section, I have taken slices across each dimension, but the aim in the following chapters is to calculate death rates for each sex-age-space-time stratum.

3.5 Summary

All-cause and cause-specific mortality estimates require data on deaths and populations. Individual death records with information of sex, age at death, cause of death, year of death, and place of residence have been taken from the ONS database and held by SAHSU in a secure

environment. I have used small-area population estimates created by the ONS. I also introduced datasets on area-level deprivation and migration which will be used to put mortality estimates into context.

Chapter 4

Modelling death rates

4.1 Overview

This chapter presents the modelling choices that are common between the proceeding analysis chapters.

4.2 Why we need to smooth death rates

For each chapter, the quantity of interest is the same: mortality in each age group, spatial unit and year. Empirically, death rates can be calculated from observed data as the number of deaths divided by the population in each strata. Formally, using a , s , and t to index age, spatial unit and time, respectively, we write

$$\hat{m}_{ast} = \frac{\text{deaths}_{ast}}{\text{population}_{ast}}, \quad (4.1)$$

where $\hat{m}_{ast}$ is the empirical death rate.

When the number of deaths becomes small, however, the empirical death rate presents an apparent variability from year to year, or from spatial unit to spatial unit, which is larger than the true differences in the risk of death. The problem is exacerbated for the young ages or causes with lower mortality, where the number of deaths might be zero, or for smaller geographical units, where the population might be very small. As a demonstration, Figure 4.1 shows the expected and simulated number of deaths for a young age group (1-4 years) and an old age group (80-84 years). The simulation assumes the hypothetical spatial unit has a population of 1000 in each age group in each year, and generates deaths using a Poisson distribution and the national

age-specific death rates for each year. Note, given the population sizes in Table 3.1, the true age-specific populations for LSOAs and MSOAs will be smaller and there will be even more zeros and even more noise in the number of deaths than in this simulation. Although there are a large number of deaths in the older age group and it is easy to visualise a curve that fits the data, the death counts for the young age group are extremely sparse and it is difficult to estimate the true underlying death rate. In this thesis, I have used Bayesian hierarchical models to obtain stable estimates of death rates by sharing information across age groups, spatial units, and years.

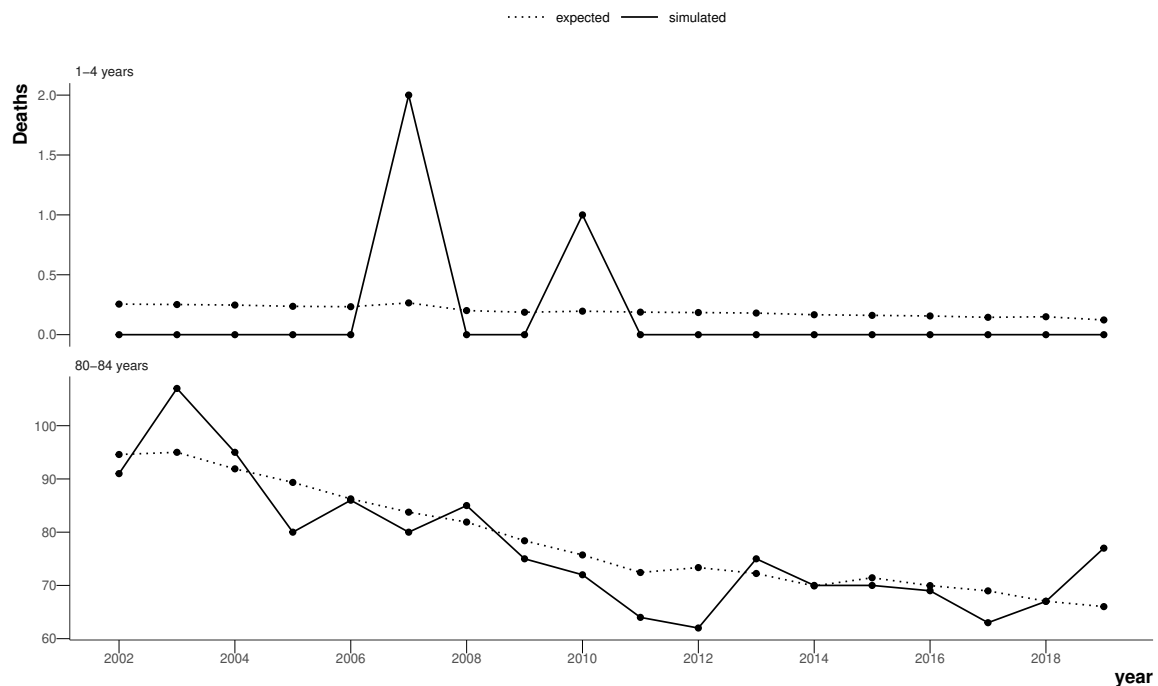


Figure 4.1: Simulated and expected deaths from 2002 to 2019 for a young age group (1-4 years) and an old age group (80-84 years), assuming the national death rates in England and a population size of 1000.

This is a regression task. We want to smooth over the data – the models are not being used for prediction. I tried to design a model that captures as much of the true variation in the data as possible using epidemiological knowledge to choose plausible effects. In other words, the model is “full”, with enough parameters to capture all the true variability. The downside of this approach is that, with more parameters, there is a risk of parametrising the noise in the data. Moreover, over-parametrised models can suffer from a lack of identifiability, which can lead to convergence issues. In contrast, models with fewer parameters, or *parsimonious* models, make inference easier but can mask some of the variance.

I used Bayesian inference methods for the smoothing models, which estimate posterior distributions for each parameter. It is easy to use samples from the posterior distribution to carry the uncertainty of different age groups through to the uncertainty in life expectancy or other nonlinear functions derived from the life table. The Bayesian approach also allows the inclusion of prior knowledge of the parameters, which I use here through spatially-structured prior distributions.

4.3 A model for smoothing death rates

Here, I will present the model for smoothing death rates used for the analysis in Chapter 5. Each chapter thereafter, I will explain how and why the model differs.

I used a Bayesian hierarchical model to obtain stable estimates of death rates by sharing information across age groups, spatial units, and years. I conducted all analyses for women and men separately because mortality levels and trends differ by sex. In the model, the number of deaths in age group $a(= 1, \dots, 19)$, spatial unit $s(= 1, \dots, 6791)$ ¹ and year $t(= 1, \dots, 18)$ follows a negative binomial distribution

$$\text{deaths}_{ast} \sim \text{Negative Binomial}(p_{ast}, r). \quad (4.2)$$

The parameter p_{ast} is

$$p_{ast} = \frac{r}{r + m_{ast} \cdot \text{Population}_{ast}}. \quad (4.3)$$

where $r \geq 0$ is the overdispersion parameter, which accounts for extra variability not captured by other components in the model, and m_{ast} is the death rate. The negative binomial² likelihood can be thought of as a generalisation of the Poisson likelihood, which allows for overdispersion, with larger values of r indicating more similarity to a Poisson distribution. A Poisson distribution is a suitable approximation to the binomial distribution for rare events.

Log-transformed death rates were modelled as a function of time, age group and spatial unit.

The model contains terms to capture the overall level and trend over time of mortality, as well

¹The dimension of s is different for LSOA- and district-level analysis. In the MSOA-level analysis, there are 6791 spatial units.

²The name “negative binomial” is quite difficult to understand in this context. A better, but less popular name, is the gamma-Poisson distribution. Here, the story is simpler: we have a mixture of Poisson distributions where the rate parameters of the Poisson distributions follow a gamma distribution (McElreath, 2020).

as age-specific and space-specific terms to allow deviations from these terms. Specifically, log-transformed death rates are modelled as

$$\log(m_{ast}) = \alpha_0 + \beta_0 t + \alpha_{1s} + \beta_{1s} t + \alpha_{2a} + \beta_{2a} t + \xi_{as} + \nu_{st} + \gamma_{at}, \quad (4.4)$$

where α_0 is the overall intercept across all age groups and spatial units. β_0 quantifies the overall trend (over time) across all age groups and spatial units. α_{1s} and β_{1s} measure deviation from the overall intercept and trend terms, respectively, for each spatial unit. α_{2a} and β_{2a} measure deviation from the global level and trend, respectively, for each age group. I used first-order random walk priors on α_{2a} and β_{2a} so that they vary smoothly over adjacent age groups, with the form $A_a \sim \mathcal{N}(A_{a-1}, \sigma_A^2)$ for both age-specific terms α_{2a} and β_{2a} . I constrained $\alpha_{21} = 0$ and $\beta_{21} = 0$ so each random walk was identifiable and centred on the corresponding overall term.

ξ_{as} is an age group-spatial unit interaction term, which quantifies space-specific deviations from the overall age group structure given by α_{2a} . This allows different spatial units to have different age-specific mortality patterns, and each age group's death rate to have a different spatial pattern. This interaction term was modelled as $\mathcal{N}(0, \sigma_\xi^2)$.

ν_{st} and γ_{at} allow space- and age group-specific nonlinearity in the time trends. For each spatial unit and age group, I again used first-order random walk priors with $\nu_{s1} = \gamma_{a1} = 0$ so that the terms were identifiable.

The spatial intercepts and slopes, α_{1s} and β_{1s} , were modelled as nested hierarchical random effects. For the MSOA analysis, MSOAs were nested in districts, which were, in turn, nested in regions. For the LSOA analysis, LSOAs were nested in MSOAs, which were nested in districts. The terms for the largest spatial unit were centred on zero to allow the spatial effects to be identifiable.

All standard deviation parameters of the random effects had $\sigma \sim \mathcal{U}(0, 2)$ priors, which were used for a previous mortality modelling study by the group (Bennett et al., 2015). I performed a sensitivity analysis using the less informative $\sigma \sim \mathcal{U}(0, 100)$ prior, to which the model was robust (the largest inferred standard deviation parameter was for the age group intercept with a mean

around 0.9). For the global intercept and slope, we used the diffuse prior $\mathcal{N}(0, \sigma^2 = 10^5)$. The overdispersion parameter r had the prior $\mathcal{U}(0, 50)$.

Table B.1 shows all model parameters, their priors and dimensions for the MSOA-level model in Chapter 5. Table 4.1 summarises the model adequacy and consistency checks performed for the analyses.

Table 4.1: Summary of model posterior checks.

Type of check	Checks performed
Model adequacy	Check all posterior death rates are between 0 and 1; scatter plots of posterior predictions of death rates against observed data by age group and year; inspection of residuals by age group and year
Model bias	Compare aggregated posterior predictions of deaths with uncertainty to national number of deaths from data each year; evaluate model shrinkage by inspecting the range of life expectancy between the top and bottom percentiles (aggregating 67 or 68 MSOAs) in 2002 and 2019 estimated using the model and the data
Consistency between models	Aggregate posterior predictions of deaths from MSOA-level model to district-level and compare to posterior predictions of deaths from the same model run at district level (and same checks for LSOA- and national-level where appropriate)

Although a random walk approach has been used here to model the J-shape age-mortality association, there are a number of alternatives. For example, Gonzaga and Schmertmann (2016) use a series of linear splines over the age dimension. Alexander et al. (2017) describe an approach using the first three principal components of standard mortality curves, where the first component represents baseline mortality, and the second and third components allow offsets for higher child mortality and higher adult mortality. However, both these approaches require the modeller to manually specify either the number of basis splines and position of the knots or the

number of principal components required to accurately describe the age-mortality relationship. This becomes more difficult when modelling several different diseases, which might not follow a J-shape, particularly those with a skew towards older ages such as prostate cancer. Random walks are more flexible and require less tuning in this respect, and are also used here to model age-specific slopes over time, for which we have no such prior demographic knowledge.

4.4 Inference

The decision was made early in my PhD research to use Markov chain Monte Carlo (MCMC) sampling methods for inference, as this is the “gold standard” with guarantees that, under mild conditions, the sequence of samples will asymptotically converge to the true posterior distribution (Roberts and Rosenthal, 2004). Although sampling approaches are the focus here, the R-INLA package, which uses approximate inference for latent Gaussian fields and has implementations of common spatial models, could also have been used.

Bayesian models can be specified in a probabilistic programming language. The starting point for this project was the NIMBLE package (de Valpine et al., 2022, 2017). NIMBLE uses the BUGS (“Bayesian inference Using Gibbs Sampling”) syntax for defining a hierarchical model, which my research group has a lot of experience with as WinBUGS, one of the earliest software packages for Bayesian analysis, was developed largely in the department for use on SAHSU studies. NIMBLE has an R interface but compiles models to C++ for speed and scalability. It also increases the sampling efficiency by automatically finding conjugate relationships between parameters in the model and marginalising over them wherever possible. The group also has a close relationship with the lead developer of NIMBLE.

Nevertheless, Bayesian inference is difficult to scale, and some of the models in this thesis had in excess of 10^6 parameters and took NIMBLE between 10 and 14 days to collect enough posterior samples. One of the main issues with NIMBLE was that the vast majority of the parameters in the model could not exploit efficient conjugate samplers, and instead used variants of basic Metropolis-Hastings samplers, which, despite numerous efforts at tuning, were inefficient. Although NIMBLE could execute a reasonable number of samples per second, the MCMC chains were struggling to explore the posterior efficiently so the *effective* sample size per second was low. This is a common problem in spatial and spatiotemporal models, where the parameters are correlated by design. To

overcome these mixing issues, the chains had to be run for longer and thinned (i.e. take every n^{th} sample so the Markov chain samples are closer to independent, which is better for computational reasons than storing a large number of correlated samples).

In an effort to increase the sampling efficiency of the models, I tested a number of alternative probabilistic programming languages across R, python and Julia, which I have detailed in Rashid (2022). In particular, I focussed on packages that have implemented the more efficient No U-Turn Sampler (NUTS) (Hoffman and Gelman, 2014). In the end, I chose to rewrite the models in NumPyro (Phan et al., 2019) because it was the fastest and inference could be performed on a GPU, rather than CPUs, which is more performant for large models (Lao et al., 2020). The major downside was that NumPyro had not been used extensively by the spatial modelling community so I had to implement the CAR distribution from Equation 2.2 myself, which has since been contributed to the source code (NumPyro documentation, 2023). Rewriting the model in NumPyro and sampling on a GPU cut the runtime down to around a day. NumPyro also has built-in methods for approximate variational inference, such as the Laplace approximation, but these failed to converge to sensible values for these models without heavy customisation of variational function, so I stuck with sampling methods.

4.5 Clean code and open source

I have paid a lot of attention to open sourcing code for all analyses during my PhD. The code is clean, version-controlled, and follows best practices for scientific software engineering. As well having contributed code to open source projects along the way, the code for [statistical models](#), [plots and analysis](#), and the [thesis itself](#) can be found on GitHub.

4.6 Summary

I presented a model for estimating death rates for small areas, using a negative binomial likelihood with intercepts and slopes for each spatial unit and age group, as well as age-space interactions and nonlinear temporal effects for each spatial unit and age group. The model contains a large number of parameters, so I have explored state-of-the-art Bayesian probabilistic programming libraries which allow sampling methods to scale to high-dimensional problems.

Chapter 5

Life expectancy trends in England at the MSOA level

This chapter is based on the peer-reviewed publication *Life expectancy and risk of death in 6791 communities in England from 2002 to 2019: high-resolution spatiotemporal analysis of civil registration data*, published in *The Lancet Public Health* (Rashid et al., 2021), for which I was first author. The paper was published under the CC-BY license, which permits reproduction provided the original work is properly cited. The figures have been reproduced for the thesis, and the text has been updated but remains much the same as the original.

There have been declines in female life expectancy in England for the most deprived deciles since the start of the 2010s (Bennett et al., 2018; Marmot et al., 2020). However, these studies have only broken England down into deciles of deprivation, rather than looking at geographical variation. And studies of trends in life expectancy have been limited to the district level (Bennett et al., 2015), which masks substantial heterogeneity. In this chapter, I address the first objective of the thesis – to estimate trends in life expectancy for small areas in England from 2002 to 2019 – by choosing a smaller unit of analysis than districts: MSOAs.

5.1 Methods

5.1.1 Mortality and population data

I performed a high-resolution spatiotemporal analysis of all deaths in England from 2002 to 2019 (8,646,878 death records, extract date: 25th June 2021), as detailed in Chapter 3. Deaths were

stratified by 19 age groups (0, 1–4, 5–9, ..., 80–84, 85+) and 6791 MSOAs. I did not use 129 death records ($<0.001\%$) for which sex was not recorded. In 48 (0.001%) age-MSOA-year combinations, the number of deaths exceeded population. Most of these were in people aged 85 years and older. In these cases, the population was set equal to the number of deaths.

5.1.2 Statistical analysis

The model for the analysis is described in Chapter 4. Here, the spatial units s are MSOAs. I found that mortality at the MSOA-level was sufficiently low such that death rates never came close to 1, so the negative binomial distribution was a suitable likelihood.

For the main analysis, the spatial intercepts and slopes, α_{1s} and β_{1s} , were modelled as nested hierarchical random effects. For comparison, I also modelled the spatial effects using a BYM model, as in Equation 2.3. The CAR component of the model requires all spatial units to have neighbours. Thus, the MSOAs containing the Isle of Wight, Hayling Island, the Isles of Scilly and Canvey Island were each joined to the nearest mainland MSOA based on road or ferry connections.

The results of the spatial model were virtually identical to the hierarchical random effects model (correlation coefficient 0.999 for female and male sexes; mean difference 0.03 years for women and 0.009 years for men; mean absolute difference 0.07 years for women and 0.09 years for men for life expectancy estimates from the two approaches). I present results from the hierarchical model for two reasons. First, it allows neighbouring MSOAs that fall in different districts to differ more than those within the same district, reflecting the relevance of district as a unit of resource allocation and policy implementation. Second, the hierarchical model was computationally less demanding, with run times about 1.4 times faster than the spatial model.

Inference was performed using MCMC in NIMBLE (de Valpine et al., 2022, 2017). I monitored convergence using trace plots and the $\widehat{R}$ diagnostic (Vehtari et al., 2021) and thinned post burn-in samples to reduce memory and storage use. For women, I ran four chains for 150,000 iterations, discarding the first 50,000 and thinning the remainder by 400 to obtain 1,000 post-burn-in draws from the posterior distribution of model parameters. For men, due to slower mixing, I ran eight chains for 150,000 iterations, discarding the first 100,000 and thinning the remainder by 400.

In 2017, the MSOA in Kensington and Chelsea, London, where Grenfell Tower is located had 119

deaths, compared with 48 in 2016 and 51 in 2018. The additional deaths were caused by a fire in a highrise residential building. This outlier year led to unstable estimates of the long-term trend in life expectancy in this MSOA, and also slightly changed estimates in other MSOAs in the district. To avoid this instability, when applying the statistical model, I replaced the number of deaths for this year with the mean of those in 2016 and 2018 for each age and sex group. When making estimates for 2017, the difference between actual and interpolated deaths was added back to the posterior estimates so that these deaths were counted in the corresponding year. The Grenfell Tower fire incident was, to the best of our knowledge, the only spatially-specific, heavy-tailed mortality event in England over the study period.

I calculated life expectancy at birth, and probability of dying at specific ages by sex and MSOA using life table methods¹.

The reported 95% credible intervals (CrIs) represent the 2.5th to 97.5th percentiles of the posterior distribution of estimated life expectancies. I also report the posterior probability that the estimated change over time in a MSOA represents an increase versus a decrease in life expectancy. Posterior probability represents the inherent uncertainty in life expectancy trends. In a MSOA in which the entire posterior distribution of life expectancy in 2019 is greater than in 2002, there posterior probability of an increase of 1.0, and hence around a 0.0 probability of a decrease, and vice versa. Posterior probabilities more distant from 0.5 indicate more certainty.

5.2 Results

5.2.1 Inequalities in life expectancy

In 2019, there was a 20.6 year (95% CrI 17.5–24.2) gap for women between the MSOA with the highest life expectancy (a MSOA in Camden, London; 95.4 years (92.4–98.7)) and the MSOA with the lowest life expectancy (a MSOA in Leeds; 74.7 years (73.4–76.2)). The gap was 27.0 years (23.4–31.1) for men, between a MSOA in Kensington and Chelsea, London (95.3 years (92.1–99.3)) and a MSOA in Blackpool (68.3 years (66.9–69.6)). The difference between the first and 99th percentiles of life expectancy in 2019 was 14.2 years (13.9–14.5) for women and 13.6 years (13.4–13.9) for men.

¹See Appendix A for full details on demographic methods used in this thesis.

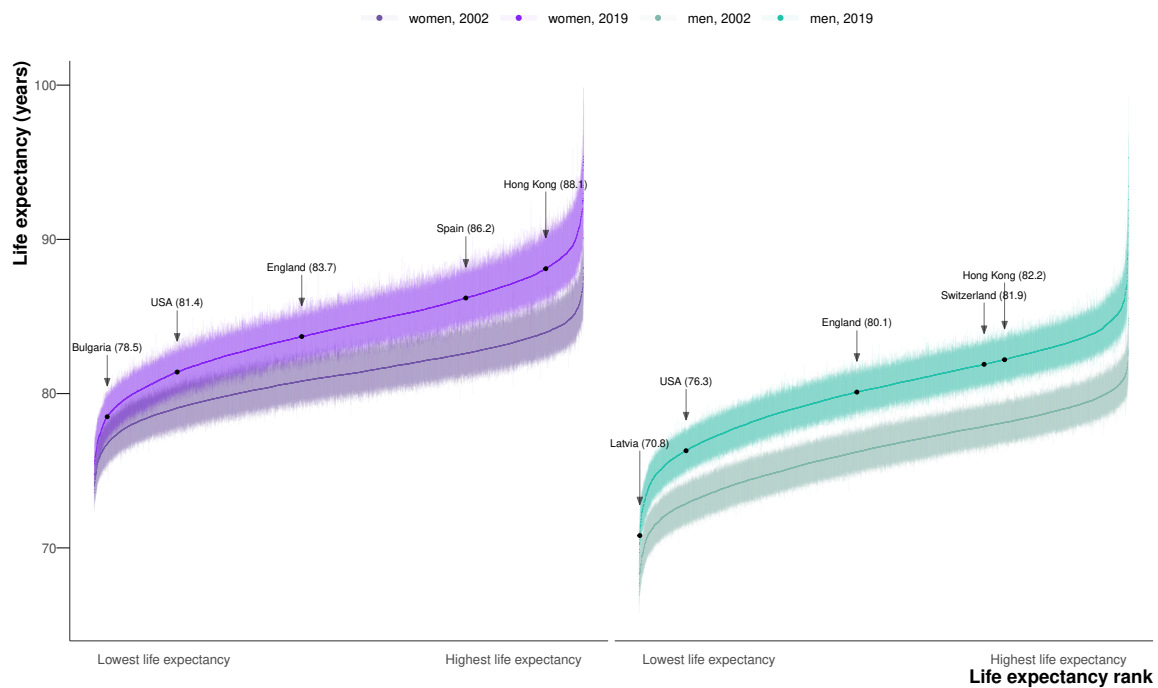


Figure 5.1: Ranked MSOA life expectancies in 2002 and 2019. Each point shows the posterior median life expectancy estimate for each MSOA, forming a curve; error bars are 95% credible intervals. Arrows indicate national life expectancies in England and selected comparator countries with life expectancies within the range of English MSOAs. Hong Kong had the highest global female and male life expectancies. In the EU, Bulgaria had the lowest and Spain had the highest life expectancies for women; Latvia had the lowest and Switzerland had the highest life expectancies for men. Life expectancy for England was calculated from the data used here, and for other countries from World Bank estimates in 2019 (World Bank, 2022).

When all MSOAs were ranked on the basis of their life expectancy, the difference between successively ranked MSOAs was particularly large for the approximately 5% of MSOAs with the lowest and highest life expectancy (seen as the sharper decline or rise at the two ends of the ranked life expectancy curve in Figure 5.1), indicating distinct groups at extreme advantage and disadvantage.

The 124 (1.8%) of 6791 MSOAs with the lowest female life expectancy and 262 (3.9%) MSOAs with the lowest male life expectancy in 2019 were located in urban areas, particularly in the north, including Blackpool, Leeds, Liverpool, Manchester, and Newcastle. Many of the MSOAs with the highest life expectancy, especially for men, were in London and its neighbouring districts (Figure 5.2²).

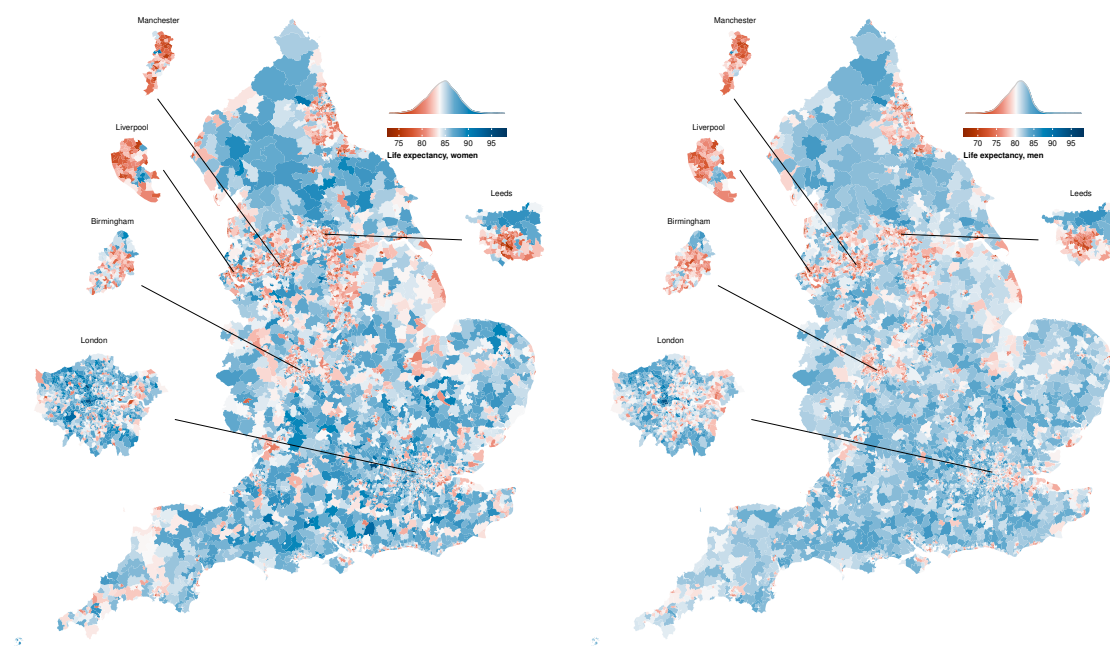


Figure 5.2: Map of life expectancy and the distribution of life expectancy in 2019. The areas in white have a life expectancy equal to the national life expectancy.

In general, as life expectancy increased, the credible interval of life expectancy increased (Figure 5.3). The widest credible interval in any sex-MSOA combination in 2019 was 10.6 years.

²For an interactive version of this figure, see the [visualisation](#). For a hex-cartogram of life expectancy in 2019, see Figure C.1.

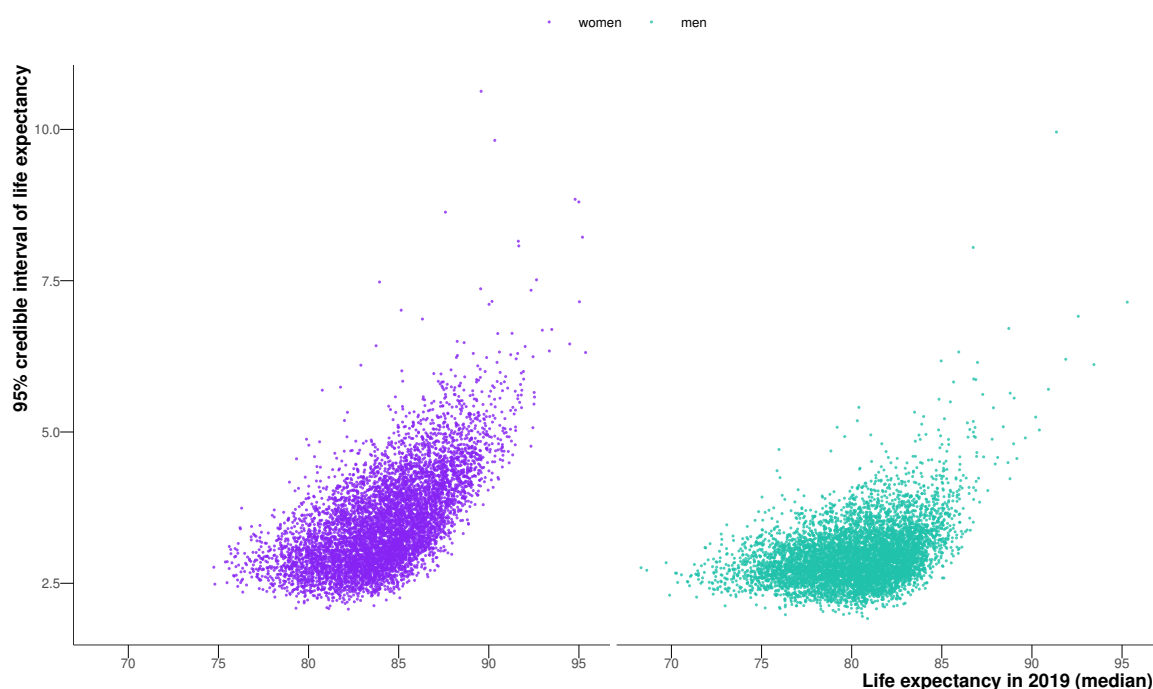


Figure 5.3: Median estimates of MSOA-level life expectancy in England in 2019 against the 95% credible intervals.

5.2.2 Change in life expectancy

Female and male life expectancy were correlated across MSOAs with a correlation coefficient of 0.87 (Figure 5.4). Female life expectancy was higher than male life expectancy in all but 15 MSOAs. The female advantage was more than 5 years in 1498 (22.1%) of 6791 MSOAs and 1–5 years in another 5187 (76.4%). From 2002 to 2019, a decline in life expectancy was more probable than an increase in 124 mostly urban MSOAs of 6791 (1.8% of all MSOAs) for women, with posterior probabilities of greater than 80% in 34 of these. The largest estimated decline of 3.0 years (0.9–5.3; posterior probability of a decline >0.99) occurred in a MSOA in Leeds (Figure 5.5, Figure 5.6).

Elsewhere, median posterior change was positive, ranging from less than 1 year in 408 MSOAs to more than 7 years in 63 MSOAs. Posterior probability of an increase in male life expectancy was more probable than a decrease in all but one MSOA in Blackpool, in which life expectancy changed by -0.4 years (-2.3 to 1.6 ; posterior probability of a decline 0.64). For the other MSOAs, the increase ranged from less than 1 year in 31 MSOAs to more than 7 years in 114 MSOAs. The largest increases in female and male life expectancies were seen in some MSOAs in and around London (e.g., in the London Borough of Camden). In 5133 (75.6%) MSOAs, male life

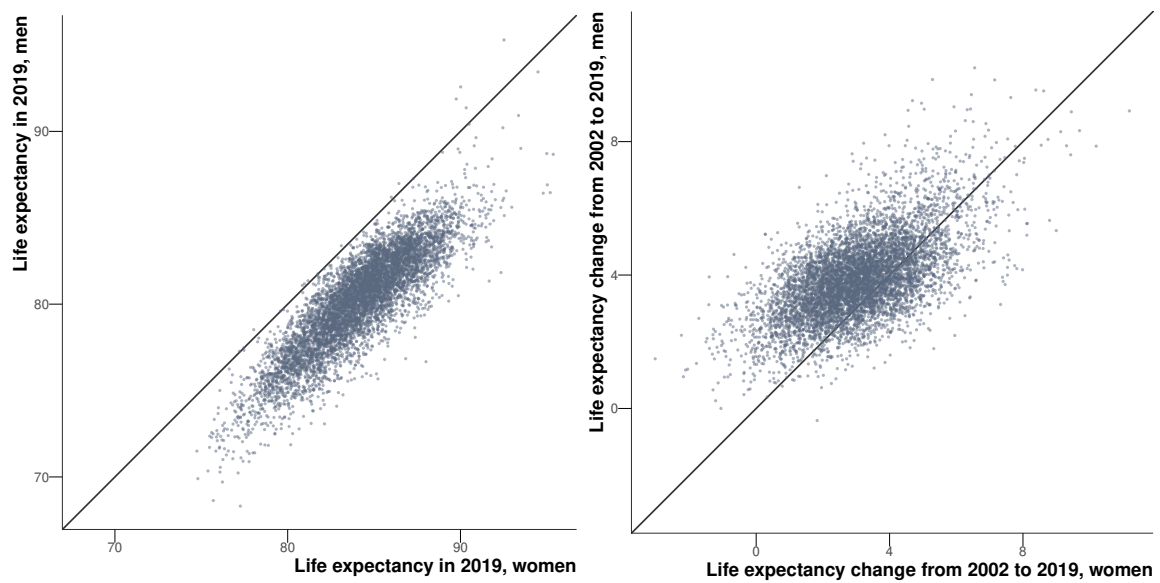


Figure 5.4: Comparison of female and male life expectancy in 2019 and change from 2002 to 2019.

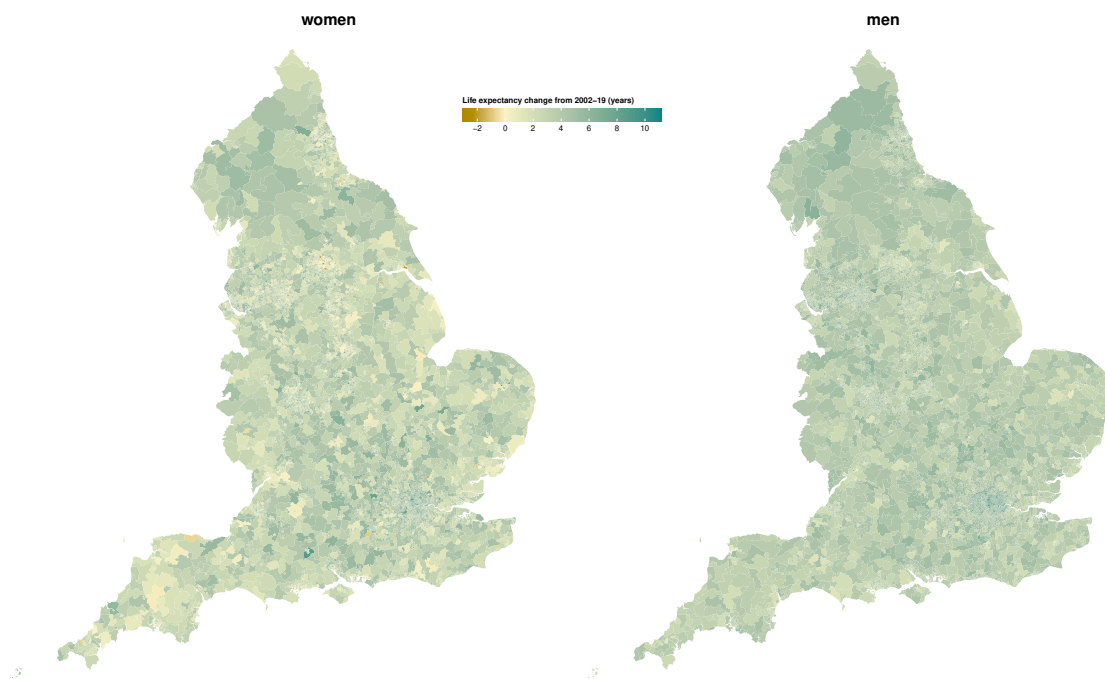


Figure 5.5: Geography of change in life expectancy from 2002 to 2019.

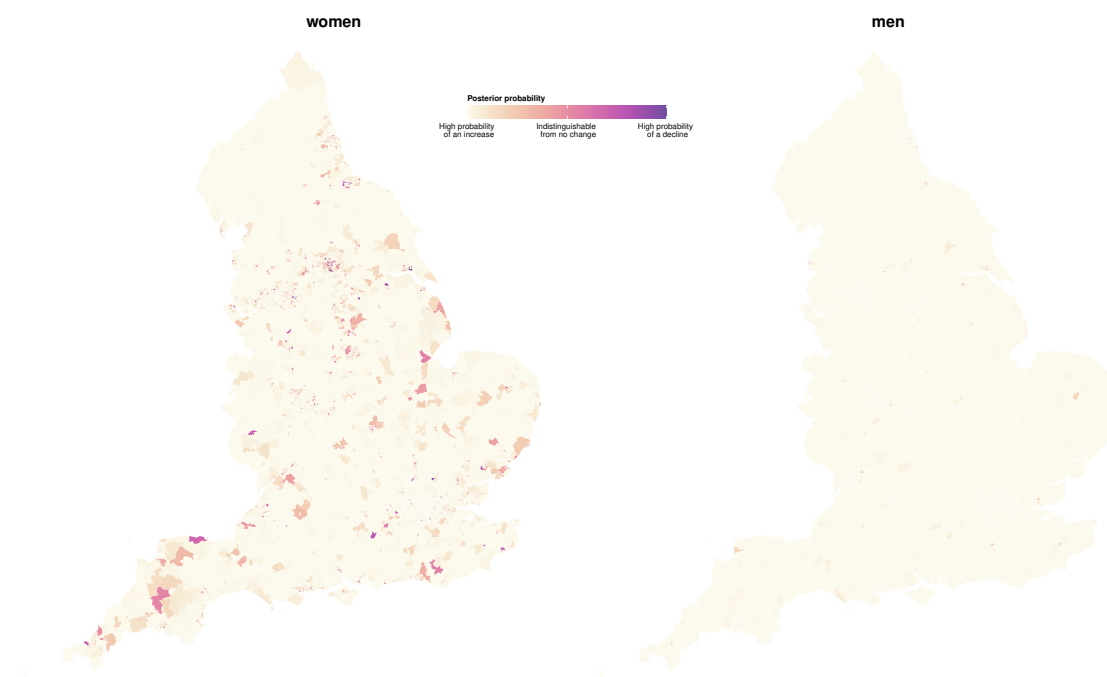


Figure 5.6: Map of posterior probability that the estimated change represents an increase in life expectancy from 2002 to 2019.

expectancy increased more than female life expectancy (Figure 5.4), leading to a closing of the life expectancy gap between female and male sexes.

The life expectancy increase from 2002 to 2019 was smaller in MSOAs where life expectancy had been lower in 2002, and vice versa, especially for women, which led to a larger life expectancy inequality across MSOAs in 2019 than in 2002 (Figure 5.7, Figure 5.8). Specifically, the aforementioned 20.6 year (17.5–24.2) gap for women and 27.0 year (23.4–31.1) gap for men between the lowest and highest MSOA life expectancies in 2019 were larger than those in 2002 by 4.3 years (–1.3 to 9.3) for women and 7.7 years (4.0 to 11.7) for men. Similarly, the gap between the first and 99th percentiles of MSOA life expectancy for women increased from 10.7 years (10.4–10.9) in 2002 to reach 14.2 years (13.9–14.5) in 2019, and for men increased from 11.5 years (11.3–11.7) in 2002 to 13.6 years (13.4–13.9) in 2019.

When broken down by time period, the vast majority of MSOAs saw a life expectancy increase in 2002–06 and 2006–10 (Figure 5.9). By contrast, women in 351 (5.2%) MSOAs had a median posterior change in life expectancy in 2010–14 that was negative. By 2014–19, the number of MSOAs with a negative median posterior change had risen to 1270 (18.7%) for women, with men in 784 (11.5%) MSOAs also showing a decline. These MSOAs tended to be places in which

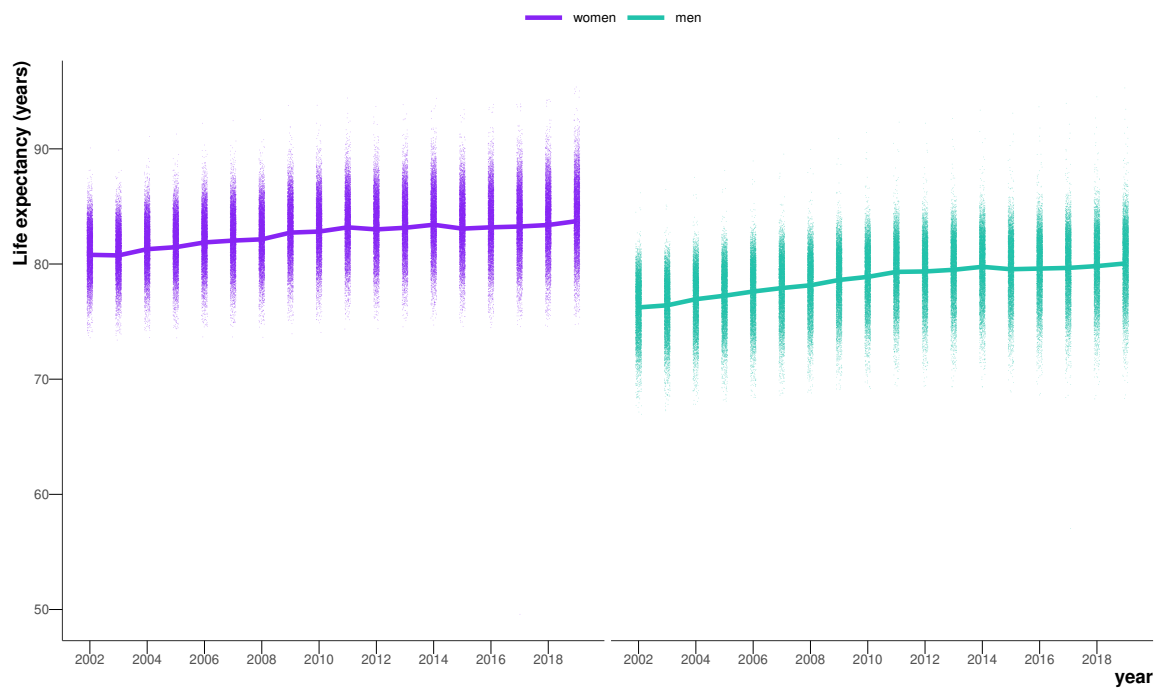


Figure 5.7: Distribution of MSOA life expectancies in each year from 2002 to 2019. Each point shows one MSOA. The central line shows national life expectancy.

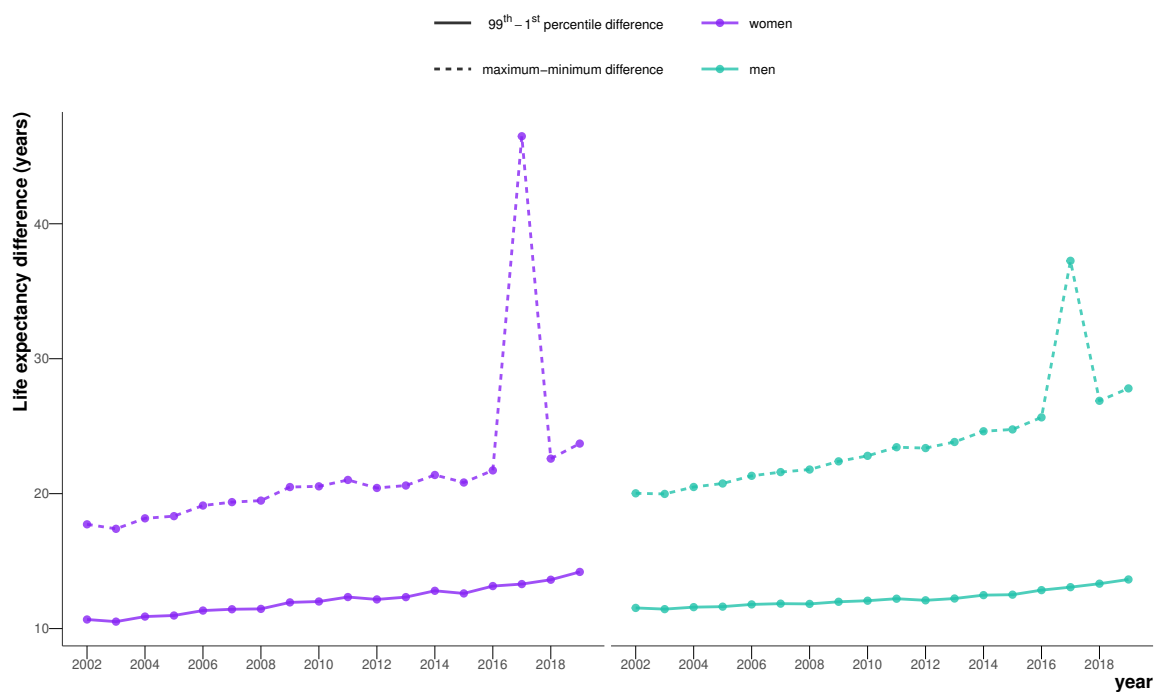


Figure 5.8: Maximum (highest) to minimum (lowest) and 99th to first percentile differences in life expectancy across 6791 MSOAs, 2002–19. The large difference in 2017 is due to the low life expectancy in the MSOA where the deaths in the Grenfell Tower (Kensington and Chelsea, London) fire took place.

life expectancy was already low.

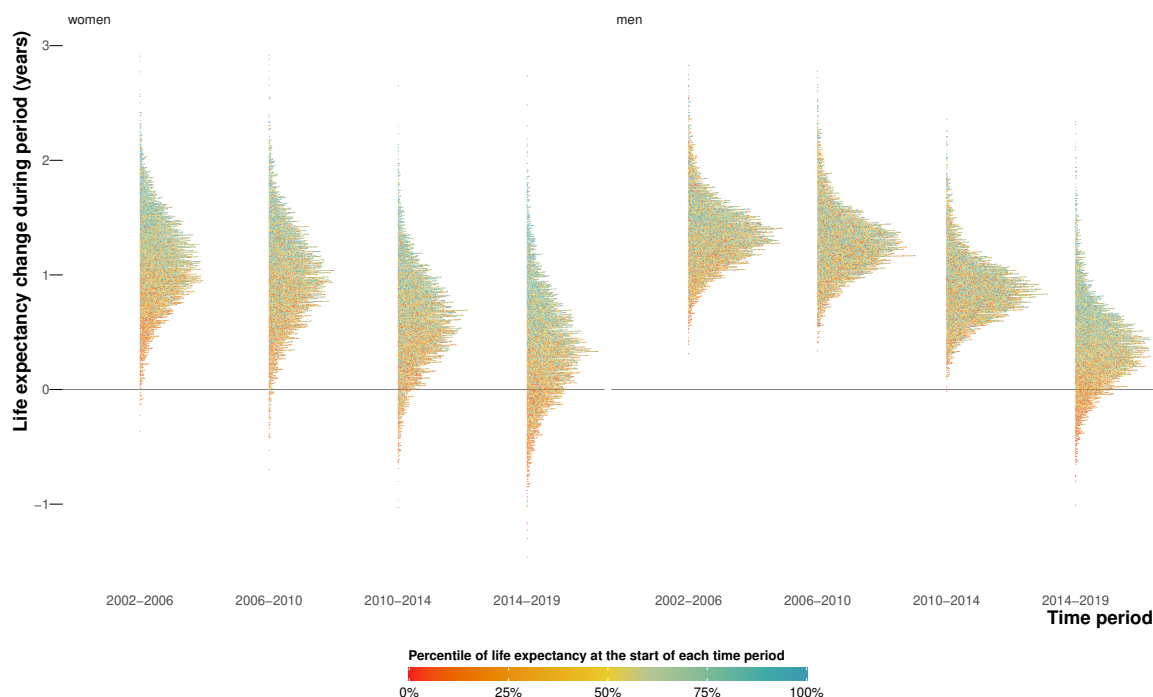


Figure 5.9: Change in MSOA life expectancy in different time periods, 2002–19. Each point shows the posterior median change in one MSOA. MSOAs are coloured by their life expectancy at the beginning of each period (e.g., for 2014–19, they are coloured by life expectancy in 2014).

5.2.3 Life expectancy and deprivation

Life expectancy at birth was inversely associated with the extent of unemployment, poverty, and low education in MSOAs in 2002 and 2019 (Figure 5.10). There was substantial variation in life expectancy across MSOAs at any level of poverty or unemployment seen in the vertical spread of points in Figure 5.10. From 2002 to 2019, there were, on average, smaller gains in life expectancy in the MSOAs with the highest levels of unemployment, poverty, and low education than in those in the lowest levels, especially for women.

5.2.4 Inequalities in probability of survival

Similar to life expectancy, there were large inequalities in the probability of surviving from birth to 80 years, which ranged from 0.42 to 0.87 in women and 0.27 to 0.85 in men across MSOAs in 2019. These large survival inequalities were present at every stage of the lifecourse including childhood and early adolescence (0–14 years), young adulthood (15–29 years), working ages (30–69 years), and older ages (70–79 years) (Figure 5.11). Specifically, the probability of dying at

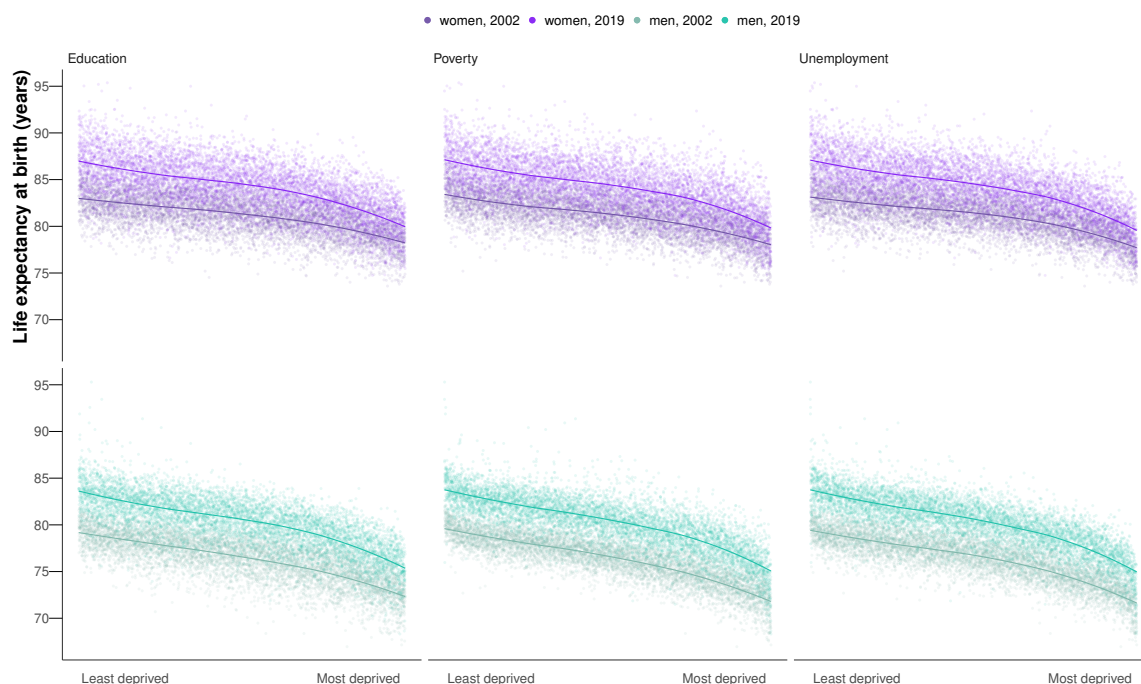


Figure 5.10: MSOA life expectancy in relation to measures of socioeconomic deprivation in the MSOA in 2002 and 2019. The socioeconomic measures are poverty, unemployment, and education. The lines show the smooth relationship fitted with locally estimated scatterplot smoothing for each year.

different stages of the lifecourse in the 99th percentile of MSOAs was between 2.6 and 3.1 times that of the first percentile for female and male sexes in 2019. From 2002 to 2019, the relative inequality across MSOAs (ie, ratio of the 99th to the first percentile) in the probabilities of dying increased at every stage of the life course; the absolute inequality (ie, difference between the 99th and first percentiles) decreased slightly in all combinations except for working age women (30–69 years). Within childhood and adolescence, there were particularly large inequalities in infant mortality (0 to <12 months), with a ratio of the 99th to the first percentile of MSOAs being 3.2 for female and male sexes in 2019. Infant mortality increased from 2014 to 2019 in 1378 (20.3%) MSOAs for girls and 888 (13.0%) for boys, many of which experienced a decline in life expectancy.

5.3 Discussion

The high-resolution analysis over space and time shows that life expectancy has not only ceased to increase, but has declined in many communities in England since 2010. The decline has accelerated since 2014, affecting the female population of 18.7% of MSOAs and the male population

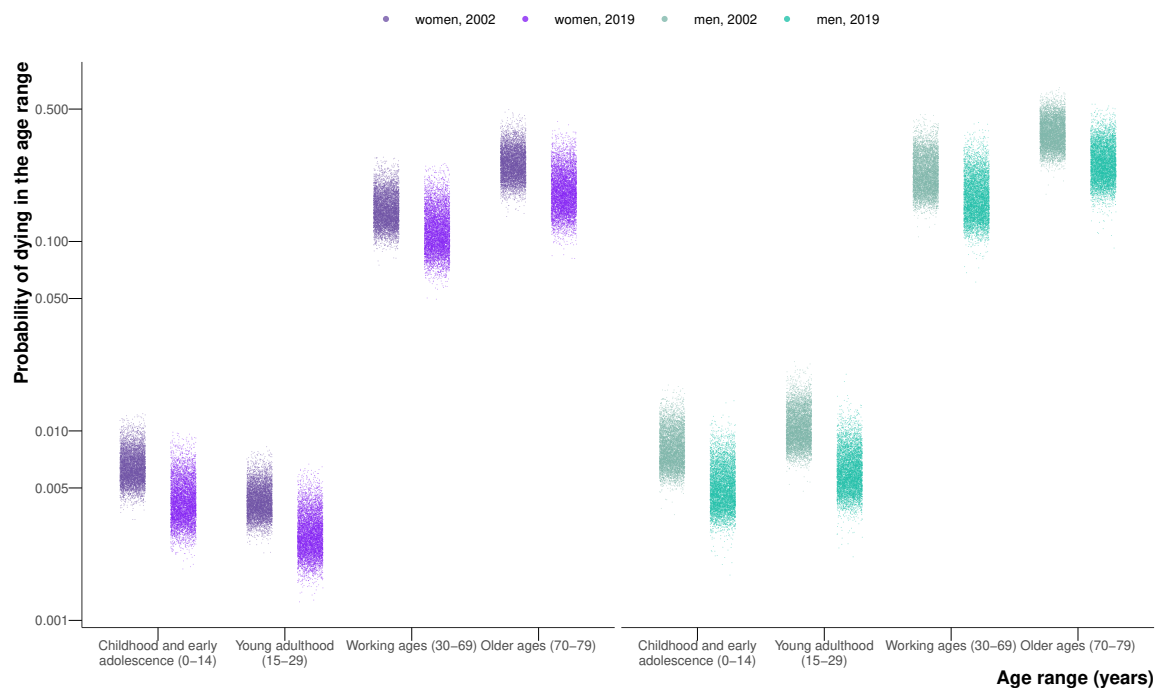


Figure 5.11: Probability of dying in specific ages in 6791 MSOAs in England in 2002 and 2019. Each point shows one MSOA. The vertical axis uses a log scale so that the large differences in survival across ages can be seen.

of 11.5% of MSOAs. In 1.8% of MSOAs, women have had a long-term decline in life expectancy over two decades. MSOAs that have gained the least in longevity since 2002 were those that started with the lowest life expectancy, located in northern urban areas with high levels of poverty and unemployment, and with relatively low education. Conversely, those MSOAs with higher life expectancies in 2002 had some of the largest gains. As a result, England has seen widening inequalities in longevity, with the life expectancy gap surpassing 20 years for women and 27 years for men.

5.3.1 Strengths and limitations

The main strength of the analysis is the presentation of high-resolution data for mortality and longevity across England over a period of substantial change in economic, health, and social care policy. By applying a hierarchical model based on patterns of mortality over age, space, and time, I obtained robust yearly estimates of mortality and life expectancy, together with the uncertainty in these estimates, for small areas. By contrast, studies that had not used a coherent model produced unstable (i.e., very large uncertainty) or implausible life expectancy estimates in some MSOAs, despite having aggregated deaths over 5 years, nor could they analyse trends at the

MSOA level (Office for National Statistics, 2015; Public Health England, 2021a). Comparison of estimates at MSOA and district level shows that the estimated MSOA life expectancy range was about 1.8 times the district-level range for women and 2.0 times the district-level range for men in 2019.

A limitation of the work in this chapter is that I did not break down age beyond 85 years, which might mask some differences in old-age mortality and survival patterns. Although MSOAs have small populations and are designed to have some socioeconomic homogeneity, there are inevitable variations in socioeconomic status and health within them. To understand life expectancy inequalities in relation to individual socioeconomic characteristics requires linking health and other data such as census records, education, and taxes, as done in countries like New Zealand and Sweden.

The people who live in each MSOA can change due to both within-country and international migration. There is some evidence that migrants tend to be healthier than those who do not move (Connolly et al., 2007). Regression of the change in life expectancy from 2002 to 2019 in each MSOA against population turnover, the proportion of households in each MSOA in 2019 who were different from those who had lived there in 2002 (van Dijk et al., 2021), was not able to explain the variation in life expectancy change for women ($R^2 < 0.001$) or for men ($R^2 = 0.01$) at the national level. Studies in both the UK (Connolly et al., 2007) and USA (Ezzati et al., 2008) have also shown that migration is not sufficient to explain the trends in health and health inequalities, and that these trends are largely due to real changes in population health. Even if rising inequalities are partly due to health-selective migration, this phenomenon has social and economic origins that should be addressed through employment opportunities, affordable housing, high-quality education, and health care.

Population and mortality statistics in the UK are generated independently from one another. As a result, we encountered a situation of having more deaths than population in a small percentage (0.001%) of age-MSOA-year combinations, a phenomenon that was more common in those aged 85 years and older. This finding might be due to errors in population estimates in years between censuses or because some people (e.g., those living in long-term care facilities such as care homes), are counted in one MSOA for the population statistic but have their death registered in another. Furthermore, care home residents might have relocated from other MSOAs, with different socioeconomic characteristics from that in which the care home is located. This factor

could attenuate the association between socioeconomic variables and life expectancy. The extent of this underestimation is modest; however, because a large part of life expectancy variation is due to deaths at earlier ages, when people are less likely to live and die in care homes (Bennett et al., 2018).

5.3.2 Comparison with previous literature

The life expectancy estimates in specific years are similar to the snapshots presented by the ONS and Public Health England (Office for National Statistics, 2015; Public Health England, 2021a), with correlation coefficients of 0.92-0.95 and mean differences of -0.004 to 0.19 years. However, these reports could not analyse trends because data were aggregated over 5 years (2009–13, 2013–17, or 2015–19). In terms of trends, studies that grouped small area units into deciles of deprivation have detected a decline in female life expectancy in the one or two most deprived deciles (Bennett et al., 2018; Marmot et al., 2020). Boulieri and Blangiardo (2020) modelled trends in life expectancy by district and decile of deprivation, using a space-time interaction term to detect local changes in the life expectancy trend. Notably, they detected lower life expectancy than expected for the district of Leeds in 2 years of the study period for women and 4 years for men. By analysing trends at the MSOA level, I could identify the specific communities in which longevity is declining and show that the decline, which began around 2010 in women in some MSOAs, has spread and accelerated since 2014.

Congdon (2019) also used spatial models to smooth over MSOA-level data of mortality, but specifically drug-related deaths and suicides between 2012-16. The author singled out the district of Blackpool as containing many of the MSOAs with the most extreme relative risks of death from these causes, including the MSOA I found had the lowest life expectancy for men in 2019.

5.3.3 Analogies with trends in the USA

Over the period of this analysis, from 2002 to 2019, national life expectancy increased in high-income countries in Australasia, Europe, and North America. Female life expectancy has stagnated or declined in various intervals since 2010 in the UK (84% of the UK population in 2019 lived in England) and in some other high-income countries including France, Germany, Italy, and the USA; the UK and USA have had some of the poorest performances in terms of the duration or extent of slowdown or reversal in longevity gain. The comparative performance of high-income

countries' longevity trends has been attributed to differences in risk factors such as smoking, health care, and social inequalities (Leon et al., 2019; National Research Council, 2010).

The declining life expectancy in numerous English MSOAs since 2014, especially those that already had a low life expectancy, resembles a trend spanning nearly three decades in the USA in two ways (Chetty et al., 2016; Dwyer-Lindgren et al., 2017a; Ezzati et al., 2008). First, in both countries, there is substantial variation in life expectancy at any level of poverty, which might be due to geographical variations in health behaviours, the public health programmes that influence these behaviours or otherwise prevent disease, and health services (Burn-Murdoch, 2023; Chetty et al., 2016). The second similarity in small area life expectancy trends is that the decline in life expectancy was more widespread in women than in men (Ezzati et al., 2008). Historically, women and men had similar life expectancies in high-income nations before a rise in traffic injuries and diseases associated with specific occupations and health behaviours such as smoking and alcohol use created a male mortality disadvantage in the 20th century (Beltrán-Sánchez et al., 2015). The closing of female and male life expectancy in the late 20th century and early 21st century in many high-income nations (Kontis et al., 2017) is partly due to the dynamics of smoking, which peaked later in women than in men, and affects causes of death such as respiratory diseases and lung cancer that have stagnated or even increased in women in deprived communities (Bennett et al., 2018; Leon et al., 2019). However, it is rare for the convergence of female and male life expectancies to occur in the form of female life expectancy decline (Ezzati et al., 2008), which might be due to a combination of the worsening economic, psychosocial (e.g., poverty, stress, and domestic violence), and behavioural (smoking and alcohol use) determinants of mortality in English women.

In both countries, the decline in life expectancy was associated with the economic trends of unemployment and insecure and low-wage employment following late 20th century deindustrialisation. In England, these economic trends led to a larger loss of jobs in the north than in London and the South East (Blundell et al., 2020; Davenport and Zarenko, 2020). These long-term changes were followed by a reduction in social support and welfare payments and in funding to the local governments during the austerity period, which increased poverty, including in-work poverty (Alexiou et al., 2021; Blundell et al., 2020), and also had larger effects in the north than in London and southern parts of the country and worsened the effects of loss of secure employment (Alexiou et

al., 2021; Davenport and Zarenko, 2020; Gray and Barford, 2018). Poverty and reduced funding to services increase mortality through health behaviours such as smoking and alcohol use, poor nutrition and living environment, psychosocial pathways, and lower provision or use of preventive and curative health care.

5.4 Summary

I performed a high-resolution spatiotemporal analysis of all deaths in England from 2002 to 2019, using a Bayesian hierarchical model to obtain estimates of age-specific death rates by age, sex, and MSOA. I used life table methods to calculate life expectancy at birth and probabilities of death in different ages by sex and MSOA.

In 2002–06 and 2006–10, all but a few (0–1%) MSOAs had a life expectancy increase for women and men. In 2010–14, female life expectancy decreased in 351 (5.2%) of 6791 MSOAs. And by 2014–19, the number of MSOAs with declining life expectancy was 1270 (18.7%) for women and 784 (11.5%) for men. The life expectancy increase from 2002 to 2019 was smaller in MSOAs where life expectancy had been lower in 2002 (mostly northern urban MSOAs), and larger in MSOAs where life expectancy had been higher in 2002 (mostly MSOAs in and around London).

These results show that numerous communities in England had begun to have a decline in longevity, mirroring an earlier trend in the USA. That these trends happened in the decade before the Covid-19 pandemic – so-called “normal times” – is worrying, and signals ongoing policy failures to tackle the social determinants of health.

Chapter 6

Life expectancy inequality in London at the LSOA level

This chapter is based on the peer-reviewed publication *Changes in life expectancy and house prices in London from 2002 to 2019: hyper-resolution spatiotemporal analysis of death registration and real estate data*, published in *The Lancet Regional Health Europe* (Bennett et al., 2023), for which I was joint first author. The paper was published under the CC-BY license, which permits reproduction provided the original work is properly cited. The paper combined two main datasets: estimates of life expectancy in London at LSOA level and estimates of house prices at the OA level. Unlike the original text, I will largely focus on the life expectancy estimates, for which I was responsible, rather than the house price estimates produced by James E Bennett.

Although the models in the previous chapter took several days of computing time, runtimes will only decrease as computing power increases. And in theory, the models can be scaled to higher and higher spatial resolutions as both hardware and inference algorithms improve. From a computational perspective, we could potentially estimate mortality for all LSOAs, OAs, or even postcodes¹ in England. In this chapter, I test this idea by modelling life expectancy at the finer LSOA level for a single region in England, its capital city London, which has a number of LSOAs of the same order of magnitude as the number of MSOAs in England.

¹Unlike OAs, postcodes do not align with other administrative boundaries, nor are they designed to contain similar numbers of people, which may lead to small number issues when modelling mortality.

6.1 Methods

6.1.1 Mortality and population data

I performed a high-resolution spatiotemporal analysis of all deaths in London from 2002 to 2019 (909,097 death records, extract date: 25th June 2021), as detailed in Chapter 3. Deaths were stratified by 19 age groups (0, 1–4, 5–9, ..., 80–84, 85+) and 4835 LSOAs. These LSOAs fit inside 33 districts: London's 32 boroughs and the City of London. I did not use 16 death records (0.002%) for which sex was not recorded. In 3267 (0.099%) age-LSOA-year combinations, the number of deaths exceeded the population. In 88% of these combinations, the deaths exceeded population by only one or two, and 96% of these combinations were in people aged 80 years and older. In these cases, the population was set equal to the number of deaths.

6.1.2 Statistical analysis

The model was largely as outlined in Chapter 4, with a few changes: First, the negative binomial likelihood from Equation 4.2 was replaced with a beta-binomial likelihood,

$$\text{deaths}_{ast} \sim \text{Beta-Binomial}(m_{ast}\rho, (1 - m_{ast})\rho, N_{ast}), \quad (6.1)$$

where m_{ast} is the death rate, N_{ast} is the population and $\rho \geq 0$ is the overdispersion parameter. The beta-binomial distribution has mean $m_{ast}N_{ast}$ and variance $N_{ast} \frac{\rho + N_{ast}}{\rho + 1} m_{ast}(1 - m_{ast})$. I found that the variability of the LSOA-level mortality data was such that death rates did near 1, violating the assumption for Equation 4.2 that mortality is low. In fact, when I tested a negative binomial or Poisson likelihood, I found the death rates for some age-LSOA-year combinations exceeded 1, which of course is impossible. The beta-binomial likelihood is a generalisation of the binomial distribution that allows for overdispersion. The death rates were logit-transformed and the model structure is otherwise the same as Equation 4.4, but using nested hierarchical random effects with LSOAs nested in MSOAs, which were nested in districts. A $\rho \sim \exp(0.1)$ hyperprior was used for the overdispersion term, and $\sigma \sim \mathcal{N}^+(1)$ hyperpriors for the standard deviation parameters, as recommended by the Stan development team (Stan development team, 2020).

Table B.2 shows all model parameters, their priors and dimensions.

Again, I used NIMBLE for MCMC (de Valpine et al., 2022, 2017), with four chains of 80,000 iterations, discarding the first 30,000, thinning the remainder by 200 and checking for convergence using trace plots and $\widehat{R}$ (Vehtari et al., 2021).

The local effect of the Grenfell Tower fire was even more pronounced at the LSOA level. In 2017, the LSOA in Kensington and Chelsea where Grenfell Tower is located had 75 deaths, compared to less than 15 in 2016 and 2018. The additional deaths were caused by a fire in a highrise residential building. I used the same treatment as in Chapter 5 to remove the outlier from the modelling stage and add the excess mortality in at a later stage before calculating life expectancies. I defined life expectancy inequality as the difference between 2.5th and 97.5th percentiles of LSOA life expectancies.

6.2 Results

6.2.1 Inequalities in life expectancy

Life expectancy at birth for London increased from 80.9 (95% CrI 80.8–81.0) years in 2002 to 85.4 (85.3–85.5) years in 2019 for women, and from 76.1 (76.0–76.2) years to 81.6 (81.5–81.7) years for men. Life expectancy inequality (difference between 2.5th and 97.5th percentiles) was 19.1 (18.4–19.7) years for women and 17.2 (16.7–17.8) years for men, in 2019 which is a substantial increase from 11.1 (10.7–11.5) years for women and 11.6 (11.3–12.0) years for men in 2002. The corresponding estimates for life expectancy inequality in 2019 calculated using MSA level analysis in Chapter 5 were 10.8 (10.3–11.3) years for women and 11.0 (10.5–11.6) years for men.

Life expectancy in 2019 was highest in LSOAs in central London districts of Kensington and Chelsea, Westminster, City of London and Camden, in the southwest (Richmond upon Thames and Kingston upon Thames) and parts of the northwest (e.g., parts of Harrow and Barnet), with life expectancy in many LSOAs surpassing 90 years (Figure 6.1). Low life expectancy was spread in LSOAs throughout the city but was more common in outer east and southeast London.

In general, as life expectancy increased, the credible interval of life expectancy increased (Figure 6.2). The credible interval was particularly wide for some LSOAs with the highest female life expectancy estimates, with some credible intervals exceeding 30 years.

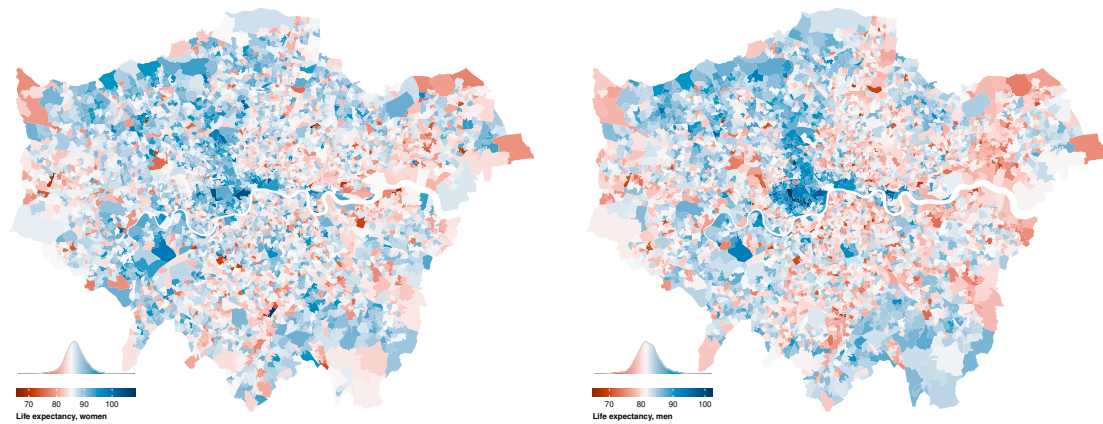


Figure 6.1: Map of life expectancy and the distribution of life expectancy in London in 2019. The areas in white have a life expectancy equal to the overall London life expectancy.

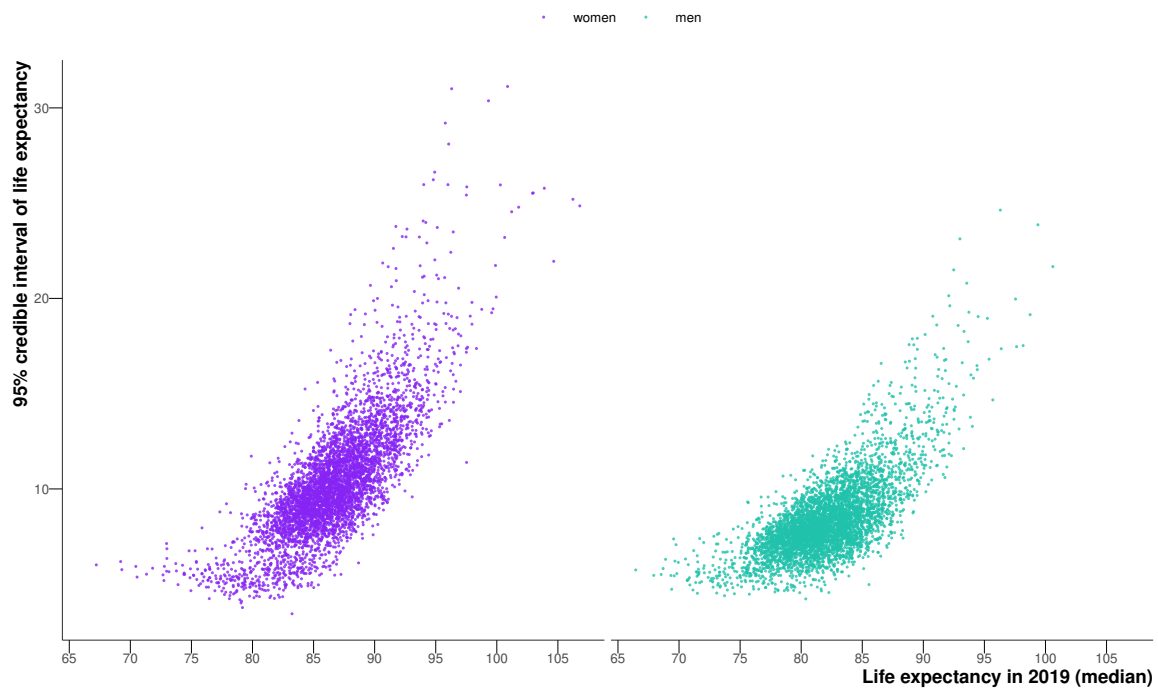


Figure 6.2: Median estimates of LSOA-level life expectancy in London in 2019 against the 95% credible intervals.

6.2.2 Change in life expectancy and inequality

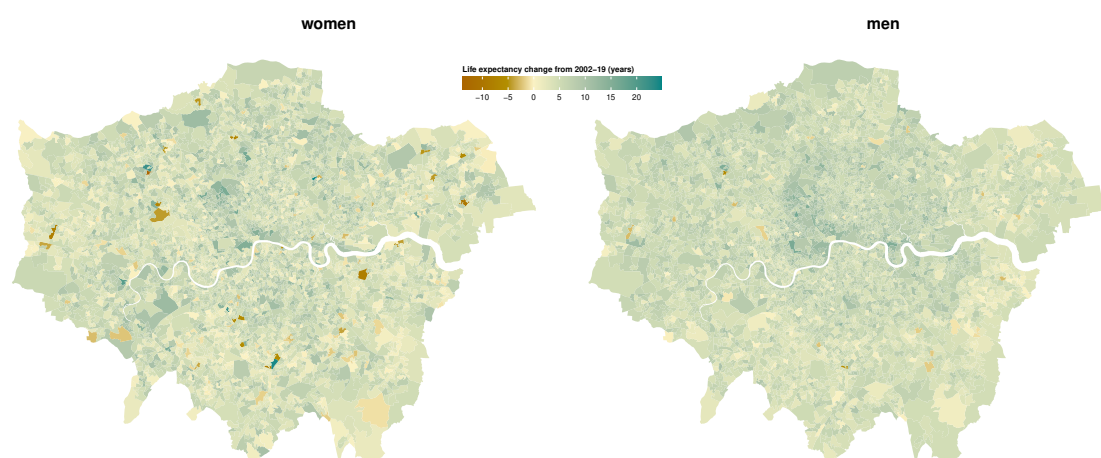


Figure 6.3: Geography of change in life expectancy in London from 2002 to 2019.

There was large variation across London in how much life expectancy increased from 2002 to 2019, ranging from <2 years in 537 (11.1%) LSOAs for women and 214 (4.4%) for men to >10 years in 220 (4.6%) for women and 211 (4.4%) for men (Figure 6.3). In 134 (2.8%) LSOAs for women and 32 (0.7%) for men, life expectancy may have declined, with a posterior probability of a decline >0.80 in 41 (0.8%) and 14 (0.3%) LSOAs for women and men, respectively. Life expectancy increased more in LSOAs in central and inner east and south London than in outer areas.

The rate of increase in life expectancy was smaller after 2010 compared to the earlier years for both sexes (Figure 6.4), and the number of LSOAs in which life expectancy may have fallen was larger, especially from 2010 to 2014. The slow-down or reversal of LSOA life expectancy after 2010 is consistent with the pattern seen throughout England's MSAs in Chapter 5, which may be partly due to increase in poverty due to low-wage employment and cuts in services in the austerity era. However, the progressive worsening of life expectancy gain and loss from 2010–2014 to 2014–2019 which was seen in other parts of England did not happen in London (Figure 6.4).

There was substantial variation in the size of life expectancy increase over short distances. As a result of this spatial heterogeneity, life expectancy inequality increased not only in London as a

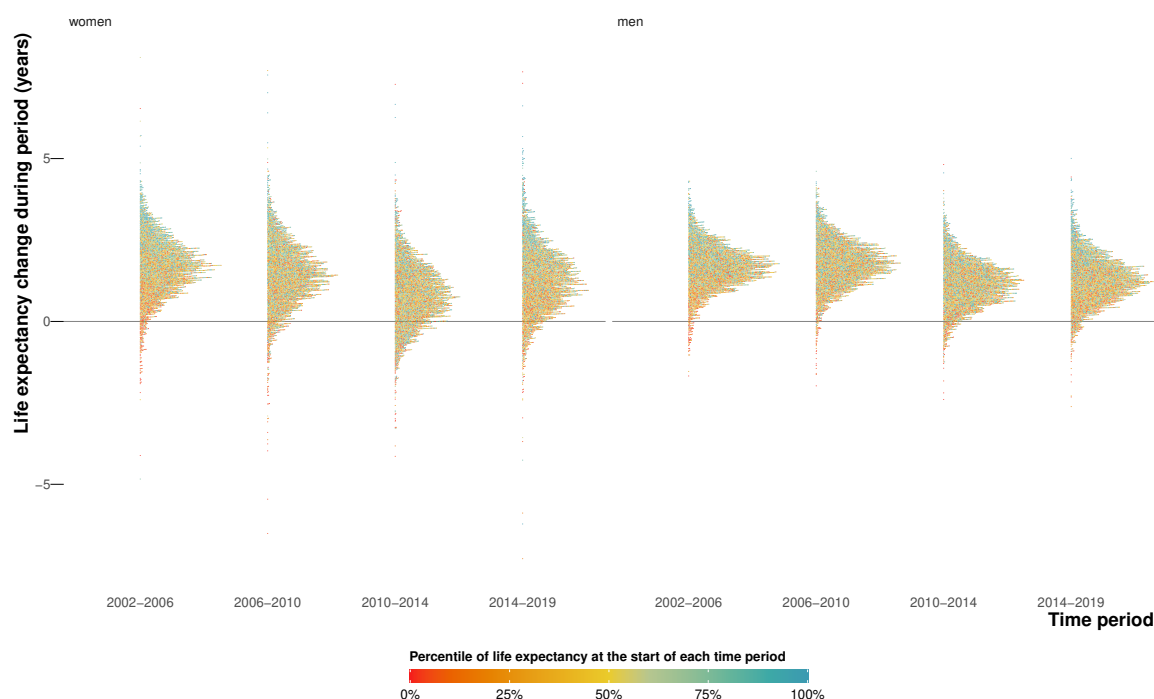


Figure 6.4: Change in LSOA life expectancy in different time periods, 2002–19. Each point shows the posterior median change in one LSOA. LSOAs are coloured by their life expectancy at the beginning of each period.

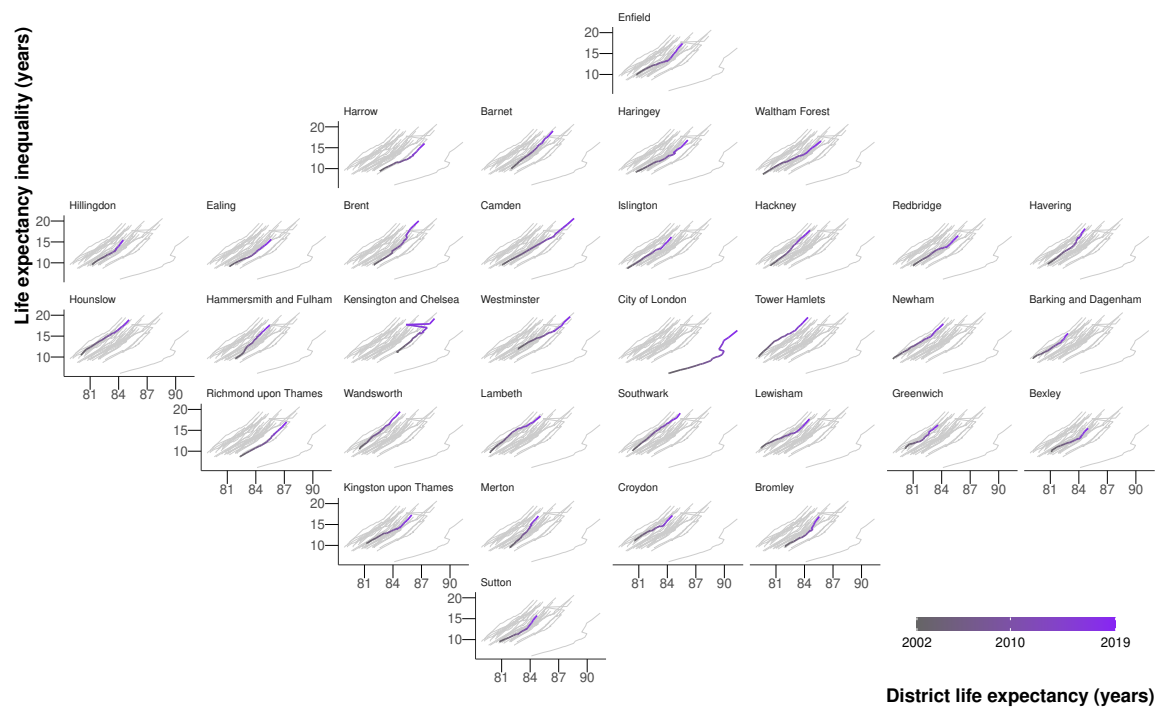
whole, but also in every district in London alongside increasing average life expectancy (Figure 6.5 and Figure 6.6).

By 2019, female and male life expectancies had a 2.5th–97.5th percentile range that was >12 years in every district, and >20 years in three districts for men and two for women. In 2002, within district inequality had been >12 years in only two districts for men and none for women. At the same time, the difference between districts with the highest and lowest life expectancy increased from 5.9 (5.2–7.7) years in 2002 to 8.5 (5.9–11.9) years in 2019 for females, and from 6.8 (5.9–8.7) years to 10.9 (8.2–14.2) years for males indicating that the rise in life expectancy inequality took place both within and between districts.

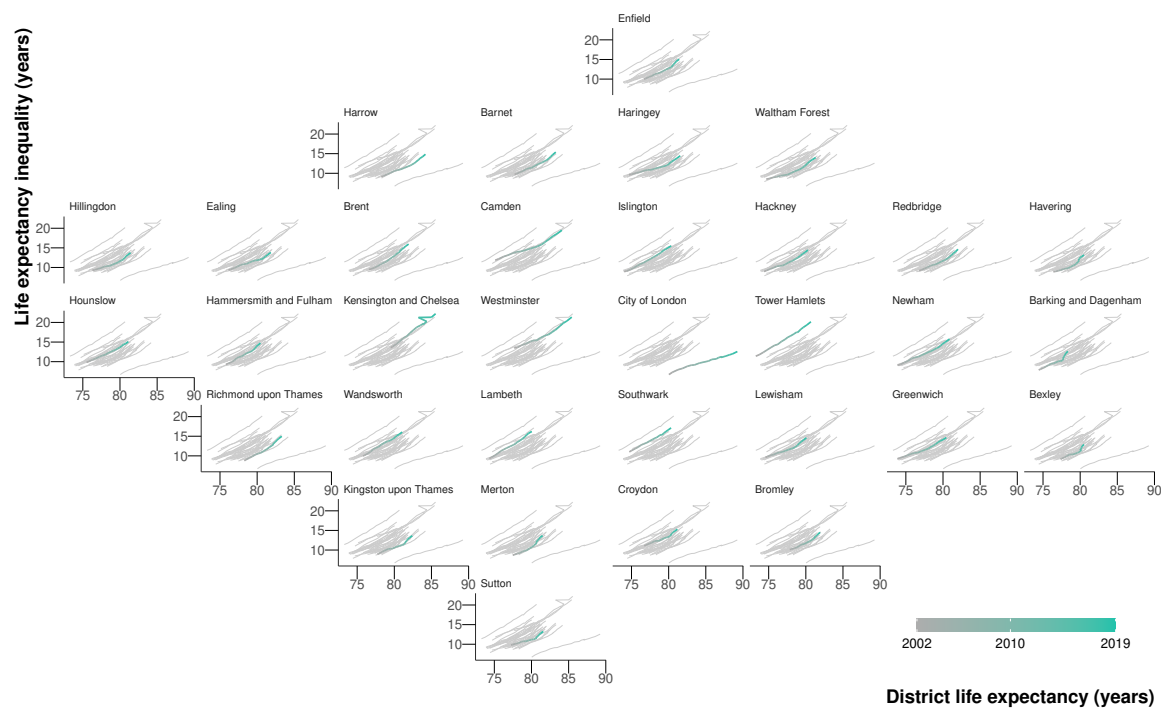
6.3 Discussion

This high-resolution analysis revealed that in London, a major global metropolis, not only has life expectancy inequality increased substantially, but this increase has happened both between districts and within every district.

Life expectancy in London has previously only been reported at MSOA level, as discussed in



(a) Women



(b) Men

Figure 6.5: Life expectancy at birth and inequality in life expectancy within London's 33 local authority districts from 2002 to 2019. Each panel plots life expectancy against life expectancy inequality over time for all 33 districts, and highlights the trend line for the district which is named at the top of the panel. The districts are arranged according to their approximate location in Greater London. The large decline in life expectancy in Kensington and Chelsea in 2017 is due to the deaths caused by the fire in Grenfell Tower.



Figure 6.6: Distribution (density plot) in 33 London districts of median estimates of LSOA life expectancy at birth for 2002 and 2019, and of the change from 2002 to 2019. Districts are ordered by the median life expectancy in 2002.

Chapter 5. When aggregated to MSOA level, the LSOA life expectancies in this work were correlated with MSOA estimates in Chapter 5 (correlation coefficient > 0.99 and mean difference ≤ 0.01 years for both sexes).

In the main paper (Bennett et al., 2023), we also looked at how life expectancy improvements varied with house prices and their changes throughout the city, and whether population change may partly mediate this association. We found the change in life expectancy had different associations with price change in more expensive versus lower cost communities: where prices were high at the turn of the century, life expectancy increased substantially independently of price change, and with little change in the resident population. In contrast, in areas that started with lower prices, life expectancy only increased proportional to price change, and was accompanied by a greater influx of new residents into the local area. This change in residents saw a greater proportion of the population in working ages, and socioeconomically better off, and an outflux of residents in retirement ages and to a lesser extent children, adolescents and young adults. This implies that a portion of the change in life expectancy may be due to change in the resident population.

6.3.1 Strengths and limitations

The main strength of the study is the presentation of very high-resolution data on life expectancy and house prices in consistent spatial and temporal units in a major global city over a period of substantial policy interest and investment in urban renewal. Carrying out the study at the LSOA level uncovered inequalities to a fuller extent.

I defined life expectancy inequality as the difference between 2.5th and 97.5th percentiles of LSOA life expectancies rather than the difference between the maximum and minimum as in Chapter 5. This was because the LSOAs with extremely low life expectancies tended to contain age-LSOA--year combinations in which the number of deaths exceeded the population. There are also some credible intervals of life expectancy in excess of 30 years for women, which questions the epidemiological plausibility of these estimates. Small-area population, which is the denominator of age-specific death rates, is estimated by the ONS for intercensal years, and may be subject to error. This is especially the case in older ages when some people live and die in a long-term care facility, and may be counted towards population (denominator of death rates) in their original

LSOA of residence and towards deaths (numerator of death rates) in the LSOA where the care facility is located. There were a higher proportion (0.099% compared to 0.001% in Chapter 5) of spatial units with this issue at the LSOA level. In 2019, 11 LSOAs for women and one LSOA for men had life expectancies in which the median estimate was greater than 100 years. Given the MSOAs with the highest life expectancies for each sex in Chapter 5 were both in London and exceeded 95 years, these estimates may be plausible for some LSOAs, which contain around a fifth of the population of MSOAs. However, upon inspection, four out of the five LSOAs with the largest increases in life expectancy, which were also LSOAs with life expectancies above 100, were adjacent to LSOAs with very large decreases. And in four out of five LSOAs with the largest estimated decreases in life expectancy, estimated population counts were lower than the recorded death counts in the older age groups in several years of the data, implying that populations or deaths may have been allocated incorrectly. Subsequent inspection showed that postcodes coinciding with, or adjacent to, LSOAs with the largest increases and decreases contained care homes, which can lead to the place of residence or death being inconsistently recorded for some of their residents.

Mortality and house prices were estimated using two separate models because a joint model would be extremely complex and would require additional assumptions about how house prices are associated with age- and LSOA-specific mortality. It was only possible to indirectly evaluate, through LSOA-level population turnover and population characteristics, whether population change is a potential mechanism for the observed change in life expectancy because routine death registration in England only records place of residence at the time of death. Further, there are currently limited time-series data on quality of housing, access to jobs, services and amenities, and other home and neighbourhood characteristics that affect health. To understand whether the change in the life expectancy of a community arises from a change in the health of the population, itself due to changes in their economic status and/or local environment and amenities, versus a change in the resident population, requires linked datasets which are able to track over time environmental characteristics of areas, together with individuals' place of residence, socioeconomic status and mortality records.

6.3.2 Comparison with previous literature

The inequalities of 19.1 (18.4–19.7) years for women in 2019 and 17.2 (16.7–17.8) years for men in London in 2019 are some of the largest for any city in the literature and the largest for women. Further, some of the inequality in the tails of the life expectancy distribution is masked by our choice of inequality measure, the 2.5th–97.5th percentile difference, rather than the full range across all LSOAs.

In a study of the 397 census tracts in King County, Washington, a county ranking in the top 5% of the US in terms of both income and life expectancy, Dwyer-Lindgren et al. (2017b) found an 18.3 year gap in life expectancy for men between the top and bottom tracts in 2014, and a 14.8 gap for women. For Vancouver, Canada, Yu et al. (2021) found a difference in 2016 of 9.5 years for men and 8.3 years for women between the 90th and 10th percentiles of 368 census tracts. Using the same measure of inequality as the Vancouver study, Jonker et al. (2012) uncovered a 8.2 year and 9.2 year gap in life expectancy for women and men for 89 small areas in Rotterdam. Across six large cities in Latin America, Bilal et al. (2019) observed the largest difference between the top and bottom decile of life expectancy at birth was 17.7 years for women in Santiago, Chile.

To some extent, as this is one of the highest-resolution studies in the literature, it is natural that we would find large inequalities. As we go to finer and finer resolutions, we will always uncover more inequality, until we reach the physical limit of the individual and the inequality is between the earliest neonatal death and the oldest person to die in the city.

In terms of work specific to London, beyond that at the MSOA level in Chapter 5, Congdon (2014) considered both ill-health and mortality for 625 wards in London, finding more than a two-fold variation in the percent of life spent in disability for men, but did not mention the absolute inequality of life expectancy in the city. Some studies have documented changes in spatial patterns of poverty in London (Bosetti, 2015; Travers et al., 2016), and have found that poverty rates have fallen in inner East London but may have risen in some outer London areas, with corresponding changes in the share of the population working in higher professional occupations. No previous study has reported how life expectancy change in small-area units coincided with change in house prices.

6.3.3 Urban policy and public health implications

Since the turn of the millennium, London's population and economy have grown substantially. The economic growth has been highly polarised with high-pay and high-skilled employment alongside low-pay and insecure jobs (Overman and Xu, 2022). As a result, despite city-wide growth in income, nearly one half of London's population fall in the bottom two quintiles of national income deprivation (Ministry of Housing, Communities & Local Government, 2019). Together with an uncontrolled property market, this has created house prices that are unfavourable to low-income families, displacing entire subgroups of the population to cheaper parts of the city, with fewer amenities and worse access to jobs, quality education, healthcare and other services (Bosetti, 2015). Many are unable to purchase (versus rent) homes and hence spend an increasing share of income on housing, and/or live in lower quality or smaller accommodation in the more desirable districts (Bosetti, 2015; Travers et al., 2016). These trends may have contributed to health inequalities both between nearby areas and across the entire city, alongside other trends such as differences in the extent of neighbourhood improvement, or provisions of health and social care services as council budgets were reduced as a result of austerity policies.

The evolution of London, and major cities in other high-income nations and emerging economies, into places where only the well-off can afford to own properties, where the balance of the city is driven by the cost of property and wealth dominates access, poses a gloomy, non-cohesive future for these cities.

6.4 Summary

I performed a hyper-resolution spatiotemporal analysis of all deaths in London from 2002 to 2019, using a Bayesian hierarchical model to obtain estimates of age-specific death rates by age, sex, and LSOA. I used life table methods to calculate life expectancy at birth by sex and LSOA.

In 134 (2.8%) of London's LSOAs for women and 32 (0.7%) for men, life expectancy may have declined from 2002 to 2019, with a posterior probability of a decline >0.80 in 41 (0.8%, women) and 14 (0.3%, men) LSOAs. The life expectancy increase in other LSOAs ranged from <2 years in 537 (11.1%) LSOAs for women and 214 (4.4%) for men to >10 years in 220 (4.6%) for women and 211 (4.4%) for men. The 2.5th-97.5th percentile life expectancy difference across

LSOAs increased from 11.1 (10.7–11.5) years in 2002 to 19.1 (18.4–19.7) years for women in 2019, and from 11.6 (11.3–12.0) years to 17.2 (16.7–17.8) years for men.

Although the analysis of life expectancy at the LSOA level is possible, there are issues with the population data for older age groups that limit its utility, particularly concerning the tails of the distribution of life expectancy.

Chapter 7

Cause-specific mortality at the district level

Where previous studies found female life expectancy in the bottom deciles of deprivation declined (Bennett et al., 2018; Marmot et al., 2020), in Chapter 5, I showed there were widespread and accelerating declines, with 1 in 5 small areas for women and 1 in 9 for men observing a fall in life expectancy in 2014-19. However, the lack of available covariates at the MSOA level makes it difficult to suggest which social and environmental effects are driving life expectancy downwards. By analysing mortality for specific causes of death, which is the second objective of the thesis and which I address in this chapter, we are able to suggest pathways that shape mortality profiles.

7.1 Methods

7.1.1 Mortality and population data

I performed a spatiotemporal analysis of vital registration data in which I extracted de-identified data for all deaths in England from 2002 to 2019 (8,648,191 death records, extract date: 17th November 2022). England was divided into 314 districts according to the boundaries in 2020. Deaths were assigned into the following age groups: 0, 1-4, 5-9, ..., 80-84, 85+ years. I did not use 130 death records (<0.001%) for which sex was not recorded. There were no age-district-year combinations in which the number of deaths from all causes exceeded the population.

7.1.2 Grouping causes of death

Each death record in the database contains a series of ICD-10 codes from the death certificate, which have been signed off by a medical professional, containing firstly the codes corresponding to the conditions leading directly to death, and then any other codes that contributed to the death. Based on this list of codes, the ONS assigns each death record in post-neonatal ages an ICD-10 code corresponding to the underlying cause of death (the disease or injury that initiated the train of events directly leading to death) using standardised computer-based selection algorithms (Office for National Statistics, 2022b). For neonates, which are not assigned an underlying cause of death, I used the ICD-10 code in the first position on the death record.

I used ICD-10 code for the underlying cause of death to assign each death to 136 cause groups of the WHO Global Health Estimates (GHE) study (World Health Organization, 2020); these groups encompass causes of death with related aetiology and clinical and public health relevance. I also grouped diabetes mellitus and nephritis and nephrosis (hereafter referred to as *diabetes*) as these deaths might have overlapping history. I used the top twelve causes of death for each sex according to the total number of deaths from 2002 to 2019 for cause-specific analysis, as well as a residual groups comprising deaths from all other cancers, all other non-communicable diseases (NCDs), all other cardiovascular diseases (CVDs), all other infections, maternal, perinatal and nutritional conditions (IMPEN), and injuries (external causes). Together, these form a mutually exclusive, collectively exhaustive list of causes of death (Figure 7.1).

The full list of ICD-10 codes for each cause group can be found in Table D.1. Any death records with an ill-defined disease (garbage code) as the underlying final cause of death were randomly assigned to one of the residual disease groups (all other cancers, all other CVDs, all other NCDs, all other IMPEN) with probability proportional to the existing age-sex-year totals in those disease groups.

There were 4,465,948 (51.6%) deaths in women and 4,182,108 (48.4%) in men in England from 2002 to 2019. 3,972,286 (45.9%) of these deaths occurred in people aged younger than 80 years of age.

Of the wider cause groups (CVDs; NCDs; cancers; maternal, perinatal, nutritional and infectious causes; injuries), CVDs had the largest proportion of deaths at 2,712,895 (31.4%) from 2002 to

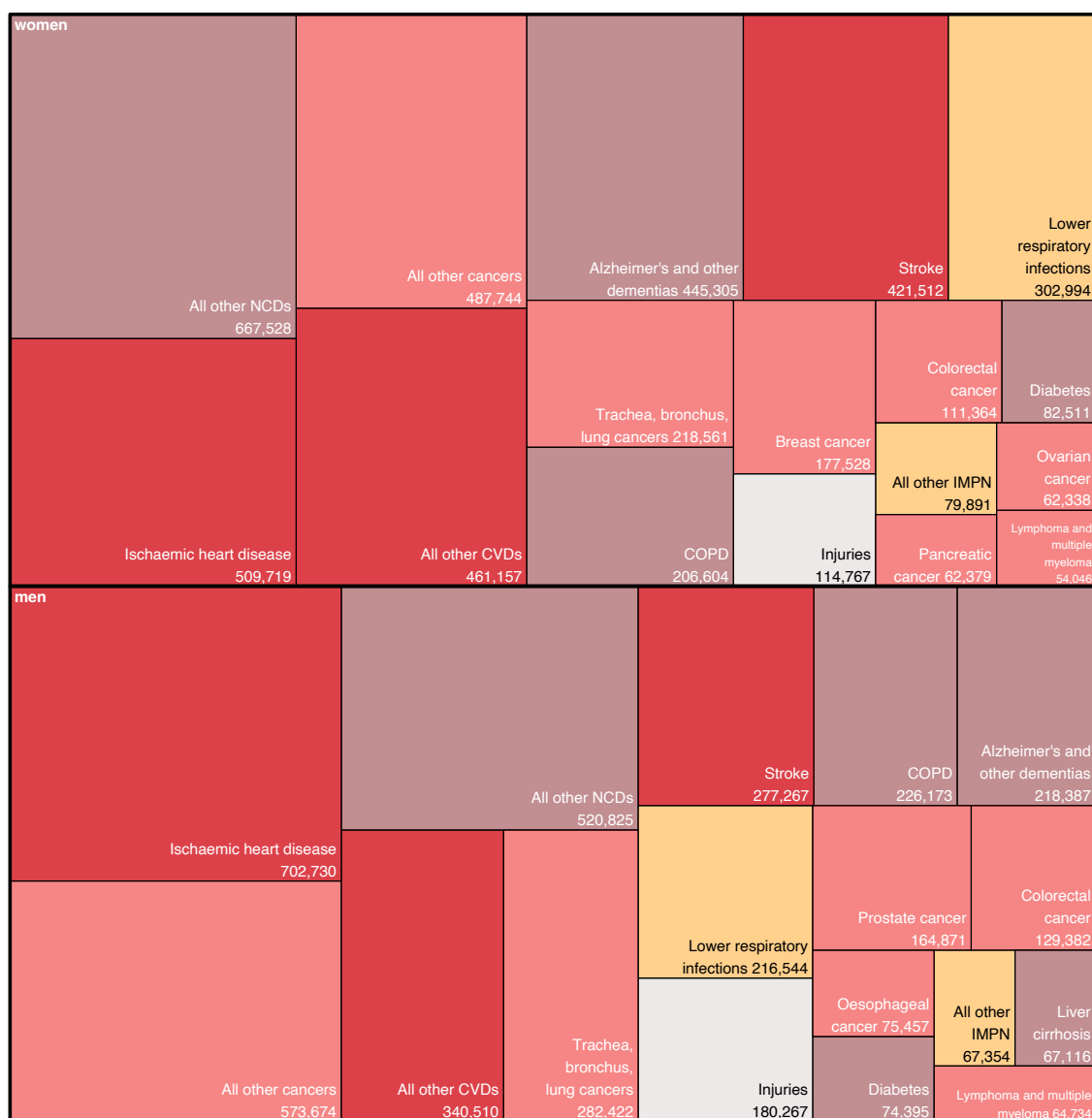


Figure 7.1: Total number of deaths for the twelve leading causes of death in England from 2002 to 2019, and the residual groups all other cancers, all other NCDs, all other CVDs, all other infections, maternal, perinatal and nutritional conditions (IMPN), and injuries. The boxes are coloured by the wider groups of CVDs; NCDs; cancers; maternal, perinatal, nutritional and infectious causes; injuries. See Table D.1 for ICD-10 codes for each category.

2019 (Figure 7.1). NCDs (excluding CVDs and cancers) were responsible for 2,508,844 (29.0%) deaths from 2002 to 2019. Cancers accounted for 28.5% of all deaths, nearly the same total mortality burden as NCDs. There were far fewer deaths from IMPN causes (666,783 (7.7%)) and injuries (295,034 (3.4%)), which are equivalent to GBD Groups 1 and 3, respectively.

7.1.3 Statistical analysis

Although the size of the spatial unit of analysis has been increased to the district-level, by stratifying by cause of death, the number of deaths in each age-district-year-cause stratum is still very small. As outlined Chapter 4, I used a Bayesian hierarchical model to obtain stable estimates of death rates. I conducted all analyses separately by sex and cause group.

Unlike the models in previous chapters, I used a binomial likelihood:

$$\text{deaths}_{ast} \sim \text{Binomial}(m_{ast}, \text{Population}_{ast}). \quad (7.1)$$

A beta-binomial was trialled but did not improve model fit. The model is the same as model Chapter 4, but without a random walk for each spatial unit (ν_{st}), which was tested but did not improve the fit despite adding a lot of complexity.

Spatial effects were modelled using ICAR priors rather than the nested hierarchy of random effects seen in Chapter 5 and Chapter 6. The nested hierarchy allowed units that fell in different districts to differ more than those within the same district, reflecting the relevance of district as a unit of resource allocation and policy. However, regions, the level above districts in the hierarchy, have no such policy implications and hence we implemented a fully spatial model.

The ICAR model requires all spatial units to have neighbours. Thus, the districts containing the Isle of Wight and the Isles of Scilly were each joined to the nearest mainland district based on road or ferry connections.

All model parameters, their priors and dimensions are in Table B.3.

I fitted the model in the probabilistic programming language NumPyro (Phan et al., 2019) and obtained 1000 draws from the posterior distribution of model parameters that were used to calculate age-specific death rates. Where possible, non-centred parametrisations were used in

the model to improve the efficiency of the NUTS samplers. I ran four chains with a warmup of 2000 iterations and then 10,000 further iterations and thinned the remainder by 40 to obtain 1000 post-burn-in draws from the posterior distribution.

The primary reporting outcome is the unconditional probability of dying between birth and 80 years of age. Unconditional probability is the probability of death in the absence of competing causes of death, and is calculated using life table methods from age-specific death rates from the cause of interest (Preston et al., 2001). I limited the age range to 80 years because the probability of death in the absence of competing causes equals 100% when the entire life course is considered. The reported 95% credible intervals represent the 2.5th and 97.5th percentiles of the posterior distribution of estimated probabilities of death.

In order to assess the effect of differences in cause-specific mortality between districts on inequalities in total mortality, I decomposed life expectancy differences into contributions from specific causes of death. This required estimates of life expectancy at the district level, for which I ran a model for a combined group of all deaths. Death rates for this total mortality group, and for injuries, were corrected in 2017 in Kensington and Chelsea for the Grenfell fire by imputing the mean deaths from injuries between 2016 and 2018 before running the model.

I calculated the contributions of deaths from each cause of death, in each age group, to both the life expectancy inequality between each district and the district with the highest life expectancy, and to the life expectancy change for each district between different time periods. I used Arriaga's method, which is widely used to decompose life expectancy differences between populations or population subgroups (Arriaga, 1984). Arriaga's method uses quantities generated from the life tables of two populations to calculate how much each age group contributes to the life expectancy difference, ${}_n\Delta_x$, by looking at differences in the number of survivors in the hypothetical cohort of the life table at different ages. It then partitions the age-specific contributions to the life expectancy gap by cause of death, ${}_n\Delta_x^i$, in proportion to the difference in cause-specific death rates between the subgroups:

$${}_n\Delta_x^i = {}_n\Delta_x \cdot \frac{{}_nm_x^i(2) - {}_nm_x^i(1)}{{}_nm_x(2) - {}_nm_x(1)}. \quad (7.2)$$

where ${}_n m_x^i$ are age-specific death rates for cause i . Arriaga showed the sum of the age- and cause-specific contributions are equal to the difference in life expectancy,

$$e_0(2) - e_0(1) = \sum_x n \Delta_x = \sum_i n \Delta^i = \sum_x \sum_i n \Delta_x^i. \quad (7.3)$$

So that the cause-specific contributions summed exactly to the differences in life expectancy as in Equation 7.3, the cause-specific death rates, which were estimated in separate model runs, were scaled such that the sum over all causes was equal to the estimate for all-cause death rate, which itself was estimated in a separate run. For this analysis, I used the sample mean death rate in each age-district-year-cause combination.

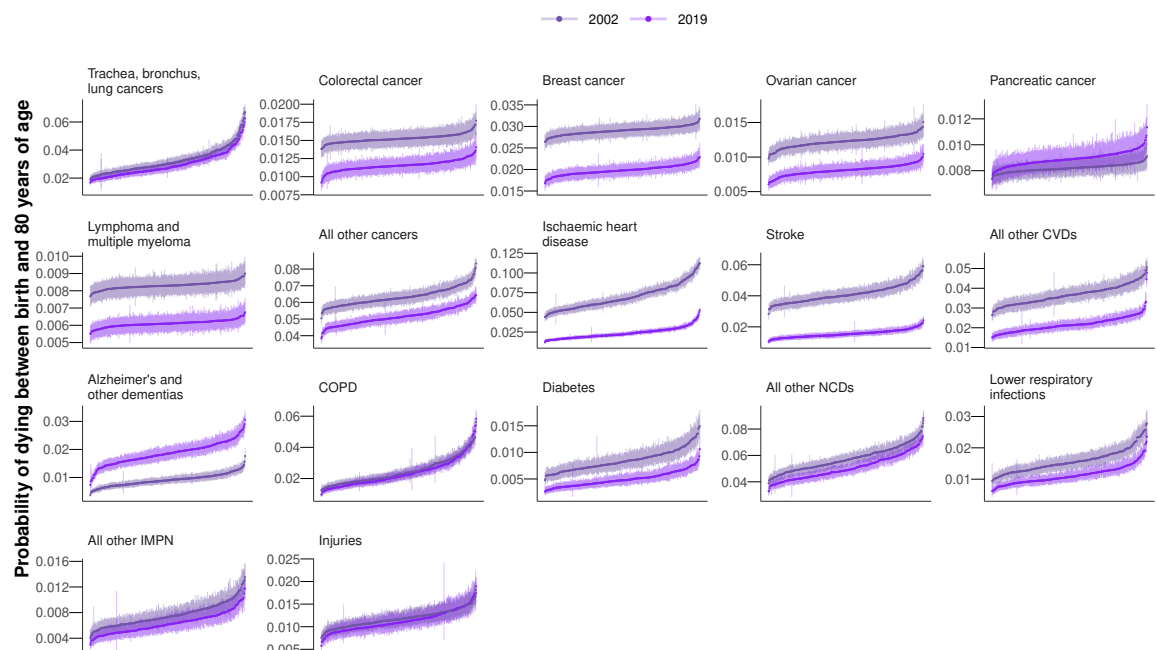
Further details on the calculations for life expectancy, the probability of dying, and Arriaga's method can be found in Appendix A.

7.2 Results

7.2.1 Variation in space

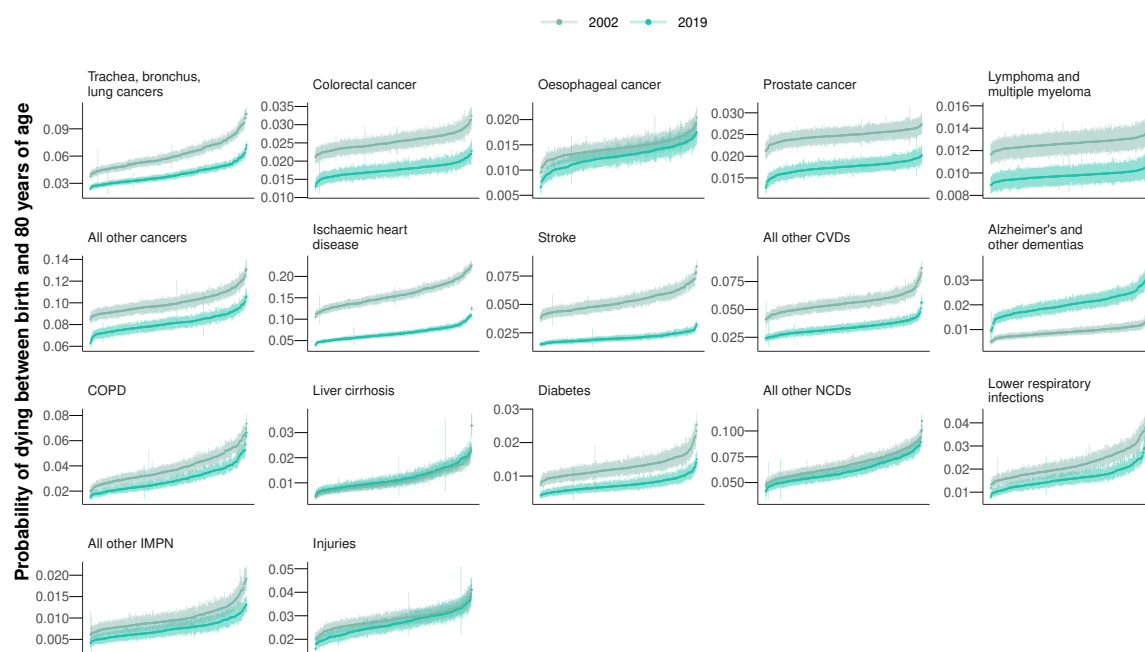
The highest probability of dying between birth and 80 years of age in 2019 for any sex-cause-district combination was for ischaemic heart disease for men in Manchester at 0.13 (0.12-0.13) (Figure 7.2 and Figure E.1). For women, all other NCDs had the highest mortality for any disease, with the highest probability of dying of 0.09 (0.08-0.09) in Blackpool (Figure 7.2 and Figure E.8).

The cause of death with the largest inequality in mortality was COPD for women, with a 6.0-fold variation between the districts with the highest and lowest probabilities of dying in 2019. COPD was also one of the most unequal causes of death for men (4.4-fold), but was exceeded in inequality by lower respiratory infections (4.8-fold) and liver cirrhosis (6.7-fold), which was in the top twelve leading causes of death for men but not for women. For women, there was large inequality in the probability of dying from the two leading causes of death by total number of deaths, Alzheimer's and other dementias (4.1-fold; referred to as *dementias* hereafter) and ischaemic heart disease (4.4-fold), although there was less variation in mortality for these causes in men (3.6-fold and 3.1-fold). All of these causes presented similar geographies, with the highest probabilities of dying clustered in urban areas, particularly in the North West (figure in Appendix E).



Ranking of probability of dying

(a) Women



Ranking of probability of dying

(b) Men

Figure 7.2: Ranked probability of dying between birth and 80 years of age in 314 local authority districts in England in 2002 and 2019 for the twelve leading causes of death and residual groups. Each point shows one district and the vertical line going through the point its 95% credible interval.

With the exception of trachea, bronchus and lung cancers (referred to hereafter as *lung cancer*; 3.7-fold variation for women and 3.1-fold for men), cancers showed the least geographical inequality, all with less than 1.7-fold variation in the probability of dying. For both sexes, the least unequal cause of death was lymphoma and multiple myeloma (1.2-fold for both sexes).

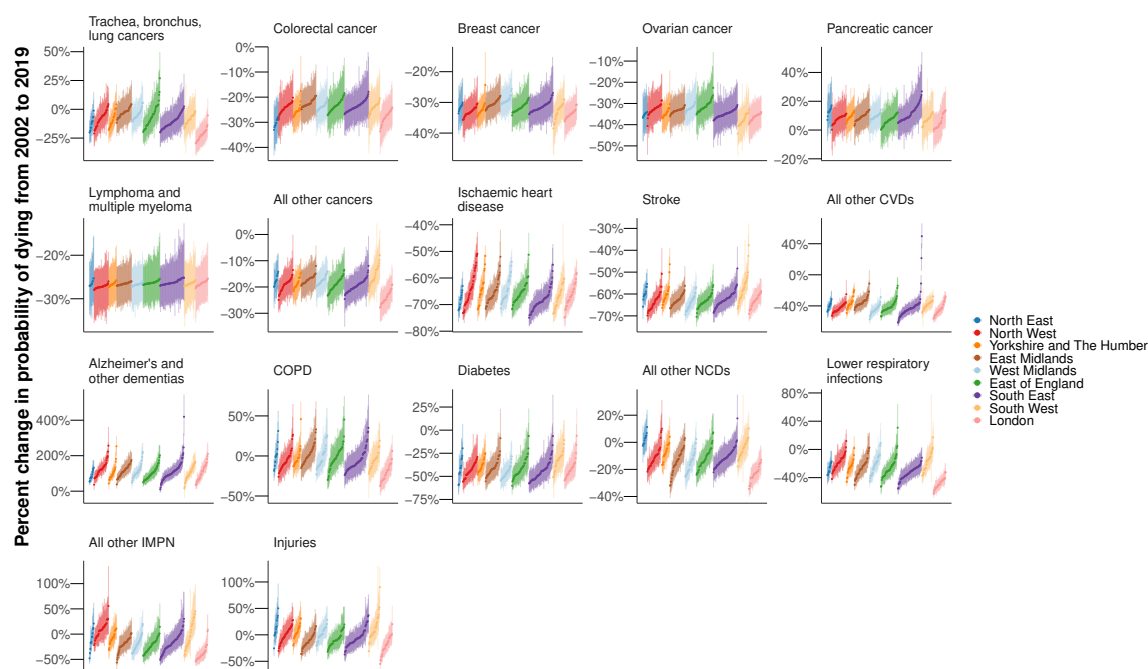
The highest probabilities of dying from diabetes for both sexes were in the east of London (Figure E.6). Specifically, the borough of Newham had the highest mortality for both sexes (0.01 (0.01-0.01) for women and 0.02 (0.01-0.02) for men). The highest probabilities of dying from injuries were in coastal areas and in the urban North West, with the highest probability of dying for men in Blackpool at 0.04 (0.04-0.05) (Figure E.26).

7.2.2 Variation over time

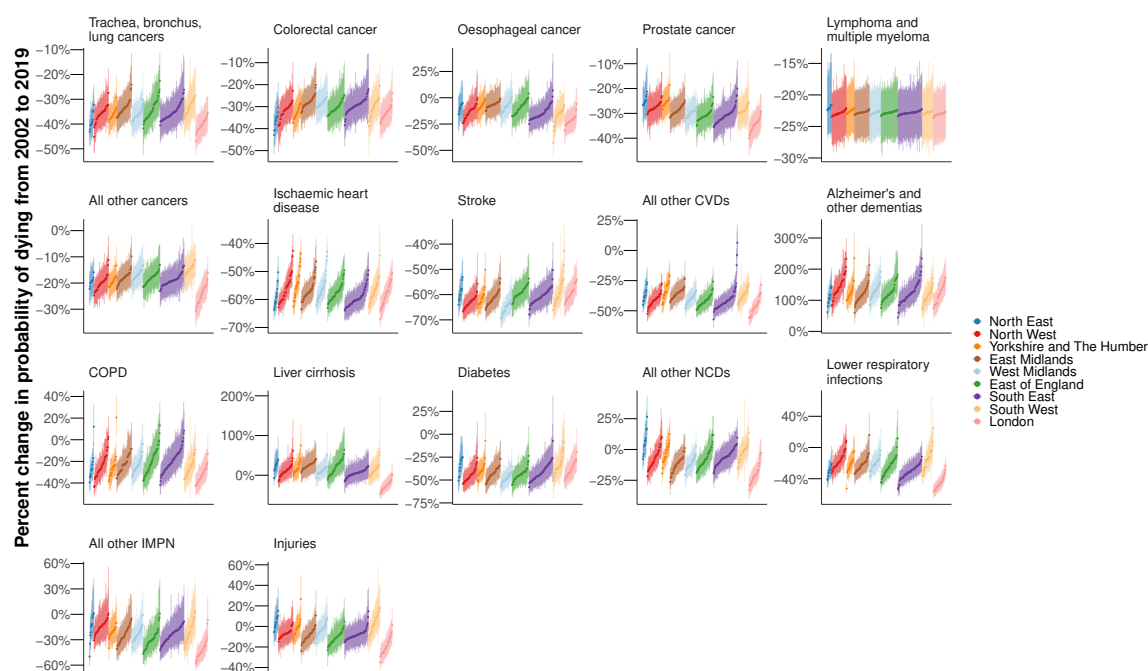
The largest declines from 2002 to 2019 were seen in CVDs (Figure 7.3). The probability of dying between birth and 80 years of age from ischaemic heart disease (median decline across districts of 67.1% for women and 58.2% for men) and strokes (62.7% and 61.3%) declined in all districts for both sexes, and mortality from all other CVDs decreased in all but the districts of Isle of Wight and Hastings (41.9% for women and 40.5% for men; Figure E.3). These increases were clear outliers which occurred in the residual group of CVDs rather than a defined leading cause of death, suggesting there are localised issues in cause of death assignment rather than true changes in population health. There were also declines in all districts for each sex for diabetes mortality (43.2% and 42.9%).

In contrast, the probability of dying from dementias saw the largest increases of any cause, with all districts observing an increase for both sexes (median increase across districts of 106.2% for women and 121.8% for men; Figure E.4). The largest increases were Eastbourne (419.4% (314.2% to 544.5%)) for women and in Kingston upon Hull (235.1% (168.8% to 316.6%)) for men.

Mortality from cancer causes of death declined in all districts and for both sexes with three exceptions: pancreatic cancer mortality, which saw increases in all districts (except a small decrease in Tower Hamlets for women), lung cancer mortality in women and oesophageal cancer mortality in men, which both saw decreases in most districts but increases elsewhere. Mortality from COPD also saw mixed trends, and with wider variability than the observed changes in lung



(a) Women



(b) Men

Figure 7.3: Ranked change in probability of dying between birth and 80 years of age in 314 local authority districts in England in 2002 and 2019 for the twelve leading causes of death and residual groups. The points are coloured by the regions in England. Each point shows one district and the vertical line going through the point its 95% credible interval. The districts were ranked first by region and then by the posterior median estimate for the probability of dying.

cancer mortality. The probability of dying declined in a larger proportion of districts for men (305 (97.1%)) than for women (215 (68.5%)). The largest increases for both sexes were seen in Kingston upon Hull (45.9% (24.8% to 67.9%) for women and 20.6% (2.9% to 40.7%) for men). See Chapter 8 for detailed results on cancer mortality.

The change in probability of dying from all other IMPN was greatly different between the sexes, with 71 (22.6%) districts seeing an increase for women, the largest in Barrow-in-Furness (55.6% (-0.8% to 135.2%)), compared to only 5 (0.02%) districts for men. There were also mixed trends for lower respiratory infections (increases in 13 (4.1%) districts for women and 9 (2.9%) for men), all other NCDs (29 (9.2%) for women and 53 (16.9%) for men), and injuries (116 (36.9%) for women and 9 (2.9%) for men).

Mortality from liver cirrhosis presented dramatic spatial variation in change in mortality over the period 2002-19. There was a contrast between huge decreases in districts in London and increases in the majority (243 (77.4%) districts) of districts elsewhere in the country. The largest increases were in the coastal South West and East of England (Figure E.7).

7.2.3 Contribution to life expectancy inequality

The district with the highest life expectancy for both sexes in 2019 was the City of London, a small and affluent district in the capital, at 90.1 years for women and 87.4 years¹ for men. For women, the district with the lowest life expectancy was Manchester, which was 10.1 years lower than the best-performing district. Blackpool had the lowest life expectancy for men, 13.0 years less than the City of London.

The major contributing cause to inequality with respect to the best-performing district was dementias for women, which was the leading contributor for 254 (80.9%) districts. For men, the all other NCDs group was the leading cause in 220 (70.0%) districts, with dementias ranking first in 64 (20.3%) districts. From the groups of cancers, lung cancer and all other cancers presented clear gradients, with the largest contributions in the worst-performing districts (Figure 7.4). Specifically, lung cancer had a difference in contribution between the 10th and 90th percentile districts of 0.6 years for women and 0.5 years for men, and all other cancers was 0.5

¹For the analyses on the contributions of different causes of death to life expectancy differences, I do not state uncertainty because I use the mean of the death rate samples. This is so the contributions from Arriaga's method sum exactly to the life expectancy difference. Hence, values of life expectancy differences and the cause-district-specific contributions towards these differences are mean estimates without uncertainty.

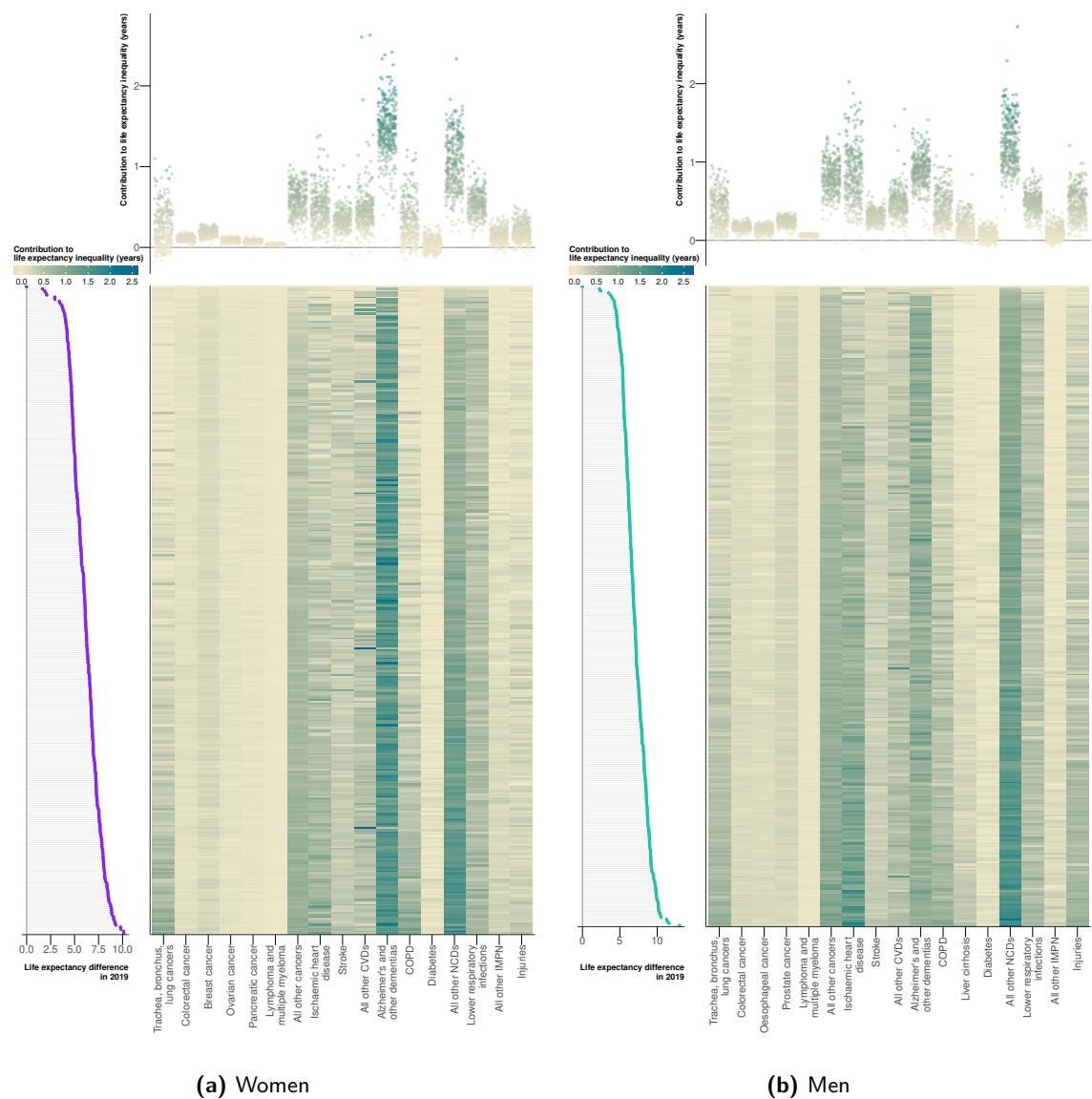


Figure 7.4: Contributions of deaths from specific causes to the life expectancy difference in 2019 between the district with the highest life expectancy (City of London for both sexes) and all other local authority districts in England. The districts are ordered from the best-performing district downwards, with the side panel showing the difference in life expectancy compared to the best-performing district, by summing over the causes. The top panel shows the contribution for each district for each cause.

years for both sexes. There were negative contributions from lung cancer to life expectancy inequality in 50 (15.9%) districts for women and two (0.6%) districts for men. There were strong gradients from the best- to worst-performing districts for all other NCDs (10th-90th percentile differences of 0.9 years for women and 0.8 years for men), ischaemic heart disease (0.5 years and 0.8 years), and COPD (0.6 years and 0.5 years). Contributions to inequality from dementias had large 10th-90th percentile differences of 0.8 years for women and 0.5 years for men. There was a more prominent gradient for injuries in men (0.5 years) than women (0.3 years).

7.2.4 Contribution to life expectancy change

From 2002 to 2019, life expectancy increased in every district for both women and men (Figure 7.5). This was driven largely by decreases in ischaemic heart disease mortality, which contributed gains of between 0.7 and 1.9 years for women between 1.4 and 4.0 years for men. In contrast, mortality from dementias acted in the other direction, forcing life expectancy downwards by between 0.2 and 1.1 years for women and between 0.4 and 1.0 years for men. Districts with the greatest improvements in life expectancy had large contributions from declines in CVD mortality as well as some of the smallest contributions from increases in mortality from dementias.

In Figure 7.6 and Figure 7.7, I break the overall trend down into two time periods around the year 2010. This both follows the analysis for Figure 5.9, but it is also the year there was a change in government, and a move in fiscal policy towards austerity.

In the period 2010-19, life expectancy increase slowed in England, as discussed at the MSOA level in Chapter 5, with three districts experiencing a decline in female life expectancy. For women, lung cancer switched from contributing negatively in most districts (277 (88.2%)) to the change in life expectancy (Figure 7.6a) to positively (312 (99.4%) districts; Figure 7.6b), unlike COPD, for which more districts (252 (80.3%)) saw negative contributions in 2010-2019 compared to 2002-2010 (95 (30.3%)). The contribution for the other cancer groups was similar for the two time periods. All districts saw positive contributions from all other NCDs to life expectancy increase in the second period compared to the first, where only 216 (68.8%) districts contributed positively. The positive contribution from all the CVD groups (ischaemic heart disease, stroke, all other CVDs) was weaker for the second time period, with the negative forcing effect of dementias strengthening massively in all districts. In the years 2010 to 2019, the *inequality in progress*

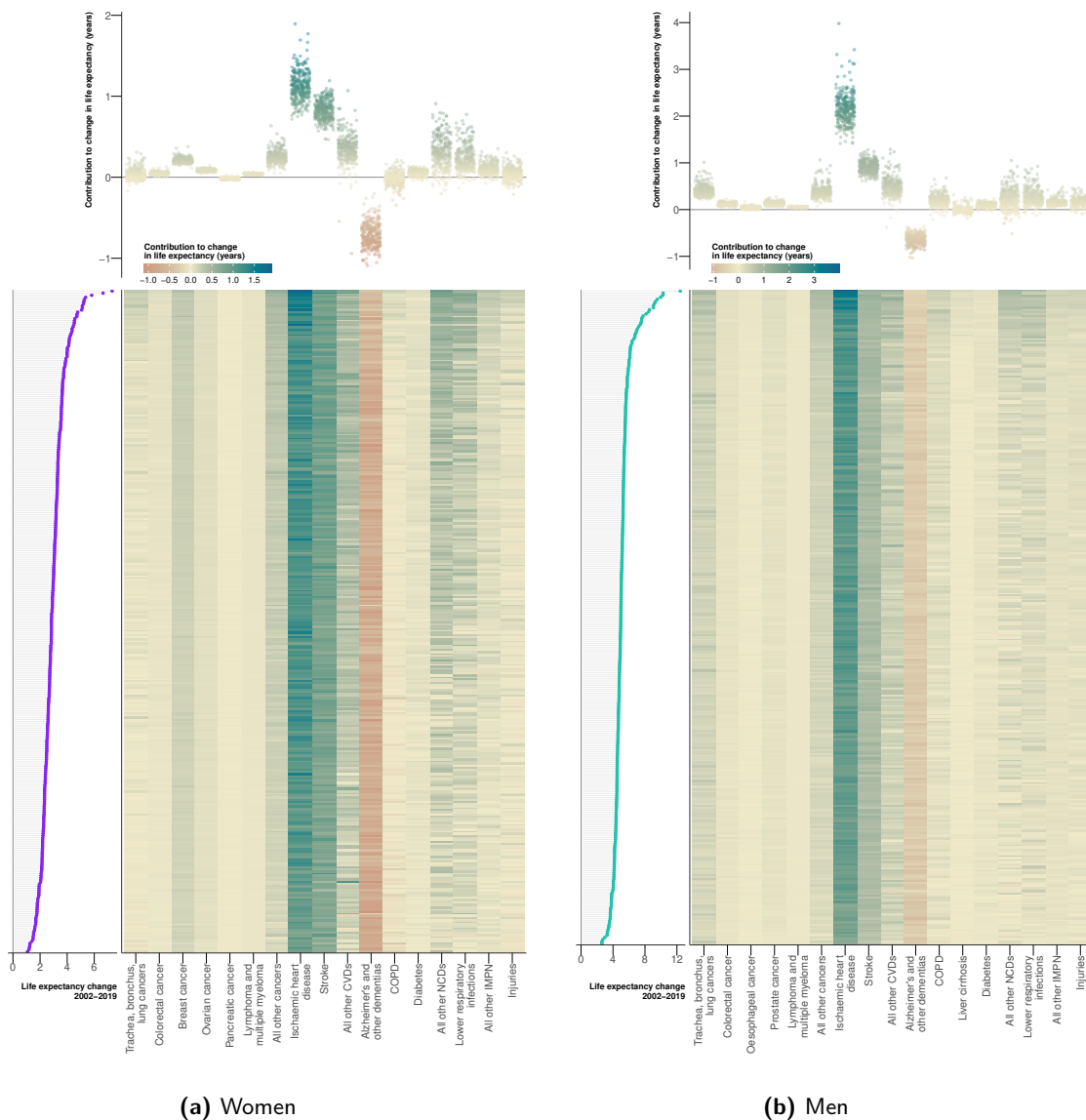


Figure 7.5: Contributions of deaths from specific causes to the difference between the district life expectancy in 2002 and 2019 in 314 local authority districts in England. The districts are ordered from the district with the largest increase in life expectancy downwards, with the side panel showing the change in life expectancy over that time, by summing over the causes. The top panel shows the contribution for each district for each cause.

between districts was driven largely by dementias and all other NCDs, both with a difference in contribution between the 10th and 90th percentile districts of 3.0 years.

For men, the contribution for all cancers groups, including lung cancer, maintained the same picture in each time period. As for women, the slowdown in progress was driven by reductions in the contribution from CVDs in the second time period and the strong negative contributions from dementias in all districts. Injuries, which contributed positively to life expectancy increases in 2002-10 for all districts, saw a reversal with all but 19 (6.1%) districts contributing negatively in the second time period. In the latter period, mortality from ischaemic heart disease, dementias, and all other NCDs were driving the inequality in progress between districts, with differences in contribution between the 10th and 90th percentile districts of 2.8, 2.9 and 3.0 years respectively. For both sexes and in both time periods, contributions from lymphoma and multiple myeloma were the smallest in magnitude of any cause group (all <0.1 years in each sex-district-time period combination).

7.3 Discussion

This analysis revealed huge inequalities in mortality for a wide-ranging variety of causes of death in England. Although there were widespread declines in CVDs, there have been huge increases in dementias, which in 2019 became one of the most unequal causes of death, with 4.1- and 3.6-fold differences for women and men between the highest and lowest districts. I found that the slowdown in life expectancy gains since around 2010 has been driven largely by dementias and all other NCDs in women, and by ischaemic heart disease, dementias, and all other NCDs in men.

7.3.1 Strengths and limitations

The work presents trends in cause-specific mortality over a period of substantial policy interest at the district level, and uncovers which causes of death have been driving the slowdown in progress since 2010.

The study was carried out at the district level rather than for MSOAs, which masks substantial heterogeneity, as shown by the all-cause mortality study in Chapter 5. However, when stratifying further by cause of death, the number of deaths in each age-space-time-cause stratum can be

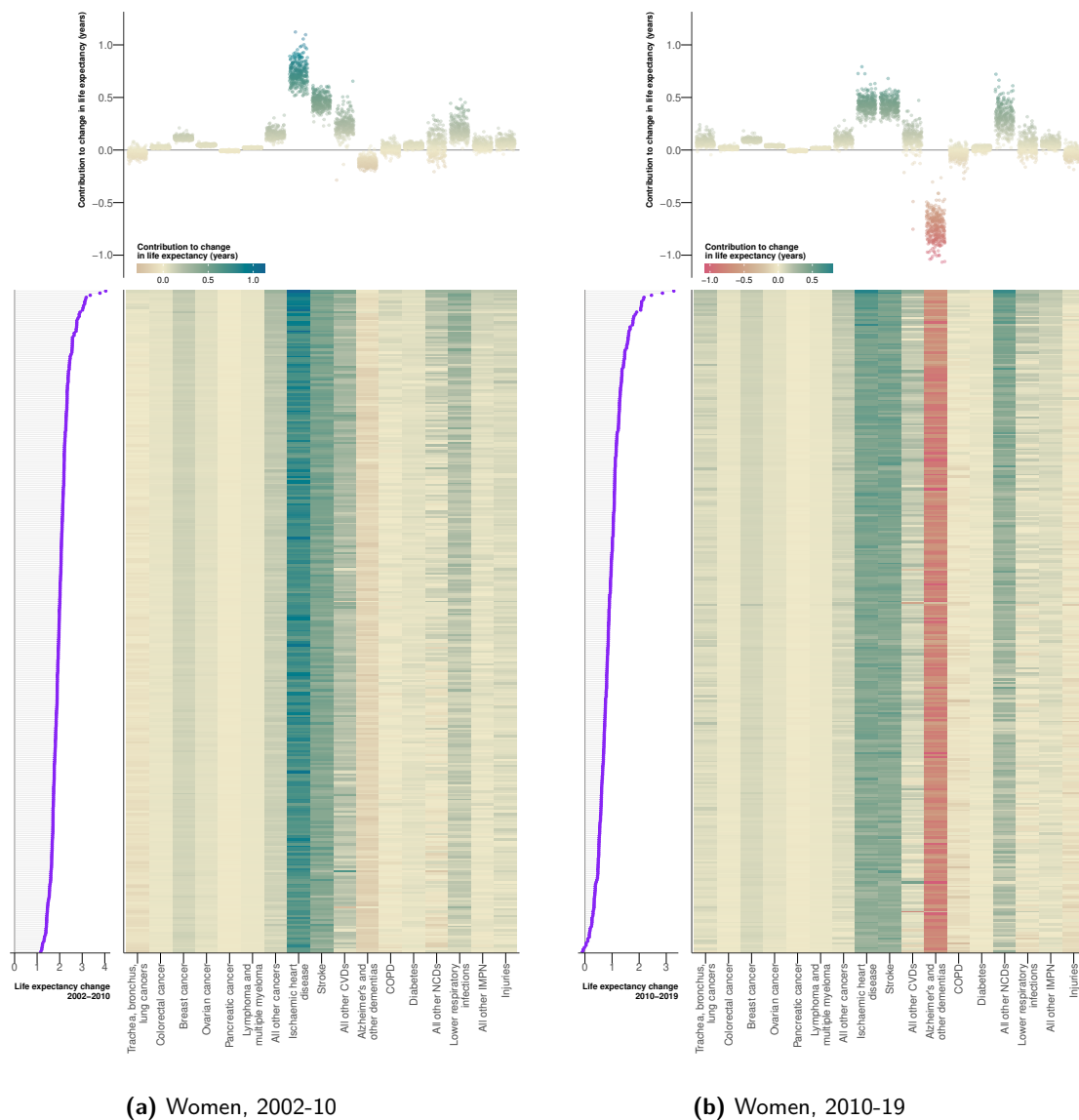


Figure 7.6: Contributions of deaths from specific causes to difference between the district life expectancy for women in 2002 and 2010, and in 2010 and 2019, in 314 districts in England. The districts are ordered from the district with the largest increase in life expectancy downwards, with the side panel showing the change in life expectancy over that time, by summing over the causes. The top panel shows the contribution for each district for each cause.

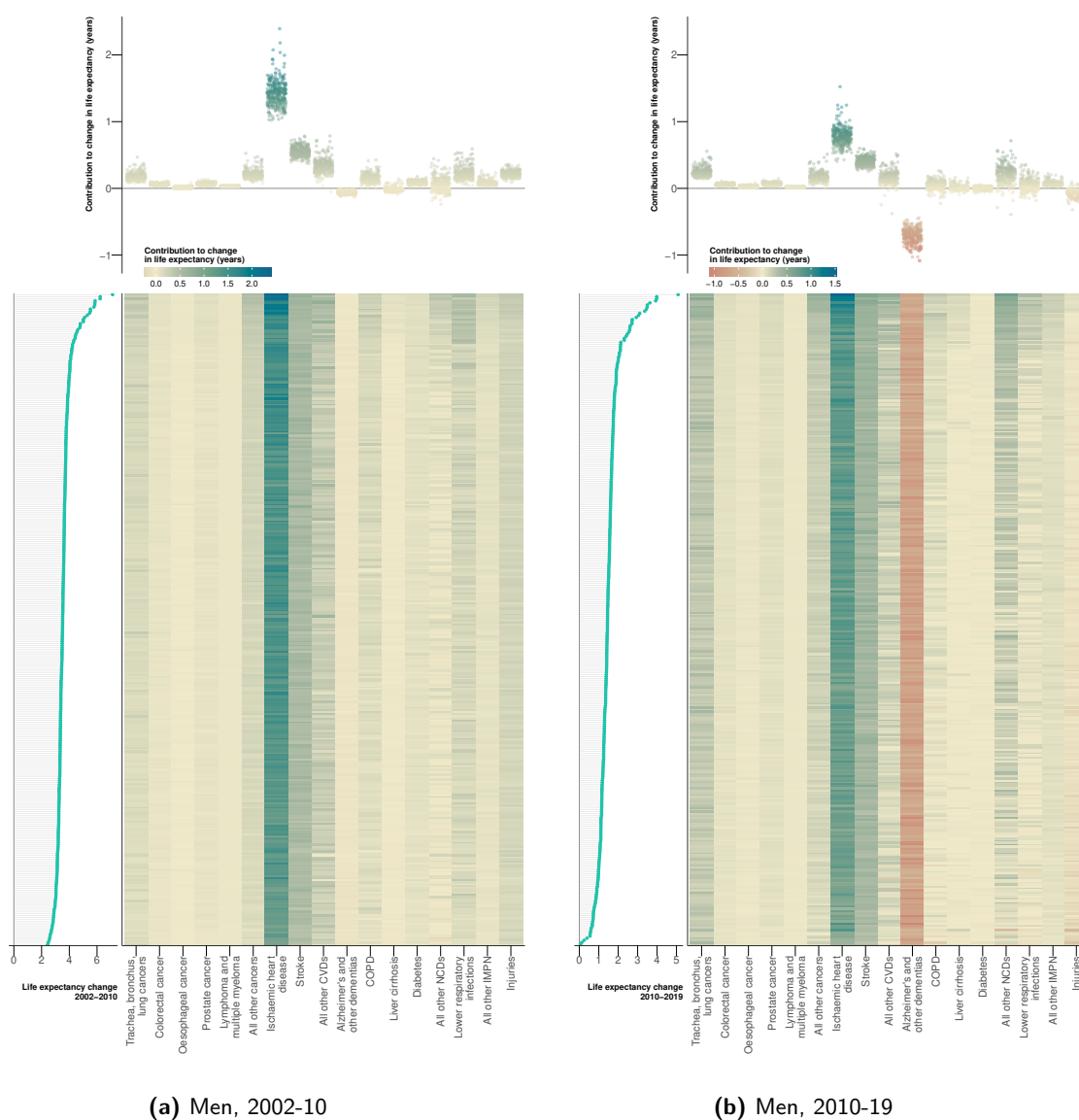


Figure 7.7: Contributions of deaths from specific causes to difference between the district life expectancy for men in 2002 and 2010, and in 2010 and 2019, in 314 districts in England. The districts are ordered from the district with the largest increase in life expectancy downwards, with the side panel showing the change in life expectancy over that time, by summing over the causes. The top panel shows the contribution for each district for each cause.

extremely small, even at the district level, which makes the estimation task difficult. Furthermore, running a model that takes a day of GPU time to reach convergence for all causes at the MSOA level for 34 cause-sex combinations is extremely computationally demanding.

This study did not look at age- *and* cause-specific contributions to life expectancy inequality, unlike Bennett et al. (2018), and instead collapsed over age groups. Although this masks some variation, the groups were selected based on the total number of deaths from 2002 to 2019, and are consequently skewed towards older ages, and the age-specific contributions of (e.g.) dementias are not particularly interesting.

Death records are subject to issues in the assignment of ICD-10 codes for the cause of death. The quality of the cause of death data in the UK has been given the highest rating based on completeness and a low share of deaths assigned to implausible and ill-defined codes (World Health Organization, 2020). There are national guidelines for death certification, and the ONS uses standardised coding algorithms and computer-based assignment of the cause of death, which assign the underlying cause of death uniformly to improve consistency between doctors when identifying the underlying cause of death (Office for National Statistics, 2022b). For example, a validation study found that that death certificates in the UK accurately identified deaths from prostate cancer (Turner et al., 2016). Despite this, the challenge of multimorbidity in older age groups makes the assignment of cause of death increasingly difficult (Meslé and Vallin, 2021). We selected 80 years of age as the upper bound to partially mitigate this effect because it covers a wide age range but does not include the very oldest ages. However, for certain diseases such as dementias where only 96,109 (14.5% of 663,692) of deaths occur in those under 80 years, the probability of dying between birth and 80 years masks variations in the age groups where the majority of deaths occur. There are also specific cases of problematic cause of death assignment in the data. For example, mortality from all other CVDs showed a downward trend in every district with the exceptions of Hastings and, more noticeably, the Isle of Wight. Looking at the raw data for these districts, there is a step increase in the use of ICD-10 code I51.5 – a garbage CVD code (Murray and Lopez, 1996) – since around 2012 in these districts, suggesting these trends result from poor coding practice.

There is always the desire to split the residual groups into more and more causes of death. For example, there are interesting patterns in the change of all other IMPN, and it would be

interesting to understand further how the component causes of all other NCDs are driving the inequality in progress. Moreover, there has been a lot of recent attention in the literature on “deaths of despair” – deaths from drug overdose, suicide or alcoholic liver disease – in the UK (Angus et al., 2023) and the US (Case and Deaton, 2015), and it would be of interest to separate injuries into whether the deaths were intentional or unintentional. However, the groupings were chosen based on a total mortality rule so that the number of deaths in each cause group allowed robust inference of cause-age-area-year-specific death rates at the district level.

It would have been desirable, albeit computationally challenging, to run all cause groups in a single joint model. When calculating contributions using Arriaga’s method, mortality for each cause group was scaled such that the sum across all causes equalled to the estimate for total mortality. A joint model of cause-specific mortality could account for correlations for diseases with common risk factors and anti-correlations for competing causes of death. Foreman et al. (2017) jointly forecasted cause-specific death rates for states in the US, but instead estimated each age group separately.

The population in each district can change because of migration, both within the country and from overseas. However, unlike Chapter 5 and Chapter 6, migration was not as much of an issue at the district level compared with the smaller levels of analyses used earlier in the thesis. The majority of moves in the UK are within the same district (van Dijk et al., 2021), and so will have little effect on the change in mortality in a given district.

7.3.2 Comparison with previous literature

The general trends of decreasing mortality from CVDs and large increases in dementias are consistent with studies on mortality in deciles of deprivation at the national level in England (Bennett et al., 2018). Despite increases in the rates of diabetes over the past decades (NHS Digital, 2023), the observation that mortality from diabetes decreased in every district for both sexes is also consistent with trends at the national level and reflects a diversification of causes of death in individuals with diagnosed diabetes (Pearson-Stuttard et al., 2021).

There are two key studies in the literature which break the population of England down into subgroups and look at the cause-composition of mortality: Firstly, from my research group, Bennett et al. (2018) divided the population by deciles of deprivation and looked at largely

similar groups of causes to this chapter. The authors found the gap in life expectancy between the most and least deprived areas was caused by the large contributions of respiratory diseases, ischaemic heart disease, and a set of preventable and treatable cancers. Unlike the present work, the authors broke the results down further by age, finding that deaths in children younger than 5 years and dementias in older ages drove the inequality between the top and bottom deciles.

Secondly, using data from one of the first GBD studies of subnational populations, Steel et al. (2018) looked at trends in mortality of 150 upper tier local authorities, which are larger than the districts used in this chapter, but considered 264 causes of death, relying heavily on covariates in the model to produce estimates for areas with little to no data on more obscure causes. As with Bennett et al. (2018) and the work here, the GBD group found the rate of improvement in life expectancy had slowed since 2010. They found, in general, the slowdown was driven by the disappearance of sustained annual improvements from ischaemic heart disease, strokes, as in the present results, but also to a lesser extent from colorectal cancer, lung cancer, and breast cancer. The authors found the annual reduction in mortality attributable to most major risk factors (tobacco, dietary risks, high blood pressure etc) had declined since 1990, but has slowed since 2010, with the exception of alcohol and drug use, which has seen little change since the 2000s.

Both these studies focus on the causes of death that have driven the inequality between subgroups of the population. I found the causes with the clearest gradients in contribution from the best- to worst-performing districts were those with modifiable risk factors, such as ischaemic heart disease, lung cancer, COPD, and injuries, and also the residual groups of all other NCDs and all other cancers. There were also contributions from dementias, but these did not follow the direction from best- to worst-performing district as obviously as the aforementioned causes. The work in this chapter extends previous studies by also considering which causes of death have driven the *inequality in progress* between subgroups of the population.

7.3.3 Explaining the variation

Asaria et al. (2012) found CVD mortality followed a persistent downward trend in nearly all wards in England from 1982 to 2006. Both mortality from ischaemic heart disease and strokes have continued to follow this trend through to 2019 at the district level.

In England, about 51% of all mortality from myocardial infarction, an acute manifestation of ischaemic heart disease, is attributable to out-of-hospital deaths (Asaria et al., 2017). Reductions in out-of-hospital mortality reflect improvements in reducing and controlling risk factors for CVD (smoking, blood glucose and diabetes, raised blood pressure, and high blood cholesterol) (Ezzati et al., 2015). Unlike cancers and COPD where the increased risk following smoking spans decades, the risk of CVD returns to the level of non-smokers within ten years after smoking cessation (Ezzati et al., 2015). Declines in CVD mortality have thus gained greatly from reductions in smoking rates, and more recently, the ban on smoking in public places, which came into force in the UK in 2007 and has rapidly decreased hospitalisation rates for acute CVD cases (Pell et al., 2008). As well as improvements in these risk factors, reductions in in-hospital CVD mortality partially reflect organisational changes to the NHS such that patients with acute myocardial infarctions are taken directly to centres with capacity to re-open coronary arteries, bypassing local accident and emergency services. Furthermore, strategies to shorten pre-hospital delays between symptom onset and a call for help, for example, the FAST (Face drooping, Arm weakness, Speech difficulties, Time) public health campaign for stroke events, can improve the efficacy of treatments such as thrombolysis.

The burden of mortality has shifted towards dementias. A number of the main risk factors for dementias are the same as for CVDs (smoking, obesity, diabetes, high blood pressure, high cholesterol) (Yu et al., 2020). However, dementia diagnoses are also heavily influenced by age, family history and education. Some part of this trend may also be due to increased efforts in improving diagnosis and cause of death coding of dementias in the UK (Hayat et al., 2022; Mukadam et al., 2014).

These contrasting trends in mortality from ischaemic heart disease and dementias could explain the finding that these causes of death have largely driven the heterogeneity of the slowdown in life expectancy gains since around 2010 at varying rates across districts, suggesting that heterogeneous dynamics of CVD risk factors within the population have influenced the inequality in progress in recent years. The sizeable contribution to the inequality in life expectancy improvement from all other NCDs is more difficult to explain without further stratifying the cause group.

The heterogeneous trends in mortality from both lung cancer, where the probability of dying

declined in all districts for men and saw mixed trends for women, and from COPD, where a larger proportion of districts experienced a decrease in mortality for women than for men, reflected that the peak in female smoking rates and smoking-attributable mortality have lagged behind that in men by about 20-30 years (Thun et al., 2012). This may also be responsible for the larger contributions of lung cancer and COPD mortality for women to inequality in life expectancy across districts.

The geography of the change in mortality from liver cirrhosis – an advanced stage of liver damage – is perhaps indicative of the contrasting dynamics of the two main risk factors for liver cirrhosis: alcohol misuse and hepatitis B/C infection. Alcohol is the main cause of liver disease, and has driven a large proportion of increases in liver cirrhosis throughout Europe (Blachier et al., 2013). Alcohol is generally consumed less in the capital, with London having the highest percentage of adults who abstain from drinking alcohol (23.6%) and the lowest percentage of adults who drink over 14 units per week (20.1%) (Public Health England, 2021b). The prevalence of hepatitis B/C has decreased over the past decades. There has been reduced incidence, in part due to vaccination, for hepatitis B. Although a large number of people acquired hepatitis C in the 1970s and 1980s, the virus has since been identified and transmission has reduced (Blachier et al., 2013).

As is the case in most high income countries, which are in the third stage of the epidemiologic transition, there were far fewer deaths from maternal, perinatal, nutritional and infectious causes (GBD Group 1) and injuries (GBD Group 3) compared to CVDs, cancers and other NCDs (GBD Group 2). Although, notably, these results consider the period before the Covid-19 pandemic, and therefore the cause-composition will likely have changed to include more deaths from infections in the most recent years. The similar profiles for men and women for the level of mortality from all other IMPN in 2019 reflects that the three main causes by number of deaths are other infectious diseases, diarrhoeal diseases and low birth weight, rather than maternal conditions, although without further disaggregating this group, it is difficult to suggest drivers of the observed trends. Mortality from lower respiratory infections, which include influenza and pneumonia, can vary greatly between years depending on the severity of the virus and whether the flu vaccination prevents the prevalent strain. It also depends on the strength of the immune system of the infected. The ability of the lungs to recover from infection can be degraded by years of industrial

exposure or smoking, which could influence the high mortality in the urban North West. The highest mortality from injuries were in coastal areas, where there are obvious natural hazards, but also include a number of “left-behind” districts following deindustrialisation and lack of investment (Whitty, 2021). The highest male mortality from injuries was in Blackpool, which agrees with both the literature on mortality from drugs and suicide (Congdon, 2019) and coincides with the MSOA level life expectancy results (Chapter 5).

7.4 Summary

I performed a spatiotemporal analysis of vital registration data on deaths from the twelve leading causes of death by sex in England from 2002 to 2019. I used life table methods to calculate the probability of dying between birth and 80 years of age by sex, cause of death, district and year, as well as the cause-specific contributions to the inequality in life expectancy between different years of the study for each district.

The most unequal causes of death were COPD for women (6.0-fold variation in mortality across districts) and liver cirrhosis for men (6.7-fold). The causes of death with the least geographical variability were lymphoma and multiple myeloma (1.2-fold for both sexes) and leukaemia (1.1-fold for women and 1.2-fold for men). There has been a slowdown in life expectancy gains since 2010, which has been driven largely by dementias and all other NCDs in women, and by ischaemic heart disease, dementias, and all other NCDs in men.

Chapter 8

Trends in cancer mortality at the district level

The work in this chapter has formed the paper *Mortality from leading cancers in districts of England from 2002 to 2019: a population-based, spatiotemporal study*, published in *The Lancet Oncology* (Rashid et al., 2024) and for which I am first author. The paper was published under the CC-BY license, which permits reproduction provided the original work is properly cited.

As seen in the previous chapter, mortality from cancers in England has declined more slowly than from other major causes of death such as CVDs, and hence the share of deaths from cancers has steadily increased. In 2019, there were more deaths from cancers (144,306 (28.9% of all deaths in 2019)) than from CVDs 126,105 (25.3%), although this was less than the number of deaths from other NCDs excluding CVDs and cancers (173,323 (34.8%)) which reflects the rise in dementia deaths in recent years. Compared with CVDs and other NCDs, which were dominated by their major causes (ischaemic heart disease and stroke, and Alzheimer's, COPD and diabetes, respectively), cancers had the largest number of individual causes in the top twelve causes of death. With over 150 different types of cancer, each with their own anatomical and molecular subtypes, cancer is extremely complex, with a specialist workforce dedicated to each unique cancer. I felt I should go deeper into the cancer story, and pay further attention to a wider array of site-specific cancers.

Subnational data on trends in cancer mortality are currently limited to large areas such as health boards or regions (Arik et al., 2021; NHS Digital, 2022). Small-area data can guide where there

is a need for primary prevention strategies to reduce incidence, and for health-care planning and delivery to improve survival. This evidence is particularly relevant for cancers, because their incidence and survival can be affected by a range of risk factors, screening to detect precancerous lesions and early-stage disease, and effective treatments.

8.1 Methods

The methodology for this section is the same as Chapter 7, but I have stratified cancer groups further. I used the top ten leading cancer causes of death according to the total number of deaths from 2002 to 2019 for cause-specific analysis, as well as a residual group comprising all other cancer deaths. The residual group also contained deaths from ill-defined diseases, which were assigned as described in Chapter 7.

There were 2,453,173 deaths from cancers in England from 2002 to 2019; of these, 1,533,703 (62.5%) deaths occurred before 80 years of age (Figure 8.1). Of cancer deaths before 80 years of age, 697,953 (45.5%) were deaths in women and 835,750 (54.5%) in men.

As I have increased the number of site-specific cancers from six (women) and five (men) to ten for each sex, the residual group in this chapter is smaller than the *All other cancers* group in Chapter 7. The results from the causes of cancer mortality that were in the twelve leading causes of death in the previous chapter are identical to those in this chapter. The full list of ICD-10 codes for each cause group can be found in Table D.2.

The modelling process for estimating cause-specific death rates is the same as Chapter 7. In addition, I ran a model for all cancer deaths combined, so that the credible intervals for overall cancer mortality are directly estimated. The results for total cancer mortality using the combined model and the sum of cause-specific models were nearly identical (correlation coefficient across all years 0.99 for both sexes; mean absolute difference in probability of dying <0.01 for both sexes).

The primary reporting outcome is, again, the unconditional probability of dying between birth and 80 years of age. I also report the mean age at death among those who died of that cause. Details on the calculations for both the probability of dying and the mean age at death can be found in Appendix A.

Men					
		Prostate 164,871 (12.8%)		Colorectal 129,382 (10.1%)	
All other cancers 333,822 (26.0%)				Stomach 45,852 (3.6%)	
		Oesophageal 75,457 (5.9%)	Pancreatic 61,558 (4.8%)	Leukaemia 39,368 (3.1%)	
Trachea, bronchus and lung 282,422 (22.0%)		Lymphoma and multiple myeloma 64,734 (5.0%)	Bladder 51,883 (4.0%)	Liver 36,257 (2.8%)	
Women					
		Trachea, bronchus and lung 218,561 (18.7%)		Colorectal 111,364 (9.5%)	Pancreatic 62,379 (5.3%)
				Ovarian 62,338 (5.3%)	Oesophageal 36,806 (3.2%)
					Leukaemia 29,465 (2.5%)
All other cancers 359,262 (30.8%)		Breast 177,528 (15.2%)	Lymphoma and multiple myeloma 54,046 (4.6%)	Corpus uteri 29,000 (2.5%)	Stomach 26,818 (2.3%)

Figure 8.1: Total number of deaths for the ten leading cancers in England from 2002 to 2019. See Table D.2 for ICD-10 codes for each category.

8.2 Results

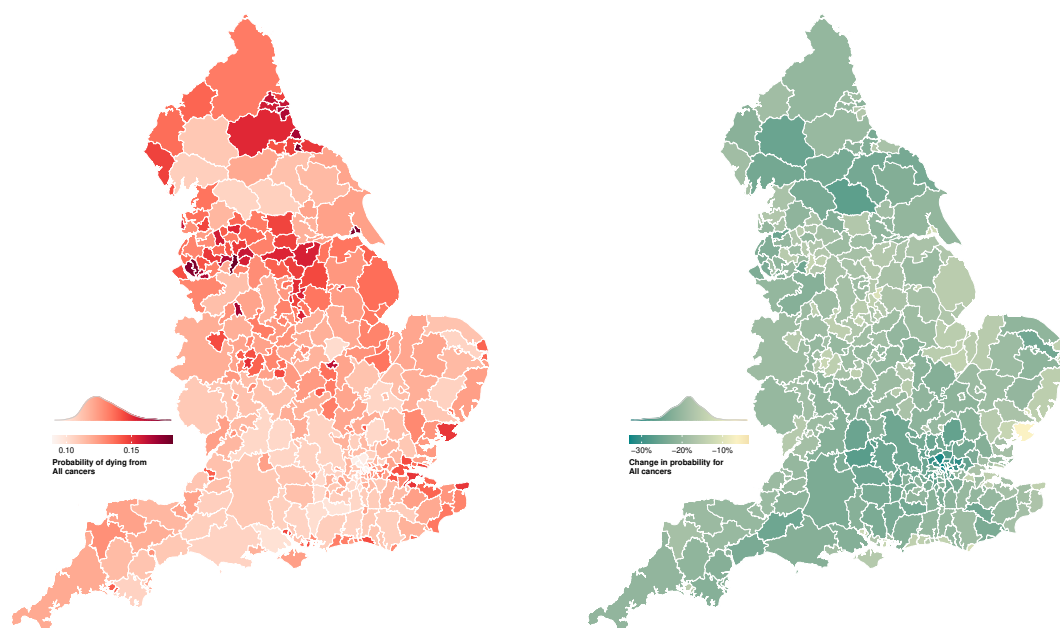
8.2.1 All cancers

In 2019, the probability of dying from a cancer before 80 years of age ranged from one in ten (0.10 (95% CrI 0.10-0.11)) in Westminster to nearly one in six (0.17 (0.16-0.18)) in Manchester for women, and from one in eight (0.12 (0.12-0.13)) in Harrow to around one in five (0.22 (0.21-0.23)) in Manchester for men (Figure 8.2). The highest probabilities of dying were predominantly in northern cities such as Liverpool, Manchester, Hull and Newcastle, and in coastal areas to the east of London. The probability of dying from a cancer declined in all districts from 2002 to 2019, but at varying rates. The district with the greatest reduction for women of 30.1% (25.6% to 34.5%) was nearly five times that of the smallest, at 6.6% (0.3% to 13.1%). For men, the largest decrease was triple that of the smallest, 36.7% (32.2% to 41.2%) compared to 12.8% (7.1% to 18.8%).

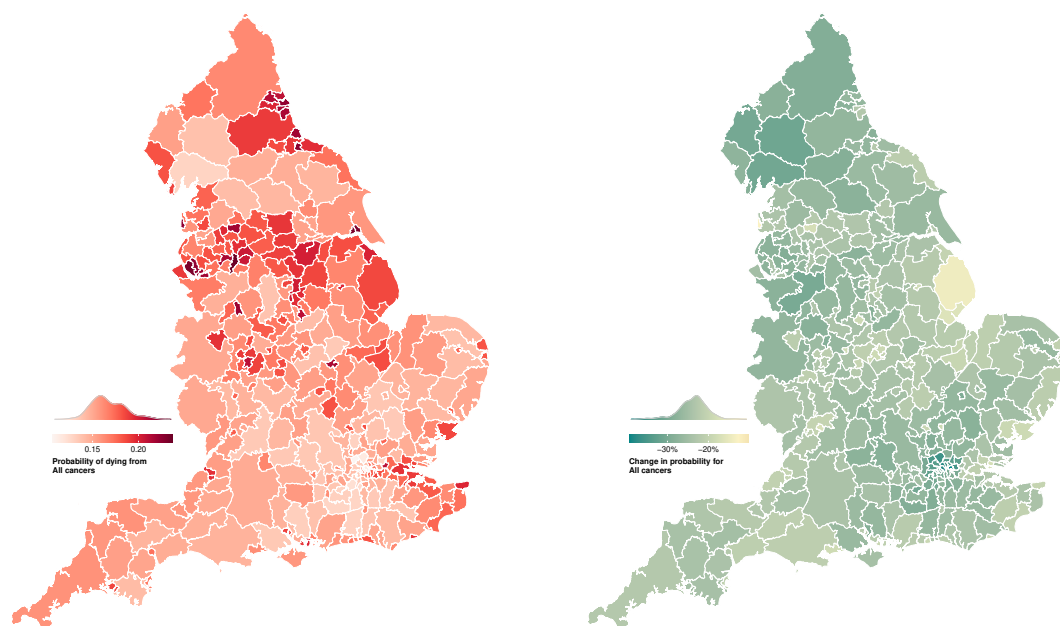
8.2.2 Site-specific cancers

The mean age at death in 2019 among those who died before 80 years of age varied from 63.0 years for breast cancer to 69.0 years for trachea, bronchus and lung cancers (hereafter referred to as *lung* cancer for brevity) in women and from 66.1 years for the combined category of all other cancers to 72.2 for prostate cancer in men.

The leading cancer cause of death for both sexes was lung cancer, causing a total of 218,561 deaths (18.7% of all cancer deaths across all ages) in women and 282,422 deaths (22.0%) in men from 2002 to 2019 (Figure 8.1). Of these, 148,551 (42.4%) and 201,862 (57.6%) occurred before 80 years of age in the two sexes, respectively. Lung cancer was also one of the most unequal cancers, with 3.7- and 3.1-fold variation across districts in 2019 in the probability of death for women and men, respectively. The highest probabilities of dying were concentrated in the urban North West and North East (Figure E.9), with the highest probability of death in 2019 being nearly one in sixteen for women (0.06 (0.06-0.07) in Knowsley) and above one in fourteen for men (0.07 (0.07-0.08) in Manchester). The distribution of district-level probabilities of dying from lung cancer was nearly identical in 2002 and 2019 for women, whereas for men the probabilities declined everywhere from 2002 to 2019 (Figure 8.3).



(a) Women



(b) Men

Figure 8.2: Probability of dying from cancer between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.

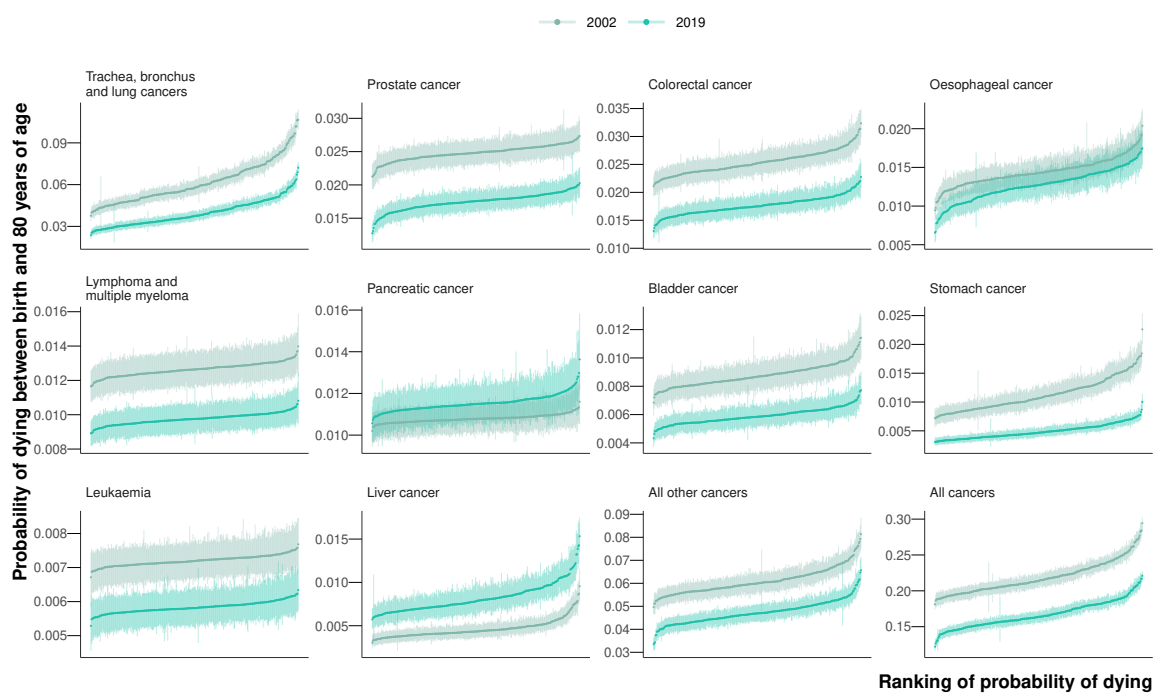
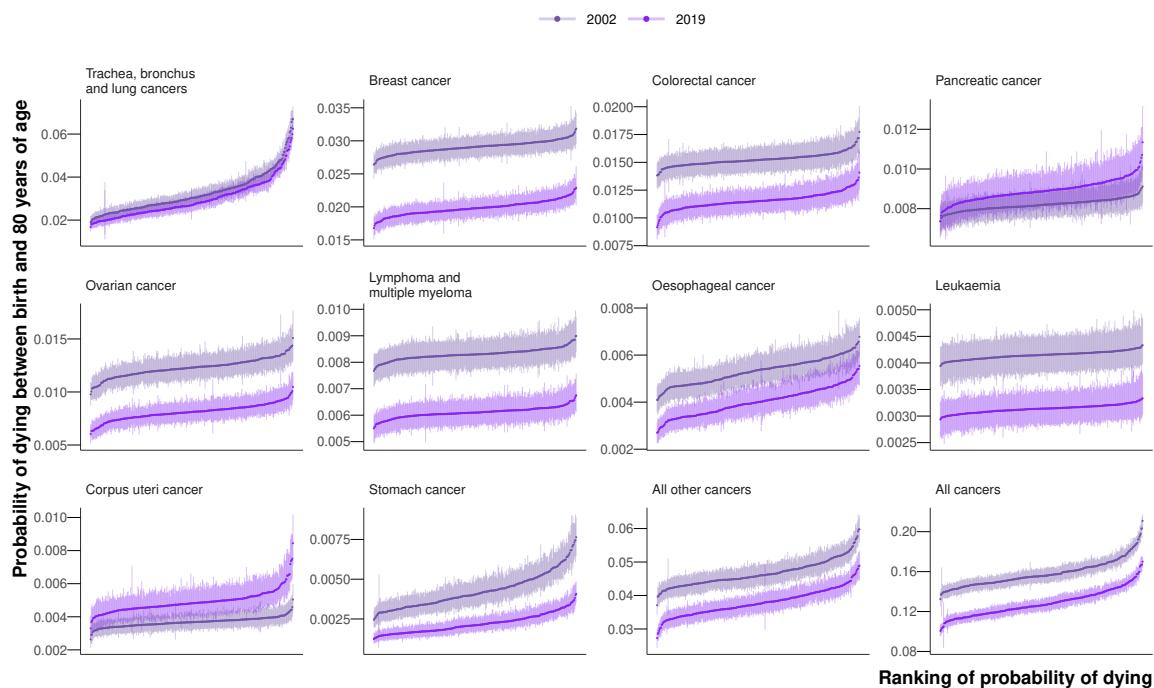


Figure 8.3: Ranked probability of dying between birth and 80 years of age in 314 local authority districts in England in 2002 and 2019 for the ten leading cancers. Each point shows one district and the vertical line going through the point its 95% credible interval.

Lung cancer was followed by the major sex-specific cancers – breast cancer (177,528 (15.2%) deaths) for women and prostate cancer (164,871 (12.8%) deaths) for men. In 2019, deaths from these cancers had less district-level inequality than from lung cancer (Figure 8.3), although the probability of dying from prostate cancer was noticeably lower in northwest London than elsewhere (Figure E.19). The probability of dying from ovarian cancer was also lower in London, in contrast to corpus uteri cancer, where east London had the highest probabilities of dying nationally (Figure E.18). High probabilities of death from liver cancer, one of the leading causes of cancer mortality among men, were very concentrated in the North West, the North East, and central London (Figure E.20). Stomach cancer mortality showed vast inequality (3.3-fold for both sexes), with the highest probabilities of dying in the urban North West (Figure E.13). The cancers with the least geographical variability were lymphoma and multiple myeloma (1.2-fold for both sexes) and leukaemia (1.1-fold for women and 1.2-fold for men).

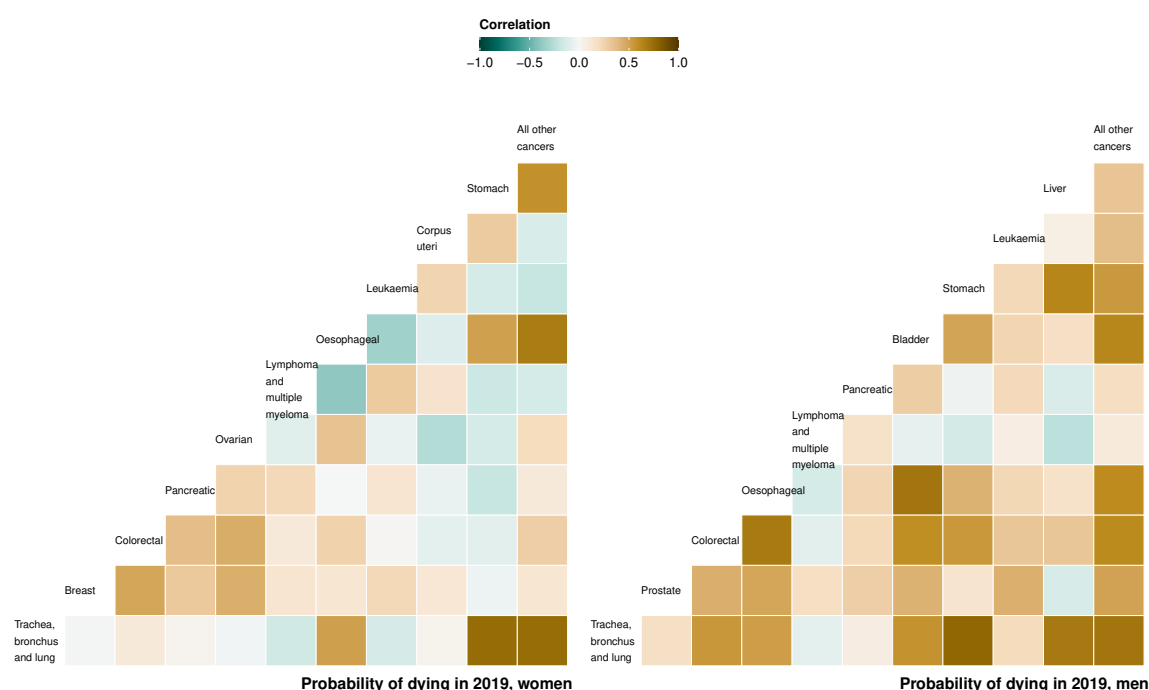


Figure 8.4: Pairwise correlations between the leading cancers for each sex in 2019 women and men.

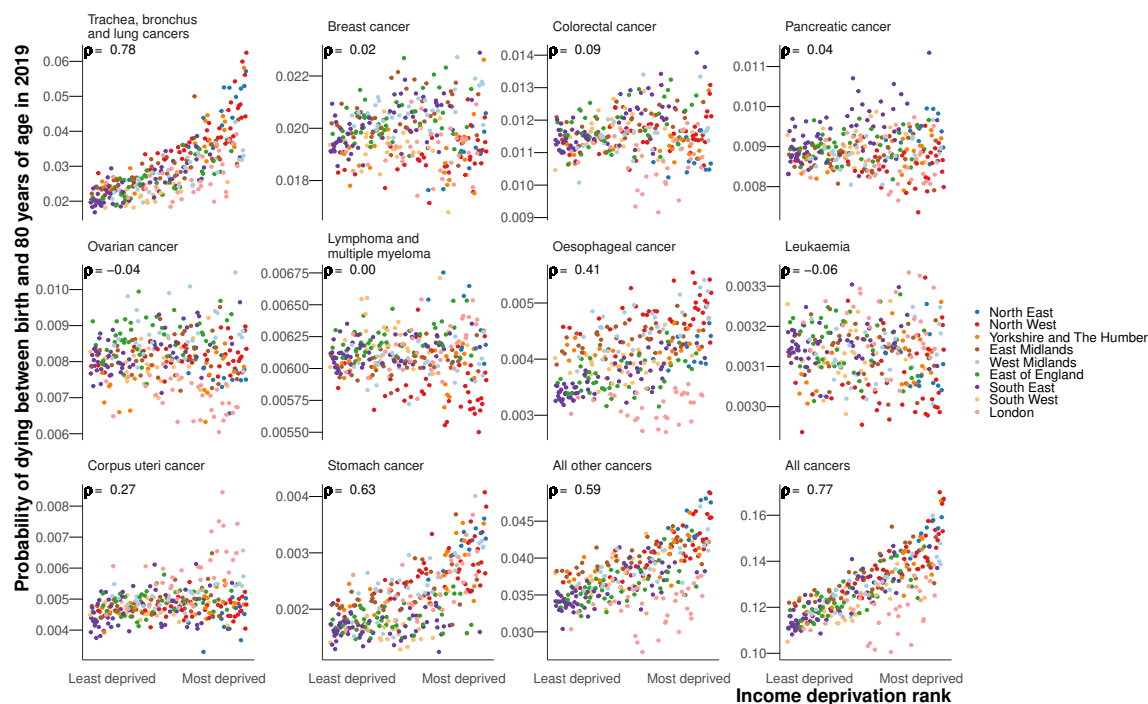
For women, the district-level probabilities of dying from stomach, oesophageal, lung and the residual category of all other cancers were correlated, with pairwise correlation coefficients ranging from 0.51 to 0.75 (Figure 8.4). For men, the probability of dying from lung cancer was correlated with the probability of dying from stomach, liver, bladder, colorectal, oesophageal, and all other cancers. There were particularly high correlations between the probabilities of dying from bladder

and oesophageal cancers (0.71), and from oesophageal and colorectal cancers (0.69). Probabilities of dying from leukaemia, lymphoma and multiple myeloma, pancreatic or corpus uteri cancer were only weakly correlated with probabilities for other categories.

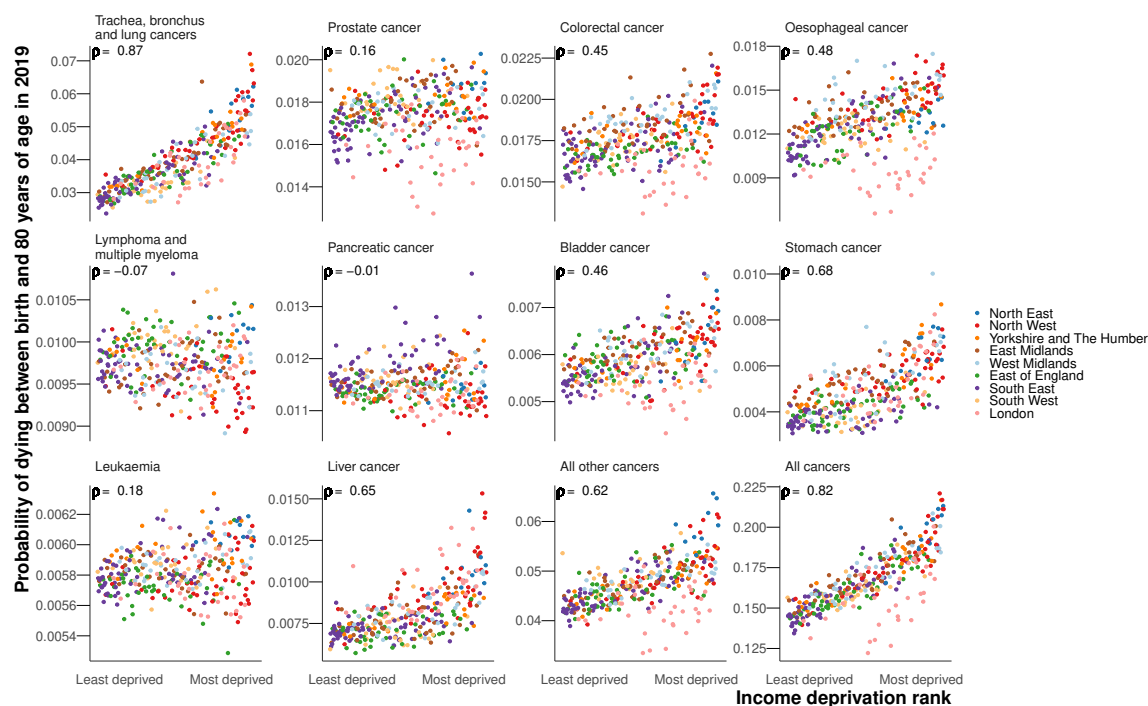
The probability of dying from a cancer was associated with poverty for both sexes, with Spearman correlation coefficients of 0.77 for women and 0.82 for men (Figure 8.5). In part, this was due to the probability of dying from the leading cancer for both sexes, lung cancer, exhibiting the strongest correlations with poverty of any of the cancer groups (0.78 for women and 0.87 for men). There were also strong correlations between poverty and mortality from stomach cancer (0.63 for women and 0.68 for men), all other cancers (0.59 and 0.62) and liver cancer (0.65 for men for whom it was analysed). The probabilities of dying from breast (0.02), prostate (0.16) pancreatic (0.04 and -0.01), lymphoma and multiple myeloma (0.00 and -0.07), leukaemia (-0.06 and 0.18), and colorectal cancer in women (0.09) showed little or no association with poverty. Those in poor districts of London had lower probabilities of dying from lung, colorectal, oesophageal, bladder (men), and all other cancers than in comparably deprived districts in the rest of the country (Figure 8.5). This London advantage dampened the association with poverty for mortality from lung cancer and all other cancers in particular, and increased the extent of variability across districts at higher levels of poverty.

8.2.3 Change from 2002 to 2019

The probability of dying from a cancer before 80 years of age declined from 2002 to 2019 in every district for both sexes, with median reductions across districts of 18.7% for women and 23.4% for men (Figure 8.6). Districts in London achieved the largest declines. Among cause categories, the largest reductions in mortality were for stomach cancer, with the declines in probability of dying before 80 years of age ranging between 39.1% and 57.2% for women across districts, and between 51.5% and 58.8% for men from 2002 to 2019, all with posterior probability >0.99 . The probability of dying from oesophageal cancer also decreased in every district for women, but varied for men from a 42.9% (27.7% to 56.0%) decrease in Plymouth to a 7.0% (-19.8% to 42.4%) increase in Gosport (Figure 8.6 and Figure E.11). The posterior probability that the district saw a decline in mortality from oesophageal cancer was >0.80 in 242 (77.1%) districts in men. Lung cancer mortality decreased everywhere for men with posterior probability >0.99 , but for women, there were mixed trends. The largest declines were in London, the strongest seen in Newham (decrease



(a) Women



(b) Men

Figure 8.5: Local authority district probability of death in 2019 in relation to poverty for each of the leading cancers. The points are coloured by the regions in England. The Spearman rank correlations (ρ) between the probability of dying from the cancer and poverty ranking are in the top left of each panel.

of 29.5% (18.5% to 38.8%)), whereas the probability of dying increased in many districts in the East of England, with the largest increase of 27.0% (6.8% to 49.7%) in Tendring (Figure 8.6 and Figure E.9). The posterior probability that the district observed a decline in mortality from lung cancer in women was >0.80 in 197 (62.7%) districts. Women in 4 (1.3%) districts experienced an increase in lung cancer mortality with a posterior probability >0.80 , and in the remaining 113 (36.0%) districts there was no clear trend at this level of posterior probability.

The probability of dying from pancreatic cancer increased (notably in the South East (Figure 8.6); median increase across districts of 8.6% for women and 6.7% for men) in all but one sex-district combination (Tower Hamlets for women; decrease of 1.0% (-12.1% to 15.7%)). Liver cancer for men (median increase 78.3%) saw increases in mortality in all districts with posterior probability >0.98 , and corpus uteri cancer for women (median increase 31.7%) saw increases in mortality in all but one district with $PP > 0.80$. There were increases in liver cancer mortality in excess of 100% in 32 (10.2%) districts, mostly in northern England (Figure 8.6).

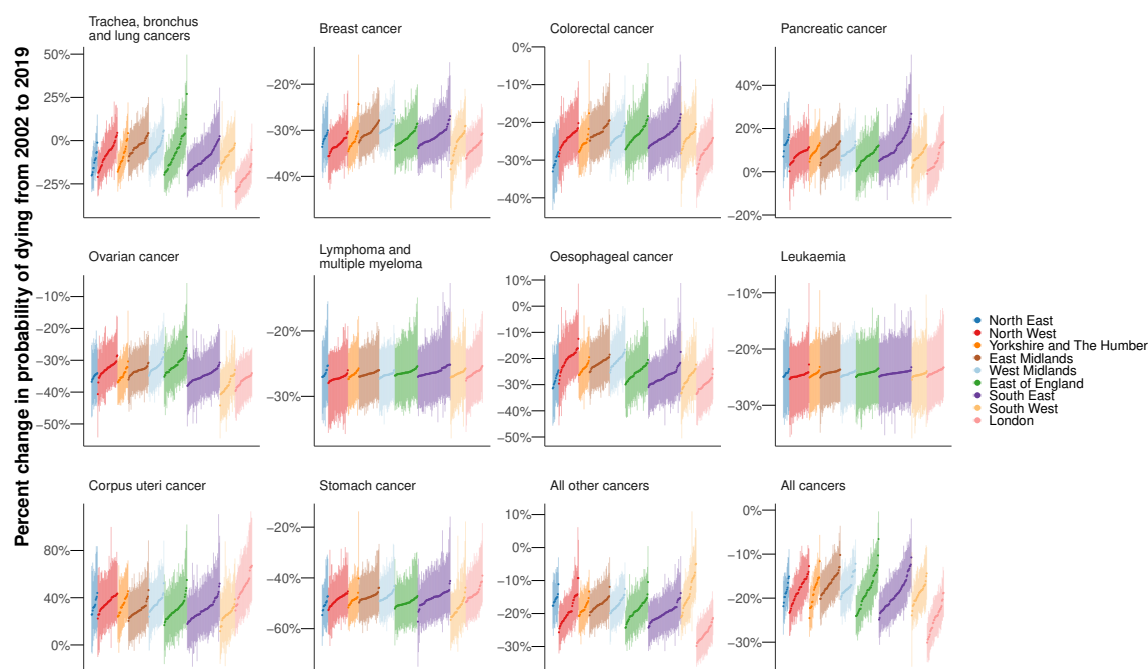
8.3 Discussion

This small-area analysis of deaths from leading cancers in England revealed that although overall cancer mortality decreased everywhere, the gains were unequal, with the largest declines almost five times that of the smallest. Preventable cancers – those with strong links to behavioural and environmental risk factors – had both the largest inequalities in mortality across districts in 2019 and the widest variation in change in mortality from 2002 to 2019. The cancers showing the least inequality across districts, lymphoma and multiple myeloma, and leukaemia, were those with weaker and more heterogeneous links to modifiable risk factors.

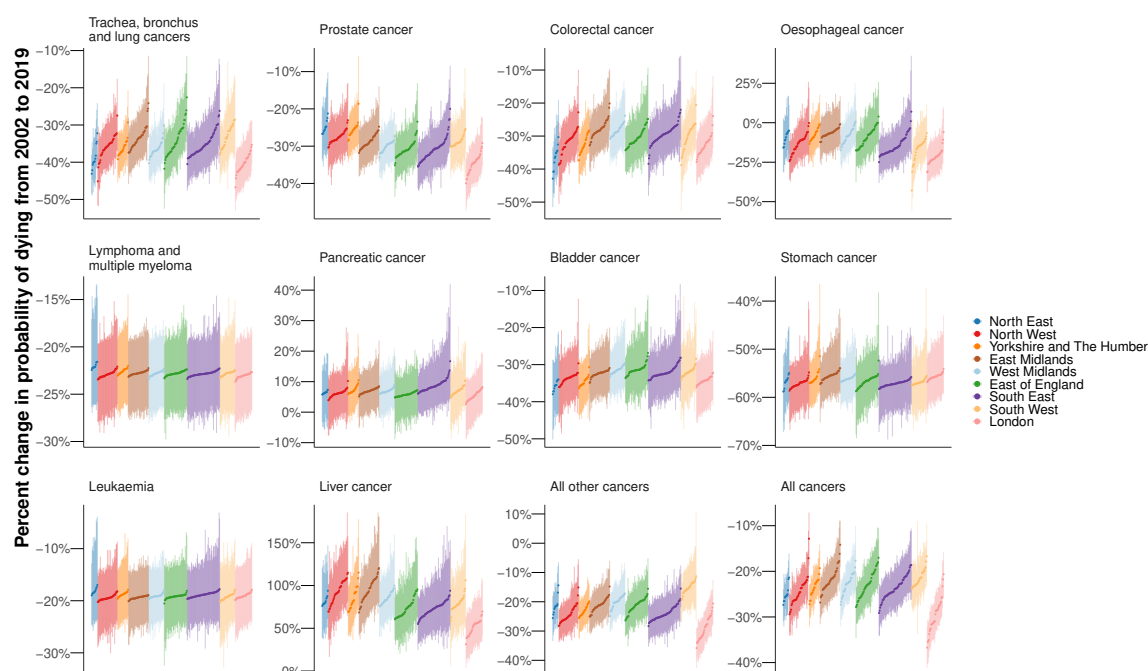
8.3.1 Strengths and limitations

A strength of our study is the presentation of mortality from leading cancers, as well as all cancers together, for subnational spatial units in England over a period of substantial change in public health programmes and healthcare technology and provision.

The study suffers the same limitations as Chapter 7: the residual group could have been split into more specific causes of death; each cause group was analysed separately rather than in a single joint model; the probability of dying between birth and 80 years masks mortality differences in



(a) Women



(b) Men

Figure 8.6: Ranked change in probability of dying between birth and 80 years of age in 314 local authority districts in England in 2002 and 2019 for the ten leading cancers. The points are coloured by the regions in England. Each point shows one district and the vertical line going through the point its 95% credible interval. The districts were ranked first by region and then by the posterior median estimate for the probability of dying.

the oldest age groups; the analysis was carried out at the district level rather than for MSOAs; the study did not account for the effects of between-district migration. Beyond this, I did not decompose mortality into incidence and survival, which would provide further insight into potential drivers of variations of mortality across site-specific cancers and geography, and should be the subject of further research.

8.3.2 Comparison with previous work

Few studies have looked at the geography of cancer mortality. In England, there have been subnational estimates for cancer survival, but these have been at large administrative units based on the organisation of the NHS (NHS Digital, 2022; Walters et al., 2011). There are mortality estimates for cancers generated by aggregating several years of data, again for large NHS boundaries rather than districts (National Cancer Intelligence Network, 2010; NHS Digital, 2022). The only study for multiple cancers at a finer resolution than district level was SAHSU's environment and health atlas, which reported risk relative to the national population in census wards (Hansell et al., 2014). The analysis aggregated data over a 25-year period and hence did not analyse variations over time. The relative risk estimates for lung cancer shared a similar geographical pattern to our results on mortality, with the highest risk in the urban North West and North East.

Two studies investigated trends in cancer mortality, although with limited resolution, either at the level of upper-tier local authorities, of which there are 150, or aggregating communities into deciles of deprivation within each region. Both studies similarly found that areas in London were exceptional, with generally lower rates of premature cancer mortality than expected for their level of deprivation (Steel et al., 2018), attenuated inequality between the most and least deprived deciles, and lower mortality for prostate cancer (Arik et al., 2022, 2021). The latter study, which only focussed on the four leading cancers, reported that in the worst performing regions inequality between the top and bottom deciles in female lung cancer mortality was similar to the 3.7-fold gap between the top and bottom districts seen in this study.

Beyond the UK, cancer atlases have been largely limited to coarse geographical units and aggregate multiple years of data rather than presenting information on trends. There have been examples of cancer atlases for smaller areas for Spain and Portugal (Fernández-Navarro et al.,

2021) and Australia (Cancer Council Queensland, 2018). A study in the USA on trends in cancer mortality at the county level found that lung cancer and stomach cancer were among the cancers with the largest inequalities in mortality across counties in 2014, but there was limited variation for leukaemia, lymphoma and multiple myeloma, which is consistent with the results presented here (Mokdad et al., 2017). The study also found that the strongest increases over time were in liver cancer mortality, with nearly all counties seeing an increase.

8.3.3 Explaining the variation and implications

Cancer mortality depends on both incidence and survival. Incidence is influenced by three modifiable factors: exposure to risk factors associated with cancer, including smoking, alcohol use, obesity, infections, and occupational and environmental exposures; the uptake of preventative treatments such as vaccinations; and the implementation of screening programmes for pre-cancerous lesions and early-stage cancers. In England, smoking, alcohol use and obesity tend to be higher where cancer mortality is highest, generally in the North (NHS Digital, 2023). The dominant risk factor in the UK remains tobacco smoking (Parkin et al., 2011), which has declined substantially since the 1960s, when it was the target of public health policy, to a prevalence among adults of 13.9% in 2019 (NHS Digital, 2023). The rise and fall in female smoking rates and smoking-attributable mortality have lagged behind that in men by about 20-30 years (Thun et al., 2012), leading to the mix of increase and decrease in female lung cancer mortality observed here, with similar patterns observed for COPD in Chapter 7. Similarly to smoking rates, after reaching a peak in the mid-2000s, alcohol consumption has declined (NHS Digital, 2023). In contrast to trends in smoking and alcohol use, obesity rates have steadily risen over the past few decades from 14.9% of adults in 1993 to 28.0% in 2019, mirrored by a rise in diabetes (NHS Digital, 2023). These trends could be partly responsible for cancers associated with diabetes and obesity increasingly co-occurring with those that are due to smoking and alcohol use, especially in men among whom smoking rates declined earlier.

I observed substantial increases in liver cancer mortality for men, especially in northern districts, which mirror a long-run trend in the USA where, in nearly all counties, there have been increases in liver cancer mortality (Mokdad et al., 2017). The major risk factors for liver cancer are infections, smoking, alcohol use and diabetes (Parkin et al., 2011). In particular, alcohol consumption has driven a large proportion of liver disease throughout Europe (Blachier et al., 2013), and the

relative risk due to diabetes of liver cancer incidence is more than two, which is the highest of any site-specific cancer (Tsilidis et al., 2015). The observed trends in liver cancer mortality could be linked to the heterogeneous trends in these risk factors, which have not been fully offset by reductions in smoking rates. Oesophageal cancer is attributed to tobacco smoking as well as other lifestyle and environmental factors such as alcohol consumption, poor diet, and obesity (Parkin et al., 2011). Differences between men and women in the temporal and geographical distribution of these risk factors may potentially account for our observation that oesophageal cancer mortality decreased in all districts for women, but increased in some districts for men. Stomach cancer exhibited the largest decreases in mortality, as well as substantial inequality in the level across districts. In part this may be due to variable prevalence of, and heterogeneous decrease in, *H pylori* infection across districts.

Survival is influenced by patient and healthcare factors including cancer awareness and utilisation of care, barriers to screening or early diagnosis, tumour factors (stage, topography, histology), comorbidities, and the quality of care. The UK, as a whole, has performed badly in terms of survival compared to other European countries (Walters et al., 2015). There were also subnational variations in survival, both in the cancers with the largest inequalities in mortality in our results (i.e. stomach and lung cancer) and for other cancers including oesophageal cancer, ovarian cancer, and leukaemia (NHS Digital, 2022). The UK has national screening programmes for breast, colorectal and cervical cancers. Cervical cancer was not in the top ten leading cancers for women by number of deaths, and so was included in the residual group. Both breast and colorectal cancers saw mortality declines in all districts, albeit at varying rates. Breast and colorectal cancer mortality in women both showed relatively low geographical inequality and no correlation with poverty, while there was moderate correlation between colorectal cancer mortality for men and poverty. The coverage of the screening programmes varies between districts, ranging from 51.2% to 79.4% for colorectal and 40.2% to 79.3% for breast (Public Health England, 2023). Despite these specific screening programmes, across all cancer sites, the proportion of cancers diagnosed at stages 1 and 2 varied from 42.0% to 61.3% (Public Health England, 2023).

Combinations of risk factors and healthcare can also explain London's better performance for some cancers. Firstly, London, and particularly the east of London, is the most ethnically diverse region in the country. Smoking rates amongst Asian (8.3% of adults) and black (9.7%) groups are lower

than the white ethnic group (14.4%) (Office for National Statistics, 2021d). Secondly, there could be differences in the quality of healthcare, either through the availability of innovative treatments such as immunotherapies, or, as London has a much higher density of hospitals, patients might be more able to travel to a specialist hospital rather than their local hospital compared to larger regions, where the hospital might be much less accessible and further from support networks. Unlike patterns in other cancers however, districts in East London had the highest mortality for corpus uteri cancer, which has higher incidence rates in the black ethnic group in England (Delon et al., 2022).

Since the turn of the century, two major national policy initiatives in the UK have aimed to improve cancer survival. Firstly, the NHS Cancer Plan of 2000 specifically targeted socioeconomic inequalities in cancer survival through an increase in expenditure, a focus on centralisation and specialisation, and a greater use of multi-disciplinary teams. Secondly, launched in 2008, the National Awareness and Early Diagnostic Initiative, addressed both patient factors, such as reducing the stigma around cancer and the barriers to seeing a doctor, and tumour factors, such as improving access to screening and optimising referral pathways so patients present at a less advanced stage. However, no consistent evidence was found of a direct impact on one-year survival after these successive initiatives (Exarchakou et al., 2018; Rachet et al., 2010), and cancer awareness explains little of the socioeconomic disparities in survival (Niksic et al., 2016). Recent NHS plans have renewed focus on survival targets and early diagnosis (NHS, 2019), as well as the roll out of a national lung cancer screening programme targeted at those with a history of smoking between the ages of 55 and 74 years (Department of Health and Social Care, 2023). The present results show heterogeneous trends in mortality, particularly for preventable cancers. To ensure these inequalities in mortality are reduced, factors affecting both incidence and survival at the local level must be considered.

8.4 Summary

I performed a spatiotemporal analysis of vital registration data on deaths from the ten leading cancers by sex in England from 2002 to 2019. I used life table methods to calculate the probability of dying between birth and 80 years of age by sex, cancer cause of death, district and year.

The probability of dying from birth to 80 years from a cancer declined in all districts from 2002

to 2019, but at varying rates: for women, the greatest reduction was nearly five times that of the smallest; for men, the largest decrease was triple that of the smallest. The most unequal cancers were lung cancer for women (3.7-fold variation in mortality across districts) and stomach cancer for men (3.3-fold). The cancers with the least geographical variability were lymphoma and multiple myeloma (1.2-fold for both sexes) and leukaemia (1.1-fold for women and 1.2-fold for men). Unlike in other regions, in London, mortality from lung, colorectal, oesophageal, and bladder cancer for men was not associated with district level poverty.

Preventable cancers with modifiable risk factors showed heterogeneous trends and the greatest geographical inequality. For some of these cancers, poorer districts in London saw lower cancer mortality than poor areas elsewhere in the country, suggesting that some features of London weaken the relationship between poverty and mortality.

Chapter 9

Discussion

9.1 Public health and policy implications of findings

Although previous studies have observed a decline in female life expectancy in the bottom two deciles of deprivation in the past decade (Bennett et al., 2018; Marmot et al., 2020), this thesis has shown both for men and women where the declines have been happening for small areas. Not only have there been declines, but the declines have been accelerating in the latter half of the decade, such that in the period 2014-19, life expectancy fell in 1 in 5 MSOAs for women and 1 in 9 for men. These trends happened in the years before the mortality-forcing event of the Covid-19 pandemic and are worse than comparable countries (Leon et al., 2019), and should be deemed a failure for public health policy. Although the complexity of policy means it cannot be proved causally, leaders within the NHS have pointed towards constrained funding for the healthcare body itself, as well as cuts to wider social determinants, such as housing and education, as the main reason for the country's poor performance (Ham, 2023).

The declines in life expectancy were sustained over a long period of time, which serves as another example where death rates for some population subgroups run contrary to the persistent mortality decline of the third stage of the Epidemiologic Transition theory, as discussed in Gaylin and Kates (1997) with the HIV/AIDS pandemic. Even if England is in the hypothesised fourth stage of the transition, the Age of Delayed Degenerative diseases (Olshansky and Ault, 1986), there are subnational patterns where degenerative diseases are *not* killing at later and later ages. This supports the growing evidence that Omran's theory acts only as a useful heuristic, and that there

is a large amount of variability, particularly in the latter stages of the transition, between broad geographical regions (Mackenbach, 2022; Sudharsanan et al., 2022).

The difference in progress between districts in the last decade was largely driven by differences in degenerative diseases - in particular, the rate of improvement for CVDs and all other NCDs, and the strength of the negative forcing effect of Alzheimer's and other dementias. Furthermore, female mortality from infectious, maternal, perinatal and nutritional conditions (GBD group 1), which dominate the second stage of the transition, increased in many districts. There is also a worrying shift towards injuries (GBD group 3) contributing negatively towards life expectancy progress, particularly for men. This is possibly driven by a rise in "deaths of despair" (Angus et al., 2023; Case and Deaton, 2015), although this would require further analysis by separating intentional and unintentional injuries. It should be noted that IMPN and injuries do not play a major role in total mortality as they accounted for only 11.1% of all deaths from 2002 to 2019, but they are almost entirely preventable and should be addressed through appropriate policy.

Cancer survival outcomes in the UK are worse than those in Europe in general (OECD, 2016). In this thesis, I found mortality showed huge spatial variation for a number of preventable cancers. Worryingly, at the same level of poverty, London performed significantly better than the rest of the country for several cancers. This suggests either regional differences in the quality of care, or differences in the populations living in deprived areas, which should both be the subject of policy discussions.

Finally, this thesis emphasises the value of small area work in informing policy. National trends in mortality are not spatially homogenous. Particular diseases have shown massive variation at the district level, such as COPD in women (6.0-fold in 2019), and by studying England at the MSOA level, I have uncovered the widest subnational gap in life expectancy of 27.0 years (men in 2019) in the literature. This rich source of data allows policymakers to work with local authorities to create targeted public health interventions. The population issues at the LSOA level in London and the coding of CVDs on the Isle of Wight show there are still limitations of the data. Nevertheless, the estimates from this thesis are already being used in the press to provide context for recent falls in US life expectancy (Burn-Murdoch, 2023), and I hope they are also being used in policy discussions.

9.2 Future work

There are many possible methodological and substantive extensions to the work presented in this thesis. Firstly, on the methodological side, with improvements both to hardware and the rise of approximate inference algorithms to replace computationally costly sampling methods, the models can, in theory, be scaled to higher and higher spatial resolutions and we could potentially estimate mortality for the entire country for LSOAs, OAs, or even postcodes. However, given the data issues in Chapter 5 and the very wide credible intervals seen for some LSOAs with high life expectancies, perhaps smaller is not better when the quality of the data is lacking.

One of the major strengths of the thesis was the use of Bayesian methods, at one of the highest spatial resolutions in the literature for a model estimating mortality. This was largely thanks to recent developments in probabilistic programming, which allow sampling algorithms to run on GPUs rather than CPUs, and is generally faster for models with over 10,000 parameters (Lao et al., 2020). The models themselves, which were based on previous work in the group (Bennett et al., 2018, 2015), followed traditional statistical approaches using intercepts and linear slopes and adding in random walk effects for nonlinearities. However, with models such as Equation 4.4, we are actually only interested in the left-hand side of the equation – the death rate – and not at all in the values of the parameters on the right-hand side. There are a vast array of flexible models which can describe complex data without needing to design heavily parametrised structures. These range from Gaussian processes (Flaxman et al., 2015; Rasmussen and Williams, 2006), where the user can use inductive biases and their own knowledge to influence how the model fits the data, to wilder options like neural networks, which completely sacrifice any remaining interpretability of the effects.

During my PhD, I became very familiar with the literature on Gaussian processes, and was involved with a paper aimed at speeding up inference in spatial settings (Semenova et al., 2022). One of the most promising avenues for the task of mortality estimation in this thesis, outlined in Flaxman et al. (2015), is the use of Kronecker-structured covariance matrices to create hierarchical Gaussian processes. In the context of this thesis, the covariance matrix would be defined as the Kronecker product of three smaller matrices: one for age effects, one for spatial effects, and one for temporal effects. We can exploit the Kronecker structure for faster inference, as we no longer need to invert the full matrix (an operation that scales cubically with the number of data points), but rather the

smaller constituent matrices of each effect. Ultimately, nothing came to fruition as the working environment, due to the sensitivity of the data,¹ was not conducive to developing novel statistical methods. Also, there was no guarantee the Gaussian process methods would be quicker, as it still requires the inversion of a large spatial matrix, nor would it produce a result much different from the original model. Instead, I allocated time to scaling existing models and testing new hardware. I believe the hierarchical Gaussian process model should be the starting point for future research because it is much more flexible and requires less testing of each effect than the models in this thesis.

Another methodological extension would be to model causes of death jointly, thereby borrowing strength across *causes*. Studies have built spatial models which use shared components between diseases, either modelling spatial patterns of any number of diseases using a weighted sum of shared components (Best et al., 2005; Held et al., 2005; Knorr-Held and Best, 2001) or pre-assigning the spatial components to the diseases based on knowledge of common risk factors (Downing et al., 2008; Mahaki et al., 2018). However, as the number of diseases in the study increases, more spatial components must be fitted, which is computationally prohibitive. Furthermore, unless the spatial components are defined *a priori*, there is a question of how many spatial components are required to parsimoniously describe the variance in the data. Alternatively, the correlations between the causes can be modelled using a multivariate normal distribution of the same dimensions as the number of cause groups, as in Foreman et al. (2017). Extending this idea, there is also the option to extend the hierarchical Gaussian process described above with another component covariance matrix describing the correlations between each of the cause groups. Ultimately, the goal would be to run a single joint model that flexibly describes interactions between sexes, age groups, spatial units, time, and causes of death.

On the substantive side, there is a reason the final year of study period was restricted at 2019 that has nothing to do with the availability of vital registration data for the proceeding years: the Covid-19 pandemic. It would be naïve to fit the same model to an extended time period that includes the pandemic years because the pandemic had such a major influence on mortality patterns. The model relies heavily on linear trends, which would be affected by the final years.

¹Working on identifiable health data requires a secure computing infrastructure. In my case, the cluster had no access to the internet and only a command line interface, which hampered method development. It is understandable why spatial statisticians continue to study the standard Scottish lip cancer dataset.

Further, pandemics are fat-tailed events, so the *Gaussian* random walk effects would struggle to capture the shock without increasing the overall variance of the effect.

Rather than extending the study to include the pandemic years and look at the trends between 2002 and 2023, a more interesting question is “had the pandemic *not* happened, how would mortality patterns be different?” This would be done using a counterfactual analysis: train a model to estimate annual mortality from 2002 to 2019 and produce a forecast for the years from 2020. A team, including myself, has already done this analysis at the national level at a weekly temporal resolution (Kontis et al., 2022, 2020). Here, the task would instead involve forecasts for each sex-age-space-time-cause stratum on annual timescales. The group has been involved in a number of mortality forecasting projects (Bennett et al., 2015; Foreman et al., 2017; Kontis et al., 2017). The models are similar to those in this thesis, but would need to be adapted and trained specifically for forecasting purposes, with the possibility of averaging multiple models to reduce bias. The results would allow us to see how the cause-composition of mortality has changed compared to the *business as usual* scenario. We would obviously expect an increase in deaths from infectious diseases, but were injuries reduced as the country was locked down in their homes? Did cancer outcomes worsen due to missed surgeries as hospitals were flooded with Covid-19 patients? Is there a longer-term effect due to strain on emergency services, which has manifested in longer waiting lists and waiting times for appointments (Dorling, 2023)? And, how have these effects varied for different age groups and different areas of the country?

References

- Alexander M, Zagheni E, Barbieri M. 2017. A Flexible Bayesian Model for Estimating Subnational Mortality. *Demography* **54**:2025–2041. doi:[10.1007/s13524-017-0618-7](https://doi.org/10.1007/s13524-017-0618-7)
- Alexiou A, Fahy K, Mason K, Bennett D, Brown H, Bamba C, Taylor-Robinson D, Barr B. 2021. Local government funding and life expectancy in England: A longitudinal ecological study. *The Lancet Public Health* **6**:e641–e647. doi:[10.1016/S2468-2667\(21\)00110-9](https://doi.org/10.1016/S2468-2667(21)00110-9)
- Angus C, Buckley C, Tilstra AM, Dowd JB. 2023. Increases in “deaths of despair” during the COVID-19 pandemic in the United States and the United Kingdom. *Public Health* **218**:92–96. doi:[10.1016/j.puhe.2023.02.019](https://doi.org/10.1016/j.puhe.2023.02.019)
- Arík A, Cairns A, Dodd E, Shao A, Streftaris G. 2022. Uneven outcomes: Findings on cancer mortality. *The Actuary*.
- Arík A, Dodd E, Cairns A, 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**:e0253854. doi:[10.1371/journal.pone.0253854](https://doi.org/10.1371/journal.pone.0253854)
- Arriaga EE. 1984. Measuring and Explaining the Change in Life Expectancies. *Demography* **21**:83–96. doi:[10.2307/2061029](https://doi.org/10.2307/2061029)
- Asaria P, Elliott P, Douglass M, Obermeyer Z, Soljak M, Majeed A, Ezzati M. 2017. Acute myocardial infarction hospital admissions and deaths in England: A national follow-back and follow-forward record-linkage study. *The Lancet Public Health* **2**:e191–e201. doi:[10.1016/S2468-2667\(17\)30032-4](https://doi.org/10.1016/S2468-2667(17)30032-4)
- Asaria P, Fortunato L, Fecht D, Tzoulaki I, Abellan JJ, Hambly P, de Hoogh K, Ezzati M, Elliott P. 2012. Trends and inequalities in cardiovascular disease mortality across 7932 English electoral wards, 1982–2006: Bayesian spatial analysis. *International Journal of Epidemiology* **41**:1737–1749. doi:[10.1093/ije/dys151](https://doi.org/10.1093/ije/dys151)
- Aylin P, Maheswaran R, Wakefield J, Cockings S, Jarup L, Arnold R, Wheeler G, Elliott P. 1999. A national facility for small area disease mapping and rapid initial assessment of apparent

- disease clusters around a point source: The UK Small Area Health Statistics Unit. *Journal of Public Health* **21**:289–298. doi:[10.1093/pubmed/21.3.289](https://doi.org/10.1093/pubmed/21.3.289)
- Bahk J, Kang H-Y, Khang Y-H. 2020. Life expectancy and inequalities therein by income from 2016 to 2018 across the 253 electoral constituencies of the National Assembly of the Republic of Korea. *Journal of Preventive Medicine and Public Health* **53**:143–148. doi:[10.3961/jpmph.20.050](https://doi.org/10.3961/jpmph.20.050)
- Barr B, Higgerson J, Whitehead M. 2017. Investigating the impact of the English health inequalities strategy: Time trend analysis. *BMJ* **358**:j3310. doi:[10.1136/bmj.j3310](https://doi.org/10.1136/bmj.j3310)
- Beltrán-Sánchez H, Finch CE, Crimmins EM. 2015. Twentieth century surge of excess adult male mortality. *Proceedings of the National Academy of Sciences* **112**:8993–8998. doi:[10.1073/pnas.1421942112](https://doi.org/10.1073/pnas.1421942112)
- Bennett JE, Li G, Foreman K, Best N, Kontis V, Pearson C, Hambly P, Ezzati M. 2015. The future of life expectancy and life expectancy inequalities in England and Wales: Bayesian spatiotemporal forecasting. *The Lancet* **386**:163–170. doi:[10.1016/S0140-6736\(15\)60296-3](https://doi.org/10.1016/S0140-6736(15)60296-3)
- Bennett JE, Pearson-Stuttard J, Kontis V, Capewell S, Wolfe I, 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. *The Lancet Public Health* **3**:e586–e597. doi:[10.1016/S2468-2667\(18\)30214-7](https://doi.org/10.1016/S2468-2667(18)30214-7)
- Bennett JE, Rashid T, Zolfaghari A, Doyle Y, Suel E, Pearson-Stuttard J, Davies B, Fecht D, Muller ES, Nathvani RS, Sportiche N, Daby HI, Johnson E, Li G, Flaxman S, Toledano MB, Asaria M, Ezzati M. 2023. Changes in life expectancy and house prices in London from 2002 to 2019: Hyper-resolution spatiotemporal analysis of death registration and real estate data. *The Lancet Regional Health – Europe* **27**:100580. doi:[10.1016/j.lanepe.2022.100580](https://doi.org/10.1016/j.lanepe.2022.100580)
- Besag J, York J, Mollié A. 1991. Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics* **43**:1–20. doi:[10.1007/BF00116466](https://doi.org/10.1007/BF00116466)
- Best N, Richardson S, Thomson A. 2005. A comparison of Bayesian spatial models for disease mapping. *Statistical Methods in Medical Research* **14**:35–59. doi:[10.1191/0962280205sm388oa](https://doi.org/10.1191/0962280205sm388oa)
- Bilal U, Alazraqui M, Caiaffa WT, Lopez-Olmedo N, Martinez-Folgar K, Miranda JJ, Rodriguez DA, Vives A, Diez-Roux AV. 2019. Inequalities in life expectancy in six large Latin American cities from the SALURBAL study: An ecological analysis. *The Lancet Planetary Health* **3**:e503–e510. doi:[10.1016/S2542-5196\(19\)30235-9](https://doi.org/10.1016/S2542-5196(19)30235-9)

- Blachier M, Leleu H, Peck-Radosavljevic M, Valla D-C, Roudot-Thoraval F. 2013. The burden of liver disease in Europe: A review of available epidemiological data. *Journal of Hepatology* **58**:593–608. doi:[10.1016/j.jhep.2012.12.005](https://doi.org/10.1016/j.jhep.2012.12.005)
- Blundell R, Costa Dias M, Joyce R, Xu X. 2020. COVID-19 and Inequalities. *Fiscal Studies* **41**:291–319. doi:[10.1111/1475-5890.12232](https://doi.org/10.1111/1475-5890.12232)
- Bosetti N. 2015. Inside out: The new geography of wealth and poverty in London. Centre for London.
- Boulieri A, Blangiardo M. 2020. Spatio-temporal model to estimate life expectancy and to detect unusual trends at the local authority level in England. *BMJ Open* **10**:e036855. doi:[10.1136/bmjopen-2020-036855](https://doi.org/10.1136/bmjopen-2020-036855)
- Burn-Murdoch J. 2023. Why are Americans dying so young? *Financial Times*.
- Cancer Council Queensland. 2018. Australian Cancer Atlas.
- Case A, Deaton A. 2015. Rising morbidity and mortality in midlife among white non-Hispanic Americans in the 21st century. *Proceedings of the National Academy of Sciences* **112**:15078–15083. doi:[10.1073/pnas.1518393112](https://doi.org/10.1073/pnas.1518393112)
- Cheshire J. 2012. Featured Graphic. Lives on the Line: Mapping Life Expectancy along the London Tube Network. *Environment and Planning A: Economy and Space* **44**:1525–1528. doi:[10.1068/a45341](https://doi.org/10.1068/a45341)
- Chetty R, Stepner M, Abraham S, Lin S, Scuderi B, Turner N, Bergeron A, Cutler D. 2016. The Association Between Income and Life Expectancy in the United States, 2001–2014. *JAMA* **315**:1750–1766. doi:[10.1001/jama.2016.4226](https://doi.org/10.1001/jama.2016.4226)
- Coale AJ, Demeny P, Vaughan B. 1983. Regional Model Life Tables and Stable Populations, 2nd ed. Academic Press.
- Congdon P. 2019. Geographical Patterns in Drug-Related Mortality and Suicide: Investigating Commonalities in English Small Areas. *International Journal of Environmental Research and Public Health* **16**:1831. doi:[10.3390/ijerph16101831](https://doi.org/10.3390/ijerph16101831)
- Congdon P. 2014. Modelling changes in small area disability free life expectancy: Trends in London wards between 2001 and 2011. *Statistics in Medicine* **33**:5138–5150. doi:[10.1002/sim.6298](https://doi.org/10.1002/sim.6298)
- Congdon P. 2009. Life Expectancies for Small Areas: A Bayesian Random Effects Methodology. *International Statistical Review* **77**:222–240. doi:[10.1111/j.1751-5823.2009.00080.x](https://doi.org/10.1111/j.1751-5823.2009.00080.x)

- Connolly S, O'Reilly D, Rosato M. 2007. Increasing inequalities in health: Is it an artefact caused by the selective movement of people? *Social Science & Medicine* **64**:2008–2015. doi:[10.1016/j.socscimed.2007.02.021](https://doi.org/10.1016/j.socscimed.2007.02.021)
- Davenport A, Zarenko B. 2020. Levelling up: Where and how? IFS Green Budget 2020. Institute for Fiscal Studies.
- de Valpine P, Paciorek CJ, Turek D, Michaud N, Anderson-Bergman C, Obermayer F, Wehrhahn Cortes C, Rodríguez A, Lang DT, Paganin S, Hug J. 2022. NIMBLE: MCMC, Particle Filtering, and Programmable Hierarchical Modeling.
- de Valpine P, Turek D, Paciorek CJ, Anderson-Bergman C, Lang DT, Bodik R. 2017. Programming With Models: Writing Statistical Algorithms for General Model Structures With NIMBLE. *Journal of Computational and Graphical Statistics* **26**:403–413. doi:[10.1080/10618600.2016.1172487](https://doi.org/10.1080/10618600.2016.1172487)
- Delon C, Brown KF, Payne NWS, Kotrotsios Y, Vernon S, Shelton J. 2022. Differences in cancer incidence by broad ethnic group in England, 2013–2017. *British Journal of Cancer* **126**:1765–1773. doi:[10.1038/s41416-022-01718-5](https://doi.org/10.1038/s41416-022-01718-5)
- Department of Health and Social Care. 2023. New lung cancer screening roll out to detect cancer sooner. GOVUK.
- Dorling D. 2023. How austerity caused the NHS crisis. *openDemocracy*.
- Downing A, Forman D, Gilthorpe MS, Edwards KL, Manda SO. 2008. Joint disease mapping using six cancers in the Yorkshire region of England. *International Journal of Health Geographics* **7**:41. doi:[10.1186/1476-072X-7-41](https://doi.org/10.1186/1476-072X-7-41)
- Duncan EW, White NM, Mengersen K. 2017. Spatial smoothing in Bayesian models: A comparison of weights matrix specifications and their impact on inference. *International Journal of Health Geographics* **16**:47. doi:[10.1186/s12942-017-0120-x](https://doi.org/10.1186/s12942-017-0120-x)
- Dwyer-Lindgren L, Bertozzi-Villa A, Stubbs RW, Morozoff C, Mackenbach JP, van Lenthe FJ, Mokdad AH, Murray CJL. 2017a. Inequalities in Life Expectancy Among US Counties, 1980 to 2014: Temporal Trends and Key Drivers. *JAMA Internal Medicine* **177**:1003–1011. doi:[10.1001/jamainternmed.2017.0918](https://doi.org/10.1001/jamainternmed.2017.0918)
- Dwyer-Lindgren L, Stubbs RW, Bertozzi-Villa A, Morozoff C, Callender C, Finegold SB, Shirude S, Flaxman AD, Laurent A, Kern E, Duchin JS, Fleming D, Mokdad AH, Murray CJL. 2017b. Variation in life expectancy and mortality by cause among neighbourhoods in King County,

- WA, USA, 1990–2014: A census tract-level analysis for the Global Burden of Disease Study 2015. *The Lancet Public Health* **2**:e400–e410. doi:[10.1016/S2468-2667\(17\)30165-2](https://doi.org/10.1016/S2468-2667(17)30165-2)
- Eayres D, Williams ES. 2004. Evaluation of methodologies for small area life expectancy estimation. *Journal of Epidemiology & Community Health* **58**:243–249. doi:[10.1136/jech.2003.009654](https://doi.org/10.1136/jech.2003.009654)
- Elliott P, Briggs D, Morris S, Hoogh C de, Hurt C, Jensen TK, Maitland I, Richardson S, Wakefield J, Jarup L. 2001a. Risk of adverse birth outcomes in populations living near landfill sites. *BMJ* **323**:363–368. doi:[10.1136/bmj.323.7309.363](https://doi.org/10.1136/bmj.323.7309.363)
- Elliott P, Toledano MB, Bennett JE, Beale L, Hoogh C de, Best N, Briggs D. 2010. Mobile phone base stations and early childhood cancers: Case-control study. *BMJ* **340**:c3077. doi:[10.1136/bmj.c3077](https://doi.org/10.1136/bmj.c3077)
- Elliott P, Wakefield J, Best N, Briggs D. 2001b. Spatial epidemiology: Methods and applications. Oxford University Press.
- Elliott P, Westlake AJ, Hills M, Kleinschmidt I, Rodrigues L, McGale P, Marshall K, Rose G. 1992. The Small Area Health Statistics Unit: A national facility for investigating health around point sources of environmental pollution in the United Kingdom. *Journal of Epidemiology & Community Health* **46**:345–349. doi:[10.1136/jech.46.4.345](https://doi.org/10.1136/jech.46.4.345)
- Exarchakou A, Rachet B, Belot A, Maringe C, Coleman MP. 2018. Impact of national cancer policies on cancer survival trends and socioeconomic inequalities in England, 1996-2013: Population based study. *BMJ* **360**:k764. doi:[10.1136/bmj.k764](https://doi.org/10.1136/bmj.k764)
- Ezzati M, Friedman AB, Kulkarni SC, Murray CJL. 2008. The Reversal of Fortunes: Trends in County Mortality and Cross-County Mortality Disparities in the United States. *PLOS Medicine* **5**:e66. doi:[10.1371/journal.pmed.0050066](https://doi.org/10.1371/journal.pmed.0050066)
- Ezzati M, Obermeyer Z, Tzoulaki I, Mayosi BM, Elliott P, Leon DA. 2015. Contributions of risk factors and medical care to cardiovascular mortality trends. *Nature Reviews Cardiology* **12**:508–530. doi:[10.1038/nrcardio.2015.82](https://doi.org/10.1038/nrcardio.2015.82)
- Fernández-Navarro P, Nuñez O, García-Pérez J, López-Abente G, Roquette R, de Sousa-Uva M, Nunes B, Dias CM. 2021. Atlas of Cancer Mortality in Portugal and Spain 2003-12. National Institute of Health Doutor Ricardo Jorge and National Institute of Health Carlos III.
- Finucane MM, Paciorek CJ, Danaei G, Ezzati M. 2014. Bayesian Estimation of Population-Level Trends in Measures of Health Status. *Statistical Science* **29**:18–25. doi:[10.1214/13-STS427](https://doi.org/10.1214/13-STS427)

- Flaxman S, Gelman A, Neill D, Smola A, Vehtari A, Wilson AG. 2015. Fast hierarchical Gaussian processes.
- Foreman KJ, Li G, Best N, Ezzati M. 2017. Small area forecasts of cause-specific mortality: Application of a Bayesian hierarchical model to US vital registration data. *Journal of the Royal Statistical Society Series C (Applied Statistics)* **66**:121–139. doi:[10.1111/rssc.12157](https://doi.org/10.1111/rssc.12157)
- Gaylin DS, Kates J. 1997. Refocusing the lens: Epidemiologic transition theory, mortality differentials, and the AIDS pandemic. *Social Science & Medicine* **44**:609–621. doi:[10.1016/S0277-9536\(96\)00212-2](https://doi.org/10.1016/S0277-9536(96)00212-2)
- Gersten O, Wilmoth JR. 2002. The Cancer Transition in Japan since 1951. *Demographic Research* **7**:271–306. doi:[10.4054/DemRes.2002.7.5](https://doi.org/10.4054/DemRes.2002.7.5)
- Gonzaga MR, Schmertmann CP. 2016. Estimating age- and sex-specific mortality rates for small areas with TOPALS regression: An application to Brazil in 2010. *Revista Brasileira de Estudos de População* **33**:629–652. doi:[10.20947/S0102-30982016c0009](https://doi.org/10.20947/S0102-30982016c0009)
- Goodman PS. 2018. In Britain, Austerity Is Changing Everything. *The New York Times*.
- Gray M, Barford A. 2018. The depths of the cuts: The uneven geography of local government austerity. *Cambridge Journal of Regions, Economy and Society* **11**:541–563. doi:[10.1093/cjres/rsy019](https://doi.org/10.1093/cjres/rsy019)
- Halonen JI, Hansell AL, Gulliver J, Morley D, Blangiardo M, Fecht D, Toledano MB, Beevers SD, Anderson HR, Kelly FJ, Tonne C. 2015. Road traffic noise is associated with increased cardiovascular morbidity and mortality and all-cause mortality in London. *European Heart Journal* **36**:2653–2661. doi:[10.1093/eurheartj/ehv216](https://doi.org/10.1093/eurheartj/ehv216)
- Ham C. 2023. The Rise and Decline of the NHS in England 2000–20. The King's Fund.
- Hansell AL, Beale L, Ghosh RE, Fortunato L, Fecht D, Jarup L, Elliott P. 2014. The Environment and Health Atlas for England and Wales. Oxford University Press.
- Hansell AL, Blangiardo M, Fortunato L, Floud S, Hoogh K de, Fecht D, Ghosh RE, Laszlo HE, Pearson C, Beale L, Beevers S, Gulliver J, Best N, Richardson S, Elliott P. 2013. Aircraft noise and cardiovascular disease near Heathrow airport in London: Small area study. *BMJ* **347**:f5432. doi:[10.1136/bmj.f5432](https://doi.org/10.1136/bmj.f5432)
- Hayat S, Luben R, Khaw K-T, Wareham N, Brayne C. 2022. Evaluation of routinely collected records for dementia outcomes in UK: A prospective cohort study. *BMJ Open* **12**:e060931. doi:[10.1136/bmjopen-2022-060931](https://doi.org/10.1136/bmjopen-2022-060931)

- Held L, Natário I, Fenton SE, Rue H, Becker N. 2005. Towards joint disease mapping. *Statistical Methods in Medical Research* **14**:61–82. doi:[10.1191/0962280205sm389oa](https://doi.org/10.1191/0962280205sm389oa)
- Hiam L, Dorling D, McKee M. 2020. Things Fall Apart: The British Health Crisis 2010–2020. *British Medical Bulletin* **133**:4–15. doi:[10.1093/bmb/ldz041](https://doi.org/10.1093/bmb/ldz041)
- Hiam L, Harrison D, McKee M, Dorling D. 2018. Why is life expectancy in England and Wales “stalling”? *J Epidemiol Community Health* **72**:404–408. doi:[10.1136/jech-2017-210401](https://doi.org/10.1136/jech-2017-210401)
- Hodgson S, Nieuwenhuijsen MJ, Colville R, Jarup L. 2007. Assessment of exposure to mercury from industrial emissions: Comparing “distance as a proxy” and dispersion modelling approaches. *Occupational and Environmental Medicine* **64**:380–388. doi:[10.1136/oem.2006.026781](https://doi.org/10.1136/oem.2006.026781)
- Hodgson S, Nieuwenhuijsen MJ, Hansell A, Shepperd S, Flute T, Staples B, Elliott P, Jarup L. 2004. Excess risk of kidney disease in a population living near industrial plants. *Occupational and Environmental Medicine* **61**:717–719. doi:[10.1136/oem.2003.010629](https://doi.org/10.1136/oem.2003.010629)
- Hoffman MD, Gelman A. 2014. The No-U-Turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research* **15**:1593–1623. doi:[10.48550/arXiv.1111.4246](https://doi.org/10.48550/arXiv.1111.4246)
- Jarup L, Best N, Toledano MB, Wakefield J, Elliott P. 2002a. Geographical epidemiology of prostate cancer in Great Britain. *International Journal of Cancer* **97**:695–699. doi:[10.1002/ijc.10113](https://doi.org/10.1002/ijc.10113)
- Jarup L, Briggs D, de Hoogh C, Morris S, Hurt C, Lewin A, Maitland I, Richardson S, Wakefield J, Elliott P. 2002b. Cancer risks in populations living near landfill sites in Great Britain. *British Journal of Cancer* **86**:1732–1736. doi:[10.1038/sj.bjc.6600311](https://doi.org/10.1038/sj.bjc.6600311)
- Jarup L, Morris S, Richardson S, Briggs D, Cobley N, de Hoogh C, Gorog K, Elliott P. 2007. Down syndrome in births near landfill sites. *Prenatal Diagnosis* **27**:1191–1196. doi:[10.1002/pd.1873](https://doi.org/10.1002/pd.1873)
- Jonker MF, van Lenthe FJ, Congdon PD, Donkers B, Burdorf A, Mackenbach JP. 2012. Comparison of Bayesian Random-Effects and Traditional Life Expectancy Estimations in Small-Area Applications. *American Journal of Epidemiology* **176**:929–937. doi:[10.1093/aje/kws152](https://doi.org/10.1093/aje/kws152)
- Kelsall J, Wakefield J. 2002. Modeling Spatial Variation in Disease Risk. *Journal of the American Statistical Association* **97**:692–701. doi:[10.1198/016214502388618438](https://doi.org/10.1198/016214502388618438)

- Knorr-Held L. 2000. Bayesian modelling of inseparable space-time variation in disease risk. *Statistics in Medicine* **19**:2555–2567. doi:[10.1002/1097-0258\(20000915/30\)19:17/18<2555::AID-SIM587>3.0.CO;2-%23](https://doi.org/10.1002/1097-0258(20000915/30)19:17/18<2555::AID-SIM587>3.0.CO;2-%23)
- Knorr-Held L, Best NG. 2001. [A Shared Component Model for Detecting Joint and Selective Clustering of Two Diseases](#). *Journal of the Royal Statistical Society Series A (Statistics in Society)* **164**:73–85.
- Konstantinoudis G, Cameletti M, Gómez-Rubio V, Gómez IL, Pirani M, Baio G, Larrauri A, Riou J, Egger M, Vineis P, Blangiardo M. 2022. Regional excess mortality during the 2020 COVID-19 pandemic in five European countries. *Nature Communications* **13**:482. doi:[10.1038/s41467-022-28157-3](https://doi.org/10.1038/s41467-022-28157-3)
- Kontis V, Bennett JE, Mathers CD, Li G, Foreman K, Ezzati M. 2017. Future life expectancy in 35 industrialised countries: Projections with a Bayesian model ensemble. *The Lancet* **389**:1323–1335. doi:[10.1016/S0140-6736\(16\)32381-9](https://doi.org/10.1016/S0140-6736(16)32381-9)
- Kontis V, Bennett JE, Parks RM, Rashid T, Pearson-Stuttard J, Asaria P, Zhou B, Guillot M, Mathers CD, Khang Y-H, McKee M, Ezzati M. 2022. Lessons learned and lessons missed: Impact of the coronavirus disease 2019 (COVID-19) pandemic on all-cause mortality in 40 industrialised countries and US states prior to mass vaccination. *Wellcome Open Research* **6**:279. doi:[10.12688/wellcomeopenres.17253.2](https://doi.org/10.12688/wellcomeopenres.17253.2)
- Kontis V, Bennett JE, Rashid T, Parks RM, Pearson-Stuttard J, Guillot M, Asaria P, Zhou B, Battaglini M, Corsetti G, McKee M, Di Cesare M, Mathers CD, Ezzati M. 2020. Magnitude, demographics and dynamics of the effect of the first wave of the COVID-19 pandemic on all-cause mortality in 21 industrialized countries. *Nature Medicine* **26**:1919–1928. doi:[10.1038/s41591-020-1112-0](https://doi.org/10.1038/s41591-020-1112-0)
- Lao J, Suter C, Langmore I, Chimisov C, Saxena A, Sountsov P, Moore D, Saurous RA, Hoffman MD, Dillon JV. 2020. Tfp.mcmc: Modern Markov Chain Monte Carlo Tools Built for Modern Hardware. doi:[10.48550/arXiv.2002.01184](https://doi.org/10.48550/arXiv.2002.01184)
- Leon DA, Jdanov DA, Shkolnikov VM. 2019. Trends in life expectancy and age-specific mortality in England and Wales, 1970–2016, in comparison with a set of 22 high-income countries: An analysis of vital statistics data. *The Lancet Public Health* **4**:e575–e582. doi:[10.1016/S2468-2667\(19\)30177-X](https://doi.org/10.1016/S2468-2667(19)30177-X)

- Mackenbach JP. 2022. Omran's "Epidemiologic Transition" 50 years on. *International Journal of Epidemiology* **51**:1054–1057. doi:[10.1093/ije/dyac020](https://doi.org/10.1093/ije/dyac020)
- Mahaki B, Mehrabi Y, Kavousi A, Schmid VJ. 2018. Joint Spatio-temporal Shared Component Model with an Application in Iran Cancer Data. *Asian Pacific Journal of Cancer Prevention* **19**:1553–1560. doi:[10.22034/APJCP.2018.19.6.1553](https://doi.org/10.22034/APJCP.2018.19.6.1553)
- Marmot MG, Allen J, Boyce T, Goldblatt P, Morrison J. 2020. Marmot Review: 10 years on. Institute of Health Equity.
- Marmot MG, Davey Smith G, Stansfeld S, Patel C, North F, Head J, White I, Brunner E, Feeney A. 1991. Health inequalities among British civil servants: The Whitehall II study. *The Lancet* **337**:1387–1393. doi:[10.1016/0140-6736\(91\)93068-K](https://doi.org/10.1016/0140-6736(91)93068-K)
- Marmot MG, Goldblatt P, Allen J, Boyce T, McNeish D, Grady M, Strelitz J, Geddes I, Friel S, Porritt F, Reinertsen E, Bell R, Allen M. 2010. The Marmot Review: Fair society, healthy lives. Institute of Health Equity.
- Marmot MG, Shipley MJ, Rose G. 1984. Inequalities in death - specific explanations of a general pattern? *The Lancet* **323**:1003–1006. doi:[10.1016/S0140-6736\(84\)92337-7](https://doi.org/10.1016/S0140-6736(84)92337-7)
- McElreath R. 2020. Statistical Rethinking, 2nd ed. CRC Press.
- Meslé F, Vallin J. 2021. Causes of Death at Very Old Ages, Including for Supercentenarians In: Maier H, Jeune B, Vaupel JW, editors. Exceptional Lifespans, Demographic Research Monographs. Cham: Springer International Publishing. pp. 69–84. doi:[10.1007/978-3-030-49970-9_7](https://doi.org/10.1007/978-3-030-49970-9_7)
- Ministry of Housing, Communities & Local Government. 2019. English indices of deprivation 2019. *Ministry of Housing, Communities & Local Government*.
- Mokdad AH, Dwyer-Lindgren L, Fitzmaurice C, Stubbs RW, Bertozzi-Villa A, Morozoff C, Charara R, Allen C, Naghavi M, Murray CJL. 2017. Trends and Patterns of Disparities in Cancer Mortality Among US Counties, 1980-2014. *JAMA* **317**:388–406. doi:[10.1001/jama.2016.20324](https://doi.org/10.1001/jama.2016.20324)
- Moriyama IM, Gover M. 1948. Statistical Studies of Heart Diseases: I. Heart Diseases and Allied Causes of Death in Relation to Age Changes in the Population. *Public Health Reports (1896-1970)* **63**:537–545. doi:[10.2307/4586527](https://doi.org/10.2307/4586527)
- Mukadam N, Livingston G, Rantell K, Rickman S. 2014. Diagnostic rates and treatment of dementia before and after launch of a national dementia policy: An observational study using

- English national databases. *BMJ Open* **4**:e004119. doi:[10.1136/bmjopen-2013-004119](https://doi.org/10.1136/bmjopen-2013-004119)
- Murray CJL, Lopez AD. 1996. The Global Burden of Disease: A comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected in 2020. World Health Organization.
- National Cancer Intelligence Network. 2010. Cancer eAtlas. *National Cancer Registration and Analysis Service*.
- National Research Council. 2010. [International Differences in Mortality at Older Ages](#). Washington (DC): National Academies Press (US).
- NHS. 2019. The NHS Long Term Plan. NHS.
- NHS Digital. 2023. Health Survey for England. *NHS Digital*.
- NHS Digital. 2022. CancerData. *CancerData*.
- Niksic M, Rachet B, Duffy SW, Quaresma M, Møller H, Forbes LJ. 2016. Is cancer survival associated with cancer symptom awareness and barriers to seeking medical help in England? An ecological study. *British Journal of Cancer* **115**:876–886. doi:[10.1038/bjc.2016.246](https://doi.org/10.1038/bjc.2016.246)
- NumPyro documentation. 2023. CAR distribution. *NumPyro*.
- OECD. 2016. OECD Reviews of Health Care Quality: United Kingdom 2016: Raising Standards. Paris: Organisation for Economic Co-operation and Development.
- Office for National Statistics. 2022a. 2011 Census geographies. *Office for National Statistics*.
- Office for National Statistics. 2022b. User guide to mortality statistics. *Office for National Statistics*.
- Office for National Statistics. 2021d. Cigarette smoking among adults. *GOVUK*.
- Office for National Statistics. 2021b. Lower layer Super Output Area population estimates. *Office for National Statistics*.
- Office for National Statistics. 2021a. Middle layer Super Output Area population estimates. *Office for National Statistics*.
- Office for National Statistics. 2021c. Population estimates by output areas, electoral, health and other geographies, England and Wales. *Office for National Statistics*.
- Office for National Statistics. 2015. Health Expectancies at Birth for Middle Layer Super Output Areas (MSOAs), England. *Office for National Statistics*.
- Olshansky SJ, Ault AB. 1986. The Fourth Stage of the Epidemiologic Transition: The Age of Delayed Degenerative Diseases. *The Milbank Quarterly* **64**:355–391. doi:[10.2307/3350025](https://doi.org/10.2307/3350025)

- Omran AR. 1977. A century of epidemiologic transition in the United States. *Preventive Medicine* **6**:30–51. doi:[10.1016/0091-7435\(77\)90003-2](https://doi.org/10.1016/0091-7435(77)90003-2)
- Omran AR. 1971. The Epidemiologic Transition: A Theory of the Epidemiology of Population Change. *The Milbank Memorial Fund Quarterly* **49**:509–538. doi:[10.2307/3349375](https://doi.org/10.2307/3349375)
- Overman HG, Xu X. 2022. Spatial disparities across labour markets. Institute for Fiscal Studies.
- Paciorek CJ, Schervish MJ. 2006. Spatial modelling using a new class of nonstationary covariance functions. *Environmetrics* **17**:483–506. doi:[10.1002/env.785](https://doi.org/10.1002/env.785)
- Parkes B, Hansell AL, Ghosh RE, Douglas P, Fecht D, Wellesley D, Kurinczuk JJ, Rankin J, de Hoogh K, Fuller GW, Elliott P, Toledano MB. 2020. Risk of congenital anomalies near municipal waste incinerators in England and Scotland: Retrospective population-based cohort study. *Environment International* **134**:104845. doi:[10.1016/j.envint.2019.05.039](https://doi.org/10.1016/j.envint.2019.05.039)
- Parkin DM, Boyd L, Walker LC. 2011. 16. The fraction of cancer attributable to lifestyle and environmental factors in the UK in 2010. *British Journal of Cancer* **105**:S77–S81. doi:[10.1038/bjc.2011.489](https://doi.org/10.1038/bjc.2011.489)
- Pearson-Stuttard J, Bennett JE, Cheng YJ, Vamos EP, Cross AJ, Ezzati M, Gregg EW. 2021. Trends in predominant causes of death in individuals with and without diabetes in England from 2001 to 2018: An epidemiological analysis of linked primary care records. *The Lancet Diabetes & Endocrinology* **9**:165–173. doi:[10.1016/S2213-8587\(20\)30431-9](https://doi.org/10.1016/S2213-8587(20)30431-9)
- Pell JP, Haw S, Cobbe S, Newby DE, Pell ACH, Fischbacher C, McConnachie A, Pringle S, Murdoch D, Dunn F, Oldroyd K, MacIntyre P, O'Rourke B, Borland W. 2008. Smoke-free Legislation and Hospitalizations for Acute Coronary Syndrome. *New England Journal of Medicine* **359**:482–491. doi:[10.1056/NEJMs0706740](https://doi.org/10.1056/NEJMs0706740)
- Phan D, Pradhan N, Jankowiak M. 2019. Composable Effects for Flexible and Accelerated Probabilistic Programming in NumPyro. doi:[10.48550/arXiv.1912.11554](https://doi.org/10.48550/arXiv.1912.11554)
- Preston SH. 1970. An International Comparison of Excessive Adult Mortality. *Population Studies* **24**:5–20. doi:[10.2307/2173259](https://doi.org/10.2307/2173259)
- Preston SH, Heuveline P, Guillot M. 2001. Demography: Measuring and Modeling Population Processes. Blackwell Publishing.
- Preston SH, Nelson VE. 1974. Structure and change in causes of death: An international summary. *Population Studies* **28**:19–51. doi:[10.1080/00324728.1974.10404577](https://doi.org/10.1080/00324728.1974.10404577)
- Public Health England. 2023. Public Health Outcomes Framework. *Public Health England*.

- Public Health England. 2021b. Local Alcohol Profiles for England. *Public Health England*.
- Public Health England. 2021a. Local Health - Small Area Public Health Data. *Public Health England*.
- Rachet B, Ellis L, Maringe C, Chu T, Nur U, Quaresma M, Shah A, Walters S, Woods L, Forman D, Coleman MP. 2010. Socioeconomic inequalities in cancer survival in England after the NHS cancer plan. *British Journal of Cancer* **103**:446–453. doi:[10.1038/sj.bjc.6605752](https://doi.org/10.1038/sj.bjc.6605752)
- Rashid T. 2022. Probabilistic-programming-packages. *GitHub*.
- Rashid T, Bennett JE, Muller DC, Cross AJ, Pearson-Stuttard J, Asaria P, Daby HI, Fecht D, Davies B, Ezzati M. 2024. Mortality from leading cancers in districts of England from 2002 to 2019: A population-based, spatiotemporal study. *The Lancet Oncology* **25**:86–98. doi:[10.1016/S1470-2045\(23\)00530-2](https://doi.org/10.1016/S1470-2045(23)00530-2)
- Rashid T, Bennett JE, Paciorek CJ, Doyle Y, Pearson-Stuttard J, Flaxman S, Fecht D, Toledano MB, Li G, Daby HI, Johnson E, Davies B, Ezzati M. 2021. Life expectancy and risk of death in 6791 communities in England from 2002 to 2019: High-resolution spatiotemporal analysis of civil registration data. *The Lancet Public Health* **6**:e805–e816. doi:[10.1016/S2468-2667\(21\)00205-X](https://doi.org/10.1016/S2468-2667(21)00205-X)
- Rasmussen CE, Williams CKI. 2006. Gaussian Processes for Machine Learning. MIT Press.
- Roberts GO, Rosenthal JS. 2004. General state space Markov chains and MCMC algorithms. *Probability Surveys* **1**:20–71. doi:[10.1214/154957804100000024](https://doi.org/10.1214/154957804100000024)
- Salomon JA, Murray CJL. 2002. The Epidemiologic Transition Revisited: Compositional Models for Causes of Death by Age and Sex. *Population and Development Review* **28**:205–228. doi:[10.1111/j.1728-4457.2002.00205.x](https://doi.org/10.1111/j.1728-4457.2002.00205.x)
- Semenova E, Xu Y, Howes A, Rashid T, Bhatt S, Mishra S, Flaxman S. 2022. PriorVAE: Encoding spatial priors with variational autoencoders for small-area estimation. *Journal of The Royal Society Interface* **19**:20220094. doi:[10.1098/rsif.2022.0094](https://doi.org/10.1098/rsif.2022.0094)
- Silcocks PBS, Jenner DA, Reza R. 2001. Life expectancy as a summary of mortality in a population: Statistical considerations and suitability for use by health authorities. *Journal of Epidemiology & Community Health* **55**:38–43. doi:[10.1136/jech.55.1.38](https://doi.org/10.1136/jech.55.1.38)
- Stan development team. 2020. Prior Choice Recommendations. *GitHub*.
- Steel N, Ford JA, Newton JN, Davis ACJ, Vos T, Naghavi M, Glenn S, Hughes A, Dalton AM, Stockton D, Humphreys C, Dallat M, Schmidt J, Flowers J, Fox S, Abubakar I, Aldridge RW,

- Baker A, Brayne C, Brugha T, Capewell S, Car J, Cooper C, Ezzati M, Fitzpatrick J, Greaves F, Hay R, Hay S, Kee F, Larson HJ, Lyons RA, Majeed A, McKee M, Rawaf S, Rutter H, Saxena S, Sheikh A, Smeeth L, Viner RM, Vollset SE, Williams HC, Wolfe C, Woolf A, Murray CJL. 2018. Changes in health in the countries of the UK and 150 English Local Authority areas 1990–2016: A systematic analysis for the Global Burden of Disease Study 2016. *The Lancet* **392**:1647–1661. doi:[10.1016/S0140-6736\(18\)32207-4](https://doi.org/10.1016/S0140-6736(18)32207-4)
- Stephens AS, Purdie S, Yang B, Moore H. 2013. Life expectancy estimation in small administrative areas with non-uniform population sizes: Application to Australian New South Wales local government areas. *BMJ Open* **3**:e003710. doi:[10.1136/bmjopen-2013-003710](https://doi.org/10.1136/bmjopen-2013-003710)
- Sudharsanan N, Aburto JM, Riffe T, van Raalte A. 2022. Large variation in the epidemiological transition across countries: Is it still valuable as a mortality theory? *International Journal of Epidemiology* **51**:1057–1061. doi:[10.1093/ije/dyac107](https://doi.org/10.1093/ije/dyac107)
- Taylor-Robinson D, Lai ETC, Wickham S, Rose T, Norman P, Bambra C, Whitehead M, Barr B. 2019. Assessing the impact of rising child poverty on the unprecedented rise in infant mortality in England, 2000–2017: Time trend analysis. *BMJ Open* **9**:e029424. doi:[10.1136/bmjopen-2019-029424](https://doi.org/10.1136/bmjopen-2019-029424)
- Thatcher AR, Kannisto V, Andreev KF. 2002. The Survivor Ratio Method for Estimating Numbers at High Ages. *Demographic Research* **6**:1–18. doi:[10.4054/DemRes.2002.6.1](https://doi.org/10.4054/DemRes.2002.6.1)
- The Economist. 2023. Britain has endured a decade of early deaths. Why? *The Economist* 19–21.
- Thun M, Peto R, Boreham J, Lopez AD. 2012. Stages of the cigarette epidemic on entering its second century. *Tobacco Control* **21**:96–101. doi:[10.1136/tobaccocontrol-2011-050294](https://doi.org/10.1136/tobaccocontrol-2011-050294)
- Toledano MB, Jarup L, Best N, Wakefield J, Elliott P. 2001. Spatial variation and temporal trends of testicular cancer in Great Britain. *British Journal of Cancer* **84**:1482–1487. doi:[10.1054/bjoc.2001.1739](https://doi.org/10.1054/bjoc.2001.1739)
- Travers T, Sims S, Bosetti N. 2016. Housing and inequality in London. Centre for London.
- Tsilidis KK, Kasimis JC, Lopez DS, Ntzani EE, Ioannidis JPA. 2015. Type 2 diabetes and cancer: Umbrella review of meta-analyses of observational studies. *BMJ* **350**:g7607. doi:[10.1136/bmj.g7607](https://doi.org/10.1136/bmj.g7607)
- Turner EL, Metcalfe C, Donovan JL, Noble S, Sterne JAC, Lane JA, I Walsh E, Hill EM, Down L, Ben-Shlomo Y, Oliver SE, Evans S, Brindle P, Williams NJ, Hughes LJ, Davies CF, Ng SY,

- Neal DE, Hamdy FC, Albertsen P, Reid CM, Oxley J, McFarlane J, Robinson MC, Adolfsson J, Zietman A, Baum M, Koupparis A, Martin RM. 2016. Contemporary accuracy of death certificates for coding prostate cancer as a cause of death: Is reliance on death certification good enough? A comparison with blinded review by an independent cause of death evaluation committee. *British Journal of Cancer* **115**:90–94. doi:[10.1038/bjc.2016.162](https://doi.org/10.1038/bjc.2016.162)
- van Dijk JT, Lansley G, Longley PA. 2021. Using linked consumer registers to estimate residential moves in the United Kingdom. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **184**:1452–1474. doi:[10.1111/rssa.12713](https://doi.org/10.1111/rssa.12713)
- Vehtari A, Gelman A, Simpson D, Carpenter B, Bürkner P-C. 2021. Rank-normalization, folding, and localization: An improved R-hat for assessing convergence of MCMC. *Bayesian Analysis* **16**:667–718. doi:[10.1214/20-BA1221](https://doi.org/10.1214/20-BA1221)
- Wakefield J, Elliott P. 1999. Issues in the statistical analysis of small area health data. *Statistics in Medicine* **18**:2377–2399. doi:[10.1002/\(SICI\)1097-0258\(19990915/30\)18:17/18<2377::AID-SIM263>3.0.CO;2-G](https://doi.org/10.1002/(SICI)1097-0258(19990915/30)18:17/18<2377::AID-SIM263>3.0.CO;2-G)
- Walters S, Benitez-Majano S, Muller P, Coleman MP, Allemani C, Butler J, Peake M, Guren MG, Glimelius B, Bergström S, Pahlman L, Rachet B. 2015. Is England closing the international gap in cancer survival? *British Journal of Cancer* **113**:848–860. doi:[10.1038/bjc.2015.265](https://doi.org/10.1038/bjc.2015.265)
- Walters S, Quaresma M, Coleman MP, Gordon E, Forman D, Rachet B. 2011. Geographical variation in cancer survival in England, 1991–2006: An analysis by Cancer Network. *J Epidemiol Community Health* **65**:1044–1052. doi:[10.1136/jech.2010.126656](https://doi.org/10.1136/jech.2010.126656)
- Whitty C. 2021. Chief Medical Officer's annual report 2021: Health in coastal communities.
- Wilkinson RG. 1992. Income distribution and life expectancy. *British Medical Journal* **304**:165–168. doi:[10.1136/bmj.304.6820.165](https://doi.org/10.1136/bmj.304.6820.165)
- Wilmoth JR. 1995. Are mortality projections always more pessimistic when disaggregated by cause of death? *Mathematical Population Studies* **5**:293–319. doi:[10.1080/08898489509525409](https://doi.org/10.1080/08898489509525409)
- World Bank. 2022. Life expectancy at birth, data. *World Bank*.
- World Health Organization. 2020. WHO methods and data sources for country-level causes of death 2000–2019.
- Yu J, Dwyer-Lindgren L, Bennett J, Ezzati M, Gustafson P, Tran M, Brauer M. 2021. A spatiotemporal analysis of inequalities in life expectancy and 20 causes of mortality in sub-neighbourhoods of Metro Vancouver, British Columbia, Canada, 1990–2016. *Health & Place*

72:102692. doi:[10.1016/j.healthplace.2021.102692](https://doi.org/10.1016/j.healthplace.2021.102692)

Yu J-T, Xu W, Tan C-C, Andrieu S, Suckling J, Evangelou E, Pan A, Zhang C, Jia J, Feng L, Kua E-H, Wang Y-J, Wang H-F, Tan M-S, Li J-Q, Hou X-H, Wan Y, Tan L, Mok V, Tan L, Dong Q, Touchon J, Gauthier S, Aisen PS, Vellas B. 2020. Evidence-based prevention of Alzheimer's disease: Systematic review and meta-analysis of 243 observational prospective studies and 153 randomised controlled trials. *Journal of Neurology, Neurosurgery & Psychiatry* **91**:1201–1209. doi:[10.1136/jnnp-2019-321913](https://doi.org/10.1136/jnnp-2019-321913)

Appendix A

Life table methods

Most of the content in this appendix is taken directly from Preston et al. (2001), but I have reproduced it here for reference and for completeness.

A.1 Period life tables

Calculating life expectancy for a cohort is possible, but you have to wait until every member of the cohort has died. Instead, demographers use period (or “current”) life tables, which consider what would happen to a hypothetical cohort that are subjected to the deaths rates in each age group at an exact period in time. Life tables can be constructed using discrete age bands starting at age x and ending at age $x + n$. We supply the age-specific death rates, ${}_n m_x$, and the average person-years lived by those dying in the interval, ${}_n a_x$, and the life table calculates the mean age at death – the life expectancy, e_x .

The probability of dying, ${}_n q_x$, is defined as the ratio of the number of people who died in the age interval, ${}_n d_x$, to the number who survived to age x , l_x :

$${}_n q_x = \frac{{}_n d_x}{l_x}. \quad (\text{A.1})$$

The age-specific death rate is defined as the ratio of the number of people who died in the age interval to the total number of person-years lived, ${}_n L_x$, which is the weighted sum of the number of person-years lived (n) by those who survived, which, in turn, is the difference between those who survived to age x and those who died in the interval ($l_x - {}_n d_x$), and the number of

person-years lived on average (${}_na_x$) by those who died (${}_nd_x$):

$${}_nm_x = \frac{{}_nd_x}{n \cdot (l_x - {}_nd_x) + {}_na_x \cdot {}_nd_x}. \quad (\text{A.2})$$

We assume the denominator of Equation A.2 can be approximated by the mid-year population, ${}_nP_x$, which leads us to recover the expression for the cross-sectional, empirical death rate in Equation 4.1. By rearranging the denominator to make the number of survivors the subject, we obtain

$$l_x = \frac{1}{n} ({}_nP_x + (n - {}_na_x) \cdot {}_nd_x). \quad (\text{A.3})$$

We can substitute this expression into Equation A.1 and divide by ${}_nP_x$ to obtain

$${}_nq_x = \frac{n \cdot {}_nm_x}{1 + (n - {}_na_x) {}_nm_x}. \quad (\text{A.4})$$

This expression, although unintuitive, allows us to convert from ${}_nm_x$ to ${}_nq_x$ with only the parameter ${}_na_x$.

In the period life table, we start with a hypothetical cohort of size $l_0 = 100,000$ and sequentially apply the probability of surviving in each age group, ${}_np_x = 1 - {}_nq_x$, to calculate the number of survivors as

$$l_{x+n} = l_x \cdot {}_np_x. \quad (\text{A.5})$$

The number of person-years lived is the sum of the number of survivors weighted by the band width and number of people who died (${}_nd_x = l_x \cdot {}_nq_x$) weighted by ${}_na_x$

$${}_nL_x = n \cdot l_x + {}_na_x \cdot l_x \cdot {}_nq_x. \quad (\text{A.6})$$

The open interval ${}_{\infty}q_x = 1$, as nobody is immortal. Using Equation A.1, it follows that the number of deaths in this interval is equal to the number who survived to the final age group, i.e. ${}_{\infty}d_x = l_x$. Since the death rate from Equation A.2 can be rewritten using the number of person-years lived, ${}_nL_x$, as the denominator and we can substitute the number of deaths with the number surviving to the final age group, we can obtain an expression for the number of

person-years lived in the open-ended age interval

$${}_xL_x = \frac{{}_x d_x}{{}_x m_x} = \frac{l_x}{{}_x m_x}. \quad (\text{A.7})$$

The total number of person-years lived above x is

$$T_x = \sum_{x=a}^{\infty} {}_xL_x. \quad (\text{A.8})$$

Then, life expectancy is given by dividing the number of person-years lived by the number of people who will live them

$$e_x = \frac{T_x}{l_x}. \quad (\text{A.9})$$

Throughout the thesis, I only consider life expectancy at birth.

A.1.1 The very young ages and the very old ages

On average, it is a good approximation to assume deaths occur halfway through the age interval: ${}_na_x = n/2$. But for younger ages, particularly at lower levels of mortality, the majority of infant deaths lie further towards the earliest stages of infancy. Coale and Demeny used regression on a series of international datasets to recommend suitable values for ${}_1a_0$ and ${}_4a_1$ instead of the midpoint (Coale et al., 1983).

The start of the open-ended age group can be many years away from some of the ages at death, particularly in ageing populations. In order to produce reliable estimates of death rates at older ages, I used the Kannisto-Thatcher method to expand the terminal age group (≥ 85 years) of the life table and adjust ${}_na_x$ above 70 years (Thatcher et al., 2002). The Kannisto-Thatcher method assumes the probability of dying is a logistic function of age. The logit-transformed probability of dying above 70 years is regressed upon age. The resulting curve is extrapolated through to 129 years before calculating the number of survivors in the cohort following the adjusted probability of dying to estimate ${}_na_x$ above 70 years.

A.2 Probability of dying

The probability of dying from a specific cause of death, i , is calculated as in Equation A.4. Equally, we can calculate the probability of dying by subtracting the probability of surviving in each age group through to that age from unity, i.e. $1 - \prod_x {}_n p_x^i$. Note, even for the smallest death rates, ${}_x q_x^i = 1 - \prod_x {}_n p_x^i = 1$ – if you live to infinity, you'll die of it eventually.

A.3 Cause-specific decomposition of differences in life expectancy

Using quantities generated from the life tables of two populations as above, Arriaga (1984) proposed a method to calculate the age-specific contributions to the difference in life expectancy between these populations as

$${}_n \Delta_x = \frac{l_x^1}{l_0^1} \left(\frac{{}_n L_x^2}{l_x^2} - \frac{{}_n L_x^1}{l_x^1} \right) + \frac{T_{x+n}^2}{l_0^1} \left(\frac{l_x^1}{l_x^2} - \frac{l_{x+n}^1}{l_{x+n}^2} \right). \quad (\text{A.10})$$

The first term on the right hand side corresponds to the “direct effect” on the life expectancy difference between the two populations in the average number of person-years lived by the survivors to that age group (${}_n L_x / l_x$). The second term represents the “indirect effect” on the number of survivors caused by the mortality changes within an age group.

We then assume the age- and cause-specific contributions are proportional to the difference in cause-specific death rates between the two populations:

$${}_n \Delta_x^i = {}_n \Delta_x \cdot \frac{{}_n m_x^i(2) - {}_n m_x^i(1)}{{}_n m_x(2) - {}_n m_x(1)} \quad (\text{A.11})$$

Arriaga showed the sum of the age- and cause-specific contributions are equal to the difference in life expectancy,

$$e_0(2) - e_0(1) = \sum_x {}_n \Delta_x = \sum_x \sum_i {}_n \Delta_x^i. \quad (\text{A.12})$$

So, we can collapse over age groups to get the cause-specific contributions to life expectancy as

$$\sum_x {}_n \Delta_x^i.$$

A.4 Mean age at death

The mean age at death among those who died from a specific cause of death was calculated as

$$\text{mean age at death} = \frac{\sum_x {}_n d_x \cdot {}_n a_x}{\sum_x {}_n d_x}, \quad (\text{A.13})$$

where ${}_n d_x$ is the number of deaths in an age band, calculated as the product of the death rate and the population.

Appendix B

Model parameters and priors

This appendix contains full information on the parameters of each model.

Table B.1: Specification of the Bayesian statistical model in Equation 4.4.

Parameter name	Symbol	Prior	Dimension
Overall intercept	α_0	$\mathcal{N}(0, \sigma^2 = 10^5)$	1
Overall slope	β_0	$\mathcal{N}(0, \sigma^2 = 10^5)$	1
Regional intercept	α_{1r}	$\mathcal{N}(0, \sigma_{\alpha_{1r}}^2)$	9
Regional intercept standard deviation	$\sigma_{\alpha_{1r}}$	$\mathcal{U}(0, 2)$	1
District intercept	α_{1d}	$\mathcal{N}(0, \sigma_{\alpha_{1d}}^2)$	314
District intercept standard deviation	$\sigma_{\alpha_{1d}}$	$\mathcal{U}(0, 2)$	1
MSOA intercept	α_{1s}	$\mathcal{N}(0, \sigma_{\alpha_{1s}}^2)$	6791
MSOA intercept standard deviation	$\sigma_{\alpha_{1s}}$	$\mathcal{U}(0, 2)$	1
Regional slope	β_{1r}	$\mathcal{N}(0, \sigma_{\beta_{1r}}^2)$	9
Regional slope standard deviation	$\sigma_{\beta_{1r}}$	$\mathcal{U}(0, 2)$	1
District slope	β_{1d}	$\mathcal{N}(0, \sigma_{\beta_{1d}}^2)$	314
District slope standard deviation	$\sigma_{\beta_{1d}}$	$\mathcal{U}(0, 2)$	1

Parameter name	Symbol	Prior	Dimension
MSOA slope	β_{1s}	$\mathcal{N}(0, \sigma_{\beta_{1s}}^2)$	6791
MSOA slope standard deviation	$\sigma_{\beta_{1s}}$	$\mathcal{U}(0, 2)$	1
Age group intercept	α_{2a}	$\mathcal{N}(\alpha_{2,a-1}, \sigma_{\alpha_{2a}}^2)$	18
Age group intercept standard deviation	$\sigma_{\alpha_{2a}}$	$\mathcal{U}(0, 2)$	1
Age group slope	β_{2a}	$\mathcal{N}(\beta_{2,a-1}, \sigma_{\beta_{2a}}^2)$	18
Age group slope standard deviation	$\sigma_{\beta_{2a}}$	$\mathcal{U}(0, 2)$	1
Age group MSOA interaction	ξ_{as}	$\mathcal{N}(0, \sigma_{\xi}^2)$	19 × 6791
Age group MSOA interaction standard deviation	σ_{ξ}	$\mathcal{U}(0, 2)$	1
MSOA random walk over time	ν_{st}	$\mathcal{N}(\nu_{s,t-1}, \sigma_{\nu}^2)$	6791 × 17
MSOA random walk over time standard deviation	σ_{ν}	$\mathcal{U}(0, 2)$	1
Age group random walk over time	γ_{at}	$\mathcal{N}(\gamma_{a,t-1}, \sigma_{\gamma}^2)$	19 × 17
Age group random walk over time standard deviation	σ_{γ}	$\mathcal{U}(0, 2)$	1
Overdispersion parameter	r	$\mathcal{U}(0, 50)$	1

Table B.2: Specification of the Bayesian statistical model in Equation 6.1.

Parameter name	Symbol	Prior	Dimension
Overall intercept	α_0	$\mathcal{N}(0, \sigma^2 = 10^5)$	1

Parameter name	Symbol	Prior	Dimension
Overall slope	β_0	$\mathcal{N}(0, \sigma^2 = 10^5)$	1
District intercept	α_{1d}	$\mathcal{N}(0, \sigma_{\alpha_{1d}}^2)$	33
District intercept standard deviation	$\sigma_{\alpha_{1d}}$	$\mathcal{N}^+(1)$	1
MSOA intercept	α_{1m}	$\mathcal{N}(0, \sigma_{\alpha_{1m}}^2)$	983
MSOA intercept standard deviation	$\sigma_{\alpha_{1m}}$	$\mathcal{N}^+(1)$	1
LSOA intercept	α_{1s}	$\mathcal{N}(0, \sigma_{\alpha_{1s}}^2)$	4835
LSOA intercept standard deviation	$\sigma_{\alpha_{1s}}$	$\mathcal{N}^+(1)$	1
District slope	β_{1d}	$\mathcal{N}(0, \sigma_{\beta_{1d}}^2)$	33
District slope standard deviation	$\sigma_{\beta_{1d}}$	$\mathcal{N}^+(1)$	1
MSOA slope	β_{1m}	$\mathcal{N}(0, \sigma_{\beta_{1m}}^2)$	983
MSOA slope standard deviation	$\sigma_{\beta_{1m}}$	$\mathcal{N}^+(1)$	1
LSOA slope	β_{1s}	$\mathcal{N}(0, \sigma_{\beta_{1s}}^2)$	4835
LSOA slope standard deviation	$\sigma_{\beta_{1s}}$	$\mathcal{N}^+(1)$	1
Age group intercept	α_{2a}	$\mathcal{N}(\alpha_{2,a-1}, \sigma_{\alpha_{2a}}^2)$	18
Age group intercept standard deviation	$\sigma_{\alpha_{2a}}$	$\mathcal{N}^+(1)$	1
Age group slope	β_{2a}	$\mathcal{N}(\beta_{2,a-1}, \sigma_{\beta_{2a}}^2)$	18
Age group slope standard deviation	$\sigma_{\beta_{2a}}$	$\mathcal{N}^+(1)$	1
Age group LSOA interaction	$\xi_{a,s}$	$\mathcal{N}(0, \sigma_\xi^2)$	19 × 4835
Age group LSOA interaction standard deviation	σ_ξ	$\mathcal{N}^+(1)$	1

Parameter name	Symbol	Prior	Dimension
LSOA random walk over time	ν_{st}	$\mathcal{N}(\nu_{s,t-1}, \sigma_\nu^2)$	4835 x 17
LSOA random walk over time standard deviation	σ_ν	$\mathcal{N}^+(1)$	1
Age group random walk over time	γ_{at}	$\mathcal{N}(\gamma_{a,t-1}, \sigma_\gamma^2)$	19 x 17
Age group random walk over time standard deviation	σ_γ	$\mathcal{N}^+(1)$	1
Overdispersion parameter	ρ	$\exp(0.1)$	1

Table B.3: Specification of the Bayesian statistical model in Equation 7.1.

Parameter name	Symbol	Prior	Dimension
Overall intercept	α_0	$\mathcal{N}(0, 10)$	1
Overall slope	β_0	$\mathcal{N}(0, 10)$	1
District intercept	α_{1s}	ICAR	314
District intercept standard deviation	$\sigma_{\alpha_{1s}}$	$\mathcal{N}^+(1)$	1
District slope	β_{1s}	ICAR	314
District slope standard deviation	$\sigma_{\beta_{1s}}$	$\mathcal{N}^+(1)$	1
Age group intercept	α_{2a}	$\mathcal{N}(\alpha_{2,a-1}, \sigma_{\alpha_{2a}}^2)$	18
Age group intercept standard deviation	$\sigma_{\alpha_{2a}}$	$\mathcal{N}^+(1)$	1
Age group slope	β_{2a}	$\mathcal{N}(\beta_{2,a-1}, \sigma_{\beta_{2a}}^2)$	18
Age group slope standard deviation	$\sigma_{\beta_{2a}}$	$\mathcal{N}^+(1)$	1

Parameter name	Symbol	Prior	Dimension
Age group district interaction	ξ_{as}	$\mathcal{N}(0, \sigma_\xi^2)$	19 x 314
Age group district interaction standard deviation	σ_ξ	$\mathcal{N}^+(1)$	1
Age group random walk over time	γ_{at}	$\mathcal{N}(\gamma_{a,t-1}, \sigma_\gamma^2)$	19 x 17
Age group random walk over time standard deviation	σ_γ	$\mathcal{N}^+(1)$	1

Appendix C

Hex-cartogram of life expectancy in 2019 at the MSOA level

This is an alternative version of Figure 5.2, using a hexagon-based cartogram where every MSOA has equal size. Hence, London, which has a lot of small MSOAs, takes up a larger proportion of the map. For an interactive version of the original figure, see the [visualisation](#).

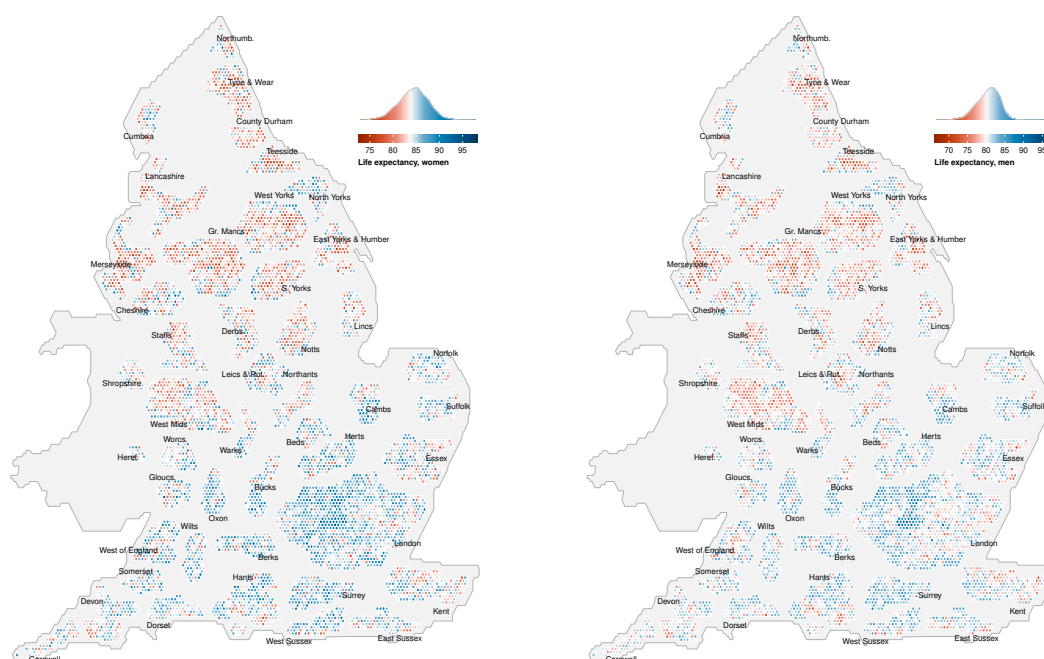


Figure C.1: Cartogram of life expectancy and the distribution of life expectancy in 2019. The areas in white have a life expectancy equal to the national life expectancy.

Appendix D

Groups of causes of death

Table D.1: Groups of causes of death used in the district-level cause-specific mortality analysis in Chapter 7 with ICD-10 codes.

Cause	ICD-10 codes
Trachea, bronchus and lung cancers	C33-C34
Breast cancer	C50
Prostate cancer	C61
Colorectal cancers	C18-C21
Pancreatic cancer	C25
Ovarian cancer	C56
Lymphoma and multiple myeloma	C81-C90, C96
Oesophageal cancer	C15
All other cancers	C00-C14, C16-C17, C22-C24, C26-C32, C37-C41, C43-C49, C51-C55, C57-C60, C62-C67, C68-C80, C91-C95, C97, D00-D48
Ischaemic heart disease	I20-I25
Stroke	I60-I69
All other CVDs	I00-I19, I26-I59, I70-I99
Alzheimer's and other dementias	F00-F03, G30
COPD	J40-44
Diabetes mellitus, nephritis and nephrosis	E10-14, N00-19

Cause	ICD-10 codes
Liver cirrhosis	K70, K74
All other NCDs	D55-D648, D65-D89, E03-E07, E15-E16, E20-E34, E65-E88, F01-F99, G06-G13, G15-G98, H00-H61, H68-H93, J30-J39, J45-J98, K00-K14, K20-K69, K71-K73, K75-K92, L00-L98, M00-M99, N20-64, N75-N99, Q00-Q99, R95, X41-X42, X45
Lower respiratory infections	J09-18, J20-J22
All other IMPN	A00-99, B00-99, D50-53, D649, E00-E02, E40-46, E50-54, G00, G03-G04, G14, H65-H66, N70-N73, J00-J06, O00-O99, P00-P96, Z353
Injuries	U00-U01, U509, V00-V99, W00-W99, X00-X40, X43-X44, X46-X99, Y00-Y01, Y10-Y36, Y381, Y40-Y86, Y870-Y872, Y88-Y89

Ovarian cancer (women) and prostate cancer (men) are sex specific. The all other cancers group also includes breast cancer for men. Liver cirrhosis was not in the top 12 leading causes of death for women, so it was included within all other NCDs.

The residual groups (all other cancers, all other CVDs, all other NCDs, all other IMPN) also contained deaths from the “ill-defined diseases” GHE group (R00-R94, R96-R99, U07, U99). There were 196,055 deaths from ill-defined diseases. These were proportionately assigned between the residual groups.

Causes of death with the ICD-10 code S00-S99 or T00-T99 are not valid underlying causes of death. These were all neonatal deaths, which are not assigned an underlying cause of death, and were imputed using the code in the first position on the death record. These deaths were classified as injuries.

Table D.2: Groups of causes of death used in the district-level cancer mortality analysis in Chapter 8 with ICD-10 codes.

Cancer	ICD-10 codes
Trachea, bronchus and lung cancers	C33-C34
Breast cancer	C50
Prostate cancer	C61
Colorectal cancer	C18-C21
Pancreatic cancer	C25
Ovarian cancer	C56
Lymphoma and multiple myeloma	C81-C90, C96
Oesophageal cancer	C15
Bladder cancer	C67
Leukaemia	C91-C95
Corpus uteri cancer	C54-C55
Stomach cancer	C16
Liver cancer	C22
All other cancers	C00-C14, C17, C23-24, C26-C32, C37-C41, C43-C49, C51-C53, C57-C60, C62-C66, C68-C80, C97 D00-D48

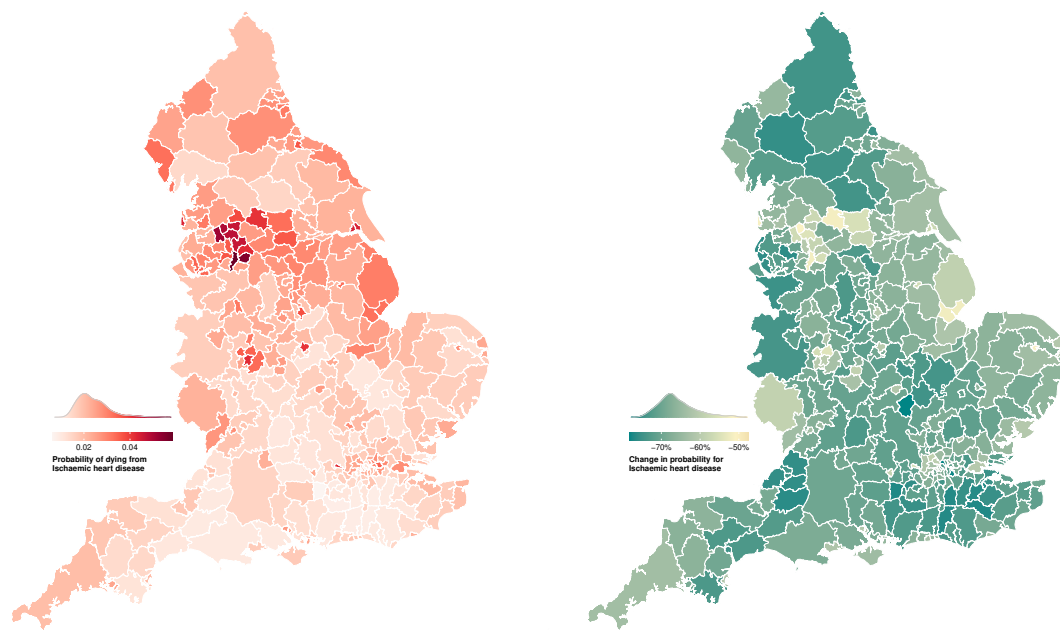
The residual group also contained deaths from the “ill-defined diseases”, which were proportionately assigned as above. The residual group also includes breast cancer for men. Ovarian cancer, corpus uteri cancer (women) and prostate cancer (men) are sex specific. Bladder cancer and liver cancer were not leading cancers for women, so they were included in the residual group.

The next leading cancers in the residual group of all other cancers were bladder cancer, brain and nervous system cancers, and liver cancer for women, and brain and nervous system cancers, mesothelioma, and melanoma and other skin cancers for men.

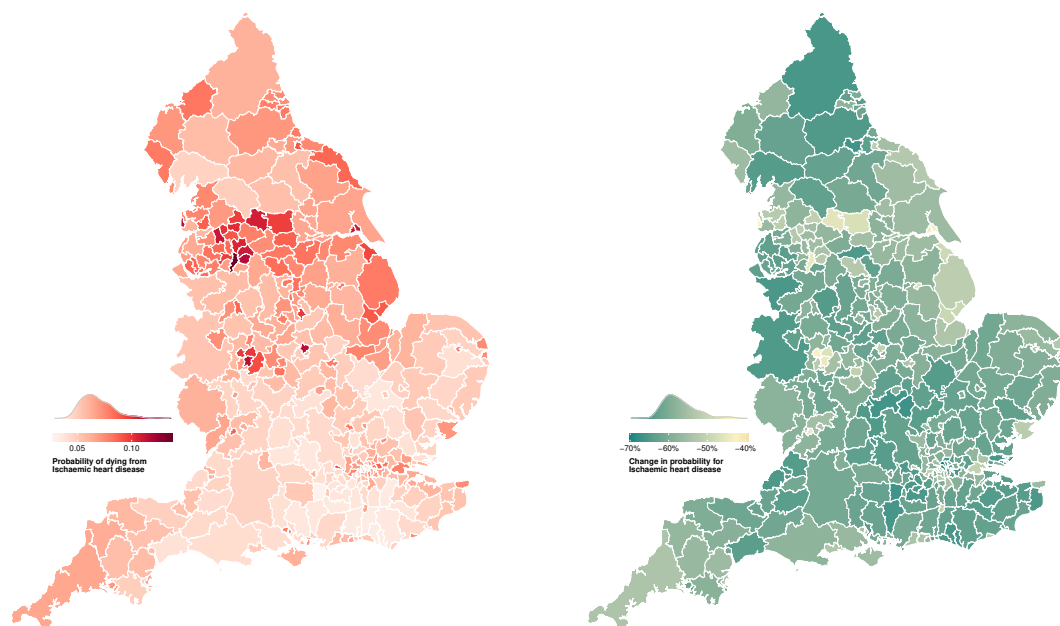
Appendix E

Maps of cause-specific mortality

This appendix contains district-level maps of the probability of dying between birth and 80 years of age in 2019 and change from 2002 to 2019 from specific causes of deaths, which supplement the text in Chapter 7 and Chapter 8. Note, there are two sets of maps for all other cancers, one for each chapter. The all other cancers group for Chapter 8 contains fewer deaths because the group from Chapter 7 had been divided to include more site-specific cancer groups.

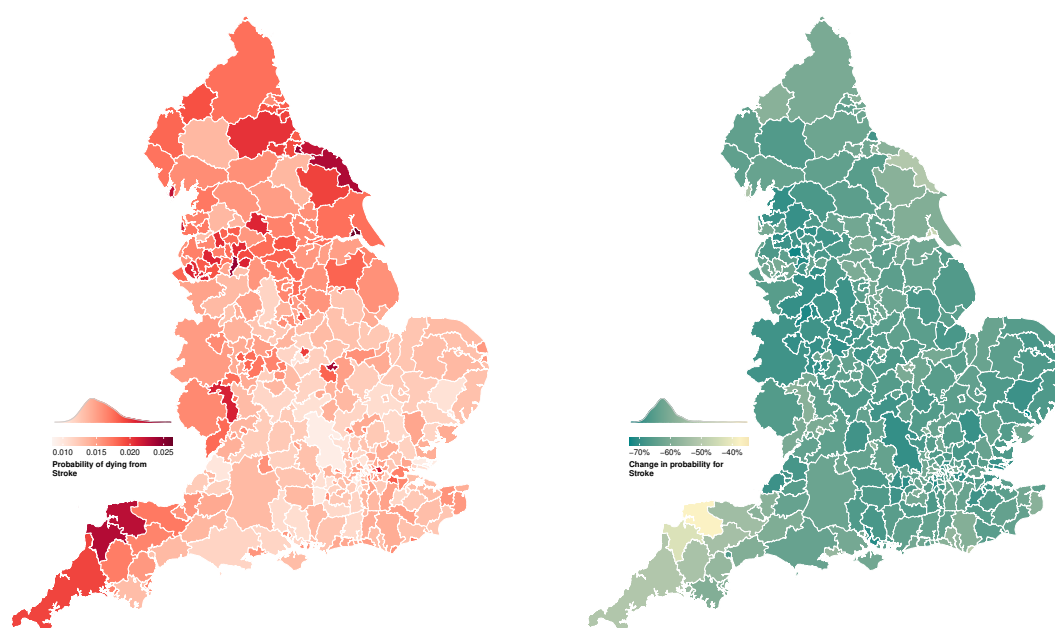


(a) Women

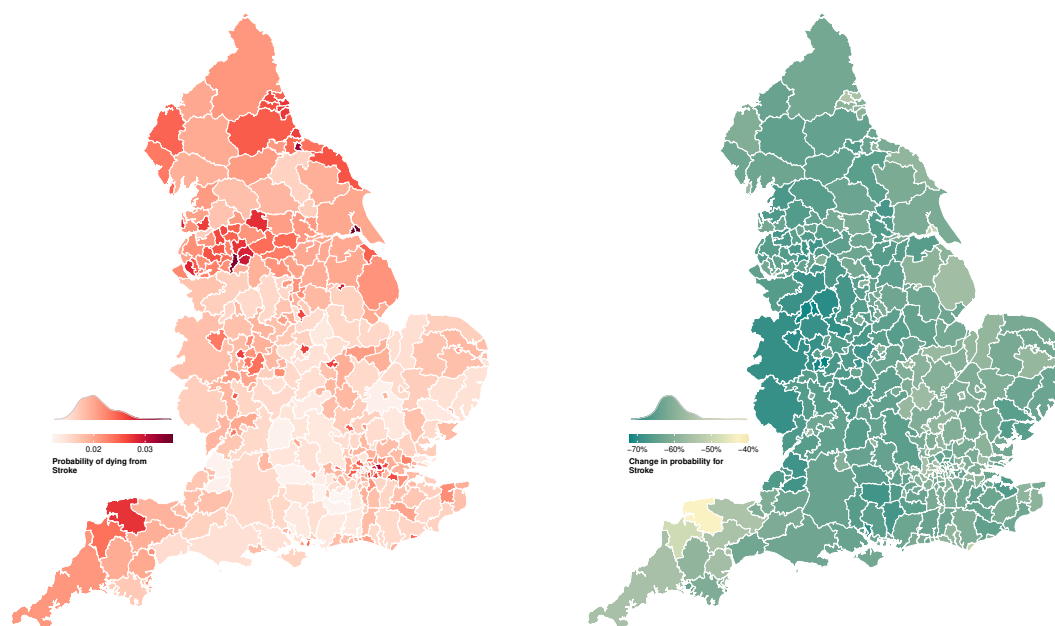


(b) Men

Figure E.1: Probability of dying from ischaemic heart disease between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.

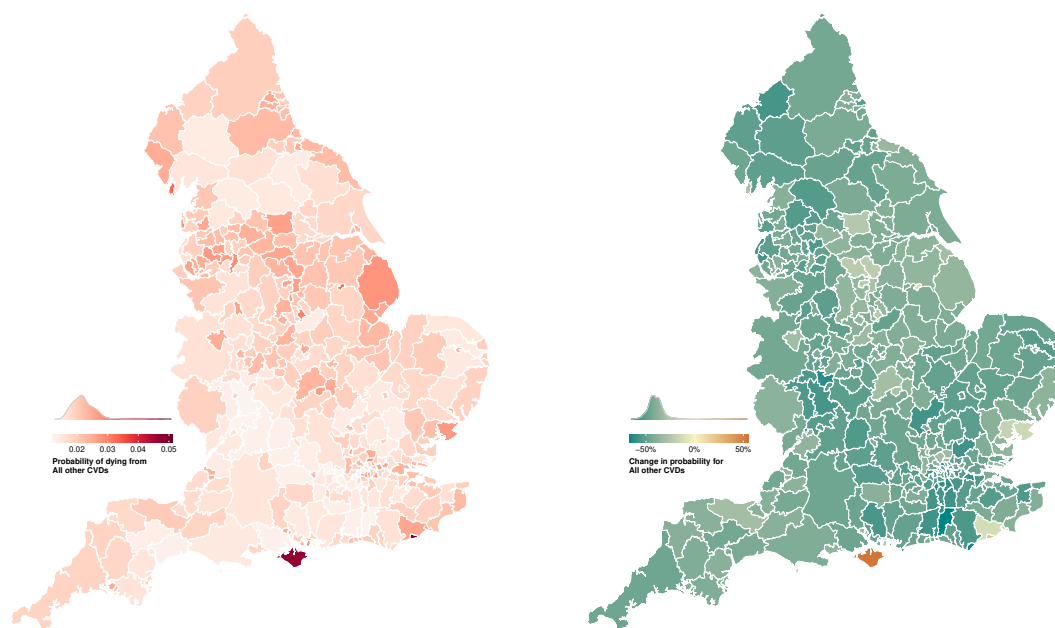


(a) Women

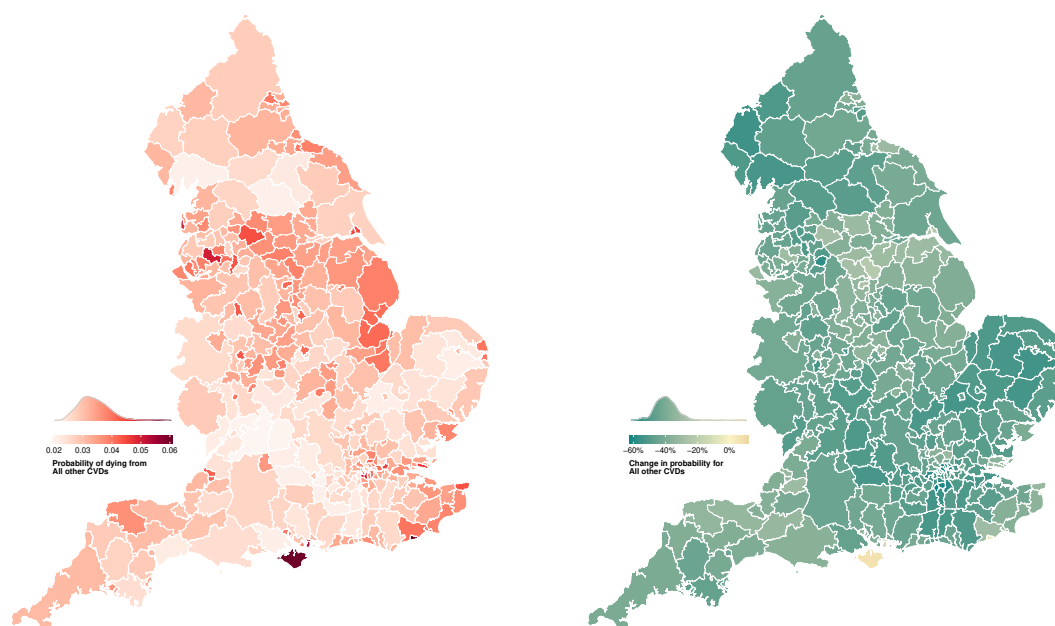


(b) Men

Figure E.2: Probability of dying from a stroke between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.

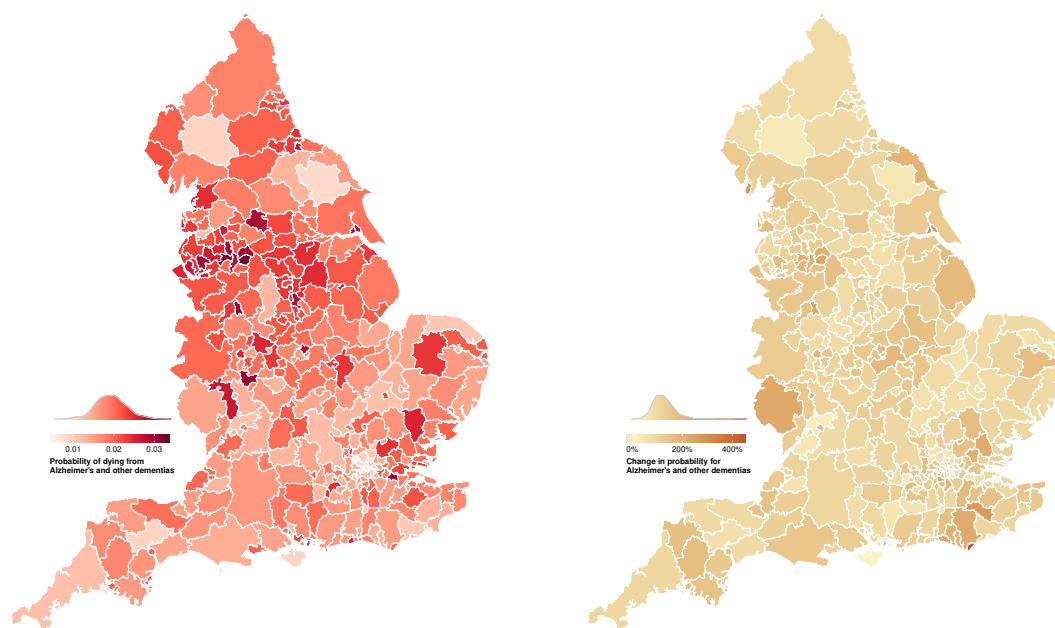


(a) Women

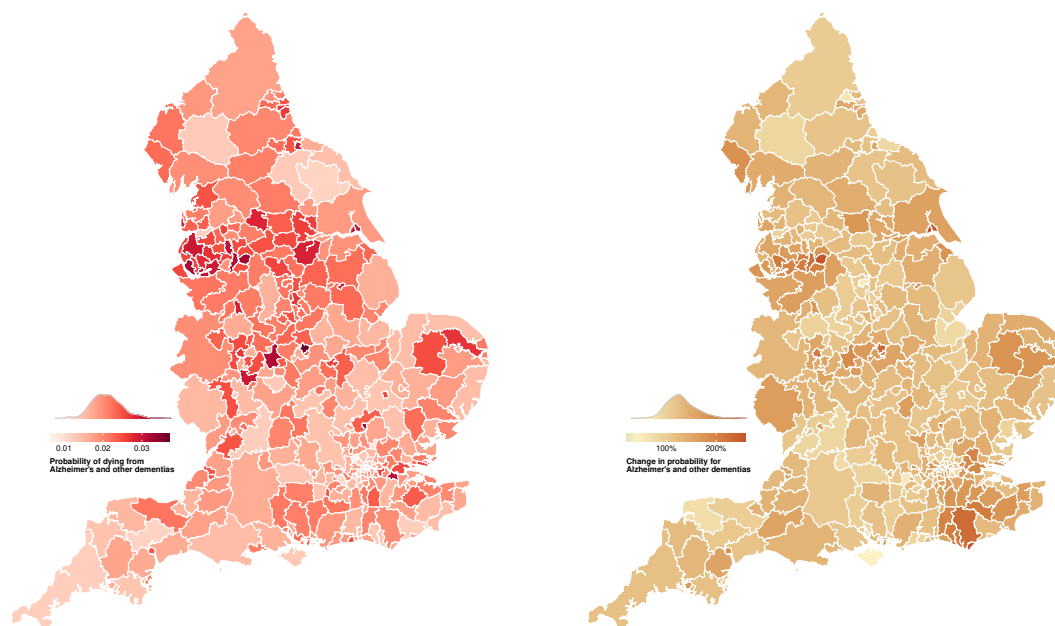


(b) Men

Figure E.3: Probability of dying from all other CVDs between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.

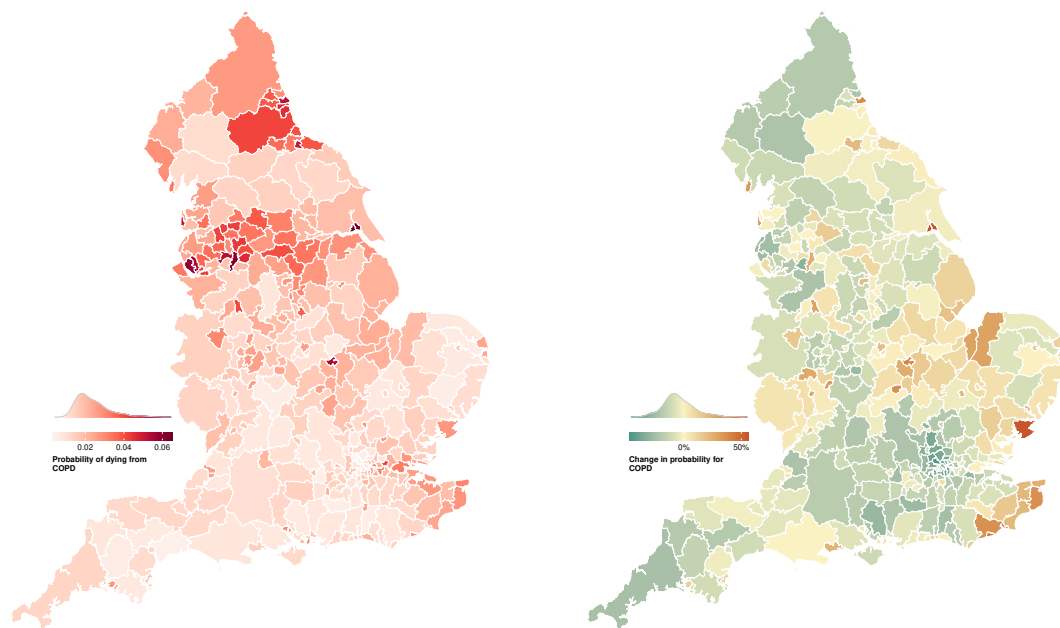


(a) Women

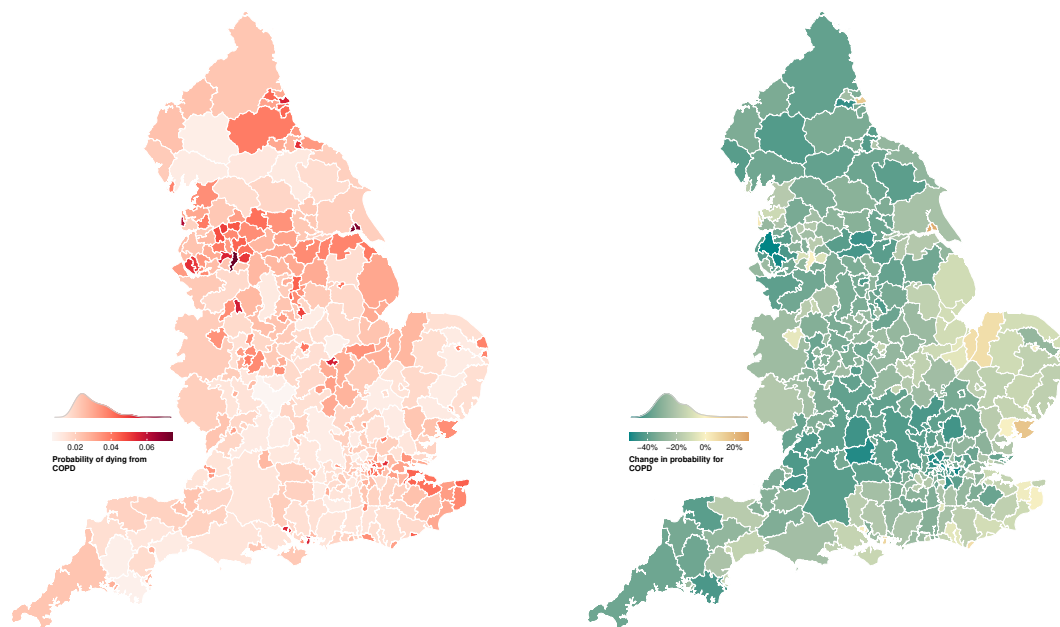


(b) Men

Figure E.4: Probability of dying from Alzheimer's and other dementias between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.

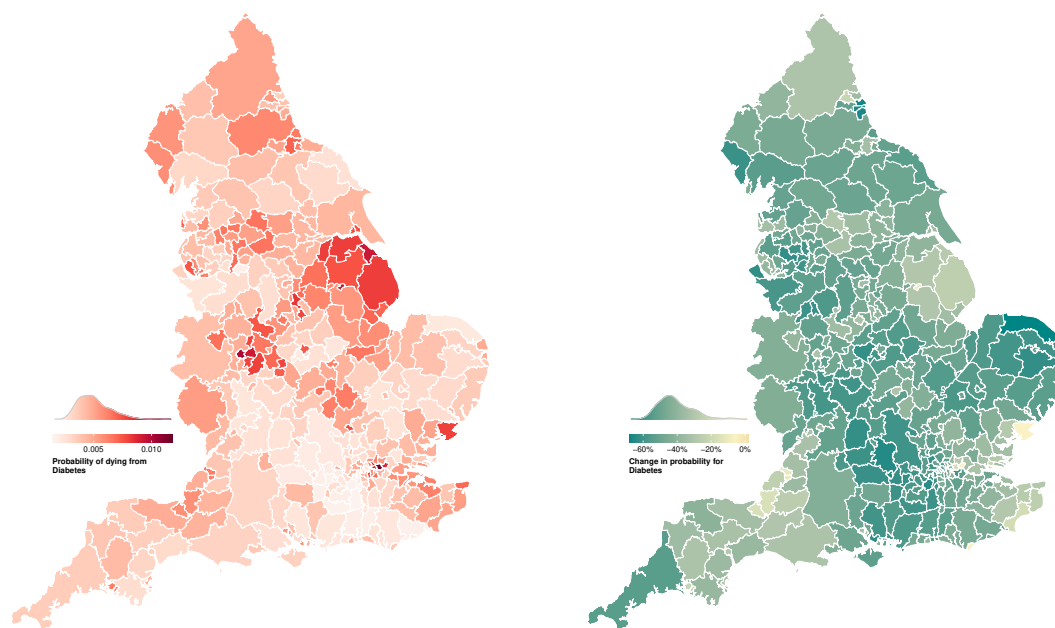


(a) Women

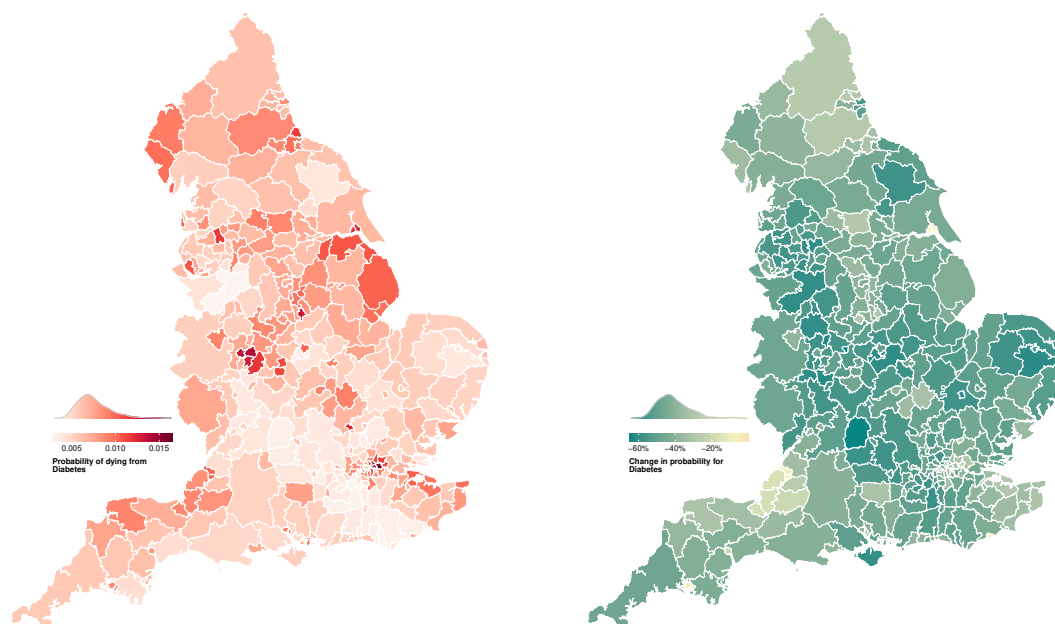


(b) Men

Figure E.5: Probability of dying from COPD between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.



(a) Women



(b) Men

Figure E.6: Probability of dying from diabetes mellitus, nephritis and nephrosis between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.

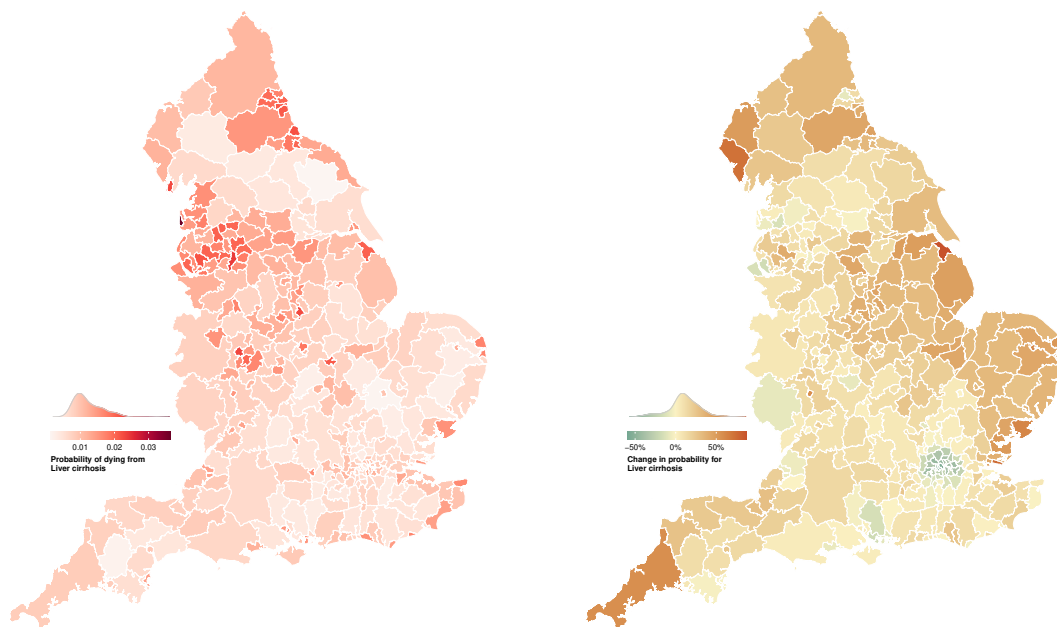
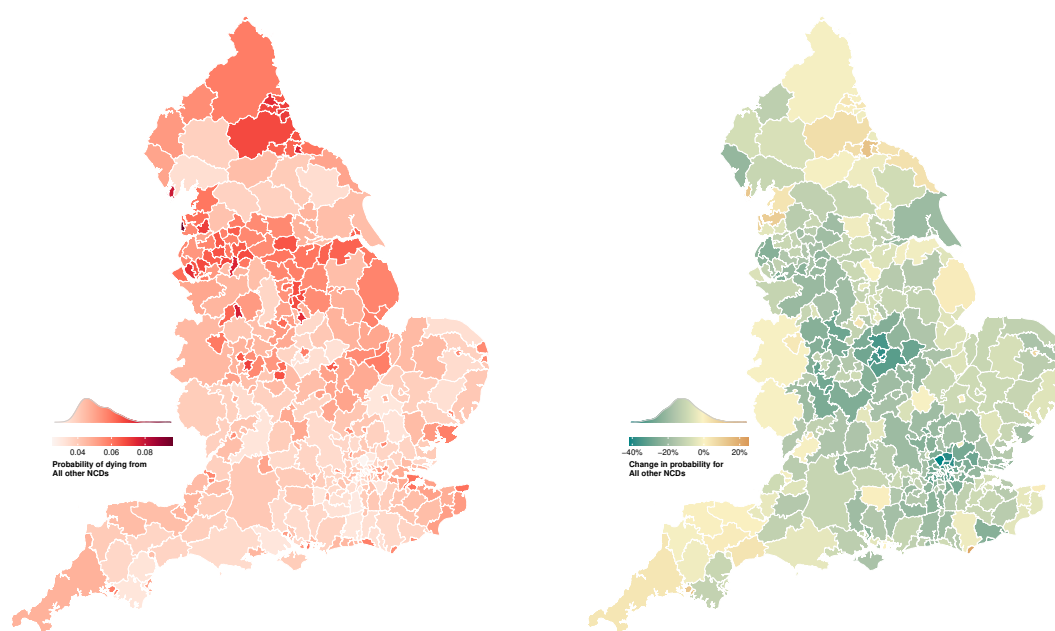
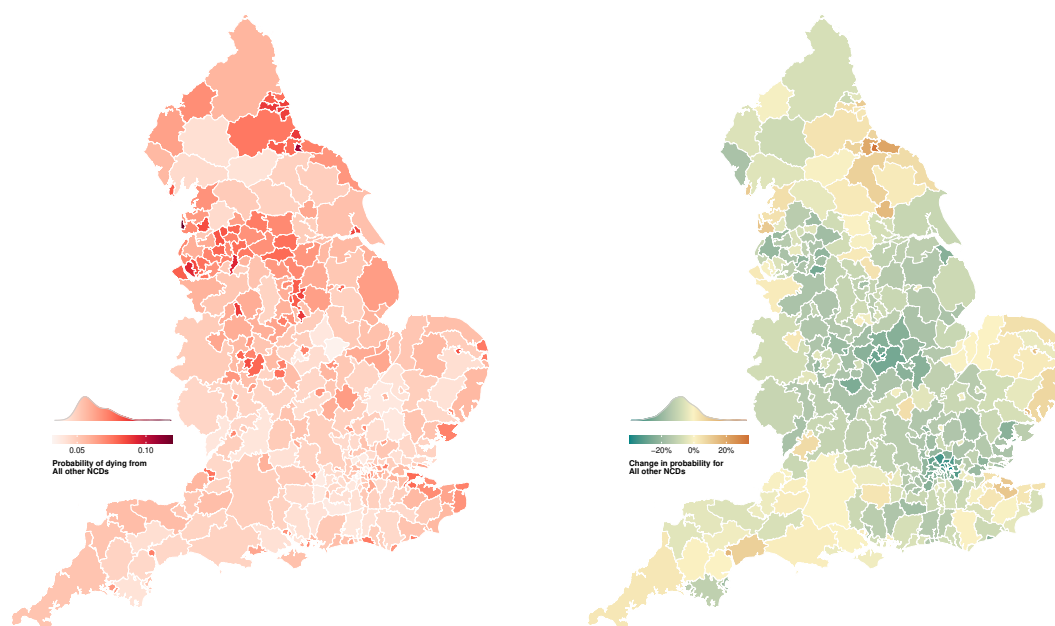


Figure E.7: Probability of dying for men from liver cirrhosis between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.

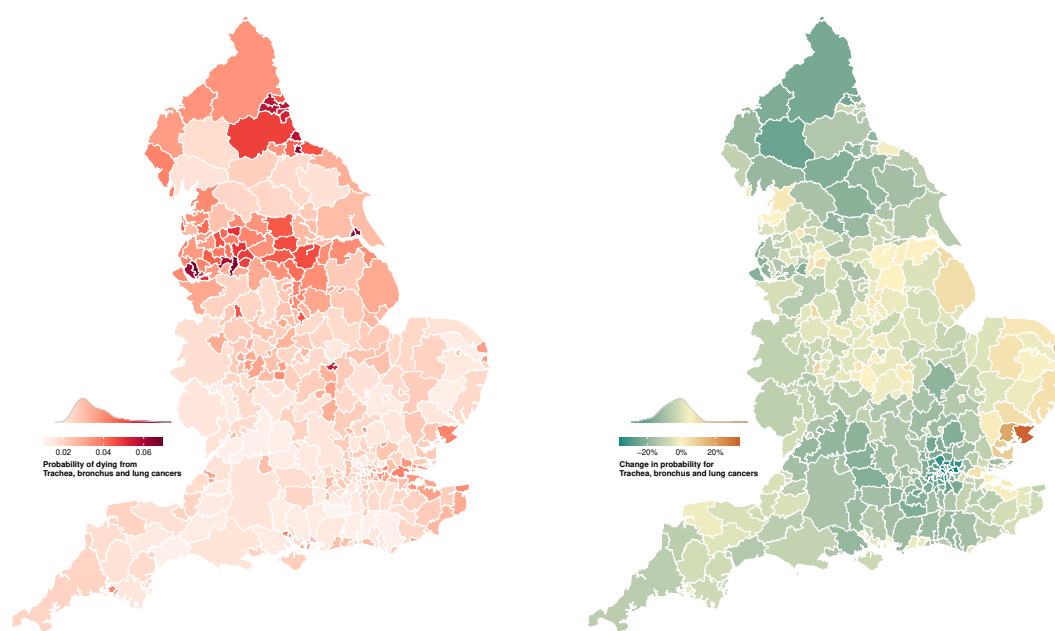


(a) Women

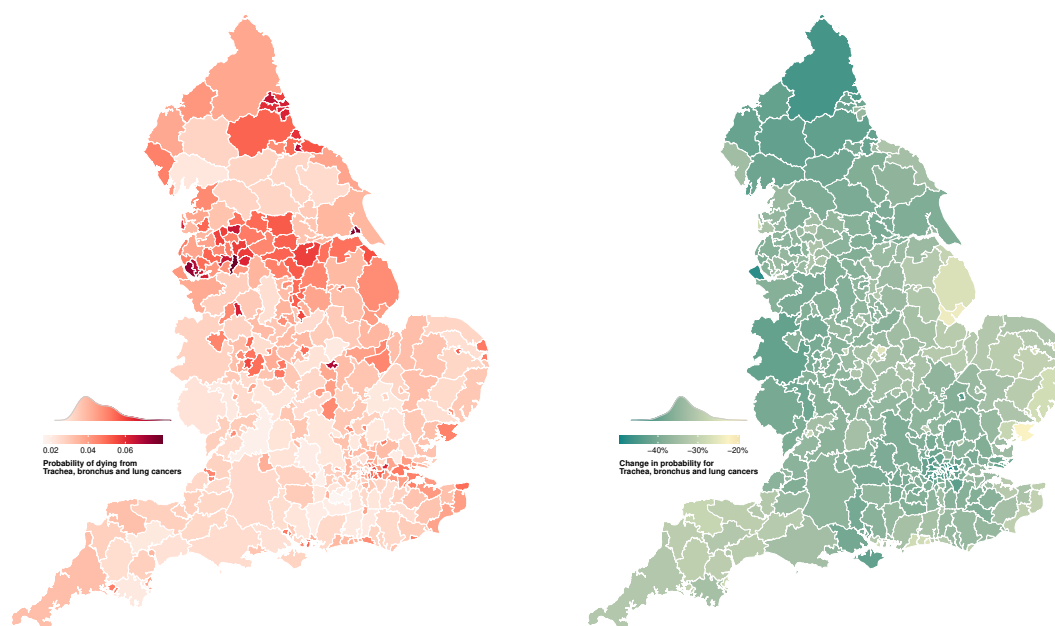


(b) Men

Figure E.8: Probability of dying from all other NCDs between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.

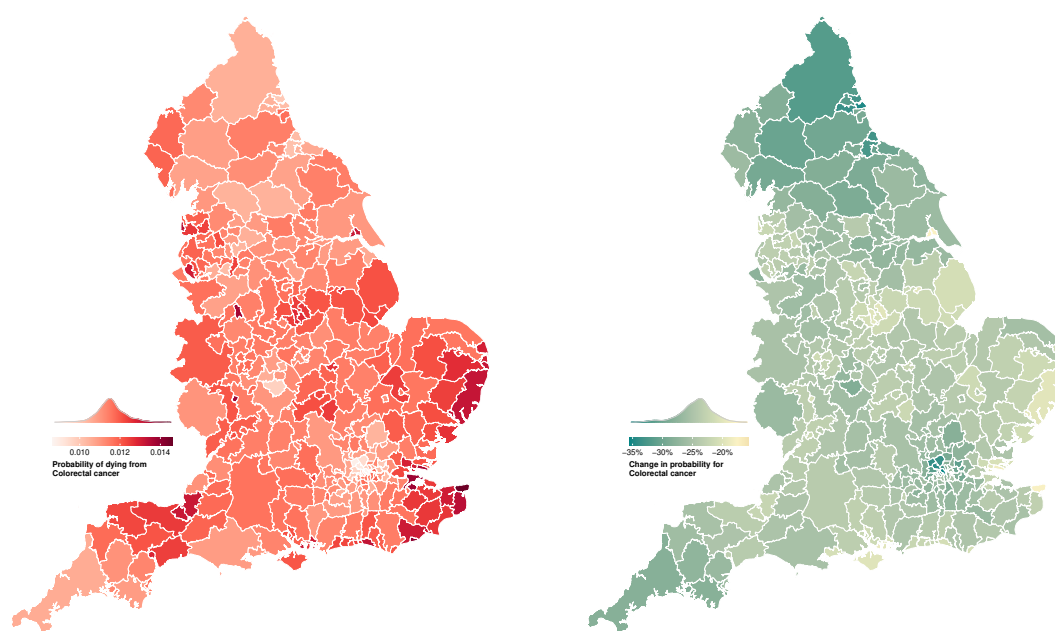


(a) Women

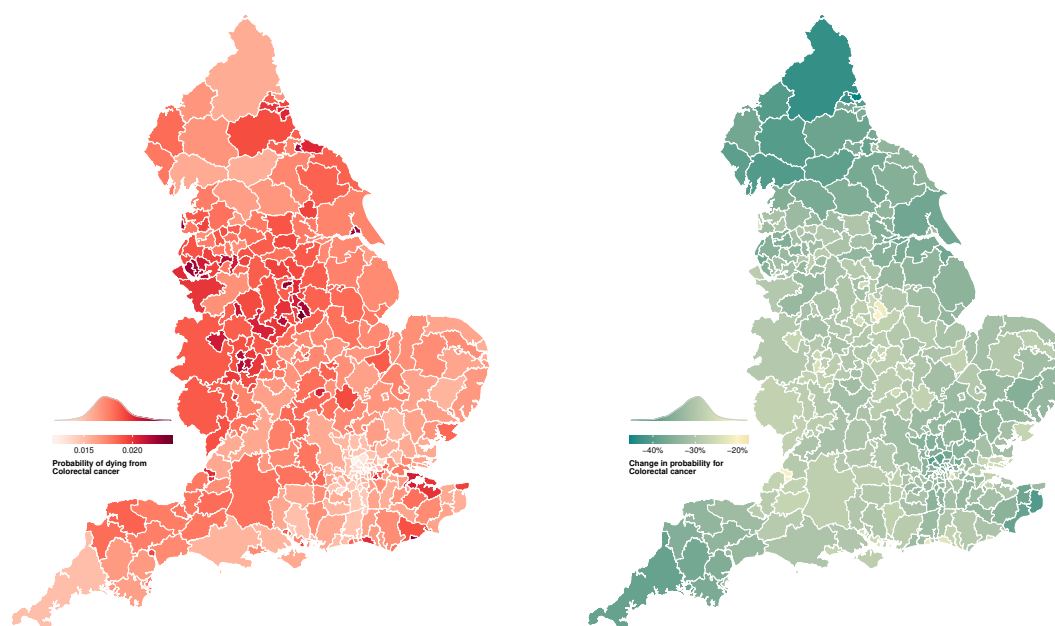


(b) Men

Figure E.9: Probability of dying from trachea, bronchus and lung cancers between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.

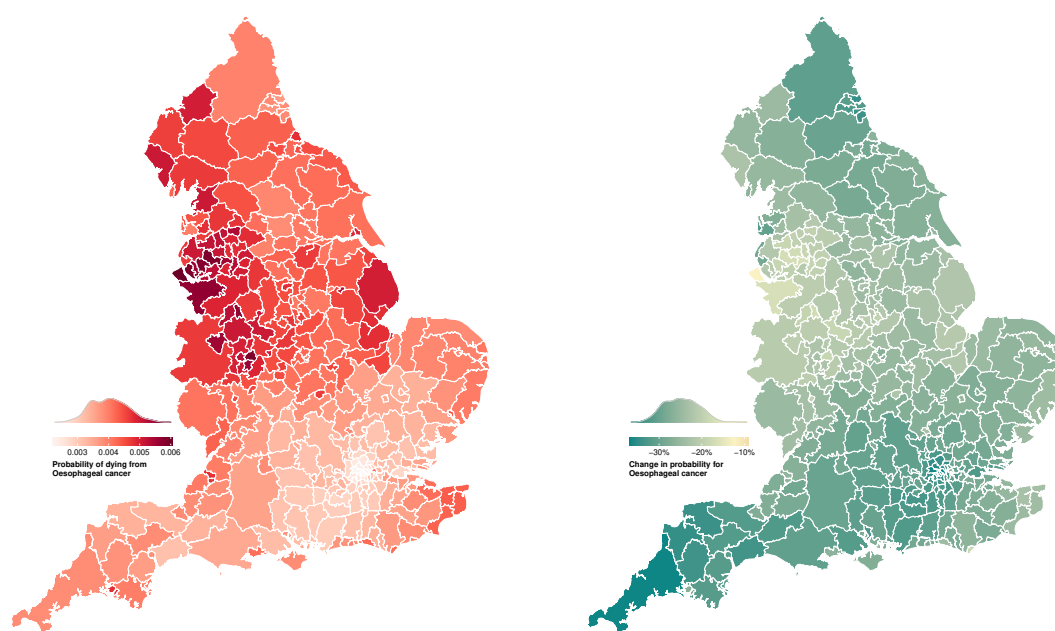


(a) Women

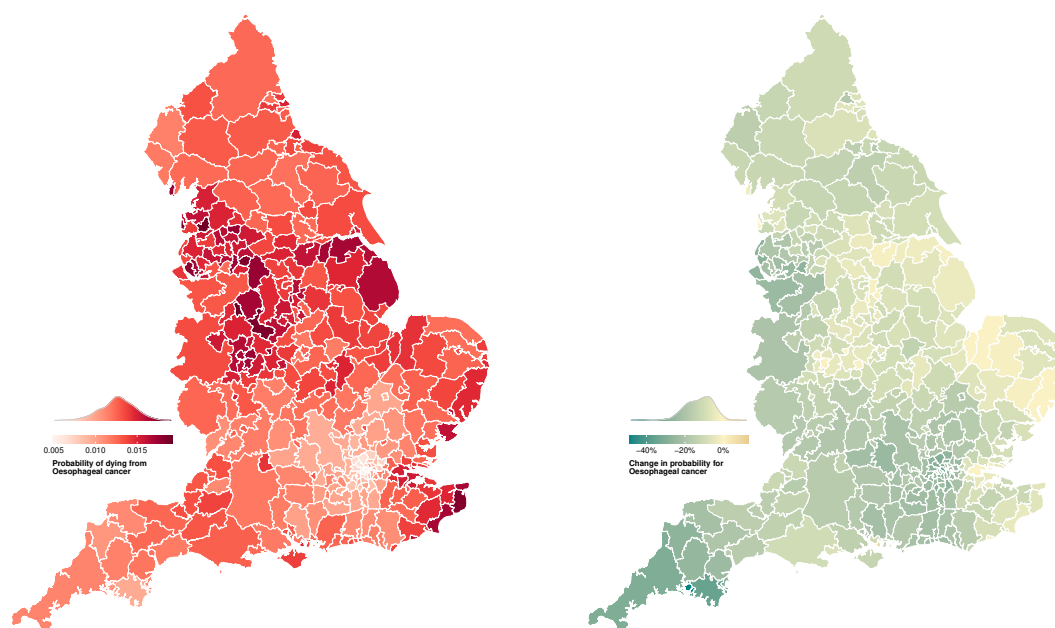


(b) Men

Figure E.10: Probability of dying from colorectal cancer between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.

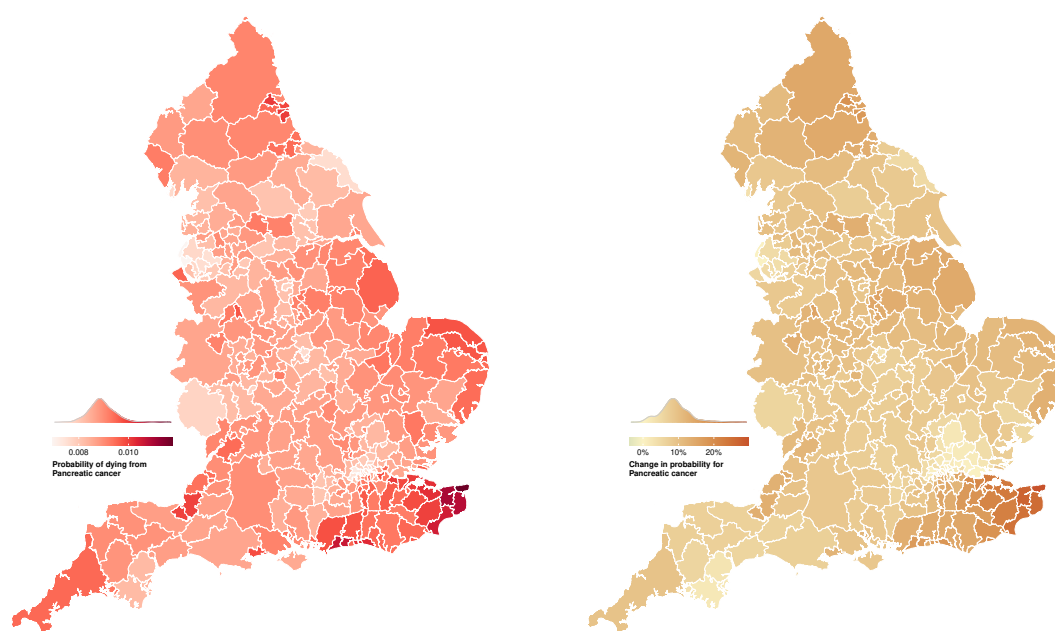


(a) Women

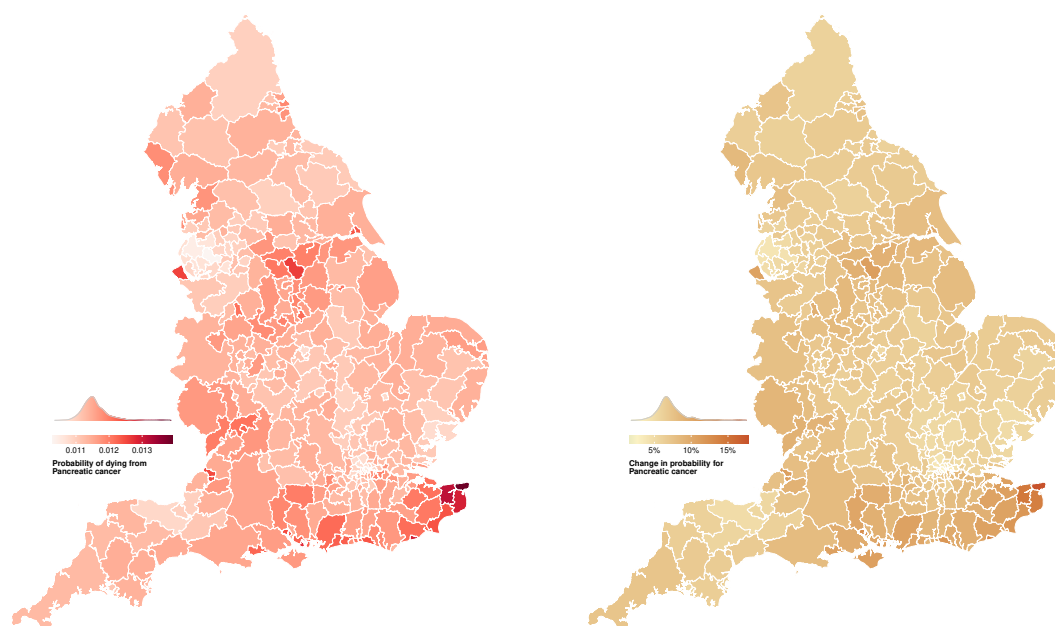


(b) Men

Figure E.11: Probability of dying from oesophageal cancer between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.

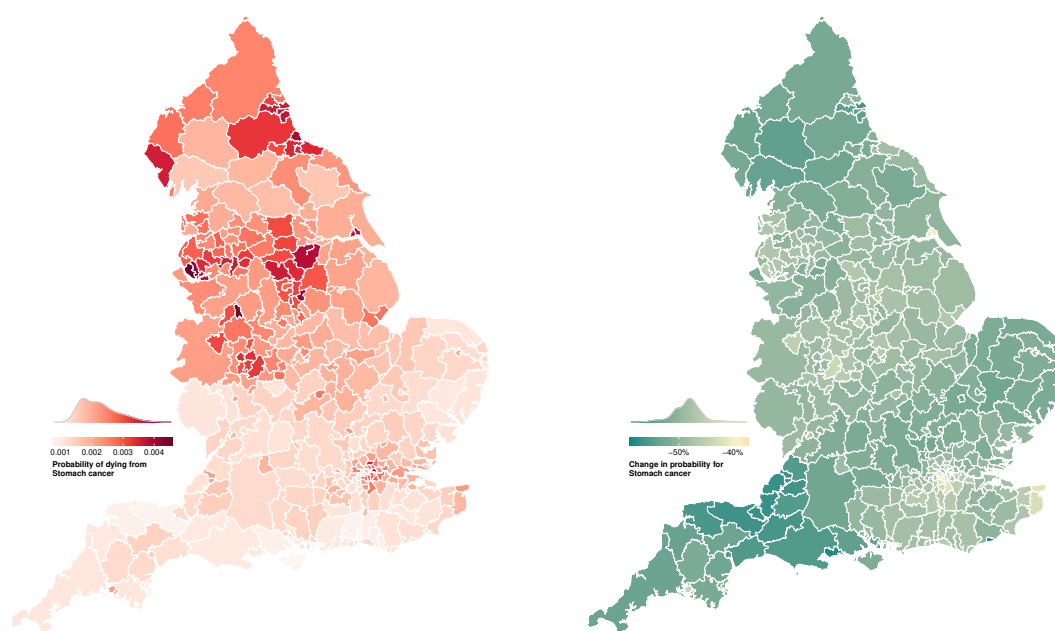


(a) Women

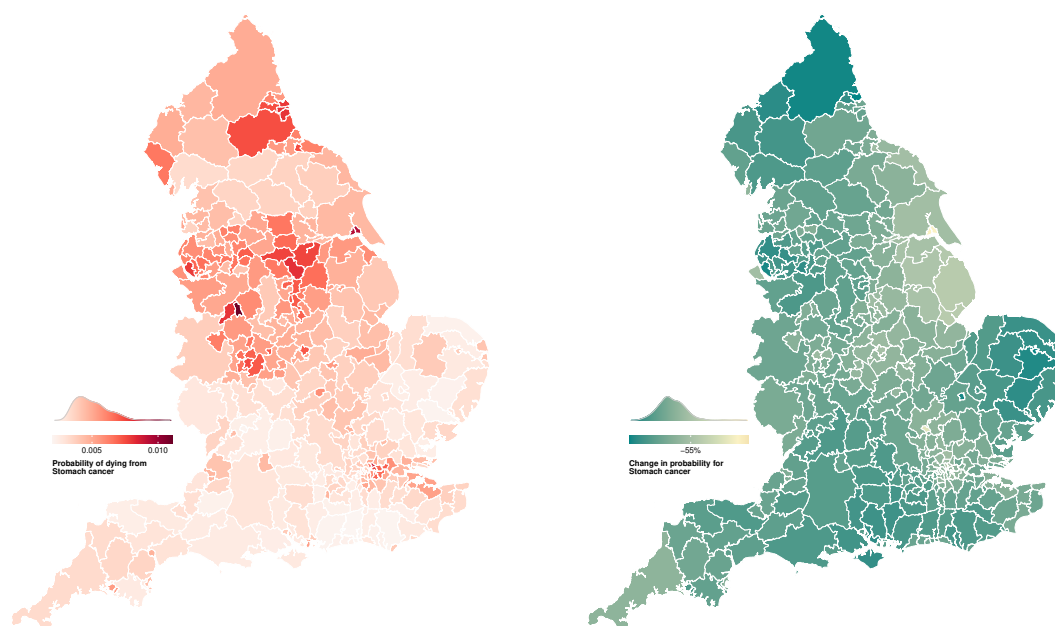


(b) Men

Figure E.12: Probability of dying from pancreatic cancer between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.



(a) Women



(b) Men

Figure E.13: Probability of dying from stomach cancer between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.

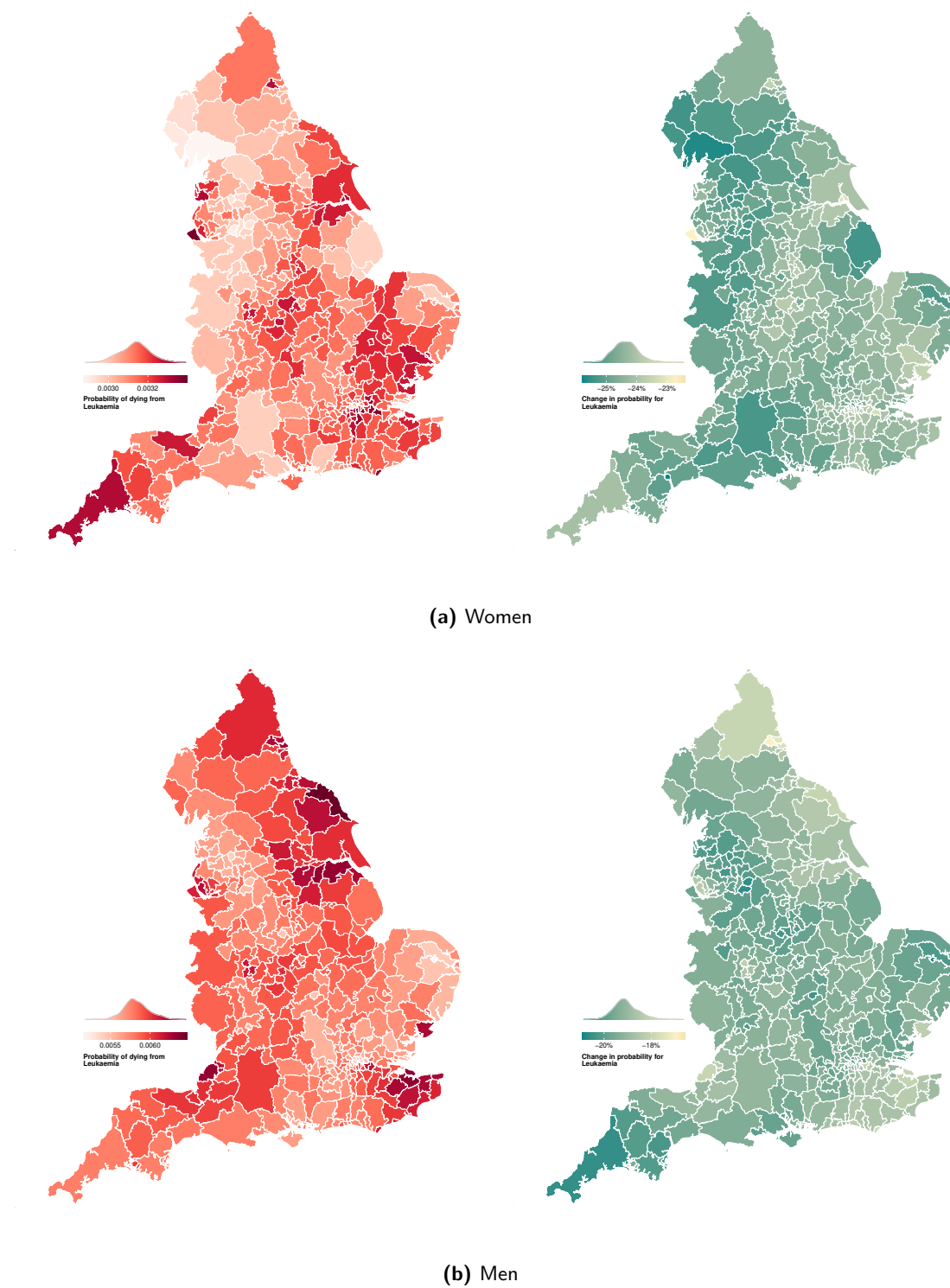
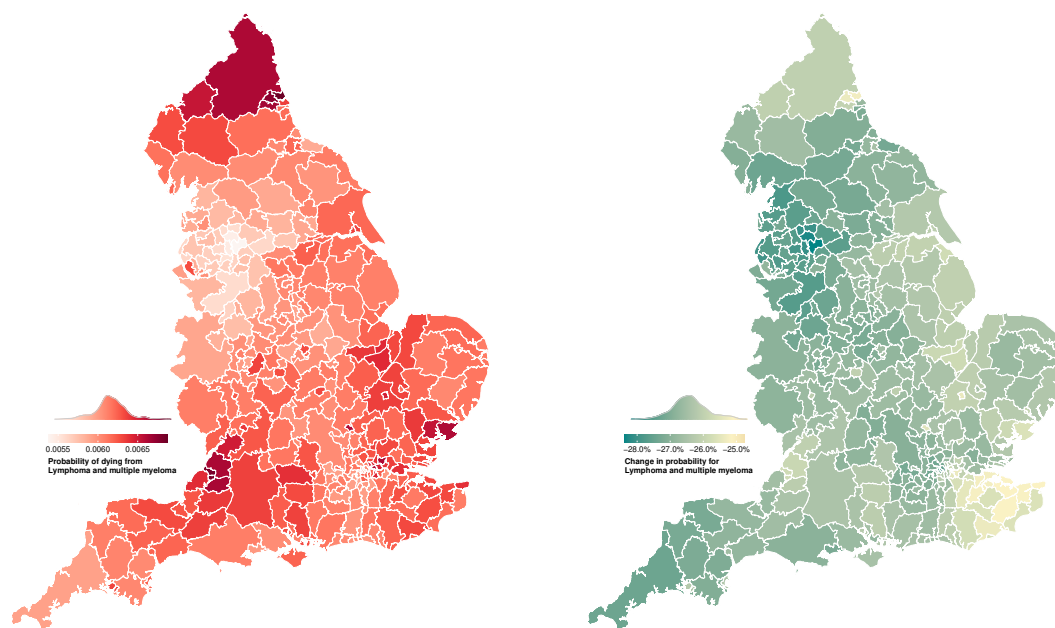
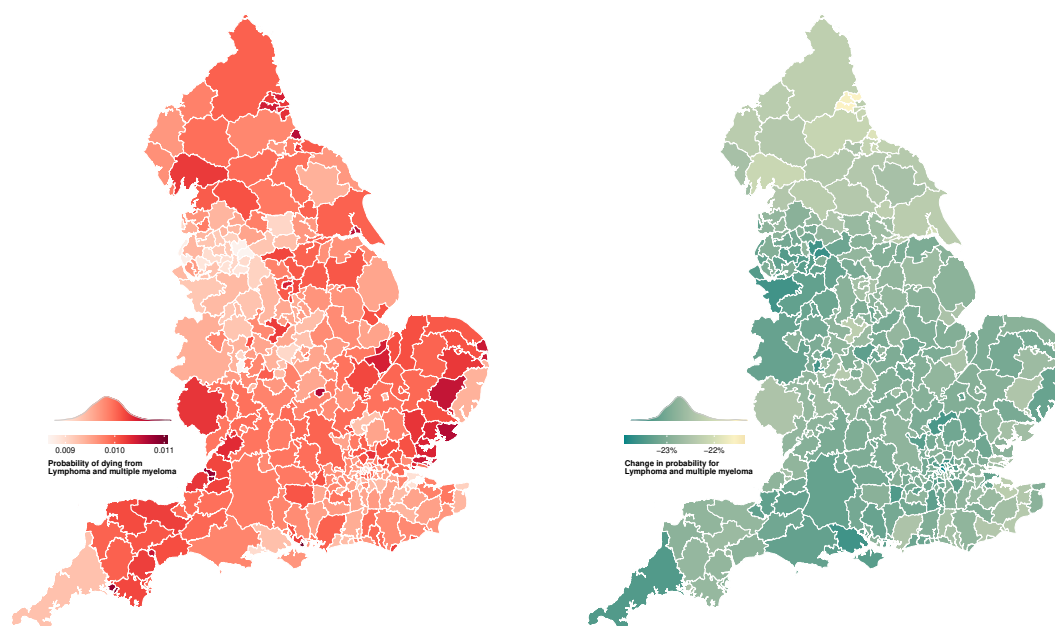


Figure E.14: Probability of dying from leukaemia between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.



(a) Women



(b) Men

Figure E.15: Probability of dying from lymphoma and multiple myeloma between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.

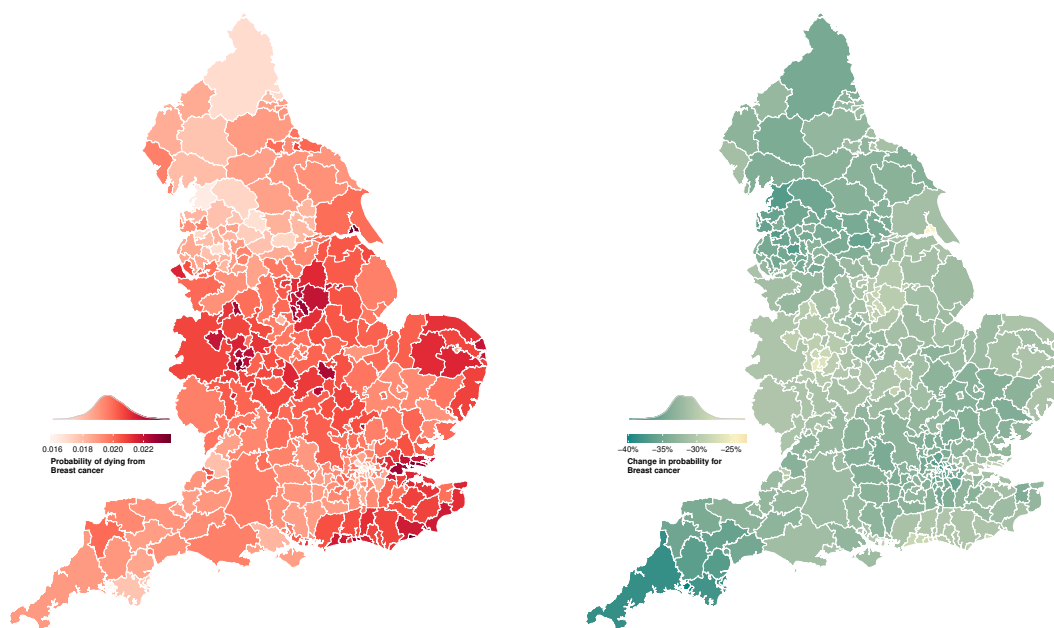


Figure E.16: Probability of dying for women from breast cancer between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.

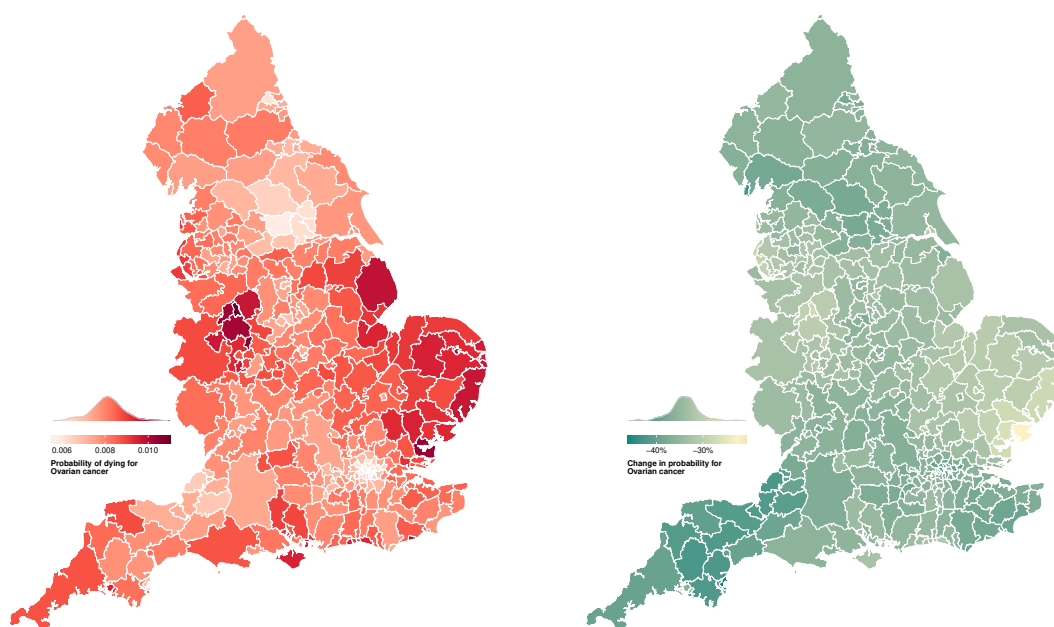


Figure E.17: Probability of dying for women from ovarian cancer between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.

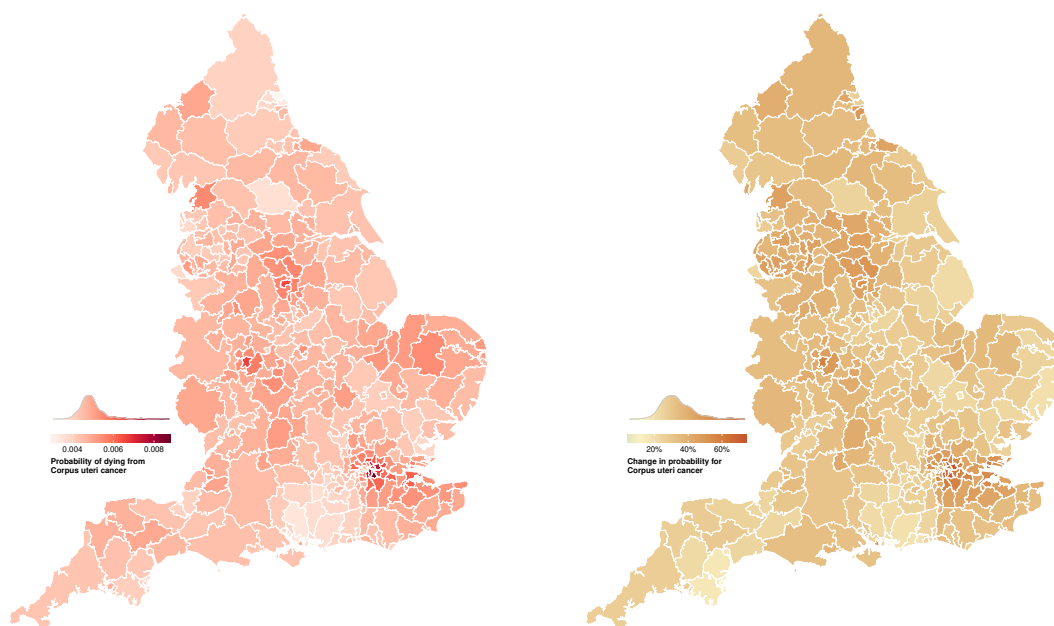


Figure E.18: Probability of dying for women from corpus uteri cancer between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.

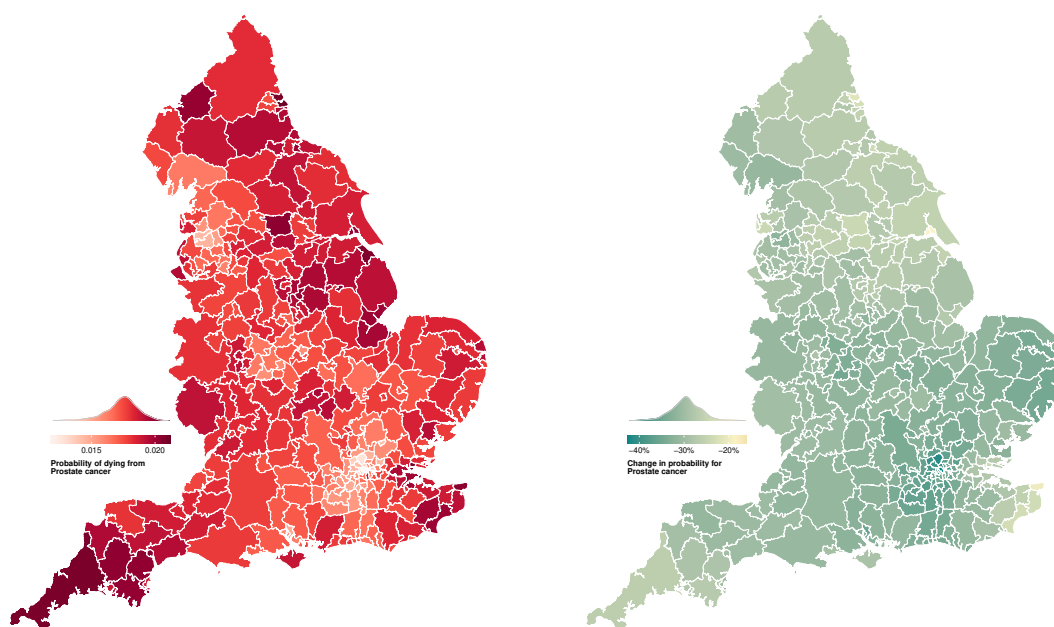


Figure E.19: Probability of dying for men from prostate cancer between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.

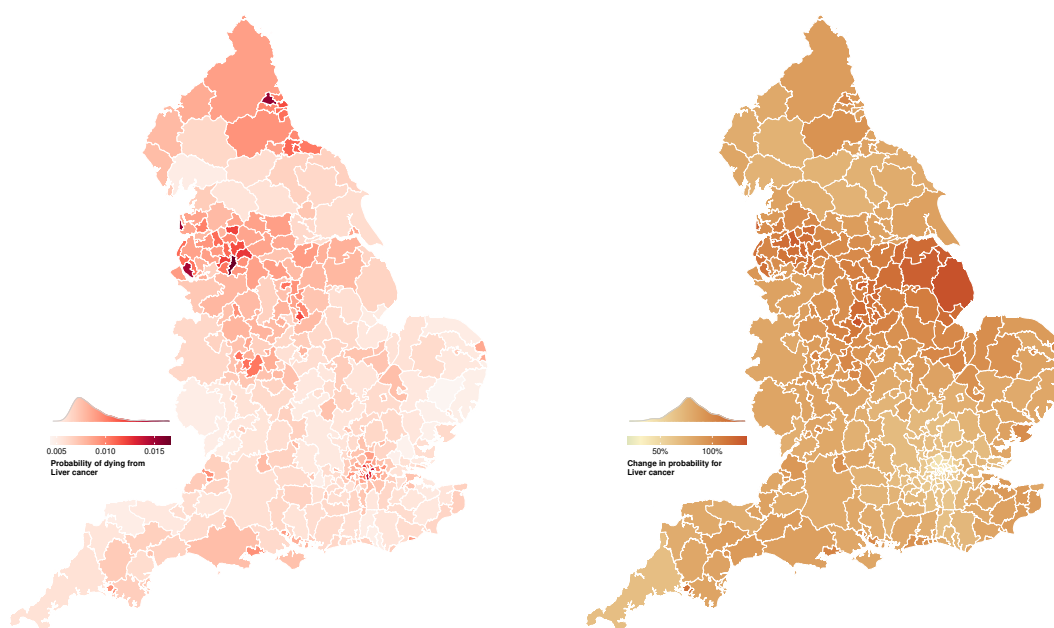


Figure E.20: Probability of dying for men from liver cancer between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.

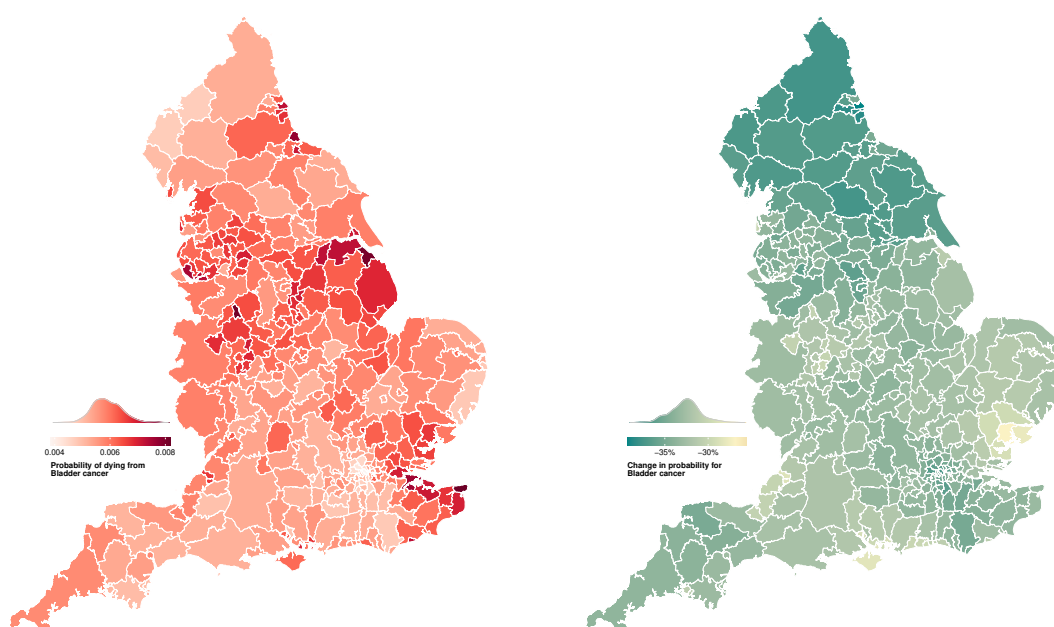
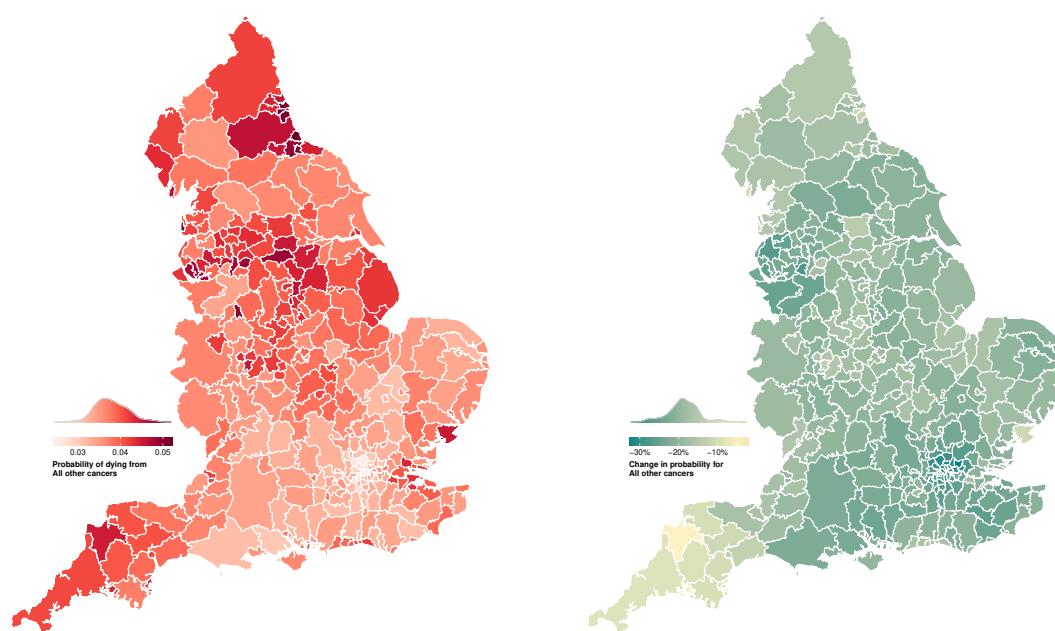
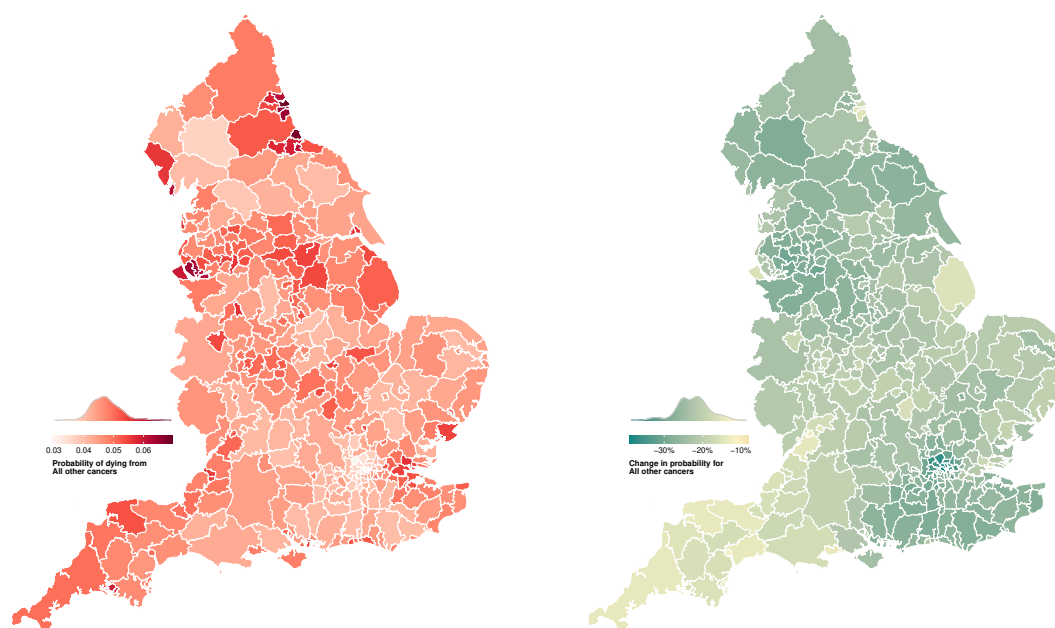


Figure E.21: Probability of dying for men from bladder cancer between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.

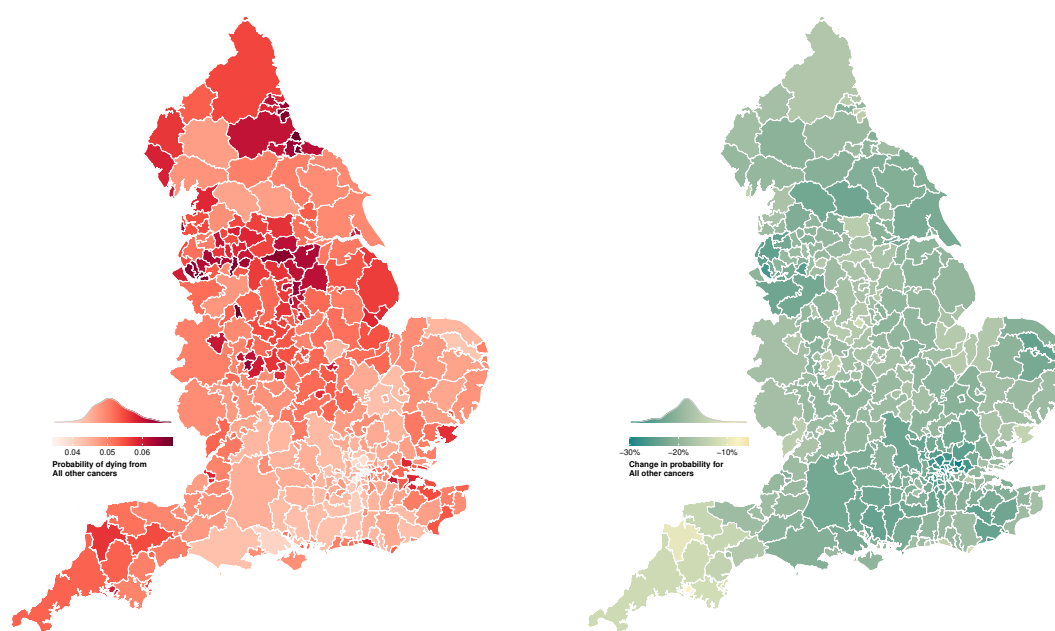


(a) Women

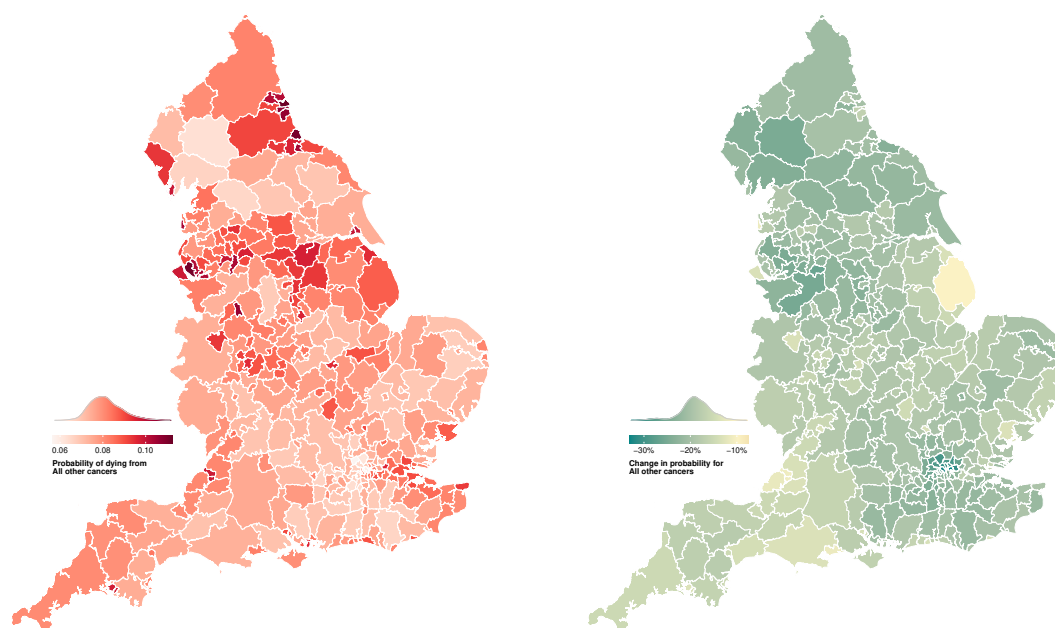


(b) Men

Figure E.22: Probability of dying from all other cancers (cancers only study) between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.

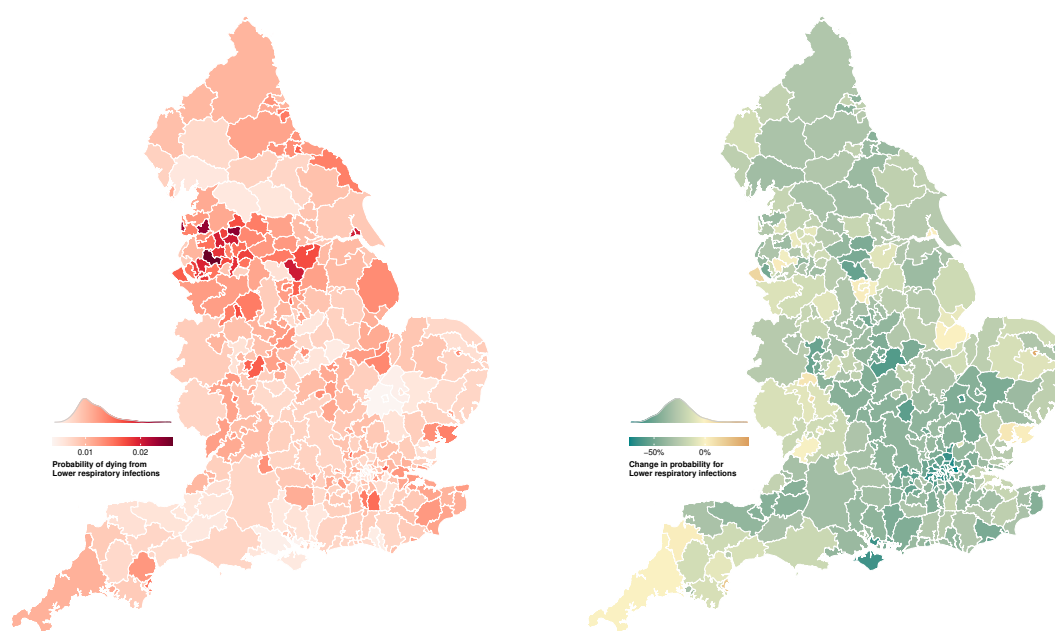


(a) Women

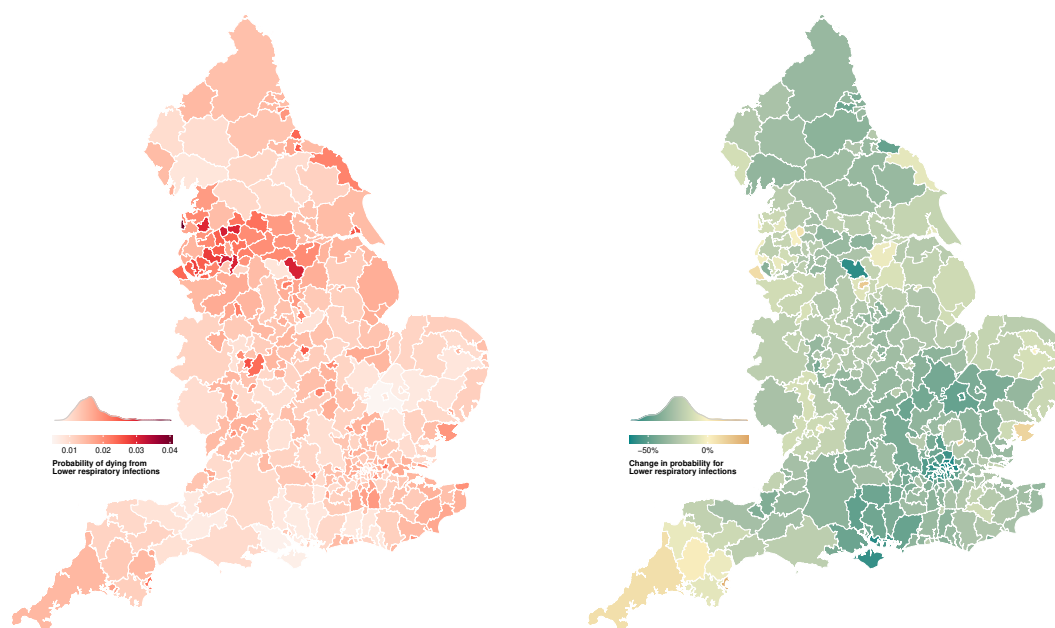


(b) Men

Figure E.23: Probability of dying from all other cancers (cause-specific mortality study) between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.

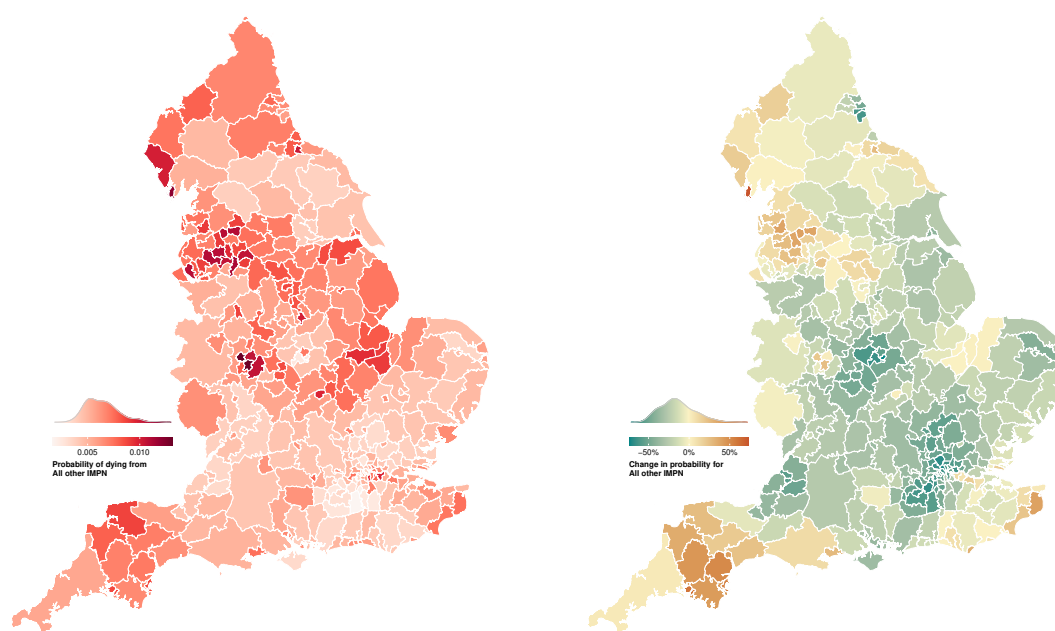


(a) Women

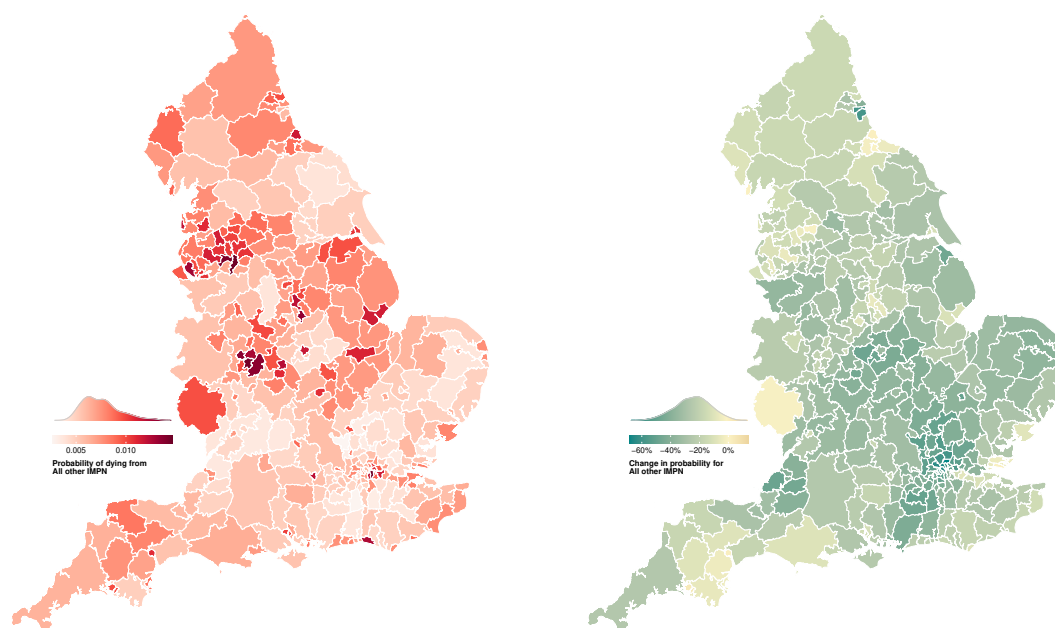


(b) Men

Figure E.24: Probability of dying from lower respiratory infections between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.

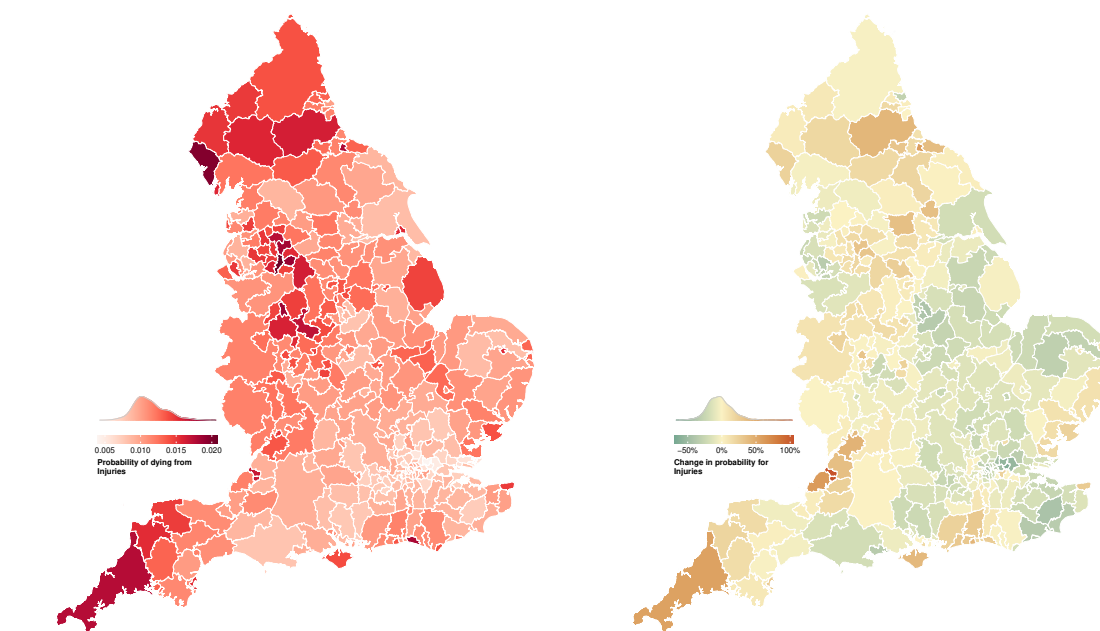


(a) Women

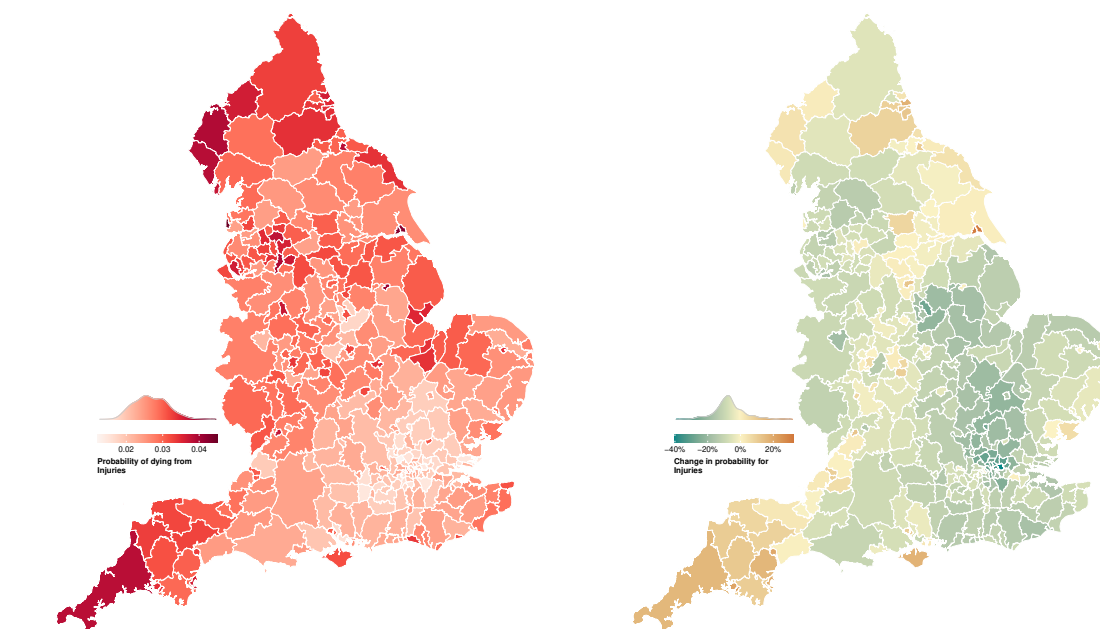


(b) Men

Figure E.25: Probability of dying from all other infections, maternal, perinatal and nutritional conditions between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.



(a) Women



(b) Men

Figure E.26: Probability of dying from injuries between birth and 80 years of age in 314 local authority districts in England in 2019 and change from 2002 to 2019.