1-1-2016

Three Essays on the Behavioral, Socioeconomic, and Geographic Determinants of Mortality: Evidence From the United Kingdom and International Comparisons

Laura Kelly
University of Pennsylvania, kellyla@sas.upenn.edu

Three Essays on the Behavioral, Socioeconomic, and Geographic Determinants of Mortality: Evidence From the United Kingdom and International Comparisons

Abstract
This dissertation contains three chapters covering the impact of behavioral, socioeconomic, and geographic determinants of health and mortality in high-income populations, with particular emphasis on the abnormally high mortality in Scotland, and the relative advantages of indirect and direct analyses in estimating national mortality. Chapter one identifies behavioral risk factors underlying mortality variation across small-areas in Great Britain, using the indirect estimation method of factor analysis on aggregate cause-of-death information from 1981-2009. Chapter two uses two indirect analytic methods to estimate the contribution of smoking to Scotland's high mortality and low sex differences in life expectancy relative to other high-income populations from 1951-2009. Chapter three performs survival analysis on first and second generation migrants using a national longitudinal study in England and Wales from 1971-2013 to quantify mortality variation by migrant status and the relative impact of socioeconomic status. The findings highlight the importance of health behaviors on aggregate mortality inequality, support the methodological advantages of indirect estimation of behavioral-attributable mortality, and exposes the importance of subgroup variation within national mortality estimates.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Demography

First Advisor
Michel Guillot

Keywords
Demography, Health behaviors, Mortality, Smoking, United Kingdom

Subject Categories
Demography, Population, and Ecology | Epidemiology | Medicine and Health Sciences

THREE ESSAYS ON THE BEHAVIORAL, SOCIOECONOMIC, AND GEOGRAPHIC DETERMINANTS OF MORTALITY: EVIDENCE FROM THE UNITED KINGDOM AND INTERNATIONAL COMPARISONS

Laura Amelia Kelly

A DISSERTATION

in

Demography

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2016

Supervisor of Dissertation

_____

Michel Guillot, Professor of Sociology

Graduate Group Chairperson

_____

Hans-Peter Kohler, Frederick J. Warren Professor of Demography

Dissertation Committee

Michel Guillot, Professor of Sociology
Samuel H. Preston, Professor of Demography and Sociology
Hans-Peter Kohler, Frederick J. Warren Professor of Demography

THREE ESSAYS ON THE BEHAVIORAL, SOCIOECONOMIC, AND GEOGRAPHIC DETERMINANTS OF MORTALITY: EVIDENCE FROM THE UNITED KINGDOM AND INTERNATIONAL COMPARISONS

ACKNOWLEDGMENTS

I would like to here acknowledge several individuals who significantly contributed to the evolution of this dissertation and to my general personal and professional well-being over the past four years. First and foremost, I am grateful for the expert mentorship and guidance of my dissertation committee: Profs. Michel Guillot, Sam Preston, and Hans-Peter Kohler. Their insightful instruction helped shape this dissertation, and their exceptional example influenced my development as a demographer and colleague. I am further indebted to the entire faculty within the Graduate Group of Demography and thankful for the tireless support of the administrative staff: Tanya Yang, Dawn Ryan, Nykia Perez, and Sheri Luckey. My cohort – Tom, Megan, Edith, David, and Ezra – has alternately been my sounding board, my study buddies, and my support group. They continually inspire me with their demographic brilliance and personal merit, and I look forward to calling them friends and colleagues for many years to come. I gratefully acknowledge the National Institute of Child Health and Human Development and the University of Pennsylvania Graduate School of Arts and Sciences for their financial support towards my degree and work-related travel. I am extremely thankful for the administrative and technical assistance from the Office for National Statistics, in particular the Longitudinal Study team, and Dr. Paul Norman. I am grateful for my second home at Penn, the Graduate Student Center, and its wonderful staff. Finally, I would like to express my gratitude for my family and friends, though I hope they already know the depth of my love and appreciation. There is no greater inspiration for me than the personal integrity, vibrant curiosity, innate generosity, sparkling humor, whole-hearted compassion, and fundamental joie de vivre displayed by my family on a daily basis. I hope they, especially, enjoy this dissertation because it would not exist without them.

ABSTRACT

THREE ESSAYS ON THE BEHAVIORAL, SOCIOECONOMIC, AND GEOGRAPHIC DETERMINANTS OF MORTALITY: EVIDENCE FROM THE UNITED KINGDOM AND INTERNATIONAL COMPARISONS

Laura Amelia Kelly

Michel Guillot

This dissertation contains three chapters covering the impact of behavioral, socioeconomic, and geographic determinants of health and mortality in high-income populations, with particular emphasis on the abnormally high mortality in Scotland, and the relative advantages of indirect and direct analyses in estimating national mortality. Chapter one identifies behavioral risk factors underlying mortality variation across small-areas in Great Britain, using the indirect estimation method of factor analysis on aggregate cause-of-death information from 1981-2009. Chapter two uses two indirect analytic methods to estimate the contribution of smoking to Scotland's high mortality and low sex differences in life expectancy relative to other high-income populations from 1951-2009. Chapter three performs survival analysis on first and second generation migrants using a national longitudinal study in England and Wales from 1971-2013 to quantify mortality variation by migrant status and the relative impact of socioeconomic status. The findings highlight the importance of health behaviors on aggregate mortality inequality, support the methodological advantages of indirect estimation of behavioral-attributable mortality, and exposes the importance of subgroup variation within national mortality estimates.

TABLE OF CONTENTS

Acknowledgements……iii
Abstract……...iv
List of Tables……..vii
List of Figures……viii

Chapter 1: Behavioral factors drive geographic mortality inequality in Great Britain …….……...1
Introduction……...1
Background……....1
Geographic variation in health in Britain……1
Health-related behaviors and mortality……..……..5
Methods for estimating behavior-attributable mortality……6
Data……....8
Methods……9
Results…….....12
Description of the study sample……12
Male factor analysis……...13
Female factor analysis……...18
Discussion……22
References……26

Chapter 2: The contribution of a history of heavy smoking to Scotland's mortality disadvantage……44
Introduction……..44
Background……..44
Scotland's life expectancy in international perspective……44
Explanations of Scotland's exceptionally high mortality……..46
Smoking in Scotland……...48
Methods for calculating smoking-attributable mortality……...49
Data……..51
Methods……51
Attributable-risk calculation in the PGW method……51
Results……..52
Smoking-attributable mortality above age 50, by sex and country, 1955-2009……52
Effect on life expectancy……...53
Discussion……55
References……56

Chapter 3: Mortality among first and second generation migrants in England and Wales …….67
Introduction……..67
Background……..70
Migration and mortality……70
Migrant health in the England and Wales……..……72
Data……..74
Sample……...74
First generation mortality: all-cause and behavior-attributable mortality…….……..…..75
Second generation mortality: infant mortality……...76
Defining migrant groups……....76

Socioeconomic variables: Individual-level……...77
Socioeconomic variables: Area-level……78
Remaining variables……..79
Methods……..…..80
First generation mortality: all-cause and behavior-attributable mortality……...…..80
Second generation mortality: infant mortality……...81
Results……..82
First generation mortality: all-cause mortality……..82
First generation preventable mortality: behavior-attributable mortality……..…..85
Second generation mortality: infant mortality……...88
Discussion……90
References……....94

Appendix……...117

LIST OF TABLES

Table 1.1. Deaths and population by sex and health geography, 1981-2009.

Table 1.2. Highest ranked causes of death by sex, 1981-2009.

Table 1.3. Rotated factor loadings 1981-2009, males.

Table 1.4. Rotated factor loadings 1981-2009, females.

Table 2.1. Life expectancy at birth, e(0), and age 50, e(50) in 2009 for Scotland and comparison countries.

Table 2.2. Estimated smoking-attributable fraction of deaths at age 50+ in 1955, 1980, 2009, by sex and country.

Table 2.3. Life expectancy at age 50 (e50) in 2009 before and after the removal of deaths attributable to smoking.

Table 3.1.A. Smoking-related diagnoses and ICD coding.

Table 3.1.b. Alcohol-related diagnoses and ICD coding.

Table 3.2. Number of England and Wales natives and first generation migrants, person-years at risk, deaths, and column percentages by covariates, England and Wales 1971-2013.

Table 3.3. Number of England and Wales natives and first generation migrants, person-years at risk, deaths, and column percentages by country of birth and mortality status, England and Wales 1971-2013.

Table 3.4. Hazard ratios of first generation all-cause mortality by covariates, England and Wales 1971-2013.

Table 3.5. Hazard ratios of first generation smoking-attributable mortality by covariates, England and Wales 1971-2013.

Table 3.6. Hazard ratios of first generation alcohol-attributable mortality by covariates, England and Wales 1971-2013.

Table 3.7. Hazard ratios of first generation all other cause mortality by covariates, England and Wales 1971-2013.

Table 3. 8. Number of second generation infants, person-years at risk, deaths, and column percentages by covariates, England and Wales 1977-2013.

Table 3.9. Hazard ratios of second generation all-cause mortality by covariates, England and Wales 1977-2013.

Table A3.1. Country of birth categories and individual countries.

LIST OF FIGURES

Figure 1.1. Age–sex bar chart by combined groups of cause of death in Great Britain, 1981–2004.

Figure 1.2. Diagram of hypothetical two factor model.

Figure 1.3. Factor 1 (Tobacco-use and Obesity) score map 1981-2009, males.

Figure 1.4. Factor 2 (Rurality and Substance Abuse) score map 1981-2009, males.

Figure 1.5. Correlation coefficient between age-specific all-cause mortality rates and factor scores 1981- 2009, males.

Figure 1.6. Factor 1 (Tobacco-use and Obesity) score map 1981-2009, female.

Figure 1.7. Factor 2 (Rurality and Substance Abuse) score map 1981-2009, females.

Figure 1.8. Correlation coefficient between age-specific all-cause mortality rates and factor scores 1981-2009, females.

Figure 2.1.A. Trends in life expectancy at age 50, e(50), among high-longevity countries, females 1950-2009.

Figure 2.1.B. Trends in life expectancy at age 50, e(50), among high-longevity countries, males 1950-2009.

Figure 2.2.A. Lung cancer death rates at ages 50 and above by country, women 1950-2009.

Figure 2.2.B. Lung cancer death rates at ages 50 and above by country, men 1950-2009.

Figure 2.3.A. Differences in life expectancy at age 50, e(50), before and after the removal of smoking-attributable mortality, Scotland vs. England & Wales, 1950-2009, women.

Figure 2.3.A. Differences in life expectancy at age 50, e(50), before and after the removal of smoking-attributable mortality, Scotland vs. England & Wales, 1950-2009, men.

Figure 3.1. Lexis diagram representing the study population.

Figure A1.1.A. Factor 2 score deciles and population-weighted rurality score deciles, males 1981-2009.

Figure A1.1.B. Factor 2 score deciles and population-weighted rurality score deciles, females 1981-2009.

Figure A2.1.A. Prevalence of daily smokers among women aged 15+ years, by year and country.

Figure A2.1.B. Prevalence of daily smokers among men aged 15+ years, by year and country.

CHAPTER 1

Behavioral factors drive geographic mortality inequality in Great Britain

Introduction

"We have made tackling health inequalities our top priority" – Nicola Sturgeon (Scottish Government, 2008)

The above statement was made in 2008 by then Cabinet Secretary for Health and Wellbeing, Nicola Sturgeon, who became First Minister of the Scottish Government in 2014. Reducing health inequality remains a policy priority for governments across the United Kingdom (UK), particularly in Scotland (Graham, 2009). Historically the predominant perspective on health inequality focused on individual or group characteristics and outcomes; however recent years have seen a growing interest in aggregate, sub-national geographies (Tyner, 2015). This ideological change is largely driven by the especially stark picture of geographic variation in mortality across Britain. As evidence of this geographic inequality, the (age-, sex-, and year-standardized) risk of mortality ranged from 50% above the national average to 76% below for 1,282 British neighborhoods measured over a 24 year period with over 14 million registered deaths (Shaw et al., 2008). Such imbalance has led to a call for population researchers to join epidemiologists and medical geographers in addressing these striking geographic health disparities (Boyle, 2004; Tyner, 2015).

Background

Geographic variation in health in Britain

The majority of UK health measures are estimated at the country (Scotland, England, Wales, and Northern Ireland) level because both administrative and non-administrative data are produced by

each independent national agency.[1] Thus UK official statistics are not easily comparable across national populations; and health statistics, in particular, are notoriously problematic. The UK Statistics Authority published a monitoring review of official health statistics in 2012, which criticized the "absence of central co-ordination of [health statistics] production and public availability" between countries, with data fluctuating in nature and quality due to "the fragmented nature of those statistics, attributable partly (but only partly) to the devolution of health policy in the UK to England, Scotland, Wales and Northern Ireland" (Dilnot, 2012). Accordingly, due to both technical and political barriers, geographic health inequalities in the UK are typically discussed in broad country-level perspectives.

Only two sources of official information exist where "statisticians have drawn together data from various sources and presented them on a consistent basis" to allow cross-national comparisons of health statistics: the discontinued United Kingdom Health Statistics (UKHS) series (Smith et al., 2010) (produced 2000, 2006, 2008, 2009, and 2010); and a 2008 cross-sectional UK Comparisons report produced using the Scottish Health Survey and the Health Survey of England (Bromley and Shelton, 2010). The UKHS series provides period estimates comparable between Scotland and England (Smith et al., 2010). According to the series, the Scottish population consistently ranks worse than the English population across time periods, data sources, and indicators. Scottish men and women have the highest all-cause mortality rates[2] in the United Kingdom; and Scottish men and women have the worst life expectancy at birth (LE), healthy life expectancy (HLE), and disability free life expectancy (DFLE) in the United

[1] Scotland, England, and Wales will be referred to as countries or nations throughout this report, according to common practice of political terminology used by the UK Office of National Statistics (Classifications and Harmonization Unit, 2013).

[2] Age-standardized death rates, standardized to the 1976 European Standard Population, by sex.

Kingdom.[3] In terms of glaring cause-specific differences, death rates for mental and behavioral disorders are 3 times higher for Scottish men relative to English and Welsh men, and 2 times higher for Scottish women relative to English and Welsh women. These mental and behavioral differences are likely driven by Scotland's high suicide rate, though death rates related to drug poisoning are also exceptionally high in Scotland (approximately 3 times higher than England for both sexes). Higher death rates due to malignant neoplasms among Scottish men and women relative to English and Welsh men and women are likely due to increased smoking-attributable mortality, particularly lung-cancer (Kelly & Preston 2016).

For health behaviors, finding comparable behavioral estimates across UK geographies is complicated or impossible due to the lack of comprehensive national surveillance systems and limited analogous regional health surveys. The Scottish Health Survey (1995+) and the Health Survey for England (1991+) allow select health behavior comparisons, as the studies are similarly structured and a few key questions asked in each survey. According to the 2008 UK Comparisons report produced using these two health surveys, the prevalence of self-reported cardio-metabolic[4] diseases is 1-2 percentage points higher for Scottish men and women relative to English men and women for all ages combined and across each age category (Bromley and Shelton, 2010). Fewer Scottish men and women report not drinking in the previous week relative to English men and women, with the gaps especially large at the youngest (16-24 years) and oldest (75+) age groups. Scottish men and women generally report a higher mean units of alcohol consumed in the previous week relative to England, with the difference greatest in the younger female age groups.

[3] HLE is calculated using the Sullivan method, with the General Household Survey (GHS) and the Census. HLE is defined as the number of years of life spent in self-reported 'Very good' or 'Good' health, and DFLE is defined as the number of years lived free from a self-reported limiting chronic illness or disability.

[4] Any CVD (angina, heart attack, stroke, heart murmur, abnormal heart rhythm, or other heart trouble); Any CVD or diabetes (the preceding category plus diabetes); IHD or stroke (angina, heart attack, and stroke).

A higher proportion of Scottish men and women report being current smokers[5] relative to England, for all ages combined and across age groups. Finally, the proportion overweight and obese is higher for Scotland relative to England for both sexes, with the difference larger for women.

These estimates describe country-level disparities in health and mortality. However, sub-national mortality variation further reveals strongly patterned inequality. Using vital statistics for 1,282 British neighborhoods aggregated from 1981 to 2004, a recent report summarized geographic mortality variation for 9 aggregate cause of death categories (Shaw et al., 2008). While deaths are expectedly driven by age and sex (Figure 1.1), the report confirmed a general "stark divide in mortality between north and south" and exposed huge mortality variation across neighborhoods [Ibid.]. For all neighborhoods combined, the average age of death in Britain was 74.4 years, 71.2 for men and 77.4 for women. However by neighborhood, the average age of death ranged from a low of 66.4 years (Glasgow, Scotland) to 80.6 years (Eastbourne, England).

This report further estimated that 6.6 million deaths (44% of total deaths) were from cardiovascular diseases, with over half of these due to chronic heart disease and heart attacks. While the risk of mortality from cardiovascular disease increases moving from rural areas to cities, the North-South divide is sustained and stark with cardiovascular disease standardized mortality ratios (SMRs) highest in Scotland, particularly Glasgow. Approximately 3.7 million deaths (25% of total deaths) were from cancer, a quarter of which were lung cancer deaths. The risk of mortality from all cancers also follows a North-South pattern, being highest in northern, Scottish cities. Suicide and deaths attributable to mental disorders are highest in cities across Britain, but unambiguously concentrated in Scotland relative to England. The SMR's for suicide

[5] Current smokers, as opposed to former smokers, are a combination of self-reported "occasional" or "regular" smokers.

and deaths attributable to mental disorders are 4-5 times higher in Scottish areas relative to English areas for both sexes.

Health-related behaviors and mortality

The largest geographic disparities in Britain seem driven by causes of death linked to health behaviors: cardiovascular disease, cancers (e.g. lung cancer), and injuries (e.g. suicide). Thus the limited available evidence on geographic variation in mortality across Britain suggests the importance of behavior-driven mortality. This study aims to comprehensively investigate the contribution of health behaviors on geographic mortality variation in Great Britain.

Health behaviors have been found to influence all-cause and so-called 'preventable' mortality globally (Ezzati et al., 2002; Ezzati et al., 2003; Danaei et al., 2005; Danaei et al., 2010; Finucane et al., 2011; Ikeda et al., 2012; Di Cesare et al., 2013; Jonker et al., 2015; Lim et al., 2015). Though precise definitions vary, preventable mortality is usually defined as deaths occurring before age 75 due to causes considered to be preventable through (a) individual behaviors and/or (b) public health measures aimed at changing behaviors or exposures to harmful environments (Hutchison et al., 2006; Wheller et al., 2007). Examples of preventable causes of death include lung cancer, substance-abuse disorders, and alcohol-induced road traffic accidents. Within this framework, health behaviors can then be considered mortality risk factors.

Behavioral risk factor information can be directly observed at the individual level by asking respondents about current or past health behaviors. Most studies focus on limited, relatively concrete behaviors in the domains of physical activity, smoking, drinking, and nutrition (Hofstetter et al., 2014). Other behavioral indicators such as psychological stress, social networks, motor vehicle behavior, sexual behavior, and illicit drug use are less reliable and less readily captured due to reporting bias (e.g. social desirability bias, recall bias), hidden populations, or insufficient survey instruments (National Research Council et al., 2015). Evidence from the US

(Ezzati et al., 2006) and the UK (Shiely et al., 2013) both find significant population-level bias in self-reported health information, sensitive to time and the data collection method.

Another added complexity is that an individual's behaviors evolve in scope and consequence across the life-course. Behavioral risk factors are multi-dimensional and cannot be captured by a single indicator at a single point in time. For example, smoking behavior could be measured by age at initiation, years since quitting, nicotine content, and inhalation. Behavioral risk factors for obesity could include temporal physical inactivity, consumption of fruits and vegetables, alcohol intake, smoking behavior, and prescription drug use.

On top of these inherent data quality issues for individual health behaviors, data availability is also a fundamental problem. A recent National Research Council Report found that few countries conduct systematic surveillance of behavioral risk factors (Crimmins et al., 2010), making comparable direct analyses of health behaviors incomplete or impossible across geographies and time.

Methods for estimating behavior-attributable mortality

An improved population approach would be to indirectly estimate the causal effect of behavioral risk factors on mortality. For example, the population-level impact of smoking behavior (Peto et al., 1992; Preston et al., 2010a; Kelly and Preston, 2016) has been estimated using aggregate cause-specific mortality information in the US.

The common epidemiological approach to estimating mortality attributable to behaviors utilizes population attributable fractions. A population attributable fraction (PAF) is defined as the proportion of population mortality burden causally explained by a specific risk factor or set of risk factors. By extension, the PAF is the proportion of mortality that would be eliminated from the population if exposure to that risk factor(s) was counterfactually removed. The PAF calculation relies on two basic inputs: the proportion of the population at each exposure level,

exposed or unexposed to the risk factor(s) of interest, and the relative risk of mortality by exposure status.

The PAF is historically a contentious measure due to its reliance on accurate estimation of risk factor distribution in the sample population, analytic sensitivity to the relative risk estimates used, and an assumption of the causal pathway between exposure and mortality outcome. Risk factor information comes from sample surveys, which again often rely on self-reported behaviors, and thus can be subject to respondent bias or limited scope. The causal pathway from exposure to outcome may also be poorly understood and thus inadequately accounted for in PAF calculations. A given PAF could greatly underestimate the mortality burden attributed to a specific risk factor if only considering excess deaths and not deaths etiologically related to the exposure (Greenland and Robins, 1988; Levine, 2007).

Tencza et al. recently conceptualized behavioral risk factors as latent variables identified through exploratory factor analysis of cause-specific mortality data (Tencza et al., 2014). Factor analysis, sometimes called latent variable analysis, uncovers latent variables or 'factors' through observed covariation, here covariation in cause-specific mortality across populations. Factors then can be conceptually interpreted in terms of known behavioral risk. Unlike the PAF approach, factor analysis offers several unique advantages in understanding the behavioral drivers of mortality inequality. First, factor analysis does not rely on a priori selection of behavioral risk factors of interest. Rather, this approach allows natural patterns to emerge and incorporates all potential mortality risks. Secondly, latent variable analysis relies only on high-quality vital statistics data. Aggregate administrative data bypasses the need for survey-derived behavioral information at the individual level with its incumbent biases and limitations.

This study uses factor analysis to understand geographic mortality variation by cause for all persons aged 15-74 in England and Scotland between 1981-2009. This specification is chosen for two reasons. First, causes of deaths at older ages are less reliably reported. Secondly, deaths

within this age range can be considered 'premature'[6] and capture preventable mortality according to established definitions. The age range begins lower than Tencza et al. (Tencza et al., 2014) due to the known mortality variation of young men, particularly in Scotland, in the UK. The age ranges extends to 74, beyond Tenzca et al., to align with the standard upper age limit in preventable mortality studies and to capture more death counts in low density areas. Understanding the behavioral drivers of preventable mortality is vital to quantifying health inequality, identifying potential areas for policy interventions, and regulating national population health. To the author's knowledge, factor analysis has never been applied to cause-specific mortality information in the United Kingdom nor has the data necessary to conduct this analysis been assembled.

Data

Death counts by decade, age, area, and cause of death were obtained by a special extraction request from the Office for National Statistics (ONS) Mortality Analysis Team. Annual mid-year population counts for health geographies were aggregated using publically available mid-year population estimates released by the Population Estimates Unit, UK Statistics Authority, ONS (UK Statistics Authority, 2014).

Due to data restrictions, only years 1981-2009 were available for release. Five year age groups were used, and temporally consistent health geographies defined the areas of analysis (25 National Health Service Areas in England, 14 National Health Service Health Boards in Scotland). For England, health geographies were originally provided as 211 clinical commissioning groups, which were manually aggregated into 25 larger National Health Services areas to create geographically consistent areas over time. In Scotland, the 14 Health Board

[6] In 2009, the life expectancy at birth in England and Wales was 82.43 years for females and 78.29 for males (Human Mortality Database, 2014). For Scotland, the life expectancies for females and males were 80.45 and 75.87 years, respectively [Ibid.].

geographies were constructed to be geographically consistent by the Vital Events Statistics team of the National Records of Scotland prior to data extraction. Thus due to special data extraction and manual manipulation, this study uniquely analyzes cause-specific mortality in UK geographies that are consistent over time.

Causes of death were coded according to a public health-oriented causes of death list recently developed by the Global Burden of Disease Study and the World Health Organization (Naghavi et al., 2010) using the International Statistical Classification of Diseases and Related Health Problems (ICD) editions 7-10. This list categorizes all deaths into one of 56 causes. The one alteration to this list is the merging of 'Alcohol-use disorders' with 'Alcohol-induced Liver Cirrhosis' into a common cause of alcohol-attributable mortality.

Deaths rates are calculated by 5-year age group, health geography, sex, and cause of death. This analysis pools across the three decades to boost death counts in low population density geographies, particularly in northern Scotland, and thus enhance stability of rate estimates. Death rates are standardized using a revised 2013 European Standard Population (ESP) according to common practice by the ONS (Olatunde, 2013; Pace et al., 2013), though the results are robust to using the original 1976 ESP. Causes of death contributing little to the overall mortality burden (those with counts less than 10,000 deaths for both sexes across all three decades for Great Britain as a whole) were dropped from analysis. The results are generally robust to alternate specifications of death counts exclusions. The analysis was repeated dropping causes with overall death counts less than 7,500 (to increase the list of causes entering the analysis) or 12,500 (to decrease the list of causes entering the analysis).

Methods

Exploratory factor analysis is a variable reduction technique that simplifies and summarizes interrelated data by extracting condensed, conceptual constructs. Exploratory factor analysis,

hereafter referred to as simply factor analysis (FA), identifies a few common factors that largely explain correlations between a larger set of observed variables.

FA is often misidentified as another variable reduction technique, principal component analysis (PCA). In PCA, observed variables are aggregated into components that explain total variance within the data. The number of extracted components equals the number of observed variables used in the analysis, and the analyst subjectively retains the minimal number of components explaining the maximum amount of total variance in the data. Alternately, FA hypothesizes underlying latent variables not directly measured in the data that causally explain shared variance (or covariance) between observed variables. PCA reduces the data into a smaller number of components whereas FA identifies what latent constructs structure the data. FA also conceptually and methodologically accounts for measurement error. In FA, each observed variable has two contributing sources of variation: a common factor and a unique factor (or specific error).

Figure 1.2 illustrates a hypothetical two factor model, with three observed variables and their associated unique factors explicitly capturing observed variable error. In the mathematical language of this analysis, observed variables are a function of simultaneous linear equations of the underlying common factors and unique factors as follows:

$$x_{ai} = f_{a1}b_{1i} + f_{a2}b_{2i} + \ldots + f_{am}b_{mi} + e_i \qquad i = 1, 2, \ldots, c \qquad (1)$$

$$a = 1, 2, \ldots, n$$

where

f_k = common factors (k = 1, 2, …, m)

b_i = factor loadings (i = 1, 2, …, c)

e_i = unique factors (i = 1, 2, …, c)

Here, a set k of underlying common factors linearly predict cause-specific death rate Xai the i^{th} cause of death among c independent causes, corresponding to the observation from area a of n independent areas. The observed data is organized as a matrix of rows corresponding to n observations (areas) and columns to c observed dimensions (cause-specific death rates). This analysis assumes that the observed variable x is linearly related to factor f, that unique factors are uncorrelated to common factors, and that unique factors are mutually uncorrelated.

Factor loadings, the set of linear coefficients in equation (1), describe both how variables are weighted for each factor and the correlation between the variables and the specific factor. The larger the factor loading, then the more that variable contributes to the dimensionality of the factor. Each variable's unique factor is also estimated, which estimates the variance unique (uniqueness) to that variable and not shared with other variables in the data. Communality describes the variance of observed variables accounted for by common factors and equals the sum of squares of factor loadings. By definition, Uniqueness equals 1-Communality. As FA aims to explain variance through common factors, variables with low communality (or conversely high uniqueness, generally 80% or higher) are usually dropped from analysis. Each factor should define a discrete cluster of interrelated variables, and thus cross-loadings (when a variable loads at 0.32 or higher across 2 or more factors) should be avoided (Costello and Osborne, 2005). To negate this interpretive issue, a statistically meaningful factor loading cut-off is used to decide how many factors to retain.

In the decision of how many factors to retain, several subjective criterion have been proposed including the Kaiser criterion (retain those factors with eigenvalues equal or higher than 1) (Kaiser, 1960) and Jolliffe's criterion (retain factors with 0.70 or higher) (Jolliffe, 1972). In subsequent years, both criterion have been criticized in overestimating the number of retainable factors (Costello and Osborne, 2005), and common practice now dictates the use of scree plots and eigenvalues. A scree plot allows a visual representation of eigenvalues against underlying

factors, and often reveals a point of inflexion in the diagonal. The number of factors before the point of inflexion suggests the number factors to be retained, i.e. factors with meaningfully large eigenvalues.

After deciding the number of factors to be retained, factor loadings are rotated to aid in interpretation. Rotation maximizes factor loadings on each variable while loading each variable on the minimal number of factors as possible to avoid cross-loading. To achieve this, varimax rotation is used, a type of orthogonal rotation where the factors are rotated 90° from each other and assumed to be uncorrelated. Varimax rotation maximizes the squared loadings for each factor.

Rotated factor loadings can then be used to create factor scores for each individual area, which are predicted factor estimates that relatively rank each individual area on the factor. The most common methods of factor scoring are regression-based, also called Thomson's regression method (Thomson, 1939), or Barlett's method (Bartlett, 1938). Both methods produce standardized factor score estimates which relatively rank each individual area on the factor (Distefano et al., 2009). The regression-based method is more widely used and essentially applies least-squares regression on the factor analysis estimated parameters to predict factor scores. Bartlett's method alternately uses maximum likelihood estimation and produces unbiased estimates of the true factor scores, though with potentially less validity than the regression-based method. The results shown here are robust to either factor scoring method used, though only regression- based results are shown for brevity. All analyses are carried out using the statistical package STATA 13.1.

Results

Description of the study sample

Table 1.1 shows the aggregate death and mid-year population counts across 1981-2009 by sex and health geography, which collectively comprise the sample population. The largest geographic

unit in England is London, with over 337,000 deaths of men and women recorded during the study period. This count is over 2.2 times higher than the cumulative death count of the largest geographic unit in Scotland, Greater Glasgow and Clyde, which had only about 150,000 deaths registered over the period. Scottish geographies are uniformly less populated than English geographies, with the smallest counts occurring in the northern Scottish island areas of the Western Isles, Shetland, and Orkney.

Table 1.2 shows the ten highest ranked causes of death, by sex, over the study period for the entire sample population. Age-specific death rates for persons aged 15 to 74 for all geographies combined were calculated for each cause and ranked for each sex. Non-communicable diseases, particularly cardiovascular diseases and malignant neoplasms, ranked the highest for both men and women as expected; these diseases contribute the greatest disease burden for high-income countries. Nine of the top ten causes were non-communicable diseases for both men and women, and ischemic heart disease was the leading cause of death. The only communicable disease which breached the top ten was lower respiratory infections. Lower respiratory infections, primarily driven by acute bronchitis and pneumonia, are established leading causes of death among the elderly and thus prominent among aging populations. Self-inflicted injury is the tenth leading cause of preventable mortality among males over this study period. Suicide has been documented to be relatively high in the UK, especially for Scottish males (Windfuhr and Kapur, 2011; McCartney, Shipley, et al., 2012).

Male factor analysis

Eigenvalues and scree plots were used to determine the optimal number of factors to retain, as discussed above. For the male sample, the first three eigenvalues were 20.62, 2.19, and 0.86. The Kaiser criterion would suggest retaining only the first two factors, having eigenvalues higher than 1. The scree plot confirmed this judgement, with the point of inflection in the diagonal occurring at the second factor. Following varimax rotation, the two-factor solution accounted for 89% of

the total variance (Factor 1: 53%; Factor 2: 36%) in cause of death rates observed across all areas. Table 1.3 shows the rotated male factor loadings. Only factor loadings higher than 0.70 (<0.001) are reported for brevity and because these loadings have the greatest contribution to the factor structure.

For males, the first factor is loaded with causes of deaths strongly associated with the behavioral risk of tobacco use, especially cigarette smoking, and to a secondary extent, diet (obesity). Medical information associated with each cause of death is obtained primarily from the Medical Encyclopedia of MedlinePlus, a National Institute of Health (NIH) online database produced by the National Library of Medicine (NIH MedlinePlus). Chronic obstructive pulmonary disease (COPD) has the highest correlation with the first factor. The leading cause of COPD is cigarette smoking. Stomach cancer has the second highest loading and affects primarily older males with a history of smoking, obesity, and a diet of salted meats. Smoking increases the risk for adult leukemia (Brownson et al., 1993), particularly for males, and with no strong hereditary risk, which strengthens attribution to the smoking behavioral component. Tobacco use (both smoking and chewing) causally raises the risk of developing peptic ulcer disease. Up to half of all bladder cancer cases are attributed solely to cigarette smoking. Lower respiratory infections, predominantly pneumonia, are strongly linked with smoking behavior. Chronic lung diseases (e.g. COPD), cigarette smoking, and co-occurring chronic heart disease, liver cirrhosis, and diabetes greatly increase the risk of death from lower respiratory infections. Trachea, bronchus and lung cancers are almost fully attributable to cigarette smoking (Peto et al., 1992), particularly in populations with a history of heavy smoking like the UK (Ezzati et al., 2003). Risk factors for heart diseases (inflammatory heart disease, ischemic heart disease, hypertensive heart disease, and cerebrovascular diseases) are cigarette smoking and conditions linked with obesity (high blood pressure, bad cholesterol, and atherosclerosis). Smoking, alcohol use, and obesity are established risk factors for colon and rectum cancers, pancreatic cancer, liver cancer, prostrate

cancer, and diabetes. Thus the first factor for males can easily be interpreted as a "Tobacco-use and Diet" factor, being heavily loaded with causes of death associated with the behavioral risks of tobacco-use, especially cigarette smoking, and malnutrition.

Figure 1.3 displays the map of factor scores for the male first factor. Metropolitan, historically industrial areas score the highest on this factor. These metropolitan regions are densely populated and were significant economic centers during the industrial age, particularly in manufacturing and mining. The English area of Birmingham and the Black Country has the highest factor score. Birmingham is the second most populous metropolitan area (after London) and exploded in size during the industrial revolution. The Black Country, immediately west of Birmingham city, is so-called for the area's famed air pollution due to the historic dominance of coal, steel, and iron production. The English areas Merseyside, Greater Manchester, South Yorkshire and Bassetlaw, West Yorkshire, and North Yorkshire and Humber have the highest factor scores after Birmingham and the Black Country. These areas comprised the major industrial belt of Northern England and encompass the cities of Liverpool, Manchester, Leeds, Bradford, Sheffield, and Hull. These cities share a similar history of booming from small towns to metropolitan hubs during the industrial age, and the English textile and coal industries were particularly concentrated in these areas. The Scottish region of Greater Glasgow and Clyde scores the next highest. Greater Glasgow and Clyde is the most populous built-up area in Scotland and was one of the preeminent centers of manufacturing, especially shipbuilding, during Great Britain's industrial revolution. West Central Scotland, particularly Glasgow, has been well documented for its historically poor health performance (Taulbut et al., 2013) in the post-industrial era. Looking broadly at top 20 areas scoring highest on the first male factor, 10 of Scotland's 14 health regions rank among the top 20. The majority of the Scottish economy is industrial and concentrated in central, mainland Scotland. The only Scottish regions scoring low on the first factor are the sparsely populated, remote northern and island areas: Western Isles;

Highland; Orkney; and Shetland. The northern, coastal regions historically relied heavily on fishing and agriculture, though the discovery of North Sea oil in the late 20th century has recently spurred the petroleum industry in the northern island regions. The majority of Scottish areas score high on the first male factor, countered only by select English regions with a historically heavy mark of the industrial era. The first male factor maps heavily on densely populated, formerly industrial areas, with a comparatively larger impact on Scotland than England relative to the number of geographies considered per country.

For males, the second factor is heavily loaded with causes of death related to injuries, drug-use, and alcohol-use (Table 1.3). The first four causes of death loading highest on the second male factor are injury mortality: other unintentional injuries; road traffic accidents; self-inflicted injuries; and falls. Of the remaining causes of death loading onto this factor, alcohol-use disorders and liver cirrhosis are almost fully attributable to alcohol abuse; and esophageal cancer is linked to alcohol-use and smoking. Endocrine disorders are largely driven by malnutrition, primarily over-weight and obesity, with a potential genetic link. Prompted by Scotland's historic homogeneity and low migration, genetic factors have been investigated to explain Scottish's poor health performance relative to England and Wales, though with no empirical support (McCartney, Collins, et al., 2012). The second male factor can be interpreted as the "Rurality and Substance Abuse" factor due to the heavy loading of injury mortality and substance abuse.

Figure 1.4 displays the map of factor scores for the second male factor. The North-South divide for this factor is alarmingly clear, with all 14 of the Scottish areas ranking among the top 15 areas scoring highest on this second male factor. Substance abuse can lead to injury mortality due to the severity of abuse and geographical location. The time to event following substance abuse can be short compared to the cumulative life-course effects of smoking-attributable mortality, for example. Harmful alcohol use or illicit drug use could cause younger age 'premature' mortality by triggering accidental injury; here evidenced by other unintentional

injuries, road traffic accidents, and falls. In other words, the more severe the substance abuse, the higher the mortality cost. Scotland has been documented to have high alcohol-related (Leon and McCambridge, 2006), drug-related (Bloor et al., 2008), and suicide (McCartney, Shipley, et al., 2012) mortality relative to the UK. Scottish people consume higher mean units of alcohol relative to the English population (Bromley and Shelton, 2010). Alcohol-related mortality is known to be highest in Glasgow (Emslie and Mitchell, 2009) within Scotland, in line with the area's elevated alcohol consumption levels after controlling for socioeconomic differences. The likelihood of mortality linked to harmful substance abuse behavior is associated with rurality due to both risk exposure and physical isolation. For road traffic accidents, for example, risk exposure signifies longer distances travelled on the road and a reduced interaction with other drivers or law enforcement. Both scenarios could lead to a less stringent adherence to the law in terms of blood alcohol levels, seat belt use, and driving behavior. Physical isolation also reduces access to emergency trauma centers or, conversely, preventative and treatment services. The mental health costs of rurality could further operate through social isolation. Suicide in Scotland is concentrated in rural areas, particularly the remotest rural areas as opposed to accessible rural areas (Levin and Leyland, 2005). Though the injury-related mortality contributes a small proportion of overall mortality, the geographic pattern in death rates due to injury-related mortality is striking. Factor scores positively correlate with degree of rurality in the expected direction (Appendix Figure A1.1.A). For the second male factor, rurality combined with higher substance abuse mortality drives geographic differences in mortality and reinforces the observed North-South health inequality in Great Britain.

Overall, this two-factor solution again accounts for 89% of the total variance in cause-specific mortality rates for males (Factor 1: 53%; Factor 2: 36%). To assess these factors' relationship with all-cause mortality, standardized all-cause mortality rates were calculated for each health area. The first and second male factors strongly correlate with all-cause mortality

rates (Factor 1 correlation: 0.81; Factor 2 correlation: 0.57). Age-specific death rates for all-cause mortality also correlate with the male factors in the expected pattern. Figure 1.5 graphs the correlation coefficients between age-specific all-cause mortality rates and factor scores for the male sample. Correlations between factor 1 and all-cause age-specific morality rates increase with age, with the inverse relationship seen for factor 2. These findings align with the older age mortality consequences of smoking and obesity and the premature mortality pattern associated with substance abuse and injuries.

Female factor analysis

For the female sample, the first three eigenvalues were 18.60, 1.07, and 0.62. The Kaiser criterion would suggest retaining only the first two factors, and the scree plot again confirms this choice. Following varimax rotation, the two-factor solution accounted for 92% of the total variance (Factor 1: 55%; Factor 2: 37%). Table 1.4 shows the rotated female factor loadings. Consistent with presentation for males, only factor loadings higher than 0.70 (<0.001) are presented in the table.

Similar to the male sample, the first female factor is loaded with causes of deaths strongly associated with primarily cigarette smoking and secondarily obesity. Lower respiratory infections and COPD have the highest correlations with the first female factor. As stated above, the leading cause of COPD is cigarette smoking; and lower respiratory infections are associated with a history of chronic lung conditions, especially COPD, and cigarette smoking. Cigarette smoking promotes chronic inflammation, which increases the likelihood of cancer and atherosclerosis (Wang et al., 2007; Walser et al., 2008).These downstream conditions collectively promote immunosuppression within an individual, increasing the risk of local chemical, carcinogenic effects. Smoking has been causally linked to stomach cancer (Chao et al., 2002), breast cancer (Reynolds, 2013), cervical cancer (Fonseca-Moutinho, 2011), pancreatic cancer (Lynch et al., 2009), and colon and rectal cancer (Slattery et al., 2004) in additional to the more obvious

connection with esophageal and lung cancers (Surgeon General's Advisory Committee on Smoking and Health, 1964). The cardiovascular diseases ischemic heart disease, hypertensive heart disease, and cerebrovascular diseases also loaded heavily on the first female factor. As stated earlier, cigarette smoking and conditions linked with malnutrition (high blood pressure, bad cholesterol, and atherosclerosis) increase the mortality risk for these cardiovascular diseases, with cigarette smoking also independently promoting atherosclerosis. A unique cause loading onto the first female factor is Alzheimer's and other dementias. While the risks for onset of Alzheimer's are unclear, obesity, lack of exercise, and social isolation increase the risk of mortality from Alzheimer's and other dementias. Overall, the first female factor can also be interpreted as a "Tobacco-use and Diet" factor. However due to the causes of death which loaded onto the first female factor, the impact of smoking, as opposed to diet, seems relatively higher for women compared to men. This result is unsurprising considering the greater impact of smoking-attributable mortality on recent female mortality differences, particularly across UK geographies (Kelly and Preston, 2016), due to the lagged uptake of smoking behavior among the female population relative to men.

Figure 1.6 displays the map of factor scores for the first female factor. Though the overall geographic distribution of the "Tobacco-use and Diet" factor is similar for the female and male samples, several Scottish areas moved up in the rankings for the first female factor. As with for the male factor, the English area of Birmingham and the Black Country has the highest factor score followed by the English area Merseyside. However, the Scottish area Greater Glasgow and Clyde scores the third highest. Furthermore, the Scottish areas of Lanarkshire, Forth Valley, and Lothian have higher factor scores for the female first factor than in the male first factor. Overall, 4 of Scotland's 14 health geographies rank among the top ten areas scoring on the first female factor. These areas are predominantly the most populous, and densely populated, health geographies in Scotland; and these are Scottish areas with the most significant industrial pasts.

Forth Valley encompasses Falkirk, the Scottish center of steel and iron production during the industrial age. Lothian contains the city of Edinburgh, Scotland's capital, and the center of printing, brewing, distilling, and engineering industries. Lanarkshire houses Stirling, the former Scottish capital which functioned initially as an inland port and trade center and later as a residential area for people working in Glasgow. In Scotland, the densely populated, industrial areas score higher for the female factor than the male. The relative rankings of the health geographies are skewed by the concentrated disadvantage in these Scottish areas. As reflected in Figure 1.5, factors scores are condensed along the Central industrial belt in Scotland, and spread more evenly in England. These results imply that the mortality consequences of cigarette smoking are peaking in formerly industrial, metropolitan areas for women, especially Scottish women.

For females, the second factor is loaded primarily with causes of death related to rurality, injuries, and substance abuse (Table 1.4). The first two causes of death loading highest on the second female factor are road traffic accidents and self-inflicted injuries. As discussed in the context of the second male factor, road traffic accidents and self-inflicted injuries are strongly linked to the behavioral risks of substance abuse, especially harmful alcohol consumption, and the degree of rurality. The only remaining cause of death with a factor loading above 0.70 (<0.001) is lymphomas and multiple myeloma. This group of causes is composed of malignant neoplasms of lymphoid, hematopoietic and related tissues, excluding leukemia. These causes, predominantly non-Hodgkin's lymphoma and multiple myeloma, account for a small proportion of the overall disease burden for either males or females. The high factor loading for lymphomas and multiple myeloma could reasonably be attributed to rurality as rural populations may have later detection following onset, fewer resources for medical care, and lower adherence to treatment due to physical barriers. Due to the overwhelming influence of injury mortality and substance abuse, the second female factor can also be labeled a "Rurality and Substance Abuse" factor.

Figure 1.7 maps factor scores for the second female factor. The picture is similar to the score map for the second male factor, with the North-South divide disturbingly well-defined. All of the Scottish areas again rank among highest scoring areas for this second female factor, and the northern, sparsely populated Scottish areas lead the relative rankings. An interesting difference between the male and female score maps is the reduced ranking of Greater Glasgow and Clyde, which allowed the English regions of North Yorkshire and Humber and Birmingham and the Black Country to jump in the relative standings. These three regions are similar in terms of their industrial histories and burden of substance-abuse mortality among the urban populations. The geographic differences between the male and female score maps align with the greater importance of road traffic accident mortality in structuring the second female factor. Road traffic accident mortality is associated with elevated substance abuse, especially alcohol-use, and rurality. While suicide is known to be higher in Scotland than England, this mortality inequality is larger for males. Thus for females, the areas most affected by road traffic mortality skew the relative rankings and subsequent geographic patterning of the second factor. Factor scores positively correlate with degree of rurality in the expected direction (Appendix Figure A1.1.B).

The female two-factor solution again accounts for 92% of the total variance in cause-specific mortality rates (Factor 1: 55%; Factor 2: 37%). As for the male sample, female standardized all-cause mortality rates were calculated for each health area; and the first and second factors strongly associate with all-cause mortality rates (Factor 1 correlation: 0.77; Factor 2 correlation: 0.64). Age-specific death rates for all-cause mortality also pattern with the female factors as expected. Figure 1.8 graphs the correlation coefficients between age-specific all-cause mortality rates and factor scores for the female sample. Similar to the male sample, correlations between factor 1 and all-cause age-specific morality rates increase with age, and the inverse relationship is found for factor 2. The gap between factor 1 and factor 2 correlations with age-specific all-cause death rates is less pronounced at older ages for the female sample relative to the

male sample. This slight different is likely for two reasons. First road traffic mortality rather than other injury mortality (self-inflicted injuries and other unintentional injuries) influences the second female factor to a greater extent than for the second male factor. Secondly, the second female factor is loaded with the causes lymphomas and multiple myeloma, which affect older ages.

Discussion

This study indirectly estimated the effect of health behaviors on geographic variation of preventable mortality using exploratory factor analysis on cause-specific mortality information. The majority of observed geographic variation in preventable mortality is driven by causes of death linked to health behaviors. Similar two-factor models explain the bulk of mortality variation for both sexes (Males: 89%; Females: 92%), and these factors are structured by causes of death linked to tobacco-use and diet (Factor 1) and rurality and substance abuse (Factor 2). The geographic patterning of each factor reinforces a North-South disparity in behavioral-driven preventable mortality for both sexes. For women, especially Scottish women, the consequences of cigarette smoking on preventable mortality are peaking in formerly industrial, metropolitan areas. Accordingly for the first female factor compared to the first male factor, the relative rankings of geographies reflect this concentrated disadvantage. The North-South disparity is particularly bleak for both the male and female second factor. Injury (unintentional and suicide) and substance-abuse mortality is unambiguously higher in Scottish areas relative to English areas.

These results align with previous studies of health disparities between England and Scotland. The prevalence of cardio-metabolic diseases, obesity, harmful alcohol consumption, and cigarette smoking is higher in Scotland than England (Bromley and Shelton, 2010). Cardiovascular disease and cancer mortality is higher in urban populations relative to rural populations in Great Britain, though the overall burden is concentrated in Scotland (Shaw et al., 2008). The impact of smoking-attributable mortality is highest in Scotland (Kelly and Preston,

2016), especially among Scottish women. Previous research has found higher mortality in Scotland due to causes of death associated with harmful alcohol consumption are (Leon and McCambridge, 2006) and illicit drug use (Bloor et al., 2008).

The prevalence of mental and behavioral disorders is higher in Scotland than England (Shaw et al., 2008; Smith et al., 2010) with suicide and substance abuse mortality particularly unequal. Suicide and substance abuse mortality are highly correlated. A comprehensive review of 22 cohort studies found that alcohol and drug use disorders strongly associate with suicide mortality for both men and women (Wilcox et al., 2004). Thus while suicide and substance abuse death rates heavily influence geographic variation in preventable mortality in Great Britain, underlying alcohol and drug use may shape both these causes of death.

The primary behavioral drivers of preventable mortality in Great Britain are tobacco-use, diet, and substance abuse. Evidence from the UK shows that all three of these behaviors follow a deprivation gradient, with more deprived individuals more likely to engage in poor behaviors (Bromley et al., 2003). However, socioeconomic differences cannot explain the higher mortality driven by these behaviors. Even after accounting for socioeconomic status, excess mortality from smoking-attributable, drug-attributable, and alcohol-attributable causes is higher in Scottish areas than English areas (Walsh et al., 2010).

No policy differences can explain the heavier tobacco use and substance abuse in Scotland relative to the rest of Great Britain. Scotland was the first country in the UK to ban smoking in enclosed public spaces and workplaces, and a Scottish regulatory relaxation on liquor availability has not caused increased alcohol-related morbidity and mortality relative to England and Wales (Duffy and Plant, 1986). Some of this inequality may be explained by the type of substance used, not merely the amount. The large impact of alcohol-use on mortality differences, for example, may be due to the type of alcohol consumed rather than solely consumption patterns. Though available evidence does find slightly higher mean alcohol consumption in Scotland

relative to England (Bromley and Shelton, 2010), previous research in English-speaking populations also suggests that the mortality consequences from spirits (e.g. scotch and whiskey) are more potent than from beer (Kerr et al., 2000).

These results suggest much of the disparity in preventable mortality may be due to Scotland's rurality compared to England. A recent paper reviewed 17 leading hypotheses suggested to explain Scotland's high mortality relative to the rest of Western Europe (McCartney, Collins, et al., 2012), and rurality was interestingly not identified as a hypothesis. The likelihood of mortality linked to harmful substance abuse can be strongly associated with rurality due to both risk exposure and physical isolation. As discussed in terms of road traffic accidents, longer distances travelled on the road and a reduced interaction with other drivers or law enforcement may lead to a reckless driving behavior and increased mortality outcomes. Previous studies have found significantly higher mortality risk following motor vehicle accidences in rural areas compared to urban areas (Kmet et al., 2003; Myers et al., 2013). Physical isolation also reduces access to emergency care, preventative care, and treatment services. For causes of death associated with both injuries and chronic conditions, rural populations may thus have fewer resources for medical care, fewer interactions with health professionals, and lower adherence to treatment due to physical barriers. The mental health costs of rurality could further operate through social isolation. Suicide in Scotland is concentrated in rural areas, particularly the remotest rural areas as opposed to accessible rural areas (Levin and Leyland, 2005).

Finally, reduced migration and residential mobility may also explain observed geographic variation in preventable mortality in Great Britain. In Scotland, there has been an increase in mortality inequality across the socioeconomic gradient over time, with the increase steepest in rural areas for both sexes (Levin and Leyland, 2006). In terms of absolute inequality, Scottish female health inequalities were greater in remote rural areas than urban areas. This evidence supports this study's findings regarding the second factor, particular for females. A similar study

in England also found an increase in mortality inequality across the socioeconomic gradient over time, but concluded that the majority of this increase was due to underlying migration for both sexes (Norman et al., 2005). Scotland may be less mobile due to its higher degree of rurality and has lower immigration than England. Collectively, reduced migration and residential mobility may help explain the variation in preventable mortality between Scotland and England. Understanding the behavioral drivers of preventable mortality is vital to quantifying health inequality, identifying potential areas for policy interventions, and regulating national population health.

References

Bartlett, Maurice S. "Methods of estimating mental factors." 1938. *Nature* 141 (3570):609-610.

Bloor, M., M. Gannon, G. Hay, G. Jackson, A. H. Leyland, and N. McKeganey. 2008. "Contribution of problem drug users' deaths to excess mortality in Scotland: secondary analysis of cohort study." *British Medical Journal* 337:a478.

Boyle, P. 2004. "Population geography: migration and inequalities in mortality and morbidity." *Progress in Human Geography* 28 (6):767.

Bromley, C, and N. Shelton. 2010. "The Scottish Health Survey: Topic Report UK Comparisons." Accessed November 2016 at: www.gov.scot/Publications/2010/08/31093025/0.

Bromley, C., K. Sproston, and N. Shelton. 2003. "The Scottish Health Survey 2003. Volume 2. Adults." Accessed November 2016 at: www.scotland.gov.uk.

Brownson, RC., TE. Novotny, and MC. Perry. 1993. "Cigarette smoking and adult leukemia: a meta-analysis." *Archives of Internal Medicine* 153 (4):469-475.

Chao, A., MJ. Thun, SJ. Henley, EJ. Jacobs, ML. McCullough, and EE. Calle. 2002. "Cigarette smoking, use of other tobacco products and stomach cancer mortality in US adults: The Cancer Prevention Study II." *International Journal of Cancer* 101 (4):380-389.

Classifications and Harmonization Unit. 2013. "National Statistics Country Classification (NSCC) Explanatory Notes and Coding Guidelines." Accessed November 2016 at: www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/national-statistics-country-classification/index.html.

Costello, AB., and JW. Osborne. 2005. "Best Practices in Exploratory Factor Analysis: Four Recommendations for Getting the Most From Your Analysis." *Practical Assessment Research and Evaluation* 10 (7):1-9.

Craig, J. 1984. "Averaging Population Density." *Demography* 21(3):405-412.

Crimmins, EM., SH. Preston, and B. Cohen (eds.). 2010. "International Differences in Mortality at Older Ages: Dimensions and Sources." In Panel on Understanding Divergent Trends in Longevity in High-Income Countries, National Research Council. Washington (DC): National Academies Press (US).

Danaei, G., EB. Rimm, S. Oza, SC. Kulkarni, CJL. Murray, and M. Ezzati. 2010. "The promise of prevention: the effects of four preventable risk factors on national life expectancy and life expectancy disparities by race and county in the United States." *PLoS medicine* 7 (3):342.

Danaei, G., S. Vander Hoorn, AD. Lopez, CJL. Murray, M. Ezzati, et al. 2005. "Causes of cancer in the world: comparative risk assessment of nine behavioral and environmental risk factors." *Lancet* 366 (9499):1784-1793.

Di Cesare, M., Y. Khang, P. Asaria, T. Blakely, MJ. Cowan, F. Farzadfar, R. Guerrero, N. Ikeda, C. Kyobutungi, and KP. Msyamboza. 2013. "Inequalities in non-communicable diseases and effective responses." *Lancet* 381 (9866):585-597.

Dilnot, A. 2012. "Review of the accessibility of official statistics on health." Monitoring Reviews: UK Statistics Authority. Accessed November 2016 at: www.statisticsauthority.gov.uk/publication/the-accessibility-of-official-statistics-on-health/.

DiStefano, C., M. Zhu, and D. Mindrila. 2009. "Understanding and using factor scores: Considerations for the applied researcher." *Practical Assessment, Research & Evaluation* 14 (20):1-11.

Duffy, JC., and MA. Plant. 1986. "Scotland's liquor licensing changes: an assessment." *British Medical Journal (Clinical Research Education)* 292 (6512):36-39.

Emslie, C., and R. Mitchell. 2009. "Are there gender differences in the geography of alcohol-related mortality in Scotland? An ecological study." *BMC Public Health* 9:58.

Ezzati, M., AD. Lopez, A. Rodgers, S. Vander Hoorn, and CJL. Murray. 2002. "Selected major risk factors and global and regional burden of disease." *Lancet* 360 (9343):1347-60.

Ezzati, M., H. Martin, S. Skjold, S. Vander Hoorn, and CJL. Murray. 2006. "Trends in national and state-level obesity in the USA after correction for self-report bias: analysis of health surveys." *Journal of the Royal Society of Medicine* 99 (5):250-257.

Ezzati, M., S. Vander Hoorn, A. Rodgers, AD. Lopez, DC. Mathers, and CJL. Murray. 2003. "Estimates of global and regional potential health gains from reducing multiple major risk factors." *Lancet* 362 (9380):271-280.

Finucane, MM., GA. Stevens, MJ. Cowan, G. Danaei, JK. Lin, CJ. Paciorek, GM. Singh, HR. Gutierrez, Y. Lu, and AN. Bahalim. 2011. "National, regional, and global trends in body-mass index since 1980: systematic analysis of health examination surveys and epidemiological studies with 960 country-years and 9.1 million participants." *Lancet* 377 (9765):557-567.

Fonseca-Moutinho, JA. 2011. "Smoking and cervical cancer." *ISRN Obstetrics and Gynecology* 847684.

Graham, H. 2009. "Understanding health inequalities." McGraw-Hill Education (UK).

Greenland, S., and JM. Robins. 1988. "Conceptual problems in the definition and interpretation of attributable fractions." *American Journal of Epidemiology* 128 (6):1185-1197.

Hofstetter, H., E. Dusseldorp, PV. Empelen, and T. Paulussen. 2014. "A primer on the use of cluster analysis or factor analysis to assess co-occurrence of risk behaviors." *Preventive Medicine* 67:141-146.

Human Mortality Database. 2014. "Human Mortality Database." Berkeley (USA) and Rostock (Germany): University of California and Max Planck Institute for Demographic Research. Accessed November 2014 at: www.mortality.org.

Hutchison, A., M. Tobias, J. Glover, C. Wright, D. Hetzel, and E. Fisher. 2006. “Australian and New Zealand atlas of avoidable mortality: Public Health Information Development Unit.”

Ikeda, N., M. Inoue, H. Iso, S. Ikeda, T. Satoh, M. Noda, T. Mizoue, H. Imano, E. Saito, and K. Katanoda. 2012. "Adult mortality attributable to preventable risk factors for non-communicable diseases and injuries in Japan: a comparative risk assessment." *PLoS medicine* 9 (1):64.

Jolliffe, IT. 1972. "Discarding variables in a principal component analysis. I: Artificial data." *Applied Statistics*: 160-173.

Jonker, MF., B. Donkers, B. Chaix, F. van Lenthe, A Burdorf, and JP. Mackenbach. 2015. "Estimating the Impact of Health-related Behaviors on Geographic Variation in Cardiovascular Mortality." *Epidemiology* 26 (6):888-897.

Kaiser, HF. 1960. "The application of electronic computers to factor analysis." *Educational and Psychological Measurement* 20:141-151.

Kelly, LA., and SH. Preston. 2016. "The contribution of a history of heavy smoking to Scotland's mortality disadvantage." *Population Studies* 70 (1):59-71.

Kerr, WC., K. Middleton Fillmore, and P. Marvy. 2000. "Beverage-specific alcohol consumption and cirrhosis mortality in a group of English-speaking beer-drinking countries." *Addiction* 95 (3):339-346.

Kmet, L., P. Brasher, and C. Macarthur. 2003. "A small area study of motor vehicle crash fatalities in Alberta, Canada." *Accident, Analysis and Prevention* 35 (2):177-182.

Leon, DA., and J. McCambridge. 2006. "Liver cirrhosis mortality rates in Britain from 1950 to 2002: an analysis of routine data." *Lancet* 367 (9504):52-56.

Levin, KA., and AH. Leyland. 2005. "Urban/rural inequalities in suicide in Scotland, 1981–1999." *Social Science & Medicine* 60 (12):2877-2890.

Levin, KA., and AH. Leyland. 2006. "A comparison of health inequalities in urban and rural Scotland." *Social Science & Medicine* 62 (6):1457-1464.

Levine, B. 2007. "Peer Reviewed: What Does the Population Attributable Fraction Mean?" *Preventing Chronic Disease* 4 (1).

Lim, SS., E. Carnahan, EC. Nelson, CW. Gillespie, AH. Mokdad, CJL. Murray, and ES. Fisher. 2015. "Validation of a new predictive risk model: measuring the impact of the major modifiable risks of death for patients and populations." *Population Health Metrics* 13 (1):27.

Lynch, SM., A. Vrieling, JH. Lubin, P. Kraft, JB. Mendelsohn, P. Hartge, F. Canzian, E. Steplowski, AA. Arslan, and M. Gross. 2009. "Cigarette smoking and pancreatic cancer: a pooled analysis from the pancreatic cancer cohort consortium." *American Journal of Epidemiology* 170 (4):403-413.

McCartney, G., C. Collins, D. Walsh, and GD. Batty. 2012. "Why the Scots die younger: synthesizing the evidence." *Public Health* 126 (6):459-70.

McCartney, G., M. Shipley, C. Hart, G. Davey-Smith, M. Kivimaki, D. Walsh, GC. Watt, and GD. Batty. 2012. "Why do males in Scotland die younger than those in England? Evidence from three prospective cohort studies." *PLoS One* 7 (7):e38860.

Myers, SR., CC. Branas, BC. French, ML. Nance, MJ. Kallan, DJ. Wiebe, and BG. Carr. 2013. "Safety in numbers: are major cities the safest places in the United States?" *Annals of Emergency Medicine* 62 (4):408-418. e3.

Naghavi, M., S. Makela, K. Foreman, J. O'Brien, F. Pourmalek, and R. Lozano. 2010. "Algorithms for enhancing public health utility of national causes-of-death data." *Population Health Metrics* 8:9.

National Research Council, Committee on Population Division of Behavioral and Social Sciences and Education, Board on Health Care Services, and Institute of Medicine. 2015. "The National Academies Collection: Reports funded by National Institutes of Health." In Measuring the Risks and Causes of Premature Death: Summary of Workshops. Washington (DC): National Academies Press (US).

NIH MedlinePlus. "A.D.A.M. Medical Encyclopedia." Available from: www.nlm.nih.gov/medlineplus/encyclopedia.html: National Library of Medicine (US).

Norman, P., P. Boyle, and P. Rees. 2005. "Selective migration, health and deprivation: a longitudinal analysis." *Social Science & Medicine* 60 (12):2755-71.

Olatunde, O.. 2013. "The impact of calculating mortality rates using the 2013 European Standard Population on causes of death." Report on Impact of 2013 ESP.

Pace, M., E. Cayotte, L. Agafitei, T. Zupanic, B. Wojtyniak, M. Gissler, G. Lanzieri, M. Glickman, and E. Grande. 2013. "Revision of the European Standard Population: Report of Eurostat's Task Force: 2013 Edition." Publications Office.

Peto, R., AD. Lopez, J. Boreham, M. Thun, and C. Heath, Jr. 1992. "Mortality from tobacco in developed countries: indirect estimation from national vital statistics." *Lancet* 339 (8804):1268-78.

Preston, SH., EM. Crimmins, and JR. Wilmoth. 2010a. "A new method for estimating smoking-attributable mortality in high-income countries." *International Journal of Epidemiology* 39 (2):430-8.

Reynolds, P. 2013. "Smoking and breast cancer." *Journal of Mammary Gland Biology and Neoplasia* 18 (1):15-23.

Scottish Government. 2008. "Equally Well: Report of the Ministerial Task Force on Health Inequalities." Scottish Government.

Shaw, M., B. Thomas, GD. Smith, and D. Dorling. 2008. "The grim reaper's road map: an atlas of mortality in Britain." Policy Press at the University of Bristol (UK).

Shiely, F., K. Hayes, IJ. Perry, and CC. Kelleher. 2013. "Height and weight bias: the influence of time." *PloS one* 8 (1):e54386.

Slattery, ML., W. Samowtiz, K. Ma, M. Murtaugh, C. Sweeney, TR. Levin, and S. Neuhausen. 2004. "CYP1A1, cigarette smoking, and colon and rectal cancer." *American Journal of Epidemiology* 160 (9):842-852.

Smith, MP., D. Sweet, and R. Holley. 2010. "United Kingdom Health Statistics 2010." London: Office for National Statistics.

Surgeon General's Advisory Committee on Smoking and Health. 1964. "Smoking and Health: Report of the Advisory Committee to the Surgeon General of the Public Health Service." Washington, D.C. (USA): US Public Health Service.

Taulbut, M., D. Walsh, S. Parcell, A. Hartmann, G. Poirier, D. Strniskova, G. Daniels, and P. Hanlon. 2013. "What can ecological data tell us about reasons for divergence in health status between West Central Scotland and other regions of post-industrial Europe?" *Public Health* 127 (2):153-63.

Tencza, C., A. Stokes, and SH. Preston. 2014. "Factors responsible for mortality variation in the United States: A latent variable analysis." *Demographic Research* 31 (2):27-70.

Thomson, G. 1939. "The factorial analysis of human ability." *British Journal of Educational Psychology* 9 (2):188-195.

Tyner, JA. 2015. "Population geography II Mortality, premature death, and the ordering of life." *Progress in Human Geography* 39 (3):360-373.

UK Statistics Authority. 2014. "Population Estimates for UK, England and Wales, Scotland and Northern Ireland". Accessed November 2016 at: www.ons.gov.uk/ons/rel/pop-estimate/population-estimates-for-uk--england-and-wales--scotland-and-northern-ireland/mid-2014/index.html.

Walser, T., X. Cui, J. Yanagawa, JM. Lee, E. Heinrich, G. Lee, S. Sharma, and SM. Dubinett. 2008. "Smoking and lung cancer: the role of inflammation." *Proceedings of the American Thoracic Society* 5 (8):811-815.

Walsh, D., N. Bendel, R. Jones, and P. Hanlon. 2010. "It's not 'just deprivation': why do equally deprived UK cities experience different health outcomes?" *Public Health* 124 (9):487-495.

Wang, Z., SJ. Nicholls, ER. Rodriguez, O. Kummu, S. Hörkkö, J. Barnard, WF. Reynolds, EJ. Topol, JA. DiDonato, and SL. Hazen. 2007. "Protein carbamylation links inflammation, smoking, uremia and atherogenesis." *Nature Medicine* 13 (10):1176-1184.

Wheller, L., A. Baker, C. Griffiths, and C. Rooney. 2007. "Trends in avoidable mortality in England and Wales, 1993-2005." *Health Statistics Quarterly* (34):6.

Wilcox, HC., KR. Conner, and ED. Caine. 2004. "Association of alcohol and drug use disorders and completed suicide: an empirical review of cohort studies." *Drug and Alcohol Dependence* 76:S11-S19.

Windfuhr, K., and N. Kapur. 2011. "Suicide and mental illness: a clinical review of 15 years findings from the UK National Confidential Inquiry into Suicide." *British Medical Bulletin* 100 (1):101-121.

Table 1.1. Deaths and population by sex and health geography, 1981-2009.

	Males		Females	
England	**Deaths**	**Population**	**Deaths**	**Population**
Arden, Herefordshire, and Worcestershire	38,224	6,835,873	40,355	7,154,869
Bath, Gloucestershire, Swindon, and Wiltshire	26,136	7,508,507	33,513	7,829,665
Birmingham and the Black Country	60,183	3,260,113	62,720	3,458,910
Bristol, North Somerset, Somerset, and South Gloucestershire	43,933	2,405,835	36,556	2,625,629
Cheshire, Warrington, and Wirral	30,436	5,206,251	42,309	5,475,243
Cumbria, Northumberland, Tyne and Wear	60,201	13,902,201	60,570	14,723,706
Derbyshire and Nottinghamshire	59,560	10,371,273	54,426	10,921,611
Devon, Cornwall, and Isles of Scilly	36,827	6,940,348	52,945	7,309,151
Durham, Darlington, and Tees	36,925	3,349,195	34,497	3,537,010
East Anglia	65,864	11,605,191	55,829	12,196,989
Essex	34,893	8,428,873	51,011	8,921,887
Greater Manchester	74,076	5,395,927	91,762	5,638,019
Hertfordshire and the South Midlands	47,850	11,819,832	51,078	12,265,045
Kent and Medway	38,073	6,933,468	46,145	7,221,364
Lancashire	45,252	10,481,699	45,077	10,984,052
Leicestershire and Lincolnshire	45,501	6,644,127	39,826	6,917,830
London	165,722	18,150,706	172,261	19,092,015
Merseyside	38,293	1,850,641	33,239	1,934,505
North Yorkshire and Humber	42,739	2,600,896	38,046	2,741,701
Shropshire and Staffordshire	37,617	6,709,577	40,715	7,029,902
South Yorkshire and Bassetlaw	37,153	3,340,842	47,343	3,507,670
Surrey and Sussex	71,718	13,069,635	80,817	13,793,911
Thames Valley	44,256	9,109,225	41,523	9,683,040
Wessex	67,309	9,843,063	74,554	10,268,073
West Yorkshire	58,646	5,285,645	55,916	5,523,329
Scotland	**Deaths**	**Population**	**Deaths**	**Population**
Ayrshire and Arran	22,093	1,781,592	24,215	1,949,204
Borders	6,420	506,711	7,353	548,737
Dumfries and Galloway	9,178	713,435	9,666	762,216
Fife	19,180	1,684,803	20,627	1,804,266
Forth Valley	15,117	1,339,496	15,872	1,435,114
Grampian	26,240	2,558,036	27,940	2,644,836
Greater Glasgow and Clyde	72,891	5,605,950	77,379	6,175,078
Highland	17,717	1,464,008	18,417	1,523,962
Lanarkshire	33,703	3,052,942	35,327	3,292,057
Lothian	39,602	3,661,559	44,264	3,973,815
Orkney	1,217	96,518	1,161	99,427
Shetland	1,175	114,976	1,184	112,392
Tayside	23,396	1,885,372	26,280	2,049,572
Western Isles	2,194	143,048	2,083	145,348

SOURCE: Death and population data is taken from the ONS and includes persons registered as English or Scottish and aged 15-74 during the period 1981-2009.

Table 1.2. Highest ranked causes of death by sex, 1981-2009.

Males, Causes of Death	**Females, Causes of Death**
Ischemic heart disease	Ischemic heart disease
Trachea, bronchus and lung cancers	Cerebrovascular disease
Cerebrovascular disease	Breast cancer
Chronic obstructive pulmonary disease	Trachea, bronchus and lung cancers
Colon and rectum cancers	Chronic obstructive pulmonary disease
Lower respiratory infections	Colon and rectum cancers
Prostate cancer	Lower respiratory infections
Stomach cancer	Ovary cancer
Alcohol-use disorders and Liver Cirrhosis	Alcohol-use disorders and Liver Cirrhosis
Self inflicted injuries	Pancreas cancer

SOURCE: Death and population data is taken from the ONS and includes persons registered as English or Scottish and aged 15-74 during the period 1981-2009.
NOTE: For each sex, causes are ranked by age-standardized death rates for the entire study sample (all geographies, all years) for persons aged 15 to 74. The highest 10 rankings are shown for each sex.

Table 1.3. Rotated factor loadings 1981-2009, males.

Cause	Factor 1	Cause	Factor 2
Chronic obstructive pulmonary disease	0.96	Other unintentional injuries	0.94
Stomach cancer	0.95	Road traffic accidents	0.85
Leukemia	0.91	Self inflicted injuries	0.85
Peptic ulcer disease	0.90	Falls	0.80
Bladder cancer	0.89	Alcohol-use disorders and Liver Cirrhosis	0.79
Lower respiratory infections	0.89	Esophagus cancer	0.77
Trachea, bronchus and lung cancers	0.89	Endocrine disorders	0.73
Inflammatory heart disease	0.84		
Ischemic heart disease	0.81		
Colon and rectum cancers	0.80		
Pancreas cancer	0.75		
Liver cancer	0.74		
Hypertensive heart disease	0.74		
Cerebrovascular disease	0.74		
Prostate cancer	0.71		
Diabetes mellitus	0.70		

SOURCE: Death and population data is taken from the ONS and includes males registered as English or Scottish and aged 15-74 during the period 1981-2009.
NOTE: Factor loadings over 0.7 ($p<.001$) are shown.

Table 1.4. Rotated factor loadings 1981-2009, females.

Cause	Factor 1	Cause	Factor 2
Lower respiratory infections	0.95	Road traffic accidents	0.90
Chronic obstructive pulmonary disease	0.90	Self inflicted injuries	0.84
Stomach cancer	0.86	Lymphomas and multiple myeloma	0.73
Cervix uteri cancer	0.84		
Endocrine disorders	0.84		
Trachea, bronchus and lung cancers	0.81		
Ischemic heart disease	0.81		
Esophagus cancer	0.79		
Hypertensive heart disease	0.79		
Alzheimer and other dementias	0.78		
Pancreas cancer	0.76		
Cerebrovascular disease	0.76		
Breast cancer	0.76		
Colon and rectum cancers	0.72		
Leukemia	0.72		

SOURCE: Death and population data is taken from the ONS and includes females registered as English or Scottish and aged 15-74 during the period 1981-2009.
NOTE: Factor loadings over 0.7 ($p<.001$) are shown.

Figure 1.1. Age–sex bar chart by combined groups of cause of death in Great Britain, 1981–2004 (Shaw et al. 2008).

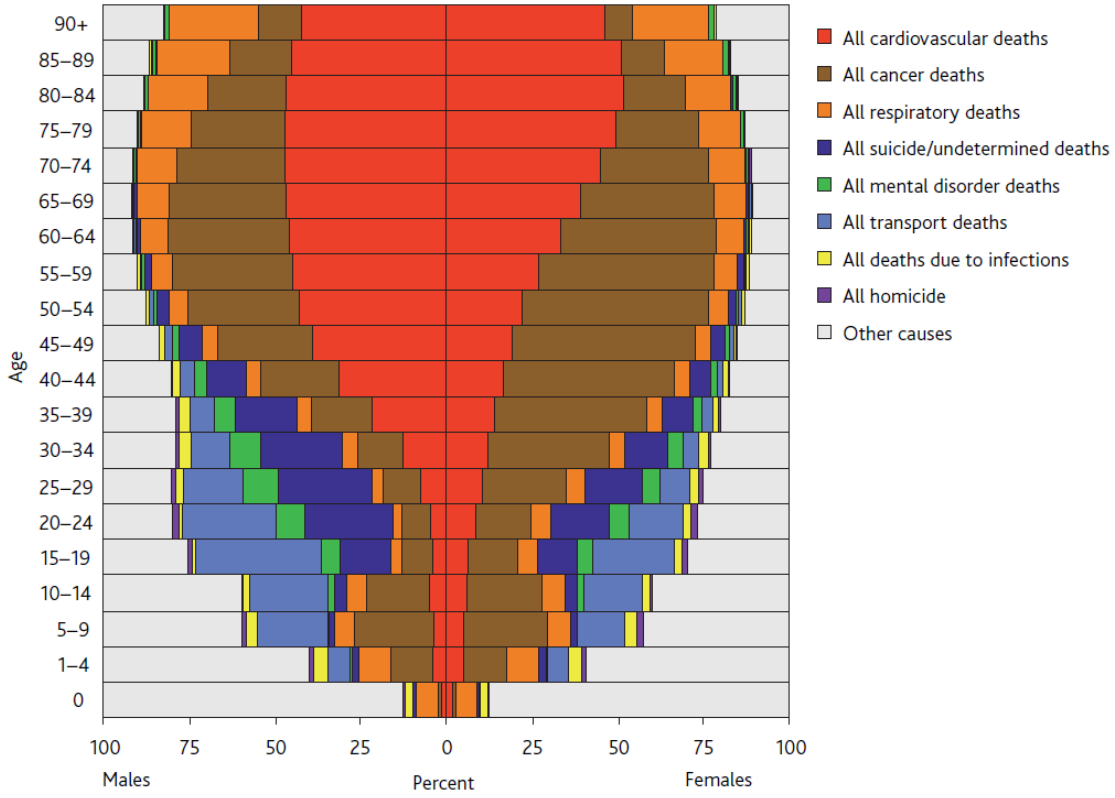


Figure 1.2. Diagram of hypothetical two factor model.

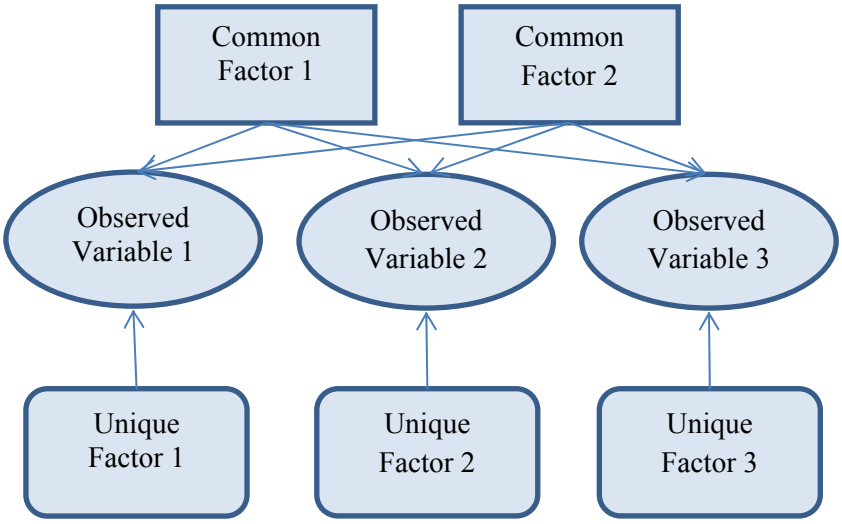
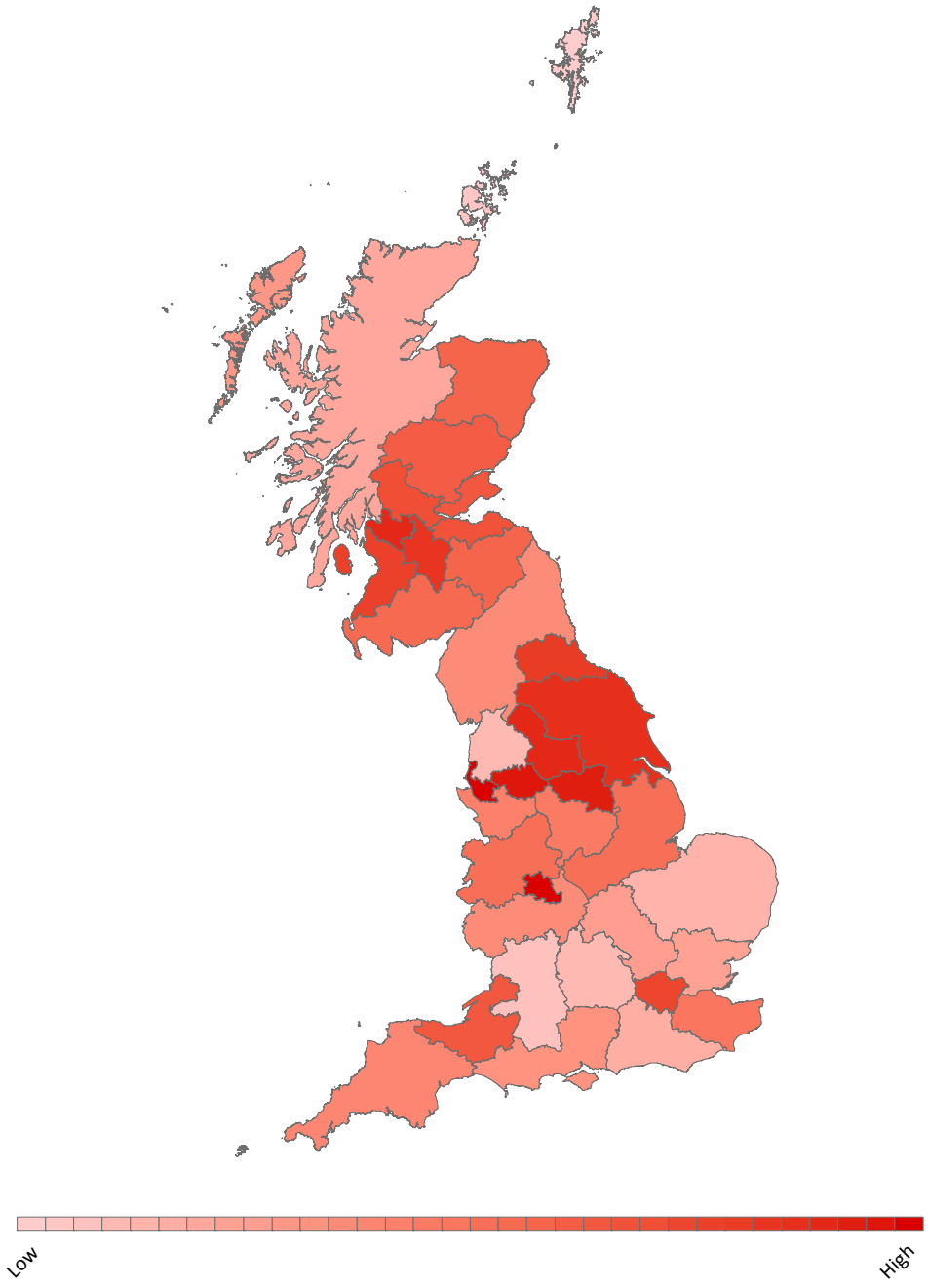
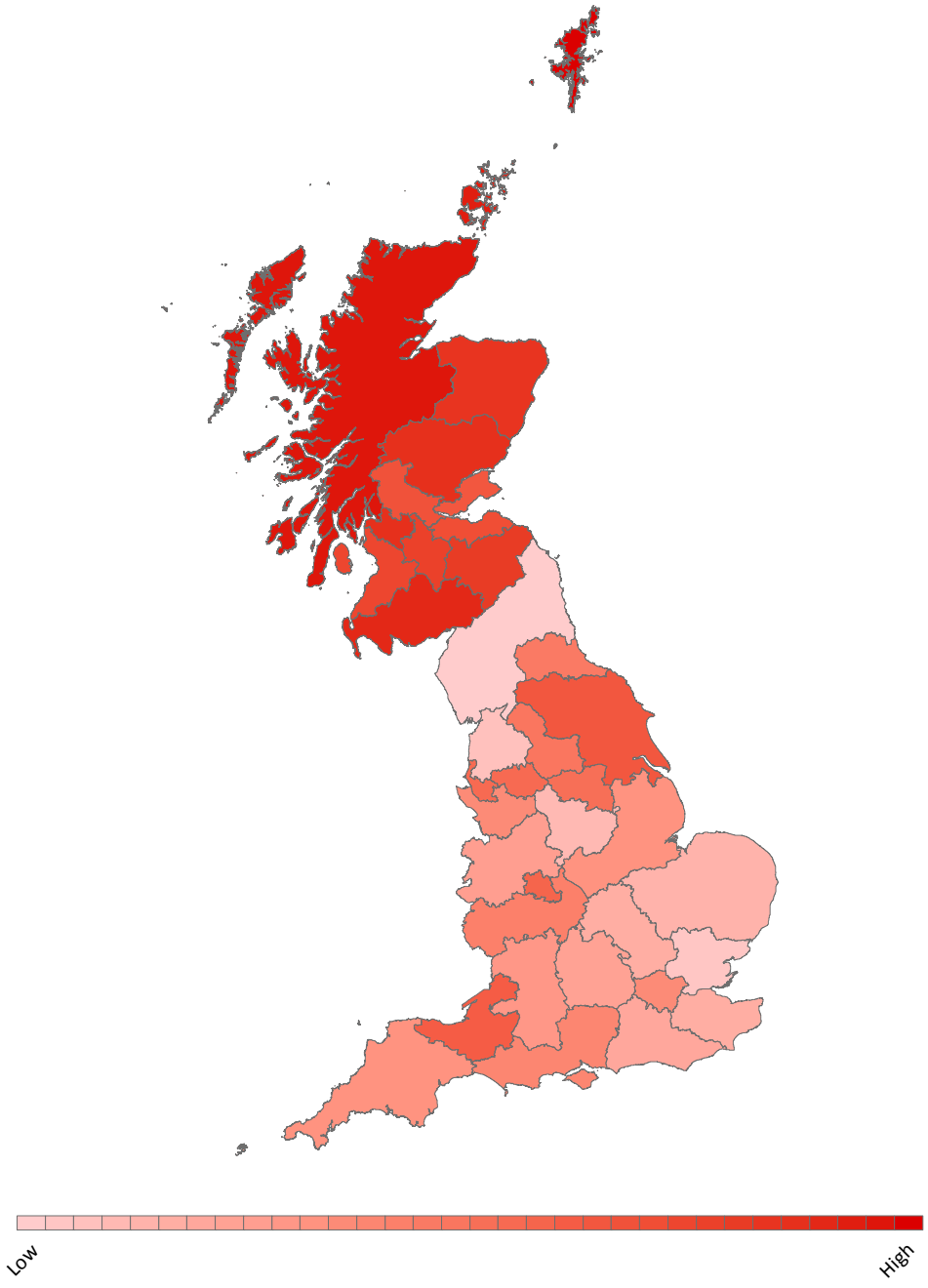


Figure 1.3. Factor 1 (Tobacco-use and Obesity) score map 1981-2009, males.



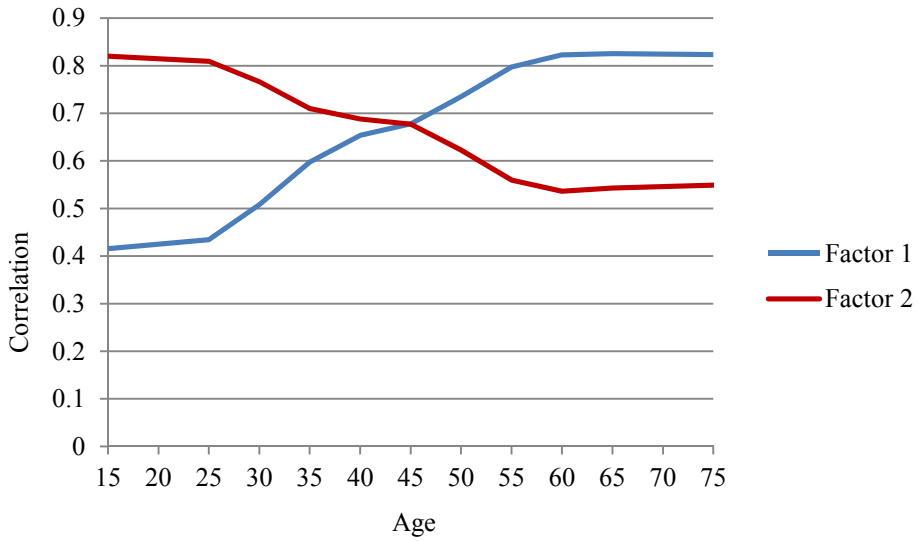
NOTE: Health geographies are shaded according to their relative rank. Darker red areas have higher factor scores, and lighter red areas have lower factor scores.

Figure 1.4. Factor 2 (Rurality and Substance Abuse) score map 1981-2009, males.



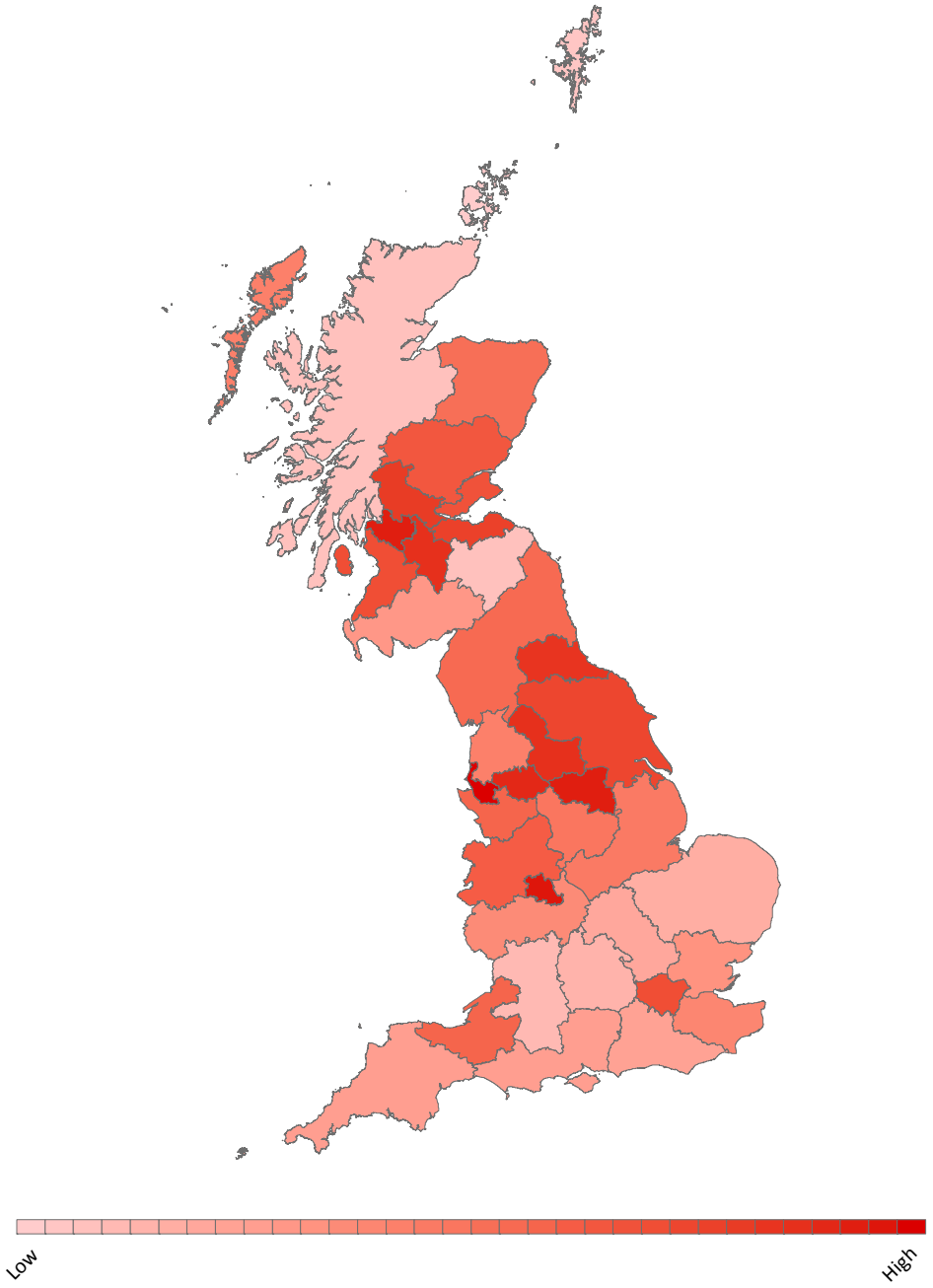
NOTE: Health geographies are shaded according to their relative rank. Darker red areas have higher factor scores, and lighter red areas have lower factor scores.

Figure 1.5. Correlation coefficient between age-specific all-cause mortality rates and factor scores 1981-2009, males.



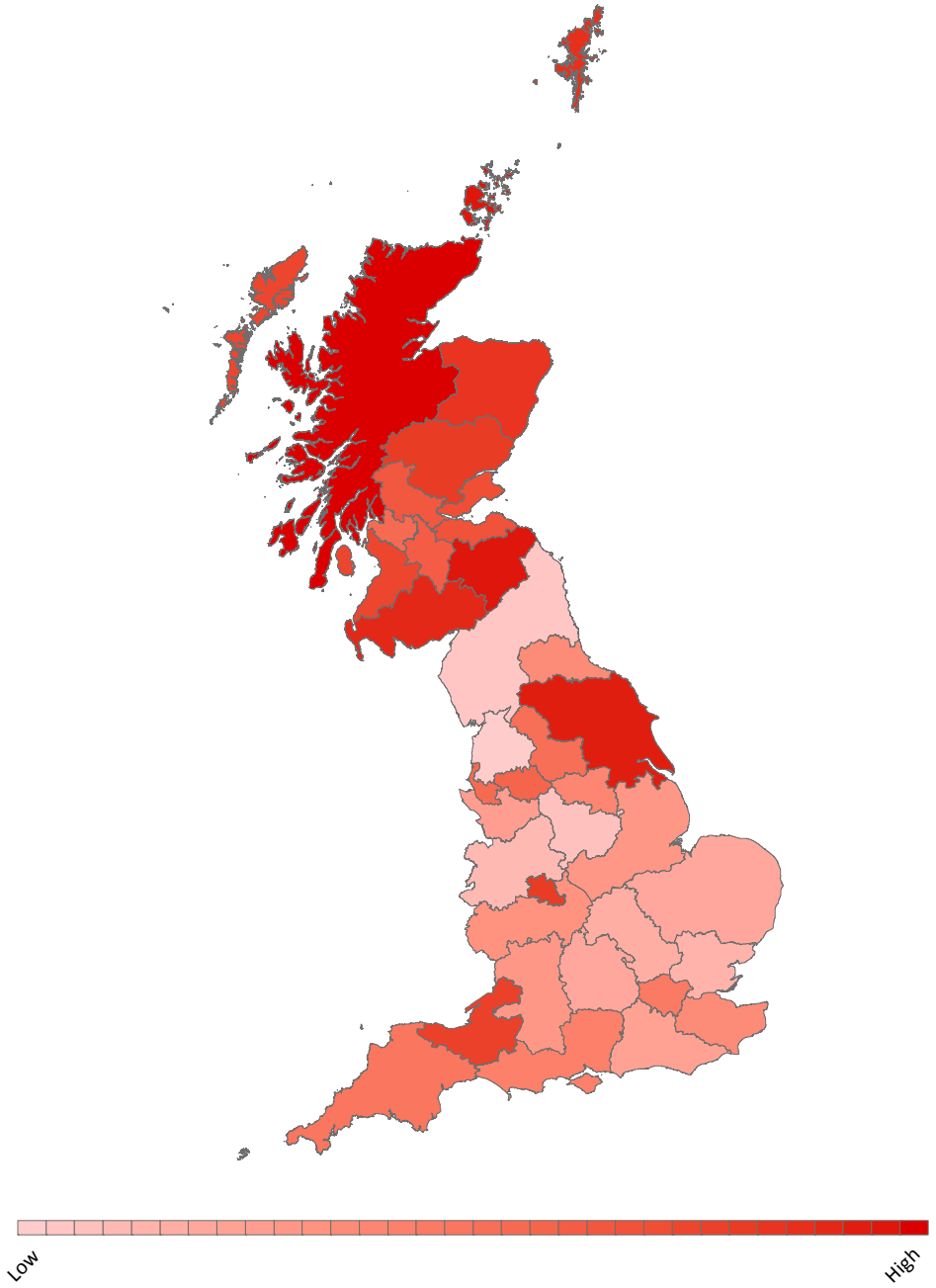
SOURCE: Death and population data is taken from the ONS and includes persons registered as English or Scottish and aged 15-74 during the period 1981-2009.
NOTE: By sex, age-standardized death rates for all-cause mortality are calculated for the entire study sample (all geographies, all years) for persons aged 15 to 74.

Figure 1.6. Factor 1 (Tobacco-use and Obesity) score map 1981-2009, female.



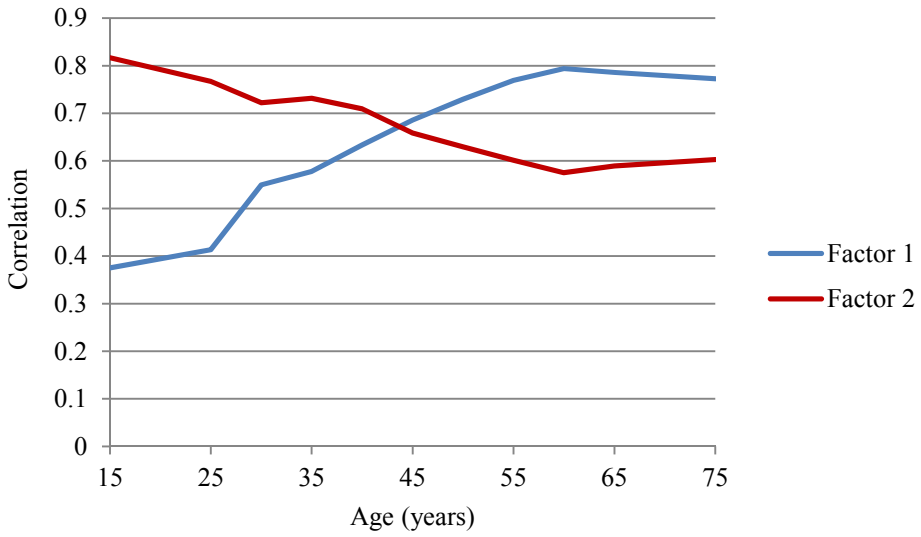
NOTE: Health geographies are shaded according to their relative rank. Darker red areas have higher factor scores, and lighter red areas have lower factor scores.

Figure 1.7. Factor 2 (Rurality and Substance Abuse) score map 1981-2009, females.



NOTE: Health geographies are shaded according to their relative rank. Darker red areas have higher factor scores, and lighter red areas have lower factor scores.

Figure 1.8. Correlation coefficient between age-specific all-cause mortality rates and factor scores 1981-2009, females.



SOURCE: Death and population data is taken from the ONS and includes persons registered as English or Scottish and aged 15-74 during the period 1981-2009.
NOTE: By sex, age-standardized death rates for all-cause mortality are calculated for the entire study sample (all geographies, all years) for persons aged 15 to 74.

CHAPTER 2

The contribution of a history of heavy smoking to Scotland's mortality disadvantage

Note: This chapter was co-authored and published (PMCID: PMC479889) with Samuel H. Preston.

Introduction

Although mortality in Scotland and other developed countries steadily declined over the last century, a widening gap in life expectancy between Scotland and other high-longevity countries emerged around 1950 and accelerated after 1980 (McCartney et al. 2012a; McCartney et al. 2012b). In this paper we document and briefly review explanations of Scotland's poor performance in the improvement of life expectancy compared with other countries. We report the results of applying two methods to estimate the contribution of deaths attributable to a history of heavy smoking to Scotland's shortfall in life expectancy compared with a set of other high-income countries.

Background

Scotland's life expectancy in international perspective

To demonstrate that Scottish mortality falls below international standards for developed countries, we chose two comparison groups with high-quality mortality data. One group consisted of other English-speaking countries: England and Wales, Ireland, the United States, Canada, Australia, and New Zealand. The second group was European: two countries from northern Europe (Sweden and Finland), two countries from central Europe (Austria and Switzerland), two countries from southern Europe (Italy and Portugal), and two countries from western Europe (France and Belgium). A non-western developed country, Japan, was added to the set of countries

when we made the broadest aggregate comparison. We relied primarily on data from the Human Mortality Database, a source of high-quality historical and contemporary data on mortality. Table 2.1. shows the value of life expectancy in 2009 for Scotland and the three sets of comparison countries. For females, Scottish life expectancy at birth in 2009 is 2.82 years below the mean for the other countries, and for males it is 2.24 years lower. The Scottish disadvantage appears in relation to all sets of comparison countries.

The bulk of Scotland's mortality disadvantage occurs above age 50 years. According to the 2009 Scottish life tables (Human Mortality Database 2014), 96.3 per cent of female births and 93.3 per cent of male births reach age 50. Compared to the mean of the 15 countries, the Scottish shortfall of 2.51 years in female life expectancy at age 50 accounts for about 85 per cent [(0.95)2.51/2.82] of the shortfall of 2.82 years in life expectancy at birth (Table 2.1.). Among men, the equivalent figure is 66 per cent. Figure 2.1. shows the trend since 1950 in life expectancy at age 50 in the countries being compared. It shows that Scottish women and men were near the bottom of rankings in 1950 and that they were always at the bottom after 1980. In fact, from 1980 onwards, life expectancy at age 50 was lower in Scotland than in any other developed country (Human Mortality Database 2014).

Next to Scotland, the lowest life expectancy at birth for both females and males in Table 2.1. is that of the United States. The poor standing of the US in life expectancy was the subject of a major National Research Council study that produced 13 background papers (Crimmins et al. 2010) and a long synthesis report (Crimmins et al. 2011). The report concluded that the single leading reason for the relatively low life expectancy of the US at age 50 was its history of very heavy cigarette smoking, and that this accounted for 78 per cent of the US shortfall relative to comparison countries for women and 41 per cent of the shortfall for men (Crimmins et al. 2011). It is not surprising that differences in smoking patterns contribute to international differences in mortality because tobacco use is the primary cause of preventable mortality in the developed

world (Ezzati et al. 2002). The 1964 US Surgeon General's report (Surgeon General's Advisory Committee on Smoking and Health 1964) decisively linked smoking to the development of lung cancer, and subsequent evidence confirmed that smokers were at increased risk of chronic obstructed lung disease, heart disease, stroke, chronic respiratory diseases, and various other cancers and circulatory diseases (Doll et al. 2004; Doll et al. 1994). The 2014 US Surgeon General's report (US Department of Health and Human Services 2014) estimates that 21 million deaths in the United States were attributed to tobacco use between 1964 and 2014, of which 24 per cent were caused by lung cancer.

Explanations of Scotland's exceptionally high mortality

Explanations of Scotland's historic mortality disadvantage have focused not on smoking but on several other causes. A recent paper reviewed hypotheses offered to explain Scotland's mortality disadvantage relative to the rest of Western Europe, distinguishing between the 1950-80 period and the period since 1980 (McCartney et al. 2012a). Seventeen hypotheses were identified in a comprehensive literature search, and each was evaluated using the Bradford-Hill criteria for causality among observational data (Hill 1965).

The review suggests that the most plausible explanation for Scotland's mortality gap within the UK between 1950 and 1980 is relative deprivation. In 1989, Carstairs and Morris developed a much-cited measure of social deprivation in the United Kingdom using four census indicators: overcrowding, unemployment among men, low social class, and not having a car (Carstairs and Morris 1989). Applying this measure to area-level data on wards and post-code sectors from the 1981 UK censuses, the authors accounted for approximately 60 per cent of Scotland's excess age-standardized and sex-standardized mortality relative to England and Wales. The McCartney review argues that the explanatory power of socioeconomic deprivation, using the Carstairs or alternate measures of deprivation, declined after 1980 (Hanlon et al. 2005; Hanlon et al. 2001; Walsh, Taulbut and Hanlon 2010). Hanlon et al. calculated the Carstairs and

Morris area-level deprivation scores using the 1981, 1991, and 2001 censuses in Scotland and England and Wales (Hanlon et al. 2005). Whereas in 1981 the Carstairs and Morris index accounted for 60 per cent of Scotland's excess mortality, the index could account for less than half of Scotland's excess mortality in 1991 and 2001. In fact, relative to England and Wales, Scotland became less deprived during the period, yet simultaneously experienced a widening mortality disparity (Hanlon et al. 2005). Following common practice, the authors refer to this unexplained excess in Scottish mortality as the 'Scottish effect'.

In light of available data, the McCartney review further evaluates, and cautiously discounts hypotheses that suggest that the Scottish mortality gap after 1980 could be explained by migration, health systems, or health-related behavior. The migration hypothesis posits that the country's higher mortality could be explained by the emigration of a higher proportion of healthy individuals from Scotland than from other Western European regions. But the limited literature on the subject shows that Scottish migrants display mortality profiles in their country of destination that are similar to those of non-migrants remaining in Scotland, particularly in death rates from lung cancer and ischemic heart disease (Connolly et al. 2011; Popham et al. 2010).

Though too little evidence is available for a full evaluation of the role of health systems in Scotland's mortality disparity, the McCartney review suggests that differentials in health systems' performance were not responsible. Self-reported use of health services in the General Household Survey of 1982 suggest that differences in use were minimal and insignificant between Scotland, England, and Wales (Haynes 1991). Since devolution of the United Kingdom in 1999, the Scottish Parliament and Government have been responsible for most areas of domestic policy, including the health system. During this period, Scotland spent a higher proportion of its national budget on healthcare than England, Wales, or Northern Ireland (Connolly et al. 2010; Sutherland and Coyle 2009). According to the OECD Regional Database (Directorate of Public Governance and Territorial Development 2014), the number of physicians

per head in Scotland was at the median of the 16 countries under review in 2010 (data not available for Finland).

Finally, the health-behavior hypothesis posits that poorer health behavior in Scotland may have driven observed mortality differentials. This hypothesis has received little attention, although a comparison of 2003 Health Surveys in Scotland and England provides little basis for the belief that behavioral differences account for Scotland's divergence from England (McCartney et al. 2012a; Shelton 2009). One study, which was focused on people at younger ages, suggests that illicit drug use may account for up to a third of Scotland's mortality disadvantage relative to England between 2002 and 2005 (Bloor et al. 2008).

Smoking in Scotland

Smoking prevalence in the United Kingdom has historically been higher than in other OECD countries, particularly among women (Figure A2.1). With few exceptions, women in the United Kingdom reported the highest prevalence of cigarette smoking over a 50-year period from 1960 (42.0 per cent) to 2009 (20.7 per cent) (Figure A2.1.A.) (Directorate of Public Governance and Territorial Development 2014).

Within the United Kingdom, several sources that permit comparisons of smoking levels in different countries routinely show Scotland to have the highest prevalence of smoking and the highest consumption of cigarettes per adult. The 2003 Health Surveys in Scotland and England found that 29 per cent of adult Scottish men were current smokers, compared to 27 per cent in England; the figures for women were 28 per cent and 24 per cent (Shelton 2009). The General Lifestyle Survey collects individual-level information on smoking status by country within the United Kingdom. It shows that between 1978 and 2009, the annual prevalence of cigarette smoking among men and women aged 16+ years was highest in Scotland. Among 2010 adult current smokers, Scottish men and women smoked an average of 14.8 and 13.1 cigarettes per day

(Gray and Leyland 2013), respectively, compared to 13.3 and 12.1 cigarettes per day in England (Eastwood 2012).

The influence of smoking on mortality depends on a number of the characteristics of the practice, including its duration, intensity, age at initiation, years since quitting, nicotine content, and how deeply the smoker inhales. These are not readily captured in a single indicator. An alternative indicator is the death rate from lung cancer. In those developed countries with high rates of smoking, 90 per cent of deaths from lung cancer have been attributed to smoking, in the sense that they would not have occurred if no one had smoked (Peto et al. 1992). On this indicator, Scotland's history of heavy smoking has left a vivid mark. Figure 2.2. shows that Scotland's death rate from lung cancer at ages 50 and above, age-standardized to the European standard population, has been significantly higher than comparison countries for women since 1970. For men, Scotland has had the first or second highest lung cancer mortality since 1960, sometimes exchanging places with Belgium in these positions.

Methods for calculating smoking-attributable mortality

To circumvent the limitations of survey-based estimates of the mortality effects of smoking, Peto et al. (1992) developed a widely documented indirect method (hereafter called the Peto-Lopez method) to estimate smoking-attributable mortality at the population level. The Peto-Lopez method uses lung cancer mortality rates as a proxy for the cumulative impact on mortality of smoking over the life-course (Peto et al. 1992). The accuracy of lung cancer death coding on death certificates is high, and as noted, approximately 90 per cent of deaths from this disease are directly attributable to the impact of smoking in countries with high rates of smoking (Peto et al. 1992). The Peto-Lopez method combines population-level rates of death from lung cancer with the cause-specific relative risk of mortality between non-smokers and smokers. It assumes that, in the absence of smoking, the set of age/sex-specific death rates for lung cancer would be those recorded in the largest prospective cohort study of the mortality hazards of smoking, the Cancer

Prevention Study II (CPS-II). This study is a US-based longitudinal study of approximately 1.2 million individuals followed from 1982 to 1988 (Thun et al. 1997). The CPS-II allowed for the calculations of the relative risks of cause-specific mortality between smokers and nonsmokers. The logic of the Peto-Lopez method is to map 'excess' lung cancer death rates onto an estimate of the prevalence of smoking in a population, and then to use the estimated prevalence to estimate the proportion of deaths from other specific causes that are attributable to smoking. The number of deaths attributable to smoking is the sum across causes of death of the cause-specific attributable deaths.

An alternative indirect method, hereafter called the PGW method, was developed by Crimmins et al. (2010) and Preston et al. (2010b). This method also uses lung- cancer death rates as an indicator of the cumulative damage caused by smoking but does not use the relative risks from the CPS-II, which may not be generalizable to all populations (Preston et al. 2010a). The PGW method relies instead solely on the macro-level statistical relationship between age-specific mortality rates for lung cancer and age-specific mortality rates for all other causes of death combined. Preston et al. estimated the parameters of this relationship using annual data for the period 1950 to 2007 for 21 high-income countries. The data set contained 9.9 billion person-years of exposure and 285 million deaths. Estimates of the smoking-mortality relations were made separately for the two sexes and 5-year age groups, controlling period effects, country effects, and interactions between the two. The analyses presented in this paper used the PGW method (Preston et al. 2010b) to estimate smoking-attributable mortality indirectly. We show that results using the Peto-Lopez method are extremely close to those using the PGW method.

Oza et al. (2011) examine time-patterns of relative mortality risks of smokers for various causes of death. Relative to the lag between smoking behavior and death for lung cancer, they found the lag structure to be longer for chronic obstructive pulmonary disease (COPD) and shorter for cardiovascular diseases. Using the Peto and Lopez method, the estimated number of

deaths attributable to smoking differed by only 1.7 per cent when cause-specific lag structures were incorporated compared to when they were not. Thus, it appears that the pattern of lung cancer lags is sufficiently similar to that for the aggregate of other causes of death that serious distortions do not arise from assuming that they are, on average, the same.

Data

Our study used the following data from the Human Mortality Database (HMD) (HMD 2014) for the period 1955-2009: annual all-cause death counts by sex in 5-year age intervals from 50-54 to 85+; population-exposure estimates; and life-tables by sex and 5-year age group. Annual death counts by cause of death, sex, and 5-year age group were taken from the World Health Organization Mortality Database for the period 1955-2009 (Mathers 2014). The International Classification of Diseases, Injuries, and Causes of Death was used. For each year-country-sex-age group, the distribution of lung cancer deaths was calculated from the World Health Organization Mortality Database. This distribution was applied to the all-cause death counts from the Human Mortality Database to determine lung cancer deaths and death rates. For the application of the Peto-Lopez method, the cause-specific death counts in the WHO Mortality Database were used to distribute deaths in the HMD by cause.

Methods

Attributable-risk calculation in the PGW method

In the PGW method, the fraction of lung cancer deaths attributable to smoking at a particular age is calculated as follows:

$$A_L = \frac{M_L - M_L'}{M_L}$$

where M_L is the observed death rate from lung cancer and M_L' is the age- and sex-specific lung cancer death rate among non-smokers in CPS-II (Thun et al. 1997). The fraction of deaths attributed to smoking from all causes of deaths other than lung cancer uses the model coefficients

produced by Preston et al. (2010b).The fraction of not-lung-cancer deaths is calculated as follows:

$$A_O = 1 - e^{\beta_L'(M_L - M_L')}$$

where β_L' is the model coefficient of lung cancer mortality.

The fraction of all deaths that is attributable to smoking is calculated as a weighted average of deaths attributable to lung cancer and not-lung-cancer deaths:

$$A = \frac{A_L D_L + A_O D_O}{D}$$

where D_L, D_O, and D are deaths attributable to lung cancer, to not-lung-cancer causes, and total deaths, respectively.

Age-specific death rates following the removal of smoking-attributable mortality are calculated as:

$$m_i^{-s} = m_i(1 - A_i)$$

for each age group i=50-54, 55-59, … , 80-84, 85+.

Sex-specific life tables are created using standard methods (Wilmoth et al. 2007) in STATA 13.1 by the use of death counts and exposure counts for each age, country, sex, and period. Age-specific mortality rates are created before and after the removal of smoking-attributable mortality.

Results

Smoking-attributable mortality above age 50, by sex and country, 1955-2009

Table 2.2. shows the estimated proportion of deaths attributable to smoking obtained by applying the PGW method to data from Scotland and comparison countries in 1955, 1980, and 2009. In 2009, Scottish women have a higher percentage of deaths attributable to smoking (28 per cent)

than any other country; the mean percentage for other countries is only 11 per cent. Other English-speaking countries have a higher attributable fraction (18 per cent) than European countries (6 per cent). The high smoking-attributable fraction in Scotland is long-standing; Scottish women had the highest fraction, 10 per cent, in 1980.

The story for men is similar but less dramatic. Scotland has the second-highest fraction of deaths attributable to smoking in 2009, just behind Belgium (26 per cent vs. 25 per cent). The mean for all comparison countries is 18 per cent. So the Scottish excess in fractions attributable to smoking is smaller for men than for women, in accordance with the narrower gap in Scottish life expectancy for men shown in Table 2.1. The relatively high smoking-attributable fraction of deaths for Scottish men was also present in 1955 (16 per cent vs. a mean of 8 per cent for comparison countries) and 1980 (33 per cent vs. 20 per cent). Reflecting the ebbing of the smoking epidemic among them, men in most countries, including Scotland, had a higher fraction of deaths attributable to smoking in 1980 than in 2009.

Also shown in Table 2.2. are the 2009 estimates made using the Peto-Lopez method. This method too shows Scotland to have the highest smoking-attributable fraction of deaths of any country. The two series track one another very closely, with a correlation coefficient between them of 0.97 for women and 0.95 for men. Furthermore, the mean of the two series for the 15 comparison countries is identical for women (0.11) and nearly identical for men (0.19 vs. 0.18). Clearly, both series similar yield the same conclusions about the importance of smoking for the Scottish shortfall in life expectancy.

Effect on life expectancy

Table 2.3. shows the impact of removing deaths attributable to smoking on the life expectancy at age 50 for each country in 2009. Smoking-attributable mortality reduces the life expectancy at age 50 by a full 3.10 years for Scottish men in 2009, compared to a mean of 2.14 years for comparison countries. Of the initial gap of 1.69 years between Scotland and comparison countries

before the removal of smoking-attributable deaths, only 0.74 years remains after their removal. Thus, smoking accounts for 56 per cent [1-(0.74/1.69)] of Scottish men's shortfall in life expectancy compared to the mean of comparison countries. Once smoking-related deaths are removed, Scottish men have nearly the same life expectancy (32.09 years) as the mean of European countries (32.30).

Among Scottish women, smoking-attributable mortality reduces life expectancy at age 50 by 3.59 years in 2009, compared to 1.37 years for the mean of women in comparison countries. The removal of smoking-related deaths reduces the original gap between Scotland and the mean of comparison countries to 0.45 years, compared to the original gap of 2.67 years. So smoking accounts for about 83 per cent of the difference in women's life expectancy between Scotland and the mean of comparison countries. When the comparison is restricted to other English-speaking countries, smoking accounts for 64 per cent of the gap. This figure is lower because women in other English-speaking countries were themselves heavy smokers, losing an average of 2.19 years of life expectancy compared to only 0.74 years for European countries.

Scottish women lost more years of life expectancy to smoking than Scottish men in 2009, 3.59 years compared to 3.10 years. One consequence is that the sex difference in life expectancy in Scotland is a relatively low 3.41 years, compared to a mean of 4.39 years in comparison countries. Once smoking-related deaths are removed, Scotland's gap in life expectancy between the sexes rises to 3.90 years, similar to the gap of 3.62 years for comparison countries. So smoking is primarily responsible for the anomalously low sex difference in life expectancy in Scotland.

More generally, smoking patterns are disturbing the international pattern of differences in life expectancy between women and men. Among European comparison countries, men lost 1.96 years to smoking compared to only 0.74 years for women. In contrast, English-speaking countries had roughly equal losses to smoking for men and women: women lost an average of 2.19 years

compared to 2.28 years for men. The correlation between men's and women's life expectancy among the 16 countries in 2009 was 0.68 before the removal of smoking deaths and 0.85 after their removal. Men's and women's life expectancies were more closely aligned after the removal of the exceptionally heavy impact of smoking on the life expectancy of English-speaking women.

In order to provide more temporal depth, Figure 2.3.A. shows the contribution of smoking to the difference in women's life expectancy at age 50 between England and Wales and Scotland annually between 1950 and 2009. The differences in life expectancy are calculated both before and after the removal of smoking-attributable deaths. A third-degree polynomial is fit to the two series. Before the removal of smoking-attributable deaths, the Scottish deficit in life expectancy grows slowly over time. The removal of smoking mortality, however, reduces the mortality gap, from about 1.5 years in 1950 to less than 0.5 years from 2000 to 2009.

Figure 2.3.B. is the corresponding figure for men. Before allowance for smoking, the disparity in life expectancy for men worsens over time in Scotland relative to England and Wales. The removal of smoking, however, largely eliminates this widening divergence and holds the Scottish disparity for men relatively constant over the period.

Smoking has little effect for either sex on the gap in life expectancy between England and Wales and Scotland in the 1950s. But its influence strengthens thereafter, especially among women after 1980. By 2010, the difference in life expectancy at age 50 is reduced to only 1.03 years for men and 0.56 years for women after the removal of smoking-attributable mortality, compared to actual differences of 1.78 and 1.76 years.

Discussion

The foregoing analysis finds strong evidence that smoking-attributable mortality is the primary driver of Scotland's large and widening mortality disadvantage relative to other developed countries. Smoking-attributable mortality in Scotland reduces life expectancy at age 50 in 2009 by a full 3.59 years for women and 3.10 years for men. The reduction for women is higher than in

any of 15 comparison countries, in which smoking reduces life expectancy by an average of 1.37 years. For men, only Belgium lost more years of life to smoking than Scotland. Smoking had an especially strong influence on women's mortality in Scotland after 1980.

The indirect estimation method used here has several limitations. One is that the model coefficients taken from Preston et al. (2010b) rely on the statistical association between lung cancer mortality and smoking. In populations where conditions other than smoking may contribute heavily to lung cancer mortality, the assumption that lung cancer mortality represents the cumulative impact from smoking on mortality may not be valid. Other contributors to lung cancer mortality include exposure to heavy pollution and such behavioral practices as indoor coal burning. Among Asians living in Asia, for example, the incidence of and mortality from lung cancer are significantly higher among nonsmokers than among their European counterparts (Thun et al. 2008). In historically heavy-smoking populations such as the United Kingdom, however, smoking is the primary contributor to lung cancer mortality (Ezzati and Lopez 2003).

Two possible sources of bias in the estimates should be borne in mind. One is differences between the samples of non-smokers used. The expected rates of death from lung cancer among non-smokers used in the PGW method are drawn from the CPS-II study, which is of a US population of predominantly middle-class, college-educated white persons (Thun et al. 1997). In contrast, the rates for non-smokers presented by (Thun et al. 2008) using the CPS-II are similar to those reported in other samples (Doll et al. 1994; Enstrom 1979). Another possible source of bias is differences in the classification of lung cancer death rates across countries, though the accuracy of lung cancer classification remains consistently high in industrialized countries (Percy et al. 1981).

The reasons for Scotland's higher prevalence of smoking, lung cancer mortality, and years of life sacrificed to smoking remain uncertain. There is nothing in government policy on tobacco use that might account for heavy smoking in Scotland relative to the rest of Great Britain.

Before devolution in 1999 and the formation of the Scottish Parliament, England, Wales, and Scotland were subject to the same policy. By 1930, Great Britain had the highest rate of lung cancer mortality in the world according to the Royal College of Physicians Action on Smoking and Health (ASH) council (Action on Smoking and Health 2014). The rising recognition of Britain's smoking epidemic led the British Parliament to introduce a massive 43 per cent tax on tobacco cigarettes in 1947. Parliament banned cigarette advertisements on television in 1965 and on radio in 1978. Upon devolution in 1999, the Scottish Parliament took a relatively stronger stance than England and Wales in legislation on smoke-free areas. In 2006, Scotland became the first United Kingdom country to ban smoking in all enclosed public spaces and workplaces.

Heavy smoking in Scotland is not inconsistent with the country being relatively deprived socioeconomically: deprivation may lead to smoking, which may then function as the 'proximate' cause of disease. That would be consistent with the widespread observations that, within developed countries, including Scotland, people of lower education or income are more likely to smoke than those without these disadvantages (Huisman et al. 2005). The 2003 Health Survey in Scotland (Bromley et al. 2003) found that, among men, 15 per cent in the highest household income quintile smoked cigarettes compared with 51 per cent in the lowest quintile. The corresponding figures for women were 13 per cent and 45 per cent.

Scotland has reduced or eliminated its economic disadvantages relative to England and Wales, but not its longevity disadvantage. As we noted earlier, the 'Scottish effect' refers to the increasing inability of relative deprivation to account for the poor ranking of Scottish longevity within the United Kingdom after 1980. Identifying smoking as the principal reason for the shortfall in Scottish longevity may help account for this anomaly. Because of the long lag between smoking behavior and it mortality consequences, smoking-attributable deaths in any particular period reflect smoking behavior over many previous decades. In the case of smoking, the past casts a long shadow.

References

Action on Smoking and Health. 2014. "Key dates in the history of anti-tobacco campaign." http://www.ash.org.uk/files/documents/ASH_741.pdf.

Bloor, M., M. Gannon, G. Hay, G. Jackson, AH. Leyland, and N. McKeganey. 2008. Contribution of problem drug users' deaths to excess mortality in Scotland: secondary analysis of cohort study, *British Medical Journal* 337:a478.

Carstairs, V., and R. Morris. 1989. "Deprivation: explaining differences in mortality between Scotland and England and Wales." *British Medical Journal* 299(6704):886-889.

Connolly, S., G. Bevan, and N. Mays. 2010. "Funding and Performance of Healthcare Systems in the Four Countries of the UK Before and After Devolution." London: Nuffield Trust.

Connolly, S., M. Rosato, H. Kinnear, and D. O'Reilly. 2011. "Variation in mortality by country of birth in Northern Ireland: a record linkage study." *Health Place* 17(3):801-806.

Crimmins, EM., SH. Preston, and B. Cohen. 2011. "Explaining Divergent Levels of Longevity in High-Income Countries." Washington (DC): National Academies Press (US) National Academy of Sciences.

Crimmins, EM., SH. Preston, and B. Cohen (eds.). 2010. "International Differences in Mortality at Older Ages: Dimensions and Sources." In Panel on Understanding Divergent Trends in Longevity in High-Income Countries, National Research Council. Washington (DC): National Academies Press (US).

Directorate of Public Governance and Territorial Development. 2014. "OECD Regional Database." Accessed November 2016 at: www.oecd.org/gov/regional/statisticsindicators.

Doll, R., R. Peto, J. Boreham, and I. Sutherland. 2004. "Mortality in relation to smoking: 50 years' observations on male British doctors." *British Medical Journal* 328(7455):1519.
Doll, R., R. Peto, K. Wheatley, R. Gray, and I. Sutherland. 1994. "Mortality in relation to smoking: 40 years' observations on male British doctors." *British Medical Journal* 309(6959):901-911.

Eastwood, P. 2012. "Statistics on Smoking: England, 2012." Leeds (UK): The Health and Social Care Information Centre, National Health Service.

Enstrom, JE. 1979. "Rising lung cancer mortality among nonsmokers." *Journal of the National Cancer Institute* 62(4):755-760.

Ezzati, M. and AD. Lopez. 2003. "Estimates of global mortality attributable to smoking in 2000." *Lancet* 362(9387):847-852.

Ezzati, M., AD. Lopez, A. Rodgers, S. Vander Hoorn, and CJL. Murray. 2002. "Selected major risk factors and global and regional burden of disease." *Lancet* 360(9343):1347-1360.

Gray, L. and AH. Leyland. 2003. "The Scottish Health Survey 2012 - Volume 1 Main Report." Accessed November 2016 at: www.scotland.gov.uk.

Hanlon, P., R. Lawder, D. Buchanan, A. Redpath, D. Walsh, R. Wood, M. Bain, D. Brewster, J. Chalmers, and D. Walsh. 2005. "Why is mortality higher in Scotland than in England and Wales? Decreasing influence of socioeconomic deprivation between 1981 and 2001 supports the existence of a 'Scottish Effect'." *Journal of Public Health* 27(2):199-204.

Hanlon, P., D. Walsh, D. Buchanan, A. Redpath, M. Bain, D. Brewster, J. Cahalmers, R. Muir, M. Samlls, and J. Willis. 2001. "Chasing the Scottish Effect: Why Scotland needs a Step-change in health if it is to catch up with the rest of Europe." Public Health Institute of Scotland, Glasgow (UK): Common Services Agency for the Scottish Health Service, Information & Statistics Division.

Haynes, R. 1991. "Inequalities in health and health service use: evidence from the General Household Survey." *Social Science & Medicine* 33(4):361-368.

Hill, AB. 1965. "The environment and disease: association or causation?" *Proceedings of the Royal Society of Medicine* 58(5):295.

Huisman, M., AE. Kunst, and JP. Mackenbach. 2005. "Educational inequalities in smoking among men and women aged 16 years and older in 11 European countries." *Tobacco Control* 14(2):106-113.

Human Mortality Database. 2014. "Human Mortality Database." Berkeley (USA) and Rostock (Germany): University of California and Max Planck Institute for Demographic Research. Accessed November 2014 at: www.mortality.org.

Mathers, C. 2014. "WHO Mortality Database." Geneva (Switzerland): World Health Organization: Innovation, Information, Evidence and Research Cluster.Department of Health Statistics and Information Systems. Accessed November 2014 at: www.who.int/healthinfo/statistics/mortality_data.
McCartney, G., C. Collins, D. Walsh, and GD. Batty. 2012a. "Why the Scots die younger: synthesizing the evidence." *Public Health* 126(6):459-470.

McCartney, G., D. Walsh, B. Whyte, and C. Collins. 2012b. "Has Scotland always been the 'sick man' of Europe? An observational study from 1855 to 2006." *European Journal of Public Health* 22(6):756-760.

Oza, S., MJ. Thun, SJ. Henley, AD. Lopez, and M. Ezzati. 2011. "How many deaths are attributable to smoking in the United States? Comparison of methods for estimating smoking-attributable mortality when smoking prevalence changes." *Preventive Medicine* 52(6):428-433.

Percy, C., E. Stanek, and L. Gloeckler. 1981. "Accuracy of cancer death certificates and its effect on cancer mortality statistics." *American Journal of Public Health* 71(3):242-250.

Peto, R., AD. Lopez, J. Boreham, M. Thun, and C. Heath, Jr. 1992. "Mortality from tobacco in developed countries: indirect estimation from national vital statistics." *Lancet* 339(8804):1268-1278.

Popham, F., PJ. Boyle, and P. Norman. 2010. "The Scottish excess in mortality compared to the English and Welsh. Is it country of residence or country of birth excess?" *Health Place* 16(4):759-762.

Preston, SH., EM. Crimmins, and JR. Wilmoth. 2010a. "A new method for estimating smoking-attributable mortality in high-income countries." *International Journal of Epidemiology* 39(2):430-438.

Preston, SH., EM. Crimmins, and JR. Wilmoth. 2010b. "Contribution of smoking to international differences in life expectancy." in SH. Preston, EM. Crimmins, and B. Cohen (eds). International Differences in Mortality at Older Ages: Dimensions and Sources. Washington (DC): National Academies Press (US) National Academy of Sciences, pp. 105-131

Shelton, NJ. 2009. "Regional risk factors for health inequalities in Scotland and England and the 'Scottish effect'." *Social Science & Medicine* 69(5):761-767.

Surgeon General's Advisory Committee on Smoking and Health. 1964. "Smoking and Health: Report of the Advisory Committee to the Surgeon General of the Public Health Service." Washington, D.C. (USA): US Public Health Service.

Sutherland, K. and N. Coyle. 2009. "Quality in healthcare in England, Wales, Scotland, Northern Ireland: An Intra-UK Chartbook." London: Health Foundation.

Thun, M., C. Day-Lally, DG. Myers, EE. Calle, WD. Flanderss, and BP. Zhu. 1997. "Trends in tobacco smoking and mortality from cigarette use in cancer prevention studies I (1959-1965) and II (1982-1988)." in D. Burns, L. Garfinkel, and J.M. Samet (eds), Changes in Cigarette-Related Disease Risks and Their Implications for Prevention and Control, Smoking and Tobacco Control. Bethesda, MD: Cancer Control and Population Sciences, National Cancer Institute, U.S. National Institutes of Health, pp. 305-382

Thun, MJ., LM. Hannan, LL. Adams-Campbell, P. Boffetta, JE. Buring, D. Feskanich, WD. Flanders, SH. Jee, K. Katanoda, and LN. Kolonel. 2008. "Lung cancer occurrence in never-smokers: an analysis of 13 cohorts and 22 cancer registry studies." *PLoS Medicine* 5(9):e185.

US Department of Health and Human Services. 2014. "The Health Consequences of Smoking—50 Years of Progress: A report of the Surgeon General." US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health.

Walsh, D., M. Taulbut, and P. Hanlon. 2010. "The aftershock of deindustrialization: trends in mortality in Scotland and other parts of post-industrial Europe." *European Journal of Public Health* 20(1):58-64.

Wilmoth, JR., K. Andreev, D. Jdanov, DA. Glei, C. Boe, M. Bubenheim, D. Philipov, V. Shkolnikov, and P. Vachon. 2007. "Methods Protocol for the Human Mortality Database." Berkeley (USA) and Rostock (Germany): University of California and Max Planck Institute for Demographic Research. Accessed November 2014 at: www.mortality.org.

Table 2.1. Life expectancy at birth, e(0), and age 50, e(50,) in 2009 for Scotland and comparison countries.

	Females		Males	
Country	e(0)	e(50)	e(0)	e(50)
Scotland	80.45	32.41	75.87	29.00
Australia	84.16	36.07	79.70	32.35
Canada	83.28	35.52	78.85	31.60
England and Wales	82.43	34.17	78.29	30.78
Ireland	82.23	33.68	77.24	29.78
New Zealand	82.34	33.98	78.37	30.96
United States of America	81.04	33.58	76.13	29.85
Austria	82.87	34.18	77.43	29.62
Belgium	82.44	33.91	77.16	29.43
Finland	83.14	34.54	76.51	29.18
France	84.46	35.96	77.80	30.35
Italy	84.24	35.72	79.22	31.20
Portugal	82.46	34.06	76.42	29.19
Sweden	83.33	34.39	79.33	31.20
Switzerland	84.20	35.39	79.63	31.56
Japan	86.39	38.63	79.55	31.80
Mean, English-Speaking Comparison Countries	82.58	34.50	78.10	30.89
Mean, Non-English Speaking European Countries	83.39	34.77	77.94	30.22
Mean, All Comparison Countries	83.27	34.92	78.11	30.59

Note: 2008 estimates are shown for New Zealand.

Source: Human Mortality Database (accessed November 2014).

Table 2.2. Estimated smoking-attributable fraction of deaths at age 50+ in 1955, 1980, 2009, by sex and country.

	Females				Males			
	PGW			Peto-Lopez	PGW			Peto-Lopez
Country	1955	1980	2009	2009	1955	1980	2009	2009
Scotland	0.02	0.10	0.28	0.25	0.16	0.33	0.25	0.26
Australia	0.00	0.04	0.13	0.12	0.07	0.22	0.16	0.17
Canada	0.01	0.06	0.22	0.21	0.07	0.22	0.22	0.22
England and Wales	0.02	0.09	0.19	0.18	0.17	0.30	0.20	0.20
Ireland	0.02	0.07	0.16	0.19	0.04	0.17	0.20	0.21
New Zealand	0.00	0.06	0.15	0.17	0.08	0.21	0.16	0.18
United States of America	0.01	0.08	0.20	0.21	0.08	0.23	0.21	0.22
Austria	0.01	0.03	0.07	0.07	0.15	0.21	0.15	0.18
Belgium	0.00	0.01	0.07	0.07	0.09	0.30	0.26	0.26
Finland	0.01	0.02	0.05	0.05	0.18	0.28	0.15	0.16
France	0.00	0.00	0.05	0.05	0.05	0.17	0.19	0.20
Italy	0.00	0.01	0.07	0.06	0.04	0.20	0.23	0.22
Portugal	0.00	0.00	0.01	0.01	0.02	0.07	0.12	0.15
Sweden	0.00	0.02	0.09	0.10	0.03	0.10	0.10	0.11
Switzerland	0.00	0.01	0.06	0.09	0.09	0.19	0.14	0.16
Japan	0.00	0.03	0.10	0.05	0.01	0.11	0.21	0.18
Mean, English-Speaking Comparison Countries	0.01	0.07	0.18	0.16	0.08	0.22	0.19	0.20
Mean, Non-English Speaking European Countries	0.00	0.01	0.06	0.08	0.08	0.19	0.17	0.18
Mean, All Comparison Countries	0.01	0.04	0.11	0.11	0.08	0.20	0.18	0.19

Note: PGW denotes estimation using the Preston-Glei-Wilmoth method. Peto-Lopez denotes use of the Peto-Lopez method. 2008 estimates are shown for New Zealand.

Table 2.3. Life expectancy at age 50 (e50) in 2009 before and after the removal of deaths attributable to smoking.

	Females			Males		
Country	With Smoking	Without Smoking	Difference	With Smoking	Without Smoking	Difference
Scotland	32.41	35.99	-3.59	29.00	32.09	-3.10
Australia	36.07	37.53	-1.46	32.35	34.23	-1.88
Canada	35.52	38.46	-2.94	31.60	34.40	-2.80
England and Wales	34.17	36.45	-2.29	30.78	33.09	-2.31
Ireland	33.97	35.85	-1.88	29.95	32.21	-2.26
New Zealand	34.24	36.05	-1.81	31.21	33.05	-1.83
United States of America	33.58	36.34	-2.76	29.85	32.42	-2.57
Austria	34.52	35.38	-0.86	29.82	31.58	-1.76
Belgium	34.25	35.19	-0.94	29.62	32.81	-3.19
Finland	34.88	35.47	-0.59	29.33	31.04	-1.71
France	36.49	37.16	-0.68	30.62	33.08	-2.46
Italy	35.72	36.49	-0.76	31.20	33.78	-2.58
Portugal	34.06	34.23	-0.17	29.19	30.59	-1.40
Sweden	34.39	35.56	-1.17	31.20	32.21	-1.01
Switzerland	35.66	36.44	-0.78	31.76	33.33	-1.57
Japan	38.63	40.02	-1.39	31.80	34.57	-2.77
Mean, English-Speaking Comparison Countries	34.59	36.78	-2.19	30.96	33.23	-2.28
Mean, Non-English Speaking European Countries	35.00	35.74	-0.74	30.34	32.30	-1.96
Mean, All Comparison Countries	35.08	36.44	-1.37	30.69	32.83	-2.14

Note: 2008 estimates are shown for New Zealand.

Figure 2.1.A. Trends in life expectancy at age 50, e(50), among high-longevity countries, females 1950-2009.

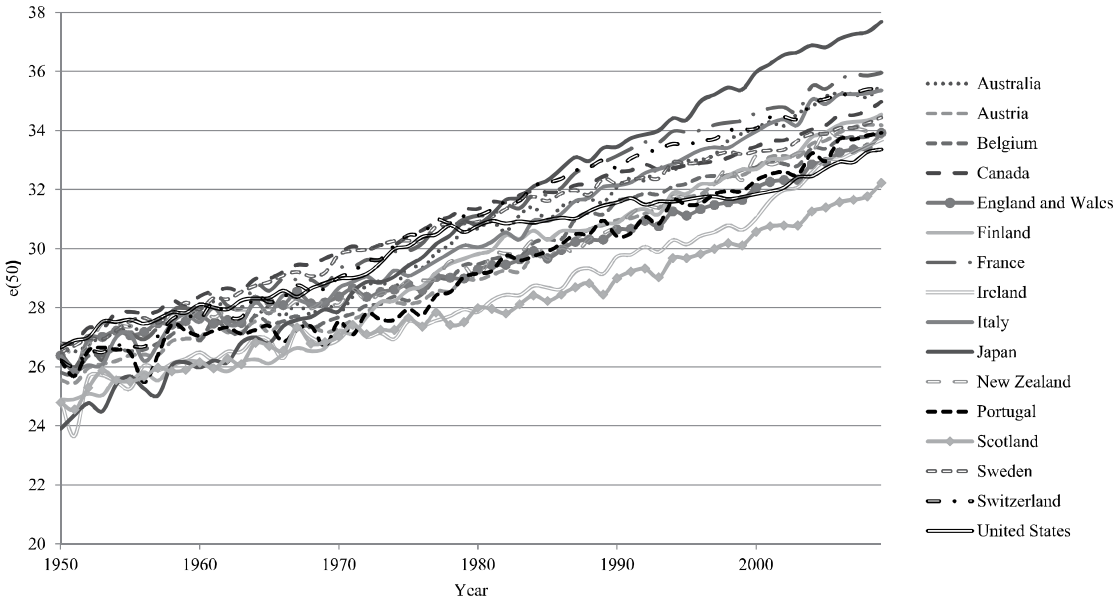
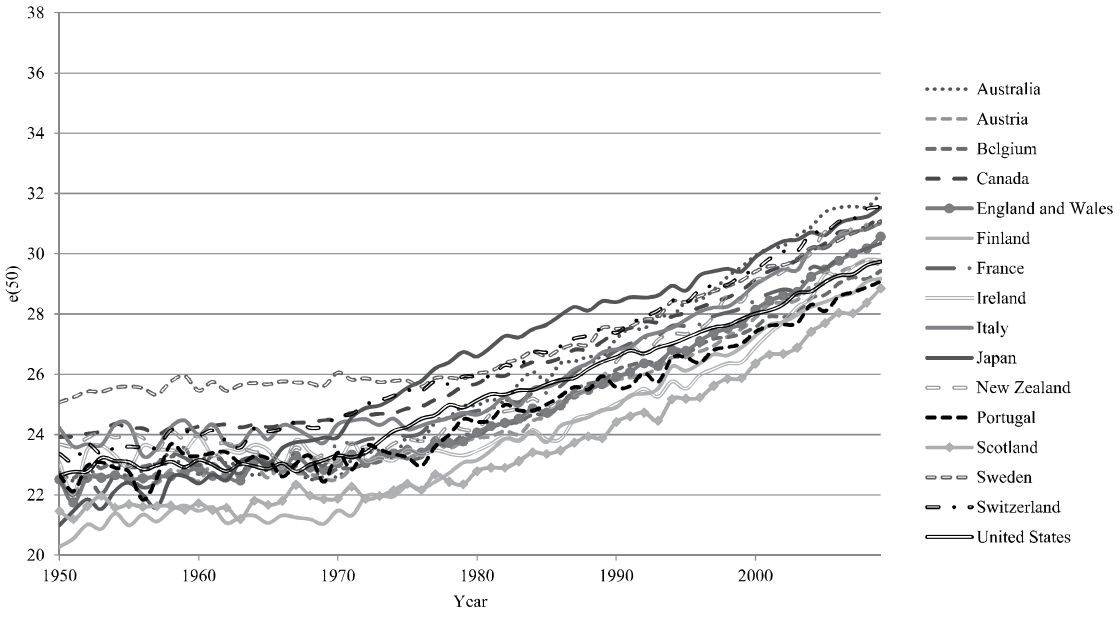


Figure 2.1.B. Trends in life expectancy at age 50, e(50), among high-longevity countries, males 1950-2009.



Source: Human Mortality Database (accessed November 2014).

Figure 2.2.A. Lung cancer death rates at ages 50 and above by country, women 1950-2009.

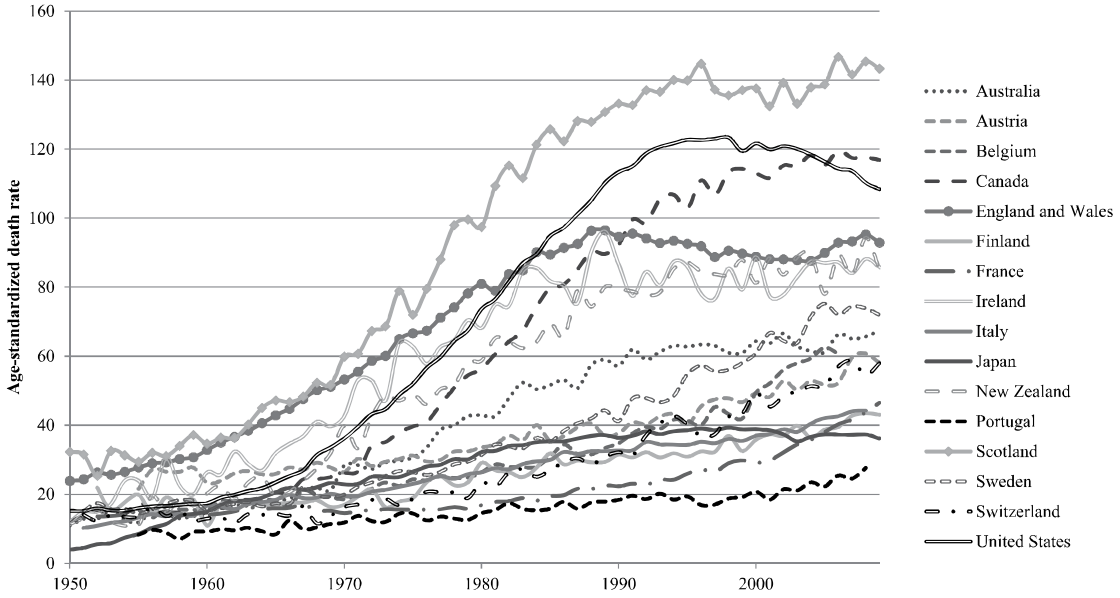
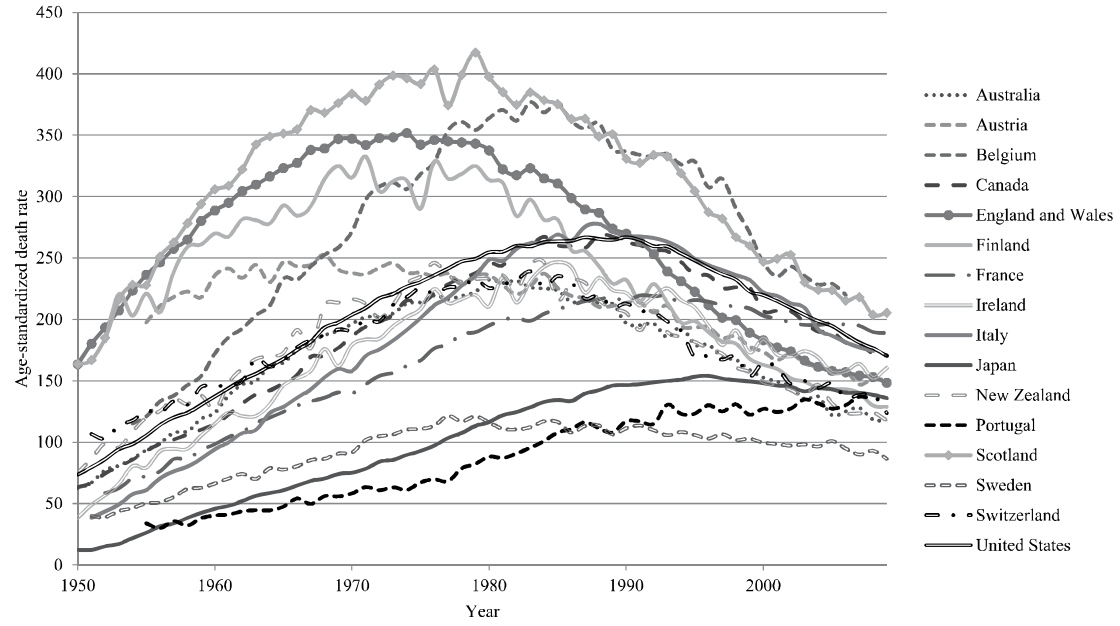


Figure 2.2.B. Lung cancer death rates at ages 50 and above by country, men 1950-2009.



Source: Human Mortality Database (accessed November 2014) and the World Health Organization Mortality Database (accessed November 2014).

Figure 2.3.A. Differences in life expectancy at age 50, e(50), before and after the removal of smoking-attributable mortality, Scotland vs. England & Wales, 1950-2009, women.

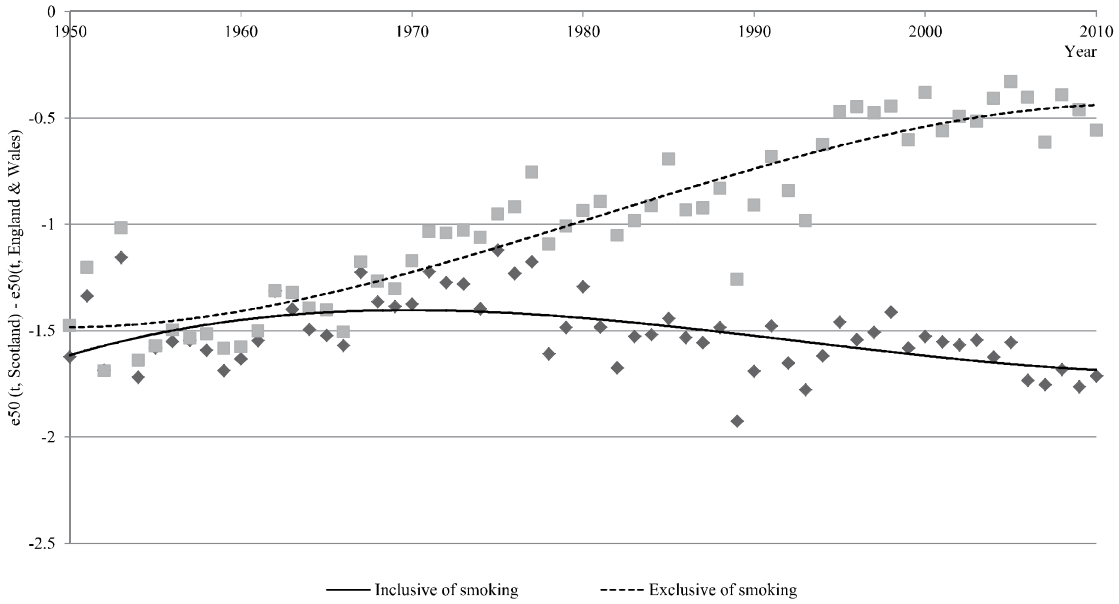
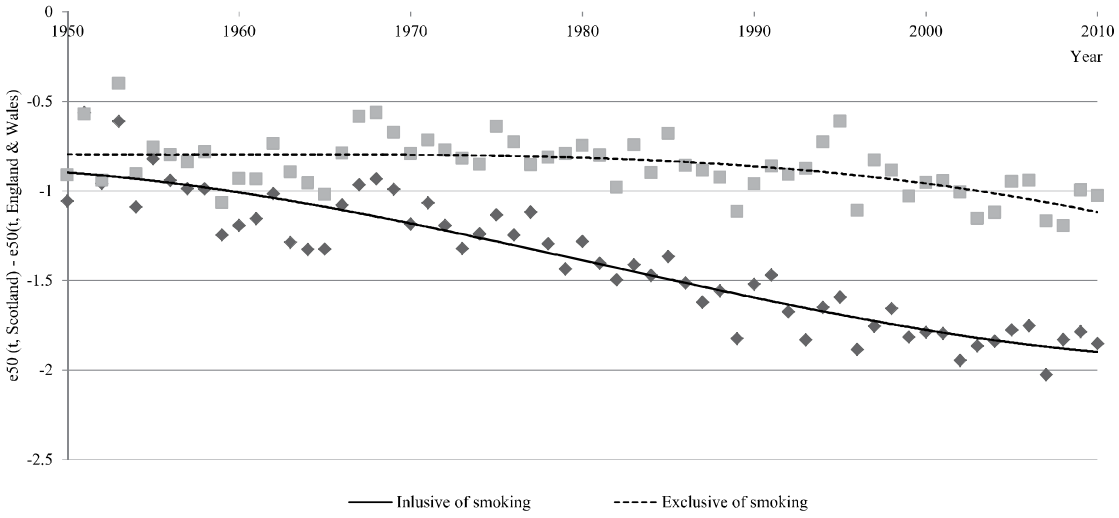


Figure 2.3.B. Differences in life expectancy at age 50, e(50), before and after the removal of smoking-attributable mortality, Scotland vs. England & Wales, 1950-2009, men.



Note: Differences in life expectancy (Scotland - England & Wales) are calculated both before and after the removal of smoking-attributable deaths. Each trend lines represents a 3rd degree polynomial.
Source: As for Figure 2.

CHAPTER 3

Mortality among first and second generation migrants in England and Wales

Introduction

Despite overall improvements in health and mortality over the past century, inequalities remain within populations. Health inequalities across socioeconomic groups are often observed, with gaps persisting and arguably increasing between those of lower socioeconomic status and those of higher status (Feinstein 1993). Group differences may exist in the reporting, understanding, and classification of health status (King et al. 2004). Mortality offers an unambiguous measure of health, aiding in the validity of observed inequality. Disparities in mortality across socioeconomic groups have been consistently observed in industrialized societies. This association is robust over time and different welfare systems (Mackenbach et al. 2008, Mackenbach et al. 2003, Feinstein 1993) including the United States (Pappas et al. 1993), Australia (Lawson and Black 1993), the United Kingdom (Vagerö and Lundberg 1989), and Nordic countries (Vagerö and Lundberg 1989, Mackenbach and Kunst 1997). While scholars debate the appropriate operationalization of socioeconomic status, the existence if not the exact extent of mortality inequalities have been undeniably observed throughout this period of overall health improvements.

Health gaps across socioeconomic groups are historically prominent in the United Kingdom (Vagerö and Lundberg 1989). The National Health Service (NHS) was established in 1948 with the primary objective to achieve universal equity in healthcare for persons resident in the United Kingdom. The NHS emerged in the aftermath of World War II, during which healthcare challenges severely overburdened the nation. Over half of the British population was uninsured at the time of the war (Bradshaw and Bradshaw 1995), and the (largely uninsured) lower social class citizens relied primarily on emergency care facilities while the wealthier

classes could pay for private service. Accordingly, British thinkers have historically conceived "equity as linked inextricably to social class" (Ibid).

In the United Kingdom and across Europe, equity debates should arguably be reconceived to include ethnicity and migrant status. In addition to highlighting socioeconomic gaps, World War II (WWII) also accelerated inward migration to Europe. In the immediate aftermath of WWII (1945-1960s), decolonization and the resettlement of approximately 20 million displaced persons triggered European in-migration (Zimmermann 1995). Labor migration characterized the majority of subsequent migrant streams from 1955-1973. Approximately 5 million labor migrants from Mediterranean countries (e.g. Greece, Turkey, Morocco, Portugal, Tunisia and Yugoslavia) migrated to northern Europe through strategic guest worker programs. Migrants from former colonial states (e.g. African countries to France; India and Pakistan to Great Britain) also bolstered the foreign-born labor supply in Europe. Migrants were important in replenishing the post-WWII labor force in Europe and contributing to economic recovery (Ibid.). European migration driven by family reunification and asylum seeking (political migration) began during the1974-1988 period, with a marked acceleration post-1988. Recent European migration largely relates to the Single European Act of 1987, the collapse of the Berlin Wall in 1989, and the formal establishment of the European Union (EU) in 1993; each of which eased migration barriers between EU countries. Currently, two-thirds of migrants in the EU come from other EU countries (Rechel et al. 2013). Migration into the United Kingdom accordingly increased in recent decades, with migrants contributing a growing share of the total population. In the United Kingdom, the proportion of the population that is foreign-born steadily rose from 6.5% in 1990 to 10.4% in 2010 (Ibid.) to 12.3% in 2013 (OECD 2016).

Collectively, migrants contributed significantly to post-WWII economic growth (Zimmermann 1995) in Europe; and migrants may revive Europe from its current low fertility challenges (Rechel et al. 2013). Despite these arguably positive effects of migration, European

countries have recently exhibited an alarming increase in anti-immigrant hostility. Several countries (e.g. France and Britain) have seen a rise in nationalist party membership. In Britain, many commentators attribute the success of the June 2016 referendum to leave the EU, and subsequent rise in hate crimes, to anti-immigrant sentiment (Tilford 2016, Frayer 2016).

Due to their increasing representation in European communities and arguable status as "vulnerable and socially excluded groups" (European Commission 2009), the quantification of migrants' health is imperative for the equity debate and for structuring population health policy. Despite this obvious need, accurate and comparable information on migrant health is limited and inconsistent (Rechel et al. 2013), with little "accepted and detailed statistics" available to understand this significant subgroup (Zimmermann 1996). A common difficulty is the actual identification of ethnic or migrant persons. Self-identification of ethnicity is often used in the literature (Stronks 2009), derived from a question on the census or sample survey which asks the respondent to self-identify with pre-coded ethnic categories. This indicator, however, relies on consistent and stable self-identification of ethnic groups. This assumption is highly problematic due to the inability to validate ethnic identity across individuals and the known fluidity of self-reported ethnicity (Alba and Islam 2009). Country of birth is an improved indicator of migrant status due to its stability and objective comparability across groups and data sources (Stronks 2009).

This paper utilizes a rich and unique source of administrative, longitudinal data to estimate migrant mortality in England and Wales, with migrants defined exclusive by country of birth. This paper estimates migrant mortality across first generation migrants groups, defined by country of birth, and second generation migrant groups, defined by mother's country of birth. The paper next analyzes the extent to which observed mortality is explained by socioeconomic status, which contributes to the existing British socioeconomic health debate. This research adds important contributions to the understanding of migrant mortality globally and to the relationship

between socioeconomic status and mortality. To the author's knowledge, no prior research has estimated first generation migrant mortality in the United Kingdom in the context of adult, preventable mortality nor incorporated such robust socioeconomic indicators. Furthermore, no previous literature investigating second generation mortality, defined exclusively by maternal country of birth, was identified nor any research estimating second generation infant mortality in the United Kingdom.

Background

Migration and mortality

Migrants have been consistently observed to have better mortality outcomes than the native-born population in nearly all developed countries including the United States (Ruiz, Steffen, and Smith 2013, Hummer et al. 2007), the United Kingdom (Wallace and Kulu 2014a), Germany (Ronellenfitsch et al. 2006), France (Boulogne et al. 2012), and Belgium (Deboosere and Gadeyne 2005, Anson 2004). This mortality advantage is often termed the Migrant Mortality Advantage (MMA).

However, research on migrant mortality faces difficult accounting issues, broadly labeled as potential data artefact biases. Data artefact describes errors in observed data due to individual misreporting, such as nationality and age, and administrative errors in counting the population at risk. Migrant events (e.g. moves and deaths) may be undercounted for a variety of reasons, including a lower incentive for immigrants to register entries and exits with the host country registrars. If a migrant's emigration is missed in official registries, then that individual will continue to age immortally in host country data, thus inflating the migrant denominator (population at risk of mortality) and decreasing observed migrant morality rates. This failure to track migrant emigrations has been termed the censoring bias, and one study in Sweden found evidence that this bias did partially explain observed low migrant mortality (Weitoft et al. 1999). Furthermore, migrant deaths may be erroneously matched to existing administrative records, due

to age and ethnicity misreporting in vital statistics. For older persons, evidence from the United States does find increased matching errors for non-white native-born and foreign-born populations (Elo and Preston 1997). Collectively, these two types of data artefact biases miscount the risk set and may be called a denominator bias. Significant research conducted in the United States (Elo et al. 2004, Palloni and Arias 2004), Germany (Kibele, Scholz, and Shkolnikov 2008), Belgium (Anson 2004), France (Khlat and Courbage 1996), and Sweden (Weitoft et al. 1999) estimated the impact of the denominator bias on migrant mortality rates. These studies find that data artefact may reduce but does not eliminate the MMA, though to varying degrees for different migrant groups. A recent study in England and Wales also found that data artefact could not explain the lower mortality rates observed among migrant groups (Wallace and Kulu 2014a).

Health selection may alternatively explain observed migrant mortality. Positive health selection may occur in either the sending event, such that only exceptionally healthy individuals migrate, or the receiving context, where a migrant retains culturally positive health behaviors relative to the native-born population. Research on Mexican migration to the United States found that Mexican immigrants have better anthropometric indicators than non-migrants in Mexico and the United States (Crimmins et al. 2005). Furthermore, migrants may be selected for positive personality traits. Individuals with increased openness, extraversion (Jokela 2009, Silventoinen et al. 2008), and risk-taking (Jaeger et al. 2010) are more likely to migrate. Migrants have also been shown to have better health behaviors than the native-born population, in terms of diet (Razum et al. 1998) and substance abuse (Abraido-Lanza et al. 1999, Blue and Fenelon 2011). Psychological and anthropometric health selection may increase the likelihood of migration and remain protective after arrival for downstream health outcomes. Evidence suggests that the impact of positive health selection decreases with length of stay and subsequent generations, as the protective effects of selection may lessen by acculturation and may not be inherited by the children of migrants (Singh and Hiatt 2006, Jasso et al. 2004).

Negative health selection may additionally affect return migration. This phenomenon, termed the 'salmon-bias,' theorizes that unhealthy immigrants are more likely to migrate back to their native country than healthy immigrants. Under the salmon-bias, unhealthy individuals are selectively removed from the risk set and the remaining migrants will display misleadingly low mortality in host country data. Research conducted on Mexican immigrants in the United States (Palloni and Arias 2004, Elo et al. 2004), Turkish immigrants in Germany (Razum et al. 1998), and internal United Kingdom migrants (Wallace and Kulu 2014b) found limited to no evidence that selective return migration affects the population-level observed MMA.

Migrant health in the England and Wales

Migrant health in England in Wales is largely understudied. Among published research on adult migrant mortality in the England and Wales, methodological decisions vary widely across studies. The most important methodological factor is the use of cross-sectional rather than longitudinal data. Such papers extract population denominators and death counts from separate administrative sources to compare (age- and sex-standardized) death rates between foreign-born migrants and the native-born population, for either a single period (Marmot et al. 1984, Wild et al. 2007) or across several time points (Harding et al. 2008, Harding et al. 2009). Reliance on cross-sectional data is particularly problematic for migrant studies due to the inherent inability to accurately track migrants and capture migrant deaths. Migrants are more likely to move relative to non-migrants; and as mentioned above, migrant moves and deaths are more likely to be undercounted relative to the native population. Population denominators taken from a single census year thus do not directly match death counts derived for a subsequent time interval, and this mismatch is more likely for migrant groups.

Longitudinal reports on migrants in England and Wales are rare, and the definition of migrant status further varies among published longitudinal research. Identified longitudinal studies primarily examine first generation migrant health, using country of birth to define

migrants (Harding and Rosato 1999, Harding 2004, Wallace and Kulu 2015). These studies generally find a significant all-cause MMA among non-UK first generation migrants, and a consistent mortality disadvantage observed among UK migrants. Cause-specific analyses suggest that the observed all-cause MMA is mostly driven by low chronic disease mortality among first generation migrant groups (Wallace and Kulu 2015), particularly cardiovascular diseases (Harding and Rosato 1999, Harding 2004). Research on subsequent migrant generations is extremely limited. A series of studies do differentiate between first, second, and third generation migrants using country of birth, though only for migrants of Irish-origin (Harding and Rosato 1999, Harding and Balarajan 1996, Harding and Balarajan 2001). One report estimated mortality among descendants of migrants in England and Wales (Wallace 2015). However, this study defined descendants using self-identified ethnicity, which relied on extremely limited pre-coded categories of ethnicity. Additionally, the report could not differentiate between "white" descendants, who make up a large share of migrants in England and Wales (e.g. Scotland, Ireland, and other UK). Most importantly, the study could not distinguish between migrant generations; thus the self-identification of ethnicity was theoretically operationalized the same for second, third, fourth, etc. generation migrants. Ethnic self-identify is known to be fluid with increased acculturation (Alba and Islam 2003), and also likely differs significantly across various ethnic groups.

No identified study investigated infant mortality among second generation migrants in England and Wales. In the context of adult mortality, no identified research estimated preventable mortality among migrants. Though precise definitions vary, preventable mortality is usually defined as deaths occurring before age 75 due to causes considered to be preventable through (a) individual behaviors and/or (b) public health measures aimed at changing behaviors or exposures to harmful environments (Hutchison et al. 2006, Wheller et al. 2007). Preventable mortality is an important indicator of health inequality within a population, particularly when comparing

migrants to non-migrants, because it captures causes of death related to behaviors or healthcare access. In other words, preventable mortality is an important indicator in the health equity debate, for both socioeconomic and ethnic minority subgroups. Comparing non-Estonians to native Estonians, preventable causes of death contributed 2.19 years to the male mortality gap and 0.78 years to the female mortality gap (Baburin and Leinsalu 2011). Compared to the native-born population, Danish and Icelandic immigrants in Sweden were found to have higher death rates due to causes associated with substance abuse behavior: liver cirrhosis, lung cancer, chronic bronchitis and emphysema, and motor vehicle accidents (Westerling and Rosén 2002). Compared to the native Dutch population, immigrants in the Netherlands have been found to have higher death rates due to almost all infectious diseases, particularly for maternal and neonatal conditions (Stirbu et al. 2006). For this paper, preventable mortality is investigated in terms of behavior-attributable mortality.

This report estimates adult all-cause and behavior-attributable mortality across first generation migrants groups, defined by country of birth. Infant mortality is investigated among second generation migrant groups, defined by mother's country of birth. The paper analyzes the extent to which observed mortality is explained by socioeconomic status.

Data

Sample

The Office of National Statistics Longitudinal Study of England and Wales (LS) is a longitudinal, representative sample of 1% of the population resident in England and Wales from 1971 to present. The LS began with the 1971 census, by selecting enumerated individuals born on one of four birth dates. These four birth dates are used to update the LS population at each decennial census (1981, 1991, 2001, and 2011) or through National Health Service (NHS) Central Register records. In the United Kingdom, individuals must register with their local NHS General Practitioner for moves both into and out of an administrative authority. Thus the NHS Central

Registrar allows linkages of migrations, cancer registrations, births, and deaths to the core census record. The NHS Central Registrar is searched annually for the four LS birth dates to identify annual event, birth, and death linkages. If an individual is un-traceable, meaning unaccounted for at census day or through NHS administrative records, then that person is dropped from this analysis. The proportion of the LS population that is un-traceable remains extremely low: 3.2% in 1971; 1.1% in 1981; 1.7% in 1991; and 0.7% in 2001. The proportion un-traceable varies by subgroup, with slightly higher proportions among the foreign-born population, though sensitivity analyses demonstrate that these differences are unlikely to skew estimated mortality rates across subgroups (Wallace and Kulu 2014a).

Collectively, an individual with one of the four birth dates can enter the LS population through registration with the local GP, birth registration, or enumeration on census day. An individual exits the LS population through a registered death, registered emigration out of England and Wales, or if un-traceable on census day. Figure 3.1. shows the Lexis Diagram depicting entry and exit into the LS population used for this paper, and the portions of an individual's life when they are considered under observation or contributing to the risk set.

First generation mortality: all-cause and behavior-attributable mortality

First generation mortality is measured by adult mortality occurring between ages 15 and 85. All-cause mortality is first estimated. Cause of death is determined using the underlying cause of death and the International Classification of Diseases, Injuries, and Causes of Death (ICD) versions 9 and 10.

Preventable mortality is investigated here as behavior-attributable deaths occurring from age 15 to before age 85. In England, male and female life expectancy at birth is 79.4 years and 83.1 years, respectively, according to 2012-2014 national life-tables (Wright 2015). The most common age at death for men and women was 86 years and 89 years [Ibid]. In light of this higher life expectancy, this analysis extends the common definition of preventable mortality to deaths

occurring before age 85. This age increase also boosts the number of deaths counts, which improves the stability of mortality estimates for minority migrant groups.

Behavior-attributable mortality is operationalized as: smoking-related diseases; alcohol-related diseases; and all other diseases. Classifications of smoking-related and alcohol-related diseases are outlined in Table 3.1. and were based on United Kingdom National Statistics definitions (Goodwin 2015, Eastwood 2012).

Second generation mortality: infant mortality

Second generation mortality is measured by infant mortality. Infant mortality is a well-measured outcome, and mothers with newborns are unlikely to migrate out of the country. Virtually all births and infant deaths are captured in administrative records in the United Kingdom, with extremely high linkage rates. In 2012, approximately 98% of infant deaths registered in England and Wales were linked to their corresponding birth record (McLaren 2014). This linkage rate has been consistent over time according to the Office of National Statistics records.

Infant mortality is a highly responsive measure of socioeconomic conditions between groups, as it involves a shorter period of time between risk exposure and mortality after birth. Infant mortality will be analyzed as a binary outcome of all-cause mortality occurring before the age of 1. In the LS, birthweight was only included with birth records from 1977 onwards, so prior years were dropped from this portion of the analysis. The proportion of birth records with birthweight recorded is high across the period under observation (around 94% from 1981 onwards). Only singleton births are considered.

Defining migrant groups

Migrants are defined by country of birth. For first generation migrants, country of birth is self-reported at the census. In cases where an individual is enumerated at multiple censuses, the first reported country is birth is assigned as the country of birth. This decision was made due to the possibility of individuals listing the host (England or Wales) country with increased length of

stay. For LS members that enter the LS population through birth rather than enumeration at the census, the maternal country of birth is reported on the birth certificate. Thus for second generation migrants, migrant groups are identified by mother's country of birth, reported on the birth record.

Country of birth categories are defined in the United Kingdom as: England and Wales; Scotland; Northern Ireland; Ireland; and other United Kingdom. The remaining countries are grouped in to thirteen regions, based on regional definitions used by the CIA and political histories relevant to the United Kingdom. These thirteen regions are: Middle East; Eastern Europe; Western Europe; Australia and Oceania; East and Southeast Asia; South Asia; Central America and Caribbean; South America; North America; Western Africa; Southern Africa; Northern Africa; Central Africa; and Eastern Africa. A full list of countries and their regional groupings can be found in the appendix (Table A3.1.). Throughout the course of the analysis, further groupings will be indicated where needed to accommodate extremely low counts for certain regional groups.

Socioeconomic variables: Individual-level

The primary socioeconomic variables are individual-level and time-varying. In the adult mortality analysis, the individual-level socioeconomic variables are self-reported at the census and include: marital status; household-head social class; household car ownership; household tenure; and household density. Except for marital status, the variables all refer to the household socioeconomic circumstances under which the individual is currently resident. The household-head social class is categorized based on occupation. Categories are based on the Registrar General's definitions for social classes by occupation and are harmonized over census years. The seven categories are: I Professional; II Intermediate; III Skilled; IV Partly Skilled; IV Unskilled; Other, including Armed Forces; Missing or Not Applicable. "Not Applicable" here indicates a household-head that is economically inactive, including: students; child under 15; retirees; those

permanently sick; or other inactive. Car ownership is an alternative socioeconomic indicator often used in the United Kingdom literature on socioeconomic deprivation. Car ownership is operationalized as a binary indicator of deprivation, indicating whether or not the household has exclusive access to at least one car. Housing tenure is another alternative socioeconomic indicator, as it is associated with both income and wealth. Housing tenure is operationalized as a binary indicator of deprivation. A household is considered deprived in the domain of housing tenure if within a government subsidized rental. Otherwise the household lives in a residence that is owner-occupied or privately rented, and the household is not considered deprived in the domain of housing tenure. Finally, household density is a binary indicator of whether or not a household is overcrowded. An overcrowded household is defined as a density of more than one person per room.

For the infant mortality analysis, the individual-level socioeconomic variable is maternal marital status at birth. The socioeconomic variables used in this portion of the analysis are limited due to the reliance on maternal information collected on the birth certificate. Maternal marital status at birth is operationalized as a dichotomous indicator of whether the mother was legally married or unmarried at the time of birth.

Socioeconomic variables: Area-level

For the adult mortality analysis, time-varying area-level socioeconomic variables are also included. Area- level deprivation is operationalized as a population-weighted Carstairs quintile. Carstairs scores are derived by combining selected variables taken from census data. Scores are described as a measure which reflects access to "those goods and services, resources and amenities and of a physical environment which are customary in society" (Carstairs and Morris 1990). The scores are derived separately using 100% census data at the smallest possible cross-sectional geography, the postcode sector. The four census variables used are: Overcrowding, or the proportion of all persons living in private households with a density of more than one person

per room; Male unemployment, or the proportion of economically active males seeking or waiting to start work; Low social class, or the proportion of all persons in private households with an economically active head with head of household in social class IV or V; and no car, or the proportion of all persons in private households which do not own a car. Population-weighting follows the method of Carstairs and Morris (1991), with each variable standardized to have a population-weighted mean of zero and a variance of one using the entire population of England and Wales as the standard reference group.

A measure of rurality was derived by county of usual residence. Each county of usual residence was grouped into one of four categories: London; Other Metropolitan; Non-metropolitan; and Missing. The Metropolitan counties are Greater Manchester, Merseyside, South Yorkshire, Tyne and Wear, West Midlands and West Yorkshire (approximately 1.2 to 2.8 million people). The remaining counties are termed Non-Metropolitan and range in population size from about 100,000 to 1.2 million people. County classifications were harmonized over the study period, according to official geographies of metropolitan and non-metropolitan counties.

Remaining variables

Sex is included as a time-invariant variable, taken from the core dataset for each individual. The census interval is also included for the adult mortality analysis. This time-varying variable indicates the period of observation for each decennial census: [1971-1981); [1981-1991); [1991-2001); [2001-2011); [2011 to end of study period). The end of study period date is December 31, 2013, which corresponds to the most current, complete date of event linkages to the LS data at the time of this analysis.

Two additional variables are derived from the birth certificate and used in the infant mortality analysis: maternal age at birth; and birth weight. The maternal age at birth is categorized to indicate the mother's age at the time of birth as: <20 years; 20-24 years; 25-29 years; 30-34 years; 35-39 years; or 40 and above years. The lowest reported maternal age was 11

years, and the highest was 54 years. Birthweight is included as a dichotomous indicator of whether the singleton birth was low birth weight (less than 2,500 grams) or not (greater than or equal to 2,500 grams).

Methods

First generation mortality: all-cause and behavior-attributable mortality

To estimate all-cause adult mortality of the first generation migrants, survival analysis is estimated with a piecewise parametric model with piecewise constant hazard functions, otherwise known as a piecewise exponential model (Friedman 1982). A piecewise exponential model was chosen for two reasons: (1) event times are precisely measured in the LS data; (2) time-varying covariates may be missing for some age intervals. Age is used as the survival time, with the constant hazard functions estimated over five year age intervals. Ideally, a smaller age interval would be specified, however the software (STATA 7.0) on which the analysis was performed did not have the memory capacity. The ONS usually provides only year of birth due to anonymity regulations of LS members. For this analysis, however, month of birth was also permitted through an extended security and analytic clearance. This extended security clearance required analyses to be remotely run by ONS staff on a secure, internal system; and this system only has STATA 7.0.

To estimate the piecewise exponential models, the data is first expanded by splitting the data according to five year age intervals. This expansion, or episode splitting, creates one row of data for each age interval for which the individual is still at risk of death. Age intervals for which the individual is not under observation (e.g. ages <15 and ≥ 85) are dropped from analysis. As covariates are time-varying according to census period, covariate values are assigned to each appropriate age interval corresponding to the census period during which that age interval falls. For example, an individual's marital status reported on census day 1971 will be constant for all age intervals for that individual from census day 1971 until census day 1981. In cases where an individual reported differing marital statuses across subsequent censuses (e.g. married in 1971

and single in 1981), the covariate value shift was assumed to occur at the midpoint between censuses. This specification applies to all time-varying covariates.

The piecewise exponential model estimates the mortality hazard rate λ_{ij} for observation i in interval j as follows:

$$\ln(\lambda_{ij}) = \alpha_j + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip}$$

where j denotes the age intervals in the expanded data (with the intervals starting at years 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, and 80); x_p indicates the covariates included in the analysis and the corresponding coefficients β_p; the baseline hazard αj denotes the constant hazard within time interval j for the baseline group as $\alpha_j = ln(\lambda_j)$, wherein the covariates take on the value zero ($x_i = 0$), and represents the underlying mortality risk for each age interval.

For behavioral-attributable mortality, competing risk survival analysis was implemented as an extension of the above all-cause analysis. As the cause-specific mortality events (smoking-related mortality, alcohol-related mortality, and all other mortality) are conditionally independent given the covariates, each cause-specific mortality event can be analyzed individually as the sole failure with the other events treated as censored in conjunction with the truly censored observations.

Years of birth and death were non-missing for all individuals. For individuals missing month of birth or death (5.56% of relevant cases), then month was assigned to the middle month of the year, July. The day of birth and death was assigned to 15 for all relevant cases.

Second generation mortality: infant mortality

A similar survival analysis model is used to estimate all-cause infant mortality of the second generation migrants. For each individual, age intervals for which the individual is not under observation (all ages one and above) are dropped from analysis. The data is expanded into twelve 1 month age intervals for the survival analysis. The decision to model infant mortality in 1 month

age intervals was made due to the known infant mortality differences between the neonatal period (deaths prior to the 28th day of life) and the post-neonatal period. Evidence from the United States suggests that a considerably higher number of infant deaths occur in the neonatal period (Alexander et al. 1999). Due to memory constraints of the current analytic software, more precise age intervals could not be performed at this time. Ideally, this analysis will be repeated with 1 day age intervals because of the known concentration of neonatal deaths on the first day (Preston, Heuveline, and Guillot 2000).

The piecewise exponential model estimates the infant mortality hazard rate λ_{ij} for observation *i* in interval *j* as follows:

$$\ln(\lambda_{ij}) = \alpha_j + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_p x_{ip}$$

where *j* denotes the age intervals in the expanded data (with the intervals starting at months 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, and 11); x_p indicates the covariates included in the analysis and the corresponding coefficients β_p; the baseline hazard αj denotes the constant hazard within time interval j for the baseline group as $\alpha_j = ln(\lambda_j)$, wherein the covariates take on the value zero ($x_i = 0$), and represents the underlying mortality risk for each age interval.

Results

First generation preventable mortality: all-cause mortality

The number of LS individuals, observations, person-years at risk, and mortality events are displayed in Table 3.2. The total number of individual LS members included in the analysis is 1,005,652, who collectively contribute 5,919,032 observations for analysis and 26,138,372 person-years at risk. The number of individuals, observations, person-years at risk, deaths, and column percentages by country of birth and mortality status are shown in Table 3.3. Of the 1,005,652 individuals comprising the study sample, 74.01% exit observation alive and 25.99% though death (smoking-related mortality: 13.77%; alcohol-related mortality: 0.39%; all other

mortality: 11.83%). Of the 196,999 observed deaths, 51.91% are due to smoking-related mortality, 1.79% due to alcohol-related mortality; and 46.30% due to all other mortality.

Table 3.4. displays the model results from the all-cause survival analysis. Table 3.4. shows first the reduced model that includes only the primary demographic predictor covariates, which provides the initial estimation of whether a MMA exists for first generation migrants in England and Wales relative to the native population. As expected, the hazard of death is about 21% lower [-21%=100%(0.79-1)] for females relative to males. A mortality advantage does appear to exist for the majority (7 of 10 groups) of first generation immigrants. Relative to those born in England and Wales, the hazard of mortality is significantly lower for those born in Western Europe (32% lower), Africa and the Caribbean (50% lower), South Asia (47% lower), East Asia (61% lower), North America, Australia, and New Zealand (36% lower), South America (48% lower), and the rest of the world (24% lower). However, there is a significantly higher hazard of death for first generation migrants from Scotland (15% higher), other UK countries (33% higher), and Eastern Europe (19% higher). The census interval covariate is included as a check that the pattern of mortality is relatively constant over the study period; the hazard rate of mortality is negligible and essentially constant across census interval periods.

Table 3.4. next shows the full model which includes the individual-level and area-level socioeconomic indicators to assess the extent to which an observed migrant mortality advantage (or disadvantage) is attenuated by socioeconomic covariates. As first generation migrants are generally socioeconomically disadvantaged relative to England and Wales natives, any observed MMA in the reduced model should deepen after controlling for socioeconomic status in the full model. For the majority of the first generation migrant groups displaying a mortality advantage relative to England and Wales, the inclusion of socioeconomic covariates essentially does not change the estimated mortality advantage or deepens it. The inclusion socioeconomic variables deepens the mortality advantage (by between four and twelve percent) for those born in Africa

and the Caribbean, South Asia, South American, and the rest of the world. The inclusion of socioeconomic variables attenuates the MMA observed for those born in North America, Australia, and New Zealand by sixteen percent (15.64%=100%(0.74-0.64)/0.64). For the three first generation migrant groups displaying a mortality disadvantage relative to England and Wales (Scotland, other UK countries, and Eastern Europe), the hazard ratio is strongly and significantly reduced with the inclusion of socioeconomic covariates. The hazard ratio for those born in Scotland relative to those born in England and Wales decreases from 1.15 in the reduced model to 1.01 in the full model, a twelve percent reduction. Similarly, the hazard ratios for those born in other UK countries and Eastern Europe are reduced by twenty-five percent. These findings suggest that the mortality disadvantage observed for these first generation migrant groups is strongly patterned on social class. These three country groups lose any mortality disadvantage (Scotland and other UK) or actually switch to a mortality advantage (Eastern Europe) after the inclusion of socioeconomic variables. In the full model, every first generation migrant group now significantly exhibits a MMA, with the exception of Scotland and other UK who converged to the native England and Wales population.

Except for rurality of area of residence, the hazard of all-cause mortality is significantly higher for those with increased deprivation along each of the individual-level and area-level socioeconomic variables. Relative to the highest household-head social class (professional), the hazard of mortality becomes increasingly higher descending the social classes scale from intermediate (22% higher), skilled (31% higher), partly skilled (51% higher), unskilled (72% higher), to other including armed forced (3.04 times higher). Relative to individuals with private access to a car, those without a car have a significantly higher (2.04 times higher) hazard of mortality. Relative to those living in a privately owned or rented home, the hazard of mortality is significantly higher for those living in government subsidized housing (10% higher). The hazard of mortality is also significantly higher for divorced or widowed individuals (58% higher) relative

to married individuals. The small area indicator of socioeconomic status, the population-weighted Carstairs quintile, retains small but significant coefficients despite the inclusion of individual-level socioeconomic variables. Relative to the those living in the least deprived areas, individuals living in areas deprived according to the Carstairs scale have a three to six percent higher hazard of mortality. Interestingly, those living in an overcrowded household have a significantly lower hazard of mortality (59% lower) relative to those not living in a crowded household. This variable may pattern with country of birth, in particular for those first-generation migrants with an observed MMA.

First generation preventable mortality: behavior-attributable mortality

Table 3.5. shows the model results from the smoking-related mortality analysis. In the reduced model, the hazard of smoking-related mortality is 30% lower for females relative to males. Intriguingly for country of birth groups, the same pattern of relative mortality disadvantage or advantage as seen in the all-cause mortality analysis appears for smoking-related mortality. For those migrant groups with an MMA, the hazard ratios are nearly identical in the smoking-related mortality survival analysis as for the all-cause survival analysis. Relative to those born in England and Wales, the hazard of smoking-related mortality is significantly lower for those born in Western Europe (35% lower), Africa and the Caribbean (55% lower), South Asia (45% lower), East Asia (65% lower), North America, Australia, and New Zealand (38% lower), South America (62% lower), and the rest of the world (28% lower). However for the country of birth groups with a mortality disadvantage, the hazard ratios are higher in the smoking-related mortality survival analysis than for the all-cause survival analysis. Relative to those born in England and Wales, the hazard of smoking-related mortality is significantly higher for those born in Scotland (22% higher), other UK countries (42% higher), and Eastern Europe (31% higher). Qualitatively comparing reduced model relationships, smoking behavior appears to contribute strongly to the observed migrant mortality profiles. Migrants with relatively higher smoking-related mortality

appear to smoke more, whereas migrants with lower smoking-related mortality seem to smoke less.

Table 3.5. next shows the full model which includes the individual-level and area-level socioeconomic indicators. Keeping in line with the findings from the all-cause mortality analysis, the inclusion of socioeconomic covariates essentially does not change the estimated mortality advantage or deepens it for the majority of the first generation migrant groups displaying a mortality advantage relative to England and Wales. The inclusion of socioeconomic variables deepens the mortality advantage (by between five and fourteen percent) for those born in Africa and the Caribbean, South Asia, South American, and the rest of the world. Similarly, the inclusion of socioeconomic variables attenuates the MMA observed for those born in North America, Australia, and New Zealand by nineteen percent. For the three first generation migrant groups displaying a mortality disadvantage relative to England and Wales (Scotland, other UK countries, and Eastern Europe), the hazard ratio is again strongly and significantly reduced with the inclusion of socioeconomic covariates. Relative to those born in England and Wales, the hazard ratio is reduced for those born in: Scotland (13.63%=100%(1.05-1.22)/1.22); other UK countries (28.56%=100%(1.02-1.42)/1.42); and Eastern Europe (27.71%=100%(0.95-1.31)/1.31). These findings suggest that the mortality disadvantage observed for these first generation migrant groups is strongly patterned on social class, which is consistent with widespread evidence that people of lower education or income are more likely to smoke than those without these disadvantages within developed countries (Huisman, Kunst, and Mackenbach 2005). In the full model, nearly all first generation migrant group now significantly exhibits a MMA. The exception is Scotland, which retains a significant 5% higher hazard of mortality relative to England and Wales.

In the full model, the hazard of smoking-related mortality is significantly higher for those with increased deprivation along each of the individual-level and area-level socioeconomic

variables. Relative to the highest household-head social class (professional), the hazard of smoking-related mortality becomes increasingly higher descending the social classes scale from intermediate (27% higher), skilled (49% higher), partly skilled (75% higher), unskilled (2.04 times higher), to other including armed forced (3.26 times higher). Relative to individuals with private access to a car, those without a car have a significantly higher (2.07 times higher) hazard of mortality. Relative to those living in a privately owned or rented home, the hazard of mortality is significantly higher for those living in government subsidized housing (17% higher). The hazard of mortality is also significantly higher for divorced or widowed individuals (47% higher) relative to married individuals. The small area indicator of socioeconomic status, the population-weighted Carstairs quintile, again has small but significant coefficients the expected direction. As with the all-cause mortality analysis, those living in an overcrowded household have a significantly lower hazard of mortality (43% lower) relative to those not living in a crowded household. Furthermore, those living in London have a slightly lower hazard of mortality (2% lower) relative to those living in a rural area. Both of these variables may again pattern with country of birth, especially for those first-generation migrants with an observed MMA. Overall, the pattern of deprivation in the smoking-related morality analysis matches the relationships observed with the all-cause mortality analysis, but to a stronger degree.

Table 3.6. shows the model results from the alcohol-related mortality analysis. Here, the results diverge in interesting patterns from the previous analyses (all-cause and smoking-related mortality). While the same three countries (Scotland, other UK, and Eastern Europe) show a migrant mortality disadvantage, the degree of the disadvantage is much stronger and less attenuated by socioeconomic variables. Relative to those born in England and Wales, the hazard of alcohol-related mortality is significantly higher for those born in from Scotland (87% higher), other UK countries (79% higher), and Eastern Europe (36% higher). In the full model, the mortality disadvantage is only slightly attenuated by the inclusion of socioeconomic variables and

remains strong and significant for Scotland and other UK countries. No other migrant groups have significantly different hazards of alcohol-related mortality relative to those born in England and Wales, in either the reduced or full model. The hazard of alcohol-related mortality generally follows the expected pattern along the socioeconomic scale, which the hazard of alcohol-related mortality significantly higher for those with increased deprivation along each of the individual-level socioeconomic variables. Overall, alcohol-related mortality seems largely relevant for the observed migrant mortality disadvantages.

Table 3.7. shows the model results from the all other cause mortality analysis. Here the results strongly pattern with the all-cause mortality analysis. The hazards of mortality for countries with an observed MMA are relatively unchanged relative to the all-cause mortality analysis. Intriguingly for the countries with an observed migrant mortality disadvantage (Scotland, other UK, and Eastern Europe), the extent of the disadvantage is less in the reduced model and switched to a significant MMA in the full model. Theoretically, only mortality not related to smoking or alcohol is considered here. As such, these results imply that the majority of the observed migrant mortality disadvantage in the all-cause analysis (for Scotland, other UK, and Eastern Europe) is due to the behavioral-driven mortality. Thus, accounting for behavioral-driven mortality and socioeconomic indicators attenuates any observed migrant mortality disadvantage.

Second generation mortality: infant mortality

The number of second generation LS individuals, person-years at risk, and infant deaths are displayed in Table 3.8. The total number of individual LS members included in the analysis is 302,738, who collectively contribute 6,358,333 person-years at risk with 1,330 observed infant deaths. Again, LS members enter this portion of the analysis through registration at birth. As such, the second generation analytic sample is primarily composed of individuals born in England and Wales (80.52%). The largest second generation migrant groups represented are Western

Europe (2.84%) and South Asia (4.78%). The majority of individuals were not low birth weight (95.16%) and born to married (71.18%) mothers. The maternal age at first birth follows the expected pattern, with a normal distribution of mothers age at birth concentrated between ages 20 and 34 years.

Table 3.9. next shows the results of the piecewise exponential model. The reduced model suggests that the majority (seven of eleven groups) of second generation migrant groups have a MMA in terms of infant mortality. However, this relationship is only statistically significant for Eastern Europe. Relative to babies born to English and Welsh mothers, the hazard of infant mortality is 50% lower for babies born to Eastern European mothers. Babies born to mothers from South Asia or Central American and Caribbean mothers display a statistically significant higher hazard of infant mortality. Relative to babies born to English and Welsh mothers, babies born to Central American and Caribbean mothers have a 2.33 times higher hazard of infant mortality. Similarly, babies that are born to South Asian mothers have a 1.32 times higher hazard of infant mortality.

Table 3.9. next includes both the individual-level variables and maternal socioeconomic indicators. As expected, low birth weight babies have a significantly higher hazard of mortality (9.25 times higher) relative to babies who are normal weight. In the full model, the MMA observed for babies born to Eastern European mothers is slightly attended (from 0.50 to 0.57) but remains statistically significant. After the inclusion of individual-level variables and maternal socioeconomic indicators, a statistically significant MMA emerges for East and South African second generation migrants. Relative babies born to mothers from England and Wales, the hazard of infant mortality is 45% lower for babies born to East and South African mothers. The migrant mortality disadvantage deepens for babies born Central American and Caribbean mothers, suggesting that the mortality disadvantage is not driven by socioeconomic differences. Alternately, the mortality disadvantage is fully attenuated for babies born to South Asian mothers

after the inclusion of socioeconomic indicators, indicating their higher infant mortality is attributable to social class. The remaining second generation groups do not have significantly different hazards of infant mortality relative to babies born to mothers from England and Wales.

Interestingly, babies that are born to unmarried mothers have a lower hazard of mortality (32% lower) relative to unmarried mothers. This finding may be patterned on low birth weight. Among low birth weight babies, 60% were born to married mothers and 40% to unmarried mothers. Since low birth weight babies have a 9.25 times higher hazard of mortality relative to normal weight babies, the unusual coefficient on maternal marital status may be driven by low birth weight babies born to married mothers.

Discussion

Several limitations should be mentioned regarding observed migrant mortality. Fist, migrant mortality estimates may be more predisposed to data artefact biases, particularly the denominator bias, as compared to the native England and Wales population. However, this analysis drops untraced individuals from analysis. Therefore, all individuals included in the analytic sample are theoretically accounted for by either a registered death or emigration. A previous analysis using LS data also concluded that the denominator bias could not explain the lower mortality rates observed among migrant groups (Wallace and Kulu 2014a). As such, the results reported in this paper should be exempt from potential data artefact effects. A final, unavoidable limitation is the inability to compare migrant mortality to the country of origin reference population. As such, this analysis cannot directly quantify sending health selection.

This paper finds evidence of a strong migrant advantage for all-cause mortality among the majority of first generation migrant groups in England and Wales. As first generation migrants are generally socioeconomically disadvantaged relative to England and Wales natives, a MMA should deepen after controlling for socioeconomic status. This analysis does find that observed MMAs strengthen with the inclusion of socioeconomic indicators. Previous research on

first generation migrant mortality (all-cause) in the UK and England and Wales agree with these results (Wallace and Kulu 2014b, a). First generation migrants from Scotland, other United Kingdom countries, and Eastern Europe are the only groups to display an adult all-cause mortality disadvantage relative to England and Wales. This finding agrees with previous literature investigating United Kingdom immigrants resident in England and Wales. Other studies (Harding and Rosato 1999, Harding et al. 2008, Harding et al. 2009, Wild et al. 2007, Wallace and Kulu 2014a, Wallace 2015) have documented higher all-cause mortality among first generation migrants from Scotland and Northern Ireland compared to England and Wales natives. For all first generation migrant groups, the all-cause mortality differences are fully attenuated (for observed migrant mortality disadvantages) or deepened (for MMAs) after accounting for socioeconomic variables.

Health selection likely does not influence Scottish and other United Kingdom migrants. As such, the lack of a MMA for Scotland and other United Kingdom migrants is not surprising. However, the strong first generation mortality disadvantage for Scotland and other United Kingdom countries is notable and seems largely related to health behaviors within this analysis. Previous evidence supports this conclusion. First generation migrants from Scotland and Northern Ireland have been found to have higher death rates from coronary heart disease (Harding et al. 2008) and cancers related to smoking and alcohol use (Harding and Rosato 1999, Harding et al. 2009). Furthermore, smoking-attributable and alcohol-attributable mortality is significantly higher in Scotland and Northern Ireland relative to England and Wales. Scotland has the highest reported smoking-attributable (Kelly and Preston 2016) and alcohol-related (Goodwin 2015) mortality in the United Kingdom, followed by Northern Ireland. These health behaviors can be conceptualized as ‘proximate causes’ of mortality, whereby deprivation leads to worse health behaviors and ultimately worse behavioral-driven mortality. In European populations, research has indeed found that individuals of lower socioeconomic status are more likely to smoke than

those of higher status (Huisman, Kunst, and Mackenbach 2005), a pattern also observed in Scotland (Bromley, Sproston, and Shelton 2003). In England and Wales, a "clear association between alcohol-related mortality and socioeconomic deprivation" has been documented, with higher alcohol-related mortality occurring among more deprived relative to less deprived individuals (Erskine et al. 2010).

However, the behavioral mortality disadvantage is reduced but not eliminated after accounting for socioeconomic variables in this analysis. In fact, Scotland and Northern Ireland have reduced their socioeconomic disadvantage relative to England and Wales in the past few decades; and area-level socioeconomic indicators no longer explain the majority of sub-national mortality differences (between Scotland and England and Wales) (Hanlon et al. 2005). In other words, achievements in area based initiatives and other policy-driven reductions in area-level socioeconomic inequalities have not eradicated mortality inequalities within the UK. Therefore, Scottish and other United Kingdom migrants seem to maintain their mortality disadvantage upon arrival in England and Wales primarily due to the retention of relatively poor health behaviors, irrespective of socioeconomic status.

The reasons why poorer health behaviors exist in Scotland and other UK countries are unclear. A possible explanation for this phenomenon is dominance of cultural norms that emphasize smoking and drinking. Additionally, the type of alcohol preferred in different populations (e.g. scotch and whiskey) may disproportionately affect mortality from alcohol-attributable causes (Kerr, Fillmore, and Mary 2000) than for deaths from beer or wine. When behavioral-mortality is removed, and only all other cause mortality is considered, all first generation migrant groups exhibit a MMA relative to England and Wales.

This analysis found some evidence of a second generation migrant mortality advantage. While there is no comparable research in the United Kingdom, studies from the United States find infant mortality differences by ethnicity and migrant status. In the United States, babies born to

Mexican-origin mothers have lower infant mortality than babies born to native non-Hispanic white mothers (Hummer et al. 2007, Powers 2013). Alternately, babies born to non-Hispanic black mothers have persistency higher infant mortality (Rossen et al. 2014) than any other ethnic group. This analysis found a migrant mortality advantage for most second generation migrant groups, with the strong exception being Central American and Caribbean second generation migrants. Generally, these migrants have ancestral origins in West Africa, with historic resettlement to British Caribbean colonies throughout the slave trade period of the 16th and 19th centuries. A shared history may underlie the infant mortality disadvantage observed among babies born to non-Hispanic black mothers in the United States and Caribbean mothers in England and Wales. According to the NHS equity mission, immigrants should immediately benefit from universal healthcare in the United Kingdom; and thus the infants of immigrants should benefit from the same obstetrics and gynecological treatment as England and Wales natives. The large variation in infant mortality among second generation migrants highlights the need to reconsider health equity in terms of ethnic and migrant minorities and to identify the underlying causes of mortality variation.

This paper identifies significant mortality variation among subgroups within England and Wales and generally finds a mortality advantage among first and second generation migrants from non-UK countries. This wide heterogeneity in mortality outcomes by country of birth highlights the need to identify and consider migrant groups within the context of healthcare equity. Improved quantification of the migrant population size and health status is critical towards this equity goal. Targeted public health initiatives must meet the clear need for enhanced administrative data (in terms of availability, coverage, and quality) on migrants. Consciously structured policies could expand integration of migrants into the welfare state and ultimately improve the population health of an increasingly diverse nation.

References

Abraido-Lanza, A.F., B.P. Dohrenwend, D.S. Ng-Mak, and J.B. Turner. 1999. "The Latino mortality paradox: a test of the" salmon bias" and healthy migrant hypotheses." *American Journal of Public Health* 89 (10):1543-1548.

Alba, R. and T. Islam. 2009. "The case of the disappearing Mexican Americans: An ethnic-identity mystery." *Population Research and Policy Review* 28(2): 109-121.

Alexander, G.R., M.E. Tompkins, M.C. Allen, and T.C. Hulsey. 1999. "Trends and racial differences in birth weight and related survival." *Maternal & Child Health Journal* 3 (2):71-9.

Anson, J. 2004. "The migrant mortality advantage: a 70 month follow-up of the Brussels population." *European Journal of Population* 20 (3):191-218.

Baburin, A., T. Lai, and M. Leinsalu. 2011. "Avoidable mortality in Estonia: Exploring the differences in life expectancy between Estonians and non-Estonians in 2005-2007."*Public Health* 125: 754-762.

Blue, L., and A. Fenelon. 2011. "Explaining low mortality among US immigrants relative to native-born Americans: the role of smoking." *International Journal of Epidemiology* 40 (3):786-793.

Boulogne, R., E. Jougla, Y. Breem, A.E. Kunst, and G. Rey. 2012. "Mortality differences between the foreign-born and locally-born population in France (2004–2007)." *Social Science & Medicine* 74 (8):1213-1223.

Bradshaw, G. and P.L. Bradshaw. 1995. "The equity debate within the British National Health Service." *Journal of Nursing Management* 3(4): 161-168.

Bromley, C., K. Sproston, and N. Shelton. 2003. "The Scottish Health Survey 2003. Volume 2. Adults." Accessed November 2016 at: www.scotland.gov.uk.

Carstairs, V., and R. Morris. 1990. "Deprivation and health in Scotland." *Health Bulletin* 48 (4):162-175.

Cozier, Y.C., J.R. Palmer, N.J. Horton, L. Fredman, L.A. Wise, and L. Rosenberg. 2007. "Relation between neighborhood median housing value and hypertension risk among black women in the United States." *American Journal of Public Health* 97 (4):718-724.

Collins, E. and R. Klein. 1980. "Equity and the NHS: self-reported morbidity, access, and primary care." *BMJ* 281(6248): 1111-1115.

Crimmins, E.M., B.J. Soldo, J.K. Kim, and D.E. Alley. 2005. "Using anthropometric indicators for Mexicans in the United States and Mexico to understand the selection of migrants and the "Hispanic paradox"." *Social Biology* 52 (3-4):164-177.

Deboosere, P., and S. Gadeyne. 2005. "Adult migrant mortality advantage in Belgium: evidence using census and register data." *Population* 655-698.

Eastwood, P. 2012. “Statistics on Smoking: England, 2012.” Leeds (UK): The Health and Social Care Information Centre, National Health Service.

Edwards, R., J. Franklin, and J. Holland. 2003. “Families and social capital: Exploring the issues.” University of London (UK).

Eikemo, T., V. Skalická, and M. Avendano. 2009. "Variations in relative health inequalities: are they a mathematical artefact?" *International Journal for Equity in Health* 8 (1):1-5.

Elo, I.T., and S.H. Preston. 1997. "Racial and ethnic differences in mortality at older ages." Committee on Population, National Research Council. Washington (DC): National Academies Press (US).

Elo, I.T., C.M. Turra, B. Kestenbaum, and B.R. Ferguson. 2004. "Mortality among elderly Hispanics in the United States: Past evidence and new results." *Demography* 41 (1):109-128.

Erskine, S., R. Maheswaran, T. Pearson, and D. Gleeson. 2010. "Socioeconomic deprivation, urban-rural location and alcohol-related mortality in England and Wales." *BMC Public Health* 10(1): 1.

Esping-Andersen, G. 2013. The three worlds of welfare capitalism: John Wiley & Sons.
Feinstein, Jonathan S. 1993. "The relationship between socioeconomic status and health: a review of the literature." *The Milbank Quarterly* 279-322.

Frayer, L. 2016. “Brexit Vote Leads To Increased Anti-Immigrant Sentiment In The UK.” National Public Radio. Accessed July 2016 at: http://www.npr.org/2016/07/05/484832444/brexit-vote-leads-to-increased-anti-immigrant-sentiment-in-the-uk.

Friedman, M. 1982. "Piecewise exponential models for survival data with covariates." *The Annals of Statistics* 101-113.

Glendinning, A., J.G. Love, L.B. Hendry, and J. Shucksmith. 1992. "Adolescence and health inequalities: extensions to Macintyre and West." *Social Science & Medicine* 35 (5):679-687.

G., Nina. 2015. “Alcohol Related Deaths in the United Kingdom: Registered 2013.” Mortality Analysis Team: UK Statistics Authority. Accessed November 2016 at: www.ons.gov.uk/.

Hanlon, P., R.S. Lawder, D. Buchanan, A. Redpath, D. Walsh, R. Wood, M. Bain, D.H. Brewster, J. Chalmers, and D. Walsh. 2005. "Why is mortality higher in Scotland than in England and Wales? Decreasing influence of socioeconomic deprivation between 1981 and 2001 supports the existence of a ‘Scottish Effect’." *Journal of Public Health* 27 (2):199-204.

Harding, S. and R. Balarajan. 1996. "Patterns of mortality in second generation Irish living in England and Wales: longitudinal study." *BMJ* 312(7043): 1389-1392.

Harding, S., and M. Rosato. 1999. "Cancer incidence among first generation Scottish, Irish, West Indian and South Asian migrants living in England and Wales." *Ethnicity and Health* 4(1-2): 83-92.

Harding, S., M. Rosato, and A. Teyhan. 2008. "Trends for coronary heart disease and stroke mortality among migrants in England and Wales, 1979–2003: slow declines notable for some groups." *Heart* 94(4): 463-470.

Harding, S., M. Rosato, and A. Teyhan. 2009. "Trends in cancer mortality among migrants in England and Wales, 1979–2003." *European Journal of Cancer* 45(12): 2168-2179.

Huisman, M., A.E. Kunst, and J.P. Mackenbach. 2005. "Educational inequalities in smoking among men and women aged 16 years and older in 11 European countries." *Tobacco Control* 14 (2):106-13.

Hummer, R.A., D.A. Powers, S.G. Pullum, G.L. Gossman, and W.P. Frisbie. 2007. "Paradox found (again): infant mortality among the Mexican-origin population in the United States." *Demography* 44 (3):441-457.

Hutchison, A., M. Tobias, J. Glover, C. Wright, D. Hetzel, and E. Fisher. 2006. "Australian and New Zealand atlas of avoidable mortality." Public Health Information Unit, University of Adelaide. Accessed November 2016 at: www.moh.govt.nz/notebook/nbbooks.nsf/0/2B8E4CB234C9BFE7CC257211006C4D10.

Jaeger, D.A., T. Dohmen, A. Falk, D. Huffman, U. Sunde, and H. Bonin. 2010. "Direct evidence on risk attitudes and migration." *The Review of Economics and Statistics* 92 (3):684-689.

Jasso, G.., D.S. Massey, M.R. Rosenzweig, J.P. Smith. 2004. "Immigrant Health: Selectivity and Acculturation." In Critical Perspectives on Racial and Ethnic Differences in Health in Late Life, edited by Norman B. Anderson, Rodolfo A. Bulatao and Barney. Cohen. Washington (DC): National Academies Press.

Jokela, M. 2009. "Personality predicts migration within and between US states." *Journal of Research in Personality* 43 (1):79-83.

Kautto, M., J. Fritzell, B. Hvinden, J. Kvist, and H. Uusitalo. 2001. "Nordic welfare states in the European context." Psychology Press: Routledge.

Kelly, L.A., and S.H. Preston. 2016. "The contribution of a history of heavy smoking to Scotland's mortality disadvantage." *Population Studies* 70 (1):59-71.

Kerr, W.C., K.M. Fillmore, and P. Marvy. 2000. "Beverage-specific alcohol consumption and cirrhosis mortality in a group of English-speaking beer-drinking countries." *Addiction* 95 (3):339-346.

Khlat, M., and Y. Courbage. 1996. "Mortality and causes of death of Moroccans in France, 1979-91." *Population* 59-94.

Kibele, E., R. Scholz, and V.M. Shkolnikov. 2008. "Low migrant mortality in Germany for men aged 65 and older: fact or artifact?" *European Journal of Epidemiology* 23 (6):389-393.

King, G., C.J.L. Murray, J.A. Salomon, and A. Tandon. 2004. "Enhancing the validity and cross-cultural comparability of measurement in survey research." *American Political Science Review* 98(1):191-207.

Kroger, H., E. Pakpahan, and R. Hoffmann. 2015. "What causes health inequality? A systematic review on the relative importance of social causation and health selection." *European Journal of Public Health* 47(4): 339-354.

Lawson, J.S., and D. Black. 1993. "Socioeconomic status: the prime indicator of premature death in Australia." *Journal of Biosocial Science* 25 (4):539-52.

Macintyre, S., and P. West. 1991. "Lack of class variation in health in adolescence: an artefact of an occupational measure of social class?" *Social Science & Medicine* 32 (4):395-402.

Mackenbach, J.P. 2012. "The persistence of health inequalities in modern welfare states: the explanation of a paradox." *Social Science & Medicine* 75 (4):761-769.

Mackenbach, J.P., V. Bos, O. Andersen, M. Cardano, G. Costa, S. Harding, A. Reid, O. Hemström, T. Valkonen, and A.E. Kunst. 2003. "Widening socioeconomic inequalities in mortality in six Western European countries." *International Journal of Epidemiology* 32 (5):830-837.

Mackenbach, J.P., and A.E. Kunst. 1997. "Measuring the magnitude of socio-economic inequalities in health: an overview of available measures illustrated with two examples from Europe." *Social Science & Medicine* 44 (6):757-771
.
Mackenbach, J.P., I. Stirbu, A.R. Roskam, M.M. Schaap, G. Menvielle, M. Leinsalu, and A.E. Kunst. 2008. "Socioeconomic inequalities in health in 22 European countries." *New England Journal of Medicine* 358 (23):2468-2481.

Marmot, M.G., A.M. Adelstein. and L. Bulusu. 1984. "Immigrant mortality in England and Wales 1970-78: causes of death by country of birth." *Studies of Medical and Population Subjects* 47.

Martikainen, P., T.M. Kauppinen, and T. Valkonen. 2003. "Effects of the characteristics of neighbourhoods and the characteristics of people on cause specific mortality: a register based follow up study of 252,000 men." *Journal of Epidemiology and Community Health* 57 (3):210-217.

McLaren, E. 2014. "Childhood, Infant and Perinatal Mortality in England and Wales: 2012." Vital Statistics Outputs Branch: UK Statistics Authority. Accessed November 2016 at: www.ons.gov.uk/.

McPherson, M., L. Smith-Lovin, and J.M. Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27:415-444.

OECD. 2016. "Foreign-born population (indicator)." doi: 10.1787/5a368e1b-en. Accessed July 2016 at: https://data.oecd.org/migration/foreign-born-population.htm.

Palloni, A., and E. Arias. 2004. "Paradox lost: explaining the Hispanic adult mortality advantage." *Demography* 41 (3):385-415.

Pappas, G., S. Queen, W. Hadden, and G. Fisher. 1993. "The increasing disparity in mortality between socioeconomic groups in the United States, 1960 and 1986." *New England Journal of Medicine* 329 (2):103-109.

Parkinson, M.. 1998. "Combating social exclusion: lessons from area-based programmes in Europe." Policy Press Bristol (UK).

Portes, A. 2000. "The Two Meanings of Social Capital." Sociological Forum 15 (1):1-12.

Powers, D.A. 2013. "Paradox revisited: a further investigation of racial/ethnic differences in infant mortality by maternal age." *Demography* 50(2): 495-520.

Preston, S.H., P. Heuveline, and M. Guillot. 2000. "Demography: Measuring and Modeling Population Processes." Wiley-Blackwell (US).

Razum, O., H. Zeeb, H.S.Akgün, and S. Yilmaz. 1998. "Low overall mortality of Turkish residents in Germany persists and extends into a second generation: merely a healthy migrant effect?" *Tropical Medicine & International Health* 3 (4):297-303.

Rechel, B., P. Mladovsky, D. Ingleby, J.P. Mackenbach, and M. McKee. 2013. "Migration and health in an increasingly diverse Europe." *Lancet* 381 (9873):1235-1245.

Ronellenfitsch, U., C. Kyobutungi, H. Becher, and O. Razum. 2006. "All-cause and cardiovascular mortality among ethnic German immigrants from the Former Soviet Union: a cohort study." *BMC Public Health* 6 (1):1.

Rossen, L.M. and K.C. Schoendorf. 2014. "Trends in racial and ethnic disparities in infant mortality rates in the United States, 1989–2006." *American Journal of Public Health* 104(8): 1549-1556.

Ruiz, J.M., P. Steffen, and T.B. Smith. 2013. "Hispanic mortality paradox: a systematic review and meta-analysis of the longitudinal literature." *American Journal of Public Health* 103 (3):e52-e60.

Sampson, R.J., J.D. Morenoff, and T. Gannon-Rowley. 2002. "Assessing" neighborhood effects": Social processes and new directions in research." *Annual Review of Sociology* 443-478.

Scanlan, J.P. 2014. "Race and mortality revisited." *Society* 51 (4):328-346.

Silventoinen, K., N. Hammar, E. Hedlund, M. Koskenvuo, T. Rönnemaa, and J. Kaprio. 2008. "Selective international migration by social position, health behaviour and personality." *European Journal of Public Health* 18 (2):150-155.

Singh, G.K., and R.A. Hiatt. 2006. "Trends and disparities in socioeconomic and behavioural characteristics, life expectancy, and cause-specific mortality of native-born and foreign-born

populations in the United States, 1979–2003." *International Journal of Eepidemiology* 35 (4):903-919.

Smith, G.D., C. Hart, G. Watt, D. Hole, and V. Hawthorne. 1998. "Individual social class, area-based deprivation, cardiovascular disease risk factors, and mortality: the Renfrew and Paisley Study." *Journal of Epidemiology and Community Health* 52 (6):399-405.

Stern, J. 1983. "Social mobility and the interpretation of social class mortality differentials." *Journal of Social Policy* 12 (01):27-49.

Stirbu, I., A.E. Kunst, V. Bos, and J.P. Mackenbach. 2006. "Differences in avoidable mortality between migrants and the native Dutch in The Netherlands." *BMC Public Health* 6 (1):1.

Strand, B.H., O.A. Steingrímsdóttir, E. Grøholt, I. Ariansen, S. Graff-Iversen, and O. Næss. 2014. "Trends in educational inequalities in cause specific mortality in Norway from 1960 to 2010: a turning point for educational inequalities in cause specific mortality of Norwegian men after the millennium?" *BMC Public Health* 14 (1):1208.

Stronks, K., I. Kulu-Glasgow, and C. Agyemang. 2009. "The utility of 'country of birth' for the classification of ethnic groups in health research: the Dutch experience." *Ethnicity & Health* 14(3): 255-269.

Thomson, H. 2008. "A dose of realism for healthy urban policy: lessons from area-based initiatives in the UK." *Journal of Epidemiology and Community Health* 62 (10):932-936.

Tilford, S. (2016). "Britain, immigration and Brexit." *CER Bulletin* 30: 64-162.

Vagerö, D., and O. Lundberg. 1989. "Health inequalities in Britain and Sweden." *Lancet* 334 (8653):35-36.

Wadsworth, M.E.J. 1997. "Health inequalities in the life course perspective." *Social Science & Medicine* 44 (6):859-869.

Wallace, M. 2016. "Adult mortality among the descendants of immigrants in England and Wales: does a migrant mortality advantage persist beyond the first generation?" *Journal of Ethnic and Migration Studies* 1-19.

Wallace, M., and H. Kulu. 2014a. "Low immigrant mortality in England and Wales: A data artefact?" *Social Science & Medicine* 120:100-109.

Wallace, M., and H. Kulu. 2014b. "Migration and health in England and Scotland: a study of migrant selectivity and salmon bias." *Population, Space and Place* 20 (8):694-708.

Wallace, M. and H. Kulu. 2015. "Mortality among immigrants in England and Wales by major causes of death, 1971–2012: A longitudinal analysis of register-based data." *Social Science & Medicine* 147: 209-221.

Weitoft, G.R., A. Gullberg, A. Hjern, and M. Rosén. 1999. "Mortality statistics in immigrant research: method for adjusting underestimation of mortality." *International Journal of Epidemiology* 28 (4):756-763.

West, P. 1988. "Inequalities? Social class differentials in health in British youth." ." *Social Science & Medicine* 27 (4):291-296.

Westerling, R., and M. Rosén. 2002. "'Avoidable mortality among immigrants in Sweden." *European Journal of Public Health* 12 (4):279-286.

Wheller, L., A. Baker, C. Griffiths, and C. Rooney. 2007. "Trends in avoidable mortality in England and Wales, 1993-2005." *Health Statistics Quarterly* (34):6.

Wild, S.H., C. Fischbacher, A. Brock, C. Griffiths, and R. Bhopal. 2007. "Mortality from all causes and circulatory disease by country of birth in England and Wales 2001–2003." *Journal of Public Health* 29(2): 191-198.

Wright, R.. 2015. "National Life Tables, United Kingdom: 2012–2014." Demographic Analysis Unit: UK Statistics Authority. Accessed November 2016 at: www.ons.gov.uk/.

Zimmermann, K.F. 1995. "Tackling the European migration problem." *The Journal of Economic Perspectives* 9(2): 45-62.

Zimmermann, K.F. 1996. "European migration: Push and pull." *International Regional Science Review* 19(1-2): 95-128.

Table 3.1.A. Smoking-related diagnoses and ICD coding.

Diagnosis	**ICD-9**	**ICD-10**
Mouth and oropharynx cancer	140-149	C00-C14
Trachea, bronchus and lung cancers	162	C33-C34
Malignant neoplasm of larynx	161	C32
Esophagus cancer	150	C15
Stomach Cancer	151	C16
Pancreas Cancer	157	C25
Cervix Uteri	180	C53
Malignant neoplasms of urinary tract	188-189	C64, C689
Malignant neoplasm without specification of site	199	C80
Myeloid leukemia	205	C92
Ischemic heart disease	410-414	I20-I25
Cerebrovascular disease	430-438	I60-I69
Other Heart Disease	390-398, 415-429	I00-I09, I26-I51
Atherosclerosis	440	I70
Aortic Aneurysm	441	I71
Other Arterial Diseases	442-448	I72-I78
Chronic obstructive pulmonary disease	490-492, 495-496	J40-J44
Asthma	493	J45-J46
Pneumonia, Influenza	481-488	J10-J18
Stomach ulcer, Duodenal ulcer	531-533	K25-K27

Table 3.1.B. Alcohol-related diagnoses and ICD coding.

Diagnosis	**ICD-9**	**ICD-10**
Alcohol-induced pseudo-Cushing's syndrome	255.0	E24.4
Mental and behavioral disorders due to alcohol use	291, 303, 305.0	F10
Degeneration of nervous system due to alcohol	331.7	G31.2
Alcoholic polyneuropathy	357.5	G62.1
Alcoholic myopathy	358.4	G72.1
Alcoholic cardiomyopathy	425.5	I42.6
Alcoholic gastritis	535.31	K29.2
Alcoholic liver disease and cirrhosis of the liver	571	K70, K74
Chronic hepatitis		K73
Alcohol-induced chronic pancreatitis	577.1	K86.0
Excessive blood level of alcohol	790.3	R78.0
Unintentional injuries	E800-949	V01-X59, Y40-Y86, Y88, Y89
Self-inflicted injuries	E950-959	X60-X84, Y870
Violence	E960-969	X85-Y09, Y871
Toxic effect of alcohol	980	T51
Accidental alcohol poisoning not classified	E860	T51

Table 3.2. Number of England and Wales natives and first generation migrants, person-years at risk, deaths, and column percentages by covariates, England and Wales 1971-2013.

Census Interval	Individuals	%	PY at Risk	%	Deaths	%
[1971-1981)	258,756	25.95	6,325,886	24.29	62,430	31.86
[1981-1991)	165,301	16.58	5,930,950	22.78	50,196	25.61
[1991-2001)	196,305	19.68	6,542,490	25.12	44,758	22.84
[2001-2011)	280,911	28.17	6,693,179	25.70	36,100	18.42
2011+	95,999	9.63	547,924	2.10	2,490	1.27
Sex						
Male	495,560	49.28	12,800,000	49.04	107,542	54.59
Female	510,106	50.72	13,300,000	50.96	89,461	45.41
Country of Birth						
England & Wales	612,426	60.90	15,800,000	60.46	157,223	79.81
Scotland	9,812	0.98	288,796	1.11	3,342	1.70
Other UK	10,755	1.07	316,642	1.21	4,281	2.17
Western Europe	6,074	0.60	163,933	0.63	1,060	0.54
Eastern Europe	8,018	0.80	196,351	0.75	2,334	1.18
Africa & Caribbean	10,024	1.00	301,361	1.15	1,447	0.73
South Asia	15,620	1.55	473,321	1.81	2,407	1.22
East Asia	3,263	0.32	91,056	0.35	326	0.17
N. America, Aus. & N.Z.	3,272	0.33	87,585	0.34	537	0.27
South America	706	0.07	20,960	0.08	103	0.05
Other, Rest of World	619	0.06	17,058	0.07	143	0.07
Missing	325,077	32.32	8,377,120	32.05	23,800	12.08
Marital Status						
Married	326,396	32.46	9,889,533	37.86	113,110	57.42
Single	291,492	28.98	6,367,188	24.37	18,899	9.59
Divorced/Widowed	66,172	6.58	1,561,048	5.98	41,904	21.27
Missing	321,606	31.98	8,306,650	31.80	23,090	11.72
Car Deprivation						
Not deprived	482,970	48.02	13,100,000	50.24	89,400	45.38
Deprived	200,063	19.89	4,637,870	17.79	84,184	42.73
Missing	322,633	32.08	8,338,947	31.98	23,419	11.89
Housing Tenure Deprivation						
Not deprived	475,447	47.28	12,500,000	47.86	103,829	52.70
Deprived	160,904	16.00	4,045,962	15.49	52,586	26.69
Missing	369,315	36.72	9,571,838	36.65	40,588	20.60
Overcrowding Deprivation						
Not deprived	628,558	62.50	16,400,000	62.70	168,407	85.48
Deprived	53,755	5.35	1,393,311	5.33	5,048	2.56
Missing	323,353	32.15	8,363,856	31.98	23,548	11.95

Household-head Social Class						
I Professional	34,137	3.39	937,273	3.59	4,630	2.35
II Intermediate	161,906	16.10	4,252,740	16.28	26,353	13.38
III Skilled	264,627	26.31	7,100,060	27.18	56,340	28.60
IV Partly Skilled	100,019	9.95	2,561,169	9.80	25,661	13.03
V Unskilled	35,567	3.54	888,414	3.40	11,561	5.87
Other/Armed Forces	24,818	2.47	464,725	1.78	11,715	5.95
Missing & NA	384,592	38.24	9,920,040	37.97	60,743	30.83
Carstairs Quintile						
Most deprived	131,974	19.29	3,550,718	19.93	29,242	16.81
Moderately	134,797	19.71	3,577,250	20.08	32,877	18.90
Midpoint	136,180	19.91	3,539,073	19.86	35,131	20.20
Next least deprived	137,295	20.07	3,529,681	19.81	37,404	21.51
Least deprived	139,763	20.43	3,510,238	19.70	38,644	22.22
Missing	4,054	0.59	110,820	0.62	616	0.35
Rurality						
Non-metropolitan	431,773	42.93	11,300,000	43.21	108,650	55.15
Other metropolitan	155,044	15.42	4,018,623	15.37	42,552	21.60
London	96,781	9.62	2,505,956	9.58	22,624	11.48
Missing	322,068	32.03	8,325,545	31.84	23,177	11.76

SOURCE: Authors calculations based on ONS LS.

Table 3.3. Number of England and Wales natives and first generation migrants, person-years at risk, deaths, and column percentages by country of birth and mortality status, England and Wales 1971-2013.

Country of Birth	Individuals	%	PY at Risk	%	Deaths	%
Alive						
England & Wales	406,187	54.57	11,700,000	55.99	0	N/A
Scotland	5,716	0.77	201,577	0.96	0	N/A
Other UK	5,662	0.76	207,998	1.00	0	N/A
Western Europe	4,599	0.62	132,076	0.63	0	N/A
Eastern Europe	5,125	0.69	134,810	0.65	0	N/A
Africa & Caribbean	8,556	1.15	263,592	1.26	0	N/A
South Asia	13,229	1.78	411,939	1.97	0	N/A
East Asia	2,945	0.40	83,420	0.40	0	N/A
N. America, Aus. & NZ.	2,575	0.35	73,113	0.35	0	N/A
South America	576	0.08	18,066	0.09	0	N/A
Other, Rest of World	325	0.04	12,112	0.06	0	N/A
Missing	288,816	38.80	7,658,566	36.65	0	N/A
Smoking-related mortality						
England & Wales	110,226	79.61	2,462,806	78.86	81,598	79.79
Scotland	2,247	1.62	52,333	1.68	1,844	1.80
Other UK	2,747	1.98	65,653	2.10	2,296	2.25
Western Europe	772	0.56	18,537	0.59	525	0.51
Eastern Europe	1,657	1.20	38,496	1.23	1,312	1.28
Africa & Caribbean	718	0.52	19,634	0.63	705	0.69
South Asia	1,335	0.96	36,334	1.16	1,355	1.33
East Asia	166	0.12	4,458	0.14	167	0.16
N. America, Aus. & NZ.	374	0.27	8,519	0.27	276	0.27
South America	57	0.04	1,385	0.04	40	0.04
Other, Rest of World	127	0.09	2,728	0.09	46	0.04
Missing	18,030	13.02	412,153	13.20	12,096	11.83
Alcohol-related mortality						
England & Wales	2,705	69.38	75,733	70.10	2351	66.32
Scotland	83	2.13	2,374	2.20	80	2.26
Other UK	84	2.15	2,253	2.09	73	2.06
Western Europe	23	0.59	526	0.49	21	0.59
Eastern Europe	36	0.92	1,007	0.93	39	1.10
Africa & Caribbean	44	1.13	1,170	1.08	45	1.27
South Asia	63	1.62	1,854	1.72	87	2.45
East Asia	<10	<0.13	222	0.21	12	0.34
N. America, Aus. & NZ.	11	0.28	332	0.31	16	0.45
South America	<10	<0.13	<10	<0.00	<10	<0.14
Other, Rest of World	<10	<0.13	<10	<0.00	<10	<0.14
Missing	835	21.42	22,550	20.87	811	22.88

Other mortality						
England & Wales	93,308	78.41	1,565,831	77.90	73,274	80.34
Scotland	1,766	1.48	32,512	1.62	1,418	1.55
Other UK	2,262	1.90	40,737	2.03	1,912	2.10
Western Europe	680	0.57	12,793	0.64	514	0.56
Eastern Europe	1,200	1.01	22,037	1.10	983	1.08
Africa & Caribbean	706	0.59	16,966	0.84	697	0.76
South Asia	993	0.83	23,193	1.15	965	1.06
East Asia	145	0.12	2,957	0.15	147	0.16
N. America, Aus. & NZ.	312	0.26	5,621	0.28	245	0.27
South America	71	0.06	1,426	0.07	61	0.07
Other, Rest of World	162	0.14	2,122	0.11	95	0.10
Missing	17,396	14.62	283,850	14.12	10,893	11.94

NOTE: Cells with counts of less than 10 (<10) cannot be disclosed according to ONS LS confidentially requirements. The cell counts and associated percentage estimates are adjusted accordingly.
SOURCE: Authors calculations based on ONS LS.

Table 3.4. Hazard ratios of first generation all-cause mortality by covariates, England and Wales 1971-2013.

Census Interval	HR	Sig.	95% CI		HR	Sig.	95% CI	
[1971-1981)	0.01	***	0.01	0.01	0.01	***	0.01	0.01
[1981-1991)	0.01	***	0.01	0.01	0.01	***	0.01	0.01
[1991-2001)	0.01	***	0.01	0.01	0.01	***	0.01	0.01
[2001-2011)	0.01	***	0.01	0.01	0.01	***	0.01	0.01
2011+	0.01	***	0.01	0.01	0.00	***	0.00	0.01
Sex								
Male	*Ref.*				*Ref.*			
Female	0.79	***	0.78	0.79	0.55	***	0.55	0.56
Country of Birth								
England & Wales	*Ref.*				*Ref.*			
Scotland	1.15	***	1.11	1.19	1.01		0.97	1.04
Other UK	1.33	***	1.29	1.37	0.99		0.96	1.02
Western Europe	0.68	***	0.64	0.72	0.71	***	0.67	0.75
Eastern Europe	1.19	***	1.14	1.24	0.89	***	0.85	0.92
Africa & Caribbean	0.50	***	0.48	0.53	0.48	***	0.46	0.51
South Asia	0.53	***	0.51	0.55	0.49	***	0.47	0.51
East Asia	0.39	***	0.35	0.43	0.42	***	0.38	0.47
N. America, Aus. & NZ.	0.64	***	0.59	0.70	0.74	***	0.68	0.81
South America	0.52	***	0.43	0.63	0.49	***	0.41	0.60
Other, Rest of World	0.76	***	0.65	0.90	0.67	***	0.57	0.79
Missing	0.29	***	0.28	0.29	1.13	***	1.05	1.22
Marital Status								
Married					*Ref.*			
Single					0.24	***	0.23	0.24
Divorced/Widowed					1.58	***	1.56	1.60
Missing					11.84	**	1.67	84.08
Car Deprivation								
Not deprived					*Ref.*			
Deprived					2.07	***	2.05	2.09
Missing					1.54	***	1.25	1.90
Housing Tenure Deprivation								
Not deprived					*Ref.*			
Deprived					1.10	***	1.09	1.11
Missing					1.98	***	1.94	2.02
Overcrowding Deprivation								
Not deprived					*Ref.*			
Deprived					0.41	***	0.40	0.42
Missing					0.45	***	0.38	0.53

Household-head Social Class					
I Professional		*Ref.*			
II Intermediate		1.22	***	1.18	1.26
III Skilled		1.31	***	1.27	1.35
IV Partly Skilled		1.51	***	1.46	1.56
V Unskilled		1.72	***	1.66	1.78
Other/Armed Forces		3.04	***	2.94	3.15
Missing & NA		3.09	***	2.99	3.19
Carstairs Quintile					
Least deprived		*Ref.*			
Next least deprived		1.03	***	1.02	1.05
Midpoint		1.05	***	1.04	1.07
Moderately		1.06	***	1.04	1.07
Most deprived		1.05	***	1.03	1.07
Missing		0.71	***	0.64	0.78
Rurality					
Non-metropolitan		*Ref.*			
Other metropolitan		0.99		0.97	1.00
London		0.99		0.98	1.01
Missing		0.59	***	0.47	0.74
Number of Subjects	1,002,377	820,040			
Number of Observations	5,876,789	4,049,578			
Log-likelihood	-295,760.25	-164,809.19			

NOTE: Significance levels at 1% (***) 5% (**) and 10% (*).
SOURCE: Authors calculations based on ONS LS.

Table 3.5. Hazard ratios of first generation smoking-attributable mortality by covariates, England and Wales 1971-2013.

Census Interval	**HR**	**Sig.**	**95% CI**		**HR**	**Sig.**	**95% CI**	
[1971-1981)	0.00	***	0.00	0.00	0.00	***	0.00	0.00
[1981-1991)	0.01	***	0.01	0.01	0.01	***	0.01	0.01
[1991-2001)	0.01	***	0.01	0.01	0.00	***	0.00	0.00
[2001-2011)	0.01	***	0.01	0.01	0.00	***	0.00	0.00
2011+	0.00	***	0.00	0.00	0.00	***	0.00	0.00
Sex								
Male	*Ref.*				*Ref.*			
Female	0.70	***	0.69	0.71	0.48	***	0.48	0.49
Country of Birth								
England & Wales	*Ref.*				*Ref.*			
Scotland	1.22	***	1.16	1.28	1.05	**	1.01	1.10
Other UK	1.42	***	1.37	1.48	1.02		0.98	1.06
Western Europe	0.65	***	0.60	0.71	0.70	***	0.64	0.76
Eastern Europe	1.31	***	1.24	1.38	0.95	**	0.90	1.00
Africa & Caribbean	0.45	***	0.41	0.48	0.42	***	0.39	0.45
South Asia	0.54	***	0.51	0.57	0.48	***	0.45	0.50
East Asia	0.35	***	0.30	0.41	0.37	***	0.32	0.43
N. America, Aus. & NZ.	0.62	***	0.55	0.70	0.74	***	0.65	0.83
South America	0.38	***	0.28	0.52	0.36	***	0.26	0.49
Other, Rest of World	0.72	**	0.54	0.96	0.62	***	0.47	0.83
Missing	0.28	***	0.27	0.28	1.16	***	1.05	1.28
Marital Status								
Married					*Ref.*			
Single					0.18	***	0.18	0.19
Divorced/Widowed					1.47	***	1.44	1.49
Missing					.	.	.	.
Car Deprivation								
Not deprived					*Ref.*			
Deprived					2.07	***	2.04	2.11
Missing					1.45	***	1.12	1.89
Housing Tenure Deprivation								
Not deprived					*Ref.*			
Deprived					1.17	***	1.16	1.19
Missing					1.96	***	1.91	2.01
Overcrowding Deprivation								
Not deprived					*Ref.*			
Deprived					0.43	***	0.41	0.45
Missing					0.43	***	0.34	0.53

Household-head Social Class					
I Professional		*Ref.*			
II Intermediate		1.27	***	1.21	1.33
III Skilled		1.49	***	1.42	1.55
IV Partly Skilled		1.75	***	1.67	1.83
V Unskilled		2.04	***	1.94	2.15
Other/Armed Forces		3.26	***	3.10	3.43
Missing & NA		3.83	***	3.65	4.00
Carstairs Quintile					
Least deprived		*Ref.*			
Next least deprived		0.97	***	0.94	0.99
Midpoint		1.01		0.99	1.03
Moderately		1.03	**	1.01	1.05
Most deprived		1.04	***	1.02	1.06
Missing		0.72	***	0.64	0.82
Rurality					
Non-metropolitan		*Ref.*			
Other metropolitan		1.01		1.00	1.03
London		0.98	*	0.96	1.00
Missing		0.54	***	0.41	0.71
Number of Subjects	1,002,377	820,040			
Number of Observations	5,876,789	4,049,578			
Log-likelihood	-214,094.74	-133,817.24			

NOTE: Significance levels at 1% (***) 5% (**) and 10% (*).
SOURCE: Authors calculations based on ONS LS.

Table 3.6. Hazard ratios of first generation alcohol-attributable mortality by covariates, England and Wales 1971-2013.

Census Interval	HR	Sig.	95% CI		HR	Sig.	95% CI	
[1971-1981)	0.00	***	0.00	0.00	0.00	***	0.00	0.00
[1981-1991)	0.00	***	0.00	0.00	0.01	***	0.01	0.01
[1991-2001)	0.00	***	0.00	0.00	0.00	***	0.00	0.00
[2001-2011)	0.00	***	0.00	0.00	0.00	***	0.00	0.00
2011+	0.00	***	0.00	0.00	0.00	***	0.00	0.00
Sex								
Male	*Ref.*				*Ref.*			
Female	0.50	***	0.47	0.54	0.45	***	0.42	0.49
Country of Birth								
England & Wales	*Ref.*				*Ref.*			
Scotland	1.87	***	1.50	2.34	1.75	***	1.40	2.19
Other UK	1.79	***	1.42	2.26	1.44	***	1.14	1.82
Western Europe	0.83		0.54	1.27	0.86		0.56	1.33
Eastern Europe	1.36	*	0.99	1.87	1.16		0.85	1.60
Africa & Caribbean	0.85		0.64	1.15	0.80		0.59	1.08
South Asia	0.99		0.80	1.23	0.95		0.76	1.18
East Asia	0.69		0.39	1.21	0.71		0.40	1.25
N. America, Aus. & NZ.	1.08		0.66	1.77	1.22		0.74	1.99
South America	0.58		0.14	2.31	0.55		0.14	2.19
Other, Rest of World	1.35		0.34	5.41	1.20		0.30	4.83
Missing	0.52	***	0.48	0.56	1.77	***	1.31	2.39
Marital Status								
Married					*Ref.*			
Single					0.18	***	0.18	0.19
Divorced/Widowed					1.47	***	1.44	1.49
Missing					.	.	.	.
Car Deprivation								
Not deprived					*Ref.*			
Deprived					2.03	***	1.84	2.23
Missing					4.68		0.58	37.64
Housing Tenure Deprivation								
Not deprived					*Ref.*			
Deprived					1.28	***	1.16	1.42
Missing					1.36	***	1.22	1.52
Overcrowding Deprivation								
Not deprived					*Ref.*			
Deprived					0.72	***	0.61	0.85
Missing					0.27		0.04	1.94

Household-head Social Class					
I Professional		*Ref.*			
II Intermediate		1.04		0.85	1.27
III Skilled		1.09		0.89	1.32
IV Partly Skilled		1.24	**	1.01	1.53
V Unskilled		1.16		0.91	1.49
Other/Armed Forces		1.52	**	1.06	2.18
Missing & NA		1.54	***	1.24	1.91
Carstairs Quintile					
Least deprived		*Ref.*			
Next least deprived		0.94		0.82	1.08
Midpoint		0.94		0.83	1.07
Moderately		1.08		0.96	1.22
Most deprived		1.10		0.98	1.23
Missing		0.70		0.38	1.27
Rurality					
Non-metropolitan		*Ref.*			
Other metropolitan		0.97		0.88	1.07
London		0.92		0.82	1.04
Missing		.	.	.	.
Number of Subjects	1,002,377	820,040			
Number of Observations	5,876,789	4,049,578			
Log-likelihood	-19,636.78	-14,536.67			

NOTE: Significance levels at 1% (***) 5% (**) and 10% (*). Cells indicated with '.' denote categories with counts too low for model estimate.
SOURCE: Authors calculations based on ONS LS.

Table 3.7. Hazard ratios of first generation all other cause mortality by covariates, England and Wales 1971-2013.

Census Interval	HR	Sig.	95% CI		HR	Sig.	95% CI	
[1971-1981)	0.01	***	0.01	0.01	0.01	***	0.01	0.01
[1981-1991)	0.00	***	0.00	0.00	0.00	***	0.00	0.00
[1991-2001)	0.00	***	0.00	0.00	0.00	***	0.00	0.00
[2001-2011)	0.00	***	0.00	0.00	0.00	***	0.00	0.00
2011+	0.00	***	0.00	0.00	0.00	***	0.00	0.00
Sex								
Male	*Ref.*				*Ref.*			
Female	0.91	***	0.90	0.92	0.65	***	0.64	0.66
Country of Birth								
England & Wales	*Ref.*				*Ref.*			
Scotland	1.05	**	1.00	1.11	0.93	***	0.88	0.98
Other UK	1.22	***	1.17	1.28	0.96	*	0.92	1.00
Western Europe	0.70	***	0.65	0.77	0.71	***	0.66	0.78
Eastern Europe	1.06	*	0.99	1.12	0.81	***	0.76	0.87
Africa & Caribbean	0.56	***	0.52	0.60	0.55	***	0.51	0.60
South Asia	0.50	***	0.47	0.53	0.49	***	0.46	0.52
East Asia	0.44	***	0.37	0.51	0.47	***	0.40	0.56
N. America, Aus. & NZ.	0.65	***	0.58	0.74	0.73	***	0.65	0.83
South America	0.69	***	0.53	0.88	0.66	***	0.51	0.85
Other, Rest of World	0.78	**	0.64	0.95	0.68	***	0.56	0.83
Missing	0.29	***	0.28	0.30	1.03		0.91	1.18
Marital Status								
Married					*Ref.*			
Single					0.29	***	0.28	0.29
Divorced/Widowed					1.71	***	1.68	1.75
Missing					36.91	***	5.17	263.33
Car Deprivation								
Not deprived					*Ref.*			
Deprived					2.07	***	2.04	2.11
Missing					1.57	***	1.11	2.22
Housing Tenure Deprivation								
Not deprived					*Ref.*			
Deprived					1.03	***	1.01	1.04
Missing					2.12	***	2.06	2.19
Overcrowding Deprivation								
Not deprived					*Ref.*			
Deprived					0.39	***	0.37	0.40
Missing					0.49	***	0.36	0.65

Household-head Social Class					
I Professional		*Ref.*			
II Intermediate		1.19	***	1.14	1.24
III Skilled		1.16	***	1.11	1.21
IV Partly Skilled		1.30	***	1.24	1.36
V Unskilled		1.47	***	1.40	1.54
Other/Armed Forces		2.75	***	2.62	2.89
Missing & NA		2.50	***	2.39	2.62
Carstairs Quintile					
Least deprived		*Ref.*			
Next least deprived		0.95	***	0.93	0.98
Midpoint		0.98	*	0.96	1.00
Moderately		0.99		0.96	1.01
Most deprived		0.97	**	0.95	1.00
Missing		0.62	***	0.53	0.72
Rurality					
Non-metropolitan		*Ref.*			
Other metropolitan		0.96	***	0.94	0.98
London		1.00		0.98	1.03
Missing		0.70	*	0.48	1.03
Number of Subjects	1,002,377	820,040			
Number of Observations	5,876,789	4,049,578			
Log-likelihood	-199,619.71	-135,775.93			

NOTE: Significance levels at 1% (***) 5% (**) and 10% (*).
SOURCE: Authors calculations based on ONS LS.

Table 3. 8. Number of second generation infants, person-years at risk, deaths, and column percentages by covariates, England and Wales 1977-2013.

Sex	Individuals	%	PY at risk	%	Deaths	%
Male	154,596	51.07	3,250,000	51.11	764	57.44
Female	148,142	48.93	3,108,333	48.89	566	42.56
Maternal Country of Birth						
England & Wales	243,767	80.52	5,250,000	82.60	1080	81.69
Scotland	4,628	1.53	110,568	1.74	20	1.51
Other UK	1,602	0.53	39,399	0.62	<10	0.61
Eastern Europe	4,253	1.40	28,597	0.45	<10	0.68
Western Europe	8,603	2.84	191,107	3.01	39	2.95
Aus, NZ, NA, & Oceania	3,319	1.10	55,569	0.87	10	0.76
N, C, S Africa & ME	5,143	1.70	74,524	1.17	21	1.59
East & South Africa	4,995	1.65	81,978	1.29	16	1.21
East & Southeast Asia	2,802	0.93	50,379	0.79	<10	0.53
South Asia	14,486	4.78	265,843	4.18	84	6.35
Central America & Caribbean	2,128	0.70	54,789	0.86	17	1.29
South America	767	0.25	12,811	0.20	<10	0.08
Missing	6,245	2.06	140,502	2.21	10	0.76
Birthweight						
Non-LBW	288,071	95.16	6,108,333	96.11	926	69.62
LBW	14,667	4.84	246,983	3.89	404	30.38
Maternal Marital Status at Birth						
Married	215,504	71.18	5,041,667	79.29	956	71.88
Unmarried	87,234	28.82	1,316,667	20.71	374	28.12
Maternal Age at Birth (years)						
<20	23,926	7.90	556,573	8.76	189	14.21
20-24	74,320	24.55	1,758,333	27.69	397	29.85
25-29	96,918	32.01	2,158,333	33.98	391	29.40
30-34	71,171	23.51	1,308,333	20.60	225	16.92
35-29	30,180	9.97	475,077	7.48	96	7.22
40+	6,223	2.06	94,534	1.49	32	2.41

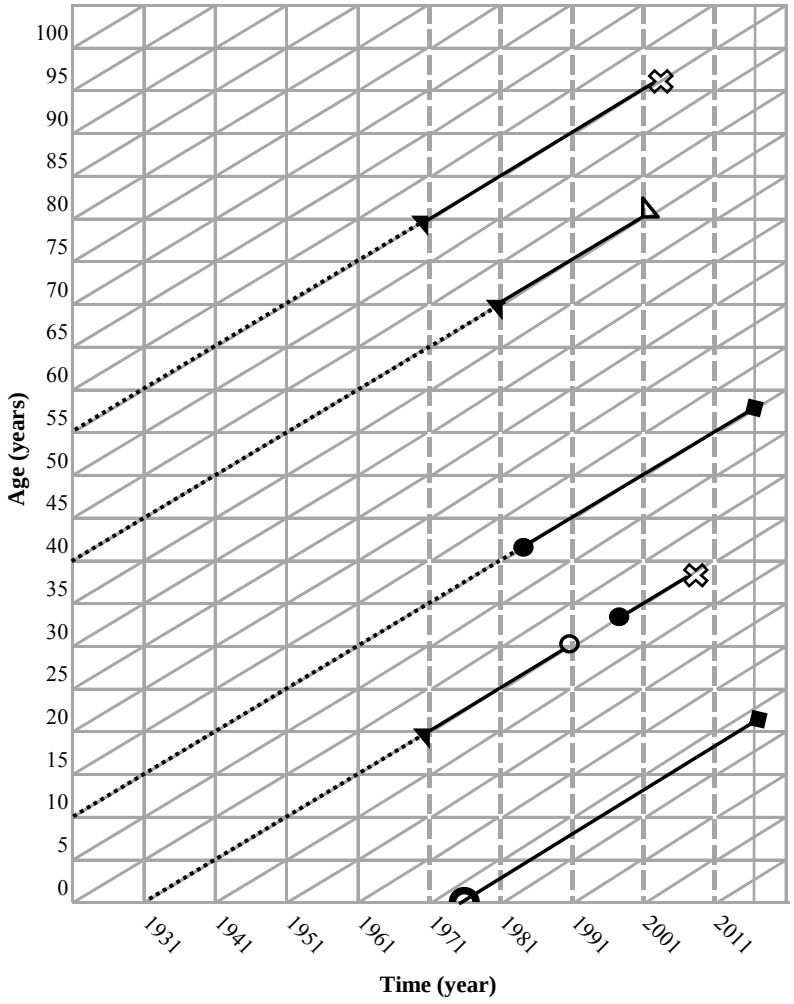
NOTE: "Aus, NZ, NA, & Oceania" signifies Australia, New Zealand, North America, and Oceania. "N, C, S Africa & ME" signifies North, Central, South Africa, and the Middle East.
SOURCE: Authors calculations based on ONS LS.

Table 3.9. Hazard ratios of second generation all-cause mortality by covariates, England and Wales 1977-2013.

Sex	**HR**	**Sig.**	**95% CI**		**HR**	**Sig.**	**95% CI**	
Male	*Ref.*				*Ref.*			
Female	0.77	***	0.69	0.86	0.74	***	0.66	0.82
Maternal Country of Birth								
England & Wales	*Ref.*				*Ref.*			
Scotland	0.98		0.63	1.52	1.03		0.66	1.60
Other UK	1.12		0.56	2.25	1.25		0.62	2.51
Eastern Europe	0.50	*	0.26	0.96	0.57	*	0.30	1.11
Western Europe	1.03		0.75	1.41	1.14		0.82	1.57
Aus, NZ, NA, & Oceania	0.69		0.37	1.28	0.74		0.40	1.38
N, C, S Africa & ME	0.93		0.61	1.44	0.99		0.64	1.53
East & South Africa	0.73		0.44	1.19	0.65	*	0.40	1.07
East & Southeast Asia	0.57		0.27	1.19	0.62		0.29	1.31
South Asia	1.32	**	1.06	1.65	1.01		0.80	1.27
Central America & Caribbean	2.33	***	1.53	3.56	2.52	***	1.65	3.86
South America	0.30		0.04	2.10	0.34		0.05	2.41
Missing	0.46	***	0.27	0.80	0.43	***	0.25	0.74
Birthweight								
Non-LBW					*Ref.*			
LBW					9.25	***	8.22	10.42
Maternal Marital Status at Birth								
Married					*Ref.*			
Unmarried					0.68	***	0.60	0.77
Maternal Age at Birth (years)								
<20					*Ref.*			
20-24					0.65	***	0.55	0.78
25-29					0.49	***	0.40	0.58
30-34					0.38	***	0.30	0.46
35-29					0.37	***	0.29	0.48
40+					0.59	***	0.40	0.86
Number of Subjects	301,836				301,836			
Number of Observations	3,569,977				3,569,977			
Log-likelihood	-14,651.47				-14,119.74			

NOTE: "Aus, NZ, NA, & Oceania" signifies Australia, New Zealand, North America, and Oceania. "N, C, S Africa & ME" signifies North, Central, South Africa, and the Middle East. Significance levels at 1% (***) 5% (**) and 10% (*).
SOURCE: Authors calculations based on ONS LS.

Figure 3.1. Lexis diagram representing the study population.



Life-lines

- At-risk of death, not yet entered observation
- At-risk of death, under observation

Entry Events

- Registered birth (after 1971 census)
- Registered in-migrant
- Registered at census

Exit Events

- Untraceable at census
- Registered death
- Registered out-migrant
- Alive, end of observation period

Appendix A1. Population-weighted rurality score deciles.

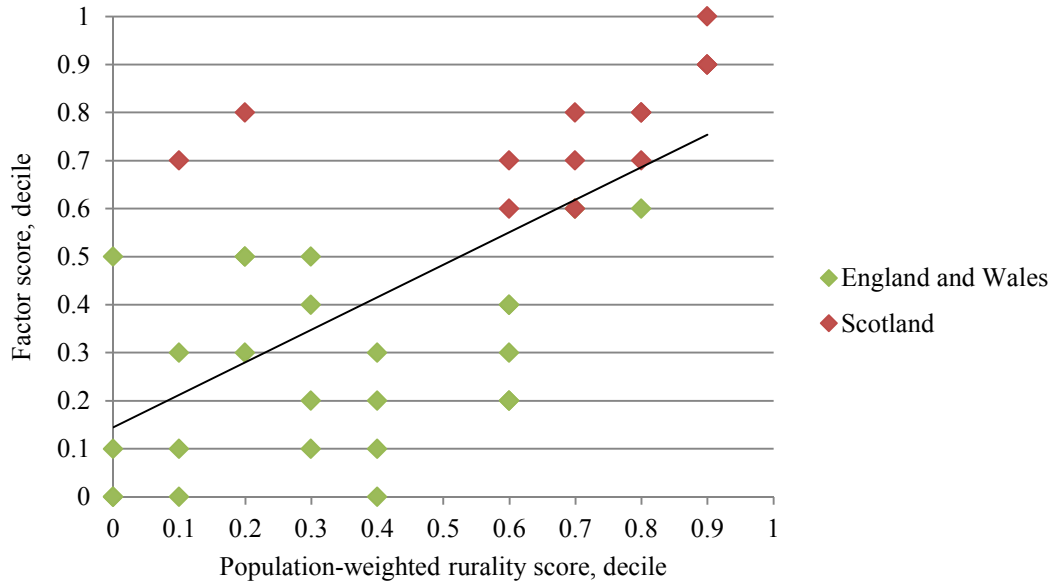
The following population-weighted rurality score deciles are calculated separately for each sex. For each geography, raw population densities are calculated as the total sex-specific population in that geography divided by the geographic shape's area; and these raw population densities are sorted into raw density deciles. For each decile, the proportion of the global population residing in that decile is next calculated by summing the total population living in that decile divided by the total population (A1). Finally, a density score is calculated for each geography weighted by the proportion of the population living in the geography's corresponding raw decile score (A2); and these weighted scores are then sorted in deciles.

$$C_d = \frac{N_d}{\sum_{d=1}^{10} N_d} \quad \text{(A1)}$$

$$S_{d,a} = \left(R_{d,a} * C_d\right) * 100 \quad \text{(A2)}$$

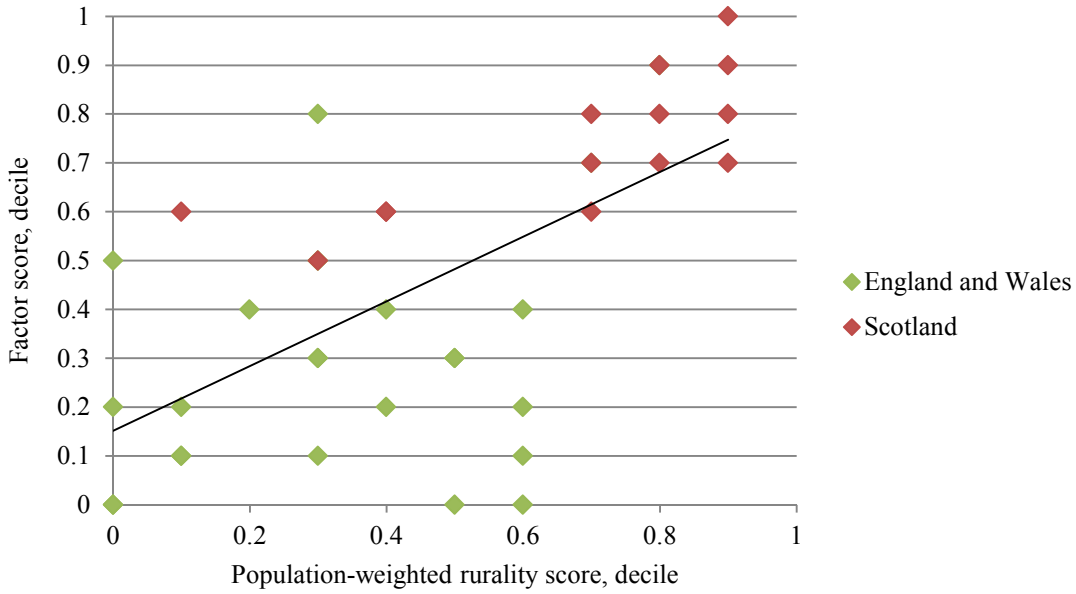
where *Cd* corresponds to the proportion of the total population living in decile *d*; *Sda* corresponds to the population-weighted score for density *d*, for each area *a*, calculated from the raw density decile of area *a*, *Rda*. The density scores are population-weighted in this way because the relative population density difference is more relevant than an absolute difference (e.g. a difference in population size between an area of 1,000 and 2,000 people is meaningful different than for two areas with a population size of 11,000 and 12,000 people, though the absolute difference is identical). This methodology is based on a density score approach used to examine population density within Great Britain (Craig 1984). The population-weighted density deciles are finally inverted to aid in interpretation. These inverted deciles are called population-weighted rurality scores, and range from 0 (the least rural decile) to 0.9 (the most rural decile).

Figure A1.1.A. Factor 2 score deciles and population-weighted rurality score deciles, males 1981-2009.



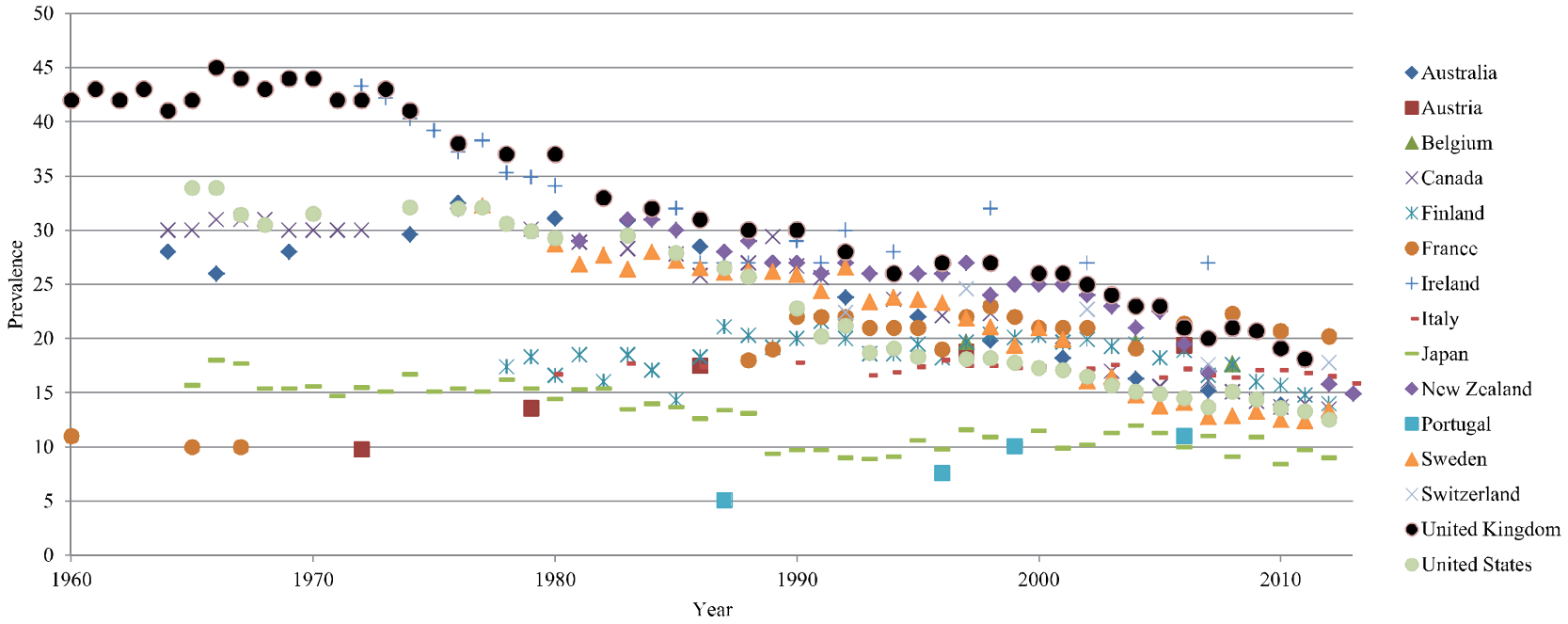
NOTE: Population-weighted rurality score deciles range from 0 (the least rural decile) to 0.9 (the most rural decile). Factor score deciles range from 0 (the lowest factor score) to 0.9 (the highest factor score).

Figure A1.1.B. Factor 2 score deciles and population-weighted rurality score deciles, females 1981-2009.



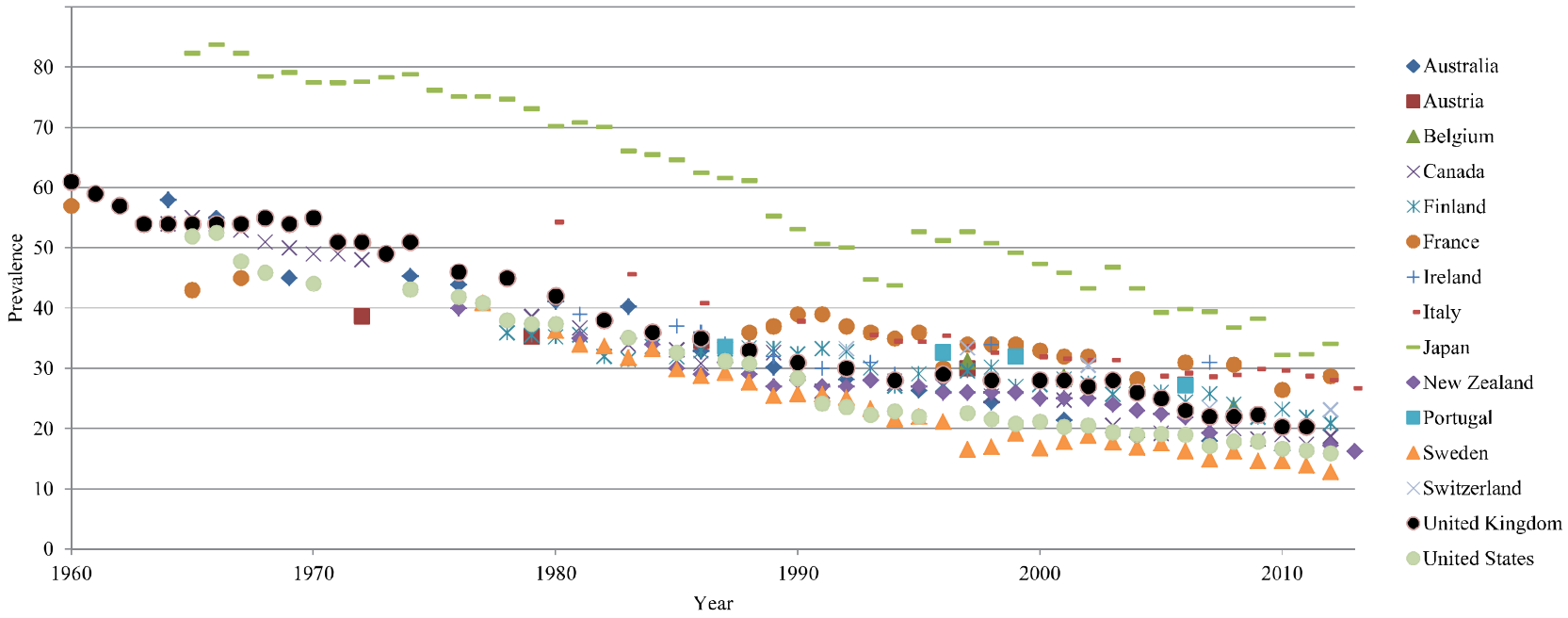
NOTE: Population-weighted rurality score deciles range from 0 (the least rural decile) to 0.9 (the most rural decile). Factor score deciles range from 0 (the lowest factor score) to 0.9 (the highest factor score).

Figure A2.1.A. Prevalence of daily smokers among women aged 15+ years, by year and country.



Source: OCCD Health Statistics 2014. June 2014. www.oecd.org/els/health-systems/health-data.htm. (accessed November 2014).

Figure A2.1.B. Prevalence of daily smokers among men aged 15+ years, by year and country.



Source: OCCD Health Statistics 2014. June 2014. www.oecd.org/els/health-systems/health-data.htm. (accessed November 2014).

Table A3.1. Country of birth categories and individual countries.

Solo Countries	**Middle East**	**Eastern Europe**	**Western Europe**
England and Wales	Bahrain	Albania	Akrotiri
Ireland	Gaza Strip	Armenia	Andorra
Northern Ireland	Iran	Azerbaijan	Austria
Rest of UK	Iraq	Belarus	Belgium
Scotland	Israel	Bosnia	Cyprus
	Jordan	Bulgaria	Denmark
	Kuwait	Croatia	Dhekelia
	Lebanon	Czech Republic	EU
	Oman	Estonia	Faroe Islands
	Qatar	Georgia	Finland
	Saudi Arabia	Herzegovina	France
	Syria	Hungary	Germany
	United Arab Emirates	Italy	Gibraltar
	West Bank	Kazakhstan	Greece
	Yemen	Kosovo	Greenland
		Kyrgyzstan	Guernsey
		Latvia	Iceland
		Lithuania	Jan Mayen
		Macedonia	Jersey
		Moldova	Liechtenstein
		Montenegro	Lithuania
		Poland	Luxembourg
		Romania	Malta
		Romania	Monaco
		Russia	Netherlands
		Serbia	Norway
		Slovakia	Portugal
		Slovenia	San Marino
		Tajikistan	Spain
		Turkey	Svalbard
		Turkmenistan	Sweden
		Ukraine	Switzerland
		Uzbekistan	Vatican
		Yugoslavia	
Australia - Oceania	**East and Southeast Asia**	**South Asia**	**Cen. Am and Caribbean**
American Samoa	Brunei	Afghanistan	Anegada
Ashmore and Cartier	Burma	Bangladesh	Anguilla
Australia	Cambodia	Bhutan	Antigua and Barbuda
Baker Island	China	British Indian Ocean	Aruba

Christmas Island	Hong Kong	Danger Islands	Bahamas
Cocos Islands	Indonesia	Diego Garcia	Barbados
Cook Islands	Japan	Eagle Islands	Belize
Coral Sea Islands	Korea	Egmont Islands	Bermuda
Fiji	Laos	India	Bonaire
French Polynesia	Macau	Maldives	British Virgin Islands
Guam	Malaysia	Nelsons Islands	Cayman Islands
Howland Island	Mongolia	Nepal	Charlotte Amalie
Jarvis Island	Paracel Islands	Pakistan	Costa Rica
Johnston Atoll	Philippines	Peros Banhos	Cuba
Kingman Reef	Singapore	Salomon Islands	Curacao
Kirbati	Spratly Islands	Sri Lanka	Dominica
Kiribati	Taiwan	Three Brothers	Dominican Republic
Marshall Islands	Thailand		El Salvador
Micronesia	Timor-Leste		Grenada
Midway Islands	Vietnam		Guadeloupe
Nauru			Guatemala
New Caledonia			Haiti
New Zealand			Honduras
Niue			Jamaica
Norfolk Island			Jost Van Dyke
N. Mariana Islands			Martinique
Pacific Islands			Montserrat
Palau			Navassa Island
Palmyra Atoll			Nicaragua
Papua New Guinea			Panama
Pitcairn Islands			Puerto Rico
Samoa			Saint Barthelemy
Solomon Islands			Saint Croix
Tokelau			Saint John
Tonga			Saint Kitts and Nevis
Tuvalu			Saint Lucia
Vanuatu			Saint Martin
Wake Island			Saint Thomas
Wallis and Futuna			St. Vin. and the Grenadines
			Sint Maarten
			Tortola
			Trinidad and Tobago
			Turks and Caicos Islands
			Virgin Gorda
			Virgin Islands

South America	North America	Western Africa	
Argentina	Canada	Benin	
Bolivia	Clipperton Island	Burkina Faso	
Brazil	Mexico	Cabo Verde	
Chile	Saint Pierre and Miquelon	Ivory Coast	
Colombia	USA	Gambia	
Ecuador		Ghana	
Falkland Islands		Guinea	
French Guiana		Guinea-Bissau	
Guyana		Liberia	
Paraguay		Mali	
Peru		Mauritania	
South Georgia		Niger	
S. Sandwich Islands		Nigeria	
Suriname		Saint Helena, Ascension & Tristan da Cunha	
Uruguay		Senegal	
Venezuela		Sierra Leone	
		Togo	
Southern Africa	**Northern Africa**	**Central Africa**	**Eastern Africa**
Botswana	Algeria	Angola	Burundi
Lesotho	Egypt	Cameroon	Comoros
Namibia	Libya	Central African Rep.	Djibouti
South Africa	Morocco	Chad	Djibouti
Swaziland	Sudan	Dem. Rep. of Congo	Eritrea
Zambia	Tunisia	Equatorial Guinea	Ethiopia
	Western Sahara	Gabon	Kenya
		Republic of the Congo	Madagascar
		São Tomé and Príncipe	Malawi
			Mauritius
			Mayotte
			Mozambique
			Réunion
			Rwanda
			Seychelles
			Somalia
			South Sudan
			Tanzania
			Uganda
			Zambia
			Zimbabwe