Diversity in causes of mortality
in the measurement of population health in Scotland

Ciaran McMonagle, MSc, BSc

SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

SCHOOL OF HEALTH AND WELLBEING

COLLEGE OF MEDICAL, VETERINARY AND LIFE SCIENCES

April 2023

Abstract

Measuring diversity in causes of mortality offers an insight into variation in health outcomes within a population. Increased diversity in causes of mortality indicates that deaths have occurred from more varied causes. This may increase diagnostic uncertainty and means health care, promotion, and prevention resources must be spread wider and these sectors must adopt a more comprehensive approach. Diversity in mortality causes has not been measured in Scotland, despite poor health outcomes relative to European comparators. Further, limited previous examination exists of differences in mortality cause diversity in sub-national population groups, divided by socioeconomic or geographic factors. Health inequalities in Scotland are large and understanding tendencies in mortality cause diversity may be valuable to addressing differential health patterns. Mortality cause diversity has been shown to be associated with increasing life expectancy and falling lifespan variation over time across nations. This relationship has not been examined between different subpopulations within a nation. Finally, the effect of the COVID-19 pandemic on diversity in causes of mortality has not been examined and this analysis may be a valuable tool to assess the ongoing impact caused by this unprecedented upheaval in population health.

In this thesis I calculate diversity in underlying and contributory causes of mortality as well as lifespan diversity using observed data and distributions extracted from multiple-decrement life tables. I propose novel methods for assessing the contribution of causes of mortality to diversity and a novel method for the measurement of lifespan diversity. I find that diversity in underlying and contributory causes of mortality increased in Scotland from 2001 to the mid 2010s when trends diverge. Trends in variation are shown to be similar across subpopulations, meaning despite socioeconomic or geographic differences reductions in the proportion of individuals who die of the most common causes and a redistribution to a wider variety of causes has occurred at similar rates. Confirming previous research, diversity in causes of mortality is found to increase as the population of Scotland lived longer and to more equal ages. However, higher diversity in causes of mortality is not necessarily found among subpopulations who live to older and more equal ages. I suggest falling mortality rates associated with the most common causes, especially at premature ages, have driven increasing diversity in causes of mortality and life expectancy and falling lifespan diversity. Diversity in causes of mortality, with COVID-19 deaths are excluded from analysis, is shown to have remained consistent with trends in previous years during the COVID-19 pandemic. Monitoring diversity in mortality causes has the potential to expand knowledge around patterns of mortality and to provide valuable insight into pressures on public health and healthcare systems.

Contents

List of Tables

List of Figures

Acknowledgements

This thesis would not have been possible without the support of a great many friends, family and colleagues.

First, to my parents, Fiona and Tony, whose unwavering and unconditional support throughout my life is a gift for which I will forever be grateful. Their advice has been ever-helpful and it is thanks to their encouragement, especially when I have struggled most, that I have been able to complete this thesis.

Second, to my supervisors who have shown me generosity, patience and wisdom in equal measure. I have grown a great deal as a researcher in during my PhD thanks to your advice and your insights have improved this work greatly.

Third, I am truly thankful for the inimitable students and staff of the SPHSU a talented and kind group whose advice and assistance have aided in my progression at every turn. I will miss the company and camaraderie of the the student community which has been invaluable to keeping me sane during my studies.

Finally, I must thank my friends for whom I consider myself incredibly lucky. They have have supported, distracted, and put up with me during even the most difficult times. Special thanks must go to those I have lived with and who were the best flatmates a stressed PhD student could ask for: Laura and especially George the best, worst influence I have ever known.

Author's declaration

I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

_____

Ciaran McMonagle

Published papers

There follows a link to the version of record of an article accepted for publication following peer-review in the completion of this thesis: McMonagle, C., Brown, D., Reeve, R., & Mancy, R (2022) 'Trends in the diversity of mortality causes and age-standardised mortality rates among subpopulations within Scotland, 2001–2019', SSM - Population Health, 19, p. 101192.

https://doi.org/10.1016/j.ssmph.2022.101192.

I declare that the research carried out for this paper was my own work and that co-authors provided editorial support.

Abbreviations

- ACF: Autocorrelation Function
- ARIMA: Auto-Regressive Integrated Moving Average
- ASMR: Age-Standardised Mortality Rate
- CHD: Chronic Heart Disease
- CSO: Chief Scientist Office (Scotland)
- DALYs: Disability-Adjusted Life Years
- GBD: Global Burden of Disease study
- HMD: Human Mortality Database
- ICD-10: International Classification of Diseases (10^{th} Revision)
- MCOD: Multiple Causes Of Death
- MRC: Medical Research Council
- NHS: National Health Service
- NPI: Non-Pharmaceutical Intervention
- NRS: National Records of Scotland
- PACF: Partial Autocorrelation Function
- SIMD: Scottish Index of Multiple Deprivations
- UK: United Kingdom
- USA: United States of America

Glossary of selected terms

- Aetiology: The reasons for or causes of a disease or condition occurring in an individual or in the population at large.
- Deaths of despair: Deaths attributed to suicide or associated with drug or alcohol abuse.
- Distribution of mortality causes: The causes of mortality which are recorded within a population. Perhaps best thought of visually as a bar graph where the x-axis is each cause of death under consideration and the y-axis is the number of deaths attributed to each cause in the population in question.
- Dominant: The most common type (cause of mortality or age at death) within a distribution.
- Fragmentation: The increased division in the causes of mortality faced by a population with a larger proportion of deaths occurring due to a greater variety of causes.
- Prominent: A type which was common but was not the dominant type in the distribution.
- Prevalence: The proportion of deaths attributed to a cause of mortality, measured as a percentage of all-cause age-standardised mortality rate (ASMR) attributed to an individual cause of mortality.
- Lifespan diversity: The diversity in ages at death within a population. A measure of variation in lifespans which I propose to use in this thesis.
- Mortality: Of or relating to death, especially at a large scale.
- Morbidity: The state of being subject to a disease or medical condition.
- Contributory cause of mortality: A disease, condition or injury recorded on the death certificate which was considered to have hastened or contributed to death but not to be the underlying reason an individual died.

- Comorbidities: The simultaneous presentation of two diseases or conditions in an individual.
- Multimorbidities: Generally used to describe the simultaneous presentation of two or more diseases or conditions in an individual. Used in this thesis to indicate cases where more than two diseases or conditions present together to differentiate from the term ”comorbidities”.
- Subpopulation: A group of individuals within a nation defined by meaningful individual or area level characteristics. Such as an income deprivation quintile.
- Underlying cause of mortality: The disease, condition, or injury recorded as the primary or main cause of death on a death certificate.

Chapter 1

Introduction

1.1 Chapter overview

This chapter introduces this thesis. It begins with a background of the measurement of diversity in causes of mortality (Section 1.2), followed by a section in which the importance of measuring health within population groups within nations is considered (Section 1.3) and another section which features a discussion of population health in Scotland (Section 1.4). In conclusion, the structure of the thesis is described (Section 1.5)[1].

1.2 Diversity and variation in the measurement of population health

Mortality patterns around the world are in a state of near constant flux (Omran, 1971; Santosa et al., 2014). Individual diseases, conditions, and indicators of ill-health become more and less common over time as their causes and risk factors emerge and are addressed by society (Mackenbach, 1994). Traditionally, research plotting these changes has, mostly, been carried out through the measurement of population-wide rates, risks, or averages. There is a recognition in the literature, however, that distributions of health outcomes around these central tendencies are

[1] Throughout this thesis, subdivisions headed by whole numbers are referred to as chapters and subdivisions headed by decimal numbers are referred to as sections.

key to understanding health across populations (Kindig & Stoddart, 2003; Etches et al., 2006). For example, variations in the distribution of lifespans have been examined to describe health inequalities in populations around the world (Seaman et al., 2016a). Greater variation in lifespans represents greater inequality within a group, as people have commonly died at a wider variety of ages.

Relatively little attention has previously been paid to variation in the distribution of causes of mortality within populations. Measures of diversity, originally from the field of ecology, have been proposed to measure this variation (Izsák, 1986; Bergeron-Boucher et al., 2020; Trias-Llimós & Permanyer, 2023). Whether or not individuals are likely to face the same cause of mortality does not, generally, factor into assessments of equality in health (Bergeron-Boucher et al., 2020). Instead, increases or decreases in variation in the distribution of mortality causes have other individual-level and population-wide implications for health.

It has been proposed that increases in diversity in mortality causes have the potential to increase diagnostic uncertainty, as the range of diseases and conditions each individual might present with broadens (Bergeron-Boucher et al., 2020; John & Innocent, 2005). At the population level, increased variation in the distribution of mortality causes is an indicator of increased fragmentation in the causes of diseases and ill-health faced by the population and may be underpinned by increases in multimorbitity (Bergeron-Boucher et al., 2020). This means that the burden of disease is more evenly spread across causes and therefore health care and public health systems must adapt. As populations become less likely to face the most common causes of illness, efficiencies and economies of scale in public health and health care systems lose their potency (Long et al., 1985; de Bruin et al., 2007). Medical research and pharmaceutical bodies must spread limited resources further to find remedies, preventive strategies, and treatments for diseases that occur relatively less frequently and, therefore, from a pharmaceutical point of view, are less profitable. In a population where diversity in mortality causes is increasing, public health resources must also be spread wider and a more holistic approach must be adopted to improve the health of the population (Bergeron-Boucher et al., 2020; Trias-Llimós & Permanyer, 2023).

A decrease in diversity in mortality causes, indicating that less varied causes of mortality have occurred, is likely caused by an increase in the proportion of deaths due to the most common causes. Although this means that the previously mentioned efficiencies of scale may become more useful and diagnostic uncer-

tainty might be reduced, this is not necessarily a desirable outcome as it suggests a failure of public health measures to address the most common causes of death. Diversity in mortality causes has not been observed to decrease in any population since the 1980s (Izsák, 1986, 1988; Hunter et al., 2003; Izsák, 1993a; Bergeron-Boucher et al., 2020; Trias-Llimós & Permanyer, 2023).

Existing research on diversity in causes of mortality has focused on trends over time. At the country level, the causes of mortality which affect the populations of relatively low-mortality countries, including the UK, have been shown to have become more varied in the 21st century (Bergeron-Boucher et al., 2020; Trias-Llimós & Permanyer, 2023). Across countries, diversity in causes of mortality has increased alongside increasing life expectancy over time and has been associated with falling variation in age at death (Bergeron-Boucher et al., 2020). Limited research has reported on trends and disparities in diversity in causes of death within subpopulations (Trias-Llimós & Permanyer, 2023), while neither the relationship between diversity in causes of mortality and life expectancy or diversity in causes of mortality and variation in age at death have been examined at the subpopulation level and, more generally, cross-sectional comparisons of these measures have not formally been examined. Furthermore, little is known about the effect of the COVID-19 pandemic on diversity in mortality causes.

1.3 The importance of measuring health outcomes at the subpopulation level

Subpopulations are defined in this thesis as groups of individuals within a nation defined by meaningful individual- or area-level characteristics. Understanding and addressing disparities in health outcomes between such subnational groups is one of the central principles of the field of population health (Kindig & Stoddart, 2003; Etches et al., 2006). Inequalities in health have been studied between subpopulations divided by such characteristics as: individual- and area-level socioeconomic position, age, educational attainment, race, and sex (Nazroo, 2003; Hartley, 2004; Silles, 2009; Brown et al., 2019; Kraft et al., 2020). Tackling inequalities between such subpopulations is the aim of many modern public health policies (Exworthy et al., 2003).

In this thesis, the population of Scotland is divided into subpopulations by: age at death, area-level income deprivation, an area-level urban-rural indicator, and the underlying cause of death faced by each individual (particularly those who died from COVID-19). In all analysis in this thesis, including in all breakdowns by subpopulation, the population of Scotland is split by sex. The following paragraphs briefly summarise existing knowledge relating to these distinctions.

For the most part around the world, males tend to face worse health outcomes than females. In the study of diversity in mortality causes, it is especially important to study males and females separately because a variety of causes of ill health and mortality are faced with a different prevalence[2] by men and women and some are faced solely by one sex (McIsaac & Wilkinson, 1997; Warnecke et al., 2008; Hosseinpoor et al., 2012).

At different ages, the risk of mortality from various causes of death changes dramatically (Wingard et al., 1989; Kuk & Ardern, 2010). Examining changes in diversity in mortality causes at different ages can inform the focus of targeted public health activities. It has been suggested that as populations live to greater ages, they are likely to face a greater variety of causes and that aging populations have driven diversification in causes of mortality in low mortality nations in the 21st century (Bergeron-Boucher et al., 2017; Bergeron-Boucher et al., 2020). Through examining diversity in mortality causes among deaths at different ages, it is possible to examine whether this assertion is true.

Socioeconomic and deprivation-related inequalities in health, in Scotland, and in the United Kingdom as a whole, are especially well studied (Smith et al., 1990). They have become a particular focus of researchers and policy makers following adverse trends in the 2010s which have been linked to various influences, although austerity policies have been blamed in particular (Holland, 2017; Miall et al., 2022).

[2]In this thesis the term “prevalence” is used to describe the proportion of the population whose deaths are attributed to a cause of mortality. This is expressed in later chapters in terms of the percentage of all-cause age-standardised mortality rates which are attributed to each cause.

Urban-rural inequalities in health are sometimes less clear-cut than those related to socioeconomic position in all-cause measures of population health (Hartley, 2004; Levin & Leyland, 2006). However, various cause-specific inequalities have been observed between urban and rural populations, suggesting that an analysis of diversity in mortality causes may be informative regarding differences in health between these areas (Levin, 2003; Levin & Leyland, 2005).

In order to address disparities in health between subpopulations it is necessary to gain knowledge on how these differences manifest as well as the underlying reasons for differences in health (Exworthy et al., 2003). The role of health determinants is an important aspect of public and population health research. However, the aims of this thesis focus on the measurement of health outcomes. Understanding how diversity in mortality causes varies between subpopulations has the potential to offer insight into the relative health of these groups (Bergeron-Boucher et al., 2020; Trias-Llimós & Permanyer, 2023). While study of diversity in mortality causes at the level of subpopulations is limited, previous research has shown higher diversity among those with lower educational attainment in the USA (Trias-Llimós & Permanyer, 2023).

1.4 Scotland as a study population

Health outcomes in Scotland have long lagged comparator nations, even when factors such as relative levels of deprivation are taken into account (Walsh et al., 2017). As a result, Scotland has been colloquialised as the "sick man of Europe" (Norman et al., 2011; Mittra et al., 2019). As well as having relatively poor health compared to other countries, the population of Scotland has also experiences relatively high inequalities in health (McLoone & Boddy, 1994; Norman et al., 2011). Previous research has established the causes of death which are factors in Scotland's comparably poor health: historically mortality rates due to drug and alcohol deaths, suicide, violence, cardiovascular disease, stroke, and cancer have all been relatively high (McCartney et al., 2011). However, the reasons why rates of these causes death are worse in Scotland than in other countries are more intractable (McCartney et al., 2012). Various hypotheses have been proposed for the comparatively poor health of the Scottish population including: the effects of deindustrialisation, especially important in Scotland where a large section of the population were historically employed in industrial professions; and deprivation,

which has for a long time been relatively high in Scotland compared to other nations (Whyte & Ajetunmobi, 2012). These "upstream" factors are thought to produce "downstream" effects such as poverty and unstable employment which interact with various other influences, including, for example, low levels of vitamin D associated with Northerly latitudes, to produce poor health outcomes (McCartney et al., 2011). Which of these influences, or which combination of influences, on health are responsible for the historic poor health in Scotland is still under debate in the literature. In the 21st Century, some progress has been made and health in Scotland has improved relative to other nations (Whyte & Ajetunmobi, 2012). However, this narrowing of the gap between Scotland and other nations has been put in danger of reversing due to slowing improvements in health in Scotland since the 2010s (Miall et al., 2022; Fenton et al., 2019a).

Scotland benefits from an universal, publicly funded health care system, the National Health Service (NHS). The NHS has, traditionally, has offered a strong generalist, primary care system with general practitioners (GPs) the first port of call for individuals seeking treatment. The UK health system, of which Scotland's NHS is a part, has fewer specialist healthcare professionals than comparable nations (McKee et al., 2021). This structure gives the ability to treat and diagnose a wide variety of conditions on the front line but means there can be issues of limited capacity for more specialised follow-up care. The strong focus on generalist care in the UK may be a benefit in the face of diversifying causes of mortality which require a more holistic approach. In the 2010s, the NHS in Scotland, and across the UK, has been under significant financial pressure. Cost-cutting measures, implemented through the austerity policies of successive right-wing governments, have left the service under strain (Kerasidou & Kingori, 2019). This lack of funding has been blamed for the UK being ill-prepared for the COVID-19 pandemic and the subsequent high mortality rate of the virus (Mellish et al., 2020). Austerity has caused Scottish healthcare system to fall behind comparable nations in terms of health care infrastructure and staffing levels (Maynard, 2017). Accessing primary care has become increasingly challenging because slowly rising GP numbers are being outpaced by an increasing patient load (Prowle & Harradine, 2014; Tucker et al., 2022; The Health Foundation, 2023).

The use of diversity in causes of mortality in population health has been promoted as a way to better understand patterns of mortality and to assess potential pressures on society and especially healthcare systems (Bergeron-Boucher et al., 2020; Trias-Llimós & Permanyer, 2023). A better knowledge of mortality patterns is important across all nations but is, perhaps, most advantageous in populations in

which health outcomes and improvements to health are lagging comparable groups. In addition, the importance of understanding the patterns of mortality in Scotland has become more pressing given the the adverse trends in various health outcomes in the late 2010s. Examining and understanding patterns of mortality in Scotland in this period may be especially important as Scotland may be a bellwether for worsening health outcomes across the UK and other nations (The Health Foundation, 2019). Understanding whether adverse trends have come alongside increases in diversity in mortality causes, or whether diversification observed at the UK level by Bergeron-Boucher et al. (2020) has reversed, can help to inform policymakers and researchers regarding how pressures on the healthcare system may manifest.

1.5 Thesis structure

In the following chapter, Chapter 2, I carry out a review of the relevant literature. It begins with a section in which I examine the role of diversity in mortality causes within the body of population health measures. This is followed by a history of the measures of diversity used in this thesis, during which the advantages conferred by their use are explained. A structured review is then presented of research which has measured diversity in mortality causes and the methods and findings of these studies are discussed. Finally, there is exploration of the gaps in the previously discussed literature and a presentation of the motivation for this thesis alongside the research questions answered within this work.

Chapter 3 presents a description of the general methods employed in this thesis as well as a description and discussion of data sources used throughout this work.

Chapters 4 to 7 report the findings of this thesis. These chapters are structured as follows: each begins with a background to the specific subject of that chapter and the specific research questions and objectives the chapter addresses, followed by details of the methods employed and a description of the results of the research. Finally, the findings of each chapter are summarised and interpreted in the context of the literature and their implications are discussed.

In Chapter 4, I examine trends in the diversity of causes of mortality across Scotland and in Scottish subpopulations over the period 2001 to 2019. I propose a measure of the contribution of causes of mortality, "additive value", to diversity. I use this to examine the causes of mortality which, through becoming relatively more or less common, have caused changes in diversity in mortality causes.

In Chapter 5, I explore the use of measures of diversity to assess variation in ages at death within the Scottish populations and Scottish subpopulations. I then examine trends and tendencies in lifespan diversity in Scotland and subpopulations in Scotland in the years 2001 to 2019. Finally, I examine the relationships between diversity in causes of mortality and both life expectancy and lifespan diversity, doing so over time and cross-sectionally between subpopulations.

I examine the impact of the COVID-19 pandemic on diversity in mortality causes in Scotland in Chapter 6. This chapter measures diversity in causes of mortality across the years 2020 and 2021 in Scotland, and in each month in these years, comparing this to trends in previous years. I explore seasonality in mortality cause diversity and use seasonal patterns to inform forecasts of monthly diversity in 2020 and 2021.

In Chapter 7, I examine diversity in contributory causes of mortality in Scotland and compare it to the diversity of underlying causes of mortality as examined in previous chapters. These are causes recorded alongside the underlying cause of death which are thought to have contributed to or hastened an individual's death. In this chapter, I then examine the diversity of contributory causes of mortality recorded alongside COVID-19 and compare this to the diversity of causes recorded alongside other leading causes of mortality in 2020 and 2021 to assess the relative health of individuals whose death was attributed to COVID-19 in these years.

Finally, in Chapter 8, I bring this thesis to a close by reflecting on the findings presented in previous chapters. I discuss the implications of this research in the context of the public and population health literature. I follow this by considering some of the strengths and limitations of the methods employed in this thesis. I discuss possible avenues of future research and present two case studies of initial work to develop methods related to the research carried out in this thesis. Finally I present the overall conclusions of this thesis.

Chapter 2

Literature Review

2.1 Chapter overview

The purpose of this narrative review of the literature was to explore the background of the measurement of variation in causes of mortality in population health; to describe the formation of the measures of diversity used in this thesis to assess this variation; and finally, to examine previous studies of diversity in causes of mortality. As such, there are three main sections of this review and a fourth section summarising the literature gaps which this thesis aimed to address:

Section 2.2 begins with a discussion of the definition of population health and the objectives of the field. This is followed by an overview of the history of population health indicators founded on the measurement of mortality and a discussion of the divide between all-cause and cause specific measures. Finally, the epidemiological transition is discussed in the context of variation in mortality causes.

In Section 2.3 a history of measures of diversity is presented. These measures are used to assess variation in distributions in a wide range of fields including ecology and computer science. This section discusses the advances in the field of diversity which led to the creation of the Reeve et al. (2016) framework for diversity which are applied to examine diversity in mortality causes and in ages at death in Chapers 4 to 7.

An exploratory review of previous studies of diversity in mortality causes is carried out in Section 2.4. This field of study is small and diverse with a variety of methodologies used to describe variation in the distribution of causes of mortality in a number of different populations.

Finally, in Section 1.5 the gaps in the literature related to diversity in mortality causes and the resultant research questions addressed in this thesis are discussed.

2.2 Measures of population health

2.2.1 What is population health?

The study of health at the population level rather than at an individual level has a long history (Declich & Carter, 1994). However, the use of the term "population health" to describe this study is relatively young having been popularised in the late 1990s and early 2000s (Arah, 2009). There is debate in the literature over whether population health is a concept of health or a field of study of health determinants (Kindig & Stoddart, 2003; Young, 2004; Kindig, 2007; Ferrara & Albano, 2022). Kindig and Stoddart (2003) propose a definition of population health focused on the measurement of outcomes suggesting that as a concept population health is "the health outcomes of a group of individuals, including the distribution of such outcomes within the group". However, they describe the field of population health as encompassing the patterns of health determinants and policies and interventions that foster and bring about these outcomes.

In addition to the debate over the definition of population health, there is disagreement on how it relates to the more historic concept of public health (Kindig, 2007). These terms may be thought of as overlapping if one takes the core concept of each term: the "health of a population" and the "health of the public" (Wallace et al., 2021) . However, in the 21st century, public health is often considered more closely related to the activities undertaken by society to assure conditions that allow people to be healthy (Turnock, 2012). On the other hand population health, in both definitions described above, is centred around the measurement of health outcomes in a given population. Some view the field of population health as en-

compassing and expanding upon the scope of public health science while taking a focus on the "upstream causes" of ill-health (Valles, 2018). No unified definition of the term "population health" has been accepted broadly, and many definitions are interdependent and are used interchangeably in much of the literature (Brackin et al., 2015; Ferrara & Albano, 2022). The study of population health is nonetheless borne out of similar motivations regardless of definition: the monitoring and improvement of health across populations and the reduction of inequalities in health within these populations (Etches et al., 2006; Krieger, 2012).

There are more philosophical debates in the literature pertaining to the concept of health itself. Many argue for "wellbeing" approaches to be adopted meaning that, on the individual level, health is not simply a matter of avoiding disease but of maximising quality of life (Sears et al., 2014; Bucks et al., 2018; Roy et al., 2018). Others suggest the use of a "life course" framing, particularly in the study of inequalities, founded on the principles that disparities in health are rooted in socially patterned factors such as socioeconomic, environmental and physical exposures from a young age and throughout life (Kuh et al., 2003; Kuh et al., 2013; Jones et al., 2019). Measures of population health based on mortality data, as used in this thesis and discussed further below, are considered under both wellbeing and lifecourse theories to be insufficient in isolation to address these concerns. There is, however, general agreement in the literature on the utility of such measures as proxies for population health (Reidpath & Allotey, 2003; Patton et al., 2009; Tapia Granados & Ionides, 2017).

2.2.2 Measurements of mortality in the history of population health outcomes

Despite disagreement in the field on which concept of population health should be adhered to, most definitions encompass broadly similar sets of measures to describe health outcomes. The stock of measures in use in the present day was formed over centuries. Some of the earliest measurements of the health of populations were facilitated by the collection of birth and death data in churches in London in the 1500s (Declich & Carter, 1994; Etches et al., 2006). These were formalised and expanded in later years, in conjunction with population data these records allowed for the calculation of rates of mortality, as the number of deaths in a population divided by the population size. These mortality rates were then

used as the basis for more complex measures of population health. For example, life tables were first accurately calculated in 1815 using parish data in England and were developed further as sources of mortality and population data became more universal (Grebenik, 1991; Etches et al., 2006).

Mortality rates and life expectancies, which are extracted from life tables, remain some of the most widespread instruments for assessing population health in the present day. Technical developments such as sex and age standardisation, in the case of mortality rates, and advanced modelling of mortality rates, in the calculation of life tables, have been developed over time (Hoem, 1977; Hougaard, 1984; Chi, 1988; Hsieh, 1991; Wilmoth et al., 2022). These developments improved the utility and applicability of mortality rates and life expectancy measures, expanding the questions they are used to address and the populations and subpopulations in which they can be calculated.

The further development of population health measures was in large part driven by increased availability of data on the health of living populations (Macran et al., 2003; Etches et al., 2006). Through combining information on health during life with mortality data, a plethora of summary measures of population health are used in the literature to represent the health of a population as a single number (Murray et al., 2000; Roy et al., 2018). These measures, generally, describe either health expectancy or health gaps. Measuring health expectancy within a population involves quantifying the amount of time the average individual spends at full health. Health gap measures are extensions of measures such as years of life lost beyond the impact of premature mortality into the amount of time spent with health less than an ideal state, an example is disability-adjusted life years (DALYs) (Murray et al., 2000; Etches et al., 2006).

The range of mortality-based measures used in the population health literature is ever expanding (Muszyńska-Spielauer & Luy, 2022; Permanyer et al., 2022). As is the variety of research questions to which they are applied. The measures mentioned to this point all provide a single number to describe the health of a population. These numbers, generally, represent what is effectively the population mean (life expectancy) or a population-wide rate, ratio or risk, arguably, this presents a limitation of these population health measures. To repeat Kindig and Stoddart's (2003) description of population health it is: "the health outcomes of a group of individuals, including the distribution of such outcomes within the group". Etches (2006) describes an ideal indicator of population health suggesting that it

should include measures of the "Central tendency (e.g., mean, median, etc.) and distribution". Kindig and Stoddart's (2003) definition arguably refers, at least in part, to inequalities in health outcomes between sub-groups within a population and the measures discussed here are used to effectively examine such inequalities (Mackenbach, 2002). However, in providing estimates of central or population-wide tendencies measures such as mortality rates and life expectancy, as well as most of their derivatives, are for the most part not directly informative regarding distributions of health outcomes within the population.

Mortality-based measures which assess the distribution of health outcomes are used to address this limitation of traditional population health indicators. The most commonly studied are lifespan variation measures, which describe variation in the distribution of lifespans in a population (van Raalte & Caswell, 2013). They have been framed as a more complete way to measure inequalities than simple differences in life expectancy, as they are able to describe the heterogeneity of ages at death within a population (Seaman et al., 2016b). Measures of lifespan variation have been adapted to measure of the variability in health expectancy through measuring as "healthy lifespan inequality" (Caswell & Zarulli, 2018; Caswell & Van Daalen, 2021; Permanyer et al., 2022).

The study of variability in lifespans and healthy lifespans is centred on the examination of inequalities in health. To simplify the argument, equity in health outcomes is greater in a population in which the variation in ages at death is minimised. This definition of health inequalities explicitly accounts for absolute differences in lifespan across a population. Measures of lifespan variation can therefore be thought of as measures of the total level of inequality in health within a population (Seaman, 2017). More commonly, health inequalities are discussed in the literature as being preventable differences in health between different groups within a population. Closing these unfair health gaps is a focus of a great deal of public health research (McCartney et al., 2019). These two definitions of health inequalities are not in conflict, it is clear that holding all else equal reducing inequalities between sub-groups within a population will in most cases reduce overall inequalities within that population.

Other sources of variability in distributions of health outcomes are less closely related to population level inequalities. The causes of mortality faced by a population form a distribution, and variation within this distribution can be expected to differ between populations. However, Bergeron-Boucher et al. (2020) observe

that “The degree to which causes of death vary across individuals is not generally a matter of health equity”. Variation in the distribution of causes of mortality has other implications for population health which are discussed further in Section 2.4 of this literature review.

2.2.3 The all-cause, cause-specific dichotomy

“The large number of ICD-10 codes raises the question of how many of these codes should be distinguished when tabulating, analyzing and publishing mortality trends.”(Mitratza et al., 2019)

A pervading theme across the use of mortality-based measures of population health is a tendency to examine population health through either an all-cause or cause-specific lens (Beltrán-Sánchez et al., 2008; Wallace, 2022). All-cause measures allow for examination of overarching trends in mortality within populations, while cause-specific measures are used to assess the burden associated with individual diseases and conditions. In some ways, all-cause and cause-specific measures are limited by their focus on either broad patterns of mortality (all-cause) or fine-grained changes (cause-specific). This is addressed, to some extent, in some studies, by the use of decomposition methods to investigate the individual causes of mortality that have driven changes at the all-cause level (Seaman et al., 2019).

The sheer range of individual causes that may be recorded in a population can complicate the study of trends in cause-specific mortality measures. The most widely adopted mortality classification system is the WHO’s International Statistical Classification of Diseases and Related Health Problems. Currently the 10th revision (ICD-10) of this classification is used by most countries having been introduced in 1990 (World Health Organization, 2019). There are more than 1700 individual ICD-10 three-character codes each refering to a single condition or disease. Three-character codes are the primary breakdown of ICD-10 code used in population health studies to identify or group causes of mortality. Some ICD-10 three-character codes refer to conditions and diseases which never appear as the underlying cause of death (the ICD-10 system encompasses description of morbidity as well as mortality). All the same, addressing all codes at the three character level in any one study is clearly challenging due to the overwhelming number (Mitratza et al., 2019). Researchers run the risk of not seeing the forest for the trees and

missing important trends. For this reason many studies examine single causes of death, potentially limiting analysis, or group individual three-character codes into broader categories (Wallace, 2022). However, grouping causes may mean that trends important to health policy are missed (Mitratza et al., 2019).

The existence of a dichotomy between the use of all-cause or cause-specific measures may be contested. Many influential studies have used a combination of these methods to effectively describe patterns and burdens of disease. These include studies that have evaluated differences across the world such as those derived from the Global Burden of Diseases, Injuries, and Risk Factors Study (GBD) as well as those that examine inequalities in health at the level of subpopulations (Singh et al., 2013; Toch-Marquardt et al., 2014; Naghavi et al., 2015; Cha & Jin, 2020; Global Burden of Disease Collaborative Network, 2020). These analyses present detailed cause-by-cause breakdowns of historical mortality patterns and projections for the future and are of great value to public health planning. They also offer indispensable information on differences and inequalities in health by subpopulations (Hosseinpoor et al., 2016). However, as patterns of mortality change through successful public health efforts and societies reduce rates of mortality, often through reducing the prevalence of the most common causes of mortality, it has been suggested that attention must be paid to ever widening ranges of causes (Bergeron-Boucher et al., 2017). This raises some of the concerns related to cause-specific measures highlighted above and means it may be advantageous to use indices that assess variation in the distribution of mortality causes alongside all-cause and cause-specific measures.

2.2.4 The epidemiological transition

Studies of all-cause and cause-specific mortality in the context of population health have captured important evolutions in the patterns of disease burden across the globe. These evolutions are described as “epidemiological transitions” and the literature related to their study is vast and, similarly to many concepts in population health research, contains many disagreements and contradictions (Andreev, 2019). Epidemiological transition theory was first promoted by Omran (1971). Under Omran’s theory of epidemiological transition populations pass through several “stages” which describe different epidemiological and demographic conditions. These stages occurred at different times and endured for variable periods in different countries and world regions.

Omran (1971) describes the first stage of epidemiological transition as the "age of pestilence and famine" and associates it with high mortality rates due to epidemics, war, and poor living conditions. Life expectancies in this stage were low and most deaths occurred due to infectious or parasitic disease (Santosa et al., 2014). Most nations moved out of this stage through medical breakthroughs and what were, effectively, public health advancements though often were not viewed through this lens at the time. "The age of receding pandemics", the second stage of transition, was also characterised by high mortality due to infectious diseases although it was a period of reduction in their prevalence. During this era, non-communicable diseases came to the fore causing increased proportions of total deaths. Omran's original theory ends with the third stage "the age of degenerative and man-made disease". Infectious diseases were drastically reduced in prevalence and diseases of the cardiovascular and cerebrovascular systems were dominant alongside cancers and chronic lung and metabolic diseases. Later updates to this theory by Omran introduced fourth and fifth stages. The fourth stage was associated with a reduction in cardio- and cerebro-vascular mortality thanks to improved care and recognition of risk factors. Unlike previous stages, the final stage is characterised by increases in prevalence of diseases, specifically the emergence of new infectious diseases such as HIV/AIDS and resurgence in more historic infectious diseases such as cholera, malaria and tuberculosis (Omran, 1998; Santosa et al., 2014). This final stage is proposed to have occurred with some regional variation, for example HIV/AIDS emerged as a significant cause of mortality across the world whereas increased prevalence of more historic diseases has been mostly limited to low/middle income countries. Resistance to existing medical treatments, difficulties in ensuring universal vaccine uptake and under-developed public health systems are among the factors linked to increased prevalence of infectious diseases in these areas (Omran, 1998; Alvarez-Uria et al., 2016; Vilar-Compte et al., 2017).

Omran's theory of transition has been critiqued for the over-simplification implicit in this model and for its foundation on mortality data and trends in Western countries (Mackenbach, 1994; Santosa et al., 2014). Omran's model generalises across subpopulations, whether split by race, socioeconomic class, or sex, is raised as problematic. Patterns of mortality differ starkly between such groups and changes and improvements to mortality occur at different rates. Therefore any epidemiological transition might be expected to manifest differently between subpopulations. Santosa et al. (2014) however, conclude that despite these concerns,

Omran's theory of epidemiological transition holds true for certain populations and can be generalised to adapt to a wider range of situations. The theory especially requires adaptation in regions which began "transitions" in more recent periods such as low-middle income countries (Andreev, 2019; Sudharsanan et al., 2022).

The aim of theories of epidemiological transition is to aid in understanding of changes in population dynamics. Observations related to mortality mostly attempt to explain reductions in all-cause mortality through cause-specific changes. Stages one through four of Omran's theory are mostly concerned with describing changes in the dominant cause or causes of mortality. Initially, infectious diseases make way for non-communicable diseases, especially cardiovascular and cerebrovascular diseases which, in turn, become less prominent in the fourth stage. These changes came alongside near-continuous reductions in the overall burden of mortality. This reduction in mortality, driven by the most common causes, has implications for variation in the distribution of causes of mortality within populations. In a period when most deaths occur due to a single cause or a small number of causes, variation in the distribution of causes of mortality would be low. However, as the most common causes of mortality affect a smaller section of the population, deaths can be expected to be spread more evenly across causes and therefore an increased variety of causes would be expected to occur in the population. Bergeron-Boucher et al. (2017) observe this effect in a study of the causes of old-age mortality in Canadians from 1979-2011. They find that the leading causes of mortality, specifically cardiovascular diseases, reduced in prevalence consistent with Omran's fourth stage of epidemiological transition. However, with some exceptions, this was not accompanied by an increase in the rate of deaths associated with other causes. They suggest this is "likely to lead to a greater diversification of causes of death" as fewer deaths occur due to the most common causes and deaths become more evenly spread across other causes of mortality (Bergeron-Boucher et al., 2017).

This framing of increased variation in the distribution of causes of mortality as increased *diversity* of causes is not novel, having been originally posited in the literature in the 1980s (Izsák, 1986). The use of the term *diversity* by Bergeron-Boucher et al. (2017) and a later study by Bergeron-Boucher and colleagues have, however, reintroduced a concept which has potential for the measurement of variation in the distribution of causes of mortality in the study of population health (Bergeron-Boucher et al., 2020). The concept of diversity in mortality causes is explored further in the next sections of this chapter.

2.3 A brief history of diversity

The concept of diversity is key to the study of many biological systems (Petchey & Gaston, 2006; Tucker & Cadotte, 2013; Chao et al., 2014b). Measuring diversity fundamentally involves describing variation within a distribution of characteristics. Diversity measures quantify the presence (or absence) and relative abundance of types (the characteristics: species, phenotypes, *causes of mortality*, etc.) within a community (the distribution: a rainforest, the gut microbiome of an individual, *a population of people*, etc.) (Simpson, 1949). In the study of diversity we may be interested in how characteristics vary over a total population, a metacommunity, or within a smaller subset of the population, a subcommunity. As an example, consider the population of Scotland (metacommunity); composed of council areas, datazones, or socioeconomic deprivation quintiles (subcommunities). Table 2.1 defines, and Figure 2.1 illustrates, these diversity-related terms alongside others used in this section and this thesis as a whole. If the "types" shown in Figure 2.1 were causes of mortality, all individuals would have faced a different cause, the distribution in causes of mortality would be completely even and diversity under the Reeve et al. (2016) framework (used in this thesis and discussed further in upcoming sections) would be maximised. If all individuals were to have faced the "blue" cause of mortality diversity, the distribution would be completely dominated by this cause and diversity under the Reeve et al. (2016) framework would be minimised.

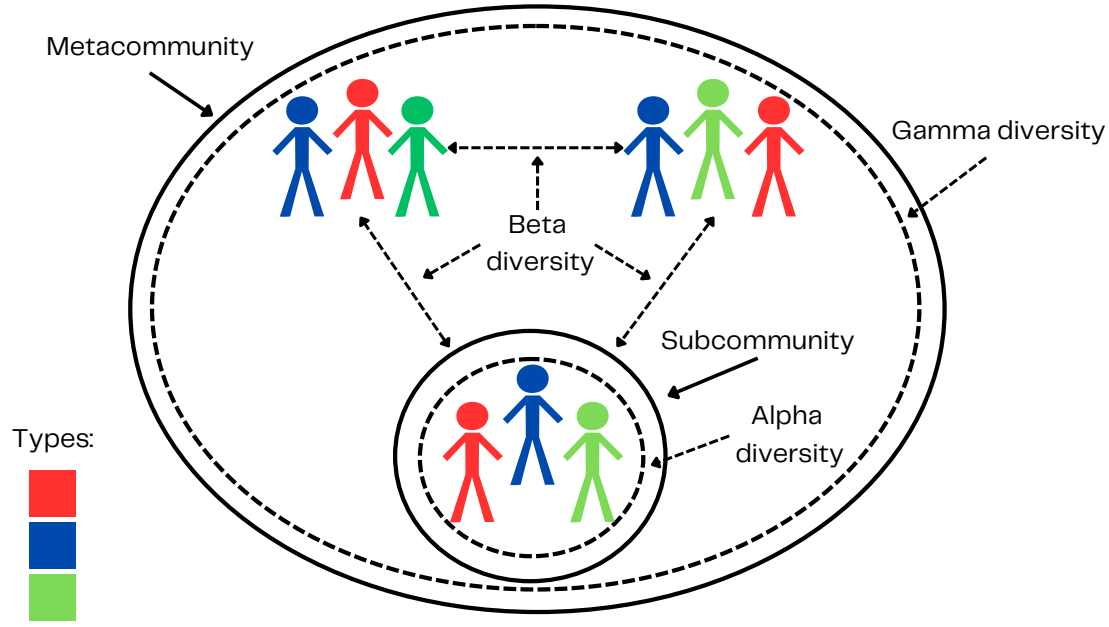


Figure 2.1: A graphic illustrating terms used in the study of diversity in this thesis. Types refer to a differentiating characteristic which defines individuals in the group, for example causes of mortality.

Table 2.1: Terms related to diversity used in this thesis and their definition.

Term	Definition
Type	A differentiating characteristic which defines individuals in a group.
Metacommunity	The total population of interest.
Subcommunity	A section of the metacommunity.
Alpha diversity	The variation of types within each subcommunity.
Beta diversity	The variation of types between subcommunities. In other words the variation across the metacommunity not explained by diversity within each subcommunity
Gamma diversity	The variation of types within the metacommunity.
Similarity	The degree of relation between types.
Naive-similarity	An assumption that types share no relation and are completely distinct from one another.

Many measures of species diversity have been proposed in the literature, they have been used in variety of disciplines. This has caused some confusion as often different terms have been used to describe the same measures (Noss, 1990). Distilling this plethora of approaches for the measurement of diversity into a formalised framework has been an aim of the field for many years (Hill, 1973; Jost, 2006; Jurasinski et al., 2009; Moreno & Rodríguez, 2010; Reeve et al., 2016). Historically, measures of diversity were often developed separately. The variety of measures, and the lack of a clear discipline wide-framework, meant that some equivalent measures have been used in conjunction without acknowledgement of their similar underlying properties (Magurran, 2021). Conversely, because the properties of some measures of diversity differ, opposing conclusions may be drawn from the application of different measures to the same distribution (Reeve et al., 2016).

Not only do these traditional measures overlap and lack categorisation but the values produced by most measures are unitless. The lack of units can make comparing diversity between populations problematic. Jost (2006) uses the example of two communities, one with eight equally abundant types and another with sixteen. The second community can clearly be thought of as twice as diverse as the first, as it contains twice as many types. However, the Shannon entropy - one of the most commonly used measures of diversity - of the communities described above is three and four respectively (Shannon, 1948; Bromiley et al., 2004; Bergeron-Boucher et al., 2020). At first glance this may imply that the second community is only one third more diverse than the first rather than twice as diverse. This does not detract from Shannon entropy which is a useful measure of diversity valued in the field of ecology. Indeed Jost (2006) argues that it is the “most profound and useful of all diversity indices”. However, because it is designed to describe uncertainty rather than diversity, its original formulation leads to results, such as those described here, which may be misleading. This problem exists across many, more historic, measures of diversity. However, solutions have been proposed (including by Jost (2006)) and the formation of measures which account for this shortcoming are described in the next section.

2.3.1 Hill numbers and the "effective number of types" formulation of diversity

In response to the limitations described above Hill (1973), building on the work of Renyi (1961), proposed a unifying notation for the measurement of diversity in 1973, later expounded by Jost (2006). Under this system, diversity is presented as the "effective number of species" (or here, types) within a community. This is the theoretical number of equally abundant types needed to produce a given value of diversity. Any value of diversity can in theory be produced by a range of communities which differ both by the number of types present and in the abundances of those types. Within this theoretical set of communities is one in which all types are equally abundant. Theoretically, despite differing in structure, all other communities in this set are equivalent in diversity to this equally abundant community, therefore they can be described as having the same effective number of types. The generalised formula for the calculation of diversity under Hill's framework is shown below in Equation 2.1. Here q is the order of diversity; R is the number of types in the distribution; and p_i is the proportional abundance of the ith type.

$$ {}^{q}D = (\sum_{i=1}^{R} p_i^q)^{1/(1-q)} \tag{2.1} $$

The effective number of types formulation is used to give a unit to diversity indices which makes it easier to compare between communities in a meaningful way. Effective number formulations are also constructed so that their mathematical properties are intuitive (Mitchell, 2019). Under effective number formulations, the diversity of a community containing S equally abundant types is S. Therefore, doubling the number of equally abundant types doubles the measurement of diversity. This is not the case for "raw" measures of diversity such as Shannon entropy as discussed above.

Hill (1973) integrated the effective number formulation into a novel family of measures diversity which are known as Hill numbers (${}_qD$). Hill numbers are differentiated, as measures of diversity, by the order of diversity - or q - which can take any value between 0 and ∞ and determines the significance placed on the abundance of types. At q = 0 all types are considered equal regardless of abundance,

therefore diversity at $q = 0$ is a count of the number of types in a community. This is equivalent to a widely used ecological concept known as the "species richness" of a community. Higher q values weight abundance more strongly and are described as more conservative. Diversity at $q = \infty$, the most conservative Hill number, essentially considers only the abundance of the most common type. It is an effective number formulation of the Berger-Parker index which is a measure concerned with the dominance the most common type.

2.3.2 Similarity of types

The development and formalisation of Hill numbers answered both the need for a unified framework for measures of diversity and the desire for indices with a meaningful unit (Routledge, 1979; Chao et al., 2014a). A key limitation of these measures, however, is that they are only applicable when it can be assumed that all types are distinct from one another. In reality this is often not true and within a system it is common for certain types to be more closely related to each other. For example, in a distribution of lifespans, individuals who died at ages 75 or 76 can be understood to be more similar to each other than individuals who died at ages 0 or 75.

Quantifying similarity between types in the measurement of diversity was first popularised by the work of Rao (1982a, 1982b). In 2012, Leinster and Cobbold (2012) built on Rao's work to produce a family of similarity-sensitive measures. These measures can be varied in their sensitivity to rare species by q, as described for Hill numbers above. Leinster and Cobbold (2012) introduce similarity through the use of a matrix of size SxS, where S is the total number of types in the community. This matrix is used to represent the degree of similarity between each pair of types. Under Leinster and Cobbold's (2012) proposed measure, similarity can be derived from any meaningful notion of relation between types. Using the matrix described above, the expected similarity between each type and an individual chosen at random from the community can be quantified and included in the calculation of diversity. This is described as the relative abundance of types that are similar to a given type (Leinster & Cobbold, 2012).

The generalised formula[1] for Leinster and Cobbold's (2012) measure of diversity is shown in Equation 2.2. Here ${}^{q}D^{Z}$ is similarity-sensitive diversity, q is the order of diversity, p_i is the proportional abundance of the ith type. $\mathbf{Zp}$ is a relative abundance of types similar to the ith type, quantified by Equation 2.3 where Z_{ij} is the degree of similarity between type i and j under the assumption $0 \geq Z_{\mathrm{ij}} \leq 1$ (Leinster & Cobbold, 2012).

$$ {}^{q}D^{\mathbf{Z}} = (\sum p_i (\mathbf{Zp})_i^{q-1})^{1/(1-q))} \tag{2.2} $$

where

$$ (\mathbf{Zp})_i = \sum_{j=i}^{S} Z_{ij} p_j \tag{2.3} $$

The development of Leinster and Cobbold's (2012) similarity-sensitive measure of diversity allowed a wide range of concepts in the field of diversity to be condensed into a single formula. Leinster and Cobbold's method is flexible in both the indices of diversity it measures (through varying q) and in the variety of systems that its definition of similarity can describe.

2.3.3 The Reeve et al. framework and the measures of diversity used in this thesis

Building on the work of Renyi (1961), Hill (1973), Jost (2006, 2007), and Leinster and Cobbold (2012), Reeve et al., in 2016, proposed a unified framework for the measurement of diversity. This framework generalised the previously proposed methods making them applicable to a wider range of diversity indices. It also addressed the problem of partitioning in the measurement of diversity. The partitioning of diversity refers to whether diversity is calculated within each subcommunity (alpha diversity), between subcommunities (beta diversity) or across the metacommunity (gamma diversity). The complexities of partitioning diversity are discussed in depth in the description of the framework by Reeve et al. (2016) and in the thesis of Mitchell (2019) who led the creation of an R package (rdiversity) to measure diversity under the framework. The measures proposed in the Reeve

[1] This formula holds for all values of q other than 1 and $\inf$ at these values of q specific formulas are used which are detailed by Leinster and Cobbold (2012).

et al. (2016) framework encompass metacommunity and subcommunity measures of alpha, beta and gamma diversity. Within each of these measures *q* (known as the viewpoint parameter in this framework) can be varied from 0 to ∞ to alter the weighting of abundance, consistent with the previous descriptions of *q* values in this section. In the research carried out in this thesis, subcommunity normalised alpha diversity is measured, and so further description is limited to this measure. However, promising applications exist for the study of beta diversity in mortality causes, and these are discussed in Section 8.5.

Subcommunity normalised alpha diversity (${}^{q}\overline{\alpha}_{j}^{Z}$), referred to in this thesis from now on as simply normalised alpha diversity, is the diversity of each subcommunity in isolation. This measure is equivalent to the measure of similarity-sensitive diversity of a single community proposed by Leinster and Cobbold (2012). Normalised alpha diversity, at any *q* value, can take values between 1 and the total number of types in the subcommunity. Under all q values (excluding q = 0) this measure can only reach its maximum if all types are equally abundant.

Normalised subcommunity alpha diversity under the Reeve et al. (2016) framework, as used in this thesis, is the diversity of each subcommunity (*j*) within a metacommunity in isolation (For discussion of the distinction between metacommunities and subcommunities see Section 2.3.3). It is calculated under the general formula shown in Equation 2.4, where *M*(1-*q*) is the weighted power mean of the order 1 - *q*; $\overline{P_{.j}}$ is a vector of the relative abundance of each type in subcommunity *j* ($P_{.j}$) divided by the proportion of the metacommunity contained within subcommunity *j* (w_j) ; and $Z\overline{P_{.j}}$ is the relative abundance of types similar to type *j*.

$$ {}^{q}\overline{\alpha}_{j}^{Z} = M_{1-q}(\overline{P_{.j}}, (Z\overline{P_{.j}})^{-1}) \tag{2.4} $$

The framework of diversity proposed by Reeve et al. (2016) provides a unified set of measures of diversity with a meaningful unit which are flexible in their use and can be used with similarity of types considered. It addresses each of the shortcomings in traditional measures of diversity discussed previously in this section. For this reason, measures under this framework are used in this thesis. Further, the notation and concepts within this framework are constant across alpha, beta, and gamma diversities meaning there is potential for further study using measures of diversity calculated under comparable notation, using the same terms.

2.4 Diversity in the context of human mortality causes and population health

The study of diversity in mortality causes involves utilising measures of diversity to examine variation in the distribution of mortality causes within a population. In the measurement of diversity, causes of mortality are treated as "types". Greater diversity in mortality causes indicates greater variation in the causes of mortality within a population with, a likely smaller proportion of deaths caused by the most common causes. On the other hand, lower diversity indicates that causes of mortality vary less within a population.

The study of diversity in mortality causes is an emerging field with a small literature base. Exhaustive searches of the literature were carried out throughout the period of research upon which this thesis is based. The following search strategy was used in November 2022 to provide a final, comprehensive, view of previous studies. The EBSCO interface was used to search the MEDLINE database to identify studies which featured the terms "diversity" or "diversification" alongside the terms "mortality cause/s" or "death cause/s" in either the title or abstract while excluding titles which contained the terms: "genetic" and "dietary". This search returned 31 results; these results were screened by title to exclude any studies which did not directly study diversity in mortality causes. This final search of the literature search returned a paper which I authored in the completion of this thesis, this paper was excluded. Three studies of diversity in mortality causes were taken forward (Izsák, 1986; Bergeron-Boucher et al., 2020; Trias-Llimós & Permanyer, 2023))[2]. Publications referenced by these three studies and studies which had cited these publications in the literature were examined for further sources. Following this search a further five articles were found. Three further studies were identified in this reference and citation search, these studies were found to be unavailable in literature searches in numerous databases including EBSCO and Google Scholar (Izsák & György, 1991; Izsák, 1982; Izsák & Juhás-Nagy, 1984) . Upon corresponding with the authors, copies of these publications were very kindly supplied and included here. Overall, this thorough literature search resulted in eleven peer-reviewed studies which are discussed in this section. Tables 2.2 and 2.4 describe these studies and their key findings.

[2]The literature search carried out in November 2022 returned a conference mansucript submitted by Trias-Llimós and Permanyer published in May 2022, in January 2023 this research was published in a peer-reviewed journal. The peer-reviewed version of this manuscript was examined in this literature review.

This section begins with a summary of the key findings and insights of these previous studies in Section 2.4.1, followed by a discussion of the motivation for examining diversity in mortality causes in Section 2.4.2, and, finally, a discussion of the selection of a definition for mortality causes in the study of diversity in mortality causes in Section 2.4.3.

Table 2.2: A description of previous studies of diversity in causes of mortality. Where more than 2 countries were studied a full list of studied nations can be found in Appendix A.1

Study	Study population and period	Format of mortality data	Diversity measure(s)	Main conclusions
Measuring the secular changes in the concentration of death causes, *Izsák*, 1986	White Americans in 1968 and 1975, all analysis performed within 5 year age groups	Raw mortality counts of underlying causes of death. Fine-grained methods - selected individual ICD-8 codes. Separate examination of all causes and of significant groups of causes, based on ICD-8 categorisation.	Hill numbers, S(m) indices, and Q index	Increased concentration of mortality causes (i.e. falling diversity) among most age groups at less conservative measures, which weight rare causes more heavily. However, the concentration fell (diversity increased) at more conservative measures which analyse the dominance of leading causes. This indicates that the dominance of the most common causes fell alongside reduced variety in the range of rarer cause likely to occur. Recommends using multiple measures of diversity to capture and describe variation among both common and rare causes of mortality.
An ecological approach to causes of death in 14 countries: the Shannon index of diversity, *Ascaso Terrén, Canela Soler, Sentís Vilalta*, 1988 (Only available in Spanish)	14 countries in 1978	Cause-specific mortality rates for underlying causes of death, standardised to the Spanish population. Fine-grained methods - selected ICD-8 codes, categorised into 15 groups based on ICD-8 groupings.	Shannon entropy (various formulations) (Hill number/q value of 1)	Show that diversity in mortality causes differs considerably between various countries and divide countries into groups with low, medium and high diversity in mortality causes. Recommends the use of the Shannon entropy as a measure of diversity in mortality causes.
Comparative study of diversity indices on mortality data., *János*, 1982 (Only available in Hungarian)	Americans in 1974 and 1975, all analysis performed within 5 year age groups	Raw mortality counts of underlying causes of death. Fine-grained methods - selected individual ICD-8 codes. Separate examination of all causes and of significant groups of causes, based on ICD-8 categorisation.	Shannon's index, Hill numbers, S(m) indices, Q index, relative Brillouin Index	Compare various measures of diversity in the measurement of diversity in mortality causes. Find results are similar across most. Show diversity in mortality causes is among those aged 20-24 years old, reducing in older age.
Diversity studies on mortality data., *János, Juhász-Nagy*, 1982	Americans separated by age and race in 1974 and 1975, all analysis performed within 5 year age groups	Raw mortality counts of underlying causes of death. Fine-grained methods - selected individual ICD-8 codes. Separate examination of all causes and of significant groups of causes, based on ICD-8 categorisation.	Relative Brillouin Index, Shannon Index	Find higher diversity in cancer deaths than in deaths within most other cause groups. Diversity shown to peak in deaths at ages younger than 35 in most cases and reduce or plateau across older ages.

Table 2.3: (Cont.) A description of previous studies of diversity in causes of mortality. Where more than 2 countries were studied a full list of studied nations can be found in Appendix A.1

Study	Study population and period	Format of mortality data	Diversity measure(s)	Main conclusions
Secular Changes of the Concentration of Neoplasm Death Causes in the Male Population of Some Countries, *Izsák*, 1988	Males in 5 countries. 1963-1985 (different years of data in different countries)	Raw mortality counts of underlying causes of death. Fine-grained methods - selected individual ICD-7, 8, and 9 codes which describe cancers	S(m) indices	Diversity in causes of mortality among deaths due to cancer increased over time in each country over the study period.
Comparative Analysis of Death Cause Diversity Curves in Various Countries, *Izsák*, 1993	5 countries. All analysis performed within 5 year age groups. 1969 - 1987 (different years of data in different countries)	Raw mortality counts of underlying causes of death. Fine-grained methods - selected individual ICD-8 and ICD-9 codes which describe cancers and diseases of the circulatory system	Reciprocal Simpson index (Hill number/q value of 2)	Diversity in causes of mortality among deaths due to cancer are, for the most part, found to be higher among younger (age 40 and younger) an older (aged 60 and older) age groups than in deaths in middle age in all countries. Diversity in causes of mortality among deaths due to diseases of the circulatory system show a similar pattern to diversity among cancer deaths in European countries (Hungary, England and Wales, Norway, Finland) but relatively consistent increases observed from younger to older ages in Japan. For both cancers and circulatory diseases, sexes are examined separately, similar patterns are observed though they occur at different ages. Highlight considerable similarity in the patterns of mortality cause diversity by age across the countries they studied and observe that differences between males and females are also consistent across nations.
Concentration analysis of diagnostic groups in a geronto-epidemiological survey., *János,György*, 1991 (Only available in Hungarian)	Reported mortality statistics in Hungary 1990	Selected mortality codes	S(m) indices	Show that diversity in causes of mortality is higher when among deaths where mortality cause was assigned by autopsy than in cases where mortality cause was assigned by attending physician
Measuring epidemiological diversity and concentration: a short review, *Iszák*, 1993	Review of various studies	Review of various studies	Shannon-Wiener index, Fishers α index and Hulberts S(m) index	A review discussing the use of diversity measures applied to causes of mortality and their merit in epidemiology and population health.

Table 2.4: (Cont.) A description of previous studies of diversity in causes of mortality. Where more than 2 countries were studied a full list of studied nations can be found in Appendix A.1

Study	Study population and period	Format of mortality data	Diversity measure(s)	Main conclusions
Changes of neoplasm concentration with geographical co-ordinates, *Hunter, Izsák, Nehaul*, 2002	English and Welsh administrative regions. All analysis performed within 10 year age groups. 1985-1990	Raw mortality counts of underlying causes of death. Fine-grained methods - selected individual ICD-9 codes which describe cancers.	Shannon index	Find that, generally, diversity in causes of mortality in deaths due to cancer is higher at older ages. Over the period 1985-1990 diversity in causes of mortality in deaths due to cancer are found to increase among most age groups in men and to decrease among most age groups in females. Higher diversity in the causes of cancer deaths were observed in more Southern regions among men aged 45+ and in more Westerly regions among men aged 65+. This relationship was found to hold in models that accounted for regional history of smoking.
Diversification in causes of death in low-mortality countries: emerging patterns and implications, *Bergeron-Boucher, Aburto, van Raalte*, 2020	15 "low-mortality" countries. 1994 - 2017 (different years of data in different countries)	Cause count distribution for underlying causes of death extracted from multiple decrement life tables. Course-grained methods -selected ICD-10 causes, categorised into 19 groups based on ICD-10 Chapter classifications.	Shannon entropy	Increasing diversity in causes of mortality across most countries. Note that reductions in the proportion of mortality associated with diseases of the circulatory system and increases in the proportion of mortality associated with diseases of the genitourinary system, mental and behavioural diseases and diseases of the nervous system. Find that in most countries, over time, increases in diversity in mortality causes were associated with increases in life expectancy and reductions in diversity in lifespans (measured using Shannon entropy). Also examine diversity in mortality causes within distinct age groups. Find that in most countries diversity in mortality causes was highest in deaths among those aged 0-19 and those aged 80+ with lower diversity in intervening age groups.
Cause-of-death diversity trends from a multiple-cause perspective in the United States, *Trias-Llimos, Permanyer*, 2023	The USA, both as a whole and divided by educational attainment. 2003 - 2018	Raw mortality records, with underlying and contributory causes of death extracted. Course-grained methods - selected ICD-10 causes, categorised into 13 groups based on ICD-10 Chapter classification.	Propose a novel method in the study of diversity in mortality causes called a "fractionalisation index". Measures average pairwise dissimilarity between the mortality causes (both underlying and contributory) faced by each individual.	Find that diversity in mortality causes increased over from 2003 to 2018 in both males and females in the USA mostly driven by increases among those aged 65 and older. Show that this increase was larger when contributory causes were given more weight in analysis (so called multiple cause of death (MCOD) analysis). Find that diversity in mortality causes was lower among those with higher educational attainment. Educational gradient particularly prominent among those aged 30 - 64, diminishing at older age groups.

2.4.1 Summary of previous findings and insights

2.4.1.1 Historic temporal trends in diversity in mortality causes

There is little coherence in the methods adopted by previous studies of diversity in mortality causes. Studies have used different measures of diversity, assessed mortality under ICD-7, 8, 9, or 10 mortality coding systems, and examined a variety of nations. The most stark divide is around the "mortality causes" over which diversity is measured. Mortality causes have been defined in two ways: as individual ICD (7, 8, 9 and 10) mortality codes, or as groups of these codes with categorisation generally informed by ICD chapter. This produces two methods for the calculation of diversity in mortality causes which I will refer to as: fine-grained cause diversity (diversity in individual ICD codes) and coarse-grained cause diversity (diversity in groups of causes). The implications of changes in diversity under both methods are similar and so they are discussed in tandem in this section.

Despite the disparate methods and settings in previous studies of diversity, some broad patterns are evident. The earliest available study, Izsák (1986), examined mortality cause diversity in the USA from 1968 to 1975. They report reductions in diversity (measured as an increase in "concentration" of death causes) under measures sensitive to rare causes but an increase in diversity in measures sensitive to the most common causes. Their findings indicate that a smaller proportion of individuals had died due to the most common causes of death in 1975 than in 1968. However, at the same time the variety of rare causes recorded in the population had fallen from 1968 to 1975. The use of different editions of the ICD system to record mortality at different time periods makes direct comparisons of diversity in mortality causes between studies problematic. Despite this the findings of each subsequent study found in this review suggest near-constant increases in diversity in mortality causes from the mid-1980s to the mid-2010s (Izsák, 1988, 1993b; Hunter et al., 2003; Bergeron-Boucher et al., 2020; Trias-Llimós & Permanyer, 2023)[3]. These increases are observed relatively consistently across

[3]Bergeron-Boucher et al. (2020) report one exception to this: among females in Finland diversity in causes of death was stagnant from 1996-2016

European countries, the USA and Japan[4]. It should be noted that all analysis of diversity in causes of mortality in the 21st century has measured diversity across coarse-grained causes of mortality while studies in the 20th century mostly assessed fine-grained cause diversity.

In their early studies of diversity in causes of mortality Izsák (1988, 1993b) suggests that increases in diversity over time may be associated with improving diagnostic accuracy. They suggest that as a result of clinicians and those who record mortality causes more precisely diagnosing each individual, a wider variety of causes may be recorded. This hypothesis is supported by the results of a study which compared diversity in mortality causes between causes of death determined by autopsy and by attending physician (Izsák & György, 1991). Diversity in causes of mortality was observed to be higher among deaths in which the cause was determined by autopsy, this finding was attributed to greater accuracy in diagnosis (Izsák, 1993b). Bergeron-Boucher et al. (2020) also suggests greater diagnostic accuracy as a possible driver of increases in diversity in mortality causes in their study of temporal trends in diversity in mortality causes from the 1990s to the 2010s in low-mortality countries, including the UK. However, they suggest that ageing populations and an older average age at death are more important reasons for increasing diversity in causes. Bergeron-Boucher et al. (2020) suggest that as individuals, on average, die at an older age the range of risk factors and diseases they are likely to face increases leading to an increase in the variety of recorded causes of death.

Diversity in mortality causes is, fundamentally, sensitive to changes in the proportion of deaths attributed to each cause of mortality. The causes of mortality which have driven changes in the diversity of mortality causes over time can be examined through studying changes in the cause-specific share of deaths. Bergeron-Boucher at al. (2020) perform this analysis and suggest that the main driver of increasing diversity in their study was improving mortality rates due to circulatory diseases. This was further tested by removing diseases of the circulatory system from analysis using cause-deleted life table methods. As a result diversity increased in fewer countries and diversification (a term for increasing diversity)

[4]Examination of diversity in mortality causes in nations outwith this group is limited to the work of Acaso-Terrén et al. (1988), who did not assess temporal trends.

occurred more slowly in countries where it was observed. The authors conclude that, despite changes in the prevalence of diseases of the circulatory system being the main driver of diversification, redistribution across the remaining causes of mortality also contributed to increases in diversity.

The analysis discussed in this section has so far focused on diversity in the underlying causes of mortality recorded for each individual in a population. However, increasingly in public health science researchers have promoted using a multiple causes of death (MCOD) approach (Fedeli et al., 2015). MCOD analysis refers to research which leverages all causes of mortality recorded on a death certificate. Alongside the underlying cause, in most countries mortality recorders can include further, contributory causes which have hastened or contributed to mortality in the deceased (Redelings et al., 2007). These contributory causes are thought to be a valuable source of information regarding co- and multi-morbidities and the health of those who have died at the time of death (Grundy & Stuchbury, 2022; Batty et al., 2019). Trias-Llimós and Permanyer (2023) promote a novel method of measuring diversity in mortality causes which takes into account multiple causes of death (MCOD), discussed further in Chapter 7. They find using this measure that in the USA in the 21th century diversity in causes of mortality has increased more quickly when contributory causes are taken into account than when underlying causes are examined in isolation.

2.4.1.2 Age-specific diversity in mortality

The causes of mortality which individuals face vary widely at different ages. Several previous studies of mortality cause diversity have sought to examine age-specific diversity in mortality causes to assess how variation in the distribution of mortality causes differs by age (Izsák, 1982; Izsák & Juhás-Nagy, 1984; Izsák, 1993a; Hunter et al., 2003; Bergeron-Boucher et al., 2020). Despite the use different methods, study populations, and periods of study, these studies report remarkably similar patterns of diversity in mortality causes at different ages. In most studies, diversity in mortality causes has been shown, in the majority of countries and settings, to be higher at younger (deaths among those aged 0 to around 39) and older ages (deaths among those aged 60 to 80+), with lower diversity in causes between these ages. This effectively creates a U-shaped curve of diversity in mortality causes if plotted across the life course. This pattern has been shown in the diversity of causes of mortality within deaths which were attributed to cancers and

to diseases of the cardiovascular system separately as well as in diversity across all causes of mortality (Izsák, 1993a; Hunter et al., 2003; Bergeron-Boucher et al., 2020). The U-shaped pattern of diversity in mortality causes is not identical across countries, in terms of the exact age at which peaks and low points in diversity occur and furthermore is not found in early studies of age specific diversity in mortality causes (Izsák, 1982; Izsák & Juhás-Nagy, 1984). These early studies find that diversity in mortality causes was highest at younger ages, before falling in deaths among those who are middle-aged but do not observe the increase in older ages. Instead diversity is shown to plateau or reduce at older ages. The increased diversity in mortality causes at found at older ages in more recent studies is a factor in the hypothesis posed by Bergeron-Boucher et al. (2020) that ageing populations have driven diversification in causes of death.

2.4.1.3 Subpopulation tendencies in diversity in mortality causes

Variation in health outcomes between subpopulations is an important aspect of population health (Newton et al., 2015). In most studies examination of diversity in mortality causes has been limited to country-wide analysis of men and women separately and at different ages at death. Hunter et al. (Hunter et al., 2003) have reported significant geographic variation in diversity in causes of mortality in deaths attributed to cancer within English and Welsh regions in the late 1980s. The observation of geographic variation in mortality cause diversity indicates that variability in causes of mortality differs by subpopulation. However, little examination of the environmental or socioeconomic factors which may have led to these geographic tendencies in the diversity of causes were examined.

Trias-Llimós and Permanyer (2023) evaluated variation in diversity in causes between subpopulations of the USA separated by educational attainment between 2003 and 2018. Educational attainment is a well-studied determinant of health in the literature in various measures of population health (Albert & Davia, 2011; Telfair & Shelton, 2012; Health, 2020). Trias-Llimós and Permanyer (2023) report that diversity in causes of mortality was lower among those with higher educational attainment in the USA. This indicates that, in general, those with fewer qualifications faced a wider variety of causes of mortality. As life expectancies are generally higher among more educated groups this indicates in the USA longer lived populations exhibit greater diversity in causes of mortality. Furthermore, diversity was partitioned in this study to examine the effect of changes within edu-

cational attainment groups (equivalent to alpha diversity) compared to changes between groups (equivalent to beta diversity). Increases in diversity in mortality causes were found to be mostly driven by larger differences in diversity between subpopulations rather than increasing diversity within subpopulations. This may be an indicator of widening inequalities between subpopulaltions.

2.4.1.4 General insights and recommendations

After authoring, and co-authoring, a series of studies on the subject in the 1980s, Izsák published a review discussing the merit of studying diversity in causes of mortality in epidemiology and population health (Izsák, 1993b). In their review, Izsák (1993b) highlight the benefit of using multiple measures of diversity together which place different degrees of significance on the prevalence. This makes it possible to assess changes in the variety of rare causes of mortality faced by populations as well as to explore the dynamics of the most dominant causes. The concurrent use of different measures of diversity is a recurring theme of Izsáks work (Izsák, 1982; Izsák, 1986, 1988, 1993a). Izák (1993b) also raises methodological issues related to the calculation of diversity in mortality causes. One key problem discussed is that the epidemiological categories over which diversity is calculated are not necessarily fixed, comparable entities. The author suggests that unlike biological taxa[5] the causes of mortality in a population are difficult to define on equal terms. This assessment is questionable, as similar issues exist in the study of diversity in any system and in reality it is rare for any distribution to be composed of completely unambiguous units (Baum, 2009; Van Rossum et al., 2020). However, the argument is not completely without merit. This is discussed further below in Section 2.4.3.

[5]The word “taxa” is used in the biological classification of organisms to describe a population or a group of populations which are considered to form a unit. For example, different species of tree in a forest would constitute different taxa.

2.4.1.5 Relationships between diversity in mortality causes and established measures of population health

A strong relationship has been reported between life expectancy and mortality cause diversity at the national level (Bergeron-Boucher et al., 2020). Increases in life expectancy were observed alongside increases in the diversity of causes of mortality across low-mortality nations, including the UK. Bergeron-Boucher et al. (2020) use this evidence to support their hypothesis that that increasing diversity in mortality causes over time are driven by ageing populations, discussed in Section 2.4.1.1. In addition, Bergeron-Boucher et al. (2020) observe a negative relationship over time between diversity in causes of mortality and diversity in age at death. This indicates that as populations died at more homogenous ages they faced a wider variety of causes of morality. Neither the relationship between life expectancy and mortality cause diversity or between diversity in causes of mortality and diversity in age at death was tested cross-sectionally. Interpreting the evidence reported by Bergeron-Boucher et al. (2020), Trias-Llimós and Permanyer (2023) suggest that diversity in mortality causes was greater in countries where life expectancy was greater and where ages at death were more varied (this second point being the opposite of the temporal relationship discussed above). Trias-Llimós and Permanyer (2023) also use their evidence of the diversity in mortality causes across subpopulations, alongside previous research into variation in lifespan in these subpopulations, to indicate that more varied ages at death are associated with greater diversity in mortality causes at the level of subpopulations as discussed above. This potentially indicates a Simpson's paradox, with increased diversity in causes of mortality found in nations with longer life expectancies but within the USA, diversity in causes of mortality found to be lowest among the most educated where lifespans are longest. This relationship has not been formally examined in previously published research outside of the work of Bergeron-Boucher et al. (2020).

2.4.2 The motivation for studying diversity in mortality causes

Few of the previous studies described in this chapter have discussed, in detail, the rationale for measuring diversity in mortality causes. In their study Bergeron-Boucher et al. (2020) present two arguments for monitoring the variability of mortality causes through the measurement of diversity. The first relates to individual

level uncertainty in cause of death. Increased Shannon entropy – as observed in by Bergeron-Boucher et al. (2020) - can be understood as a reduction in the certainty with which the cause of mortality of any individual picked at random from the population can be predicted (Masisi et al., 2008). The same is generally true for a number of measures of diversity. However, Shannon entropy is specifically designed to quantify this uncertainty. This increasing uncertainty in the cause of mortality each individual may face is proposed to have the effect of making medical diagnoses more difficult. This is referred to further in this thesis as increasing diagnostic uncertainty. Diagnostic uncertainty is a term that has been widely discussed in the literature with a range of definitions proposed as well as many concepts for its measurement (Simpkin & Schwartzstein, 2016; Bhise et al., 2018). Whether increased diversity in causes of mortality has an impact on diagnostic uncertainty may be arguable. Indeed, as discussed in previous sections, various studies have suggested that increasing diagnostic precision may contribute to increasing diversity in causes of mortality meaning that in fact increased diversity in causes of mortality may be caused by increased *certainty* in diagnoses (Bergeron-Boucher et al., 2020; Izsák & György, 1991). However, it has been suggested that increasing prevalence of a wide range of conditions, especially those with overlapping symptoms, may contribute to diagnostic uncertainty in medical practice (Mishel, 1988; Pugh et al., 2009; Kelly & Panush, 2017). This impact is likely to be limited to assessments and diagnoses performed by physicians. In the 21st century diagnoses are increasingly assisted by software-based tools and improved diagnostic equipment (Dilsizian & Siegel, 2013; Lai et al., 2006). As these devices become more widespread, accurate and trusted, they will be increasingly used in place of more subjective judgements made by human physicians, which, through reducing the need for subjective judgements, may limit the impact of increasing diversity in mortality causes on medical diagnoses.

The second rationale for monitoring variability in causes of mortality proposed by Bergeron-Boucher et al. (2020) is as a measure of fragmentation of causes at the population level. Increased diversity in mortality causes indicates that deaths are spread more evenly across causes, this means there is greater division in the causes of the burden of disease. Therefore, in order to make continued improvements to health a wider variety of causes will need to be addressed. This has implications for the production of new pharmaceuticals; treatment by healthcare systems; and preventative public health policies. In each of these cases, increased fragmentation of disease burden (indicated by increasing diversity) will mean that already limited resources must be spread to address a wider variety of causes of mortality. However, the impact of fragmentation is not likely to be uniform across health

promotion and care sectors. Prevention policies and wider public health systems are likely to be impacted least because generally their interventions target risk factors rather than specific causes of mortality. Reducing the prevalence of a risk factor, for example tobacco smoking, can have wide ranging effects reducing mortality from heart disease as well as cancers and other causes of mortality (Critchley & Capewell, 2003). By addressing wide ranges of mortality causes, prevention strategies may be, to an extent, protected by fragmentation in mortality causes. Fragmentation in mortality causes might be expected to have a greater impact on treatments by healthcare systems. Generally speaking, medical departments and healthcare professionals, even specialists, treat a range of illnesses and causes of ill health meaning they are likely prepared, to an extent, for an increasing diversity of mortality causes (Smyth et al., 2022; Horwood et al., 2018). Despite this, as diversity in the causes of ill health facing a population increases, new specialisms and increased diagnostic staffing may be required to combat the fragmentation of disease (McKee et al., 2021). Finally, the production of pharmaceuticals is likely to be most impacted by fragmentation. Most pharmaceutical treatments are designed to cure only one disease or a small range of diseases or to treat a small number of causes of ill health (Taylor, 2016). As diversity and fragmentation in causes of mortality increases, an increasing number of pharmaceutical treatments will be needed and these treatments may be useful to a smaller proportion of the population. To summarise, fragmentation is likely to impact prevention the least, have more impact on healthcare and treatment systems and have the largest impact on direct cures for causes of ill health and mortality.

At the societal level, diversification in mortality causes may also be a sign that economies and efficiencies of scale in the treatment and prevention of disease are at risk. When many deaths occur due to common causes, the health of a large portion of the population can be improved by focusing on prevention of risk factors associated with these common causes. Similarly, in hospitals and medical care centres, as many patients will present with the same cause, strategies can be employed to treat common causes more efficiently. However, as causes diversify and a greater variety of diseases, conditions, injuries and causes of ill-health appear in the population these economies of scale will become less powerful.

None of the studies of diversity in mortality causes discussed in this section since Izsak (1986) have observed significant reductions in diversity in mortality causes over time. Reductions in the diversity of mortality causes is likely to come alongside an increasing proportion of deaths attributed to the most common causes. The mortality rates of these causes are likely either increasing, stagnating or reducing

at a slower pace than other less common causes. Many public health resources are generally focused on the most common causes: in general, diseases which are more common receive more research funding, though certain diseases such as breast cancer receive more generous funding than other causes of ill health with similar prevalence (Gross et al., 1999; Carter & Nguyen, 2012). Further, it has been shown, in the USA, that public health policy is generally geared towards the most common causes (Pilar et al., 2020). Therefore falling diversity in mortality causes likely indicates that public health research and policy are to an extent failing in this regard, at least relative to efforts in other less common causes of mortality.

The dynamics of diversity in mortality causes have potentially important implications for population health and for public health practitioners and health care systems in general. Using diversity in mortality causes to examine historic trends we can learn how changes in patterns of mortality were achieved. It can further provide insight into the diagnostic burden on individual physicians and the burden of greater variation in mortality causes on public health and healthcare systems.

2.4.3 How should mortality causes be defined?

In creating somewhat arbitrary groups of causes of mortality or excluding certain individual causes of mortality each of the studies of diversity in mortality causes discussed in this chapter has used a different definition of the distribution of mortality causes. In fine-grained cause diversity, studies have excluded certain ICD mortality codes from analysis with little consistency between studies and limited explanation of an exclusion strategy (Izsák, 1986, 1988, 1993a; Hunter et al., 2003). Using coarse-grained cause diversity none of the three previous studies discussed in this chapter (Ascaso Terrén et al., 1988; Bergeron-Boucher et al., 2020; Trias-Llimós & Permanyer, 2023) calculated diversity over the same groups of causes. Only Bergeron-Boucher et al. (2020) discuss the selection of appropriate groups over which to calculate diversity. They compare three different ways to group causes of mortality and found that diversity in mortality causes increased during their study period under each system and that the conclusions of their study were upheld.

The selection of the causes of mortality over which diversity is calculated is, of course, fundamental to the outcome of any measure. However, the rationale for the use of course-grained methods as opposed to fine-grained methods (or vice versa) is rarely discussed in detail in previous studies of diversity in mortality causes. These studies are likely to have made these decisions to ensure with aetiological standards and relevant literature or to simplify further analysis.

It is suggested by Trias-Llimós and Permanyer (2023) that measuring coarse-grained cause diversity may lessen the effect that changes in coding practices may have on diversity in causes of mortality. This may be true if changes in coding practice are mostly confined to reassigning ICD codes within each group. However, using Scotland as an example, the two major changes in coding practice (National Records of Scotland, 2011; National Records of Scotland & Scottish Government, 2017) in the 21st century both lead to a change in the ICD-10 chapter that certain deaths were recorded under. Therefore, these changes would likely have an effect on diversity measured across both fine-grained causes and coarse-grained causes of mortality. The effect of changes in coding practice are, to a degree, unavoidable when examining mortality data without large scale re-coding of records to create a consistent dataset. Furthermore, changes in coding practices generally occur for a reason. A better understanding of aetiology may have developed or treatment policies may have changed (National Records of Scotland, 2011). Therefore, in order to reflect the true diversity of causes of mortality faced by the population and treated by healthcare systems it is desirable to reflect these changes in coding practice in analysis.

Another argument for the use of coarse-grained cause diversity could be the creation of more meaningful units of comparison. The current systems of classification of morbidity and mortality such as ICD-10 coding are designed to aid the reporting and study of mortality records and to allow comparisons between populations. Comparison between fine-grained ICD codes may be considered problematic as each is designed to describe a specific condition, some of which may be interpreted more broadly than others. However, grouping causes does not necessarily address this concern. In fact, the categorisation process arguably creates groups which are less comparable than fine-grained ICD codes. Most cause-grouping methodologies are based on ICD chapters. This generally means that all cancers are treated as one cause of mortality, and all diseases of the circulatory system are treated as another. A greater number of individual three-character ICD-10 codes (the primary breakdown of mortality code in this system) exist to describe cancers (around 150) than diseases of the circulatory system (around 100). Therefore,

a greater variety of causes of mortality might be expected within the "cancers" group than within the "diseases of the circulatory system" group. Supporting this, diversity in fine-grained ICD codes has been shown to be higher in deaths attributed to cancers than in deaths attributed to diseases of the circulatory system (Izsák, 1988, 1993a). Furthermore, in using coarse-grained methods a degree of variation in the distribution of causes may be lost.

The selection of a definition for mortality causes depends on the motivation for measuring diversity. Diversity in mortality causes is a measure of uncertainty at the individual level and of fragmentation at the population level. Although in both cases understanding changes in the distribution of causes by chapter may be informative; understanding the dynamics of diversity at a greater level of detail is clearly beneficial. Trias-Llimos et al (2023) argue that "studying the effect that more granular classifications can have on cause-of-death diversity measures would be of great added value in future research". ICD chapters are grouped by body system and type of disease; therefore, any uncertainty associated with diagnosis is likely to be higher between causes within the same ICD Chapter. In practice, it may be more likely that one form of cancer is misdiagnosed for another cancer than for a cause in a different ICD-10 Chapter. In grouping causes together by ICD Chapter, it is not possible for measures of diversity to address this uncertainty. For these reasons diversity is measured in this thesis at the level of fine-grained individual ICD-10 three-character codes.

This section discusses the motivations for, and difficulties in, selecting a definition for "mortality causes" in the measurement of diversity in mortality causes. As discussed there is no clear consensus on whether the fine-grained three-character code approach adopted in this thesis or the coarse-grained ICD-10 Chapter approach is optimal. Finding an appropriate medium between these approaches could allow for greater insight than coarse-grained analysis and make visualisation and interpretation more simple than in fine-grained analysis. Using ICD-10 Blocks, an intermediate step between three-character codes and Chapters, or even larger groups of codes could accomplish this. Another option could be the use of similarity-sensitive measures of diversity which explicitly account for for the inherent similarity between certain causes of mortality. A case study reporting a possible route towards this second approach is discussed in Section 8.5.

2.5 Motivation for this thesis and research gaps

The previous section has summarised studies in the literature that have used diversity in causes of mortality in the study of population health. As indicated by the limited number of previous studies in this field, important research gaps remain.

Increases in diversity over time in the late 20th and early 21st century have been observed in numerous studies across various nations, as discussed in this chapter. These increases have been associated with reductions in the prevalence of circulatory diseases observed across various countries. In the 2010s, improvements in mortality associated with cardiovascular disease have slowed in many countries and, in some nations, have reversed (Wilson et al., 2017; Lopez & Adair, 2019; Cheema et al., 2022). The dynamics of diversity in mortality causes during this period of change in mortality trends have not been examined previously. Scotland presents an interesting case study for this research as a nation where this resurgence in cardiovascular mortality has come alongside increasing mortality associated with deaths of despair and other increases in conditions such as degenerative diseases (Brown et al., 2019; Allik et al., 2020).

Previous research into the causes which have, through changes in prevalence, driven diversification in mortality causes in the 21st century has used coarse-grained cause methods. As a result, the individual causes that have been responsible for changes in the diversity of causes have not been examined. Further to this, methods for examining the drivers of diversification are limited. Understanding which causes are driving change in mortality cause diversity is key to addressing diversification or concentration (the opposite effect). In this chapter I introduce "additive value" a measure of the contribution of each fine-grained cause of mortality to overall diversity in mortality causes.

Most of the previous studies of mortality cause diversity discussed in this chapter have been carried out at the national level. Limited previous evidence has been presented into how this measure may vary within countries by socio-economic or geographic factors. This is despite the importance of inequalities in health within countries on overall population health (Wyper et al., 2019). No available research

has examined how diversity in causes of mortality varies by subpopulations divided by deprivation or across the urban-rural gradient despite important all-cause and cause-specific variations in mortality between these groups (Levin & Leyland, 2006; Brown et al., 2019).

These research gaps are the focus of Chapter 4 of this thesis, which aims to answer the questions:

- What were the national temporal trends in diversity in mortality causes in deaths across all ages in Scotland from 2001 to 2019?
- What were the national temporal trends in diversity in mortality causes at different ages at death in Scotland from 2001 to 2019?
- Did diversity in mortality causes differ between subpopulations, grouped by area-level deprivation and urban-rural class, in Scotland from 2001 to 2019?
- What were the subpopulation level trends in diversity in mortality causes in Scotland from 2001 to 2019?
- Which causes of mortality drove trends in the diversity of mortality causes at the national and subpopulation level in Scotland from 2001 to 2019?

The relationship between diversity in causes of mortality and variation in ages at death within populations has previously been examined by Bergeron-Boucher et al. (2020). They use the same measure of diversity to examine variation in the distribution of causes of mortality and ages at death. A number of measures exist for the study of variation in lifespans. Measuring variation in the distribution of ages at death using measures of diversity has the potential to be advantageous within this field beyond simply improving comparability to diversity in mortality causes. Comparisons of measures of diversity in age at death to established measures of lifespan variation have not been made, nor have the advantages of using measures of diversity in the study of variation in lifespans been fully explored (van Raalte & Caswell, 2013; Aburto & van Raalte, 2018; Aburto et al., 2021).

It has been observed that, over time, increased diversity in causes of mortality is associated with increased life expectancies and reduced diversity in age at death at the national level (Bergeron-Boucher et al., 2020). It has also been hypothesised that because there are varied cause-specific mortality risks at differ-

ent ages (Robine et al., 2007; Case & Deaton, 2015; Brown et al., 2019; Murphy, 2021); increased diversity in causes of mortality might be expected in populations where ages at death are more varied (Trias-Llimós & Permanyer, 2023). Understanding whether increased diversity is associated with older populations or with populations where people die at more varied ages can help to aid our understanding of how the distribution of mortality causes will change as population health progresses. In order to explore this it is important to explore both cross-sectional relationships between diversity in mortality causes, life expectancy and diversity in age at death and relationships over time; despite this such cross-sectional examinations are lacking.

The aim of Chapter 5 is to explore these research gaps and answer the following questions:

- Is normalised alpha diversity in age at death (lifespan diversity) well-correlated with existing measures of lifespan variation?
- Can lifespan diversity calculated from observed death counts be considered a reliable proxy for lifespan diversity calculated from the lifetable death distribution?
- What were the temporal trends in lifespan diversity in Scotland and Scottish subpopulations from 2001-2019?
- What is the relationship between diversity in causes of mortality and lifespan diversity in Scotland, firstly, across time and secondly, between subpopulations?

Previous studies of diversity in mortality causes have suggested that the COVID-19 pandemic may have had a disruptive effect on the distribution of mortality causes (Bergeron-Boucher et al., 2020; Trias-Llimós & Permanyer, 2023). These studies have not analysed the diversity of mortality causes during the pandemic. The high burden of mortality related to COVID-19, as the most common cause of mortality in many countries in 2020 and 2021, will have had a substantial effect on the distribution of mortality causes. However, the COVID-19 pandemic has been proposed to have wider direct and indirect effects. Examining diversity in mortality causes offers an opportunity to understand these effects in the context of the entire distribution of mortality causes. Through diversity analysis, it is possible to examine pandemic related changes in the prevalence of the most common causes

of mortality and to understand whether variation in the range of rare causes of mortality recorded in the population shifted. Understanding these dynamics can provide a way to contextualise the impact of the COVID-19 pandemic on population health.

Trias-Llimós and Permanyer (2023) suggest that changes to the distribution of mortality causes related to the COVID-19 pandemic are especially important in nations that have lagged comparators in health improvements. They frame this around mortality in the USA where various health outcomes have been slow to improve compared to other Western, low-mortality countries. These adverse trends have also affected Scotland where improvements to a number of mortality-based measures of population health (e.g. life expectancy) have slowed or reversed in recent years (Fenton et al., 2019b). This followed decades of slower improvements to health in Scotland compared to other constituent nations of the United Kingdom and comparable European nations (McCartney et al., 2012). Understanding how the effects of the COVID-19 pandemic presented alongside these existing adverse trends in population health in Scotland may be especially important to understand the impact of the pandemic on population health and to planning the recovery from this impact.

In Chapter 6, research gaps related to the COVID-19 pandemic are addressed, with the aim of answering the following questions:

- What was the effect of the COVID-19 pandemic on diversity in mortality causes in Scotland across in the years 2020 and 2021?
- Does the diversity in causes of mortality in Scotland exhibit a seasonal pattern?
- What was the effect of the COVID-19 pandemic on diversity in mortality causes during the peaks of the COVID-19 pandemic in Scotland?
- Was the impact of the COVID-19 pandemic on diversity in mortality causes consistent between groups of causes and in deaths at different ages?

In Chapters 4 to 6, diversity is measured in the underlying causes of mortality recorded in Scotland and Scottish subpopulations. Underlying causes of death are the foundation on which most studies of mortality in population health are built (Bishop et al., 2022). Despite this, they often do not tell the full story of an in-

dividual's health at the time of death. As discussed in Section 2.4.1.1 to gain a wider perspective, a growing field of research uses methods which consider other indicators of ill-health recorded on the death certificate. These methods are referred to as MCOD analysis and involve assessing contributory causes of morbidity and mortality alongside the underlying cause of death (Désesquelles et al., 2014).

In their study, Trias-Llimós and Permanyer (2023) use MCOD analysis in the context of diversity in mortality causes. They use methods which assess differences in the specific set of mortality causes faced by each individual. This groups the underlying cause and contributory causes together and offers insight into the range of multimorbidities faced by the population. Little is known regarding variation in the distribution of contributory factors in isolation (i.e. without taking into account underlying causes) or what effect changes in the diversity of underlying causes have on diversity in contributory causes of morbidity and mortality. This is important to understand as changes in the diversity of contributory causes may compound the effects of diversification in underlying causes of mortality. If diversity in contributory causes of mortality increases alongside diversity in underlying causes it would indicate that, in addition to dying from a wider variety of causes, the population had also faced a wider range of diseases and conditions at the time of death. This has the potential to exacerbate uncertainty in diagnosis, as overlapping causes become more likely to appear together. This may increase treatment and prevention costs as causes of morbidity and mortality are treated in the same way regardless of where they appear on the death certificate.

Chapter 7 aims to examine diversity in contributory causes of mortality and answer the following questions:

- What are the national trends in the diversity of contributory causes of morbidity and mortality in Scotland?
- What is the relationship between the diversity of contributory causes and the diversity of underlying mortality causes in Scotland?
- Were the contributory causes of mortality recorded alongside COVID-19 more or less diverse than those recorded alongside other leading causes of mortality in 2020 and 2021?

Chapter 3

Methods

3.1 Chapter overview

This chapter introduces and discusses the data and methods used throughout this thesis, as well as some methods used in specific chapters. All results chapters (Chapters 4 to 7) use the following datasets: firstly, mortality records and secondly, population data provided by National Records of Scotland (NRS). In addition, Chapters 4 and 5 make use of Scottish Index of Multiple Deprivations (SIMD) income deprivation data and Scottish Government urban-rural class data. This chapter begins with a section outlining these data sources (Section 3.2).

The methods discussed in this chapter are described in the following paragraph. Using mortality and population data referred to above, life tables were constructed for the Scottish population and for Scottish subpopulations; these life tables were used in Chapters 4, 5, and 7. The calculation of life tables is described in Section 3.3.1. Cause-specific and all-cause age-standardised mortality rates (ASMRs) are used across results chapters to examine the burden of disease associated with various causes of death, and their calculation is described in Section 3.3.2. In Section 3.3.3 the measures of diversity used in this thesis are discussed. These measures, namely subcommunity normalised alpha diversity at q values of 1, 2, and ∞ under the Reeve et al (2016) framework, are used across results chapters. In Chapter 5 a number of established measures of variation in age at death are compared to a novel measure, lifespan diversity, proposed in this thesis. These established measures are introduced in Section 3.3.4. Finally, auto-regressive integrated moving average (ARIMA) models are introduced in Section 3.3.5 these models are used

in Chapter 6 to forecast diversity in causes of mortality during the COVID-19 pandemic given previous trends. This makes it possible to compare a counterfactual scenario with observed diversity in the years 2020 and 2021. For the most part, this Chapter 3 covers methods and data that apply to multiple chapters. Chapters 4 to 7 contain their own specific methods sections.

3.2 Data

This section presents and describes the four main sources of data used in this thesis: mortality (3.2.1) and population data (3.2.2), both obtained from NRS, as well as area-level income deprivation (3.2.3) and urban-rural classification data (3.2.4) obtained from the Scottish Government. These represent the core data upon which analysis in this thesis is based; further supplementary sources of data used in specific chapters are described in those chapters.

3.2.1 Mortality data

Scottish mortality data used throughout this thesis were extracted from individual-level mortality records produced by NRS and held by the MRC/CSO Social and Public Health Sciences Unit. These records contain all deaths registered in Scotland meaning that deaths of Scottish residents which occurred outside of Scotland are not included. The NRS dataset is the most complete record of mortality in Scotland and is the basis for official statistics and government studies.

From this dataset, the year of death, sex, age at death, data on location of residence and cause of death were extracted for each recorded death. Mortality records were assigned to datazones through postcode of residence data linked to each record and Scottish Government lookup tables. Datazones are small area geographies used by the Scottish Government in the production of a number of statistical outputs. Here mortality records were linked to datazones to make it possible to link population, area-level deprivation, and urban-rural class data to each record. The resources used to match postcodes to datazones were the Scottish Postcode Directory held by the NRS and the 2006 and 2016 SIMD postcode-level

release. Neither of these resources contained all postcodes in Scotland within which deaths had been recorded. However, through combining these datasets, it was possible to match a datazone to 99.9% of mortality records. The boundaries of datazones in Scotland change following census years to account for changes in population. During the study period two sets of datazones were in use, namely, those formed in 2001 and 2011. Records of deaths in the years 2001 to 2009 were assigned to 2001 datazones and those in the years 2010 to 2021 were assigned to 2011 datazones. There are 6,505 2001 datazones and 6,976 2011 datazones. Both sets of datazones were designed to contain a population of between 500 and 1000 individuals each.

3.2.1.1 Cause of death information and the ICD-10 classification system

The cause of death information extracted from NRS mortality data included both the underlying cause of death and up to nine additional contributory causes. These were in the form of ICD-10 codes, which have been in use in Scotland throughout the study period having been introduced in 2000 (National Records of Scotland, n.d.). All recorded ICD-10 codes for the underlying and contributory causes of mortality were converted into three-character codes (which I term fine-grained causes) for further analysis by removing any further characters. The ICD-10 classification system and its structure are described below.

ICD-10 Categorisation

The ICD-10 classification system consists of individual codes designating a certain cause of morbidity, mortality, or ill health which are classified into Chapters together with codes denoting similar conditions or diseases. There are 22 ICD-10 Chapters which generally contain either: diseases which pertain to an individual body system (i.e. Chapter IX: Diseases of the cardiovascular system) or a certain category of disease or cause of injury or mortality (i.e., Chapter II: Neoplasms or Chapter XX: External causes of morbidity and mortality).

ICD-10 Code Structure

ICD-10 codes consist of between three and seven characters, with a larger number of characters indicating greater specificity. Codes within the ICD-10 system for the classification of causes of morbidity and mortality consist of three main parts.

The first part of an ICD-10 code is a single letter from A to Z that signifies the body system or the category of disease. These letters are used to group codes into ICD-10 Chapters. Most Chapters consist of codes beginning with a single letter; for example, codes which begin with the letter I are all found within Chaper IX: Diseases of the circulatory system. However, some Chapters contain codes that begin with a different letter (Chapter II contains codes beginning with both letters C and D), and some letters are found in more than one chapter (codes beginning with the letter D are found in both Chapters II and III).

The next part of an ICD-10 code consists of two characters that are always numeric and signify the specific disease, condition, or cause of injury. For each initial letter these characters range from 00 to 99, each coding for a different cause of mortality.

Together, the first and second parts of an ICD-10 code form a three-character code. Examples include codes: “C34: Malignant neoplasm of bronchus and lung” or “X70: Intentional self-harm by hanging strangulation and suffocation”. These three-character codes are the primary level of specificity used in population health research. The last part of the ICD-10 code is four alphabetic or numeric characters that come after a decimal place and are used to increase the specificity of the code, allowing for the position within the body or the specific presentation of a disease or condition to be recorded. These are not used in the analysis reported in this thesis.

3.2.1.2 Changes in mortality coding practices

The coding of causes of mortality in each country follows a protocol set by national or regional authorities. These differ slightly from country to country and region to region to suit the needs and capabilities of different healthcare systems. In Scotland, mortality coding practices are set nationally. Over time, as knowledge

around the aetiology of diseases and diagnostic procedures improve, new causes of mortality arise and technological innovations occur, such that it is necessary to update coding practices. Two such changes in coding practice are reported by the NRS during the years which are studied in this thesis. These occurred in 2010 and 2017; both changes were associated with new software systems used in the recording of mortality causes (National Records of Scotland, 2011; National Records of Scotland & Scottish Government, 2017). The first was estimated to have affected the first character of the ICD-10 code in 2% of deaths. NRS suggest that this change in coding practice had effects across ICD-10 Chapters. One of the most significant effects was a consolidation of deaths due to alcohol and drug abuse, suicides, and violence within Chapter XX: External causes of morbidity and mortality, increasing the number of deaths recorded in this Chapter. This meant a reduction in the number of deaths recorded in Chapter IV: Mental and behavioural disorders, where some of these deaths had previously been categorised. An increase in the number of deaths assigned to degenerative diseases (namely various types of dementia and Alzheimer's disease) also occurred; these deaths would mostly previously have been assigned to causes within Chapters X: Diseases of the respiratory system or XIV: Diseases of the genitourinary system (National Records of Scotland, 2011). NRS report that following the second change in coding practice, 96% of deaths were assigned to the same ICD-10 Chapter as they would have been previously (National Records of Scotland & Scottish Government, 2017). They report the largest change to be a further increase in the number of deaths recorded as dementia rather than within ICD-10 Chapter X. The NRS advice on these coding changes is that while they are small, they should be taken into account when examining trends over time. This is the approach taken in this thesis and the potential impact of these coding practices are discussed throughout this work.

Alongside the changes in coding practice discussed in the previous paragraph, the Scottish Government introduced legislation in 2015 which changed the process of mortality recording in Scotland (Scottish Government, 2011). The Certification of Death Act (2011) aimed to increase scrutiny and therefore accuracy in mortality cause recording and introduced a new form on which mortality details were recorded. However, the NRS and Scottish Government have not identified that this act is likely to have had a major effect on the recording of underlying cause of death (Earle, 2010; National Records of Scotland Web Team, 2013). While there is no reason to discount this, evidence presented in Chapter 7 shows that after 2015, the average number of contributory causes recorded on each death certificate increased. The potential impact of this change is also discussed throughout this thesis.

3.2.1.3 "Garbage" mortality codes

A number of three-character codes defined within the ICD-10 classification system are considered to be impractical for use in population health research (Ellingsen et al., 2022). These codes are referred to as "garbage codes", mostly because they are vague and uninformative meaning they cannot be used effectively in research or reporting (Johnson et al., 2021). In most years garbage codes are assigned to between 10% and 15% of deaths in Scotland (Trends in the share of deaths in Scotland and within the Scottish subpopulations examined in this thesis are shown in Appendix Figure B.6). In population health research it is common to redistribute deaths assigned to garbage codes to "valid" mortality codes (Ellingsen et al., 2022). This generally involves complex algorithms which use aetiological information or multiple-cause-of-death methods to reassign garbage code deaths to causes of death considered to be more useful for further analysis (Johnson et al., 2021). In the analysis of diversity in mortality causes presented in this thesis garbage codes are retained within the distribution of mortality causes due to the complexity involved in accurately redistributing mortality causes. Therefore, in Appendix B.2 I test the sensitivity of the results in this thesis to the presence of garbage codes. In this Appendix, a simple redistributive algorithm is used to reclassify all deaths assigned to garbage codes to valid causes of mortality. With these garbage codes redistributed, diversity in causes of mortality was calculated as described below in Section 3.3.3. Trends in the diversity of mortality causes with garbage codes redistributed was then compared to trends in the diversity of mortality causes in the observed distribution of mortality causes as described in Chapters 4 and 7. The trends in diversity in mortality causes reported in this thesis are upheld in most cases under this analysis suggesting that the conclusions of this thesis are not sensitive to the presence of garbage codes. In Chapters 4 to 7, where the analysis described in Appendix B.2 suggests trends would differ with garbage codes redistributed, the implications of this difference are discussed. In addition to this the potential impact of garbage codes is discussed throughout this thesis.

3.2.1.4 Data description

The total NRS mortality dataset used in this thesis consisted of records for 1,186,827 individual deaths in Scotland in the years 2001 to 2021. Of these records, 198 were excluded due to missing data (namely records lacking an age at death (n=2)) or problems with matching to administrative records (n=196). Although it would have been preferable to have included these records, any biasing effect of their exclusion on temporal analysis was considered negligible, as they occurred with relative uniformity throughout the study period, as shown in Figure 3.1.

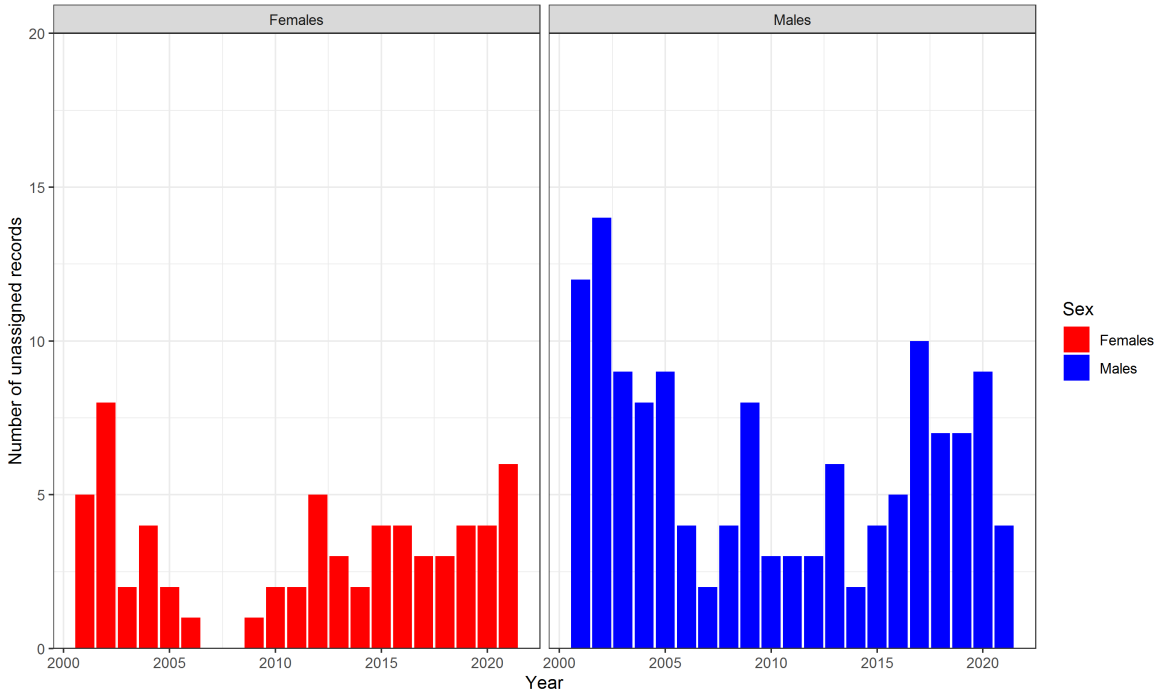


Figure 3.1: The number of excluded mortality records in males and females in Scotland in each year 2001 to 2021.

Following exclusion of incomplete records, data for 1,186,629 deaths were encluded in analysis. Between 53,464 and 64,073 deaths occurred in Scotland in each year during the study period. Figure 3.2 shows the raw number of deaths in males and females separately in each year from 2001 to 2021 in Scotland, as well as the median age at death in these years. The largest number of deaths was recorded in 2020 (64,073), followed by 2021 (63,420), reflecting the toll of the COVID-19 pandemic in those years. Before this, the peak in terms of raw deaths in Scotland was

in 2017 (58,272). The fewest deaths were recorded in Scotland in 2011 (53,464). During the study period, the average age at death among males was lower than among females, although the median age at death in both sexes increased over the study period from 80 to 83 in females and from 74 to 77 in males.

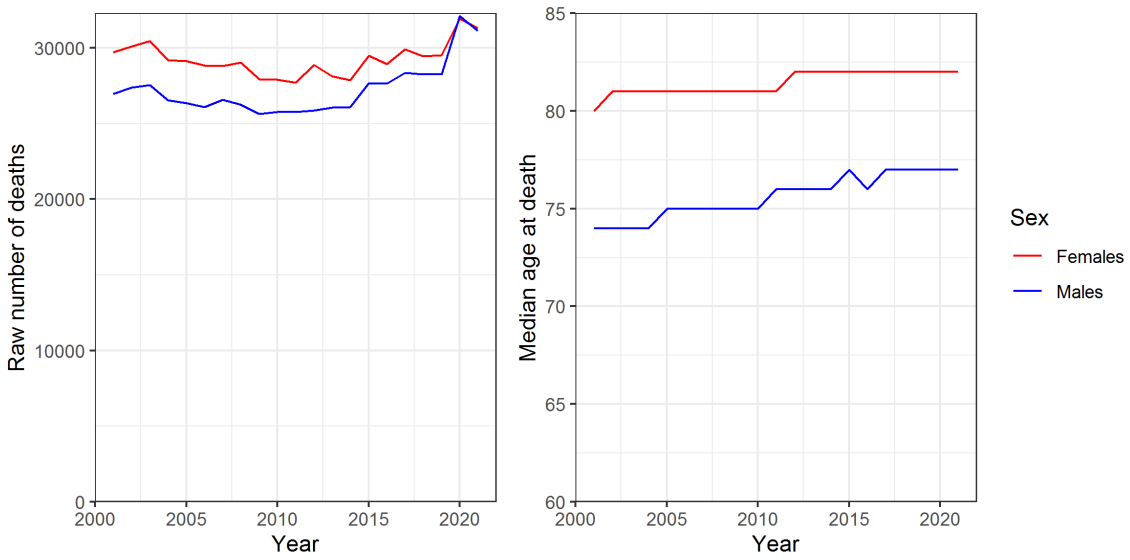


Figure 3.2: The raw number of deaths in each sex in Scotland from 2001 to 2021 and the median age at death in each sex in those years.

A larger number of individual ICD-10 three-character codes were recorded in males than in females in each year 2001 to 2021. However, the difference in the number of causes recorded in each sex was small. During the study period the number of recorded causes changed little, lying between 529 and 596 in both sexes. Figure 3.3 shows the trends in the total number of ICD-10 codes recorded as the underlying cause of death in each sex in Scotland. Of over 1,700 ICD-10 three character codes, only a relatively small proportion were used to classify deaths within Scotland in this period.

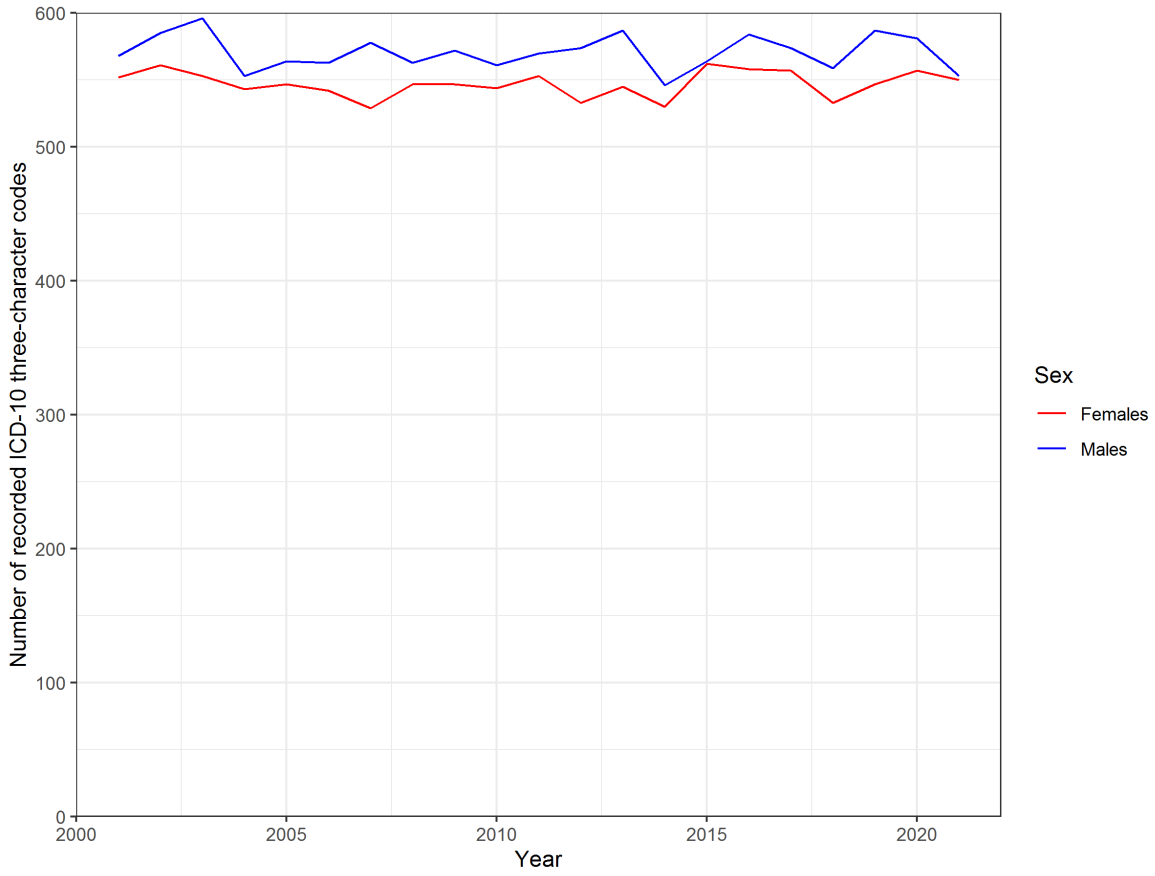


Figure 3.3: The number of ICD-10 three character codes recorded among males and females in Scotland in each year 2001 to 2021.

3.2.2 Population data

In this thesis, population data is used in the calculation of mortality rates, life tables and age-structure diversity. NRS mid-year small-area population estimates for the years 2001 to 2021 are used (National Records of Scotland, 2021). Mid-year population estimates are modelled projections of the size of the population which is usually resident in an area. These estimates are provided for males and females separately at single years of age from 0 to 89 with a further open-ended age class 90+. The total population of Scotland according to these estimates increased over the study period from 5,064,200 in 2001 to 5,479,900 in 2021. The population size in each year of the study period in males and females separately is shown in Figure 3.4.

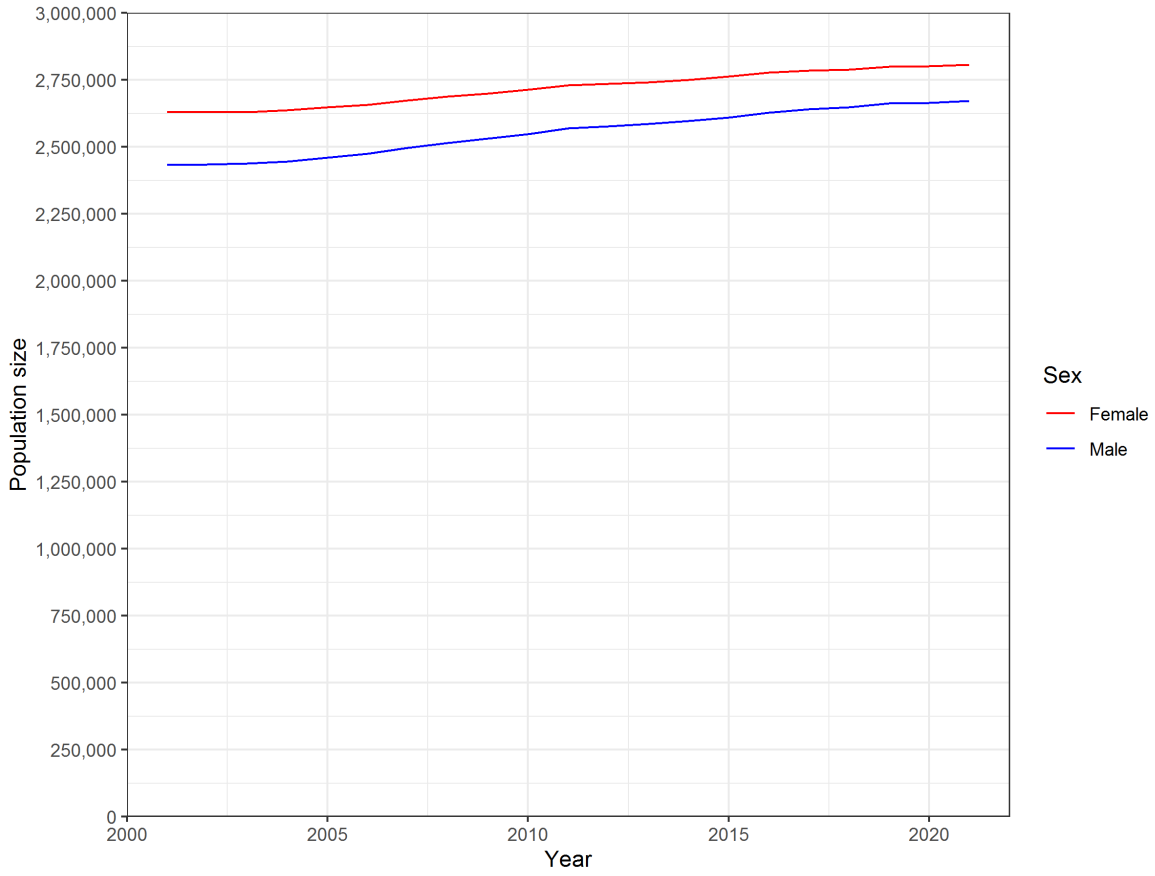


Figure 3.4: The male and female population in each year of study.

Figure 3.5 shows population pyramids for the male and female populations of Scotland in 2001 and 2019; with the stationary population structure extracted from lifetables overlaid. The calculation of the lifetables used in Figure 3.5 and throughout this thesis is discussed further in Section 3.3.1. The population structure of Scotland is shown here to have changed a little over the study period, from 2001 to 2019 the Scottish population has aged with a greater proportion living to older ages and relatively fewer younger residents.

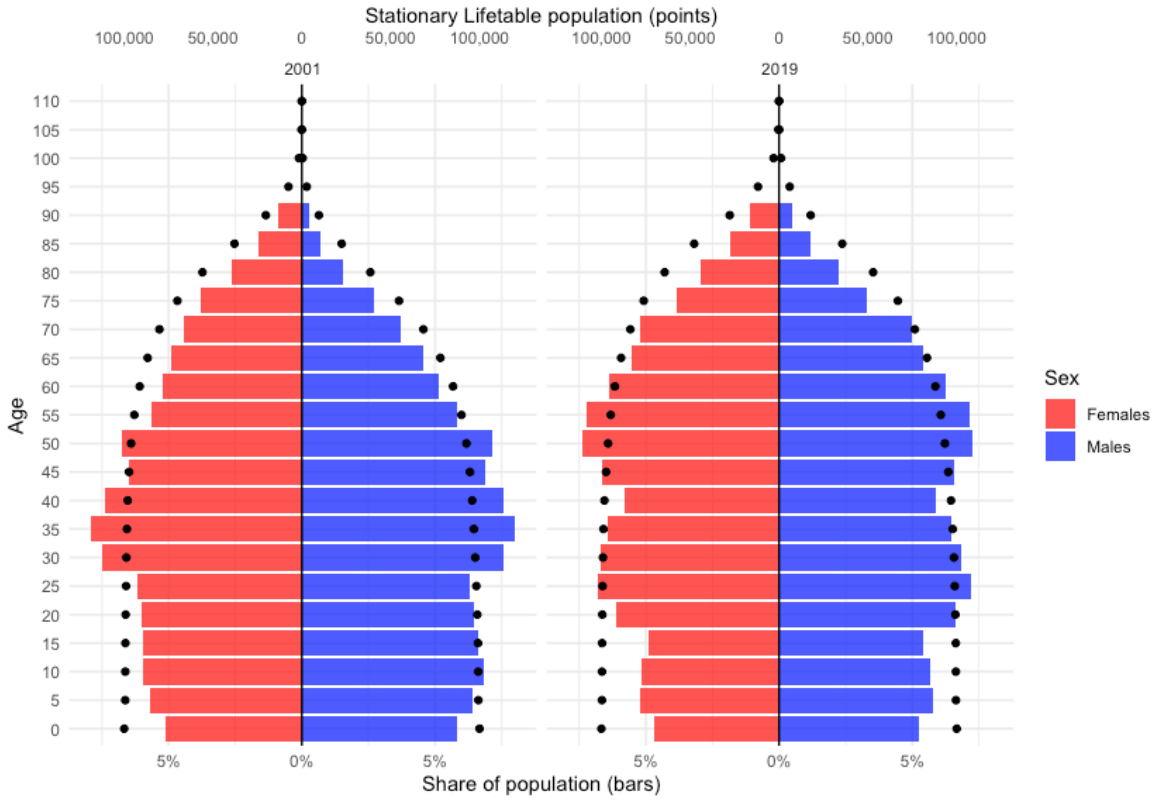


Figure 3.5: Population pyramids for Scotland in 2001 and 2019 with lifetable stationary populations overlaid.

3.2.3 Income deprivation data

In Chapters 4 and 5, socioeconomic deprivation is measured using an area-level measure of deprivation, the Scottish Index of Multiple Deprivations (SIMD). No single measure of deprivation is universally used in the study of population health. The two most prominently used area-level measures in Scotland are Carstairs score and the SIMD (Carstairs & Morris, 1989; The Scottish Government, 2017). These measures are considered broadly comparable in their scope and applicability to population health research (Seaman, 2017). The SIMD was chosen over Carstairs score for two reasons. The first relates to the size of the administrative geography over which these measures are calculated. The SIMD is calculated at the level of datazones in Scotland which had an average population size of 786 in 2021, whereas the 2001 release of Carstairs score is available only at the level of postcode sectors which have populations of roughly 5000 individuals (Brown et al., 2014; National Records of Scotland, 2022b). The SIMD is a relative, area-level measure of deprivation and so not all individuals identified in an area of high deprivation will personally face high levels of deprivation (Lupton & Tunstall, 2003; The Scottish

Government, 2017). The SIMD is available at the datazone level throughout the study period meaning it is calculated among smaller, more homogeneous, populations which minimises the limitation of misclassification by degree of deprivation. Secondly, the Carstairs score is calculated using census data which is only produced once a decade. During the study period, census years were 2001 and 2011. The 2021 census was delayed until 2022 because of disruption related to the COVID-19 pandemic. This limits the applicability of this measure to studies over time. The SIMD is calculated at more regular intervals in Scotland. Only two releases were used to assess deprivation in this thesis (2006 and 2016) however, because they represent the mid-point of the periods to which they are applied (2001 to 2009 and 2010 to 2019 respectively) they were considered to have greater relevant coverage over the study period than census data would.

The SIMD is calculated across seven domains providing information on different aspects relating to deprivation: income, employment, education, housing, health, crime and geographical access to services. The SIMD health domain takes into account several indicators of the health of populations. It has been suggested the inclusion of health data in this way might skew analysis of other health measures such as those presented in this thesis (Bradford et al., 2022). In order to avoid both this and the effect of methodological changes, such as changes in the weighting of domains, the SIMD income domain is used in isolation in this thesis, as is common practice (Brown et al., 2019; Bradford et al., 2022). It has been shown that the income domain is highly correlated with the SIMD as a whole and therefore is an appropriate proxy for deprivation (Bradford et al., 2022).

The SIMD income domain is calculated by estimating the proportion of individuals within each datazone who are income deprived. Income deprivation is determined by the Scottish Government using the following indicators obtained from from the United Kingdom Department of Work and Pensions and His Majesty's Revenue and Customs (The Scottish Government, 2017) :

- Number of adults receiving income support or income-based employment and support allowance, and the number of adults receiving jobseeker's allowance.
- Number of adults receiving Guaranteed Pension Credit.

- Number of children dependent on a recipient of income support or income-based employment and support allowance, and the number of adults receiving jobseeker's allowance.
- Number of adults in paid employment receiving Universal Credit.
- Number of adults and children in Tax Credit families on low incomes.

The total number of individuals who are determined to be classified within the above indicators is divided by a population estimate for each datazone to calculate the proportion of the population within a datazone who face income deprivation (Scottish Government, 2020). Income deprivation proportions at the datazone level were extracted from the income domain of the SIMD (2006 and 2016 releases). Datazone level SIMD releases contain population estimates for each datazone. These are the populations used in the SIMD to calculate the proportion of the population who face income deprivation. Datazones were ranked by these proportions and the SIMD population estimates were used to create population-weighted quintiles. Quintiles are approximate fifths of the population. In this case quintile 1 represents the fifth of the population who live in the most income deprived areas and quintile 5 represents the fifth of the population who live in the least income deprived areas. The resultant subpopulations represent around a fifth of the total Scottish population in each year. Figure 3.6 shows the mid-year population estimates of the datazones in these quintiles. Note that the populations of these quintiles are not equal because quintiles were calculated across the total population, without separating out males and females. For example, the top row of Figure 3.6 shows that there are more men in the least deprived areas than the most deprived; in contrast, there are more females in the most deprived areas than the least deprived. Further, quintiles were calculated using a single year of population, meaning ongoing changes in population dynamics within Scotland are not captured. However, quintile populations varied by less than 6% from the median quintile by population in each year. Quintiles varied more by number of deaths, with more deaths recorded in more income deprived areas, irrespective of sex, as shown in the bottom row of Figure 3.6.

Following 2010 a sharp change in the populations of each SIMD income deprivation quintile can be observed in the top row of Figure 3.6. This change in population size occurs because different editions of the SIMD are used for 2001-2009 and 2010-2019. The use of these different editions means that different areas, with slightly different population sizes are included within each quintile. This change is likely to have little effect on analysis in this thesis thanks to the use of life tables which standardise populations in all analysis of subpopulations.

Figure 3.6: The population size and raw death count of each SIMD income deprivation quintile in each year 2001 to 2019.

3.2.4 Urban-rural classes

There are a number of ways to determine where on the urban-rural spectrum an area lies (Rashed & Jürgens, 2010). Often population per unit area, or residences (such as addresses or households) per unit area, are used (Verheij et al., 2008). This method is often used in the literature related to urban-rural effects of health. It does not however, take into account place-related aspects of urban/rurality which

are thought to be relevant to health. Place-related aspects of health include varying levels of access to both health services and to health improving factors. These factors may be more important than simply the population density of an area (Long, 1993; Corburn, 2017). Instead, some studies use indicators of urban/rurality which take into account geographical features as well as the number of people who live in an area (Levin, 2003). In Chapters 4 and 5, one such indicator, produced by the Scottish Government, is used to assign mortality records to urban-rural classes. The Scottish Government produces urban-rural classes for small areas in Scotland based on settlement size by population, and accessibility based on drive time to large settlements. In this thesis the six-fold classification produced at the datazone level was used (3.1). The six fold classification is not as comprehensive as other proposed methods for the classification of urban/rural features in Scotland, which encompass more detailed data on social characteristics (Levin, 2003). However, the Scottish Government urban-rural classes are used often in the study of health in Scotland and are generally considered to be a good proxy for the urban-rural characteristics of small areas (Pateman, 2011; McKenzie et al., 2013; Rushworth et al., 2015).

The 2006 release of the Scottish Government urban-rural classifications was linked to records from 2001 to 2009 and the 2016 release was linked to records from 2010 to 2019. As shown in Figure 3.7, using these releases results in a change in the population size of certain classes from 2009 to 2010. This change is unlikely to significantly impact analysis of urban-rural differences in this thesis as all such comparisons are made using life-tables which standardise for population size.

Table 3.1: Descriptions of the Scottish Government Urban/Rural Classifications.

Class	Class Name	Description
1	Large Urban Areas	Settlements >= 125,000 people
2	Other Urban Areas	Settlements of >10,000 and <125,000 people
3	Accessible Small Towns	Settlements of >3,000 and <10,000 within 30 minutes' drive of a settlement of >10,00 people.
4	Remote Small Towns	Settlements of >3,000 and <10,000 with more than a 30 minutes' drive of a settlement of >10,00 people.
5	Accessible Rural Areas	Settlements of <3,000 within 30 minutes' drive of a settlement of >10,00 people.
6	Remote Rural Areas	Settlements of <3,000 with more than a 30 minutes' drive of a settlement of >10,000 people.

The Scottish population is not evenly split between the rural-urban classifications, with the urban areas, classes 1 and 2, home to greater proportions of the population. These differences in population may bias some of the measures of diversity studied in this thesis. At smaller sample sizes, measurements of diversity may be lower simply because there are fewer individuals to measure. This is discussed further in Section 3.3.3.2. The minimum number of observations needed to provide a reliable measurement is dependent on the number of possible types (causes of mortality or ages at death) and the total number of deaths recorded (which is related to population size). In Section 3.3.3.2, a diversity accumulation curve is presented indicating that sample sizes > 1000 are sufficient to provide reasonably reliable measurements of diversity in mortality causes in Scotland. In the smallest subpopulations by urban-rural status, a minimum of approximately 1,250 deaths occurred each year; therefore it can be assumed that any biasing effects of sample size in mortality cause diversity between these subpopulations in deaths across all ages are likely to be small. This applies to all analysis of populations and subpopulations when analysing deaths across all ages, as is the case in the majority of analyses performed in this thesis. The issue of sample size in the calculation of diversity in mortality causes is discussed further in Section 3.3.3.2.

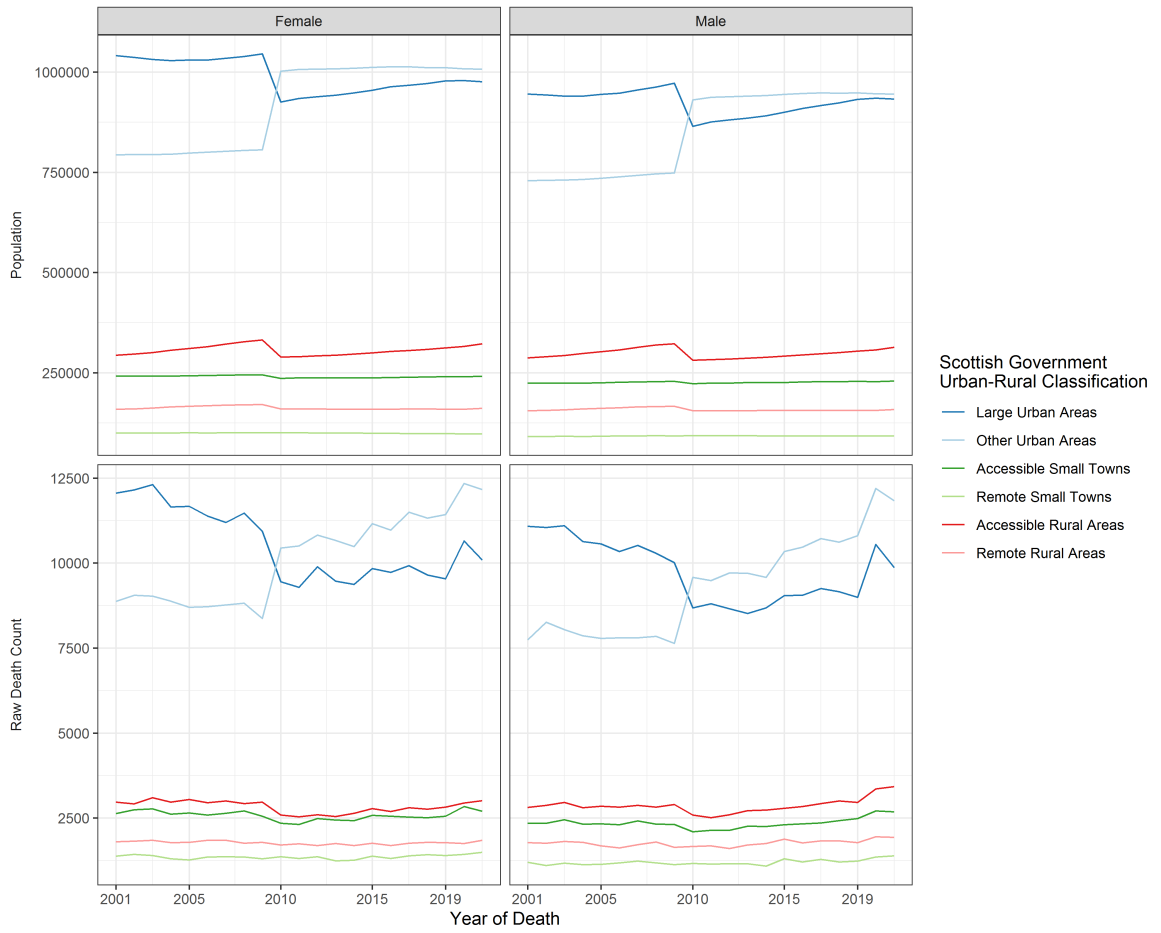


Figure 3.7: The population size and raw death count of each Scottish Government urban-rural class in each year 2001 to 2019.

3.3 General methods

The following section covers an explanation and discussion of several methodologies used in analysis across chapters: life tables (Section 3.3.1), the calculation of ASMRs (Section 3.3.2), and normalised alpha diversity (Section 3.3.3). This section also contains descriptions of methodologies used in specific chapters, namely: the measurement of variation in lifespan in Chapter 5 (Section 3.3.4) and predictive ARIMA models used in Chapter 6 (Section 3.3.5). Other methodologies used in this thesis are described in the chapter in which they are used.

3.3.1 Life tables

Period life tables are a demographic tool used to describe the force of mortality within a population in a given period of time (Vaupel et al., 1998). Life tables are used to examine the probability of death at any age as well as to calculate life expectancy within a population. There are two types of life table: period and cohort. Cohort life tables, as the name suggests, examine the demographic patterns within a defined cohort of the population. The cohort is tracked from birth to death and the precise age of death is recorded for each individual. This allows for a detailed and exact demographic description of the given population (Preston et al., 2001). This type of life table can only be calculated once everyone in a given cohort has died. In the fast-paced world of health and demographic research it is clearly desirable to have data on the health of the current population without this delay. Period life tables meet this need as they, in effect, simulate a cohort and apply the force of mortality at each age over the study period in the population of interest. Both period and cohort life tables are almost exclusively calculated for males and females separately because of differences in demographic processes and timings between sexes (Andersson & Philipov, 2002).

Period life tables were constructed, in this thesis, following established methods (Preston et al., 2001). Each life table in this thesis was calculated using data for a single year. A simulated population of 100,000 (the so-called 'radix' of the life table) individuals is set at the start of each table. Single-year age-specific mortality rates (m_x) in the study population for a single year are used to calculate the probability of death at age x (p_x). These probabilities are calculated as 1-q_x, the probability of survival from one age to the next. Values for q_x are calculated according to Equation 3.2 where (a_x) is the average length of time, in (partial) years, lived within age period x by individuals who died at age x (Wilmoth et al., 2022).

$$p_x = 1 - q_x \tag{3.1}$$

where

$$q_x = \frac{m_x}{1 + (1 - a_x) \times m_x} \tag{3.2}$$

It is common practice to assume that, on average, individuals die at the mid-point of the year and to set a_x to be 0.5 for all age classes other than age 0. This approach is also used in this thesis. At age 0, this assumption cannot be upheld due to distinct mortality dynamics among infants. In this thesis, a_0 was calculated using Human Mortality Database (HMD) methods (Wilmoth et al., 2022). These methods use a formula (Equations 3.3 and 3.4) to calculate a_0 given values of (m_0) (the rate of mortality at age 0).

$$Males : a_0 = 0.14929 - 1.99545 \times m_0 \tag{3.3}$$

$$Females : a_0 = 0.14903 - 2.05527 \times m_0 \tag{3.4}$$

This calculation was performed for each population in which life tables were to be calculated and found to be approximately 0.14 in each case. Thus, for simplicity and consistency a_0 was rounded to 0.14 for the calculation of each life table[1]. The transformation of mortality rates to probabilities of survival, under Equation 3.2 using a_x, is performed to account for the fact that individuals die at different points within each age class, rather than all individuals dying as soon as they reach age x.

Probabilities of death (p_x) are applied to the simulated population progressively to find the number of surviving individuals at each age (l_x) (Eq. 3.5).

$$l_{x+1} = l_x \times p_x \tag{3.5}$$

The number of deaths at each age (d_x), the number of person-years lived at each age (L_x) and the number of person-years lived (T_x) after each age can then be found according to Equations 3.6, 3.7, 3.8, where i is the maximum age in the life table.

[1]This value is exactly the value of a_0 found in publicly available HMD life tables for the population of Scotland as a whole in the years 2001-2021.

$$d_x = l_x - l_{x+1} \tag{3.6}$$

$$L_x = l_x + (a_x \times d_x) \tag{3.7}$$

$$T_x = \sum_{x}^{i} L_x \tag{3.8}$$

These values are then used to calculate life expectancy at age x (e_x) under Equation 3.9.

$$e_x = T_x / l_x \tag{3.9}$$

In this thesis d_x is used in the creation of multiple-decrement life tables and for the calculation of diversity in age at death and (e_0, life expectancy at age 0) is compared against these measures. Other life table variables are used in the calculation of lifespan variation indices.

Small population sizes can reduce the precision of transforming mortality rates to probabilities of survival (Preston et al., 2001). This is thought to be especially a risk in so-called 'complete life tables', in which each age is treated separately. To avoid this, it is possible to use abridged life tables, in which ages are grouped together to pool mortality. However, abridged life tables do not completely avoid the risks of measurements in small populations because they are still based on the same small number of deaths. Furthermore, complete life tables, as used in this study, offer a more precise account of mortality processes. There are no widely agreed upon minimum required population sizes for the calculation of life tables. Although it has been suggested that populations of greater than 50,000 - 100,000 individuals are of a sufficient size (Huang et al., 2020a). All subpopulations assessed in this study are larger than this proposed minimum.

3.3.1.1 Projecting old age mortality rates

Limitations of the NRS population data used in this study (which contain single year populations from 0 to 89 and an open ended class 90+) meant it was only possible to calculate mortality rates up to age 89. While it is possible to calculate life tables using an open ended age class such as 90+, doing so is not generally considered to be best practice (Missov et al., 2016). To overcome this, many in the field of demography use extrapolation techniques to project mortality rates at older ages. In this thesis, Kannisto-Makeham logistic models were used to extrapolate mortality rates from ages 90 to 110 (Vaupel et al., 1998). Mortality rates between ages 50 and 89 were used to inform these models. These methods are standard in the field of demography (Németh & Missov, 2018; Huang et al., 2020b; Fu et al., 2021). The mortality rates produced for the population of Scotland as a whole through this process were validated against those produced by the Human Mortality Database (HMD). This resource is considered one of the best sources of demographic data globally and offers, upon registration, open access to life tables for a wide range of nations around the world (Vaupel et al., 2011; Human Mortality Database, n.d.). The mortality rates at ages 95 to 110 produced by the HMD are smoothed at older ages, using a Kannisto model of old age mortality (Wilmoth et al., 2022). This method uses the number of deaths and population at each age to produce smoothed older age mortality rates, as population data was not available in this study it was necessary to use modelling techniques. They are compared to those calculated for this thesis in 2019 in Figure 3.8. Although these projections do not conform precisely, they were considered to be well correlated and to validate the use of these Kannisto-Makeham models. Other methods for projecting mortality rates were examined, including logistic models which have been used in previous studies, with none considered to be methodologically superior or found to be better aligned with HMD projections (Seaman, 2017). It was not possible to simply use HMD life tables in this thesis because they are released only at the level of Scotland as a whole. Therefore, any analysis of subpopulations would not have been possible. Version 1.9.4 of the package *MortalityLaws*, in *R* version 4.1.3 was used to calculate Kannisto-Makeham logistic models in this thesis (Pascariu, 2022; R Core Team, 2022).

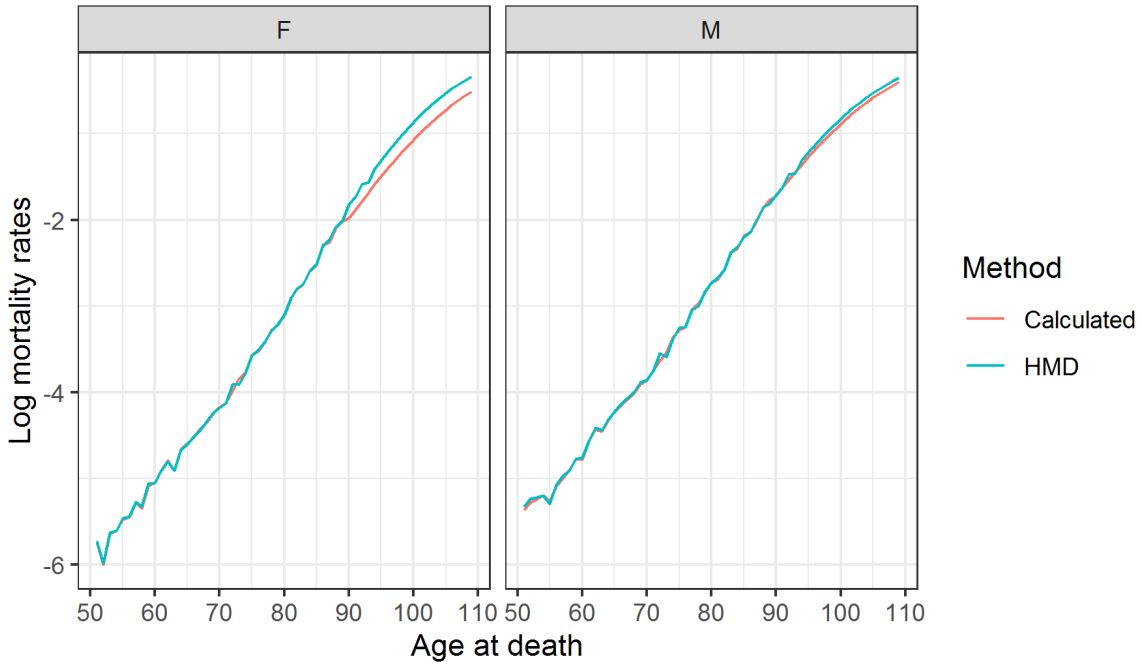


Figure 3.8: Mortality rates used in the calculation of life tables in this thesis for males and females in 2019 in Scotland compared to HMD values. Rates from age 90 onwards were modelled in this thesis while HMD data are modelled from age 95 onwards. HMD modelling takes into account deaths only at ages 80 and older, while modelling carried out in completion of this thesis includes all deaths from age 50 onwards. This may account for some of the disparity between extrapolated mortality rates.

3.3.1.2 Life tables constructed for this thesis

In this thesis, life tables were calculated for each year 2001 to 2021 for males and females separately across the population of Scotland as a whole. These were complete life tables constructed with single-year age classes from ages 0 to 109 and an open-ended age class at 110+. Mortality rates for ages 0 to 89 were calculated using NRS mortality and population data as previously described. Mortality rates for ages 90 to 110 + were calculated using the methods described above in Section 3.3.1.1. Sex-specific life tables were also calculated separately for the population of each SIMD income deprivation quintile and Scottish Government urban-rural class in each year from 2001 to 2019.

The computed all-cause life tables (known as "single-decrement" life tables) were used to create multiple-decrement life tables for analysis of diversity in mortality causes. In the single-decrement life table discussed to this point individuals may leave the cohort for one reason: death. Multiple-decrement life tables are those that describe different ways in which an individual might leave a cohort, for example, by facing different causes of mortality (Artzrouni, 2005). The number of life table deaths at each age (d_x) is multiplied by the proportion of observed deaths due to each three-character ICD-10 cause of mortality in the study population. This produces the number of individuals who, theoretically, would have faced that cause of mortality in the period life table. In effect a vector of (d_x) values is turned into a matrix of ($d_{x,\,i}$) where i is each observed cause of mortality. This process was repeated for each of the life tables described above. Due to the modelling of old age mortality there are ages, particularly at 90+ years, at which no deaths occurred in reality but for which d_x was imputed for the purpose of constructing the single-decrement life table (Section 3.3.1). As it is not possible to create multiple decrements for these ages they are excluded from analysis of diversity in mortality causes. Through techniques such as projection of cause-specific mortality rates it is possible to produce multiple decrement life tables which overcome this lack of age-specific data. To use these techniques appropriately considerable knowledge is required of the aetiology of each cause and consideration must be given to each in turn. Applying such techniques to each of the large range of ICD-10 three character causes recorded in Scotland during the study period was considered outwith the scope of this work. Furthermore, modelling techniques would likely be based on projecting cause-specific mortality rates from ages at which data exists it is therefore probable that any effects on diversity would be small. This is because the measures of diversity used in this thesis assess the relative prevalence (or abundance) of each cause of death. It can be expected that modelling techniques based on existing data would change the prevalence of each cause to a relatively small extent.

3.3.2 Age-standardised mortality rates

Age-standardised mortality rates (ASMRs) are a commonly used measure in population health research. They weight mortality rates at different ages, attributed to all causes or specific causes, by the size of the standard population at that age.

To calculate ASMRs, mortality records were grouped into five year age bands: 0-4, 5-9, 10-14, …, 85-89 and 90+, with bands chosen to match those of the 2013 European Standard Population (as is common practice (Ahmad et al., 2001; Levi et al., 2001; Eurostat, 2013; Wilson et al., 2017)). The population within these age bands was found using the population data described in Section 3.2.2 and used to calculate the age-specific mortality rate for each 5-year age band as the number of deaths divided by the population size, multiplied by 100,000. Each rate was then weighted by multiplying by the proportion of the European Standard Population belonging to the specific age band in question. The resulting age-weighted rates were summed to generate the ASMR for the population in question. The resulting measure is the mortality rate per 100,000 that would be expected in a population if the population age structure corresponded to that of the reference European Standard Population.

The specific populations for which ASMRs are calculated are described in the relevant chapter and each chapter explains whether calculations were performed across all causes or for specific causes of mortality.

3.3.3 Normalised alpha diversity

Diversity in causes of mortality and in age at death is calculated in this thesis using normalised subcommunity alpha diversity under the Reeve et al. (2016) framework which is equivalent to Leinster and Cobbold's diversity of a single community ((Leinster & Cobbold, 2012; Mitchell, 2019)). The rationale for selecting this method for the calculation of diversity is explained in detail in Section 2.3. The equation used in the calculation of normalised subcommunity alpha diversity is expressed using general notation in Equation 2.4. Diversity was calculated in this thesis using version 2.0 of the *rdiversity* package in *R* version 4.1.3 (Mitchell et al., 2020; R Core Team, 2022). In order to calculate diversity in this package, mortality data is transformed into a matrix where rows are ICD-3 character causes of mortality; columns are years of death; and each element is the number of deaths attributed to each cause in each year. In the case of diversity in age at death, individual ages at death replaced causes of mortality as rows. The *metacommunity()* function (*rdiversity*) is then used to create a metacommunity object from this matrix. Normalised subcommunity alpha diversity is then calculated from this metacommunity object using the *norm_sub_alpha()* function (*rdiversity*), with order of diversity (q) determined by the variable "qs".

Three specific measures of diversity are used in this thesis. These are normalised alpha diversity at q values of 1, 2, and ∞. These measures are "effective number of types" formulations of diversity, as discussed in Section 2.3. This means that they produce a value equivalent to the number of equally abundant types needed to create an equally diverse community. These measures can each take values between 1 and the total number of types (causes of mortality or ages at death) which occur in the dataset. Figure 3.9, illustrates the how these measures behave for different simlualted distributions. Figure 3.9A indicates a distribution completely dominated by 1 type (cause of mortality or age at death) this is the least diverse distribution possible, and diversity under all measures is minimised. Figures 3.9B and C indicate more diverse distributions leading to Figure 3.9D where all types have equal abundance (the same number of people die due to each cause of mortality or at each age at death). Figure 3.9D diversity is maximised and all measures of diversity take the same value, which is equal to the total number of types, here 10.

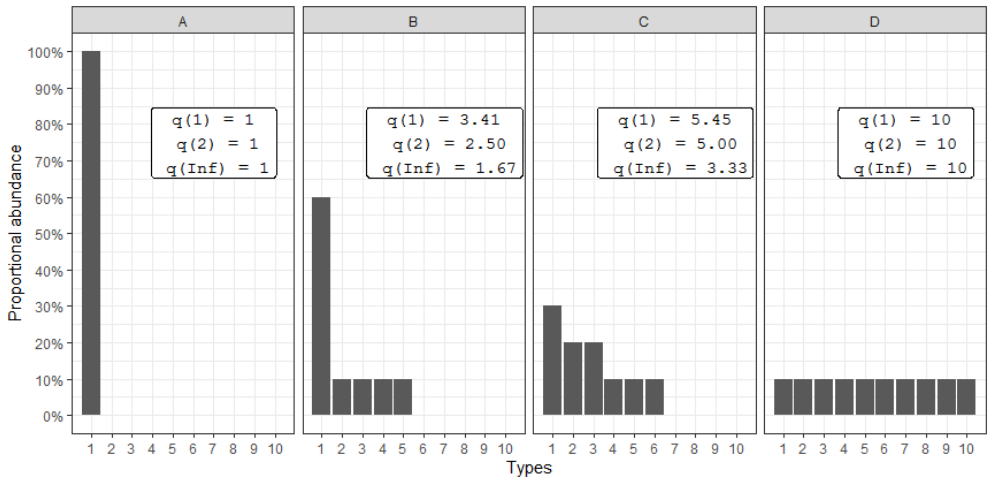


Figure 3.9: The normalised alpha diversity at q = 1, q = 2 and q = Infinity. Each distribution contains ten types of illustrative distributions. In distribution A the entire distribution is concentrated in one type, illustrating the lowest possible level of diversity. In distribution D an equal share of the distribution is attributed to each type, indicating the maximum possible diversity.

The three q values chosen in this thesis were selected because they represent a range of measures that differentially weight the abundance of causes and, therefore, allow different aspects of a distribution to be examined (Izsák, 1986). These measures are related to formulations of three well-known diversity measures that are used often in the literature in the field of ecology. These are Shannon entropy (q = 1), Simpson's index (q = 2), and the Berger-Parker index (q = ∞). The Shannon entropy is a measure of the average uncertainty associated with predicting the

type of an individual chosen at random from the metacommunity (Shannon, 1948). Simpson's index is the probability that two randomly chosen individuals will share the same type (Simpson, 1949). Finally, the Berger-Parker Index is a measure of the dominance of the most common type within the metacommunity (Berger & Parker, 1970). These measures, when used in conjunction, are capable of providing insight into variation in the range of rare types in a community as well as the dominance of the most common types. This is especially an advantage in the case of diversity in causes of mortality where often the most common causes are those which receive the most attention, but changes in the range of rarer causes are also of interest (Carter & Nguyen, 2012; Bergeron-Boucher et al., 2020).

As discussed in Section 2.4, the methods employed by previous studies of diversity in mortality causes are varied. The most closely comparable in terms of time frame and aims to this thesis is the work is Bergeron-Boucher et al. (2020) who use a normalised measure of Shannon entropy. In this thesis comparisons of diversity in mortality causes are made between different subpopulations in Scotland. Comparisons between such groups are made more robust through the use of normalised alpha diversity as it is expressed consistently as an effective number of types throughout. Furthermore, in this thesis I aimed to present an examination of variation in different aspects of the distribution of mortality causes. It would be possible to achieve this through the use of the measures discussed in the previous paragraph in their original form (or even to use a more advanced formulation of these measures). However, using their equivalents in the Reeve et al. (2016) framework allows for them to be calculated and expressed using the same terms and notation which aids in the interpretation of results across different measures.

3.3.3.1 The use of life tables in the calculation of diversity in causes of mortality

In all cases in Chapters 4, 5, and 7, diversity in causes of mortality is calculated using the distributions of causes of mortality extracted from multiple-decrement life tables (Section 3.3.1). In Chapter 6, diversity is measured in counts of the directly observed causes of mortality in the Scottish population without transformation through the calculation of life tables. Life tables were used for the most part in this thesis because of the advantages they offer in comparison between populations. In creating a standard starting cohort, life tables effectively standardise

population structures (Jansen et al., 2012). This is suggested in the literature as an approach to make comparisons between populations with differing population structures more robust (Coale et al., 2013). For further commentary on the merit of the use of life tables, see Chapter 5. In Chapters 4 and 5, where life tables were used, comparisons were made between subpopulations such as income deprivation quintiles and urban-rural classes. These methods were used to ensure comparisons between these subpopulations were reliable.

Life tables are not used in Chapter 6 because within that chapter, diversity is measured within individual months, meaning that the number of deaths in each population was small. Calculating life tables with such small numbers of deaths is problematic due to the need for a mortality rate for each age class to properly capture demographic forces across the life table (Preston et al., 2001). Methods to overcome this exist; however, they are generally founded upon modelling and extrapolating mortality rates in ages where no deaths occurred. There are several limitations to these methods, especially in for calculating multiple-decrement life tables, because when mortality rates are projected into an age in which no deaths occur, there is no cause of death information for this age which can be used in the calculation of multiple decrements. The objective of this chapter was to study the effect of the COVID-19 pandemic on the causes of mortality in Scotland. Therefore, the analysis was sensitive to the effects of the introduction of this cause. Although our knowledge is improving over time, our understanding regarding those who are most vulnerable to COVID-19 remains imperfect and any technique to model its effect across ages would be vulnerable to this. Furthermore, any techniques for modelling life tables based on previous years would be inappropriate as these years would not be informative regarding mortality during the pandemic. Figure 3.10 shows diversity in mortality causes in Scotland from 2001 to 2019 in males across all ages calculated using the distribution of mortality causes drawn from life tables and from observed raw death counts in those years. It demonstrates that these methods produce trends that are qualitatively indistinguishable. Given that the effect of using life tables compared to observed data is shown to be small, it was considered appropriate to simply measure diversity in observed distributions of counts of mortality causes in Chapter 6 rather than employ more complex methods to model life tables.

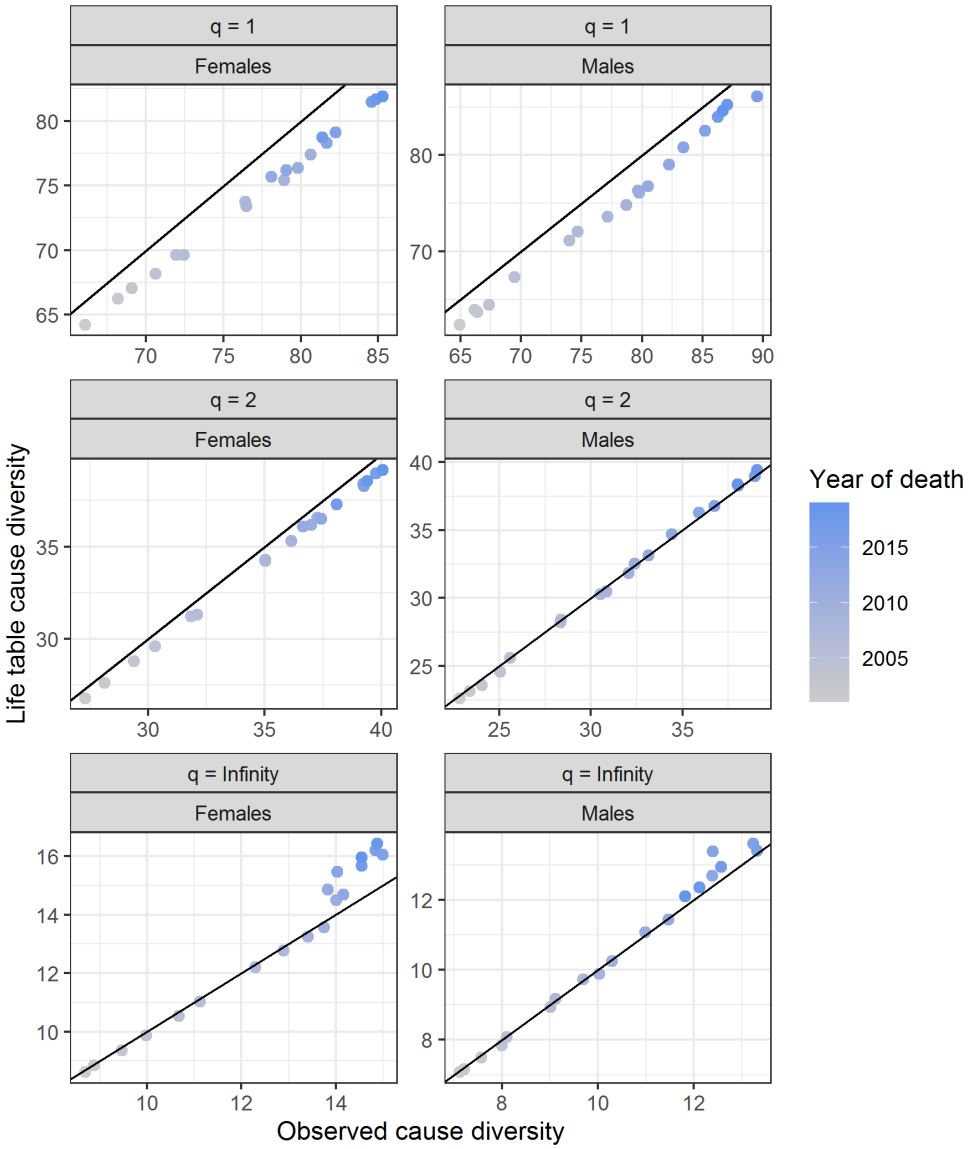


Figure 3.10: Comparison between diversity in causes of mortality, calculated from the observed distribution of causes of mortality (x-axis) and from the distribution of causes of mortality extracted from multiple-decrement life table data (y-axis), in males and females over the period 2001 to 2019. The diagonal line indicates a situation under which diversity in causes of mortality calculated from multiple-decrement life table data is exactly equal to diversity in causes of mortality calculated from the observed distribution of causes of mortality.

Previous studies of diversity in causes of mortality have assessed the distribution of mortality causes extracted from life tables (Bergeron-Boucher et al., 2020; Trias-Llimós & Permanyer, 2023) and from observed mortality data (Izsák, 1986). My work suggests that both methods are appropriate for the measurement of diversity in causes of mortality and that the choice of a method should depend on the objective and situation of research. Calculating diversity in distributions of mortality causes extracted from multiple-decrement life tables can be advantageous

when comparisons are to be made between populations with differing age structures. Diversity calculated in observed distributions of mortality causes may be useful if life tables cannot be accurately calculated or if one aims to examine the distribution of mortality causes without any potential biasing effect introduced by the transformations inherent in the calculation of by life tables.

3.3.3.2 The calculation of diversity with small sample sizes

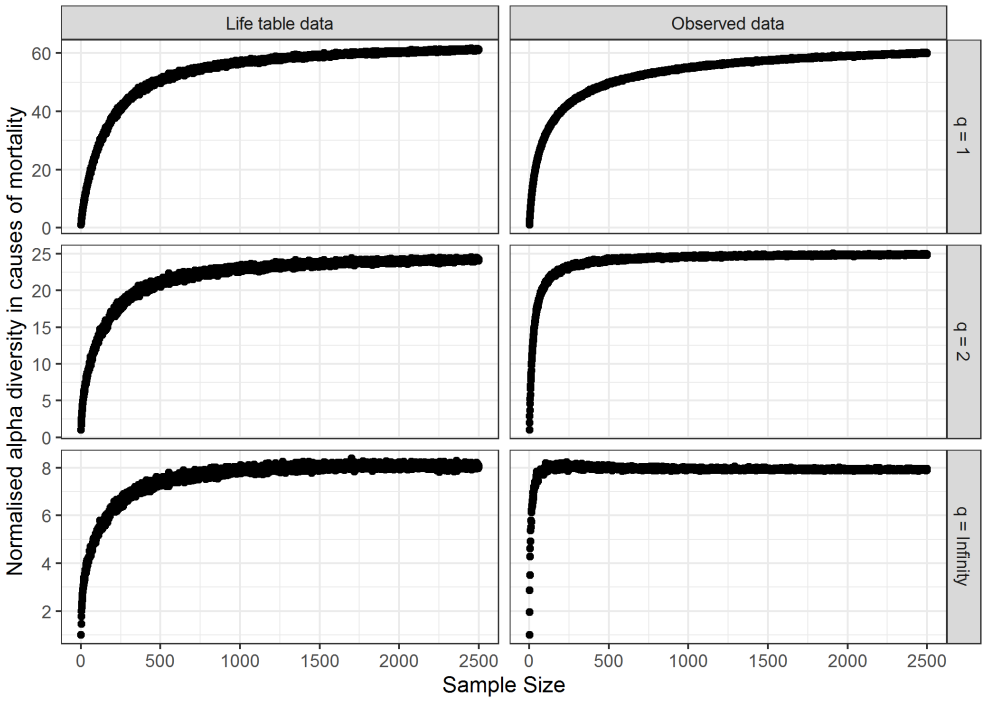


Figure 3.11: Diversity accumulation curve showing normalised alpha diversity at *q* = 1 in mortality causes for samples with increasing numbers of deaths drawn from the deaths among the female population of Scotland in 2001. Each point represents the mean of diversity calculated from 200 simulated populations (samples) with the corresponding number of deaths.

Diversity measures are limited by their sensitivity to small sample sizes. Inherently, a smaller number of individuals can have a reducing effect on the variety of types which appear in a population. Figure 3.11 demonstrates this by showing a diversity accumulation curve, equivalent to a species accumulation curve approach that is commonly used in ecology (Leinster & Cobbold, 2012). To create this curve, the female mortality records in Scotland in 2001 were randomly sampled with sample size from 1 to 2000 and the normalised alpha diversity of the causes of mortality in this simulated subpopulation was calculated. This process was repeated 200 times

for each sample size and the mean of calculated diversities was found. Diversity increases dramatically at sample sizes of <250 before beginning to level off and does not completely plateau even for sample sizes around 2000. This pattern is characteristic of diversity accumulation curves (Yurkov et al., 2011). The use of life tables does not correct for this issue because the distribution of causes of mortality from which multiple decrements are calculated are derived from small numbers of deaths.

In most cases throughout this thesis, diversity is measured in populations in which the number of deaths was greater than 1000. Figure 3.11 suggests if the number of deaths is greater than 1000, any effects of sample size on diversity would be small. As discussed above, this applies to all analysis of diversity in mortality causes across all ages in the population of Scotland and all Scottish subpopulations in this thesis. In analysis of diversity in mortality causes within populations divided by age at death there are cases in which this threshold is not met. Analysis of diversity in mortality causes in the Scottish population with deaths divided into twenty-year age ranges is presented in Chapter 4 alongside evidence suggesting that despite smaller number of observations, the conclusions of this analysis are upheld. Analysis of diversity in mortality causes within subpopulations divided by age is also performed and is presented in Appendix C and discussed in Chapter 4. For this analysis all deaths among those aged 0-39 are pooled to increase sample sizes.

3.3.3.3 Alternative measures of diversity

Various other methods for the measurement of diversity in mortality causes. In order to test the sensitivity of the results reported in this thesis to other methods of calculating diversity in mortality causes, in Appendix B.1 three alternative measures of diversity are assessed. These are Shannon's index, Shannon's entropy and a measure related to measures of evenness referred to here as uniformity. The formulations of these methods are described in Appendix B.1. Uniformity is distinct as a method for measuring diversity from those used in this thesis in that it accounts for the entire distribution of mortality causes in each year. Uniformity explicitly includes causes of mortality which have zero values (i.e. a cause of mortality for which a death has not occurred in a given year) in the calculation of diversity which makes this measure sensitive to causes of mortality which do not occur in every year. Diversity in underlying causes of mortality is measured using

the three measures described here for deaths at all ages in the population of Scotland as a whole, in SIMD income distribution quintiles and Scottish government urban-rural classes. Diversity in underlying causes of mortality is also measured among deaths within twenty year age ranges across Scotland. Finally, diversity in contributory causes of mortality is measured in deaths across Scotland. Further details of the methods used in this analysis are detailed in Appendix B.1. In each case diversity is measured in distributions of causes of mortality extracted from multiple decrement life tables. Cases where trends under this alternative analysis differ from those calculated under normalised alpha diversity are discussed in Chapters 4 and 7. The potential impact of these differential trends are discussed throughout this thesis.

3.3.4 Lifespan variation measures

A number of indices of variation in age at death are used in the literature, and six of these are calculated in Chapter 5 to allow comparison between diversity in age at death (lifespan diversity) and established measures in this field. Several studies have previously made comparisons between the existing indices, showing very high degrees of correlation, and indeed, van Raalte and Caswell ((van Raalte & Caswell, 2013)) suggest many are "apparently interchangeable" (Kannisto, 2000; Evans, 2001; Cheung et al., 2005; Vaupel et al., 2011; van Raalte & Caswell, 2013). Table 3.2 describes the indices examined in Chapter 5. These were calculated using the spreadsheet developed by Shkolnikov and Andreev (2010) which uses values from life tables to calculate indices of variation in lifespan, the calculation of which described in detail by those authors. The measures of diversity in age at death proposed in this thesis are also described in Table 3.2; the calculation of these measures is discussed in detail in Chapter 5.

Table 3.2: The names and definitions of the indices of variation in age at death examined in Chapter 5

Specialised measures of variation in lifespan	
Lifespan variation (e†)	Lifespan variation is calculated as the sum of the residual life expectancy at each age weighted by the proportion of deaths which occur at this age. Greater lifespan variation means that individuals are dying at a younger age than would be expected (Shkolnikov et al., 2003; van Raalte & Caswell, 2013; Seaman et al., 2019).
Lifetable entropy	Lifetable entropy is a relative formulation of lifespan variation. It is, effectively, lifespan variation standardised by life expectancy which rescales this measure to between 0 and 1 (Keyfitz & Caswell, 2005; Shkolnikov & Andreev, 2010; Aburto et al., 2019).
Gini coefficient	The Gini coefficient can take values from 0 to 1 with lower values indicating greater equality in lifespan (Keyfitz & Caswell, 2005). It is the mean of the interindividual differences in lifespans, divided by life expectancy (Evans, 2001; Shkolnikov et al., 2003; Shkolnikov & Andreev, 2010; van Raalte & Caswell, 2013).
Traditional statistical measures of variation	
Inter-quartile range	A measure of the distance (in years) between the lower and upper quartiles of the distribution of ages at death (Wilmoth & Horiuchi, 1999; Shkolnikov & Andreev, 2010).
Standard Deviation	The traditional statistical standard deviation of the distribution of ages at death (Shkolnikov & Andreev, 2010).
Coefficient of variation	The ratio of the standard deviation of the ages at death to the mean age at death (Shkolnikov & Andreev, 2010).
Diversity measures proposed in this thesis	
Lifespan diversity	Similarity-sensitive normalised alpha diversity at $q = 1$ in age at death. This measure can take any value between 1, when all deaths occur at a single age, and the total number of ages at death within the population, when an equal number of deaths occur at each age.

3.3.5 ARIMA models

In Chapter 6, Autoregressive Integrated Moving Average (ARIMA) models were used to assess seasonality in diversity in causes of mortality and to produce a counterfactual projection of diversity in causes of mortality in each month in 2020 and 2021 given trends in previous years. This was performed to give a prediction of

diversity in mortality causes in the absense of the COVID-19 pandemic. It is known that mortality rates associated some causes of death are seasonal, as are all-cause mortality rates, and it is therefore necessary to consider the potential for a seasonal pattern in diversity in causes of mortality (Gemmell et al., 2000). ARIMA models were chosen for use in this chapter as they can be applied flexibly, meaning that they could capture this seasonality, but also that even in situations where seasonality was not observed they could be used to estimate a non-seasonal projection. ARIMA models have previously been used in analysis of a range of outcomes, although most commonly outcomes expressed as rates (e.g. per 1,000 population) or counts; however, in the literature, it has been suggested that these models can be generalised to any continuous outcome (Wu et al., 2007; Ramezanian et al., 2019; Alzahrani et al., 2020), making them appropriate for the analyses of diversity here. Unlike other methods for time series analysis that typically regress an outcome variable on time, ARIMA models predict future values of a variable by regressing on the values of that variable at previous time points i.e. following an auto-regressive approach (Schaffer et al., 2021). The specific cases in which ARIMA models were used are discussed in Chapter 6. These models were fit according to the methodology described by Schaffer et al. (2021).

ARIMA models are described using the following notation: ARIMA$(p,d,q)(P,D,Q)_s$. They are composed of seasonal (s) and non-seasonal models, each made up of three components: the autoregressive model (AR) in which the dependent variable is predicted by lagged values of itself (p in the nonseasonal model, P in the seasonal model); the moving average model (MA) in which the dependent variable is predicted by lagged values of an error term (q,Q); and differencing, which is used to create a stationary time series (d, D) (Schaffer et al., 2021). Through varying p, q, P, and Q, the order of the autoregressive and moving average parts (i.e. the number of lags to include) of the ARIMA model can be changed. ARIMA models are selected and fit to data through a process of choosing the order of these parts and by choosing the degree of differencing, achieved by varying d and D (Hyndman et al., 2022).

The seasonal component of an ARIMA model is introduced in the model by predicting the dependent variable from lagged values of itself at a regular interval which is defined as the season, where s is the number of observations which make up a full year, i.e. 12 for monthly observations (Hyndman et al., 2022). The ARIMA models used in this thesis are all constructed with a season of 12 meaning that in the seasonal component of the model values of diversity in mortality causes in each month are regressed on values of diversity in that month in previous years.

The Schaffer et al. (2021) methodology suggests the following steps to fit an ARIMA model. 1) The data should be plotted to understand trends and to examine outliers. 2) Data should be transformed to stabilise variance if such transformation is necessary. 3) The third step is model selection, a process which can be aided by the use of statistical packages in *R* and should be performed using Autocorrelation function (ACF) and partial autocorrelation function (PACF) plots; to do this, first, if necessary, the data should be made stationary by adding a first order difference (d=1) and if seasonality is present a seasonal difference should be applied (through varying D). The ACF/PACF plots should then be used to decide on the order of the AR/MA models (i.e. the number of lags to include) and p, P, q, and Q should be adjusted accordingly. For example, positive autocorrelation at lag 1 indicates that an AR term is needed and negative autocorrelation at this point indicates MA terms are needed. Up until this point, all that is determined is whether the model should contain seasonal terms and which AR/MA terms are needed. The next step is to estimate the model a variety of AR/MA terms and use information criteria to find the best fit model. 4) The residuals of the chosen model are then checked for autocorrelation using a Ljung-Box test. The null hypothesis of these tests is that no autocorrelation exists in the residuals. If autocorrelation is suggested by this test then the model in question gives a poor fit and step 3 should be repeated. This process is described by Schaffer et al. (2021) as iterative and it should be used to find the most parsimonious model with the smallest AR/MA terms, which is well fit and does not generate autocorrelated residuals.

The process described above was used in this thesis to fit ARIMA models as described in Chapter 6. In some cases where this proved difficult, the function *auto_arima()* from the *forecast* was used to inform decisions, as suggested by Schaffer et al. (2021). Where the automated process was used, the models produced were compared against those produced manually and the model which was best fit was taken forward for analysis.

Chapter 4

Trends in normalised alpha diversity in mortality causes in Scotland and Scottish subpopulations

4.1 Background

Advances in medical and public health sciences have allowed for almost continual reductions in total mortality rates, and resultant increases in life expectancies, in developed countries since at least the 1950s (Moser et al., 2005). These improvements in all-cause mortality rates have been explained through theories of epidemiological transition(Omran, 1971; Mackenbach, 1994). The most recent of these transitions in Scotland and most developed nations has been characterised by falling mortality rates associated with the most common causes of mortality; heart disease and cancers. This has meant these causes have accounted for a smaller proportion of deaths. Meanwhile, the prevalence of other less common causes of mortality such as degenerative diseases and other diseases of old age increased. This represents fundamental changes in the distribution of mortality causes.

Stalling improvements in various measures of population health such as life expectancy and mortality rates have been observed in Scotland in the 2010s (Ramsay et al., 2020; Walsh et al., 2020b; Wraw et al., 2020). Ramsay et al. (2020) link stalling mortality rates in this period to slowing improvements in mortality rates associated with circulatory diseases. They further observe worsening mortality rates associated causes such as dementia and Alzheimer's disease as well as in-

creased rates of drug-related deaths. The authors suggest that greater research is required to understand changing mortality trends across ages and causes in Scotland. Through examining diversity in mortality causes it is possible to examine how changing mortality patterns have affected variation in the distribution of mortality causes.

Previous research has shown diversification in causes of mortality at the national level in high income countries, indicating an increase in the range of causes likely to be faced by the population (Bergeron-Boucher et al., 2020). The consequences of such an increase are discussed in detail in Section 2.4.2 but are closely linked to uncertainty at the individual-level and fragmentation at the societal level. In both cases an increase in the diversity of mortality causes may have negative implications for population health and public health and health care systems.

Previous studies of diversity in mortality causes are mostly limited to examination of trends in diversity at the national level, with examination of subpopulations limited to groupings by sex and age. Although, a gradient has been observed by education in the USA with lower diversity in mortality causes among groups with the highest educational attainment (Trias-Llimós & Permanyer, 2023).

Health outcomes vary widely between subpopulations, whether they are separated by socioeconomic or geographic factors. In this chapter, diversity in mortality causes is measured within Scottish subpopulations defined by SIMD income deprivation quintile and Scottish Government urban-rural class. The effect of deprivation on health in Scotland has been described in a large number of studies; compared to the population as a whole, those in more deprived areas face higher rates of all-cause mortality. Further mortality rates in deprived areas have improved more slowly over time. Additionally, wide variation has been found in cause-specific mortality rates across deprivation quintiles (Brown et al., 2019). Cause-specific mortality has also been shown to vary across the urban-rural gradient in Scotland, with key differences in the burden of suicide and heart disease (Levin & Leyland, 2005; Walker, 2021; Casant & Helbich, 2022). The dynamics of change over time in cause-specific mortality within Scottish subpopulations are complex and analysing diversity in mortality causes allows for an examination of overarching trends. Further, recognising populations in which diversification is occurring most quickly makes it possible to examine where pressures on health care and public health systems may be felt most acutely and thus be better prepared to face this pressure.

4.2 Research questions and objectives

This chapeter had the following objectives:

- To examine the trends in normalised alpha diversity of mortality causes within Scotland.
- To explore differences by sex, age at death, SIMD income deprivation quintile, and Scottish Government urban-rural class.
- To identify the causes and groups of causes driving trends in normalised alpha diversity of mortality causes.

Therefore, this chapter aimed to answer the following research questions:

- What were the national temporal trends in diversity in mortality causes in deaths across all ages in Scotland from 2001 to 2019?
- What were the national temporal trends in diversity in mortality causes at different ages at death in Scotland from 2001 to 2019?
- Did diversity in mortality causes differ between subpopulations, grouped by area-level deprivation and urban-rural class, in Scotland from 2001 to 2019?
- What were the subpopulation level trends in diversity in mortality causes in Scotland from 2001 to 2019?
- Which causes of mortality drove trends in the diversity of mortality causes at the national and subpopulation level in Scotland from 2001 to 2019?

4.3 Methods

4.3.1 Data

Mortality data for the years 2001 to 2019 was extracted and processed as described in Methods Section 3.2.1. Annual mid-year small area population estimates for 2001 to 2019 for Scotland as described in Methods Section 3.2.2 were used.

4.3.2 SIMD income deprivation quintiles

SIMD income deprivation quintiles were created, as described in Section 3.2.3, to examine the effect of deprivation on diversity in mortality causes in this chapter. Mortality records for the years 2001 to 2009 were linked to SIMD 2006 and those for the years 2010 to 2019 were linked to SIMD 2016.

4.3.3 Scottish Government urban-rural classes

The Scottish Government six-fold urban rural classes, as described in Methods Section 3.2.4, were used in this chapter. Datazones were used to link mortality records to urban-rural class; 2005/2006 classes were used for the years 2001 to 2009 and 2016 classes were used for 2010 to 2019.

4.3.4 Life tables

Multiple Decrement life tables were constructed, as described in Section 3.3.1, for Scotland as a whole and for each deprivation quintile and urban-rural class. Life tables used single-year age classes for males and females separately, and were calculated for each year from 2001-2019.

4.3.5 Normalised alpha diversity

Normalised alpha diversity was calculated under the Reeve et al. (2016) framework described in Section 3.3.3 from distributions of mortality extracted from multiple-decrement life tables. Naïve-similarity normalised alpha diversity in mortality causes at the level of ICD three-character codes is examined at q values of 1, 2, and ∞. Diversity was calculated in each year from 2001 to 2019 in males and females separately across Scotland as a whole and in each deprivation quintile and urban-rural class. Calculations were performed for deaths across all ages and for deaths in 20-year age groups. These ranges were 0 to 19, 20 to 39, 40 to 59, 60 to 79 and 80+ for analyses across Scotland, whereas for analysis of subpopulations, the first two age groups were combined to ensure sufficient data in each subpopulation.

4.3.6 Additive value

Assessing the contribution of individual, fine-grained causes of mortality, or indeed groups of causes, to normalised alpha diversity is complicated by the nature of this measure and the way it takes into account the proportional prevalence of causes. Previously Bergeron-Boucher et al. (2020) have used cause-deleted life table methods (effectively removing an individual cause from the calculation of multiple-decrement life tables) to assess the contribution of cardiovascular diseases to national trends in diversity in mortality causes by examining these trends with this cause removed. I have developed a novel measure, additive value, which I use in this chapter to examine the contribution of fine-grained causes of mortality and ICD-10 Chapters, to diversity in causes of mortality. In this chapter, I introduce the use of 'additive value' to examine the contribution of individual, fine-grained causes of mortality and groups of causes (namely ICD-10 Chapters), to diversity in causes of mortality. Additive value is a simple but useful measure of the effect each cause has on diversity, essentially describing the value diversity would take if all deaths due to a particular cause of mortality were excluded from analysis.

To calculate additive value, each cause (or group) in turn is removed from the distribution of causes and diversity is recalculated without this cause. Additive value is calculated by dividing the cause-excluded diversity by all-cause diversity and subtracting the resultant proportion from 1. Therefore additive value can, in theory, take any value from $\frac{S-n}{S}$ to $-\frac{S-n}{S}$ where *S* is the total number of individual causes of mortality and *n* is the number of individual causes excluded.

The additive value can be either positive or negative. A positive value means that the cause in question makes the distribution of causes more even and removing this cause reduces diversity. Negative values occur for dominant causes the exclusion of which increases diversity by increasing the evenness of the distribution. To use an example, if cause *x* had an additive value of 0.1 it would indicate that diversity in mortality causes was reduced by 10% with cause *x* excluded. Therefore, if cause *y* had an additive value of 0.2 it would mean that the removal of cause *y* from the distribution reduced the calculated value of diversity twice as much as the removal of cause *x*. The additive value of causes of mortality is not additive, meaning that diversity in mortality causes is not equal to the sum of the additive value of all causes of mortality.

The additive value of each of the ICD-10 three-character codes recorded during the study period, as well as each ICD-10 Chapter, was calculated across Scotland and in each of the subpopulations examined in this chapter in each year from 2001 to 2019. Additive value was calculated across in deaths all ages and in each of the twenty-year age groups described in section 4.3.5.

4.3.7 Age-standardised mortality rates and proportional prevalence

All-cause and cause specific age-standardised mortality rates (ASMRs) were calculated for Scotland as a whole and in each subpopulation in each year from 2001 to 2019 as described in Section 3.3.2. Both all-cause and cause-specific ASMRs were calculated for deaths across all ages and in each twenty-year age group. Cause-specific ASMRs were calculated for each three-character ICD-10 code which was recorded in each population in a given year.

The proportional prevalence (referred to further as prevalence) of each ICD-10 mortality cause (and each ICD-10 Chapter) recorded during the study was calculated by dividing the cause-specific ASMR by all-cause ASMR. These calculations were carried out separately for males and females across Scotland as a whole, and for each subpopulation. This was repeated for deaths across all ages and in the twenty-year age groups described in Section 4.3.5.

4.3.8 Software

Analysis in this chapter was carried out in *R* version 4.1.3 (R Core Team, 2022). Data manipulation and processing was performed using version 1.3.1 of the *tidyverse* package, plots were created using version 3.3.5 of the *ggplot2* package and version 1.1-2 of the package *RColorBrewer* (Neuwirth, 2014; Wickham, 2016; Wickham et al., 2019). Calculation of diversity was carried out using version 2.0 of the *rdiversity* package (Mitchell et al., 2020).

4.4 Results

The results section of this chapter is structured as follows. All analyses are conducted separately by sex. Firstly, to assess overall trends, diversity in causes of mortality is examined across Scotland as a whole. Then, I consider trends according to the following three breakdowns: (1) income deprivation quintiles; (2) urban-rural class; and (3) 20-year age group. For all of these, I examine diversity at q values of 1, 2, and ∞. In each case, diversity in mortality causes is calculated from distributions of mortality causes extracted from multiple decrement life tables.

The values of q of 1, 2, and ∞ are used as they represent a range in the weight placed on the prevalence of causes. At $q = 1$, which is the least conservative measure, all causes are weighted exactly by their prevalence and relatively rare causes are thus given prominence. Increasing q values indicate a greater weight on the prevalence of causes and at $q = \infty$, the most conservative measure, only the most common causes are considered. This provided the ability to study different

aspects of the diversity in distributions of mortality causes. In each case, the additive value to diversity of causes of mortality, both individually and grouped as ICD-10 Chapters, is examined. Through additive value I was enabled to examine the mortality causes which have driven changes in diversity in causes of mortality.

4.4.1 Trends in normalised alpha diversity at q = 1, 2, and ∞ in deaths across all ages

4.4.1.1 Scotland

In this section, the evolution over time of normalised alpha diversity in causes of mortality, calculated in Scotland from 2001 to 2019 is calculated for males and females separately across deaths at all ages using distributions of mortality causes extracted from multiple decrement life tables. Diversity is studied at three q values - 1, 2, and ∞. First, to help illustrate the distributions discussed in this section, Figure 4.1 shows the share of deaths attributed to each of the 100 most common causes of mortality in males and females in Scotland in 2001 and 2019 separately. The share of deaths assigned to each cause was extracted from multiple-decrement life tables for each population. Figure 4.1 shows a significant reduction in the share of deaths assigned to the most common causes in both males and females over this period.

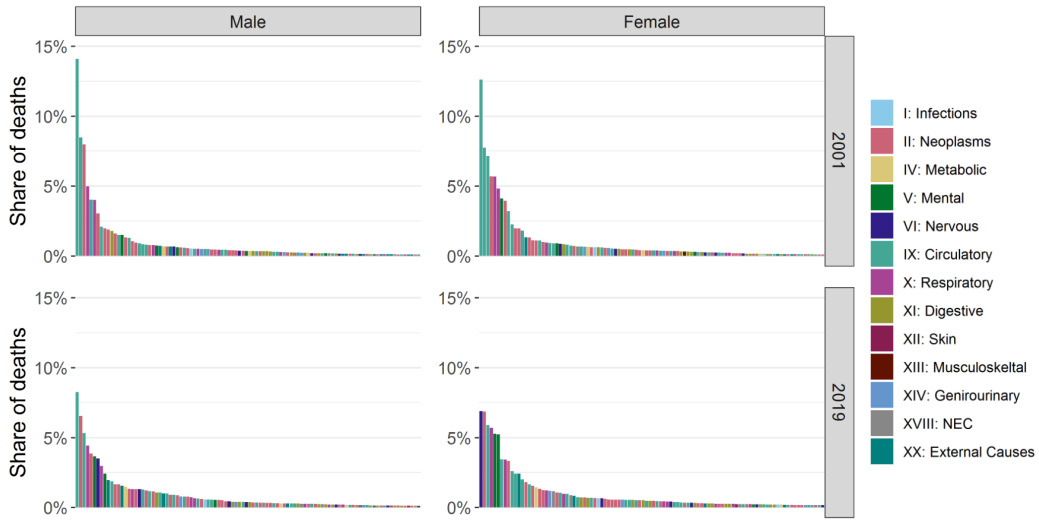


Figure 4.1: Bar plot showing the share of deaths, calculated from multiple-decrement life tables, attributed to each of the 100 most common mortality causes in males and females in 2001 and 2019 in Scotland. Bars are coloured by the ICD-10 Chapter in which the cause resides.

The first measure of diversity considered in this chapter, normalised alpha diversity at $q = 1$ is an equivalent measure to the exponent of Shannon entropy (Shannon, 1948), which describes the average uncertainty associated with predicting the type of a single individual picked randomly. It is the least conservative of the measures of diversity considered here in that it weights rare causes of mortality most heavily in the calculation of diversity and is therefore the most sensitive to changes in the number of deaths caused by these rare of causes of mortality.

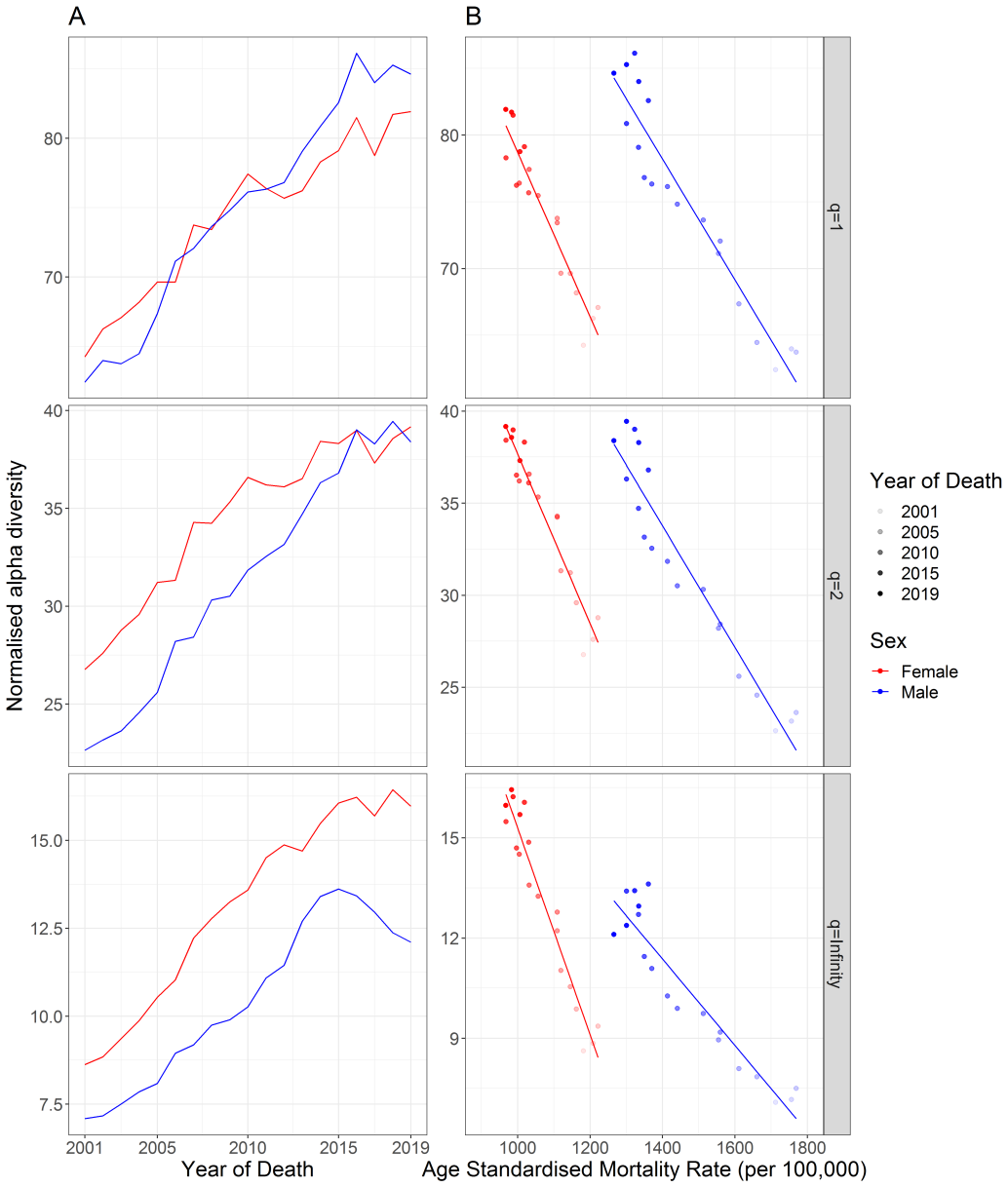


Figure 4.2: Trends in normalised alpha diversity at values of *q* equal to 1, 2 and Infinity against A) Year of death and B) age-standardised mortality rate, in males and females in Scotland across all ages. In panel A lines and in panel B points, represent values in each year from 2001-2019.

Figure 4.2A (top panel) shows trends in the normalised alpha diversity at q=1 of causes of mortality. Initially, diversity was higher for females who faced 64.2 effective causes of mortality in 2001 compared to 62.4 for males. This means that diversity at q = 1 in the distribution of mortality causes was effectively equivalent to approximately 64 equally prevalent causes of mortality among females and 62 among males [1]. However, diversification occurred more quickly among males

[1]The definition of the "effective number of types" formulation is explained further in previous chapters

than females and after 2012 a more diverse distribution of mortality causes was observed in males in every year under diversity at q = 1. The top panel of Figure 3B shows the relationship between normalised alpha diversity at q = 1 in causes of mortality and all-cause age-standardised mortality rates from 2001 to 2019, in males and females, for deaths across all ages. A clear negative relationship is present in both sexes meaning that as rates of mortality have fallen, diversity in mortality causes at q = 1 has increased.

Normalised alpha diversity at q=2 is equivalent to the inverse of Simpson's index (Simpson, 1949), a measure of diversity which describes the probability that two randomly chosen individuals will die of the same recorded mortality cause. Diversity at q = 2 is described as more conservative than at q = 1, with more abundant types given more weight in the calculation of diversity and less significance placed on rare types.

The second panel of Figures 4.2A and 4.2B show normalised alpha diversity of mortality causes at q=2 against year of death and age standardised mortality rate respectively. Similarly to diversity at q=1, males faced a less diverse distribution of mortality causes in 2001 and increases to diversity in mortality causes occurred more quickly in males than females. Under this measure, the greater rate of diversification in males lead to a convergence later in the study period than under diversity at q = 1. This indicates that faster diversification in males is likely associated with rare causes of mortality becoming more prominent more quickly in males than females. At the level of the most dominant causes trends were more similar between sexes.

The last measure of diversity considered in this chapter is the most conservative, with the most weight given to abundant types and, in effect, rare types disregarded completely. Normalised alpha diversity at $q = \infty$ is equivalent to the inverse Berger-Parker Index (Berger & Parker, 1970), a measure of how dominant the most common type is within a population. As with the previously discussed measures, the maximum possible value is equal to the number of types in the population. As with all measures other than q = 0, it can only reach this number if all types are equally distributed. Trends under diversity $q = \infty$ are important because a small number of causes of mortality account for a large proportion of all-cause mortality rates. For example, among females in 2001, 13% of the recorded causes of death when combined represent for 86% of the total ASMR (Appendix A.2).

Under normalised alpha diversity in mortality causes at $q = \infty$ males again faced less diverse mortality causes in 2001 than in 2019 as shown in the lower panel of Figure 4.2A and B. However, males continued to face less diverse causes across the period. In both sexes, relatively steady increases in the diversity of mortality causes at $q = \infty$ are observed up to the early 2010s. In the following years, diversity has stalled in females and fallen slightly in males. This measure is more sensitive to trends in the proportion of deaths attributed to the most common cause than diversity at $q = 1$ and $q = 2$. Therefore, trends in diversity at $q = \infty$ following 2015 indicate that among males an increased share of deaths has been attributed to the most common causes.

The trends noted here mean that over time, for most of the study period, fewer deaths occurred due to the most dominant mortality causes and that a redistribution occurred with deaths attributed to a range of other causes of mortality, creating a more even distribution. Taken together, the trends under diversity in mortality causes at $q = 1$, $q = 2$, and $q = \infty$ indicate that while the proportional prevalence of rare causes of mortality has continued to increase following 2015 (as shown by changes in $q = 1$ and $q = 2$), the prevalence of the most common causes is no longer reducing (as seen for $q = \infty$).

In Appendix B.1 alternative methods for the calculation of diversity are used to measure diversity in causes of mortality in males and females in Scotland from 2001 to 2019. This analysis suggests that under a measure related to evenness, which explicitly accounts for zero values in the distribution of mortality causes (causes of mortality which do not occur in every year in the study period), the trend for increasing diversity throughout the study period, as described above, is not observed. Instead no clear trend in diversity in mortality causes is observed from 2001 to 2014 in females while in males diversity increases slightly over this period. A marked increase in diversity in mortality causes is observed from 2015 onwards in both sexes. This leads to greater diversity in 2019 than in 2001 as observed under normalised alpha diversity above but indicates different trends during the study period when zero values are accounted for.

4.4.1.2 Scottish subpopulations

Here, trends in diversity in mortality causes are compared between deprivation quintiles and urban-rural class. Comparison of the relationship between diversity and mortality rates (as in Figure 4.2B) are not shown because the relationship differs little across subpopulations.

Deprivation Quintiles

To illustrate the distributions of mortality causes in SIMD income deprivation quintile, Figure 4.3 shows ridge plots of the share of deaths attributed to each of the 100 most common causes of death in males and females separately in 2001 and 2019. As across Scotland above, this shows a reduction in the share of deaths attributed to the most common causes from 2001 to 2019 in most cases.

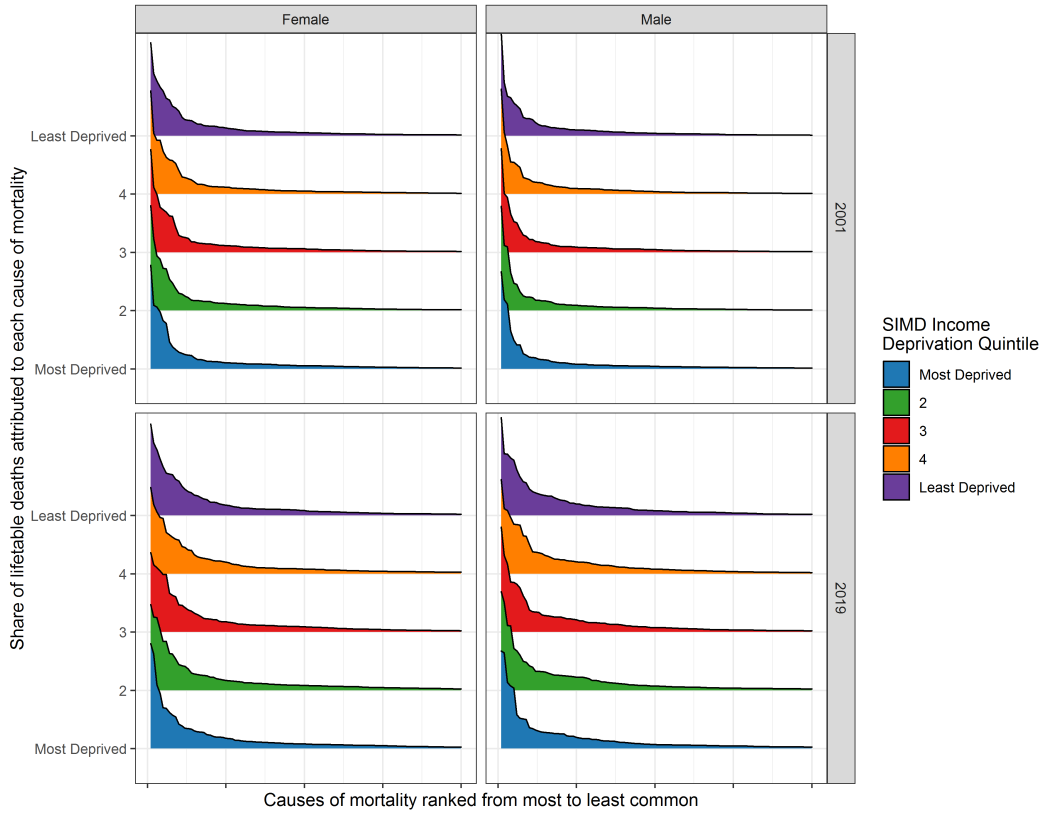


Figure 4.3: Illustrative ridge plot showing the share of deaths, calculated from multiple-decrement life tables, attributed to each of the 100 most common mortality causes in males and females in 2001 and 2019 in each SIMD income deprication quintile.

Across deprivation quintiles, broadly the same trends in diversity in causes of mortality are observed as in the Scottish population as a whole at each value of q. However, some important insights are found by studying differences and divergences in trends between quintiles shown in Figure 4.4.

First, at $q = 1$, all deprivation quintiles are found to have less diverse causes of mortality than across the country as a whole in most years. Because this pattern is not noted at higher values of q, it suggests that it is driven by differences in rare causes. Specifically, it suggests that rare causes of mortality are not spread equally across deprivation quintiles. Rather, individuals in each quintile face different sets of rare causes which combine to produce more diversity across the country.

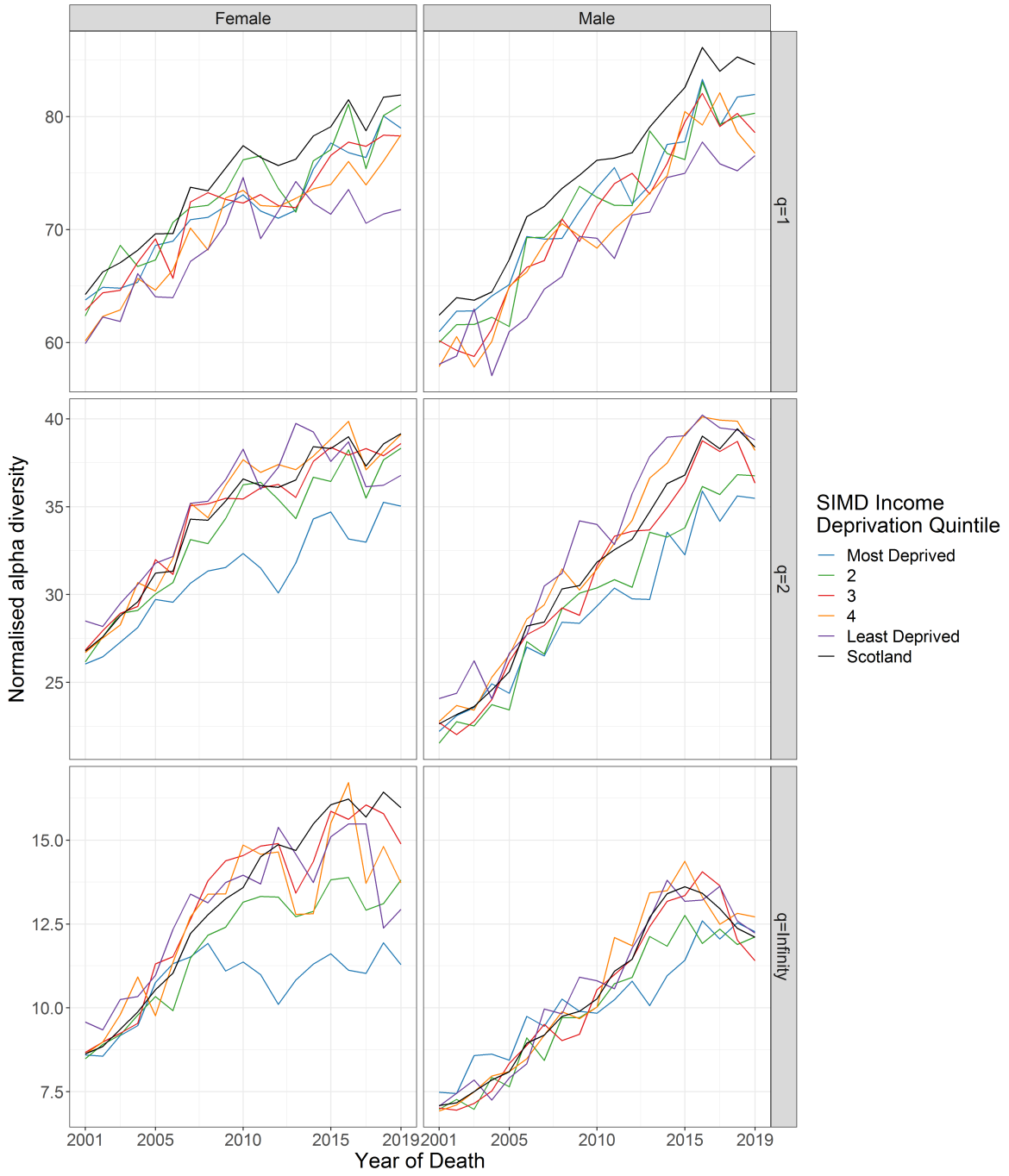


Figure 4.4: Diversity in mortality causes in deaths across all ages in SIMD income deprivation quintiles, plotted against year of death for the years 2001 to 2019. Plots are separated by sex with a separate line for each deprivation quintile.

In most years, at $q = 1$ diversity in mortality causes is lower in less deprived areas and highest in more deprived areas. Over the study period, the linear models plotted in Figure 4.4 show that diversification occurred at similar rates across income deprivation quintiles. The least deprived quintile is shown to have had the least diverse set of mortality causes in almost every year in both sexes. Notably, among females in the years from 2015 onwards, diversity changed little in the least

deprived quintile while continued increases occurred elsewhere. The difference in diversity between the least deprived fifth of the population and other subpopulations has therefore increased slightly among females from 2001 to 2019 but has not changed markedly among males.

This pattern is reversed at q values of 2 and ∞ with a tendency for higher diversity to be found in less deprived areas. Increases to diversity in mortality causes have occurred most slowly in the most deprived areas. This slower diversification has led to increased disparities between the least deprived quintile and the rest of the population, especially in females. These differences in diversity between quintiles at different diversity measures (different values of q) suggest that a greater number of rare causes of mortality occur in more deprived areas. Yet, in these deprived areas the most common causes are nonetheless more dominant (i.e. proportionally more prevalent) than in less deprived areas.

The least consistent trends in diversity are found under normalised alpha diversity at $q = \infty$. Among females across Scotland, a plateau in diversity is noted following 2015. Trends, however, diverge across deprivation quintiles. The plateau in diversity is observed to begin around 2009 in the most deprived and around 2012 in quintile 2, while a great deal of variation is observed in the least deprived areas. In males, the trends noted across the country generally hold, with a peak in 2015 followed by reductions in diversity at $q = 2$ and $q = \infty$.

When an alternative measure of diversity which explicitly accounts for zero values in the distribution of mortality causes is used in the measurement of diversity in mortality in SIMD income deprivation quintiles, Appendix B.1 shows trends differ slightly from those discussed above. Under this measure, uniformity, in most quintiles there is no clear trend diversity in mortality causes in females and a slight increase in diversity in males from 2001 to the mid 2010s. From 2015 onwards a marked increase in uniformity is observed in both sexes. As across Scotland, this means that diversity in mortality causes under this measure was higher in 2019 than in 2001 but that the general trend for increasing diversity observed in most cases under normalised alpha diversity is not found when zero values are explicitly accounted for.

Urban-rural classes

Figure 4.5 illustrates the distribution of mortality causes in Scottish Government urban-rural classes in the years 2001 and 2019. As across Scotland and in SIMD income deprivation quintiles, here there is a reduction in the share of deaths attributed to the most common causes between 2001 and 2019.

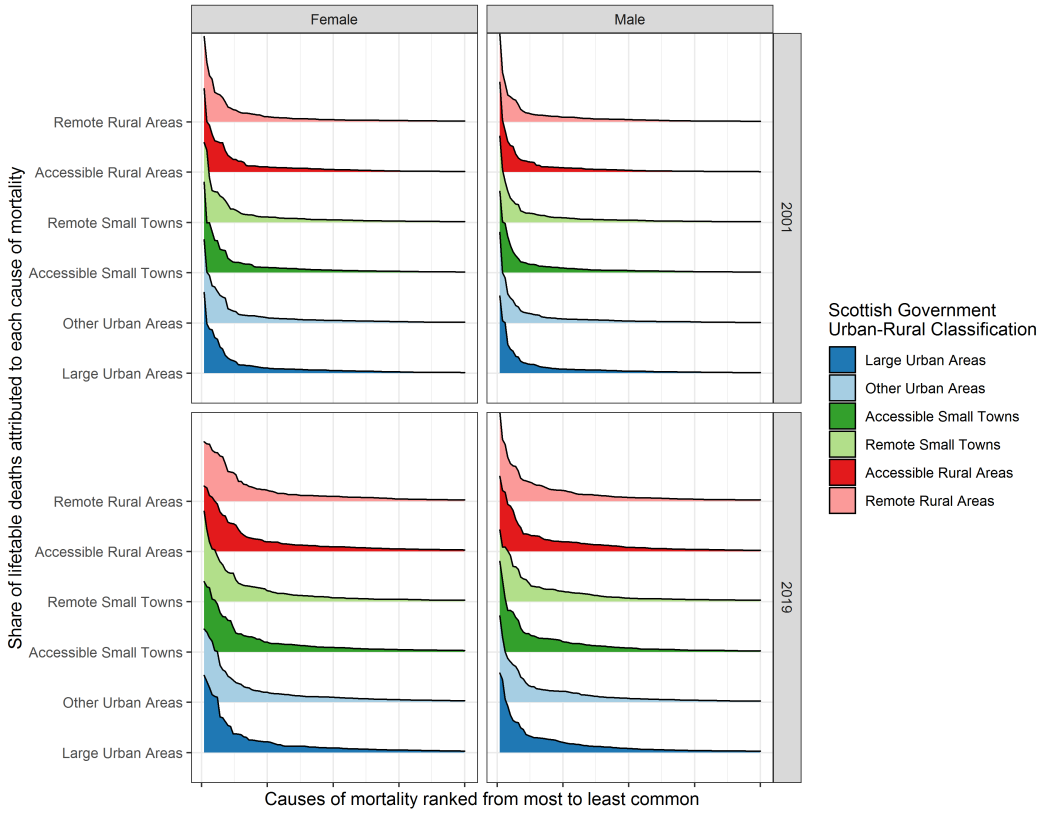


Figure 4.5: Illustrative ridge plot showing the share of deaths, calculated from multiple-decrement life tables, attributed to each of the 100 most common mortality causes in males and females in 2001 and 2019 in each Scottish Government urban-rural classes.

Regarding urban-rural classes, under diversity at q =1, Figure 4.6 shows that individuals in urban areas faced the most diverse causes of mortality and displayed trends most similar to those across Scotland as a whole. Urban areas are home to most of the population, which may explain the similarities to diversity across Scotland. Rural areas and small towns faced less diverse causes than cities. In both sexes, diversity increased more quickly in rural areas than small towns, and in urban areas, diversity increased more slowly than across the country.

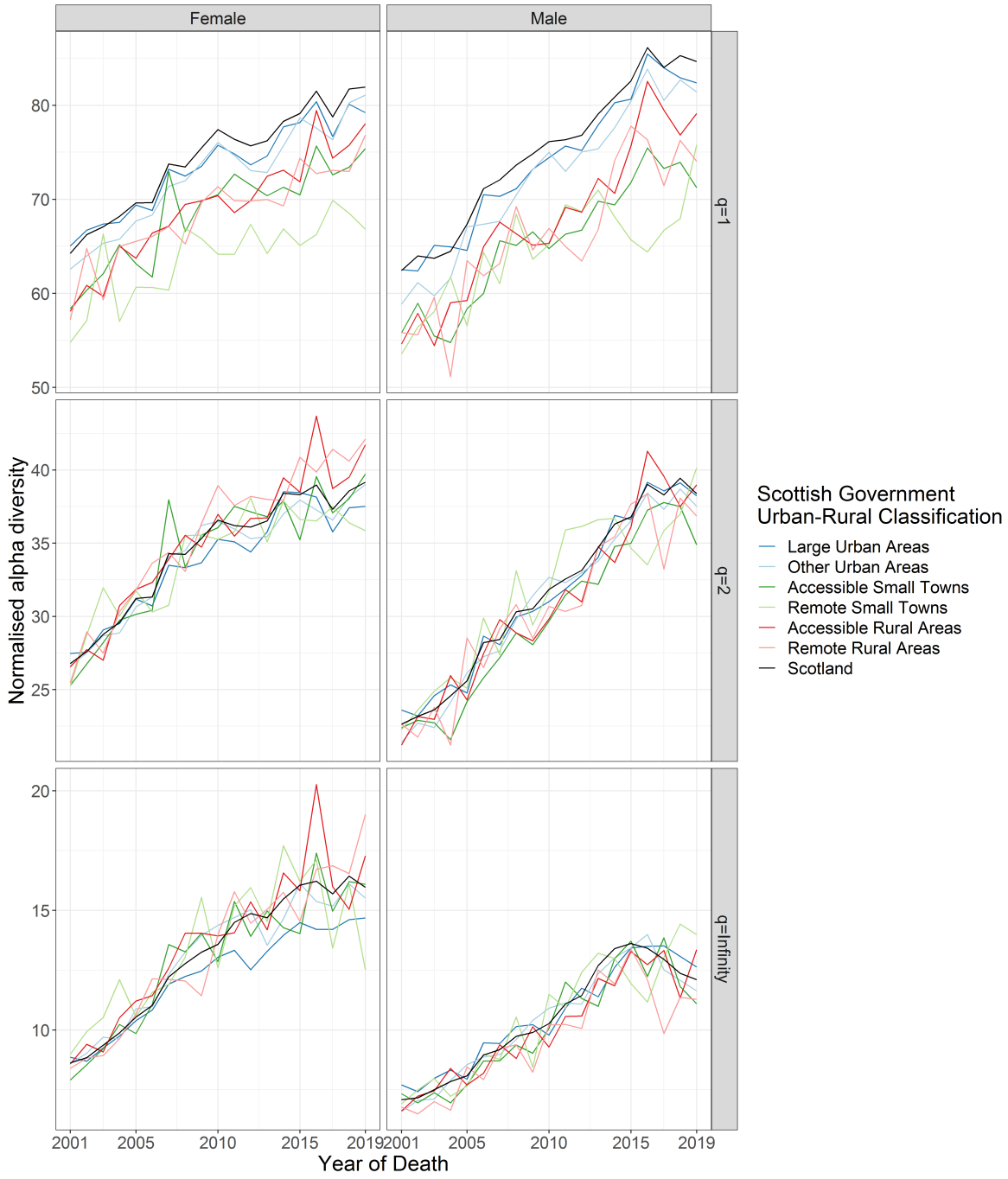


Figure 4.6: Diversity in mortality causes in deaths across all ages in Scottish Government urban-rural classes, plotted against year of death for the years 2001 to 2019. Plots are separated by sex with a separate line for each urban-rural class.

At $q = 2$, among males, urban-rural subpopulations differ little from the country as a whole. Greater variation is observed in diversity at $q = 2$ between urban rural classes than between deprivation quintile. Urban-rural populations vary more in both population and the number of deaths recorded and therefore greater variation

might be expected in diversity here. However, the observed pattern suggests that deprivation is more important as a differentiator of diversity in mortality causes. In females, however, increasing divergence is noted with diversity in mortality causes at $q = 2$ increasing more quickly in rural areas than in urban areas or small towns.

In diversity at $q = \infty$, among males, for the most part, each of the urban-rural subpopulations follows trends across the country, with the exception of remote urban areas, a possible outlier with slightly less diverse causes of mortality. However, in females, a greater degree of variation around the national trends is noted, and specifically, increases to diversity have occurred more slowly in urban areas than in rural areas or small towns.

Combined, these trends under different diversity measures show complexity in how the distributions of mortality causes differ between areas. Higher diversity at q=1 indicates a greater number of rare causes of mortality are recorded in urban areas. However, when the prevalence of causes is given more weight and focus is placed on more common causes, diversity in urban areas is found to be more similar to rural areas and small towns, especially in males. This suggests that at the level of the most common causes, distributions are more uniform across subpopulations than for $q = 1$. Females in urban areas are noted to have less diverse mortality causes than rural areas and small towns under more conservative measures of diversity, suggesting that in urban areas, deaths are more dominated by the most common causes.

Appendix B.1 shows that when an alternative measure related to evenness, which explicitly accounts for zero values in the distribution of mortality causes, is used in the measurement of diversity, trends in some Scottish government urban-rural classes differ to those discussed here. In rural areas and small towns no clear trend in diversity is observed across the study period. Meanwhile, in urban areas a similar trend is observed to that across Scotland with no clear trend in females and a slight increase in males from 2001 to the mid 2010s followed by an increase in the final years of the study period in both sexes. Therefore using this alternative measure which explicitly accounts for zero values in the distribution of causes of mortality, the trends reported here may not be apparent.

4.4.2 The causes of mortality driving trends in diversity in causes across all ages

4.4.2.1 Scotland

Underlying the trends in diversity noted in the previous section are the dynamics of more than 500 individual ICD-10 3-character causes, and the increases, decreases, or stagnation in their associated prevalence[2]. To assess which causes are driving changes in diversity, I assess both the additive value to diversity of each ICD-10 3-character cause and ICD-10 Chapter (the additive value of ICD-10 Chapters was determined by deleting all three-character codes within a Chapter simultaneously) alongside their prevalence. Causes and Chapters are excluded one by one from the calculation of diversity and the effect this has on computed diversity values is measured.

The dynamics within four ICD-10 Chapters are of particular note and are explored here. Chapter II: Neoplasms, where a highly positive additive value remained across the study period; Chapter V: Mental and Behavioural Disorders where the additive value has transitioned from positive to negative; Chapter IX: Diseases of the Circulatory System where the opposite has occurred; and Chapter XX: External causes of morbidity and mortality, where the largest gradient between q values is observed.

Figure 4.7 shows that diseases of the circulatory system (Chapter IX) had a negative additive value in 2001 across diversity measures in both sexes. This indicates that calculated values of diversity would have been higher in this year without deaths within this Chapter. The negative additive value of ICD-10 Chapter IX is due to the dominance of acute myocardial infarction (ICD-10 3-character code I21) as a cause of mortality. This was the most prevalent cause in both sexes in 2001 accounting for 14% and 11% of all-cause mortality rates in males and females respectively. In this year it was dominant over other causes in the distribution, with the second most common cause accounting for only 9% in males and 7% in females. Acute myocardial infarction had the largest individual negative additive

[2] As discussed in Section 4.3.7 prevalence values are computed as the proportion of the all-cause ASMR within a given population (i.e., males or females in Scotland as a whole or a subpopulation, in a single year) attributed to a given cause.

value across measures of diversity in 2001. Figure 4.8 shows that for most of the study period, the prevalence of acute myocardial infarction fell; however, following 2015, this reversed, causing the additive value of this cause at $q = 2$ and $q = \infty$ to decrease further. The ASMR of cause I21 in males across Scotland fell from 242 per 100,000 population in 2001 to 97 per 100,000 in 2014 before increasing to 105 per 100,000 in 2019. Associated with this resurgence in the mortality rate of acute myocardial infarction its additive value became more strongly negative following 2015 in males driving the reduction in diversity at $q = \infty$ observed in males in Figure 4.2. This contributed to a reduction in the net contribution of Chapter IX to diversity, observed as a lower positive additive in Figure 4.7.

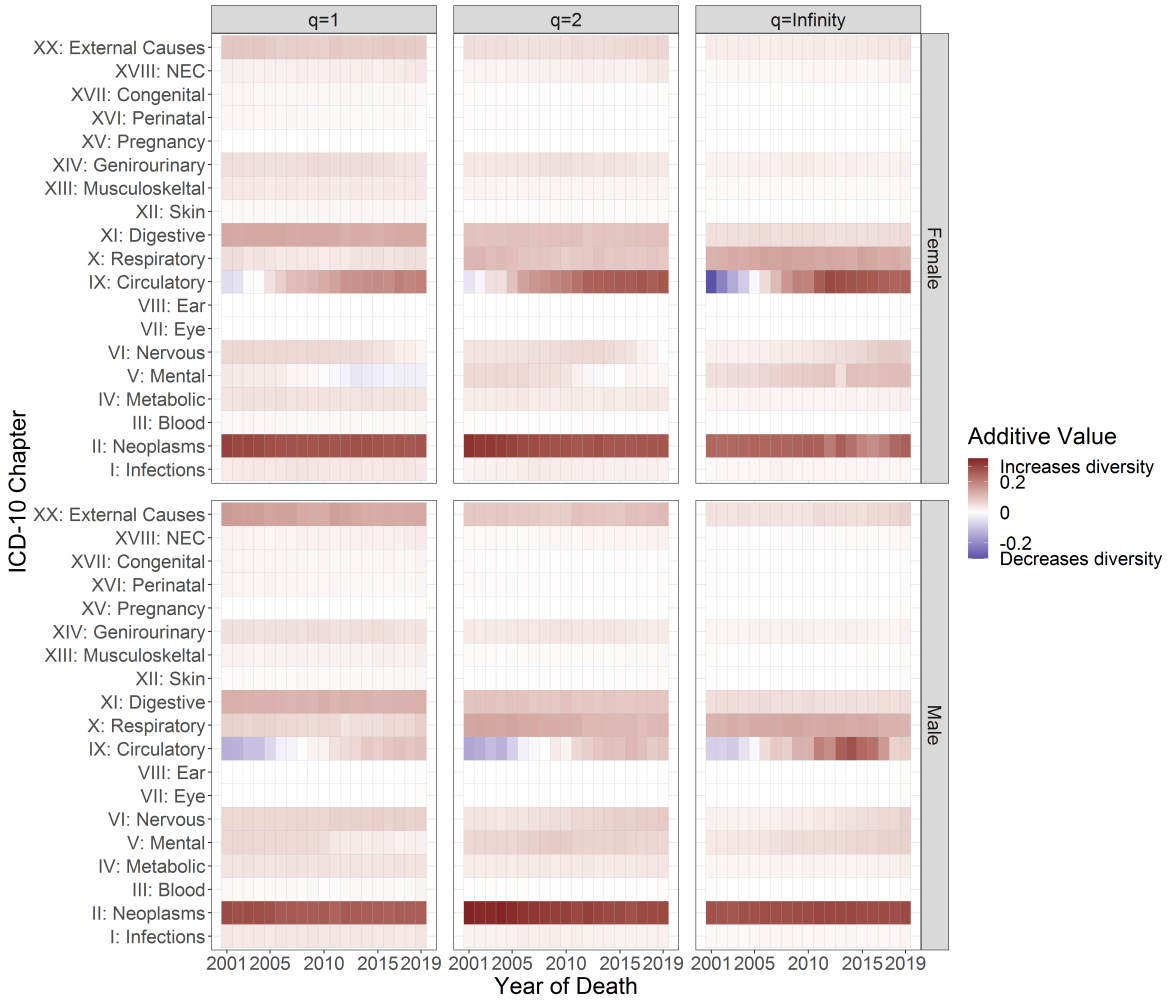


Figure 4.7: Additive value to diversity in mortality causes of ICD-10 Chapters in deaths across all ages among males and females in Scotland[3].

[3]Chapter XVIII: NEC indicates causes "Not elsewhere classified".

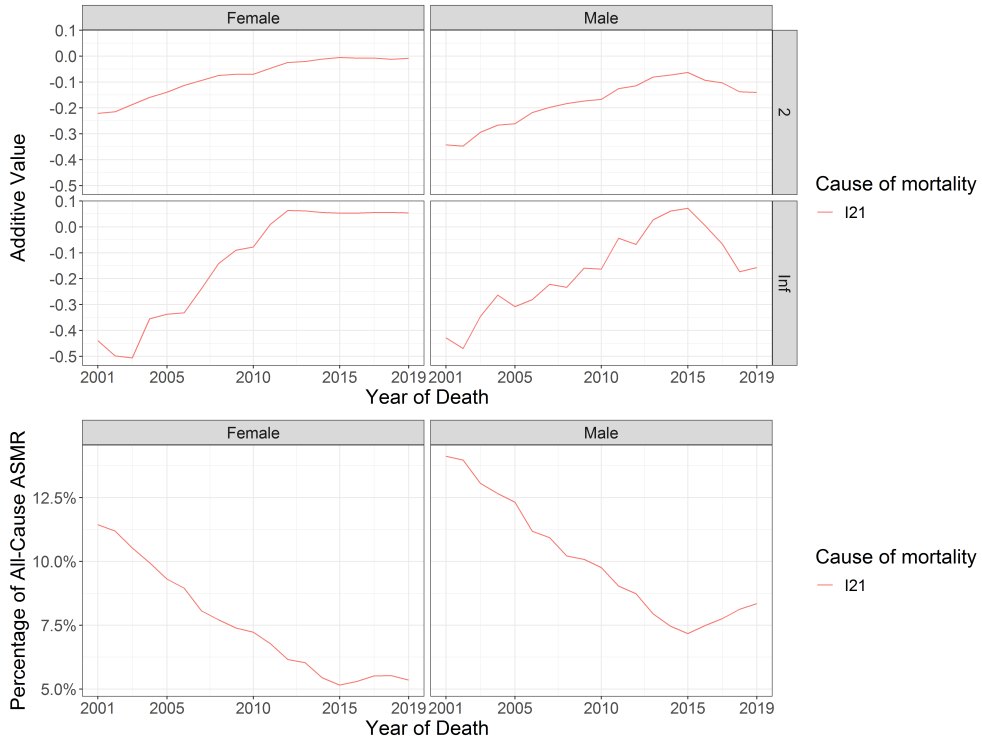


Figure 4.8: The additive value at $q = 2$ and $q = \infty$ and proportion of all-cause ASMR assigned to ICD-10 cause I21 - acute myocardial infarction in males and females across all ages in Scotland.

Chapter II, which represents cancers, has the largest positive additive value in most years. This means that excluding Chapter II from the measurement of diversity in mortality causes leads to a greater reduction in diversity than the removal any other Chapter. In other words, this chapter contributes most to diversity in mortality causes in both males and females. Contributing to diversity at $q = 1$ indicates that the ICD-10 codes within this Chapter make up a significant proportion of the relatively rare causes of mortality across the total distribution. Chapter II accounted for the largest number of three character ICD-10 codes recorded in females and the second largest in males in both 2001 and 2019 (Appendix A.2). The breadth of causes within Chapter II compared to other ICD-10 Chapters is a function of the way diseases and conditions are grouped within the ICD system. Most other ICD-10 Chapters describe diseases and conditions within a body system whereas Chapter II covers all cancers and all presentations of cancers. This emphasises a key advantage of examining diversity in mortality causes at a fine-grained cause level because with causes grouped together the strong contribution of variation in Chapter II to overall variation in mortality causes could be lost. In total, the causes within Chapter II form collectively one of the most prevalent groups in each year and therefore cancers represent a significant portion of the total mortality distri-

bution. Due to the distribution of 3-digit causes within Chapter II, excluding this group of causes results in a significant reduction in diversity. At more conservative measures of diversity, the prominent additive value of Chapter II indicates that a number of causes of mortality within the Chapter are among those which are the most proportionally prevalent in the total distribution of mortality causes. Appendix A.2 shows the number of causes in each chapter which accounted for more than 0.25% of total mortality in 2001 and 2019 (totalling around 10% of the total number of individual mortality causes). Chapter II contains more of these causes than any other Chapter, in both males and females, and in both years (2001 and 2019). Excluding causes in Chapter II therefore makes the distribution of mortality causes less even and lowers diversity.

Across Scotland, only Chapter V – Mental and Behavioural Disorders in females in diversity at q=1 moves from a net contribution to diversity to a reduction. This is primarily driven by Vascular Dementia (F01) and Unspecified Dementia (F03); both of these causes transitioned from having a positive to a negative additive value over the study period. This was caused by increases in the prevalence of both causes. This increase was especially pronounced in Vascular Dementia, increasing from 0.8% of all-cause mortality in 2001 to 4.2% in 2019. Part of this increase is likely associated with changes in coding practices in Scotland discussed in Chapter 3 which involved a larger number of deaths assigned to ICD-10 codes F01 and F03.

Table 4.1: The five causes in ICD-10 Chapter XX which increased in additive value to diversity at $q = \infty$ most from 2001 to 2019 in males and females across Scotland and the difference in additive value between years.

ICD-10 3-Character Code	ICD-10 Description	Additive Value at $q = \infty$		
		2001	2019	Absolute Difference
	Female			
W19	Unspecified fall	1.20E-02	2.20E-02	0.0096
X42	Accidental poisoning by and exposure to narcotics and psychodysleptics [hallucinogens], not elsewhere classified	3.00E-05	7.20E-03	0.0072
X41	Accidental poisoning by and exposure to antiepileptic, sedative-hypnotic, antiparkinsonism and psychotropic drugs, not elsewhere classified	0.00E+00	1.60E-03	0.0016
X70	Intentional self-harm by hanging, strangulation and suffocation	1.20E-03	2.50E-03	0.0013
X45	Accidental poisoning by and exposure to alcohol	2.20E-16	5.00E-04	0.0005
	Male			
X42	Accidental poisoning by and exposure to narcotics and psychodysleptics [hallucinogens], not elsewhere classified	4.60E-04	1.90E-02	0.0181
W19	Unspecified fall	6.70E-03	1.60E-02	0.0089
X70	Intentional self-harm by hanging, strangulation and suffocation	7.20E-03	1.00E-02	0.0028
X41	Accidental poisoning by and exposure to antiepileptic, sedative-hypnotic, antiparkinsonism and psychotropic drugs, not elsewhere classified	9.00E-05	2.60E-03	0.0025
X45	Accidental poisoning by and exposure to alcohol	0.00E+00	1.00E-03	0.001

Chapter XX – External causes of morbidity and mortality had a more positive additive value to diversity at q = 1 than at more conservative measures of diversity. This indicates that Chapter XX contains a number of relatively rare causes, increasing diversity at q =1 but few causes with high proportional prevalence. Appendix A.2 shows that among males, the largest number of individual causes of mortality were within Chapter XX in both 2001 and 2019, despite the Chapter accounting for only 4.0% and 7.2% of the all-cause mortality rate in these years. For reference, the Chapter with the second most recorded causes (Chapter II: Neoplasms) accounted for 27.5% and 29.3% of all-cause ASMR in these years. The increasing prevalence of both Chapters together likely contributes to the rapid increase in diversity across all causes in males at q values of 1 and 2. Fewer causes within Chapter XX occur among females, and it accounts for a smaller proportion of all-cause ASMR which may be a factor in the slower diversification in causes observed among females.

The additive value of Chapter XX to diversity at $q = \infty$ increased towards 2019, especially in males. This indicates that the contribution of this Chapter to diversity at the level of the most common causes has increased. The five causes in Chapter XX which increased their additive value to diversity at $q = \infty$ most are shown in Table 4.1. For the most part, these are dominated by causes generally classified as deaths of despair, deaths attributed to suicide, or associated with alcohol or drugs (Allik et al., 2020a). During the study period, deaths of despair have been reported in the literature to have increased (Brown et al., 2019; Allik et al., 2020); however, some changes (discussed in Chapter 3) to coding practices may contribute to specific increases in this Chapter.

4.4.2.2 Scottish Subpopulations

Deprivation quintiles

There are some differences in the contribution of ICD-10 Chapters to diversity in mortality causes between deprivation quintiles. Figure 4.9 shows the additive value of each Chapter, in each year, in the most and least deprived fifth of the Scottish population. Overall, similar trends can be noted to those across the country in Figure 4.7; however, there are some key differences.

Unlike the population as a whole, or those in more deprived areas, among males in the least deprived quintile, Chapter IX: diseases of the circulatory system persisted as having a negative additive value beyond 2005 under more conservative measures. Under $q = \infty$ Chapter IX had a negative additive value in most years in the least deprived areas. This indicates that the causes in this Chapter remained dominant for longer in the least deprived areas. Trends in the contribution of Chapter IX to diversity in mortality causes are largely driven by acute myocardial infarction (I21) which was the most common cause of mortality, in Chapter IX and across all causes, in both the most and least deprived quintile in 2001 and 2019. Table 4.2 shows it accounted for roughly equal proportions of the total mortality rates in both quintiles. Differences arise when a comparison is made to the second most common cause, which was Malignant Neoplasm of the Bronchus and Lung (ICD-10 C34) – lung cancer – in both the most and least deprived quintile. In the most deprived quintile, the two causes account for a more even proportion of deaths, while among those in the least deprived quintile, cause I21 accounts for more of the total mortality rate. acute myocardial infarction is therefore more dominant as the leading cause in the least deprived areas, giving it a greater negative additive value.

Table 4.2: The percentage of all-cause age standardised mortality rate (ASMR) assigned to ICD-10 causes I21 – acute myocardial infarction and C34 – malignant neoplasm of the bronchus and lung in males in the most and least deprived population quintiles in 2001 and 2019.

	All-cause ASMR	**I21**		**C34**	
		Cause-specific ASMR	**Percentage of all-cause ASMR**	**Cause-specific ASMR**	**Percentage of all-cause ASMR**
			2001		
Most Deprived	2203.2	304.3	13.8%	207.2	9.4%
Least Deprived	1258.7	175.2	13.9%	68.0	5.4%
			2019		
Most Deprived	1766.7	146.5	8.3%	142.5	8.1%
Least Deprived	943.3	79.2	8.4%	51.7	5.5%

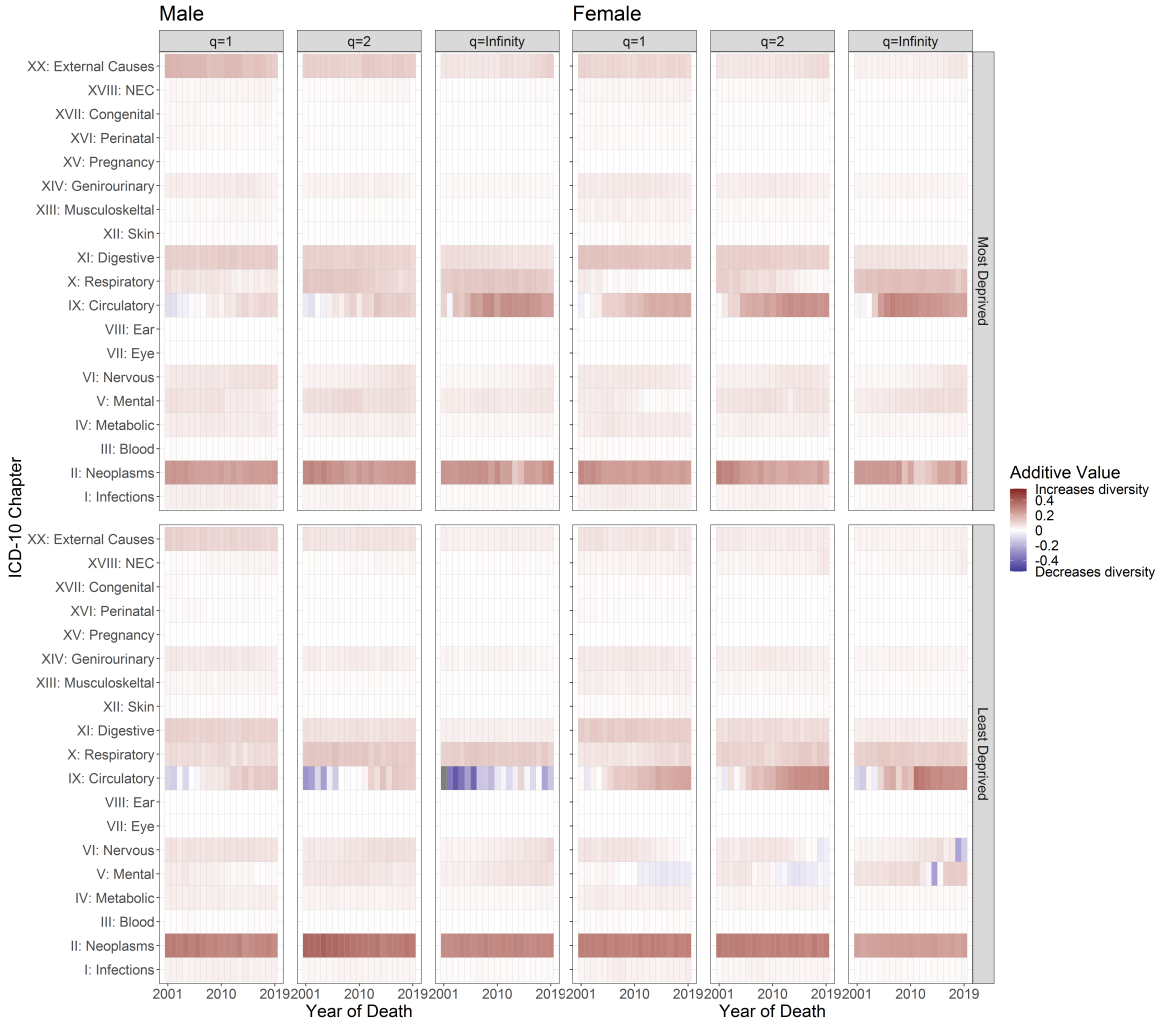


Figure 4.9: Additive value to diversity in mortality cause of ICD-10 Chapters in deaths across all ages among males and females in the most and least deprived deprivation quintiles.

Across the total female population of Scotland the additive value of Chapter V: Mental and Behavioural Disorders transitions from 2001 to 2019 from positive to negative as causes within the Chapter became more prominent. This trend can be seen strongly in the least deprived areas but it is not observed among those in the most deprived areas. This reflects the disparate burden of mortality associated with dementia in these areas. The prevalence of causes F01 and F03 in the least deprived areas is almost double that of the most deprived, as shown in Table 4.3. However, the prevalence of the two causes increased at a similar rate in both quintiles over the study period. If this continues, it is likely that the trends observed in the least deprived quintile here will also occur in the most deprived areas as degenerative diseases become relatively more dominant.

Similarly to the trends reported for Chapter V, a transition from positive to negative additive values is observed in Chapter VI: Diseases of the Nervous System among females in the least deprived areas. This is driven primarily by increases in the prevalence of Alzheimer's Disease (ICD-10 code G30) which rose from around 1% of all-cause ASMR to 7% between 2001 and 2019. Again, increases also occurred in more deprived areas, but at slower rates.

Table 4.3: The percentage of all-cause age standardised mortality rate (ASMR) assigned to ICD-10 causes F01 – Vascular Dementia and F03 – Unspecified Dementia in females in the most and least deprived population quintiles in 2001 and 2019.

	All-cause ASMR	F01		F03	
		Cause-specific ASMR	Percentage of all-cause ASMR	Cause-specific ASMR	Percentage of all-cause ASMR
			2001		
Most Deprived	1384.3	8.1	0.6%	33.4	2.4%
Least Deprived	927.1	10.4	1.1%	41.7	4.5%
			2019		
Most Deprived	1288.9	41.3	3.2%	39.6	3.1%
Least Deprived	718.1	39.0	5.4%	33.0	4.6%

Finally, Chapter XX: External Causes of Morbidity and Mortality, has a more strongly positive additive value in the most deprived areas . Table 4.4 shows that in the most deprived areas, a greater number of individual causes of mortality was recorded in this Chapter and it accounted for a greater percentage of all-cause ASMR in both 2001 and 2019. Among those in the most deprived areas, deaths of despair and violent deaths were prominent among the causes with the largest individual additive value, while among the least deprived subpopulation accidental deaths featured more (Appendix A.3).

Table 4.4: The number of ICD-10 three-character codes under which deaths were recorded in Chapter XX External Causes of Morbidity and Mortality and the percentage of all-cause ASMR associated with Chapter XX in the most and least deprived population quintiles in males and females in 2001 and 2019.

	Number of Individual ICD-10 causes in Chapter XX	All-cause ASMR	Chapter XX ASMR	Percentage of all-cause ASMR
		2001		
		Female		
Most deprived	54	1384.3	44.2	3.2%
Least deprived	30	927.1	28.2	3.0%
		Male		
Most deprived	79	2203.2	101.1	4.6%
Least deprived	54	1258.7	48.0	3.8%
		2019		
		Female		
Most deprived	39	1288.9	77.8	6.0%
Least deprived	30	718.1	22.4	3.1%
		Male		
Most deprived	70	1766.7	178.8	10.1%
Least deprived	53	943.3	43.4	4.6%

Urban-rural classes

There are also differences in the contribution of ICD-10 causes and Chapters to diversity in causes of mortality between the populations of the six urban-rural classes. Figure 4.10 shows the additive value of selected ICD-10 Chapters to diversity in mortality causes in each of the urban-rural classes.

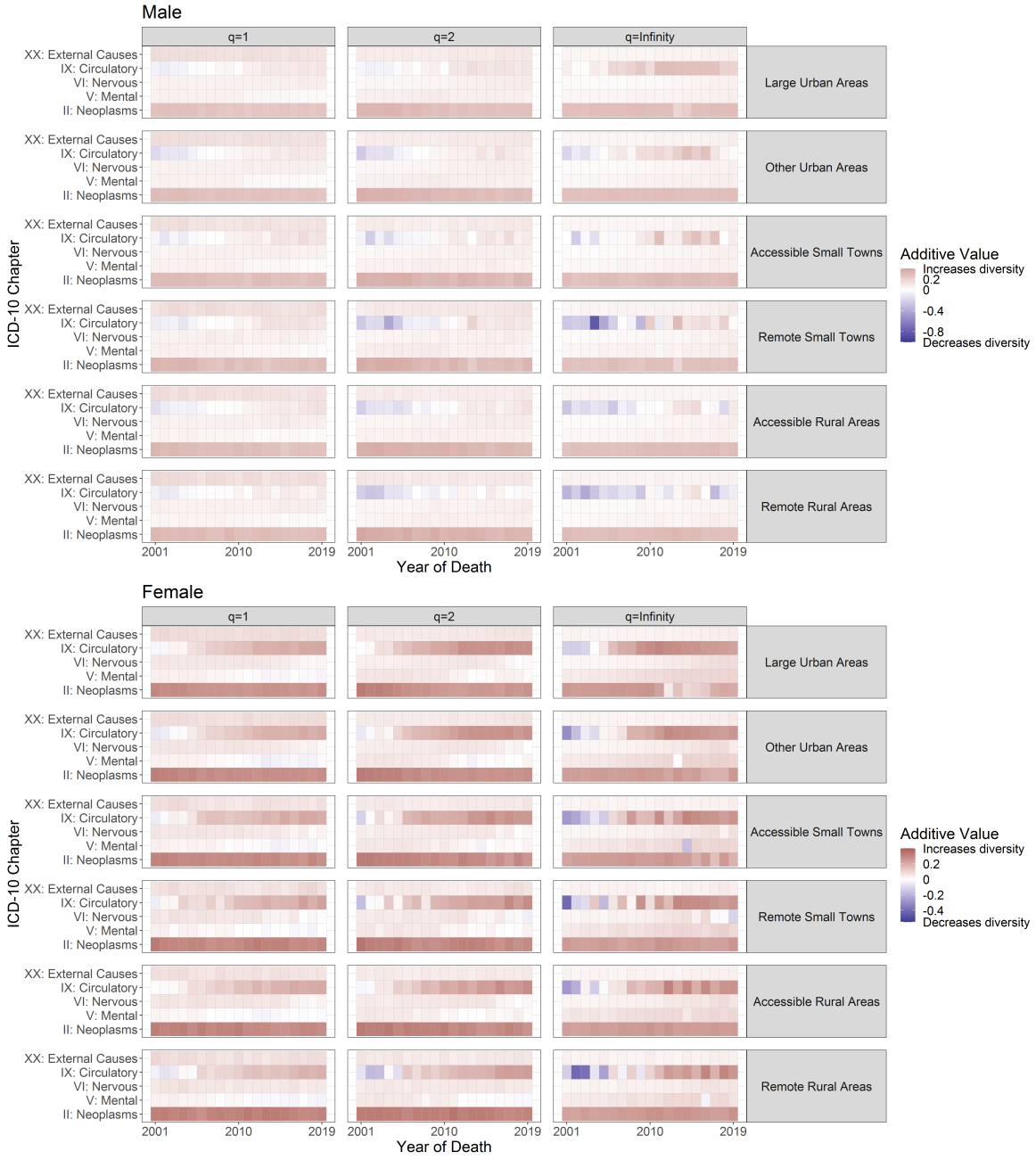


Figure 4.10: Additive value to diversity in mortality cause of selected ICD-10 Chapters in deaths across all ages among males and females in urban-rural classes.

In small towns and rural areas, and especially the remote classes of these areas, Chapter IX is observed to have a more negative additive value to diversity later into the study period than in urban areas or across the country. In these more rural areas, mortality rates and therefore the prevalence of certain causes within this Chapter, most notably acute myocardial infarction , fell more slowly than in more urban areas. This pattern is similar to the observations among the least

deprived population quintile, as previously noted. Income deprivation is patterned intricately across the urban-rural divide though more deprived areas tend to be concentrated in urban locations and rural areas are often less deprived. Therefore, the tendencies across urban-rural classes and deprivation quintiles may be linked.

Findings presented above suggest that the contribution of Chapter XX (External causes of morbidity and mortality) to all-cause diversity is higher in more deprived areas. However, across the urban-rural classes the additive value of Chapter XX is less clearly patterned. Previously, similarities were noted in trends in diversity between more deprived areas and urban areas, and in other sections of this analysis, this similarity holds. However, it is not the case for the additive value of Chapter XX. The differences noted are likely a consequence of the complex interplay of social, economic and geographic factors which are risk factors for causes in this Chapter. Chapter XX covers both deaths of despair, such as suicide and drug-related mortality, as well as mortality associated with accidents such as falls or road traffic accidents. A number of these causes have been shown to be associated with more urban areas and deprived communities. In the research presented here, due to the limits of population sizes, these subpopulations are not further split to reflect both deprivation related factors and urban-rural geography.

4.4.2.3 Section summary

I have demonstrated that, across Scotland, diversification in causes of mortality from 2001 to 2019 was driven by changes in the prevalence of a number of causes of death. Reductions in the proportion of deaths attributed to cardiovascular diseases and specifically acute myocardial infarction is responsible for much of the increase in diversity especially at more conservative measures. However, increasing prevalence of a wide variety of cancers was also responsible for increased diversity. As was the increasing rate of deaths due to external causes, especially in males. The reductions in diversity observed at $q = \infty$ in males in later years of the study period are shown to have been related to increases in mortality attributed to acute myocardial infarction in males following 2015.

Across Scottish subpopulations, variation in the burden of diseases and their relative contribution to diversity is observed. In urban and more deprived areas increased prevalence of external causes of mortality are shown to have a greater contribution to diversity than in other areas. Meanwhile, increases in the prevalence of degenerative diseases were observed to have a greater influence on diversity in less deprived areas than other subpopulations. Diseases of the cardiovascular system were observed to have remained a dominant cause in less deprived areas further into the study period. These trends indicate that despite increases in diversity occurring at broadly similar rates in most subpopulations, and under most measures of diversity, the causes driving diversification differ between groups.

4.4.3 Trends in normalised alpha diversity at q = 1, 2, and ∞ in deaths within twenty-year age groups

The leading causes of mortality and the distribution of mortality causes varies greatly by age (Naghavi et al., 2015). In work carried out during the completion of this thesis, findings published in McMonagle et al. (2022) show that increases in the diversity of mortality causes at q=1 have occurred at similar rates in Scotland from 2001 to 2019 for both deaths among those younger than 75 (premature mortality) and deaths among those aged 75 and older. Here, the population is further divided into twenty-year age groups and the diversity of mortality causes within each of these subpopulations is examined.

4.4.3.1 Scotland

Figure 4.11 shows trends in normalised alpha diversity in mortality causes at q values of 1, 2, and ∞ over time in Scotland among males and females aged 0 to 19, 20 to 39, 40 to 59, 60 to 79, and 80+ separately.

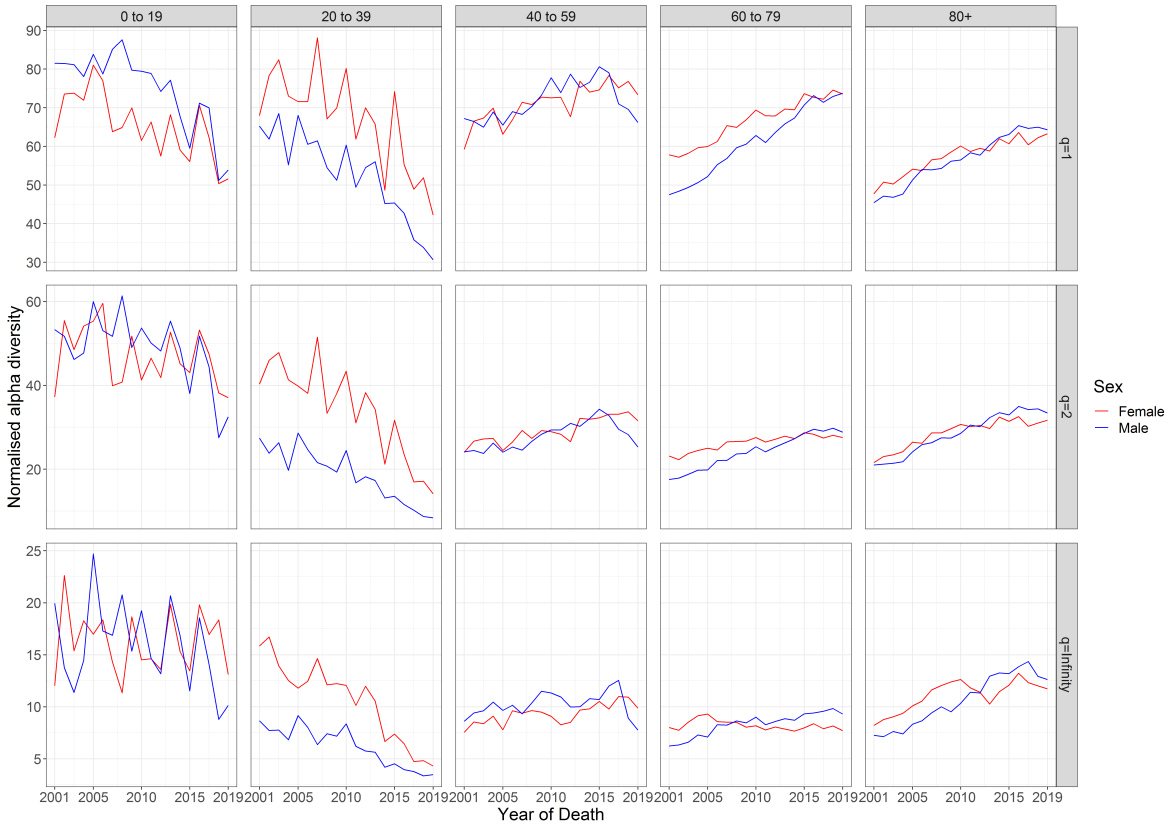


Figure 4.11: Trends in diversity in mortality causes in deaths across twenty-year age groups across Scotland from 2001 to 2019.

For most age groups and *q* values, males faced less diverse mortality causes than females in 2001, with trends generally leading to increased similarity in males compared to females over the study period. The last three columns of Figure 4.11 show that in deaths among those aged 40 and older, similar trends to those observed across all ages are noted with increasing diversity over the study period (although with some exceptions discussed below). However, in deaths among those aged 0 to 19 and 20 to 39 (first two columns of Figure 4.11), diversity in mortality causes fell over the study period.

In deaths among those aged 20 to 39, females faced a more diverse set of mortality causes at each value of *q*, although diversity fell more quickly in females than males at q=2 and q= ∞. The distribution of causes of mortality has become more dominated by the most common causes of deaths in this age group. The proportion of deaths attributed to rare causes has fallen while that attributed to the most common has increased.

Among those aged 0 to 19 however, while diversity in mortality causes has fallen at q values of 1 and 2, the trend at $q = \infty$ is less clear. While the proportion of deaths attributed to rare causes is falling, this has not been met by an increase in the proportion of deaths assigned to the most common causes. This suggests a more complex redistribution at these ages.

Both among those aged 0 to 19 and 20 to 39, it is possible that the observed reduction in diversity is linked to the number of recorded deaths in these populations. Among these age groups, the number of deaths is relatively small and is within the range at which the number of deaths can be considered a factor in diversity calculations (See Section 3.3.3). Figure 4.12 shows the relationship between the number of deaths in each year and diversity in mortality causes. If the sample size was a major influence, we would expect a strong linear correlation. In deaths among those aged 0 to 19 under diversity at $q = 1$ and $q = 2$ this relationship appears to exist. In other panels, diversity does not seem to be linked to the number of deaths and this suggests trends in diversity have occurred despite changes to the number of deaths. Appendix A.4 shows trends in diversity in mortality causes with deaths grouped across years to increase population size and supports the trends noted in Figure 4.12. Trends in younger age groups should be treated carefully because of the small number of deaths which occur. However, the evidence presented here suggests that changes in diversity have not been closely linked to changes in the number of deaths meaning that conclusions drawn in this section can be considered robust.

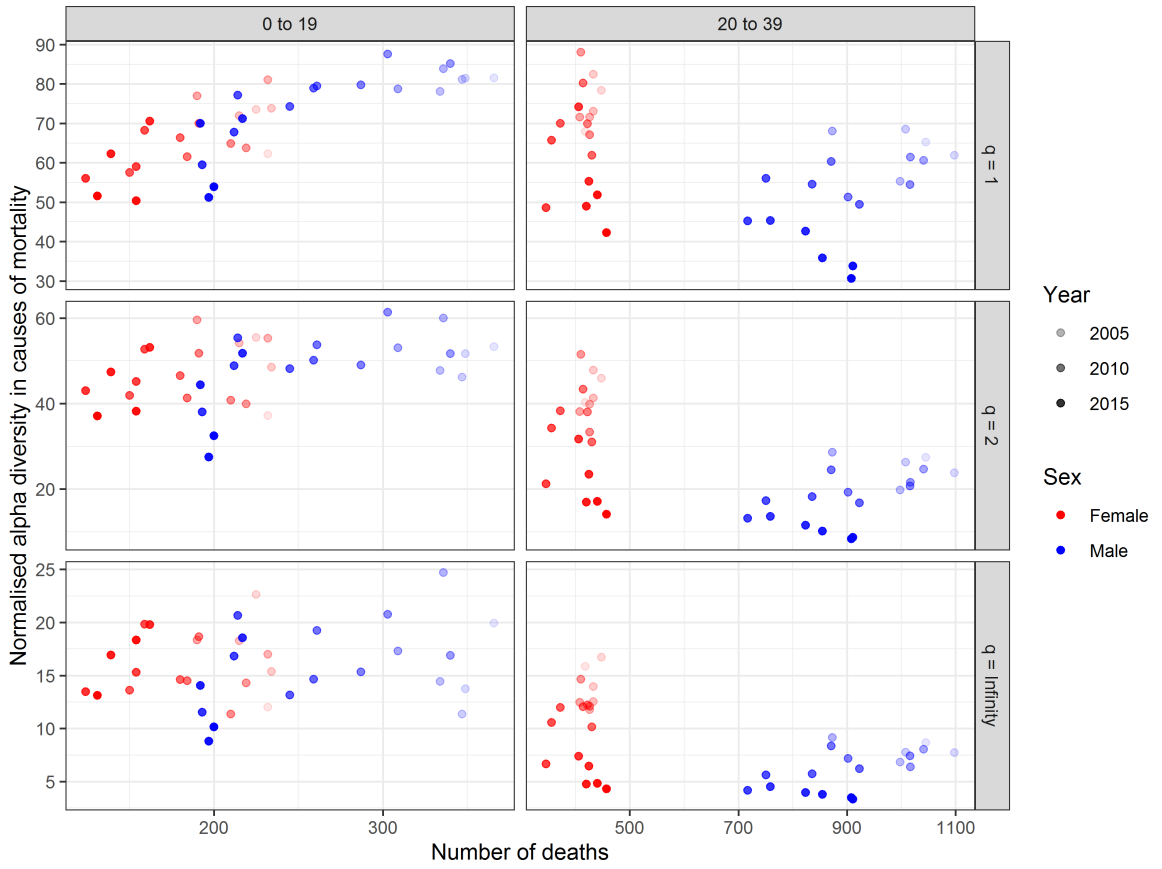


Figure 4.12: Diversity in causes of mortality among those aged 0 to 19 and 20 to 39 across Scotland plotted against the number of deaths in each year from 2001 to 2019.

Diversity in mortality causes increased at all values of q in those aged 80 and older, with diversity increasing more quickly among males than females. This was also true for deaths among those aged 60 to 79 at q values of 1 and 2; however, at $q = \infty$ diversity in mortality causes among females fell over the study period. This indicates that mortality causes have become increasingly dominated by the most common causes in deaths in this age group, while at the same time, the range of relatively rare causes to which deaths are assigned has increased. The reductions in diversity after 2015, noted at $q = \infty$ across the total population, are only observed in deaths among those aged 40 to 59 and 80+.

In Appendix B.1 an alternative measure of diversity which is more closely related to evenness and explicitly accounts for zero values in the distribution of mortality causes is used to assess trends in the diversity of mortality causes in deaths within twenty year age ranges in Scotland. In deaths at ages 0 to 19, 60 to 79 and 80+ trends in mortality cause diversity under this alternative method are found to

be similar to those reported above under normalised alpha diversity above. This suggests the trends in these age ranges discussed in this section are robust to this alternative method for the calculation of diversity. In deaths among those aged 20 to 39, no clear trend is noted when diversity is measured using a measure which explicitly accounts for zero values in contrast to the negative trend noted above. Similarly using a measure which accounts for zero values, no clear trend in mortality cause diversity is found in deaths among those aged 40 to 59 despite the positive trend found when measured using normalised alpha diversity in this chapter. The results of this analysis should be considered when examining the trends in diversity among those aged 20 to 39 and 40 to 59 reported above.

4.4.3.2 Scottish Subpopulations

The overall tends previously noted by age group hold within both deprivation quintiles and the populations of each urban-rural class. Specifically, diversity in mortality causes fell in those younger than age 40 and generally increased in older age groups. For older age groups, less variation in diversity is noted for both levels of deprivation and urban-rural class. Further explanation of these trends and tendencies is presented in Appendix C.

4.4.4 The causes of mortality driving these trends in diversity by age group

In this section additive value analysis, as described in Section 4.4.2, is carried out on the diversity in mortality causes within twenty-year age groups. This section mostly focuses on trends across Scotland as a whole, with some references made to differences in additive value and the prevalence of causes in subpopulations. Further analysis of additive value to diversity in mortality causes within specific age groups among deprivation quintiles and urban-rural classes is described in Appendix C.

In deaths among those aged 0 to 19, Figure 4.13 shows the key contributors to diversity in mortality causes were Chapters: XVI: Certain conditions originating in the perinatal period; XVII: Congenital malformations, deformations, and chromosomal abnormalities; XVIII: Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified; and XX: External causes of morbidity and mortality. Over the study period the contribution of Chapters XVI, XVII, and XVIII remained high, reflecting that a large proportion of mortality in this age group is concentrated at birth and among those aged under 1 year old. The strong negative additive value in certain years in Chapter XVIII are associated with R95: Sudden infant death syndrome. This cause of mortality has varied between contributing 5% and 10% of ASMR in this age group, with no clear trend. In both males and females, the additive value of Chapter XX to each diversity measure trended negatively over the study period and this is likely associated with relatively older individuals within this age group. This increase appears to be closely associated with suicide, particularly cause X70: "Intentional self-harm by hanging, strangulation and suffocation" which increased in proportional prevalence and therefore additive value across diversity measures from 2001 to 2019.

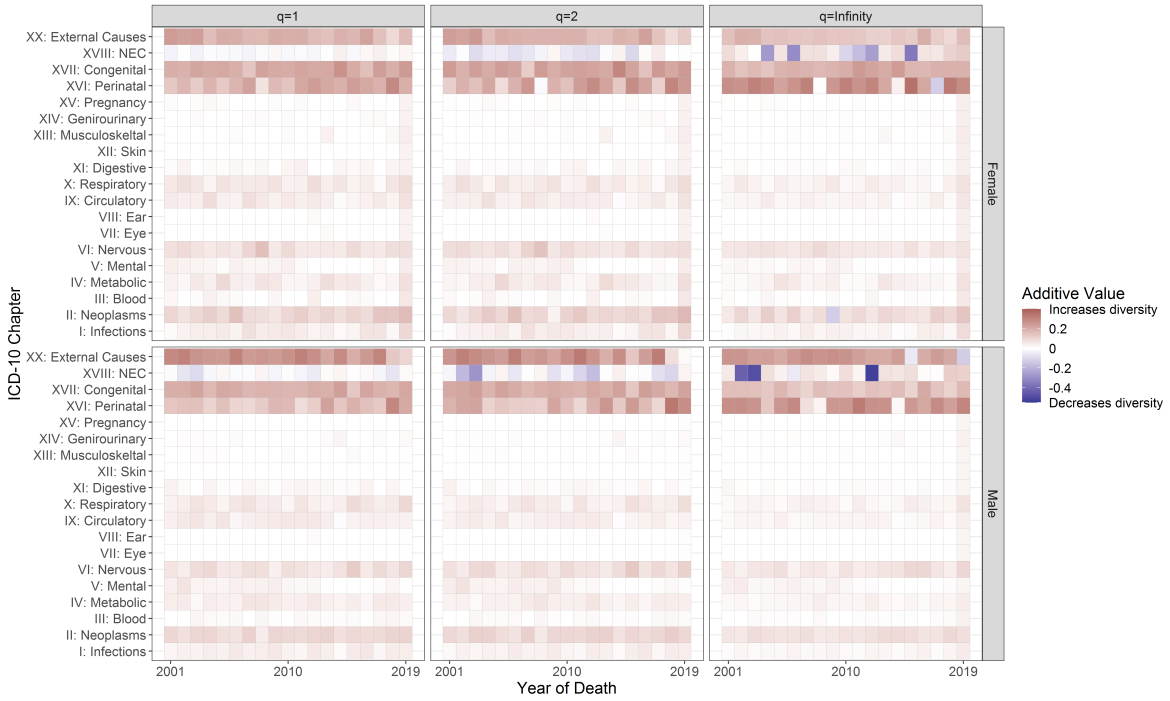


Figure 4.13: Additive value to diversity in mortality cause of ICD-10 Chapters in deaths among males and females aged 0 to 19 in Scotland.

The additive value of Chapter XX to diversity has also increased markedly over the study period among those aged 20 to 39 as shown in Figure 4.14. In 2019, causes within this Chapter were responsible for 46% of all-cause mortality rates in females aged 20 to 39 and 67% of the all-caues mortality rate in males aged 20 to 39. These deaths was further concentrated among a small number of common causes of mortality within Chapter XX. The five most common causes of mortality in Chapter XX in both sexes were almost exclusively those defined as deaths of despair or violent deaths and cumulatively accounted for more than 38% of all-cause mortality rates in females aged 20 to 39 and 53% in males as shown in Table 4.5. In deaths among those aged 20 to 39, Chapter XX is found to be more prominent in more deprived areas. Previous research has clearly shown strong deprivation-related effects on mortality in Chapter XX with specifically deaths of despair linked to more deprived areas and this is clearly evident in this analysis.

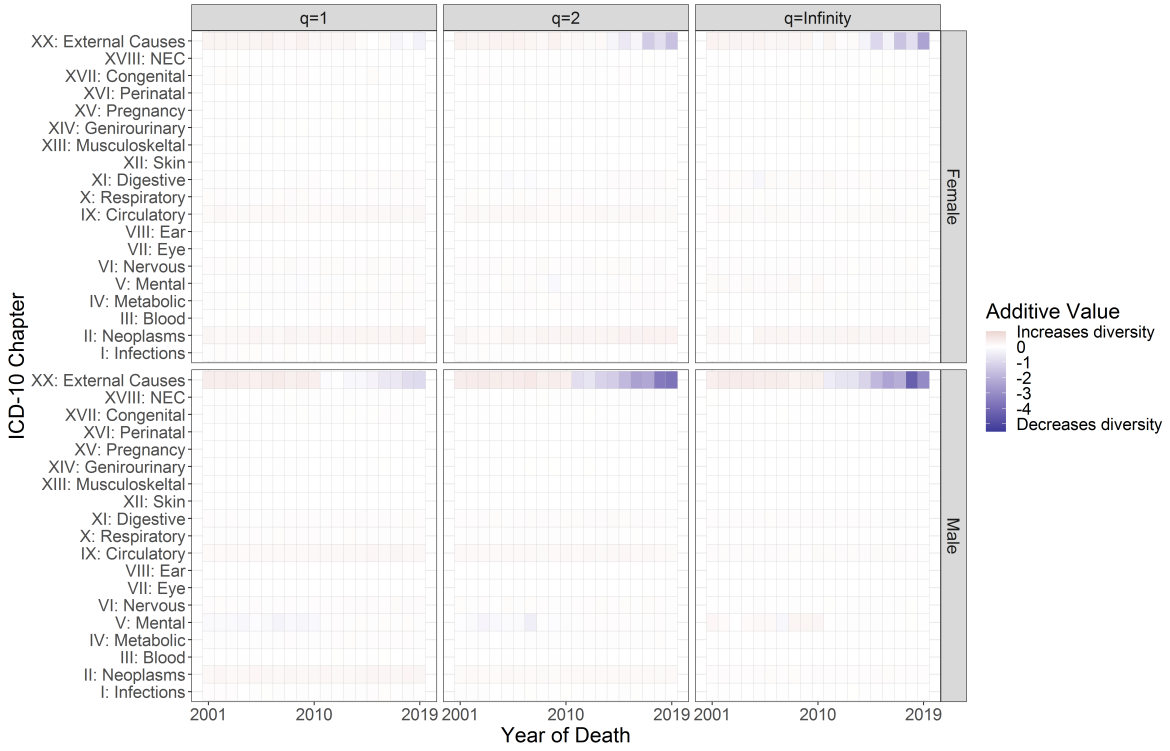


Figure 4.14: Additive value to diversity in mortality causes of ICD-10 Chapters in deaths among males and females aged 20 to 39 in Scotland.

Figure 4.15 shows that several of the trends in additive value reported for deaths across all ages are evident among those aged 40 to 59. The trends of increasingly negative additive value in Chapter XX noted at younger ages can be noted in males. In females however, Chapter II: Neoplasms had a negative additive value in certain years at $q = 2$ and $q = \infty$ indicating that causes within this Chapter were dominant over the distribution of mortality causes. Breast cancer (ICD-10 code

C50) was primarily responsible for this dominance causing as much as 13% of mortality in females at this age in certain years. Trends observed in cancers (Chapter II) at ages 40 to 59 in females are more prominent in the least deprived areas where all-cause diversity stagnated across the study period. The specific cancers which were most prominent in each quintile differs with breast and ovarian cancers (ICD-10 codes C50 and C57) most common in the least deprived areas and lung cancer (C34) more dominant in more deprived areas.

Table 4.5: The five causes in ICD-10 Chapter XX with the highest percentage of all-cause ASMR in males and females aged 20 to 39 across Scotland in 2019.

ICD-10 Code	ICD-10 Description	All-cause ASMR	Cause-specific ASMR	Percentage of all-cause ASMR (%)
	Females			
X42	Accidental poisoning by and exposure to narcotics and psychodysleptics [hallucinogens], not elsewhere classified	64.5	12.9	19.9%
X70	Intentional self-harm by hanging, strangulation and suffocation	64.5	4.4	6.9%
X41	Accidental poisoning by and exposure to antiepileptic, sedative-hypnotic, antiparkinsonism and psychotropic drugs, not elsewhere classified	64.5	1.4	2.2%
X80	Intentional self-harm by jumping from a high place	64.5	1.2	1.9%
V43	Car occupant injured in collision with car, pick-up truck or van	64.5	0.9	1.3%
	Males			
X42	Accidental poisoning by and exposure to narcotics and psychodysleptics [hallucinogens], not elsewhere classified	129.5	32.4	25.1%
X70	Intentional self-harm by hanging, strangulation and suffocation	129.5	19.2	14.8%
X41	Accidental poisoning by and exposure to antiepileptic, sedative-hypnotic, antiparkinsonism and psychotropic drugs, not elsewhere classified	129.5	4.1	3.1%
X99	Assault by sharp object	129.5	1.4	1.1%
X80	Intentional self-harm by jumping from a high place	129.5	1.2	0.9%

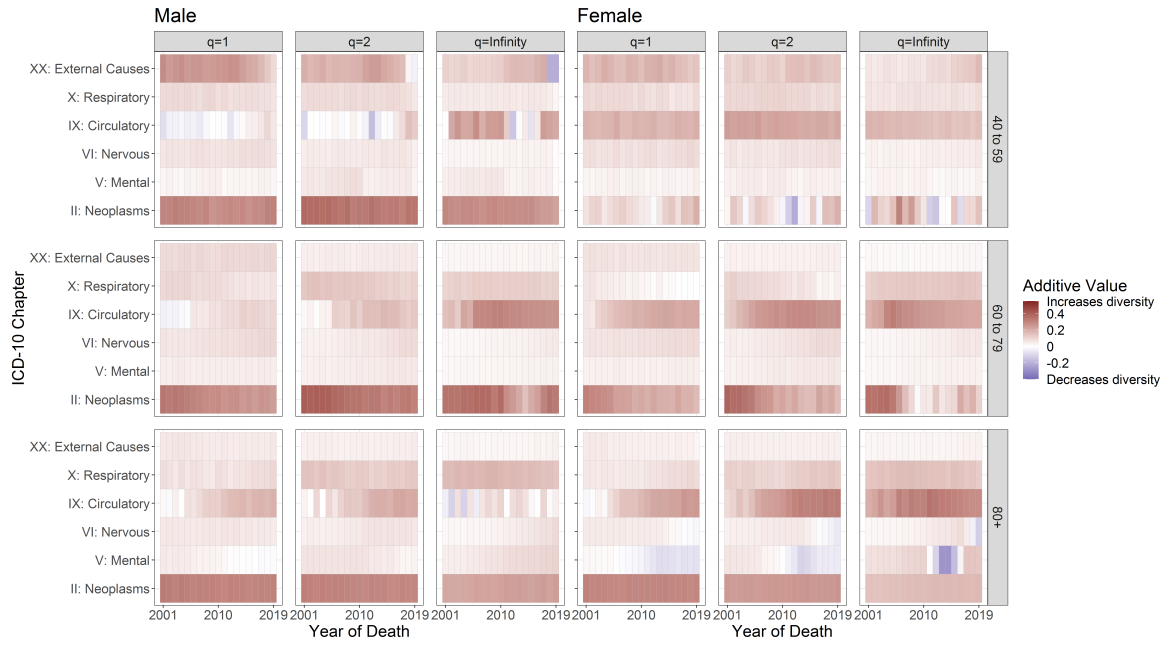


Figure 4.15: Additive value to diversity in mortality cause of certain ICD-10 Chapters in deaths among males and females aged 40 to 59, 60 to70 and 80+ in Scotland.

In those aged 60 to 79, trends in additive value are mostly similar to the trends across all ages previously discussed. However, similar to findings in those aged 40 to 59, the contribution of cancers to diversity became less positive later in the study period. In both sexes this is associated with lung cancer (ICD-10 code C34) which contributed ~10% of all-cause ASMR across the study period with increases noted in females towards 2019. The reduction in additive value is especially prominent at $q = \infty$ in females. This increasing prevalence may be closely associated with the observed reductions in diversity at $q = \infty$ among females aged 60 to 79 in Figure 4.11.

Finally, the most prominent trends among aged 80+ are observed in Chapters V and VI in females. Within these Chapters are the codes for the degenerative diseases noted previously as causing a significant proportion of mortality in females: dementias (ICD-10 codes F01 and F03) and Alzheimer's disease (G30). These causes collectively were attributed to 8.4% of all-cause ASMR in this age group in 2001 rising to 22.0% in 2019 making them the three most common causes of mortality in females 80+. The proportion of mortality associated with these causes is shown in Figure 4.16. The peak of mortality associated with Unspecified Dementia (F03) between 2010 and 2015 corresponds to reductions in diversity at $q = \infty$ in those aged 80+ as this was the most dominant cause in these years causing it to have a

large effect on diversity at $q = \infty$. During this period the Scottish Government introduced the Certification of Death Act (Scottish Government, 2011), which changed the process for the certification of death in Scotland. This law enacted increased scrutiny on death certificates and aimed to increase the accuracy of recording. The reduction in prevalence of cause F03 "Unspecified dementia" while increases continued in the more defined cause F01 "Vascular dementia" may be associated with this change.

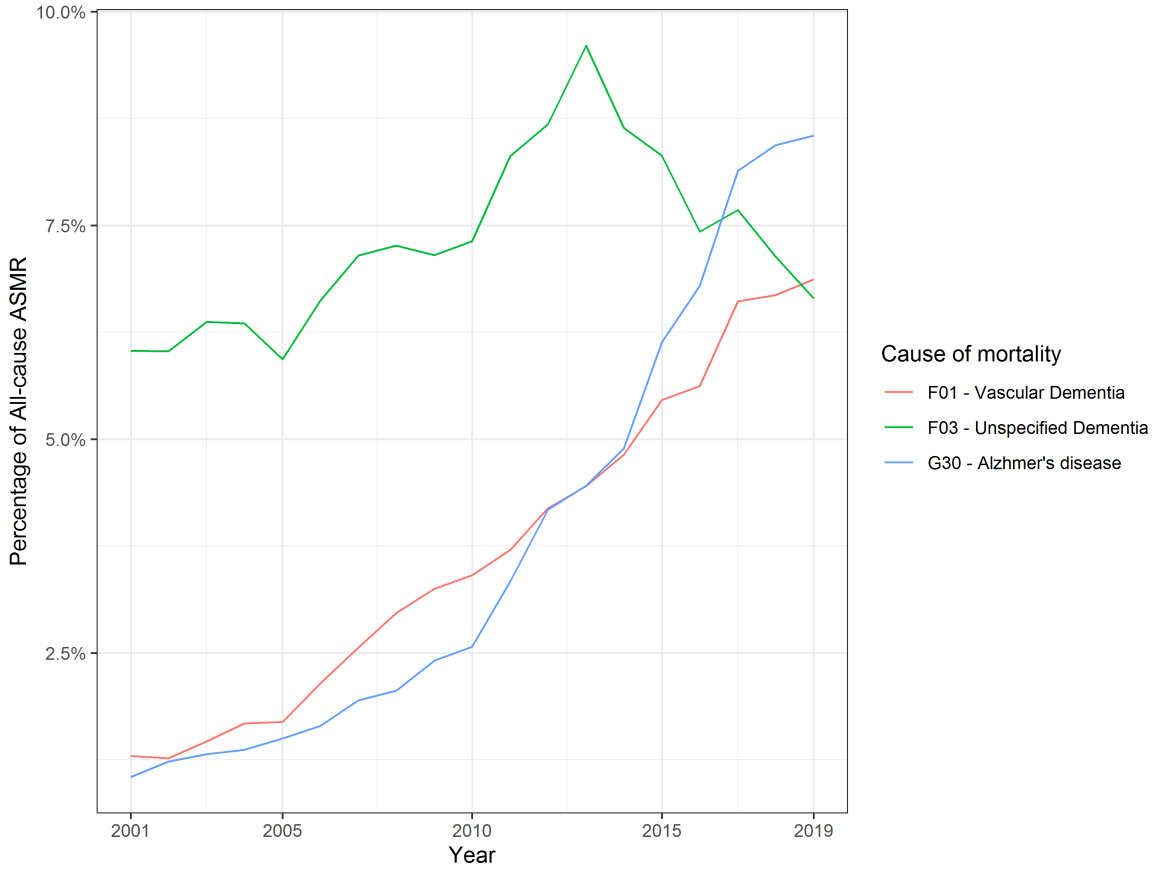


Figure 4.16: The percentage of all-cause ASMR among females aged 80+ attributed to ICD-10 causes F01 – Vascular Dementia, F03 Unspecified dementia and G30 Alzheimer's diseases across Scotland from 2001 to 2019.

4.5 Discussion

This chapter aimed to examine trends in the diversity of mortality causes across Scotland, and within subpopulations distinguished by level of income deprivation, urban-rural class and twenty-year age group.

4.5.1 Principal findings

4.5.1.1 What were the national temporal trends in diversity in mortality causes in deaths across all ages in Scotland from 2001 to 2019?

For deaths across the Scottish population divided only by sex, diversity in mortality causes increased in both males and females from 2001 to 2019. However, a change in trends is evident in the 2010s, especially in males under diversity at $q = \infty$ where following 2015 reductions in diversity are observed. Higher diversity is observed in 2019 than in 2001 under each measure of diversity examined in this chapter, from q=1 where relatively rare causes are heavily weighted, to the most conservative measure $q = \infty$. Diversity in causes of mortality increased more quickly among males than females under diversity at $q = 1$ and $q = 2$. This means that variation in the relatively rare causes faced by the male population increased more quickly than among those faced by females. When alternative methods for the measurement of diversity, which explicitly account for zero values in the distribution of mortality causes, are used in the measurement of mortality cause diversity trends differ slightly. No clear trend is observed in female deaths and a slight increase in male deaths are observed from 2001 to 2014 before an increase in the final years of the study period. This means that diversity in causes of mortality under this alternative measure was higher in 2019 than 2001 despite the difference in trends.

4.5.1.2 What were the national temporal trends in diversity in mortality causes at different ages at death in Scotland from 2001 to 2019?

The observed diversification in mortality causes across the population is shown to be restricted to deaths in those aged 40 and older. Indeed, diversity in mortality causes actually fell from 2001 to 2019 in the age groups 0 to 19 and 20 to 39. The only exception to the trend for increasing diversity in mortality causes at older ages is among females aged 60 to 79, where reductions in diversity at $q = \infty$ are observed. This indicates that, in this age group, the dominant causes of mortality increased

their prevalence over the study period becoming relatively more common. Under alternative methods for calculation diversity, which account for zero values in the distribution of mortality causes, no clear trend is found in diversity in mortality causes among those aged 20 to 39 and 40 to 59.

4.5.1.3 Did diversity in mortality causes differ between subpopulations, grouped by area-level deprivation and urban-rural class, in Scotland?

Under diversity at q = 1, diversity tended to be highest in more deprived areas and in urban areas, indicating that in these areas the population was likely to face greater variation in mortality causes. In urban areas and deprived areas, external causes of mortality, such as deaths of despair and violent deaths, were prominent and contributed more to diversity, especially at q = 1 than in less deprived areas and rural areas. At more conservative values of q in both males and females across deprivation quintiles and in females in urban-rural classes a reversal is observed and diversity was shown to be higher in less deprived areas and more rural areas. Together tendencies between subpopulations across q values indicate that those in more deprived areas and in urban areas faced a more diverse range of relatively rare causes of mortality, while being the most likely groups to die due to the most common causes.

4.5.1.4 What were the subpopulation level trends in diversity in mortality causes in Scotland from 2001 to 2019?

In males, without notable exception, national trends in the diversity of mortality causes from 2001 to 2019 are observed to hold within each Scottish subpopulation. Namely, diversity increased from 2001 to around 2015 following which reductions in diversity are observed. In males trends in the diversity of mortality causes were similar across subpopulations from 2001 to 2019, meaning that across measures of diversity disparities between subpopulations changed little. Under alternative diversity measures which account causes of mortality which do not occur in every

year, diversity is observed to increase slightly from 2001 to the mid 2010s before increasing in the final years of the study period across SIMD income deprivation quintiles and in urban areas. In rural areas no clear trend is observed in mortality cause diversity from 2001 to 2019 under this alternative measure.

Among females, national trends in the diversity of mortality causes are also observed to hold across subpopulations under diversity at $q = 1$ and 2. However, in the most deprived areas diversity at $q = \infty$ is observed to plateau after 2009. This indicates slower reductions in the prevalence of the most common causes of mortality in more deprived areas than their less deprived counterparts. This caused increasing deprivation related disparities in mortality cause diversity under $q = \infty$ in females in the 2010s. Under the alternative measure of diversity assessed in Appendix B.1 no clear in diversity trend is observed from 2001 to the mid 2010s in each SIMD quintile and in urban areas followed by an increase in diversity towards in the final years of the 2010s. In rural areas no clear trend in diversity is observed when a measure sensitive to zero values in the distribution of mortality causes is used.

4.5.1.5 Which causes of mortality drove trends in the diversity of mortality causes at the national and subpopulation level in Scotland from 2001 to 2019?

Increases in diversity across Scotland were driven both by improvements in mortality associated with the most dominant causes of mortality and corresponding increases in the prevalence of a range of other causes. The rate of deaths due to the most dominant causes of mortality (most notably, acute myocardial infarction and lung cancer) reduced over the study period meaning fewer individuals faced the most common cause in 2019 compared to 2001. These reductions in rate were more rapid than reductions in all-cause mortality rates meaning that the most common causes were also responsible for a smaller percentage of deaths in 2019. This is demonstrated by the increase in diversity at $q = \infty$. Cancers; degenerative diseases such as dementia and Alzheimer's disease; and external causes of mortality were among the causes which increased in prevalence. These effects together create a more even distribution of mortality causes which is measured as an increase in diversity in the measures used in this chapter.

The falling diversity in mortality causes among men after 2015 is shown to be closely linked to the prevalence of acute myocardial infarction among males in Scotland. Both the mortality rate and prevalence of acute myocardial infarction fell up to 2015 before increasing again in the period up to 2019. Despite earlier improvements in both rate and prevalence, acute myocardial infarction remained the most dominant cause of mortality among males in each year from 2001 to 2019, meaning diversity at $q = \infty$ was very sensitive to its fluctuations.

Disparities in the contribution of various ICD-10 Chapters and mortality causes are found between deprivation quintiles and urban-rural classes. Despite lower absolute mortality rates associated with cardiovascular diseases in the least deprived areas, a slightly larger proportion of deaths were attributed to these causes of mortality. Further, acute myocardial infarction was more dominant in less deprived areas because other common causes were less prevalent than in more deprived areas.

Increased diversity in mortality causes at $q = 1$ in more deprived areas may be associated with a greater contribution of external causes of death in these areas. Differences in the additive value of ICD-10 Chapters are noted to be smaller between urban and rural areas than between deprivation quintiles. Nonetheless, tendencies for cardiovascular disease deaths to have a greater positive additive value at $q = \infty$ and therefore be more dominant are noted in remote rural and small-town areas.

4.5.2 Interpretation

4.5.2.1 National trends and disparities in the diversity of mortality causes

Increasing diversity in mortality causes in Scotland was driven by reductions in the proportion of deaths attributed to the most common causes of mortality. These causes receive a great deal of public health resources and research funding (Luengo-Fernández et al., 2006; Mays & Smith, 2011; Luengo-Fernandez et al., 2012). This focus is likely to be a key cause of the reduction in mortality rates across the study period which lead to the falling prevalence of causes of mortality

such as acute myocardial infarction and lung cancer. Indeed, some key risk factors for both cardiovascular disease and cancers such as tobacco smoking and alcohol consumption reduced in Scotland in the late 20th and early 21st century (Meier, 2010; Brown et al., 2019; Petrou & Kupek, 2019; Batty et al., 2020; Robinson et al., 2021). However, it should be noted that other risk factors such as obesity have become more common (Keaver et al., 2020). Treatment regimens have also improved through increased research and funding (Lee et al., 2011; Glover et al., 2014; Bhatnagar et al., 2015). The reduction in mortality rates associated with cardiovascular disease in the 2000s continued a consistent trend which began as early as the 1950s in Scotland (Mitchell, 2005).

In addition to the reduction in rate and prevalence associated with the most common causes, increases in the proportion of deaths due to less common causes of mortality are noted. These causes include cancers and deaths of despair. Overall, shifts in the distribution of mortality causes took the form of reductions in the most common causes and corresponding increases in the prevalence of less common causes. This increased prevalence was driven by increases in mortality rates associated with some less common causes while the mortality rate of many less common causes changed little but reductions in overall mortality rates meant they increased in prevalence.

Together, common causes affecting proportionally fewer individuals and other causes occurring relatively more often leads to a more even distribution of causes. Measured here as an increase in diversity, this more even distribution has the potential to challenge health care and public health systems. Increased diversity in mortality causes may affect economies of scale in the prevention and treatment of common causes of mortality as the patients they apply to make up a smaller proportion of the population. Diversification also increases the burden on public health systems and officials as increased resources must be directed to an increasingly broad range of significant causes of mortality.

At the individual level, higher diversity means greater uncertainty in the causes of mortality that one might face which may have the potential to increase the diagnostic uncertainty in physician led diagnosis in medical practice (Bhise et al., 2018; Bergeron-Boucher et al., 2020). It has however, been suggested that increasing diagnostic accuracy is a driver of increasing diversity in morality causes (Bergeron-Boucher et al., 2020). If this is the case then an increase in diversity could not credibly lead to an increase in diagnostic uncertainty. That is to say

that if diversity is increasing because physicians are capable of making more specific diagnoses the increase in diversity is unlikely to make these diagnoses more difficult. Despite increasingly advanced diagnostic techniques, evidence from this chapter and other studies suggests that a reduction in deaths associated with common causes is the driver of much of the diversification in Scotland and developed countries (Bergeron-Boucher et al., 2020; Trias-Llimós & Permanyer, 2023). It is, therefore, credible that the increasing diversity in mortality causes observed in these countries could lead to greater uncertainty for physicians in diagnosing causes of ill-health.

From 2001 to 2009, diversity at $q = \infty$ increased at near-identical rates in males and females. This means that the prevalence of the most common cause, acute myocardial infarction in both sexes, reduced at a similar rate. Acute myocardial infarction, as well as other forms of heart disease, was, and remains, a focus of long running public health campaigns in Scotland (Scottish Government, 2014). Reductions in the prevalence and mortality rate of acute myocardial infarction suggests that these campaigns were successful and caused similar reductions in the percentage of the population who faced this cause of death. These trends have been observed despite different prevalences (acute myocardial infarction has, within the study period, always been more common in men) and some proposed differences in presentation and outcome in heart disease between sexes (Steingart et al., 1991; Savarese & D'Amario, 2018).

In deaths at younger ages, the findings in this chapter show falling diversity in mortality causes. These trends are driven by the increasing dominance of a number of causes of mortality including deaths of despair. Deaths of despair, grouped in ICD-10 Chapter XX, are a public health concern, especially for young males, with a previous study noting increasing rates in Scotland (Allik et al., 2020). A number of public health interventions aimed at these causes have been attempted in Scotland; however, evidence of their effectiveness is varied (Bird et al., 2016; Boniface et al., 2017). It is not the case, therefore, that falling diversity in mortality causes is a sign of easing strain on the healthcare and public health systems. Which would be a reversal of the challenges brought about by diversification of causes described above. Rather, as mortality reduces from other causes, as it has done in the younger age groups, the pressures of addressing an increasingly dominant group of common causes of mortality brings different challenges.

4.5.2.2 Diversity in mortality causes within Scottish subpopulations

Trends in the diversity of fine-grained causes of mortality at the subpopulation level have not previously been widely examined. The similarity in trends in diversity found in this chapter is, perhaps, surprising given that the analysis of the additive value of causes presented here shows that the burden of disease is markedly different across subpopulations. Despite disparate common causes of mortality, the similar rates of diversification across subpopulations shows that uncertainty in the cause of mortality each individual might face has increased to a similar degree across subpopulations. The reasons for this similar increase in diversity are unclear. It may be that improvements to diagnosis and accuracy in recording causes of death have occurred uniformly across the population. This increased specificity would increase diversity across the population. There may be more profound reasons for the similar increase in the variation in causes faced by subpopulations which may be a valuable direction for future research.

The results of this research reveal that the causes which are driving changes in diversity, and therefore must be addressed to improve health, are different across subpopulations. In more deprived areas, ICD-10 Chapter XX: external causes of morbidity and mortality, is shown to be associated with increases in diversity while in less deprived areas cancers and degenerative diseases have driven diversification. Therefore, while a more comprehensive approach in public health and health care systems may be required across Scotland, it should be tailored to suit the specific causes driving changes in diversity in different areas.

A possible reason for variation in the diversity of mortality causes within a nation is irregular mortality recording practices (O'Malley et al., 2005). In most countries, including Scotland, there are national standards for how mortality should be recorded on death certificates. Nonetheless, the physicians, coroners and mortality coders who record mortality causes for each death are human and a level of differential judgement, perspective and bias can be expected between individuals (Danilova, 2016). In addition to variation at between individual recorders of mortality, regional variation in the recording of mortality causes has been described in the literature (Lanska & Peterson, 1995). There may be some logic behind an expectation of differential mortality coding practices across both the socio-economic and urban-rural gradients. Despite central funding and a directive to provide high

quality care for all, the provision of care by the Scottish National Health Service (the NHS) is known to vary (Brown et al., 2010). Research has shown increased difficulty in accessing healthcare in more deprived areas in Scotland, with longer waiting times and shorter appointments despite greater healthcare needs in these areas (Katikireddi et al., 2018).

Across the urban-rural gradient issues of access are also posited as having an effect on the provision of care. Although this generally is an issue of geographical access rather than the overburdened or under resourced systems blamed for poor access in more deprived areas (Sutton, 2002). Rural areas tend to have fewer general practitioners per-individual and almost by definition are likely to be further from centres of advanced medical care such as hospitals (Farmer et al., 2005; Farmer et al., 2010).

Issues of access, both geographical and socioeconomic, have the potential to impact the accuracy of recording of mortality causes, meaning that deaths are recorded under an incorrect mortality code. It is possible that in less deprived areas and in more urban areas accuracy of recording would be higher due to greater access to healthcare. This may allow for increased certainty in diagnosis and make it more likely unusual and rare conditions are identified. Better access to healthcare may therefore be expected to increase diversity in mortality causes. The results of this chapter do not necessarily uphold this hypothesis as diversity at $q = 1$ (the measure which would be most likely to identify a greater range of rare causes recorded in an area) is shown to be lower in less deprived areas. Diversity at this measure is shown to be higher in urban areas which maybe associated with differential accuracy in mortality cause recording. The potential impact of this differential recording accuracy should therefore be considered when assessing the results presented here. However, it is likely impossible to address variable recording accuracy, nor was it within the scope of this thesis to assess to what degree it exists in Scotland.

4.5.2.3 Variation in mortality causes in the literature

Measuring the degree of uncertainty and variation in mortality causes faced by populations is not a topic that the public and population health literature have considered widely. Nor does it generally factor into public health practice where interventions are mostly designed to address defined risk factors or causes of ill-health. The implications of research which assesses diversity in mortality causes may however be invaluable in future especially as populations age[4]. Some of the findings of this chapter, such as that those in more deprived areas face worse rates of mortality due to external causes or that heart diseases mortality rates have increased in Scotland in the late 2010s, are well reported (Brown et al., 2019; Ramsay et al., 2020; Public Health Scotland, 2022b). The study of diversity adds a dimension to these findings showing that, for example, alongside increasing heart disease mortality rates in males following 2015, it is likely variation in rare causes within the population has also increased. This changes the context of public health aims. It means that while resources should continue to be directed towards acute myocardial infarction as the most common cause of mortality, an increasing variety of rare causes must also be catered for. Similarly, the need for public health messaging and interventions to focus on different causes of ill-health among different subpopulations are well discussed (King et al., 2008; Moore et al., 2014). I have shown, however, that in Scotland subpopulations divided by deprivation and urban-rural class have faced similar rates of diversification in causes of mortality. This has the potential to make the design of public health interventions more difficult as ever-wider subsets of causes of ill-health must be accounted for within each subpopulation. Here, I have presented two examples of the ways analysis of diversity in mortality causes can expand understanding of patterns of mortality. Increasing variation in mortality causes has potentially hazardous implications for health at the population and subpopulation level and should be addressed to a greater degree by the public health community.

[4]Further exploration of the relationship between age at death within a population and diversity in mortality causes is carried out in the next chapter.

4.5.3 Comparisons with existing studies

Previous study of diversity in causes of mortality is limited and no previous analysis has been performed solely within Scotland. However, the findings in this chapter broadly conform to findings reported by Bergeron-Boucher et al. 2020 across high-income countries and Trias-Llimos and Permanyer (2023) in the USA, namely that diversity in causes of mortality has increased over the first decades of the 21st century. These studies measured diversity in coarse-grained causes of mortality and both used only one measure of diversity. Bergeron-Boucher et al. 2020 utilise Shannon entropy (related to the measure $q = 1$ used in this chapter) while Trias-Llimos and Permanyer (2023) use a novel measure they propose. The findings of this chapter show that fine-grained causes of mortality have also diversified from 2001 to 2019 in Scotland. This indicates a greater degree of increased variation in mortality causes and fragmentation of diseases as deaths became more evenly distributed across fine-grained causes of mortality rather than groups of causes. This is likely to make both individual- and society-level implications of diversification in mortality causes more severe. As discussed in Section 2.4.3 the selection of a definition for causes of mortality is fundamental to the study of diversity in mortality causes. Both coarse-grained cause diversity and fine-grained cause diversity have merit as methods for assessing variation in causes of mortality. However, the greater granularity offered by fine-grained cause diversity makes it possible to include variation within groups of causes such as chapters.

Bergeron-Boucher et al. (2020) used a different age-breakdown to the one used in this chapter. The findings presented here show similar trends among those in older age groups; however, in younger age groups, findings diverge. The previous study showed little change in diversity in the UK, or other high-income countries, among those aged 0 to 19 or 20 to 49. In this chapter, causes of mortality associated with deaths of despair are identified as having been the drivers of falling diversity in deaths among these age groups. Deaths of despair are known to be increasing more quickly in Scotland than in comparable countries, which may explain some of this disparity.

4.5.4 Strengths and limitations

The research in this chapter presents the first application of the measure described here as additive value in the study of diversity in causes of mortality. Additive value allows for analysis of the contribution of causes and groups of causes to diversity. It is applicable across the measures of diversity examined here; however, due to the differential weighting of proportional prevalence the additive value of causes changes across different values of q. This complicates extracting findings from additive value when examining a range of diversity measures because the additive value of a cause can change dramatically between diversity at different values of q.

Using the Reeve et al. (2016) framework for the calculation of diversity, diversity is calculated as an "effective numbers of types" within each population. This allows for straightforward interpretation and comparison, something which can be problematic under other systems of diversity measurement. In this chapter, diversity in causes of mortality is calculated at, and compared across, a variety of measures (through the use of different q values) for the first time. This allows for a more complete analysis of variation in the distribution of mortality causes making it possible to distinguish between increased variation in relatively rare causes of mortality and changes in the prevalence of the most common causes.

In analysis within this chapter "garbage codes" are retained in the distribution of mortality causes when diversity in mortality causes is calculated, as discussed in Section . These causes of mortality are considered, in the literature, to have the potential to bias population health research. Methods to accurately redistribute these causes of mortality can be intricate and were not carried out in this case. Instead, the sensitivity of the results in this chapter to garbage codes is tested in Appendix B.2 using a simple redistribution. Garbage codes are redistributed proportionally by age, gender, year of death and ICD-10 Chapter and diversity in mortality causes is recalculated. Under this analysis the overall trends in diversity in mortality causes reported in this chapter for deaths at all ages across Scotland and in Scottish subpopulations and in deaths within twenty year age ranges across Scotland are upheld. Using an alternative, more intricate method for redistributing garbage codes may have produced different outcome here nonetheless, the analysis in Appendix B.2 suggests that the results reported in this Chapter are robust to the impact of garbage codes.

4.5.5 Research implications and next steps

This chapter presents a more thorough analysis of diversity in fine-grained mortality causes in the 21st century than previous studies by measuring and comparing across various measures of diversity and examining variation in fine-grained causes of mortality. Using different measures of diversity allowed me to identify adverse trends in the prevalence of common causes of mortality and show that they were not necessarily associated with similar trends at the level of rarer causes. To provide a direct examination of the causes and ICD-10 Chapters driving trends in diversity, I introduce additive value as a measure of the contribution of each cause, and group of causes, to diversity. This chapter builds on work I published during the completion of this thesis that shows trends in diversity in mortality causes at the level of subpopulations in Scotland by incorporating different measures of diversity (through varying q), expanding study of diversity within different age groups and including different subpopulations (McMonagle et al., 2022). Previous research has suggested that life expectancy and variation in lifespans within a population may, to an extent, explain trends in the diversity of mortality causes (Bergeron-Boucher et al., 2020; Trias-Llimós & Permanyer, 2023). To explore this possibility, the next chapter introduces the use of a measure of similarity-sensitive normalised alpha diversity in age at death in the study of variation in lifespan comparing this novel measure to existing, established measures in the field. I then examine the relationships between diversity in causes of mortality, the measure of diversity in age at death I propose and life expectancy. I explore these relationships over time and cross-sectionally between subpopulations, presenting the first formal analysis of such relationships at a sub-national level.

Chapter 5

The relationship between lifespan diversity and diversity in causes of mortality

5.1 Background

This chapter covers four main themes which are each introduced in turn in separate sections: the development of normalised alpha diversity as a measure of variation in lifespans; trends in lifespan variation in Scotland and Scottish subpopulations from 2001 to 2019 and disparities in lifespan variation between subpopulations in these years; and finally, the relationship between diversity in causes of mortality and variation in lifespan.

5.1.1 Applying diversity measures to lifespan variation

The first theme addressed in this chapter relates to the measurement of lifespan variation. Variation in lifespan refers to the degree of variability in the distribution of lifespans within a population. Variation in lifespans is a source of uncertainty in health outcomes as well as a measure of inequalities within a population. Life expectancy is, fundamentally, a measure of the average length of life within a population and can be used to understand differences between populations in expected lifespans. Around this average, however, lifespans within a population can

vary and the degree of variation can be an indicator of total inequalities in health within a population. It has been shown that greater variation in age at death, and in most cases, also lower life expectancies are seen among those in more deprived areas, those with lower educational attainment and those employed in manual labour (van Raalte et al., 2013; Permanyer et al., 2018; Van Raalte et al., 2018; Seaman et al., 2019; Hiam et al., 2021). In addition to implications for health inequalities, greater variation in age at death can impact the provision of care because health care systems must provide health and end-of-life care for individuals at a wider range of ages. In this chapter the term "lifespan variation" is used to refer to any measurement of variation in the distribution of lifespans within a population and "lifespan diversity" is used to refer to measurement of variation in lifespan performed using measures of diversity.

In this chapter, normalised alpha diversity is applied to the distribution of ages at death, creating a measure I call "lifespan diversity". Diversity measures have been applied to ages at death previously (Bergeron-Boucher et al., 2020), although comparisons with more established measures of variation in lifespan have not been reported. Measures of variation in age at death in the literature are used, for the most part, interchangeably and mostly produce similar results when measured in the same populations (van Raalte & Caswell, 2013). In this chapter I validate lifespan diversity by comparing several measures of diversity with established measures of variation in age at death. From these, a single measure of diversity in age at death is selected for analysis: similarity-sensitive normalised alpha diversity at $q = 1$.

In addition to comparing lifespan diversity to previous measures in the field, in this chapter I measure diversity in ages at death using data obtained from observed mortality counts and from life tables. Life tables are a standard tool used in demography and public health in the calculation of life expectancy and variation in lifespan. However, they rely on a considerable amount of population data making them difficult to calculate in some situations, for example, in historical settings or groups for which population data is not routinely collected. Measuring diversity in the observed ages at death in a population relies only on mortality data and therefore may be a useful measure in situations where life tables cannot be easily or reliably calculated.

5.1.2 Trends in lifespan diversity

To address the second theme of this chapter, I investigate trends in lifespan diversity in Scotland and in two sets of Scottish subpopulations, namely SIMD income deprivation quintiles and Scottish Government urban-rural classes. Prior research has indicated that Scotland has higher variation in age at death than many otherwise comparable countries, with greater lifespan inequality observed in more deprived areas (Seaman, 2017; Seaman et al., 2019). This previous research has shown narrowing inequalities in lifespan variation between deprivation quintiles from 2001 to 2011, following decades of widening differences. However, lifespan variation has not been reported in Scotland or in deprivation quintiles since 2011. As discussed in the previous chapter, improvements in various health outcomes have slowed or even reversed in the 2010s in Scotland (Fenton et al., 2019a). Examining lifespan variation can give vital insight into how these adverse trends have impacted inequalities in health.

No studies have examined variation in age at death across the urban-rural gradient in Scotland, despite well-studied differences in health outcomes between such areas (Levin & Leyland, 2005). While few international studies exist, it has been demonstrated that non-metropolitan populations in the USA experience greater lifespan variation than their urban counterparts (De Ramos et al., 2022). To reduce inequalities in health it is essential to understand how they manifest across nations, between subpopulatians and within subpopulations. By examining lifespan diversity and directly measuring the degree of variation in lifespans, we can reveal the underlying inequalities in health, ultimately helping us to address them.

5.1.3 The relationship between variation in lifespan and diversity in causes of mortality

The third theme addressed in this chapter is the relationship between variation in age at death and diversity in causes of mortality. This relationship is unclear despite previous research in this area (Bergeron-Boucher et al., 2020; Trias-Llimós & Permanyer, 2023). Two hypotheses have been proposed for this relationship in the literature.

5. Lifespan diversity and diversity in causes of mortality

Bergeron-Boucher et al. (2020) propose the first hypothesis. In a study of mortality in developed countries including the UK, they observed that over time, diversity in causes of mortality increased. Over the same period, they also observed an increase in life expectancy and a decrease in variation in lifespan. These trends in life expectancy and lifespan variation imply that the population was ageing, and specifically, that more individuals were living to - and dying at - older ages. They therefore suggest that the increase in diversity in causes of mortality is driven by the ageing of the population: living to an older age means exposure to more potential diseases and risk factors, and thus a greater range of causes of mortality.

The second hypothesis is proposed by Trias-Llimós and Permanyer (2023) who suggest that in subpopulations "the higher the lifespan inequality, the higher the cause of death diversity, which coheres with the fact that deaths at more dissimilar ages are likely to be caused by a more variegated set of factors than deaths occurring at more similar ages." That is, by sampling from a larger range of ages, one should sample from a larger range of causes. This study examined subpopulations in the USA and found greater diversity in causes of death in less educated groups. While they do not measure lifespan variation, they use evidence from previous research to argue that, cross-sectionally (i.e. in a single year), diversity in mortality causes should be greater in subpopulations where ages at death were more varied.

These hypotheses are, to an extent, contradictory. The first suggests increasing diversity in causes of mortality is caused by falling lifespan diversity while the second indicates diversity in causes of mortality would be highest where lifespan diversity is greatest. However, it may be argued they could plausibly co-exist because they are suggested to apply to temporal and cross-sectional relationships respectively. I suggest that these hypotheses indicate two potential mechanisms which might explain the relationship between lifespan variation and diversity in causes of mortality.

The first mechanism, suggested by the hypothesis of Bergeron-Boucher et al. (2020), is linked to the fact that falling lifespan diversity is generally linked to populations dying at more homogeneous, older ages. As people reach an advanced age, they are exposed to a wider variety of diseases and risk factors, which may increase the diversity in mortality causes (Bergeron-Boucher et al., 2020).

Under the second mechanism, it is possible that in subpopulations where individuals die at more varied ages, a greater diversity in causes of mortality would also occur. The causes of mortality that an individual is likely to face vary widely at different ages and a number of causes exclusively (or near-exclusively) occur at a restricted set of ages (Plana-Ripoll et al., 2022). For example, neonatal deaths occur only among those aged 0 and degenerative diseases mostly affect older individuals aged 60+. Therefore, a greater variety of ages at death, and age-specific causes of death, occurring in a population would be expected to increase diversity in causes of mortality. This might manifest most strongly under less conservative measures of diversity (such as $q = 1$) because they are sensitive to rare causes of mortality. However, at more conservative measures, this relationship might not be apparent as common causes are weighted more heavily and many age-specific causes of death which cause a relatively small number of deaths are, effectively, ignored.

To explore these mechanisms and hypotheses, I examine the relationship between variation in lifespan and diversity in causes of mortality in this chapter. I first investigate this relationship over time in the population of Scotland and in Scottish subpopulations. I then test the relationship cross-sectionally between subpopulations. This work represents the first formal analysis the relationship between variation in lifespan and diversity in causes of mortality both temporally and cross-sectionally in the same subpopulations. Therefore, it is the first to explore the two hypotheses discussed in this section together.

5.2 Research questions and objectives

This chapter had the following research objectives:

- To develop a suitable measure of lifespan diversity and make comparisons with existing measures of variation in age at death.
- To explore trends over time in normalised alpha lifespan diversity within Scotland.
- To explore disparities in lifespan diversity by sex, age at death, income deprivation quintile, and Scottish Government urban-rural classification.

- To examine the cross-sectional and temporal relationships between: lifespan diversity, life expectancy and diversity in causes of mortality.

As such it aimed to answer the following questions:

- Is normalised alpha diversity in age at death (lifespan diversity) well-correlated with existing measures of lifespan variation?
- Can lifespan diversity calculated from observed death counts be considered a reliable proxy for lifespan diversity calculated from the lifetable death distribution?
- What were the temporal trends in lifespan diversity in Scotland and Scottish subpopulations from 2001-2019?
- What is the relationship between diversity in causes of mortality and lifespan diversity in Scotland, firstly, across time and secondly, between subpopulations?

5.3 Methods

5.3.1 Data

Mortality data for the years 2001 to 2019 was extracted and processed as described in Section 3.2.1. Annual mid-year small area population estimates for 2001 to 2019 were obtained for Scotland as described in Section 3.2.2.

5.3.2 Subpopulation geographies

Income deprivation quintiles were calculated using the SIMD income domain as described in Section 3.2.3. Populations of Scottish Government urban-rural classes were grouped as described in Section 3.2.4.

5.3.3 Life tables

Multiple-decrement life tables were constructed as described in Section 3.3.1 for Scotland as a whole, as well as for each SIMD income deprivation quintile and Scottish Government urban-rural class. Life tables were constructed for single years of age from 0 to 109 with an open-ended age class of 110+. These life tables were created separately for males and females in each year from 2001-2019 for the population of Scotland and for each subpopulation.

5.3.4 Indices of variation in lifespan

Six indices of variation in age at death are calculated in this chapter to allow comparison between diversity in age at death and established measures in this field. These indices are described in Section 3.3.4, and were calculated using values from the life tables constructed as above. Each index was calculated separately for males and females in each year from 2001 to 2019.

5.3.5 Normalised alpha diversity

5.3.5.1 Diversity in causes of mortality

Normalised alpha diversity in causes of mortality was calculated as described in Section 3.3.3 using distributions of mortality causes extracted from multiple-decrement life tables. For each year it is calculated in the population of Scotland as a whole and in each SIMD income deprivation quintile and Scottish Government urban-rural class across all ages. In each case diversity at $q = 1$, $q = 2$, and $q = \infty$ was calculated separately for males and females in each year 2001 to 2019. Calculations were performed within each population or subpopulation with years as subcommunities.

5.3.5.2 Lifespan diversity

Lifespan diversity (normalised alpha diversity in age at death) is calculated under Equation 5.1 for each population and subpopulation where *M*(1-*q*) is the weighted power mean of the order 1 - *q*; $\overline{P_{\cdot j}}$ is a vector of the proportion of deaths which occur at each age in year *j* ($P_{\cdot j}$) divided by the proportion of the deaths within the population in the study period which occurred in year *j* (w_j); and $(Z\overline{P_{\cdot j}})_i$ the proportion of deaths which occur at ages similar to age *i* with *Z* calculated as described below.

$$ {}^{q}\overline{\alpha}_j^Z = M_{1-q}(\overline{P_{\cdot j}}, (Z\overline{P_{\cdot j}})_i^{-1}) \tag{5.1} $$

Up to this point in this thesis, diversity has been calculated under the assumption of naïve similarity. This is because under naïve-similarity diversity, it is assumed that each type, for example, each year of age, is completely (and thus equally) distinct from all other types. In reality, this is clearly not true, and it may be desirable to acknowledge that deaths in those 75 years old are more similar to deaths at 74 years old than they are to deaths at 25 years old. In the calculation of diversity, a similarity matrix can be used to quantify these similarities and make explicit the connection between deaths at similar ages. The concept of similarity in the measurement of diversity is discussed in Section 2.3.2.

Following work by Leinster and Cobbold (2012) similarity is incorportated into the Reeve et al. (2016) framework for the calculation of diversity through a pairwise-matrix *Z* with each entry Z^{ij} representing the similarity between ages *i* and *j*. Here, a similarity matrix was calculated under a formula proposed by Mitchell (2019) for developing similarity matrices in age distributions. The absolute value of *i* minus *j* (d_{ij}) is transformed through the formula $Z^{ij} = e^{-kd_{ij}}$, where *k* is a scaling factor. In developing these similarity matrices *k* can take any positive value. Here, to calibrate *k*, candidate values were used to create a range of similarity matrices which differentially weight distance between ages. These matrices were then used to calculate the similarity-sensitive metacommunity gamma diversity at q=1 ($^1G^Z$) of an age distribution. This measure can be interpreted as the effective number of age classes in the age distribution (Mitchell, 2019).

5. Lifespan diversity and diversity in causes of mortality

For this analysis, the similarity matrix was calibrated to the age structure of the Scottish population. This total age structure was the male and female population estimates for Scotland with single year age classes from age 0 to 89 and an open-ended age class 90+ in the years from 2001-2019 combined. ${}^1G^Z$ was calculated for this population with ages as types (Reeve et al., 2016; Mitchell, 2019). Using this distribution, rather than the distribution of ages at death, means that the range of effective age classes at which people die can be compared to the number of effective age classes in the living population.

Here, k was gradually reduced from a value of 1 (where Z^{ij} is equal to the negative exponential of the distance in age) to create similarity matrices which produced different values of ${}^1G^Z$ (Appendix Figure A.5). Ultimately, a value of k was chosen which produced a similarity matrix that when applied to population data as described above produced an ${}^1G^Z$ of approximately 6 ($k = 0.12962$). This value was chosen to align with the six twenty-year age classes used in Chapter 4. Values of diversity calculated using this similarity matrix can be considered a count of the effective number of age classes in which deaths occurred. To use an example, if subpopulation A was found to have a diversity of 3 it would imply deaths within this subpopulation died within 3 effectively distinct age classes meaning that deaths occurred at effectively half of the age classes which make up the population of Scotland. Say that a second diversity value of 4.5 was produced for subpopulation B, this would imply that in subpopulation B individuals died within effectively 50% more age classes than in subpopulation A. It is important to note that the age classes mentioned here do not refer to classifiable ranges of ages within the population rather to the effective number of age classes within the population.

Similarity sensitive diversity in lifespans at q =1 and q = 2 were calculated from both life table data and observed mortality data at each age of death extracted from NRS mortality data. These calculations were performed in males and females across the Scottish population for each year from 2001 to 2019. A single measure was chosen following an exploration of these measures and their correlation with the established indices of variation in age at death described previously. This measure, similarity-sensitive lifespan diversity at q = 1 calculated using life table data, was used to examine the distributions of age at death further in this chapter.

5.3.6 Analysis of correlation

The Pearson correlation coefficient is used to assess linear covariance in a number of places in this Chapter. To answer Research Question 1, required analysing the correspondence between diversity in age at death and the pre-existing indices of variation in lifespan; to do this, pairwise correlation was measured in the relationships between each of the pre-existing measures of variation detailed in Section 5.3.4 and each measure of normalised alpha lifespan diversity described in Section 5.3.5.

To answer the fourth research question addressed in this chapter, two relationships were tested, between: diversity in causes of mortality and life expectancy; and diversity in causes of mortality and lifespan diversity. The Pearson correlation coefficient of these relationships were both measured temporally across the years 2001 to 2019, for deaths across Scotland as a whole and in each subpopulation described in Section 5.3.2. Pearson's correlation coefficient was also measured for these relationships cross-sectionally within each year 2001 to 2019, between subpopulations. Only selected years are presented in this chapter to give an overview across the study period while avoiding repeating analysis[1]. These years were: 2001, 2005, 2009, 2011, 2015, 2019. For both temporal and cross-sectional analysis correlation was measured separately for each sex. The correlation coefficient and p-values are displayed in the figures where the relevant relationship is assessed, values are repeated in Appendix Tables A.8, A.9, and A.11 for ease of access.

5.3.7 Software

Analysis in this chapter was performed in *R* version 4.1.3 (R Core Team, 2022). Data manipulation and processing was performed using version 1.3.1 of the *tidyverse* package, plots were created using version 3.3.5 of the *ggplot2* package and version 0.92 of the package *corrplot* alongside version 1.1-2 of the package *RColorBrewer* and version 0.4.0 of the package *ggpubr* (Neuwirth, 2014; Wickham, 2016;

[1]Correlation coefficients and p-values of the cross-sectional relationship between subpopulations in each year 2001-2019 are shown in Appendix Table A.11.

Alboukadel Kassambara, 2020; Taiyun Wei et al., 2021). Calculation of diversity was performed using version 2.0 of the *rdiversity* package (Mitchell et al., 2020). Analysis of correlation was performed using version 2.2.9 of the package *psych* (Revelle, 2022).

5.4 Results

The results section of this chapter is structured as follows. In all analysis in this section, deaths of males and females are analysed separately. First, Section 5.4.1 addresses the calculation of lifespan diversity as a measure of variation in lifespan, using both life table and observed data . Next, in Section 5.4.2 lifespan diversity is established as a measure of variation in lifespan through comparison with existing measures in use in the literature. Through insight gained in these sections a single measure of lifespan diversity is chosen for further analysis. Using this measure trends and disparities in lifespan diversity across Scotland and Scottish subpopulations are examined in Section 5.4.3. The final results sections of this chapter present an examination of two relationships: between diversity in causes of mortality and life expectancy (Section 5.5.2); and between diversity in causes of mortality and lifespan diversity (Section 5.4.5). In order to investigate the fourth research question addressed in this chapter, each of these relationships is first examined temporally across Scotland as a whole and in Scottish subpopulations. Then these relationships are examined cross-sectionally (i.e. within individual years) between subpopulations. The fourth research question examined here focuses on the relationship between diversity in causes of mortality and lifespan diversity however, understanding the relationship between diversity in mortality causes and life expectancy helps to contextualise this analysis.

5.4.1 Calculation of a measure of normalised alpha lifespan diversity and comparison to existing measures

This chapter considers two measures of similarity-sensitive normalised alpha lifespan diversity: diversity at $q = 1$ and $q = 2$. Diversity at $q = 2$ is the more conservative in that it more strongly weights ages at which deaths occur more often. In Chapter 4, three measures of normalised alpha diversity in causes of mortality were studied.

Diversity at $q = \infty$ in age at death is not measured due to its strong focus on the most common type. It would amount to, effectively, a measure of the proportion of deaths which occur at the modal age. For causes of mortality, measuring diversity at $q = \infty$ gives valuable insight into the mortality pressures associated with the most common causes. However, it was considered important to use measures which address the entire distribution of ages at death rather than effectively excluding rare ages of death.

Trends in lifespan diversity, calculated using observed data and life table data, across Scotland in males and females are shown in Figure 5.1. Calculating lifespan diversity using life tables rather than from observed data had an effect on both overall levels of diversity and trends over time. Specifically, using life tables resulted in lower diversity values at both $q = 1$ and $q = 2$; it also revealed a slight reduction in lifespan diversity over the study period not seen in the analysis using raw death data. The difference in trends between data sources is more pronounced in females. These trends in lifespan diversity are examined in more detail later in this chapter following the selection of a single measure.

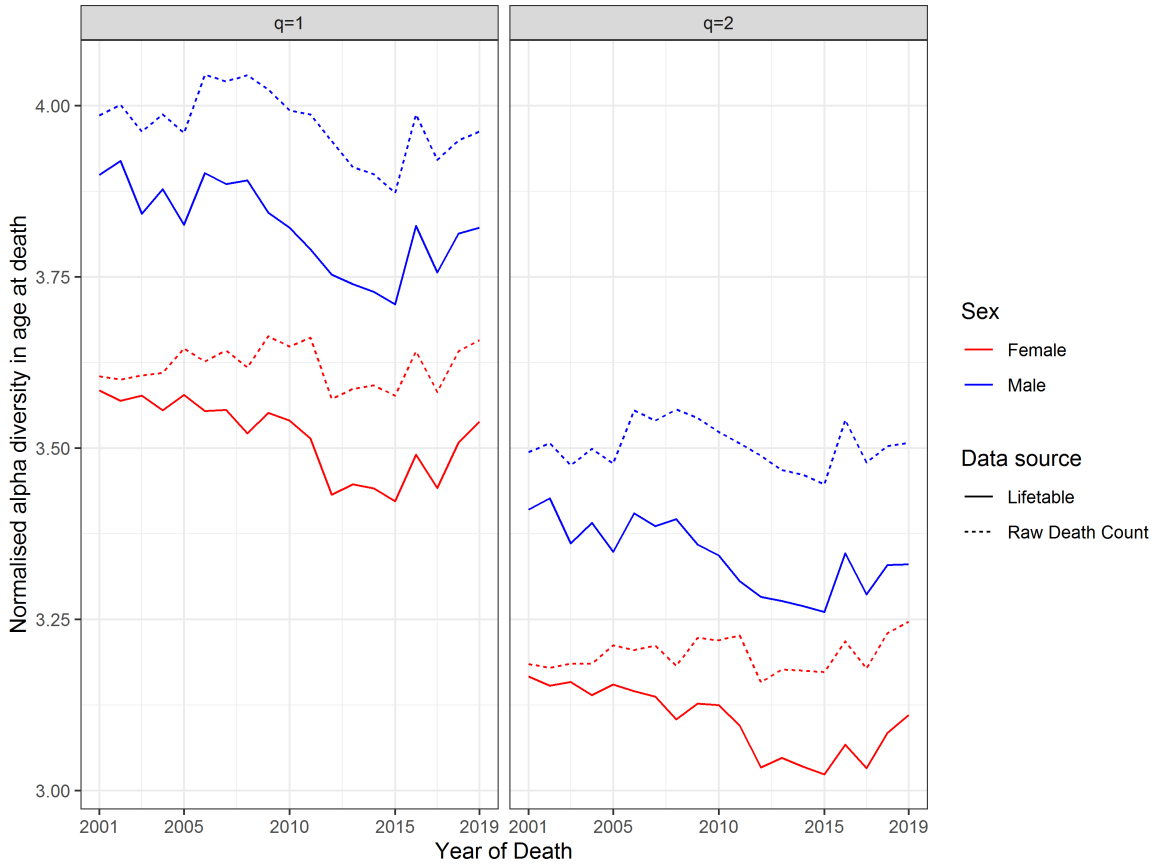


Figure 5.1: Similarity-sensitive normalised alpha lifespan diversity in each year 2001 to 2019 across Scotland in males and females using data from life tables and observed raw death counts.

To understand the differences in lifespan diversity between observed data and life tables, Figure 5.2 shows the distributions of age at death in the Scottish population using each method in 2001 and 2019. Differences in the distributions can be observed, reflecting the larger emphasis on deaths at older ages introduced through the calculation of life tables. In this calculation – for populations in which fewer than 100,000 deaths occur per year – the cohort size is increased and through modelling old age mortality an increased proportion of the population is assumed to live to an older age. Appendix Figure A.6 replicates Figure 2 but further includes life table data from the Human Mortality Database, demonstrating that this shift in the distribution is also found in what is generally considered to be the highest standard of life table data (Wilmoth et al., 2022). Despite the difference in overall trends produced using these different data sources, it is notable that much of the year-to-year patterns in diversity are consistent (Figure 5.1). For example, a spike in diversity in 2016 in both sexes can be observed comparably across measures.

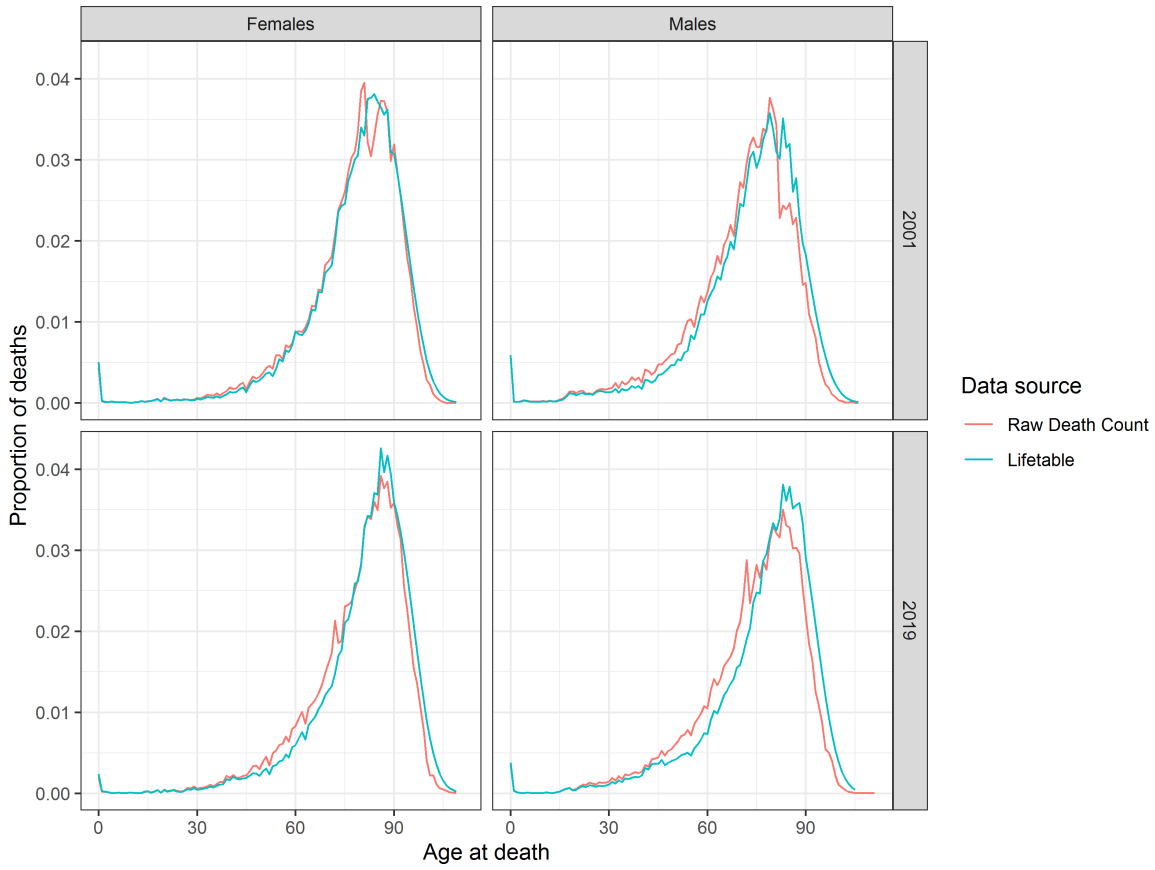


Figure 5.2: Distributions of age at death in the Scottish population from life tables and observed raw death counts in males and females in 2001 and 2019.

5.4.2 Establishing lifespan diversity against existing measures of variation in lifespan

Variation in the distribution of age at death within populations has been examined using various measures previously. Using data from France and Russia, Van Raalte et al. (2013) show a high degree of correlation between measures including e†, the Gini coefficient and inter-quartile range (IQR). Correlation between the measures of lifespan diversity introduced in this chapter and a number of indices of variation in lifespan are examined in Figure 5.3. More detail on the formulation and interpretation of these indices can be found in Section 3.3.4.

Figure 5.3 shows that lifespan diversity calculated from life table data is highly correlated with most of the measures examined here. Measures of lifespan diversity calculated from observed data are less highly correlated (alhough note that p-values of all correlations were <0.05, Appendix Table A.12). However, the difference between diversity calculated from these two data sources is slight, suggesting that lifespan diversity calculated using observed data may be a useful tool in situations where population structure data is insufficient to construct reliable life tables.

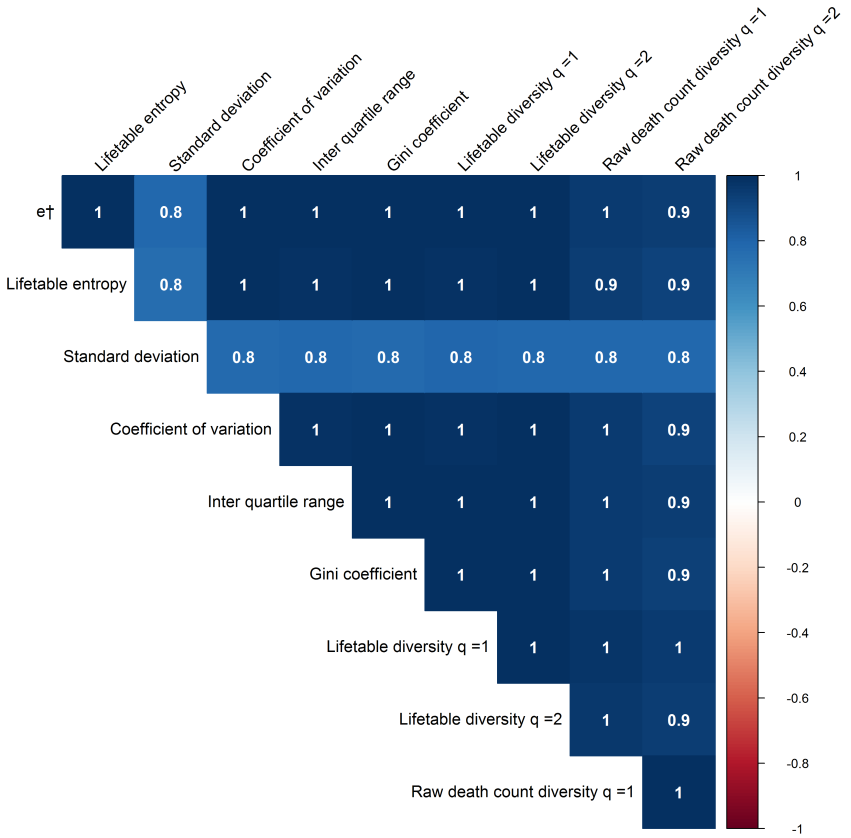


Figure 5.3: Correlation matrix showing the relationship between various measures of variation in age at death and measures of lifespan diversity.

Each of the measures of normalised alpha lifespan diversity calculated from life table data are found to be well correlated to each other and with the established measures examined here (these measures are discussed in Section 3.3.4). Alongside this, evidence from Figure 5.1 shows that the measures of diversity calculated from life tables introduced in this chapter display similar trends and do not present distinct insights. Little difference is observed in trends in lifespan diversity under q = 1 or 2. To avoid repetition of analysis only one measure was taken forward, diversity at q = 1. Therefore, similarity-sensitive normalised alpha diversity at q = 1 calculated from life table data is taken forward for analysis in this chapter.

5.4.3 National and subpopulation temporal trends in lifespan diversity

In the 2000s, lifespan diversity across the Scottish population, calculated from distributions of age at death extracted from life tables, (black lines in Figure 5.4) was relatively constant. This persisted until around 2008 in males and 2010 in females. After these years diversity in lifespans fell until 2015 in both sexes after which an increase in diversity can be observed to 2019. Despite the increase in lifespan diversity both sexes after 2015, the Scottish population as a whole died at slightly less diverse ages in 2019 than in 2001. This represented a reduction from effectively 3.58 age classes to effectively 3.53 in females and 3.89 to 3.82 in males.

Lifespan diversity was observed to be around 15-20% higher among males in the most deprived areas of Scotland than in the least deprived areas (Top panels of Figure 5.4). In all years, this disparity was smaller among females. Deaths occurred at less diverse ages in rural areas than urban areas (Lower panels of Figure 5.4). However, differences are found to be smaller between these areas than between deprivation quintiles. Lifespan diversity is higher in large urban areas than other urban areas but for the most part diversity is similar in remote and accessible classes for both small towns and rural areas.

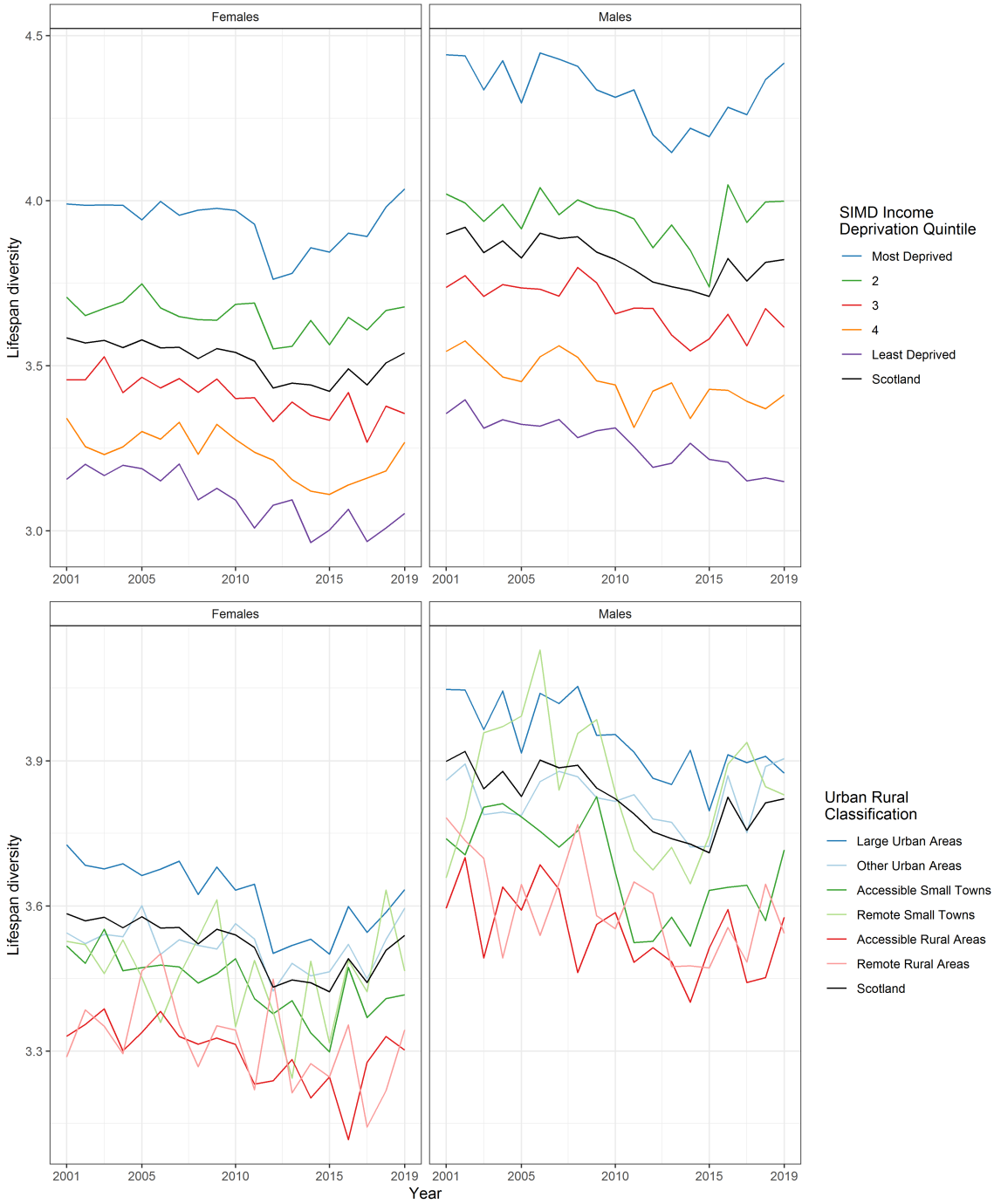


Figure 5.4: Trends in normalised alpha lifespan diversity at q=1 in Scotland in the years 2001 to 2019 in males and females, across the country as a whole (black line) and in deprivation quintiles (top row) and urban/rural classes (bottom row).

Previous study has suggested that from 2001 to 2011, a slight reduction occurred in deprivation related inequalities in lifespan inequality in Scotland (Seaman et al., 2019). However, variation in ages at death has not been studied since, despite stalling improvements in a number of measures of population health in the 2010s (Fenton et al., 2019a). Figure 5.4 shows plateaus in lifespan diversity from

2001 to the mid-2000s across deprivation quintiles. After this lifespan diversity fell across quintiles, especially in more deprived areas. Differences in lifespan diversity between quintiles were, therefore, slightly smaller at the early in the 2010s than they were early in the 2000s, as observed by Seaman et al. (2019). However, trends in diversity in age at death diverge between quintiles after 2015. In both males and females, deaths occurred at increasingly varied ages in quintiles 1 and 2 - the most and second-most deprived quintile - towards the end of the study period. In contrast, lifespan diversity continued to decrease in the least deprived areas in both sexes. Iequalities in lifespan diversity between the most and least deprived fifths of the population have therefore increased in the 2010s in both males and females. As a result the difference in lifespan diversity between the most and least deprived quintile was wider in 2019 than it was in 2001.

Yearly fluctuations in lifespan diversity are larger in the populations of Scottish Government urban-rural classes than in deprivation quintiles. These larger fluctuations are likely linked to smaller population sizes in some of these groups. Despite this, inequalities between urban-rural classes changed little over the study period with the largest difference between urban rural subpopulations remaining at a lifespan diversity of around 10%. Trends in each urban-rural subpopulation are mostly similar to those across the country as a whole.

5.4.4 The relationship between diversity in causes of mortality and life expectancy

Trends over time in diversity in causes of mortality within Scottish subpopulations are examined in Chapter 4. Figure 5.5 plots diversity in causes of mortality, calcualted from distributions of mortality causes extracted from multiple decrement life tables against life expectancy across the Scottish population; in deprivation quintiles (upper panels of Figure 5.5); and in Scottish Government urban-rural classes (lower panels of Figure 5.5). In each case, diversity in causes of mortality was greater in years with higher life expectancies. As these populations experienced longer lifespans, variation in the causes of mortality they were likely to face increased.

For the most part, this relationship is consistent across the subpopulations examined here. However, in more deprived areas, a weaker correlation is observed at more conservative measures of diversity in causes of mortality. This suggests that over time a higher average age at death across all subpopulations has been associated with greater variation in the relatively rare causes of mortality they face. However, in more deprived areas, higher life expectancies have been associated with less of a reduction in the dominance of the most common causes of mortality than in other subpopulations.

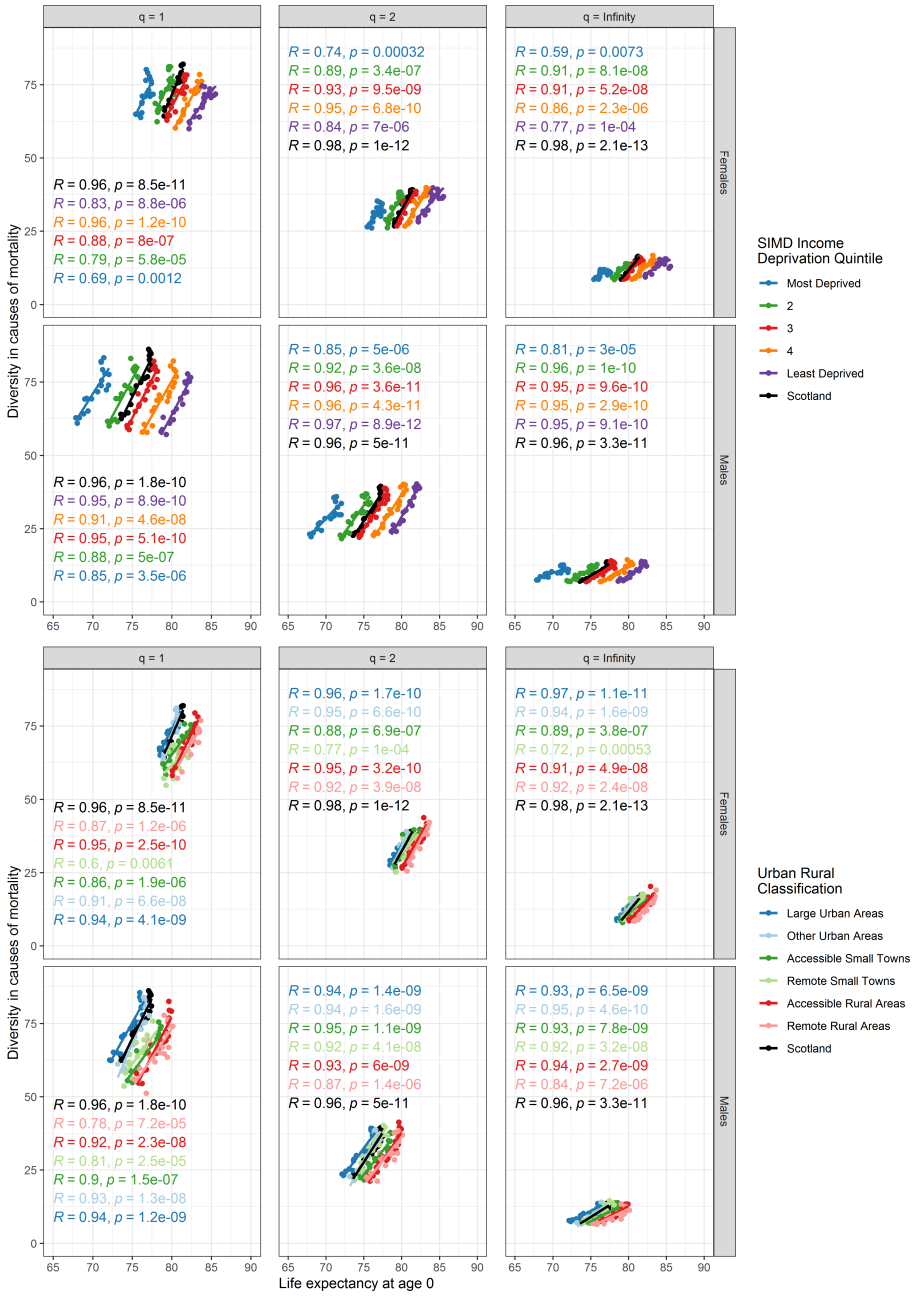


Figure 5.5: The relationship between normalised alpha diversity at q=1, q = 2 and q = infinity in causes of mortality and life expectancy at age 0 in Scotland in the years 2001 to 2019 in males and females, across the country as a whole and in deprivation quintiles and urban/rural classes. Pearson's correlation coefficient and associated p-values are presented, colour coded, for each relationship.

In spite of the positive relationship between diversity in mortality causes and life expectancy when observed by year in subpopulations in Scotland (and across the total population), mortality cause diversity is not, generally, observed to be higher among subpopulations with longer life expectancies (Appendix Table A.10).

In fact, in a number of years diversity in causes of mortality is found to be lower in subpopulations with higher life expectancies. This confirms the supposition, posed in Section 2.4.1.3, that there may be a Simpson's paradox in the relationship between diversity in causes of mortality and life expectancy.

The relationship between variation in lifespan and life expectancy has previously been well studied. A consistent tendency for reduced variation in lifespan as life expectancy increases has been observed across the world (Nigri et al., 2021). Furthermore, a strong, but not universal, tendency for lower lifespan variation in populations and subpopulations with longer lifespans is known to exist. The relationship between lifespan diversity and life expectancy measured in this study follows this pattern (Appendix Figure D).

5.4.5 The relationship between lifespan diversity and diversity in causes of mortality

This section presents an examination of the relationship between normalised alpha diversity at q = 1, 2 and ∞ in causes of mortality and similarity-sensitive lifespan diversity at q = 1 with both measures calculated using life table data. As discussed in Section 5.1, two hypotheses are tested in this section. The first hypothesis is that diversity in causes of mortality increases as a population lives to longer and more homogenous ages. The second hypothesis is that in populations with more heterogenous ages at death diversity in causes of mortality would be higher. To explore these hypotheses, first, I examine temporal relationships assessing the contention that as the population lives to a longer age on average and lifespan diversity narrows around this average, diversity in causes of mortality would increase. Then I examine cross-sectional relationships to evaluate the claim that greater lifespan diversity is associated with greater diversity in causes of mortality.

Over time in Scotland and across Scottish subpopulations, reductions in lifespan diversity have been associated with increasing diversity in causes of mortality. This upholds the first hypothesis raised in the previous paragraph and means that when people died at less varied ages they died due to a wider variety of causes. This

relationship is slight although it is generally more evident at more conservative values of q in diversity in causes of mortality. This suggests a reduced likelihood of facing the most common causes of mortality when lifespans were more equal across the population.

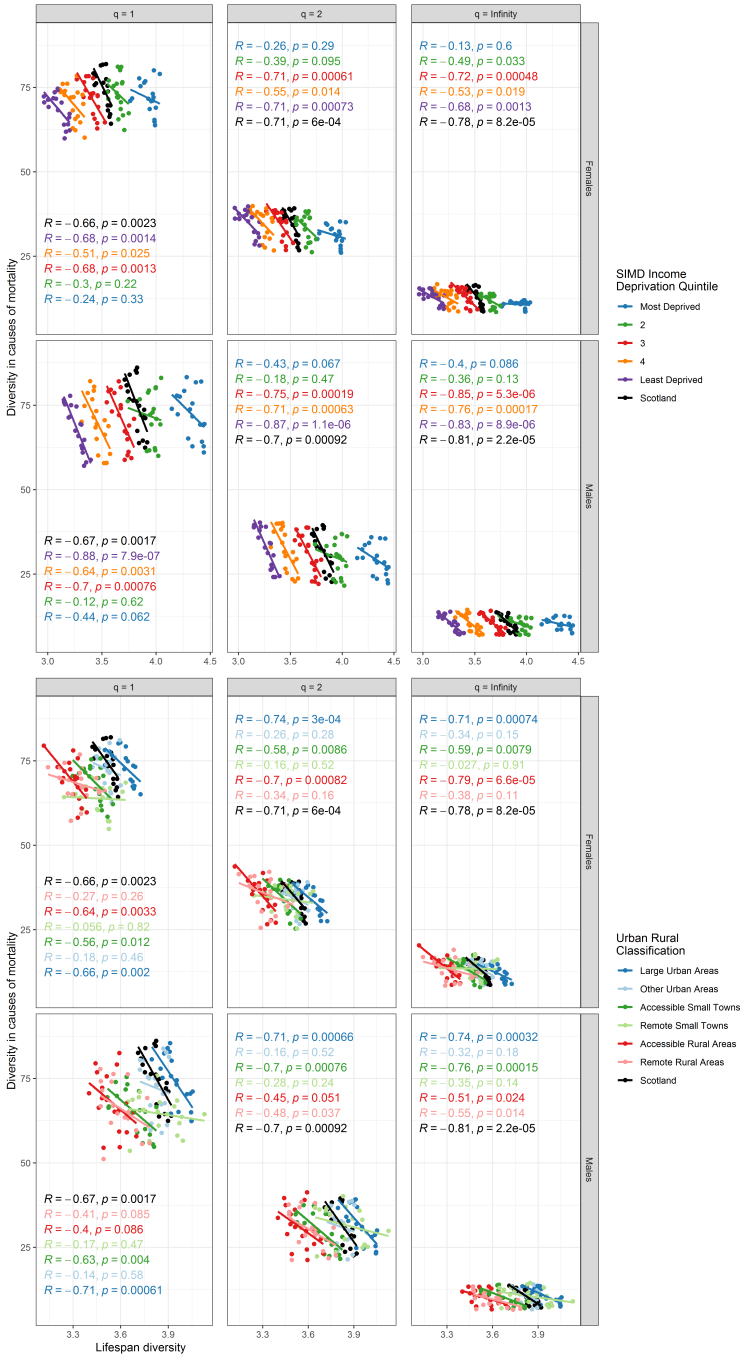


Figure 5.6: The relationship between lifespan diversity at q=1 and normalised alpha diversity at q=1, q = 2 and q = ∞ in causes of mortality in Scotland in the years 2001 to 2019 in males and females, across the country as a whole and in deprivation quintiles and urban-rural classes. Pearson's correlation coefficient alongside p-values are presented, colour coded, for each relationship.

In the breakdowns by subpopulation, increasing diversity in causes of mortality over time tended to occur alongside falling lifespan diversity in most SIMD quintiles, although evidence for this trend generally weakens in more deprived areas (upper panels of Figure 5.6). This relationship is also observed in large urban areas, accessible small towns and accessible rural areas, while in the remaining urban-rural classes it is weaker (lower panels of Figure 5.6).

The cross-sectional relationship between diversity in causes of mortality and lifespan diversity between deprivation quintiles and urban-rural classes is shown in Figures 5.7 and 5.8 respectively. In both Figures, there is some inconsistency between years. Despite this, some overall conclusions can be drawn from this analysis.

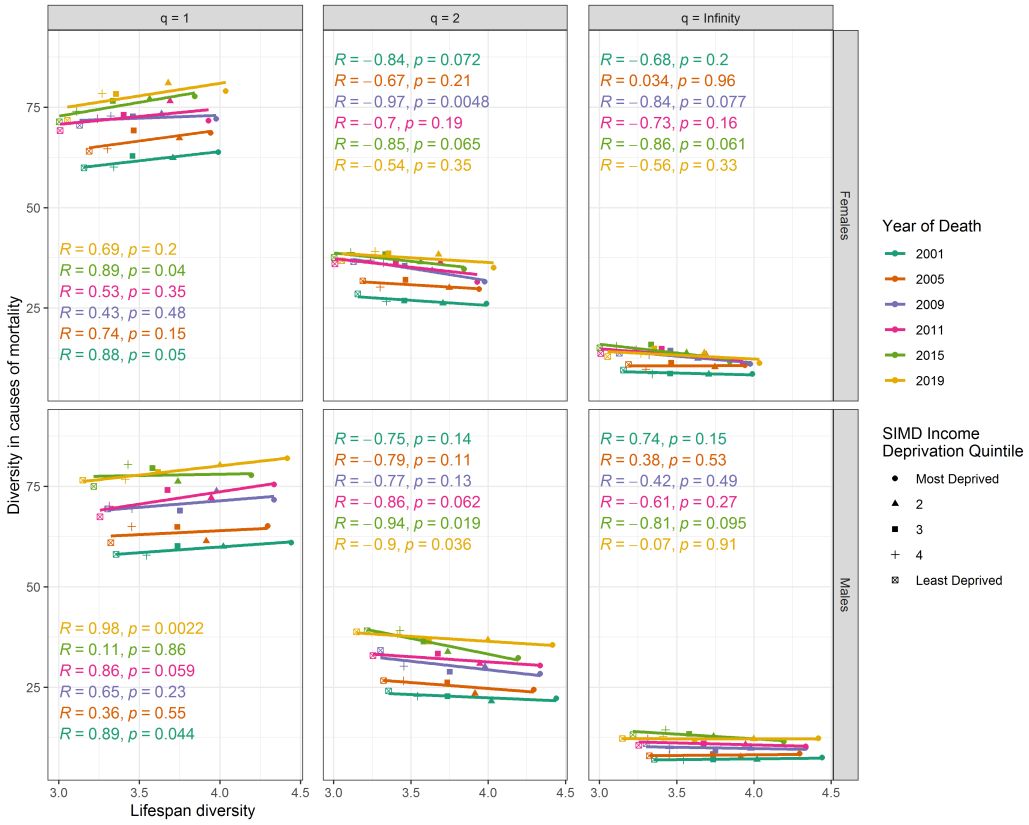


Figure 5.7: The relationship between normalised alpha lifespan diversity at q=1 and normalised alpha diversity at q=1, q = 2 and q = infinity in causes of mortality across deprivation quintiles in Scotland in selected years in males and females. Pearson's correlation coefficient and associated p-values are presented, colour coded, for each year. Different subpopulations in each year are denoted by the shape of points.

In deprivation quintiles a positive relationship can be observed between diversity in causes of mortality at q = 1 and lifespan diversity. This indicates that in quintiles with greater variation in age at death more varied causes of death occurred. The strength of correlation between lifespan diversity and diversity in causes of mortality varies over years but generally this conclusion holds. This upholds the conclusion of Trias-Llimós and Permanyer (2023), the second hypothesis raised at the beginning of this section. However, this hypothesis is not confirmed at more conservative measures of diversity in causes of mortality in the analysis show in Figure 5.7. In fact, when rare causes are discounted and common causes are weighted more heavily a negative relationship is observed between lifespan diversity and diversity in causes of mortality. This relationship at more conservative measures of diversity means that in populations in which lifespan variation was higher the most common causes of death were more prevalent. In other words where individuals died at more varied ages they were more likely to face the most common causes.

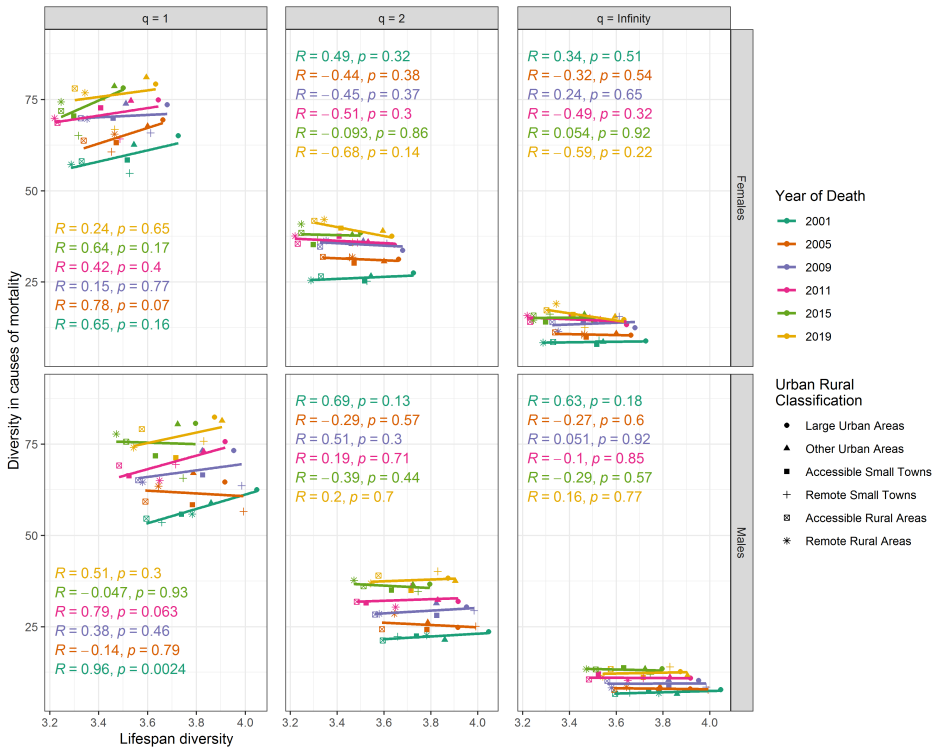


Figure 5.8: The relationship between normalised alpha lifespan diversity at q=1 and normalised alpha diversity at q=1, q = 2 and q = infinity in causes of mortality across urban-rural classes in Scotland in selected years in males and females. Pearson's correlation coefficient and associated p-values are presented, colour coded, for each year. Different subpopulations in each year are denoted by the shape of points.

When examining the cross-sectional relationship between diversity in causes of mortality and lifespan variation in urban-rural classes (Figure 5.8), a positive relationship is observed at $q = 1$. This is similar to the finding I report above for deprivation quintiles. It adds further evidence to the claim that in populations with more heterogeneous ages at death more varied causes of mortality will occur. However, at more conservative measures of diversity in causes of mortality no clear cross-sectional relationship is observed. This indicates that there is some complexity in this relationship and it is not necessarily true that in subpopulations more varied ages the most common causes are more prevalent.

5.5 Discussion

This chapter aimed to calculate and validate a measure of lifespan diversity and use that measure in the examination of the relationship between life expectancy, lifespan diversity and diversity in causes of mortality. Here, I summarise, interpret and discuss the findings of this chapter in four sections which are structured around the research questions addressed in this chapter. The first findings section relates to the suitability of lifespan diversity as a measure of variation in lifespans (Section 5.5.1.1). In the second findings section I explored the utility of lifespan diversity calculated using observed data as opposed to life table data Section 5.5.1.2) . The third section covers trends in lifesan diversity in Scotland and Scottish subpopulations as well as disparities between subpopulations (Section 5.5.1.3). Finally, the fourth findings section relates to the temporal and cross-sectional relationships between diversity in causes of mortality and lifespan diversity in Scotland and Scottish subpopulations (Section 5.5.1.4). Following discussion of the findings of this chapter, I propose potential future research directions (Section 5.5.2) and discuss the implications of this research (Section 5.5.4).

5.5.1 Principal findings and interpretation

5.5.1.1 Is normalised alpha diversity in age at death (lifespan diversity) well-correlated with existing measures of lifespan variation?

Four measures of similarity-sensitive lifespan diversity are explored in this chapter, using different values of q and measuring from life tables and observed data. Following this, a single measure of diversity was chosen for analysis of lifespan diversity in this chapter. This measure was lifespan diversity at q = 1 calculated from life table data. Life table data was chosen to align with the calculation of diversity in mortality causes and with measures in use in the literature. Little difference is found between between overall trends in diversity at q = 1 and q = 2. Only one measure was taken forwards to avoid effective repetition of analysis; q = 1 was chosen as it is related to Shannon entropy a measure which has been used previously in study of diversity in lifespans (Bergeron-Boucher et al., 2020).

Each of the measures of lifespan diversity assessed in this chapter is shown to be highly correlated with established measures of variation in lifespan. The established measures assessed in this chapter included specialised demographic measures such as lifespan inequality (e†) and statistical measures like the IQR. These established measures represent a number of the standard measures used in the literature for analysis of variation in lifespan (van Raalte & Caswell, 2013; Shkolnikov & Andreev, 2010). These findings suggest that lifespan diversity has merit as a measure of variation in lifespan.

5.5.1.2 Can lifespan diversity calculated from observed death counts be considered a reliable proxy for lifespan diversity calculated from the lifetable death distribution?

Diversity in age at death is calculated in this chapter using life table mortality data and observed raw death counts. The calculation of reliable life tables requires a considerable volume of detailed population data, whereas lifespan diversity calculated from observed death counts requires only data for age at death within the population. Diversity in age at death using raw counts is therefore explored here as a possible proxy for more established measures of variation in lifespan in cases when such detailed information is not available. The findings in this chapter support the use of lifespan diversity calculated using observed data by demonstrating that it is relatively well correlated with established measures calculated from life table data. In contemporary studies of developed nations it may be rare to have mortality data in the absence of accompanying, and reliably comparable, population data. However, particularly in historical contexts it is possible for mortality data to be more reliable than population data (Asah et al., 2017; Villegas & Hiabu, 2021). However, the evidence presented in this chapter suggests that lifespan diversity calculated using observed data may be considered a reliable proxy for measures which rely on more extensive data. These conclusions are based on study of a population where population structures have changed relatively little over time, see Section 3.2.2. In situations where population structures are changing rapidly measures which assess only mortality data may not be appropriate.

Calculating lifespan diversity using life tables and observed data is shown to produce qualitatively similar, but not identical, temporal trends across the study period, despite consistency in capturing yearly fluctuations. The slight differences in trends are caused by the effects of life table calculation on the distribution of ages at death. Measures of variability in lifespan are near-universally calculated using data from life tables and indeed, a number of advantages are conferred by the use of life tables (Murray et al., 2000). Period life tables create uniform starting populations to which the mortality pressures of real-world populations are applied. In effect, this has the benefit of countering biases relating to population structure and size between populations.

The shifts in the distribution of ages at death introduced by the calculation of life tables and the resultant effects on lifespan diversity pose the question: what is of interest in the calculation of these measures of variation in lifespan? The use of life tables provides, in theory, analysis of the distribution of death within constant populations. To use the example of lifespan diversity across Scotland as shown in Figure 5.1, analysis using life tables suggests a slight reduction in lifespan diversity from 2001 to 2019 despite an upwards trend from 2015 onwards. The distribution of ages at death in the life tables used to calculate this measure has narrowed

and individuals were likely to die at less varied ages. In contrast, when calculated from observed death counts, lifespan diversity was near stagnant in males and slightly increased in females from 2001 to 2019. This means that, in fact, females in Scotland were likely to die at slightly more varied ages in 2019 than in 2001; in apparent contradiction to the evidence from life tables.

In principle, life table results imply that had population structures been identical in 2001 and 2019, then females would have died at less varied ages in 2019. However, over time, population structures change and, as previously mentioned, results from diversity calculated using observed data show that in the true population, the opposite trends are evident. Studies measuring variation in age at death are often concerned with making comparisons and investigating inequalities between populations or subpopulations (Seaman et al., 2019; Brønnum-Hansen et al., 2021; Hiam et al., 2021). For this use, life tables certainly offer an advantage as they ensure theoretically equal study populations. Nonetheless, the use of measures of variation in lifespans as indicators in population health is founded upon the consequences of greater or lesser variation in age at death within a community. This may be at the level of the society, in implications for inequalities in health and the impacts of a greater range of ages at death on health care systems, or at the level of the individual, in effects on financial or health related decisions (van Raalte et al., 2013; Sasson, 2016) . The evidence presented in this chapter suggests that the use of life tables may mask underlying trends in variation in ages at death in the observed population. The second section of this chapter focuses on lifespan diversity calculated using life table data. This method was chosen because comparisons are made between subpopulations and this approach exploits a key advantage of the use of life tables. Further, it aligns with both the calculation of diversity in causes of mortality in Chapter 4 of this thesis, performed using life tables, and with existing literature in the field.

5.5.1.3 What were the temporal trends and cross-sectional tendencies in lifespan diversity in Scotland and Scottish subpopulations from 2001-2019?

Across Scotland, in both sexes, lifespan diversity computed from life tables fell from 2001 to 2015. After 2015, this reversed and a trend for increasing lifespan diversity is observed from 2015 to 2019. The increase in lifespan diversity after 2015 has counteracted much of the decrease from 2001 to 2015 in females but in males, a reduction between 2001 and 2019 is still evident.

In temporal analysis of subpopulations, trends in lifespan diversity are observed to be remarkably similar from 2001 to 2010, with mostly flat trends in females and slight reductions in males. After this, between 2010 and 2015, reductions in lifespan diversity are observed across subpopulations but are observed to be most acute in more deprived areas and urban areas. Finally, from 2015 to 2019, more divergence is observed in trends between subpopulations. Increases in lifespan diversity are confined to more deprived groups and urban groups in females with relatively little change in the least deprived areas and urban areas. Among males, lifespan diversity continued to fall from 2015 to 2019 in the least deprived areas while increasing in the most deprived areas.

When examining cross-sectional tendencies in lifespan diversity, both males and females in more deprived areas and urban areas are shown to have faced greater lifespan diversity than their counterparts in less deprived areas and rural areas. Further, deprivation-related inequalities are observed to have widened between the most and least deprived population fifths from 2001 to 2019.

Urban-rural differences in lifespan variation have not previously been examined in Scotland. A previous study in the USA has suggested lower variation in lifespans in metropolitan areas compared to all other areas combined (De Ramos et al., 2022). The findings of this chapter suggest the opposite tendency in Scotland, with greater diversity in lifespan observed in urban areas. Variation in lifespan is used as a measure of inequalities in health and this finding suggests greater inequalities in cities (large urban areas) than other areas of Scotland. This finding is consistent with what is already known about Scotland, and specifically that cities encompass areas with some of the highest and lowest life expectancies, representing significant heterogeneity (Seaman et al., 2015). The large differences in life expectancy across cities such as Glasgow have been linked to deprivation and socio-economic inequality within cities although, especially in Glasgow, such effects are not thought to fully explain inequalities in health (Walsh et al., 2010a).

5. Lifespan diversity and diversity in causes of mortality

Previous analysis of variation in lifespans across the socioeconomic gradient in Scotland by Seaman et al. (2019) showed consistently higher variation in more socioeconomically deprived areas from 1981 to 2011. These authors showed that these inequalities were at their highest in 2001 but that they reduced between 2001 and 2011. Trends in lifespan variation in Scotland have not previously been reported for the years following 2011. This chapter shows a similar reduction in inequalities between income deprivation quintiles in the 2000s and early 2010s. In the following years, however, the work presented here shows that this trend reversed and inequalities in lifespan diversity are observed to have grown from 2011 to 2019.

The implications of these findings are that inequalities in health outcomes have grown in the 2010s after a period of relative improvement in the 2000s. Increases in lifespan diversity across Scotland as a whole and increased disparities between certain subpopulations during the 2010s indicates greater inequality in health outcomes. Increasing variation in lifespan is not inherently considered to be a negative outcome in the literature as it may suggest that individuals are living to an older age. However, in Scotland, trends in life expectancy stalled in the late 2010s (Ramsay et al., 2020). This suggests that increases in lifespan diversity have been driven by increases in deaths at younger ages, rather than a significant proportion of individuals achieving extremely long lives.

Widening inequalities in a number of health outcomes have been reported in Scotland after 2015 and the findings of this section support this evidence of regression in progress towards health equality (McCartney et al., 2022; Miall et al., 2022). It is suggested in the literature that inequalities in health in Scotland may be becoming more intractable as, increasingly, social determinants of health are intricately patterned across the population. The increasing diversity in lifespan observed in more deprived quintiles, in this section and in previous work, supports this, suggesting that even in the most deprived areas, health outcomes have become more heterogeneous.

5.5.1.4 What is the relationship between diversity in causes of mortality and lifespan diversity in Scotland, firstly, across time and secondly, between subpopulations?

The relationship between diversity in causes of mortality and lifespan diversity was examined in this chapter to assess the relevance of two previously proposed hypotheses. First, that over time, increasing diversity in causes of mortality is caused by an ageing population with falling lifespan diversity. Secondly, that, cross-sectionally, in populations in which individuals die at a wider variety of ages, more diverse causes of mortality would occur.

As I discuss in Section 5.1.3, the two hypotheses examined in this chapter are based on different mechanisms for the relationship between lifespan diversity and diversity in causes of mortality. Under the first hypothesis, it is assumed that more diverse causes of mortality occur at older ages. As reductions in lifespan diversity are often associated with a greater proportion of the population reaching older ages, it is expected individuals would accumulate more varied risk factors and diseases over the lifecourse and therefore die of more diverse causes. This mechanism is supported by findings of a positive correlation over time between diversity in causes of mortality and life expectancy both in this chapter and in previous work (Bergeron-Boucher et al., 2020). The second mechanism assumes that a greater diversity of causes of mortality occurs when a wider range of ages at death occur due to the varied age-specific risks associated with different causes of death.

The findings of this chapter support both hypotheses, to an extent. I present evidence that over time, across Scotland and in each subpopulation, as lifespan diversity has fallen and life expectancy has increased, diversity in mortality causes has increased, consistent with the first hypothesis. I also show that, generally, cross-sectionally in subpopulations with greater lifespan diversity, greater diversity in mortality causes occurred at $q = 1$, but not for other values of q. As I note in Section 5.1.3, it might be expected that this second hypothesis would only be observed at less conservative measures of diversity, such as $q = 1$.

5. Lifespan diversity and diversity in causes of mortality

In supporting the two hypotheses raised in this chapter, the evidence I present highlights the logical inconsistency between them. It appears contradictory to expect increasing diversity in causes of mortality to be associated with falling lifespan diversity *as well as* higher diversity in causes of mortality being found in populations with higher lifespan diversity. This contradiction may not be irreconcilable: it is possible that the reasons for the temporal relationship observed here are distinct from the causes of the observed cross-sectional tendencies. However, it suggests that the mechanisms which I discuss in Section 5.1.3 are not wholly appropriate and may need refining.

In particular, I challenge the contention that an older population with less variation in age at death is inherently associated with greater diversity in causes of mortality. My findings support those of Bergeron-Boucher et al. (2020) finding that over time as populations lived longer and died at more homogeneous ages, they died due to more diverse causes. However, this previous study proposed that this increased diversity was caused by the population living to older ages. Instead, I suggest that, in Scotland, the observed increases in diversity in causes of mortality and life expectancy and the reductions in lifespan diversity have been driven by same the underpinning trend: a reduction in the prevalence of the most common causes of mortality, especially cardiovascular diseases. I have shown in Chapter 4 that reductions in the mortality rate and prevalence of cardiovascular disease were responsible for much of the increase in diversity in mortality causes from 2001 to 2019 in Scotland. Reductions in cardiovascular disease mortality rates and prevalence have occurred across all ages in these years, but were largest among those younger than 75 (Brown et al., 2019; Ramsay et al., 2020). This reduction in rate of death at younger ages has been shown to have contributed to both increasing life expectancy *and* decreasing lifespan variation in Scotland in the 21st century (Ramsay et al., 2020; Seaman, 2017). Lung cancer, another common cause in Scotland, also reduced in rate and prevalence in this period, again particularly at younger ages, further contributing to increases in mortality cause diversity and life expectancy and decreases in lifespan diversity. As premature deaths due to the most common causes have reduced, this has led to both longer and more homogeneous lifespans, but also to an increase in diversity in mortality causes through a reduction in the prevalence of the most common causes. This is further corroborated by the timing of increases in lifespan diversity in Scotland which I show in this chapter occurred after 2015. In the same years, 2015 to 2019, increases are observed in the prevalence and mortality rate of acute myocardial infarction which was the most common cardiovascular cause of death in both sexes throughout the

study period (discussed in Chapter 4). The results of this thesis cannot directly link this resurgence in cardiovascular disease to the increases in lifespan diversity, and this may be an valuable avenue of future research. However, the coincidence of the timing of these trends is remarkable.

The conclusion of Bergeron-Boucher et al. (2020), discussed above, could be consistent with the conclusion of the previous paragraph. As fewer premature deaths occur due to cardiovascular disease, deaths are redistributed to older ages and to more diverse causes. The subtle difference in arguments between what I propose here and the supposition of Bergeron-Boucher et al. (2020), is the causal agent. Rather than an increased age at death in the population[2] causing diversity in mortality causes to increase; increases in both life expectancy and diversity in causes of mortality and reductions in lifespan variation have been driven by reduced mortality rates, especially premature mortality rates, due to the most common causes. This is also supported by my findings in the previous chapter which show that that it is not necessarily true that causes of death are more diverse at older ages. Analysis in Chapter 4 (Figure 4.11) shows that diversity in causes of mortality at $q = 1$ was, in fact, lower among those aged 80+ than any other age group in both sexes in 2001. Furthermore, diversification in causes of mortality occurred in almost all age groups older than 40 in both sexes[3]. This is because not only have cardiovascular diseases historically been the most common causes of death at all ages 40+, but they have reduced in prevalence across all ages (Brown et al., 2019). These reductions have driven the increases in diversity observed across older age groups because the reduction in cardiovascular mortality was proportionally larger than the reduction in all-cause mortality. A way to confirm that changes in the prevalence of cardiovascular diseases have driven trends in both diversity in causes of mortality and lifespan diversity would be the use of decomposition analysis. This is likely a valuable avenue for future research.

Reframing this temporal mechanism for the relationship between diversity in causes of mortality and lifespan diversity makes it more consistent with the cross-sectional hypothesis explored in this chapter. If we accept that older, more homogeneous ages at death are not necessarily associated with more diverse causes of mortality, as I suggest in the previous paragraph, then it is conceivable that

[2]both on average, indicated by increased life expectancy, and proportionally, indicated by reduced lifespan diversity.

[3]The exception being deaths among women aged 60 to 79 among whom diversity at $q = \infty$ fell from 2001 to 2019. However, in this age group diversity in mortality causes increased under $q = 1$ and 2

in populations with greater lifespan diversity, more diverse causes of mortality occur. I show evidence for this relationship, at least under diversity at $q = 1$ in mortality causes. In capturing deaths at a wider range of ages, populations with greater lifespan diversity may capture a more diverse set of causes. However, I have presented evidence that over time, improvements in population-level health outcomes have been achieved in the same way, through relatively large reductions in the mortality rate of the most common causes, especially at premature ages.

5.5.2 Future areas of research

In order to more fully explore the drivers of trends in lifespan diversity noted in this chapter, the measure of additive value described in Chapter 1 could be applied to lifespan diversity. Measuring the additive value of each year of age would allow for an examination of which ages at death are driving changes in lifespan diversity across the population. Further to this, a possible avenue of research is the measurement of the additive value of each *cause of mortality* to lifespan diversity. In this case, positive additive value would indicate a cause of mortality was associated with the ages at which the population died most often. Through this analysis, it would be possible to determine which causes of death contribute to differences in lifespan diversity between populations. Through this method, it might be possible to test my hypothesis that the increases in lifespan diversity from 2015 in Scotland were driven by increasing cardiovascular mortality in these years.

I have suggested that lifespan diversity calculated from observed mortality data may be a reliable proxy for more established and developed measures of lifespan variation. However, it is important to note a caveat that I have only confirmed this in the study of a population is relatively close to a stationary population structure[4]. This measure is most likely to be useful in contexts in which appropriate and reliable population data is unavailable. These situations are likely to be rare but may include studies of historical data or of low-middle income nations. Further research should confirm how lifespan diversity, calculated from observed data, compares to established measures in populations which are in flux, for example experiencing epidemiological transition.

[4]Population pyramids in Chapter 3 show that the population of Scotland is close to but does not match the stationary population extracted from life tables.

In Section I show that my results add to the evidence that there may be a Simpson's paradox between life expectancy and diversity in causes of mortality. It has been shown both in my research and in the literature that diversity in causes of mortality and life expectancies have increased together in a number of settings in the 21st Century (Trias-Llimós & Permanyer, 2023; Bergeron-Boucher et al., 2020). However, it has also been shown it is often the case that in populations with the greatest life expectancies have the lowest diversity in causes of mortality. Further research may be warranted to tease out the cause of this inverse relationship.

Previous comparison of measures of variation in lifespan has examined the sensitivity of measures to changes in the distribution of ages at death (van Raalte et al., 2013). Performing similar analysis on the measure of lifespan diversity proposed in this chapter would give further insight into how this measure fits within the range of established measures.

5.5.3 Strengths and limitations

In Appendix B.1, the use of alternative measures of diversity (namely those which explicitly account for zero values in the distribution of mortality causes) is shown to have an effect on trends in mortality cause diversity across Scotland and in Scottish subpopulations. Overall diversity is found to be higher in 2019 than in 2001 in most cases - as found under normalised alpha diversity, the measure used in this chapter - however, trends during the study period differ under the alternative measure which accounts for zero values. The analysis of mortality cause diversity in this chapter was not repeated with diversity calculated under alternative measures of diversity. The use of this alternative measure may impact the conclusions drawn in this chapter. To test the robustness of the conclusions drawn here it may be useful in future research to replicate the analysis discussed in this chapter using a method for the measurement of mortality cause diversity which accounts for zero values in the distribution of mortality causes.

5.5.4 Research implications and next steps

Previous research has not examined variation in lifespans in Scotland since 2011. On average, the population of Scotland in 2019 died at an older age than in 2001 and this chapter has shown that there is less slightly variation around that average. However, in the 2010s, while life expectancies have stalled across Scotland, lifespan diversity has increased. This trend is more pronounced in more deprived areas. It is possible to use insights gained from the study of diversity in both lifespans and causes of mortality to learn about pressures on the healthcare system. As lifespan diversity decreases within a population, and people die with less variation in age at death, resources can be focused on health and end-of-life care for those at more homogeneous ages. In contrast, if the observed trends continue, then public health infrastructure should be prepared for individuals to need care at a wider variety of ages, even if life expectancies do not change dramatically.

Rising diversity in causes of mortality means that the healthcare system must diagnose and treat a more diverse range of diseases and conditions, potentially impacting on economies of scale in treatment. While the relationship between them is relatively weak, the findings confirm previous research that over time, increases in diversity in mortality causes have been associated with reductions diversity in age at death. This is shown to be consistent in findings across all ages in the Scottish population and in each subpopulation examined in this chapter. Therefore, potential gains in health economics from a more homogeneous age at death within the population may be counteracted by the necessity to remedy a greater variety of conditions. These combined implications for public health have not previously been explored widely but the results of this chapter, especially the adverse trends in the later years of the 2010s, suggest that variation in the distributions of health outcomes examined here require greater attention.

This thesis has, so far, addressed the diversity in mortality causes and in ages at death in the years 2001-2019. In the 2000s, trends consisted of relatively consistent increases in mortality cause diversity and reductions in the diversity of age at death, followed by more uncertain trends in both measures up to 2019. This study period ends directly before the beginning of the COVID-19 pandemic in Scotland.

This pandemic represented disruption to population health and daily life in Scotland that has been unprecedented in the modern era. The next chapter explores the application of diversity in mortality causes in the measurement of the impact of the COVID-19 pandemic on population health in Scotland.

Chapter 6

Diversity in causes of mortality in the study of the COVID-19 pandemic in Scotland

6.1 Background

The previous chapters of this thesis have presented research on diversity in causes of mortality as a measure of population health focusing on Scotland and Scottish subpopulations in the years 2001-2019. During this period, diversity in causes of mortality in deaths across all ages increased up to 2015 when trends diverge with a plateau in diversity among females and a reduction among men across Scotland as a whole and in all subpopulations. From 2001 to 2019 life expectancy in Scotland increased, although this was mainly driven by increases in the 2000s and in the 2010s stalls in life expectancy have been widely reported (National Records of Scotland, 2022a). Alongside this I have shown that lifespan diversity was slightly lower in Scotland in 2019 compared to 2001. In other words, at the end of the period, the population died at older ages and with less variation around this age at death, while facing more varied causes of mortality.

The 21st century has been a period of gradual changes in population health in Scotland, marked by steady improvements in various indices like life expectancy and mortality rates in the 2000s (Fenton et al., 2019a; Ramsay et al., 2020). However, these improvements slowed and even, in some cases, reversed in the 2010s (Fenton et al., 2019a; Fenton et al., 2019b). The research presented in this

thesis shows a similar change in the pace of diversification in mortality causes, with slower increases in diversity and more volatile year-to-year variation in the 2010s. This period of gradual changes in population health preceded the most acute public health crisis in recent history: the COVID-19 pandemic.

The first case of COVID-19 in Scotland was recorded on the 28th of February 2020 (Public Health Scotland, n.d.-a). This would be followed by the imposition of a national lock-down, which restricted the movements of the population, within a month, the first of a series of non-pharmaceutical interventions (NPIs) designed to limit disease spread (Hadjidemetriou et al., 2020). The COVID-19 pandemic in Scotland in 2020 and 2021 can be divided into three distinct periods of high prevalence: the "first wave" from March to June 2020; the "second wave" from October 2020 to February 2021; and finally, the "third wave" from late Summer 2021 onwards (Phin, 2021). The first two waves of COVID-19 infections were accompanied by significantly increased mortality in Scotland. Docherty et al. (2020) estimated excess mortality to be 68% higher in Scotland than expected by 6 May 2020[1]. The population began to be vaccinated in December 2020 and early 2021. The vaccines used in Scotland were effective in reducing mortality and this meant that despite a much higher number of confirmed cases of COVID-19 during the third wave than in previous waves, fewer deaths were recorded (Andrews et al., 2021; Bernal et al., 2021; Phin, 2021; Public Health Scotland, n.d.-c).

Significant societal effects were associated with the NPIs employed during the pandemic and because of the large number of patients hospitalised due to COVID-19, healthcare and public health systems faced huge stressors (Imai et al., 2020; Alfonso Viguria & Casamitjana, 2021). Together, these effects have been proposed to have had a substantial impact on population health. Concerns have been raised regarding increased mortality associated with a wide range of causes of mortality. Indeed, Docherty et al. (2020) suggest that only 73% of excess mortality in Scotland in the first wave of the pandemic was associated with COVID-19 deaths (Docherty et al., 2020).

[1]Note that there is debate in the literature regarding the utility of excess mortality as a measure of the burden of the pandemic, especially among populations with slowing improvements in mortality in the 2010s, such as Scotland (Corrao et al., 2021; Helleringer & Queiroz, 2022; Walkowiak & Walkowiak, 2022).

The literature links this increase in non-COVID-19 mortality to both direct and indirect effects of the pandemic (Figueroa et al., 2020b). Direct impacts indicate situations in which infection with COVID-19 hastened or contributed to death in patients. Indirect impacts refer to the possible negative spill-over effects of the pandemic on society and health care systems, for example increased difficulty in accessing health care, especially specialist care. In cardiovascular diseases, for example, co-morbid infection with COVID-19 has been shown to increase the risk of death in patients (Banerjee et al., 2021). However, compared to this direct effect, Wu et al. (2021) argue that in England and Wales, indirect effects, such as delays in seeking treatment, had a more significant impact on mortality. Compounding this, it has been suggested that disturbance in the supply and demand structures for acute cardiovascular care affected the provision of treatment during the pandemic in the UK (Banerjee et al., 2021). Studies have also indicated risks in cancer mortality due to disruption in screening and detection procedures during the pandemic both in Scotland and around the world (Campbell et al., 2021; Figueroa et al., 2021). Increased rates of death attributed to degenerative neurological diseases such as dementia were also recorded during the COVID-19 pandemic, especially during the first wave of infections. This has been closely linked to policies in the UK regarding the release of patients from hospitals to old age care homes (Burton et al., 2021; Paplikar et al., 2022). Increased mortality in care homes may have arisen from COVID-19 infection hastening death among some patients and it is possible some care home deaths in patients infected with COVID-19 were attributed to previous chronic conditions.

The studies noted above represent a small fraction of the literature published in the wake of the pandemic which attempt to quantify its impact on population health. These impacts are thought to affect various causes of mortality with potential consequences across the distribution of mortality causes. By considering the relative prevalence of each mortality cause, diversity in mortality causes is uniquely suited to examining the impact of the pandemic. Many deaths were attributed to COVID-19 during the years 2020 and 2021 and these deaths will have shifted the distribution of causes of mortality. Measuring diversity in mortality causes makes it possible to examine this shift while taking into account all other causes. If the pandemic led to an increase in mortality associated with common causes such as cardiovascular disease, cancers or degenerative diseases, without affecting rarer causes, a reduction in the diversity of causes would be expected, when examining the distribution of causes with COVID-19 deaths removed. Conversely, the pandemic may have caused the relative prevalence of rarer causes to increase, causing continued diversification of mortality causes once COVID-19

deaths are removed. A more uniform change across the distribution would have caused stagnation in the diversity of causes. It might be possible to be observe some of these trends through study of publicly available information. For example, Public Health Scotland data indicates that rates of mortality due to Lung cancer in 2020 and 2021 were similar to those observed over the past five years and that coronary heart disease mortality rates rose slightly in 2020 and 2021 (Public Health Scotland, 2022a, 2022b). This suggests the prevalence of the most common causes of death in Scotland, stagnated and potentially increased during the COVID-19 pandemic which might be expected to decrease diversity. By measuring diversity in mortality causes it is possible to assess overall trends which may be missed through analysis of cause-specific mortality rates. It may also be the case, given the pressure on healthcare systems during the pandemic that the precision of mortality coding changed throughout waves of the pandemic leading to less accurate recording of mortality causes. On the other hand the presence of a new, highly prevalent cause of mortality may have led to a wider range of diagnostic tests being performed for some symptomatic individuals, increasing diagnostic precision. These effects may impact the diversity of mortality causes during periods of high pressure on the healthcare system during the pandemic.

This chapter aims to apply measures of diversity in causes of mortality to the study of the impact of the COVID-19 pandemic on population health. The knowledge of these measures gained in previous chapters is used to estimate counterfactual scenarios based on previous trends. These scenarios are then used to assess the diversity of mortality causes during the years 2020 and 2021 both with COVID-19 deaths included and excluded from analysis. This analysis is performed across all ages and within twenty-year age groups as previously described in this thesis. This is especially important in analysis of the COVID-19 pandemic as not only were COVID-19 deaths more common in older ages, but the impact of direct and indirect effects of the pandemic on older age mortality has been highlighted in the literature (Douglas et al., 2020). Due to increased health risks older individuals are more likely to need to access the health care interventions that the pandemic restricted. Added to this at older individuals more commonly face chronic and multimorbid health conditions which may be exacerbated by COVID-19 infection (Salive, 2013).

The pressures of the pandemic on society and the healthcare system were varied throughout the years 2020 and 2021 with the times of greatest pressure being felt during the waves of infection outlined above. These periods of high prevalence of COVID-19 mortality would be expected to reduce diversity in mortality causes, as COVID-19 deaths represent a new and highly prevalent cause of mor-

tality. However, with COVID-19 deaths it is unclear how the pandemic may have affected the distribution of mortality causes. Above, I discuss studies which have proposed significant impacts on cardiovascular treatment in Scotland during the COVID-19 pandemic especially during the three waves of high prevalence of the virus (Banerjee et al., 2020; Wu et al., 2021). Cardiovascular diseases are among the most common causes of death in Scotland, were they to increase in mortality rate we might, again, expect to see reductions in mortality cause diversity in the distribution of mortality causes with COVID-19 mortality excluded. To study the impact of the pandemic during these periods of increased COVID-19 prevalence and mortality, this chapter presents diversity in causes of mortality by month of death rather than year. This also makes it possible for seasonality in diversity in mortality causes to be assessed. Seasonal trends in population health tend to be associated with increased older age mortality in Winter months driven by increases in cardiovascular and degenerative disease prevalence (Seretakis et al., 1997; Gemmell et al., 2000; Barnett et al., 2008; Liddell et al., 2016). This tends to lead to higher mortality rates among the most common causes of mortality in Winter. This might be expected to reduce diversity in mortality causes in Winter compared to summer months when mortality rates due to the most common causes of death are often lower. Assessing the seasonal aspects of diversity in mortality causes makes it possible to improve our understanding of pressures on the healthcare systems associated with wider ranges of prevalent diseases and conditions.

6.2 Research questions and objectives

The research presented in this chapter had the following objectives:

- To measure the diversity of causes of mortality in 2020 and 2021 in Scotland.
- To examine the impact of the COVID-19 pandemic on the diversity of mortality causes in Scotland in 2020 and 2021 across the population and at different ages.
- To investigate the seasonal dynamics of monthly diversity in mortality causes from 2001 to 2019.
- To produce forecasts of monthly diversity in mortality causes in 2020 and 2021 accounting for any seasonality in this measure.

- To assess the impact of peaks in pandemic prevalence on the distribution of mortality causes using forecasts of monthly diversity in mortality causes.

Using these objectives this chapter aims to answer the following research questions:

- What was the effect of the COVID-19 pandemic on diversity in mortality causes in Scotland across in the years 2020 and 2021?
- Does the diversity in causes of mortality in Scotland exhibit a seasonal pattern?
- What was the effect of the COVID-19 pandemic on diversity in mortality causes during the peaks of the COVID-19 pandemic in Scotland?
- Was the impact of the COVID-19 pandemic on diversity in mortality causes consistent between groups of causes and in deaths at different ages?

6.3 Methods

6.3.1 Data

Mortality data for the years 2001 to 2021 was extracted and processed as described in Section 3.2.1. Annual mid-year small area population estimates for 2001 to 2019 were obtained for Scotland as described in Section 3.2.2.

6.3.2 COVID-19 mortality

During the COVID-19 pandemic, the WHO released seven ICD-10 emergency use codes to classify morbidity and mortality due to infection with the novel coronavirus. Of these, two were used as 'underlying causes' of mortality for records of individuals who died in Scotland (Public Health Scotland, n.d.-b). These codes, their description, and the number of deaths in males and females in each year 2020 and 2021 associated with them are shown in Table 6.1.

Table 6.1: ICD-10 codes used to classify COVID-19 deaths in Scotland and the number of deaths recorded under them in males and females in 2020 and 2021.

		Number of deaths with each code			
Code	**Description**	**Males**		**Females**	
		2020	**2021**	**2020**	**2021**
U07.1	COVID-19 infection confirmed by laboratory testing	2237	2165	2534	2450
U07.2	COVID-19 infection diagnosed clinically or epidemiologically where laboratory confirmation is negative, inconclusive or not available/performed	761	42	633	32

Deaths associated with COVID-19 are classified in this chapter as any death with codes U07.1 or U07.2 recorded as an underlying cause of mortality in 2020 or 2021[2].

6.3.3 Normalised alpha diversity in causes of mortality

Normalised alpha diversity in causes of mortality is measured in this chapter as described in Section 3.3.3. In previous chapters, diversity in mortality causes has been measured using distributions of mortality causes extracted from life tables. The relative advantages of this approach have been explored elsewhere in this thesis. A key limitation of the use of life tables is their instability at small population sizes. In this chapter, the diversity of mortality causes among deaths within individual months is measured. These represent very small numbers of observations over which to calculate life tables (Huang et al., 2020a). Fortunately, the measurement of diversity in mortality causes has been shown to be insensitive to the use of observed data (raw counts of mortality attributed to each cause of death) as compared to life table data in Section 3.3.3. Therefore, throughout this chapter diversity in mortality causes is measured using observed mortality data.

[2]Alongside these codes, Public Health Scotland (PHS) suggest 2 others which might have been used to record COVID-19: B34.2 "The term 'coronavirus' without any further specification or other specified types of coronavirus, that is, not COVID-19 coronavirus" and B97.2 "The term 'coronavirus' without any further specification (as the cause of other diseases) or other specified types of coronavirus, that is, not COVID-19 coronavirus (as the cause of other diseases)". Unlike U07 codes these were extant prior to the COVID-19 pandemic and may be listed for a death due to any coronavirus not specifically COVID-19. Codes B34.2 and B97.2 are not included in the definition of COVID-19 mortality in this chapter under PHS recommendation that deaths recorded under them are excluded from analysis.

Normalised alpha diversity at $q = 1$, $q = 2$ and $q = \infty$ was measured in the causes of mortality in each year from 2001 to 2019 for deaths across the Scottish population and within the following age groups: 0 to 19, 20 to 39, 40 to 59, 60 to 79 and 80+. Diversity was measured separately for males and females in each of these groups. These calculations were performed for mortality across all causes and for mortality in all causes of mortality excluding COVID-19 codes as described in Section 6.3.2.

6.3.3.1 Monthly diversity in causes of mortality

Normalised alpha diversity at $q = 1$, $q = 2$ and $q = \infty$ were also applied to causes of mortality grouped by month of death in each month from January 2001 to December 2021. Diversity was measured separately for males and females in mortality across all ages and in each twenty-year age group discussed above for all causes and in all causes of mortality excluding COVID-19. The diversity of causes within selected ICD-10 chapters was also measured for deaths across all ages in males and females separately in these months. These chapters included those determined through additive value analysis to have a noteworthy impact on all-cause diversity in Results Chapter 4: II: Neoplasms; V: Mental and behavioural disorders; VI: Diseases of the nervous system; IX: Diseases of the circulatory system; X: Diseases of the respiratory system; and XX: External causes of morbidity and mortality. Chapter I: Certain infectious and parasitic diseases was included to assess any spillover effects of COVID-19 interventions on mortality associated with other communicable diseases.

To assess the dynamics of seasonality in monthly diversity in mortality causes, a rolling year-wide measure of diversity was calculated. Mortality records from a 12-month period, with each month in the years 2001 to 2019 as a centre-point were collated. For example, for January 2001, diversity is calculated using all mortality records in Scotland from August 2000 to July 2001. The diversity in causes of mortality in each of these 12 months was then calculated in males and females across the Scottish population. A comparison between the trends in this measure and trends in diversity within each year is shown in Figure 6.1.

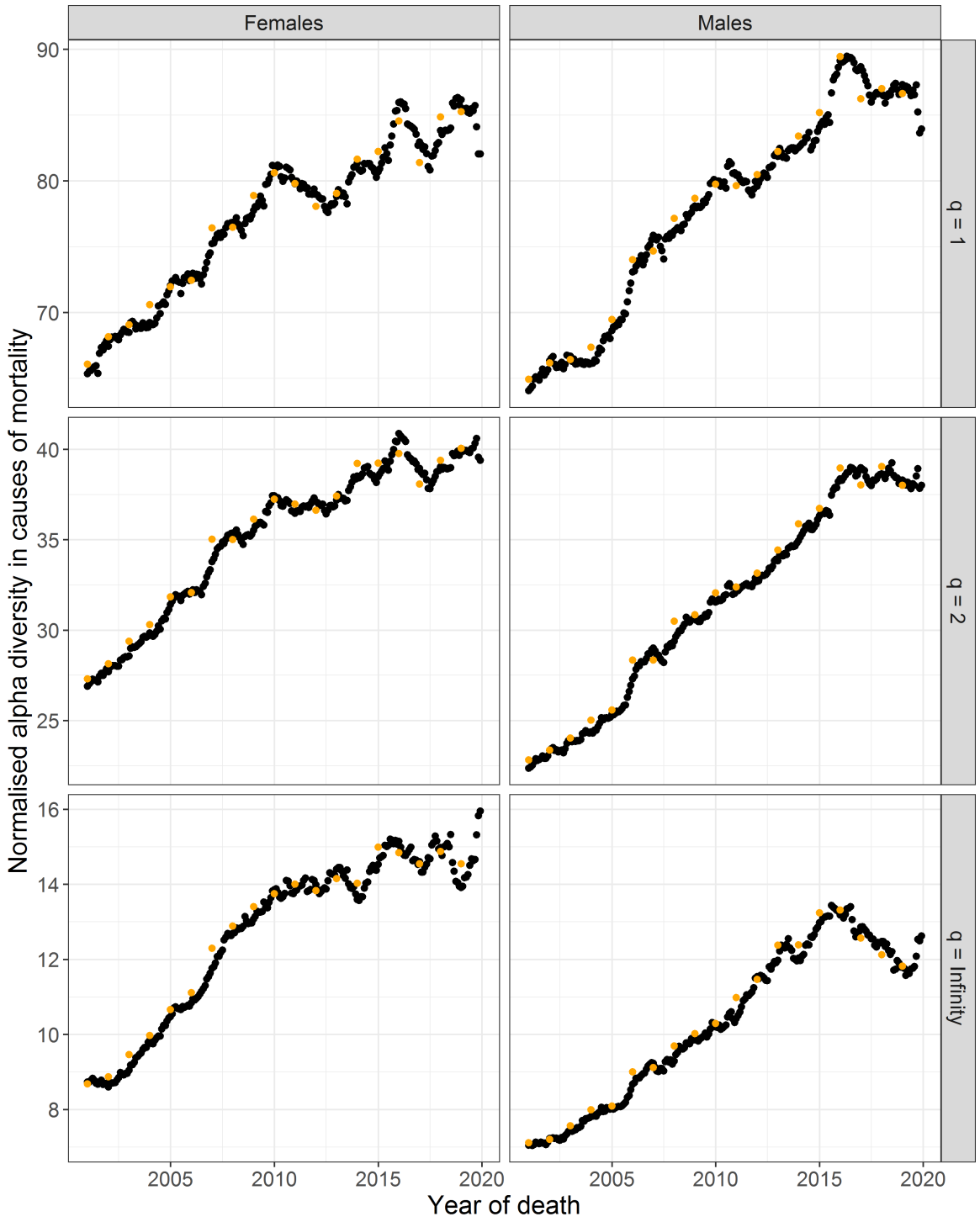


Figure 6.1: Comparison of yearly measurements of diversity in causes of mortality (orange points) at $q = 1$, q= 2 and $q = \infty$ with a rolling yearly measure (black points) used to assess seasonality. N.B. The years 2020 and 2021 are excluded from this analysis; therefore, in the final 5 months of 2019, diversity is not measured across a full 12-month period and this is likely associated with extreme points in this period in this figure.

6.3.4 Age-standardised mortality rates

All-cause and cause specific age-standardised mortality rates (ASMRs) were calculated as described in Section 3.3.2 for Scotland in each year from 2001 to 2021. Both all-cause and cause-specific (at the level of individual ICD-10 three character codes) ASMRs were calculated for deaths across all ages and in each twenty-year age group.

6.3.5 Additive value

The additive value of each ICD-10 chapter to diversity in mortality causes was calculated as described in Section 4.3.6. This calculation was performed on observed mortality data rather than life tables as discussed previously. Additive value was calculated in each ICD-10 Chapter for males and females separately in each year from 2001 to 2021.

6.3.6 Statistical modelling

Linear models were fit to the relationship between diversity in mortality causes and year of death in the years 2001 to 2019. Models were fit separately for males and females at each q value discussed in Section 6.3.3 for mortality across all ages. However, as has been noted previously (Chapter 4), trends in the diversity of mortality causes in Scotland are not consistently linear from 2001 to 2019; instead, a change in yearly trends occurs in the 2010s. To assess this change in trends, segmented regression analysis was applied to the same data for which the previously described linear models were fit. Across measures of diversity, a breakpoint was estimated in 2009 for females and 2016 for males; this is discussed further in Section 6.4.1. Both linear and segmented regression models of the relationship between diversity in mortality causes and year of death were extended to predict diversity in mortality causes in the years 2020 and 2021. Table 6.2 shows the R^2

values of the linear and segmented regression models. For all values of q in both sexes, segmented regression models a greater degree of the variation in diversity in mortality causes was explained by the year of death, but that values differed only slightly.

Linear models of the relationship between diversity in mortality causes and ASMRs were also fit for males and females separately across all ages. These models were extended to predict diversity in mortality causes given ASMR in 2020 and 2021 both with and without COVID-19 deaths included in analysis.

Table 6.2: R^2 values of linear and segmented regression models of the relationship between diversity in mortality causes and year of death among males and females in Scotland.

Diversity q value	R^2 (Adjusted)
Females	
Linear	
1	0.9171
2	0.8991
Infinity	0.8729
Segmented	
1	0.9565
2	0.9741
Infinity	0.8991
Males	
Linear	
1	0.9463
2	0.9767
Infinity	0.8914
Segmented	
1	0.9712
2	0.9882
Infinity	0.9862

6.3.7 ARIMA models

Autoregressive integrated moving average (ARIMA) models were constructed to forecast monthly diversity in mortality causes in 2020 and 2021 based on trends from 2001 to 2019. ARIMA models were chosen because they are adaptable to seasonal and non-seasonal data and because they are suited to non-stationary data. Models were fit following the methodology proposed in Schaffer et al (2021) described in Section 3.3.5. Seasonality was determined by applying ARIMA models with seasonal and non-seasonal components separately to diversity in mortality causes.

Separate models were constructed for diversity in mortality causes at q values of 1, 2 and ∞ in males and females across all ages and within each twenty-year age group described in Section 6.3.3. Models were also fit to the monthly diversity in mortality causes among deaths across all ages in males and females separately for selected ICD-10 chapters described in Section 6.3.3.

The parameters for each individual ARIMA model are listed in Appendix Table A.13.

6.3.8 Software

Analysis in this chapter was performed in *R* version 4.1.3 (R Core Team, 2022). Data manipulation and processing was performed using version 1.3.1 of the *tidyverse* package, plots were created using version 3.3.5 of the *ggplot2* package alongside version 1.1-2 of the package *RColorBrewer* (Neuwirth, 2014; Wickham, 2016; Wickham et al., 2019). Calculation of diversity was performed using version 2.0 of the *rdiversity* package (Mitchell et al., 2020). For segmented regression analysis version 1.6-0 of the *segmented* package was used (Muggeo, 2008). Time series analysis (including ARIMA models) was performed using the packages *zoo* (Ver. 1.8-10) and *forecast* (Ver. 8.17-0) (Zeileis & Grothendieck, 2005; Hyndman et al., 2022; Hyndman & Khandakar, 2008).

6.4 Results

The results section of this chapter is structured as follows. In section 6.4.1 I measure yearly diversity in causes of mortality in 2020 and 2021 to examine the impact of the COVID-19 pandemic on trends in this measure. In this section I analyse diversity of mortality causes across the population in deaths across all ages combined, analysing separately deaths among males and females, and measuring diversity at $q = 1$, $q = 2$, and $q = \infty$. I use linear models and segmented regression analysis to examine trends in the relationship between diversity in mortality causes and year at death and linear models to assess the relationship between diversity in mortality causes and ASMRs. I also examine trends in the diversity of mortality cause by year of death in deaths among twenty-year age groups to assess the effect of the pandemic on these groups.

Next, in Section 6.4.2, I examine the impact of the pandemic on diversity in mortality causes in each month in 2020 and 2021. Section 6.4.2.1 begins with analysis of seasonality in monthly diversity in mortality causes using ARIMA models in the years 2001 to 2019 in males and females separately across all ages. I then extend this to deaths among distinct age groups and ICD-10 Chapters in Section 6.4.2.2. In Section 6.4.2.3, I use these ARIMA models to forecast diversity in mortality causes in each month in 2020 and 2021. I then compare the observed diversity in mortality causes across all causes and in all causes excluding COVID-19 to these forecasts.

6.4.1 The COVID-19 Pandemic and trends in diversity in mortality causes

In 2020, 6165 deaths occurred in Scotland with COVID-19 recorded as the underlying cause of mortality; in 2021, this reduced to 4689. In both years, deaths associated with COVID-19 were responsible for the highest cause-specific ASMR of any individual three-character ICD-10 mortality cause in both males and females. Figure 6.2 compares the ASMR of the leading causes of mortality in males and females in

Scotland between 2016 and 2021. In 2020, these rates dwarfed even those from the second most common cause: acute myocardial infarction in men, and lung cancer in women. COVID-19 continued to have the highest ASMR of any cause of mortality in 2021.

Figure 6.2: Cause-specific ASMRs of leading causes of mortality across the Scottish population in the years 2016 to 2021 in males and females.

It is reasonable to expect the introduction of this novel and dominant cause of mortality to have had a significant impact on the diversity of the distribution of mortality causes faced by the population. The large proportion of deaths associated with a common cause would be expected to reduce diversity as it represents a reduction in the variety of causes each individual was likely to face. The trends over time in mortality cause diversity from 2001 to 2019 have been examined in Chapter 4. To examine the effect of the COVID-19 pandemic on the distribution of causes of mortality, in Figure 6.3, linear models of these trends are extended to the years 2020 and 2021. Diversity in mortality causes in the years 2020 and 2021 for all-causes and for all-causes excluding COVID-19 mortality are plotted as different shapes.

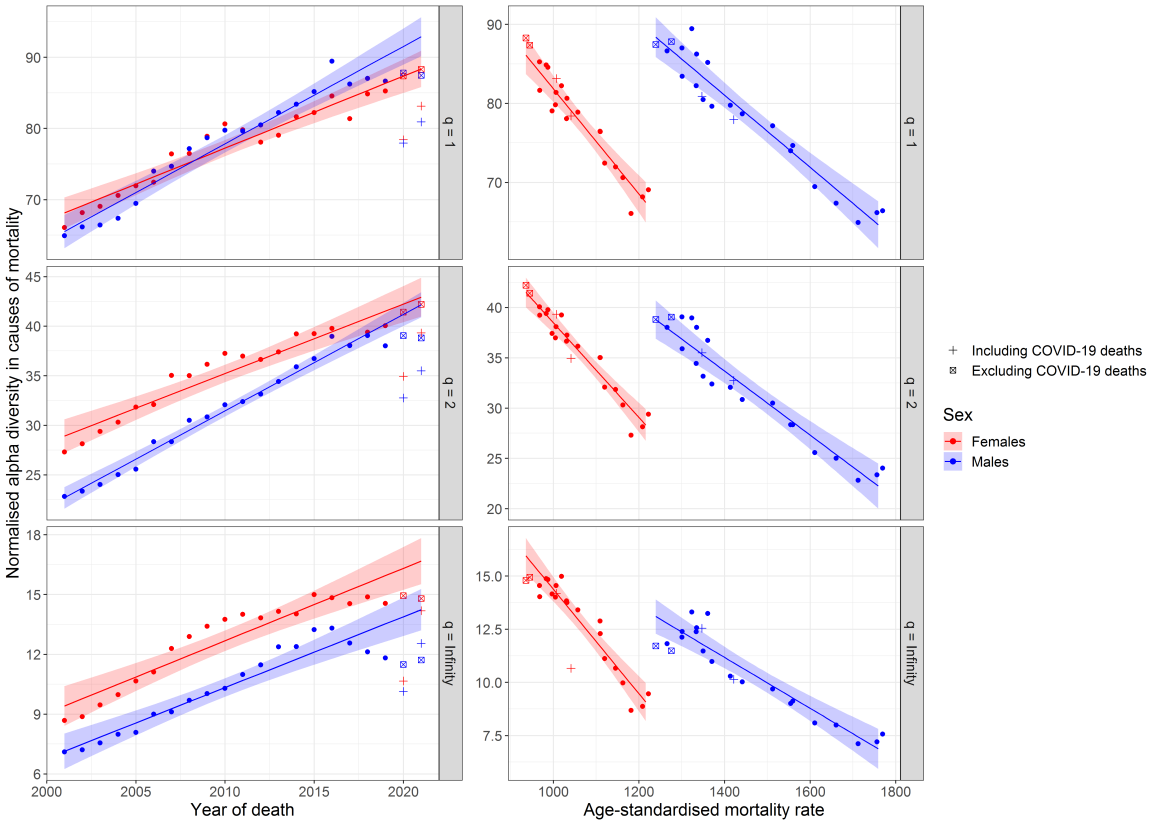


Figure 6.3: Diversity in mortality causes in males and females in Scotland plotted against year of death and age-standardised mortality rate. Solid lines and confidence intervals represent linear model estimates on 2001 to 2019 data extended to encompass the years 2020 and 2021.

Observing the crosses in the left-hand column of Figure 6.3, and as expected during the COVID-19 pandemic, all-cause diversity in mortality causes was markedly reduced due to the large proportion of deaths associated with COVID-19. As noted above, this dominance meant that variation in causes of mortality within the population decreased. Since around 1 in 10 deaths occurred due to COVID-19, the probability that a death would occur from any other cause decreased.

The reduction in diversity observed here is what would be expected given the introduction of a novel and dominant cause of mortality. It is, however, more difficult to predict the effect of the pandemic on the remaining causes of mortality within the population. Several studies have noted changes in the prevalence of mortality causes during the pandemic, and there has been a great deal of scientific

interest in excess mortality and its sources during the period (Docherty et al., 2020; Banerjee et al., 2021; Wu et al., 2021). In Figure 6.3, indicative linear models of the trend in diversity in mortality causes over the years 2001 to 2019 are plotted and extended to 2020 and 2021.

Among females, under diversity at $q = 1$ and $q = 2$, diversity in mortality causes in the years 2020 and 2021 excluding COVID-19 falls within the confidence intervals of these models (upper and middle panels of Figure 6.3). However, for males for all q values, and for females for q=infinity, diversity in causes of mortality excluding COVID-19 deaths in these years is found to be outwith model projections. Despite failing to align with these projections, it can be observed in Figure 6.3 that values of mortality cause diversity in 2020 and 2021 with COVID-19 deaths excluded were consistent with those in previous years. It is clear visually in Figure 6.3 that in the latter half of the 2010s the plotted linear models poorly fit the data in certain cases. This is most obvious at $q = \infty$ (lower panel of Figure 6.3) in both sexes but can also be observed in males under $q = 1$ and $q = 2$. This deviation from previous trends at $q = \infty$ is discussed in Chapter 4. To assess whether a change in trends occurred during this time segmented regression analysis was performed on the relationship between diversity in mortality causes and year of death in the years 2001 to 2019. This analysis was carried out separately for males and females under diversity at $q = 1$, 2, and ∞.

Segmented regression analysis of the relationship between diversity in mortality causes and year of death suggested a single breakpoint in temporal trends in each measure of diversity in both males and females. In males, this breakpoint was in the year 2016, with a trend for falling diversity shown in the following years under $q = \infty$. Under diversity at $q = 1$ and $q = 2$, the second section of the segmented regression models shown in the upper and middle panels of Figure 6.4 also suggests a possible trend for reducing diversity in mortality causes from 2016 to 2019 in males. Although under these less conservative measures, this trend is not found to be statistically distinguishable from a flat trend or slightly increasing diversity in this period. Among females, segmented regression analysis suggests a breakpoint in 2009 across q values; in the years after 2009, it suggests that diversity in mortality causes continued to increase, albeit at a slower rate.

Projections into 2020 and 2021 were performed using segmented regression models for males and females under each diversity measure discussed above. With COVID-19 mortality excluded, diversity in mortality causes in 2020 and 2021 were found to fall within the 99% confidence intervals of these projections in all cases except in deaths among females in 2021 under diversity at $q = 2$ where diversity was found to be fractionally higher than the expected range (39.1 to 42.1) at 42.2. This small difference and the fact that this data point falls within the projected confidence intervals of a comparably well-fit linear model in Figure 4 implies that it can be assumed diversity in this year was consistent with previous trends. Therefore, the analysis presented in Figures 6.3 and 6.4 indicates that with COVID-19 mortality excluded trends in the diversity of mortality causes in Scotland changed little during the pandemic. It should be noted that the second portion of the segmented regression models in males uses data from only 4 years. Confidence intervals for males are larger than among females due to the increased uncertainty introduced by this small range of datapoints.

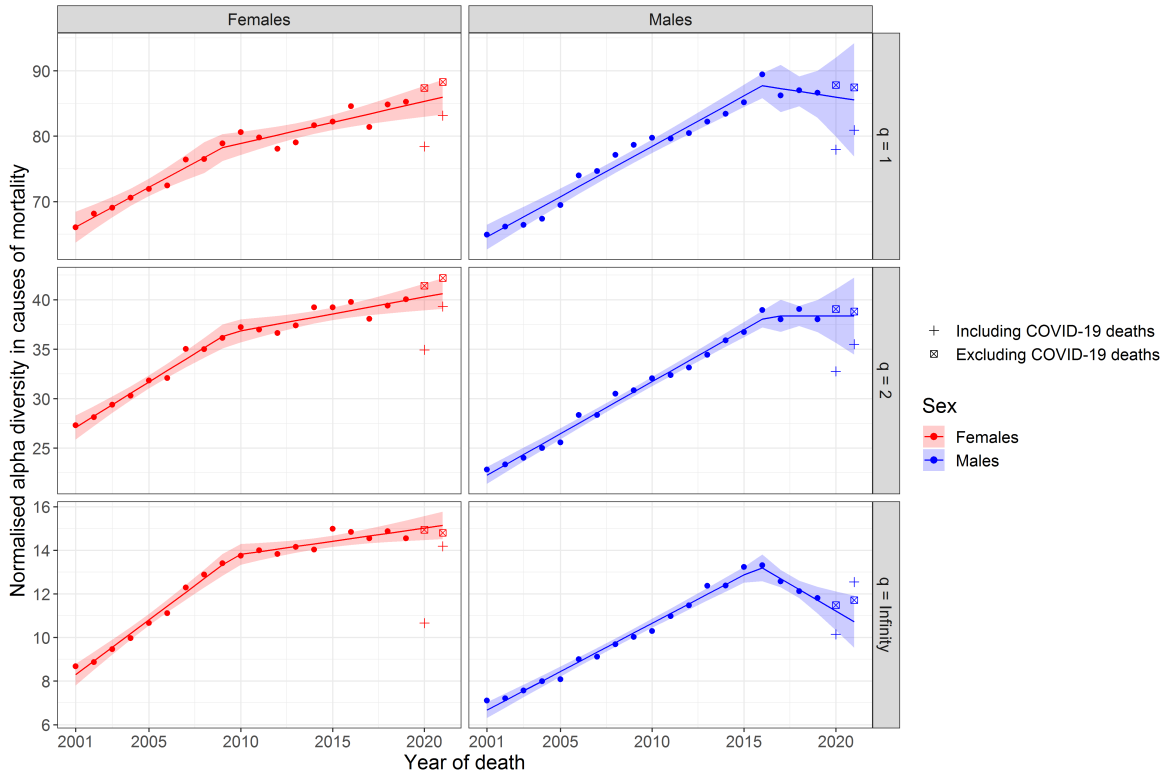


Figure 6.4: Trends in the diversity of mortality causes at q values of 1, 2, and ∞ in males and females in Scotland. Segmented regression lines are plotted, with breakpoints at 2016 across measures in males and 2009 in females. Diversity in mortality causes in the years 2020 and 2021 are plotted in separate shapes with COVID-19 mortality included and excluded from analysis.

The relationship between diversity in mortality causes and ASMR in 2020 and 2021 is also consistent with previous years as shown in the right-hand column of Figure 6.3. Mortality rates excluding deaths associated with COVID-19 were lower than in previous years, though within the range suggested by previous trends. Under diversity at $q = 1$ and $q = 2$, the continuation of trends in diversity meant that mortality in 2020 and 2021 fit within the relationship as modelled in the years 2001 to 2019 (upper and middle panels of Figure 6.3). All-cause mortality rates (i.e. including COVID-19 mortality) in 2020 and 2021 were similar to those in the early 2010s in Scotland, an increase compared to later years of the decade. Notably, the reduction in diversity brought about by deaths associated with COVID-19 meant that diversity in mortality causes in 2020 and 2021 was also comparable to the early 2010s.

The additive value of ICD-10 chapters to diversity in mortality causes was also measured in 2020 and 2021 as in Chapter 4. Across Chapters, little difference to the trends noted previously in this thesis are found. Plots comparing this additive value with previous years can be found in Appendix Table A.13.

Across the distribution of mortality causes, evidence presented in this section suggests that during the COVID-19 pandemic, previous trends in the diversity of other mortality causes continued for the most part unchanged. This is despite the prevalence of COVID-19 infections, non-pharmaceutical interventions with significant social effects and extreme strain placed on the healthcare system. However, the pressures of the COVID-19 pandemic were not felt evenly throughout the population. At older ages, death after COVID-19 infection was much more common. Furthermore, some governmental and social responses to the pandemic, such as care homes and hospital discharge policies, affected the elderly more severely. Figure 6.5 examines the trends in diversity at different ages.

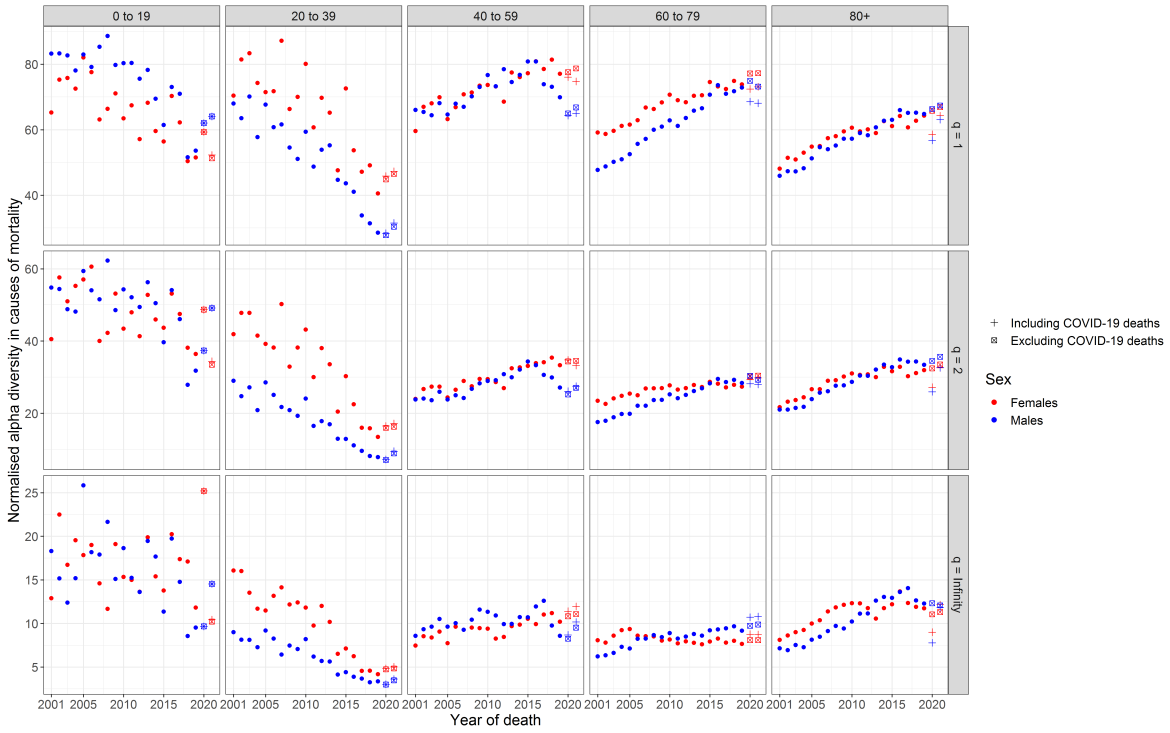


Figure 6.5: Diversity in mortality causes in males and females in deaths within twenty-year age groups plotted against year of death from 2001 to 2021.

In most age groups shown in Figure 6.5 , trends in the diversity mortality causes continued into 2020 and 2021 when COVID-19 is excluded. COVID-19 mortality had the greatest effect on diversity in mortality causes in deaths among those aged 80+ (rightmost column of Figure 6.5). However, even in deaths among this age range, changes to the distribution of all other mortality causes were consistent with previous trends. The same is true for deaths among males aged 40 to 59, where a reduction in diversity during 2020 and 2021 with COVID-19 excluded is not separable from a fall in diversity in 2017 to 2019 (central column of Figure 6.5). Regression lines were not plotted as they were not informative beyond observable trends.

In females aged 0 to 19, diversity in mortality causes at $q = \infty$ was markedly higher than in previous years in 2020 (bottom left panel of Figure 6.5). This marks the only major exception to previous trends in deaths among any age group. Diversity at $q = \infty$ reflects only changes at the level of the most common causes. Appendix Table A.15 shows the mortality rates associated with the most common causes of mortality in each age group from 2016 to 2021. Most of the leading causes of death among females 0 to 19 reduced in prevalence from 2019 to 2020. The largest reduction among females aged 0 to 19 was associated with intentional

self-harm by hanging, strangulation and suffocation (ICD-10 code X70) which fell from an ASMR of 2 per 100,000 to 0.9 per 100,000. As shown in Figure 6.5, diversity in mortality causes fell from 2020 to 2021 among deaths in this group returning to previous levels. Analysis of cause-specific rates showed that this was not driven by an increase in deaths associated with intentional self-harm by hanging, strangulation and suffocation which remained at an ASMR of 0.9 per 100,000 in 2021. Rather, increases are observed in the prevalence of several causes of mortality associated with the perinatal period and new-born infants. Notably, the rate of deaths attributed to sudden infant death syndrome (ICD-10 code R95) more than doubled from 2020 to 2021 from 0.9 per 100,000 to 1.9 per 100,000 and the mortality rate associated with ICD-10 code P07 "Disorders related to short gestation and low birth weight, not elsewhere classified" increased almost 3-fold from 0.9 per 100,000 to 2.6 per 100,000. The mortality rates associated with these conditions tend to be low and fluctuate between years with no clear trend; however, the 2021 figures are considerably higher than in previous years.

Overall, in the years 2020 and 2021, this section indicates that, with COVID-19 deaths excluded, diversity in causes of mortality in Scotland was consistent with trends from 2001 to 2019. This is found consistently in most age groups and across ICD-10 chapters.

6.4.2 Monthly diversity in mortality causes: Seasonality and the COVID-19 pandemic

In Section 6.4.1, evidence is presented suggesting that the COVID-19 pandemic and associated disruption had little spillover effect on variation in the distribution of non-COVID mortality causes in the years 2020 and 2021. However, the pressures on the population and the healthcare system in Scotland were not constant across these years. Significant peaks in mortality associated with COVID-19 in the spring of 2020 and from October 2020 to January 2021 tested the capacity of the healthcare system and were addressed by increasingly stringent NPIs (Douglas et al., 2020). This section explores diversity in mortality causes in each month to examine the effect of these increases in the prevalence of COVID-19 on the health of the population.

6.4.2.1 Accounting for seasonality in mortality cause diversity

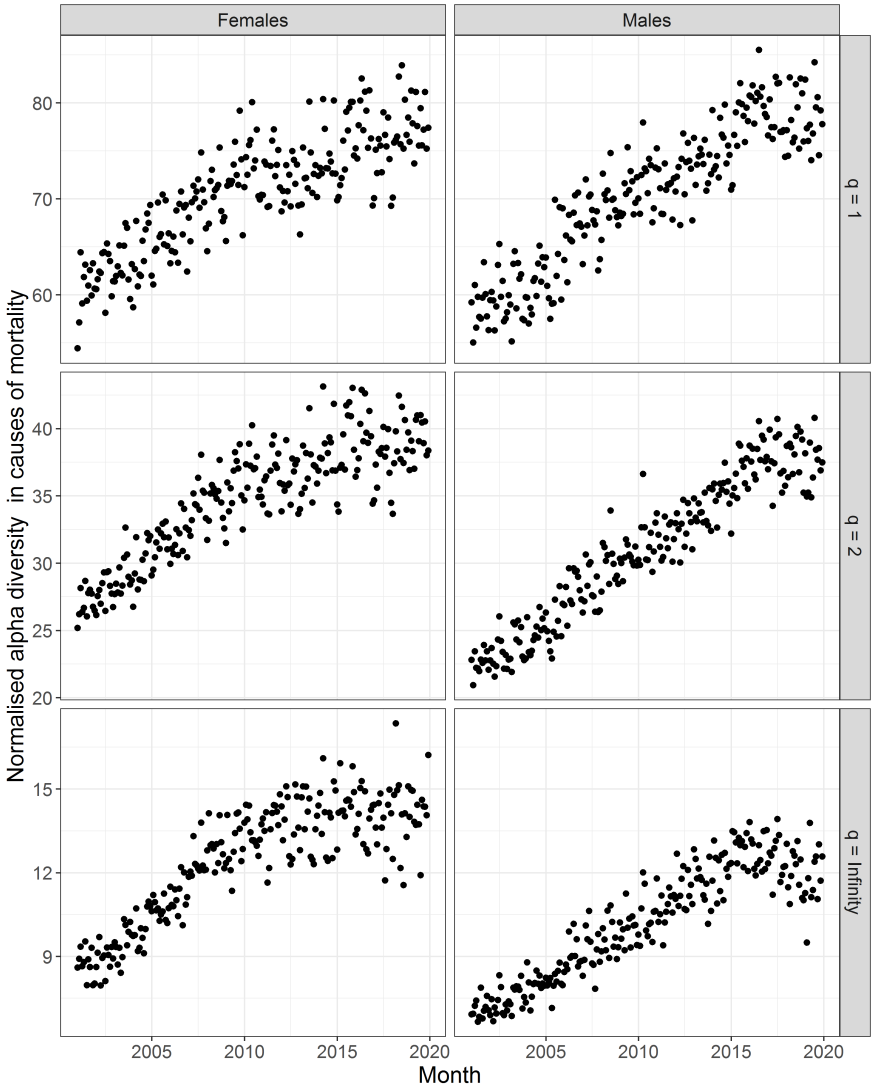


Figure 6.6: Diversity in mortality causes in each month from January 2001 to December 2019 plotted separately in males and females.

Previous studies of temporal trends in diversity in mortality causes have presented evidence with years as a temporal unit (Izsák, 1986; Bergeron-Boucher et al., 2020; Trias-Llimós & Permanyer, 2023). To examine the diversity of mortality causes within years it is important to consider the possibility of seasonal patterns. This is because, as mentioned previously, mortality patterns are known to vary significantly by season. In Winter months, rates of death are higher, driven by increased mortality among the elderly and associated with specific causes of mortality, especially infectious diseases such as influenza and associated conditions such as pneumonia, as well as cardiovascular and degenerative diseases. The regularly changing prevalence of these causes of mortality is likely to be associated with concurrent

fluctuations in the diversity of mortality causes. In Figure 6.6, the diversity of mortality causes is shown in each month from January 2001 to December 2019 for both males and females in Scotland. A general trend of increasing diversity, as noted in Section 6.4.1, can be observed across the results in each month. However, variation from month to month in this measure is clearly evident, suggesting shifts in the distribution of causes of mortality faced by the population within each year.

To test for seasonality in mortality cause diversity, these data were examined using ARIMA models as described in Section 6.3.7 ARIMA models with a seasonal component were used to examine monthly diversity in males and females across Scotland separately. The results of this analysis provide sufficient evidence to suggest seasonality in monthly diversity in mortality causes. Individual model parameters can be found in Appendix Table A.13.

This seasonality presents as increased diversity in mortality causes in Summer months and reduced diversity in Winter under diversity at $q = 1$ and $q = 2$ (upper and middle Figure 6.7). To demonstrate this, diversity was measured across a 12-month rolling period with each month as the centre point as described in Section 6.3.3. This 12-month diversity measure was subtracted from the observed diversity in mortality causes in each month (as shown in Figure 6.6). This produces a measure of the difference in diversity in each month from the year around it. The distribution of this measure is shown in Figure 6.7. There are two possible explanations for increased diversity in mortality causes in Summer months either: common causes have an increased prevalence in Winter, reducing diversity in those months; or, for reasons which are unclear, an increased variety of causes occurs in Summer. The first explanation is what might be expected given the previous understanding of seasonality in mortality in the literature(Liddell et al., 2016; Barnett et al., 2008). As common causes such as cardiovascular diseases and degenerative diseases such as dementia and Alzheimer's disease increase in prevalence in Winter, a reduction in diversity would be expected. However, it is questionable whether a seasonal effect exists under diversity at $q = \infty$ in the analysis displayed in the lower panel of Figure 6.7[3]. This measure would be most sensitive to changes in the prevalence of the most common causes. This counters the argument made in the previous para-

[3]Despite this seasonality was suggested by the ARIMA analysis for diversity in mortality causes under diversity at $q = \infty$ in both sexes.

graph changes in the prevalence of the most common causes would be expected to manifest as changes in diversity at $q = \infty$. The small seasonal pattern in diversity at this measure indicates that instead that the variety of rare causes changes more significantly over the year, supporting the second explanation raised above.

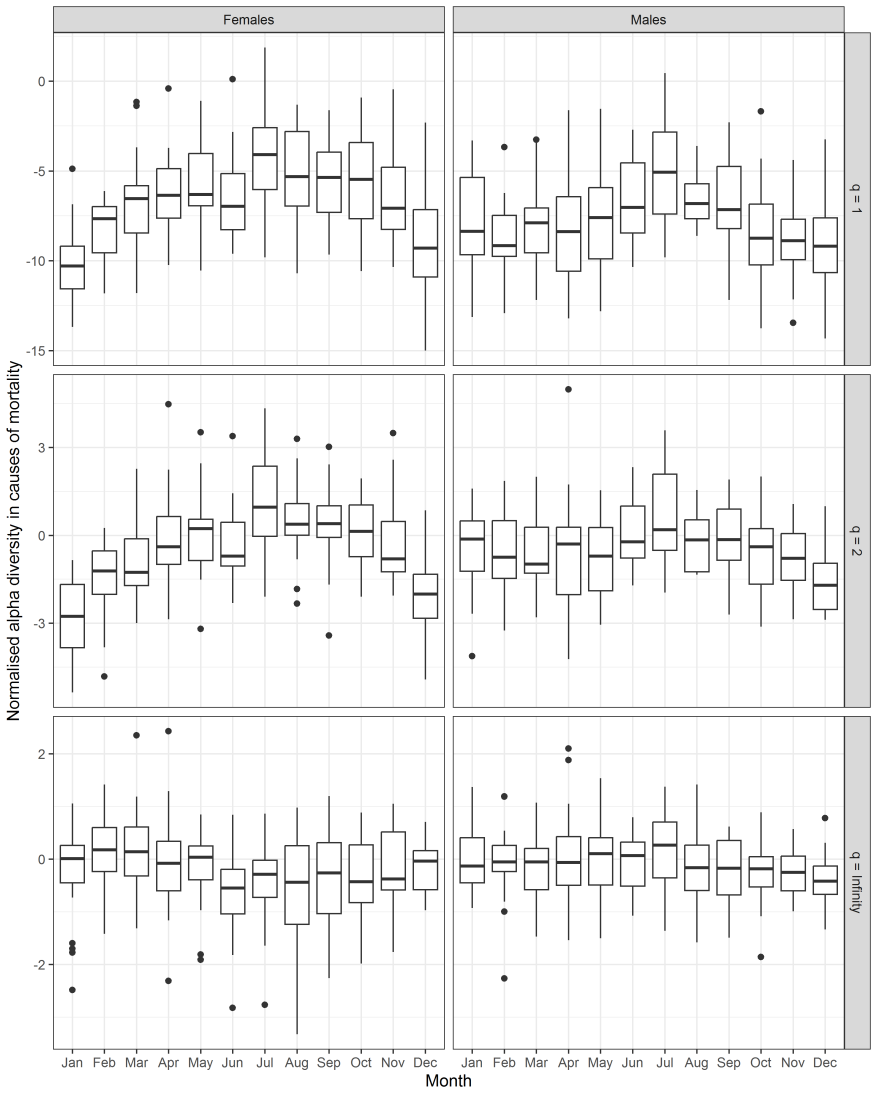


Figure 6.7: Boxplot of the differences between monthly diversity in mortality causes at $q = 1$, $q = 2$ and $q = \infty$ and rolling year diversity in each month from January 2001 to December 2019 showing seasonality in mortality cause diversity.

6.4.2.2 Forecasting by age group and ICD-10 Chapter

Diversity in mortality causes in each month was also tested within the selected ICD-10 chapters described in Section 6.3.3. Seasonality was observed in most Chapters at $q = 1$ but occurred in a smaller number of chapters under other measures of diversity. This suggests a degree of seasonal change in variation in relatively rare

causes of mortality within these Chapters, while the prevalence of more common causes of mortality shows less seasonal variation, similar to that noted across all causes above. In most Chapters models were fit using data from each month from 2001 to 2019; however, in several Chapters, ARIMA models did not converge over this period. These were Chapter VI: Diseases of the nervous system in both sexes and Chapter II: Neoplasms in females. Examining the diversity of causes within Chapter IV over the years 2001 to 2019 reveals a significant change in trends in diversity in the early 2010s coinciding with a change in coding practice mentioned previously (Section 3.2.1) (National Records of Scotland, 2011). This change in coding practice led to an increase in the number of deaths attributed to causes in this chapter, specifically Alzheimer disease (ICD-10 cause G30). For this ICD-10 Chapter, models were fit to data from the years 2010 to 2019 and under this method, it was possible to fit a model. No such change in trends is observed in the diversity within Chapter II in females; however, limiting the ARIMA model to the years 2010 to 2019 also allowed for the model to be fit, therefore this method was used. Further details of ARIMA models used these chapters, can be found in Appendix Table A.13.

Seasonality in the diversity of mortality causes at q values of 1, 2, and ∞ was also tested for deaths in each of the twenty-year age groups described in Chapter 4. The methodology for fitting ARIMA models and testing for seasonality as explained above was followed for each age group. In most age groups, this analysis suggested seasonality in monthly diversity cause mortality. Non-seasonal ARIMA models were fit in age groups where no seasonality was found; Appendix Table A.13 shows the parameters and any conditions on the ARIMA models. No systematic patterns in non-seasonal cases is found. All models were fit using data from the years 2001-2019 with the exception of those for deaths among females aged 80+, where models did not converge using data over this period. Causes within Chapter IV, discussed above, make up a significant portion of mortality among females within this age range. Following the analysis of this Chapter described above, ARIMA models for deaths among females aged 80+ were fit using data from 2010 to 2019, over which period it was possible to fit a satisfactory model to the observed data.

6.4.2.3 Monthly diversity in mortality causes during the COVID-19 Pandemic

Using the seasonal ARIMA models of diversity in mortality causes described above, the expected diversity at q = 1, 2, and ∞ in each month in 2020 and 2021 was forecast based on trends from 2001-2019. Figure 6.8 shows this forecast for males and females separately across all ages, alongside the observed monthly diversity in causes of mortality both including and excluding COVID-19 deaths. Alongside this, the ASMR of deaths attributed to COVID-19 in each month is plotted to provide context for the peaks of mortality during the pandemic.

Across diversity measures, when COVID-19 deaths are included, two periods of significant deviation from forecast diversity are observed: April and May 2020 and November 2020 to February 2021. In these months, a notable reduction from the expected range of diversity is observed. In April 2020, observed diversity at q = 1 was found to be less than half of the central estimate of forecast diversity (upper panel of Figure 6.8A). This indicates that, in effect, the range of causes from which deaths were likely to occur was halved. These two outlier periods correspond to the first two waves of the COVID-19 pandemic in Scotland when the largest number of deaths associated with the virus occurred, as discussed previously in Section 1. As explained previously, the reductions in diversity observed here are the effect that we might expect given the dominance of COVID-19 as a cause of mortality.

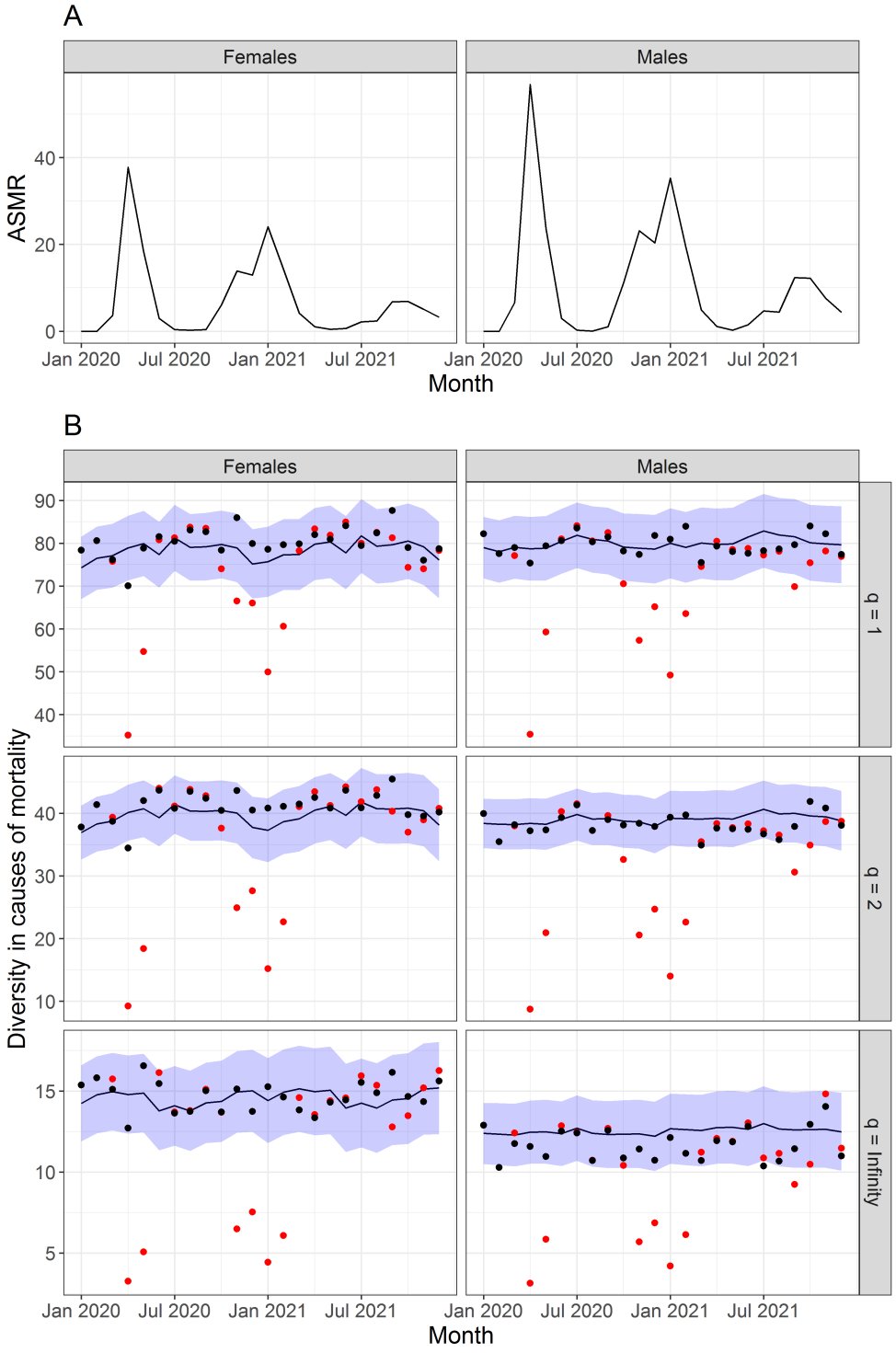


Figure 6.8: A: Age-standardised mortality rate of deaths associated with COVID-19 in each month in 2020 and 2021 in males and females separately. B: Diversity in mortality causes at $q = 1$, $q = 2$ and $q = \infty$ in each month in 2020 and 2021 in males and females, with COVID-19 mortality included (red) and excluded (black). Black line represents the central estimate of ARIMA model forecasts with blue ribbons indicating 99% confidence intervals.

With COVID-19 removed as a cause of mortality, the COVID-19 pandemic had no significant effect on the diversity of non-COVID-19 mortality causes in almost every month during 2020 and 2021. The only notable case in which diversity in mortality causes excluding COVID-19 is found to be outside of the 99% confidence intervals of the ARIMA model forecasts is females in April 2020. Observed diversity for females in April 2020 is slightly lower than the expected values only in diversity at $q = 1$ and $q = 2$ (upper and middle panels of Figure 6.8A). This corresponds to the initial peak of COVID-19 prevalence in Scotland. To examine the causes which may be linked to this, the mean ASMR of individual mortality causes in April in the years 2016-2019 was found and subtracted from the ASMR of these causes in 2020. The five causes with the largest and smallest absolute difference are shown in Figure 6.9. The largest increases are among degenerative diseases such as dementia and Alzheimer's disease. These causes have been noted in the literature as increasing in prevalence in this period due to issues surrounding old age care homes, this is discussed further in Section 6.5.2.

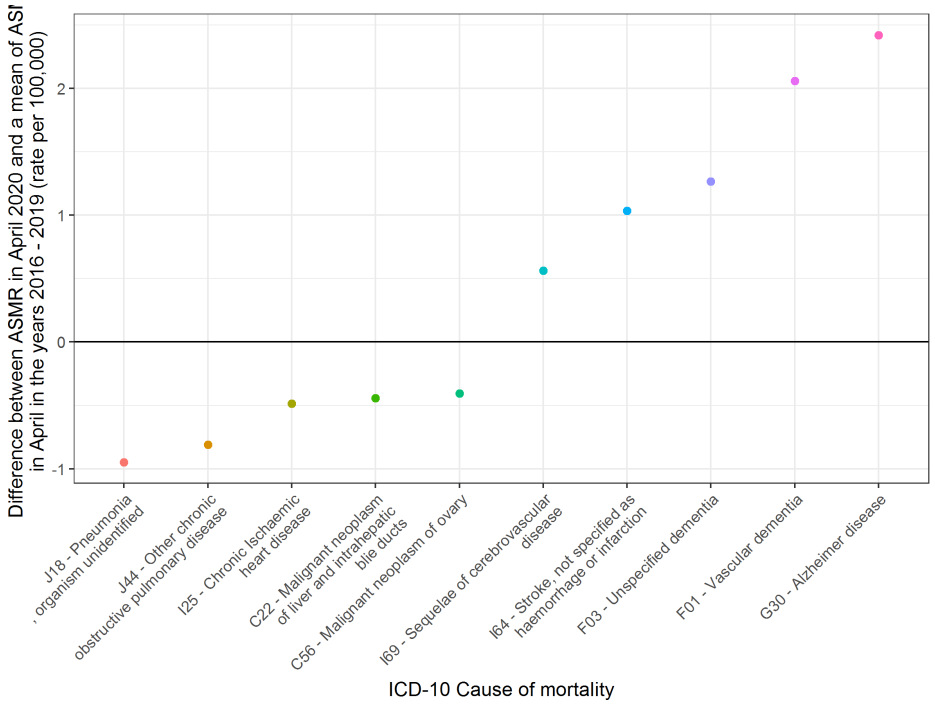


Figure 6.9: The difference in ASMR in April 2020 from a mean of ASMRs in April in the years 2016-2019 in females across all ages in Scotland the five causes which increased most and the five causes which decreased most from the previous mean in 2020.

6. The impact of the COVID-19 pandemic

The analysis in this section presents evidence that diversity across all causes of mortality excluding COVID-19 deaths in each month in 2020 and 2021 was similar to what would be expected given trends in the previous 19 years. To examine whether the impact of the pandemic was limited to certain ICD-10 Chapters or age groups, mortality is next grouped by age and ICD-10 chapter.

For the analysis by age group, monthly trends in the diversity of mortality causes in 2020 and 2021 for deaths among males and females and in twenty-year age groups are shown in Appendix Figure A.16. With COVID-19 deaths excluded from analysis, monthly diversity in mortality causes is within the forecast confidence intervals without notable exception. The increase in diversity at $q = \infty$ found among females aged 0 to 19 years in the year 2021 is not reflected in the monthly analysis in Appendix Figure A.16. The finding was, therefore, likely not linked to a brief phenomenon but rather to year-wide dynamics.

The reduction in diversity outside of expected bounds in deaths during April 2020 among females at all ages shown in Figure 6.8 is not observed in any age group. Given that Figure 6.9 shows an increase in the prevalence of several degenerative diseases, it might be expected that among females aged 80+ this effect would be more pronounced. That it is not observed at all among this group and is only observed strongly at $q = 1$ and $q = 2$ across all ages suggests that this finding was not associated with an increase in the prevalence of common causes. Rather it is likely a consequence of fewer deaths attributed to a range of relatively rare causes being recorded in this period.

For the analysis by selected ICD-10 chapters, the diversity in mortality causes in each month in 2020 and 2021 for the relevant chapters is presented in Appendix Table A.17 alongside forecast estimates. In most Chapters, diversity estimates are remarkably well predicted by the forecast values obtained from ARIMA models. Few cases are found with outliers beyond 99% confidence intervals with none deviating remarkably. Within Chapter I: “Certain infectious and parasitic diseases”, diversity in mortality causes is observed to reduce from mid-2020 onwards and is found to be consistently around the lowest values of the 99% projected confidence intervals. This reduction may be associated with spillover effects from public health measures aimed to reduce the spread of COVID-19 such as lockdowns, mask wearing, and hand washing.

The previously observed increases in prevalence among degenerative diseases among females in April 2020 is not observed to cause diversity in mortality causes to fall outwith ARIMA model confidence intervals. These causes are found within Chapters V and VI where diversity in causes is low due to dominance by dementias (Chapter V: F01 and F03) and Alzheimer's disease (Chapter VI: G30). As these causes already represent a large proportion of mortality within these Chapters, their further increase in prevalence has a negligible effect on diversity.

6.5 Discussion

This chapter presents findings on the diversity of mortality causes during the COVID-19 pandemic. This section begins with a summary of the findings of this chapter (Section 6.5.1), followed by an interpretation of these findings (Section 6.5.2). I then discuss the strengths and limitations of this work (Section 6.5.3) and future research directions in this area (Section 6.5.4). Finally, the implications of this research are explored (Section 6.5.5).

6.5.1 Principal findings

The findings of this chapter are presented in this section. They are structured by the research question, posited in Section 6.2, which they address.

6.5.1.1 What was the effect of the COVID-19 pandemic on diversity in mortality causes in Scotland across in the years 2020 and 2021?

The COVID-19 pandemic had a considerable direct and indirect toll on the population of Scotland in 2020 and 2021. During this period, deaths following infection with the COVID-19 virus were the most common cause of mortality, associated with around 1 in 10 deaths. This caused a significant reduction in the diversity of mortality causes in these years. This finding of lowered diversity is consistent with what would be expected with the introduction of a novel and dominant cause of mortality.

Despite the burden of COVID-19, diversity in the distribution of mortality causes with COVID-19 deaths excluded was remarkably similar projected counterfactual scenarios based on trends in previous years. However, the continuation of previous patterns highlights a change in trends over time in the diversity of mortality causes. In males, diversity in mortality causes has fallen across measures following 2016, though most notably at $q = \infty$, with trends at less conservative measures less clear. Among females, the rate of diversification in causes of mortality is shown to slow following 2009.

6.5.1.2 Does the diversity in causes of mortality in Scotland exhibit a seasonal pattern?

A seasonal pattern is observed in the diversity of causes of mortality across the Scottish population in deaths among both males and females. Diversity in mortality causes is observed to be higher in Summer months and lower in Winter months from 2001 to 2019. This finding is more pronounced at less conservative measures of diversity which indicates that a greater variety of rare causes of death occur in Scotland in Summer rather than seasonal trends being driven by the most common causes.

6.5.1.3 What was the effect of the COVID-19 pandemic on diversity in mortality causes during the peaks of the COVID-19 pandemic in Scotland?

The peaks of COVID-19 mortality in 2020 and 2021, associated with waves of high prevalence of the virus in the population, are clearly visible in analysis of monthly diversity in causes of mortality. Deaths attributed to COVID-19 caused a significant reduction in diversity in causes of mortality during both the initial wave of the pandemic and the second wave as defined in Section 6.1. With COVID-19 deaths excluded from analysis, the diversity of mortality causes in each month was generally found to be within the expected range of forecasted values, based on trends in previous years. A minor exception was observed among females in April 2020 where diversity fell slightly below expected values which is, partly, suggested to be associated with increased degenerative diseases prevalence. However, diversity within those aged 80+ (the age range most affected by degenerative diseases) was found to be within expected values in this month. Therefore, the the reduction in diversity across all ages among women in April 2020 was likely caused by a reduced variety of causes of mortality occurring across all ages, not within a specific age group. Further, this reduction in the variety of causes was not limited to one ICD-10 Chapter. The reasons for a lesser variety of causes being recorded in April 2020 are unclear but may be associated with delayed treatment or overstretched resources in the initial wave of the pandemic.

6.5.1.4 Was the impact of the COVID-19 pandemic on diversity in mortality causes consistent between groups of causes and in deaths at different ages?

The effect of the COVID-19 pandemic, in terms of mortality as well as indirect impacts, were not equal across all ages. Indeed, the impact of COVID-19 mortality on diversity in mortality causes is shown to be greatest in deaths among those aged 80 and older. However, excluding deaths attributed to COVID-19, previous trends in diversity in causes of mortality are, mostly, observed to have continued undisturbed. The only notable departure from trends in diversity in mortality causes in the years 2001 to 2019 is in diversity at $q = \infty$ in females aged 0 to 19. Here, an increase in 2020 is suggested to have been associated with a reduction in the prevalence of intentional self-harm by hanging, strangulation and suffocation, previously the leading cause of mortality among this group. A reduction in diversity in the following year to levels similar to those in 2019 was caused by an increase in mortality in the perinatal period. This finding, although notable, was associated with a relatively small number of deaths, with fewer than 15 deaths among females aged 0-19 associated with each of these causes in any year 2019-2021.

Beyond the findings in all-cause diversity, no notable differences are observed from previous trends in the additive value of ICD-10 chapters in 2020 and 2021. This indicates that, under this analysis at the level of ICD-10 Chapters, no cause-specific effect of the pandemic on diversity in mortality causes was observed. This is especially notable because of the various proposed cause-specific effects of the COVID-19 pandemic, a point to which I return below. In analysis of diversity in mortality causes in each month with deaths broken down by age at death and ICD-10 Chapter, no notable deviation from the forecast values of diversity in mortality causes is observed.

6.5.2 Interpretation

To understand the impact of the COVID-19 pandemic on population health, it is important not only to examine direct mortality associated with the virus, but also to assess the changes in the overall burden of mortality during the pandemic (Figueroa et al., 2020b; Mansfield et al., 2021). Several studies have examined different aspects of this in Scotland, from measuring changes in mortality rates associated with specific (non-COVID-19) causes to calculating excess deaths excluding COVID-19 mortality (Figueroa et al., 2020a; Burton et al., 2021). The aim of this chapter was to use diversity in mortality causes to add to this growing knowledge base on the influence of the pandemic on population health.

Findings across this chapter suggest that during 2020 and 2021 changes to the distribution of mortality causes were consistent with counterfactual projections of trends in diversity in mortality causes in 2020 and 2021. This is evident across the population, in deaths within distinct age groups, and across ICD-10 chapters; further, these findings are consistent both across the years 2020 and 2021 and in each month within these years. The counterfactual scenarios used as comparators to observed mortality cause diversity in this chapter are based on trends in diversity in mortality causes in the years prior to the pandemic. They therefore, rely on the assumption that diversity in causes mortality in 2020 and 2021 would not have differed markedly from previous years, in the absence of the pandemic. My findings suggest that year on year trends in mortality cause diversity in Scotland are relatively consistent however, it is important to give consideration to potential exogenous influences on mortality cause diversity not considered in this chapter.

6. The impact of the COVID-19 pandemic

Certain conditions and causes of mortality have been more closely related to pandemic-related disruption than others. Studies have recommended monitoring mortality associated with cardiovascular disease due to changes in the behaviour both of clinicians and patients during the pandemic (Ball et al., 2020). Furthermore, at the times of the greatest acute pressure on the healthcare system, many cardiac resources were redirected toward COVID-19 care. Regardless of this, the contribution of ICD-10 chapter IX: Diseases of the circulatory system to all cause diversity showed little deviation from prior trends. Similarly, even though Cancer detection, screening and treatment systems were affected by the COVID-19 pandemic (Campbell et al., 2021; Baxter et al., 2021; Figueroa et al., 2021), little deviation from previous trends is observed in the diversity of cancers that affected the population in 2020 and 2021. This finding is what might be expected because most cancer deaths do not occur suddenly. Following detection of cancer individuals may live for a number of years (Wishart et al., 2010; Sud et al., 2020). Therefore, if the pandemic affected detection and screening services, it might be expected that any impact of this disruption might be observed in the years following the pandemic rather than in 2020 and 2021.

Some of the effects of the pandemic may not be apparent for years to come. This is especially true of diseases and conditions for which testing and screening procedures were disrupted by the pandemic, such as disruption to cancer screening services. Beyond disruption to screening services, many sources report delayed or forgone care for a variety of conditions during the pandemic (Chen & McGeorge, 2020; Rattka et al., 2021; Malagón et al., 2022). Alongside potential long-term effects of infection with the COVID-19 virus itself, a range of other potential health risks have been proposed which may become apparent in the coming years (Arthi & Parman, 2021). These potential long-term consequences are mentioned here to highlight that the effects of the COVID-19 pandemic are unlikely to be limited to the years studied in this chapter. Therefore, in conjunction with other measures of population health, the diversity of mortality causes should be monitored further in the coming years to assess the impact of the pandemic. This will be complicated by the previously discussed slowing to improvements in mortality noted in Scotland and other countries in the late 2010s.

Age-standardised mortality rates excluding COVID-19 fell from 2019 to 2020 and into 2021. This continued a long-running trend which had slowed in the later years of the 2010s (Ramsay et al., 2020). Evidence of reductions in mortality associated with a host of conditions have been observed in the literature, especially during the first UK lockdown in Spring 2020 (Mansfield et al., 2021). The various govern-

mental and societal responses to the pandemic likely had innumerable interwoven consequences on the health of the Scottish population. As noted, these combined to produce lower mortality rates than in previous years and a commensurate increase in the diversity of mortality causes. Variation in the causes of mortality other than COVID-19 recorded during the pandemic increased. However, this increase is as would have been expected given a similar decrease in mortality rates in other circumstances according to trends from 2001 to 2019.

The findings of this chapter suggest that trends affecting variation in the distribution of mortality causes changed little during the COVID-19 pandemic. However, this serves only to highlight a possibly worrying change in trends in recent years. In deaths across all ages, diversity in mortality causes has fallen since 2016 in males while the pace of diversification has slowed since 2009 in females. This thesis has warned of the potential pressures on various aspects of public health and healthcare systems associated with continued diversification in mortality causes. However, this should not be taken to mean that a reduction in diversity is necessarily a positive sign. A reduction which is observed more strongly under diversity at more conservative measures implies increasing prevalence in the most common causes of mortality. This should be a concern for public health officials as it implies strategies to mitigate mortality from these causes are reducing in effectiveness. For example, an increase in the mortality rate associated with acute myocardial infarction (ICD-10 code I21) can be observed in Figure 6.2 from 2016 onwards. An increase in mortality rates attributed to heart attacks (using a broader definition including both ICD-10 codes I21 and I22) has been previously reported by public health officials from 2016 onwards (Public Health Scotland, 2022b).

Although this chapter shows that the COVID-19 pandemic had little impact on variation in causes of mortality, it is important to note that the pandemic did affect population health. Throughout 2020 and 2021, there were changes in cause-specific mortality rates for conditions such as cardiovascular and degenerative diseases (as discussed in Section 6.1). These changes might have been expected to alter the diversity of mortality causes during the pandemic. However, the limited impact I report in this chapter can be attributed to two probable reasons. Firstly, mortality causes in Scotland are highly dominated by the most common causes, [4]. Therefore, changes in the prevalence of other causes may have had limited impact on the diversity of mortality causes. Secondly, changes in the prevalence of vari-

[4]The two most common causes of mortality account for approximately 20% of the total ASMR in each year from 2001 to 2019 (Chapter 4)

ous causes likely cancelled each other out to an extent, with some becoming more common and others less common. This second mechanism might be considered a limitation of measures of diversity in mortality causes. However, the lack of a distinct trend in diversity suggests that no systemic patterns were observed in the distribution of mortality causes. In other words, rare causes did not become more or less common overall, and the prevalence of the most common causes of death did not significantly change from previous years in 2020 and 2021 (see Figure 6.2).

In this chapter, I find a seasonal pattern in diversity in mortality causes in Scotland, with increased variation in causes of mortality in Summer months, especially at less conservative measures of diversity. The seasonal pattern in mortality cause diversity is shown to be weaker in measures which weight the most common causes more heavily. Therefore, the findings I present in this chapter suggest that an increased variety of relatively rare causes of death occur in Summer in Scotland. This phenomenon has not previously been examined or discussed. There may perhaps be benefit to future public and population health research examining the reasons for greater variation in causes of death in Summer months and whether addressing this increased variation can aid in public health planning. Given that all-cause mortality rates are generally worse in Winter months, it may be that the case that Summer months are more representative of a healthy population.

While this chapter suggests limited impact of the COVID-19 pandemic on diversity in causes of mortality in Scotland, I have reported some cause-specific effects. I show that degenerative disease mortality among women was higher than a 4 year average in April 2020. I have noted that this is, likely, associated with well-reported failings around the protection of care home residents in Scotland (Reilly et al., 2020). A relatively high rate of infection with COVID-19 occurred in care homes (Dutey-Magni et al., 2021). Subsequently those with degenerative diseases, such as dementia, have been found to be likelier to face severe COVID-19 outcomes including death (Numbers & Brodaty, 2021). There is continuing debate on the number of deaths in care homes which should be attributed to COVID-19 as evidence has suggested many deaths occurred in those who were likely to be infected with the virus but were not tested (Dutey-Magni et al., 2021). Therefore, the rise in degenerative disease mortality rate in April 2020 may be partly due to deaths which might, following testing, have been attributed to COVID-19.

6.5.3 Strengths and limitations

In this chapter, ARIMA models are used to forecast monthly diversity in mortality causes into the years 2020 and 2021. A great variety of methods for the analysis of time series data such as this exist in the literature. ARIMA models were chosen for analysis in this chapter due to their general flexibility and particularly because they are well suited to both seasonal and non-seasonal data. Given the varied trends in the data assessed in this chapter, this flexibility was deemed an asset.

ARIMA models were fit separately to diversity data in each situation in which they were employed, descirbed in Section 6.3.7 following the same methodology. The quality of these models, in terms of how well they fit the given data, varied, resulting in wide confidence intervals in some forecasts. It may have been possible to overcome this using different techniques for example TBATS methods (De Livera et al., 2011). The relative advantages of such methods compared to ARIMA models are debated and for example, methods such as TBATS are also susceptible to producing wide confidence intervals and are not always considered advantageous to ARIMA models (Hyndman & Athanasopoulos, 2018). It would, perhaps, have been advantageous to support the findings of this chapter by applying other time series analysis methods to the research questions addressed here. This was, however, beyond the scope of this thesis. Given the change in the observed trends in diversity in mortality causes during the 2010s it may have been advantageous to improve the ARIMA models used in this chapter by applying a two-stage fitting. This could have consisted of applying segmented regression analysis to the monthly measurements of diversity in mortality causes to find a breakpoint. ARIMA models could then have been applied to the second segment of this regression ensuring a more consistent trend during the training period of the ARIMA model.

In this chapter, an increase in the ASMR from certain degenerative diseases is noted in April 2020 compared to previous years. This reduction in April 2020 is not found when examining individual ICD-10 Chapters as might have been expected. This is because within their respective ICD-10 Chapters, dementia and Alzheimer's disease are very dominant, and therefore changes in their prevalence generate only limited changes in diversity. This highlights that, when the number of causes under examination is small and the distribution is heavily skewed towards a small number of dominant causes, diversity measures can be insensitive to changes in prevalence.

As discussed in previous chapters, the measurement of diversity in mortality causes is, to an extent, sensitive to the measure of diversity utilised. As shown in Appendix B.1 the use of a measure which explicitly accounts for zero values in the distribution of mortality causes can have an impact on trends in diversity. It is therefore, possible that the use of this measure could impact the measurement of diversity in mortality causes in analysis within this chapter. It may be valuable to replicate the analysis in this chapter using a measure which accounts for zero values to assess whether the conclusions of this chapter are sensitive to the impact of the prevalence of causes of mortality which do not occur in every year.

Garbage codes are ICD-10 codes which denote causes of mortality which are considered vague or unhelpful in population health research. In analysis within this chapter garbage codes were retained in the distribution of mortality causes. As discussed in Sections 4.5.4 and 4.5.4 this may have the effect of biasing the measurement of diversity. The sensitivity of yearly trends in diversity to this bias is tested in Appendix B.2 using a simple method for redistributing garbage codes to valid causes of mortality. This analysis garbage codes have little impact on these yearly trends, it is however, possible that garbage codes would impact the monthly analysis discussed in this chapter. Future research into diversity in mortality causes should consider the potential impact of garbage codes.

6.5.4 Future areas of research

As noted previously, various studies on the effect of COVID-19 on population health have suggested that measures should be monitored into the future. Several of the potential consequences of the pandemic are likely to have a delay before their effects are fully felt. In particular, the prevalence of different cancers has the potential to be strongly affected in the coming years given that the pandemic had an impact on screening services (Campbell et al., 2021). Given the large contribution of cancers to overall diversity in causes of mortality (shown in Chapter 4) this may impact diversity across all causes of mortality as well as the diversity of cancers. In the ongoing study of the impact of the COVID-19 pandemic on cancer mortality in Scotland, it may be useful to measure diversity in causes of cancer deaths. Neoplasms are the most diverse range of mortality causes in any ICD-10 chapter. The diversity of causes within this chapter has increased in recent years as common cancers such as lung cancer have reduced in prevalence (and rate) and the distribution has become more even (See Appendix Figure A.18 and discussion

in Chapter 4). Reductions in the diversity of cancers in coming years would therefore indicate an increase in the relative prevalence of the most common cancers. Under the assumption that patterns in cancer mortality would have continued in the absence of the pandemic, this might imply that the COVID-19 pandemic had a negative impact on the detection and treatment of common cancers. This trend could also be observed through an examination of mortality rates among these common cancers. However, the large number of mortality codes describing varied cancers would make it more challenging to detect overall trends across rare cancers using traditional methods. In this case the measurement of diversity could prove useful. Were the diversity of causes of mortality among cancer deaths to increase markedly following the pandemic it might indicate an increased prevalence of these rare causes of mortality. This may be expected as it is likely that common cancers, such as breast or colon cancer, were less affected by the pandemic than rare causes. This is because these common causes have national screening services which are well established and were able to adapt to the conditions of the pandemic (Campbell et al., 2021). More rare cancers lack this infrastructure and rely more on in-person diagnostic tests. While many of these tests continued, with mitigations, during the pandemic they may have been more impacted because of the more bespoke nature of such treatment.

6.5.5 Research implications and next steps

In this chapter, diversity in mortality causes is used to assess the effect of COVID-19 pandemic on the distribution of mortality causes in Scotland. Overall, variation in causes of mortality fell, because around 10% of deaths were attributed to COVID-19. However, with COVID-19 mortality excluded, there is little evidence to suggest that, compared to counterfactual projections, the pandemic had an effect on diversity in the distribution of mortality causes. For diversity measured over years and in each month, values were within ranges which would be expected given patterns in previous years. This finding is shown to hold across age groups and within selected ICD-10 Chapters. Despite the varied pressures of the COVID-19 pandemic, patterns of diversity in mortality causes are shown in this chapter to have been remarkably robust. This is significant not only to understanding the effects of the pandemic but also in understanding how to focus the recovery. Reductions in diversity in mortality causes, observed from 2016 onwards in males across measures, continued during the pandemic. These reductions in diversity in mortality causes are suggested to have been associated with increases in the proportion of the pop-

ulation who face the most common causes of death. Understanding that these trends appear to have continued means that in recovering from the COVID-19 pandemic healthcare and public health systems must continue to focus on treatment and prevention of the most common causes of mortality and recognise that they are becoming, relatively, more common. Furthermore, they must consider how the resurgence of these causes, and particularly ischaemic heart disease the most common cause, will interact with COVID-19 and whether comorbidities will worsen trends in the prevalence of these causes. While some of this insight could have been gained from a study of specific rates of mortality using diversity measures allows us to take a wider focus and parse out these overall trends.

There has been much debate over how, going forwards, the COVID-19 pandemic will affect population health. The virus itself remains endemic across the world and, as previously noted, many effects of societal interventions during the pandemic may only present in years to come. The pandemic has altered the distribution of mortality causes through the addition of COVID-19 as a cause of mortality and it remains unclear how patterns of mortality will evolve in the future.

Due to Scotland's, comparatively, poor health outcomes prior to the COVID-19 pandemic it might be expected that the impact of the pandemic would be felt more acutely than in other nations. Despite this, the findings of this chapter suggest little effect on diversity in the distribution of mortality causes, excluding COVID-19. Across the world various different NPI strategies were employed in the face of COVID and different countries faced largely different mortality rates due to COVID-19. The impact of the pandemic on diversity in mortality causes may diverge drastically in such varied situations.

Previously in this thesis diversity has been measured in the underlying causes of mortality recorded for each individual. However, further information regarding the health of an individual at the time of death can be gleaned from analysis of contributory causes of mortality recorded on the death certificate. The following chapter examines diversity in the contributory causes of mortality recorded in Scotland from 2001 to 2021 and assesses the diversity of contributory causes among those whose underlying cause of death was COVID-19 relative to those who died from other leading underlying causes of mortality.

Chapter 7

Diversity in contributory causes of morbidity and mortality: national trends and the disease burden associated with COVID-19

7.1 Background

Previously in this thesis, diversity has been measured in the underlying causes of mortality faced by the population of Scotland and by Scottish subpopulations. Diversification in causes of mortality has been observed from 2001 to 2021 in deaths across all ages; however, mortality associated with COVID-19 is shown to have reduced the diversity of underlying causes in 2020 and 2021. These findings are based on study of the underlying cause of mortality recorded for each death in Scotland in these years. However, underlying causes of mortality do not tell the full story of the health of each individual at the time of death and can exclude other causes which contributed to death.

Alongside the underlying cause of death, most mortality recording systems allow for a number of contributory factors to be listed. These are, generally, ICD-10 codes for diseases and conditions that an individual has faced which contributed to or hastened death but were not thought to be the direct cause of mortality (National Records of Scotland, n.d.). These causes are considered a valuable resource

in the study of population health in the literature (Fedeli et al., 2015). Contributory codes allow a great deal of additional information to be recorded on the health of the deceased such as information regarding multimorbidities or further detail in the case of complex diagnoses (Grundy & Stuchbury, 2022).

Contributory causes of morbidity and mortality have been shown to be informative with respect to risk factor associations. Batty et al. (2019) show that associations between risk factors, such as smoking and educational attainment, and causes of mortality, such as cardiovascular disease, cancers, and dementia, hold no matter where the cause of mortality appeared on the death certificate. Partly because of such associations, it has been proposed that relying on underlying causes of mortality can lead to undercounts and conceal trends in population health (Bishop et al., 2022; Trias-Llimós & Permanyer, 2023). This can be addressed through the use of "multiple causes of death" (MCOD) approaches which take into account the contributory causes of death recorded on death certificates (Désesquelles et al., 2014; Piffaretti et al., 2016; Rao, 2020).

Diversity in mortality causes has previously been assessed from a MCOD perspective by Trias-Llimós and Permanyer (2023). They use novel methods to assess differences in the set of mortality causes faced by each individual in the USA from 2003 to 2018. The measure of dissimilarity used by Trias-Llimós and Permanyer (2023) is not explicitly a measure of diversity; however, for simplicity, I have and will continue to refer to it as such throughout this work. Their measure quantifies the dissimilarity in causes of mortality (both underlying and contributory) between all possible pairs of individuals in the population and takes the average of these differences. This proposed method is flexible in that it allows for different weights to be given to the underlying cause compared to contributory causes. Trias-Llimós and Permanyer (2023) found diversification in mortality causes over their study period and showed that diversity increased more rapidly with multiple causes of death accounted for than when studying underlying cause diversity alone. This previous study included underlying causes in all analyses and therefore the diversity in contributory causes in isolation has not previously been assessed.

This chapter aims to examine the temporal trends in the diversity of contributory causes of mortality in Scotland from 2001 to 2021. These trends are then compared to trends in the diversity of underlying causes of mortality as described in previous chapters of this thesis and the relationship between these measures is assessed. Greater diversity in contributory causes, similar to underlying causes,

may be a sign of increasing pressure on healthcare systems indicating that a wider variety of conditions require care. Alternatively, were the diversity in these additional causes to have fallen it would indicate that despite dying due to more varied causes of mortality (as shown in previous chapters), the population had faced a more uniform set of co-morbid conditions. This might indicate increasing prevalence of specific causes of morbidity and ill health which may be important to address to improve population health.

The previous chapter of this thesis examines the effect of the COVID-19 pandemic on the distribution of mortality causes in Scotland. The health status of those whose deaths were attributed to COVID-19 in 2020 and 2021 is important to understanding the impact of the pandemic. Many conditions have been found to be risk factors for COVID-19 mortality including cardiovascular diseases and diabetes (Chadeau-Hyam et al., 2020; Elliott et al., 2021). Examining how the variation in causes recorded alongside COVID-19 compared to other leading causes can offer an insight into the relative health of those who died due to the novel coronavirus. If the diversity of contributory causes were to be higher in COVID-19 deaths, it would indicate that among the subpopulation of those who succumbed to this disease, greater variety in conditions and co-morbidities existed compared to other common underlying causes. On the other hand, if those who died following COVID-19 infection were found to have less diverse contributory causes, it might suggest that a set of common conditions appeared alongside COVID-19. To examine this, the contributory causes of mortality in COVID-19 deaths in 2020 and 2021 are investigated and the diversity of these causes is compared to those recorded alongside other leading causes in these years.

7.2 Research questions and objectives

This chapter aimed to address the following objectives:

- To examine temporal trends in the diversity of contributory causes of morbidity and mortality (defined as those appearing alongside the underlying cause of death on mortality records) in Scotland between 2001 and 2019.
- To assess the relationship between diversity in underlying causes of mortality and diversity in contributory causes of morbidity and mortality.

- To compare the diversity of contributory causes of mortality recorded alongside COVID-19 deaths with those recorded alongside other leading causes of mortality in 2020 and 2021.
- To describe the contributory causes of mortality most commonly listed alongside COVID-19 deaths.

As such the following research questions are answered:

- What are the national trends in the diversity of contributory causes of morbidity and mortality in Scotland?
- What is the relationship between the diversity of contributory causes and the diversity of underlying mortality causes in Scotland?
- Were the contributory causes of mortality recorded alongside COVID-19 more or less diverse than those recorded alongside other leading causes of mortality in 2020 and 2021?

7.3 Methods

7.3.1 Data

Scottish mortality data for the years 2001 to 2021 was extracted and processed as described in Section 3.2.1. Annual mid-year small area population estimates for the years 2001 to 2019 were obtained for Scotland as described in Section 3.2.2.

7.3.2 Life tables

Multiple-decrement life tables were constructed as described in Section 3.3.1 for all of Scotland, for males and females separately, for each year from 2001 to 2021. Life tables are calculated using single year age classes for ages 0 to 109 and an open-ended age class 110+.

7.3.3 Diversity in underlying causes of mortality

Normalised alpha diversity in underlying causes of mortality was calculated for males and females at all ages in Scotland in each year from 2001 to 2021 separately with year of death as subcommunities. Diversity was calculated under the Reeve et al.(2016) framework as described in Section 3.3.3 using distributions of mortality cause extracted from multiple-decrement life tables.. Diversity at q values of 1, 2, and ∞ was calculated in the distributions of individual ICD-10 three-character codes extracted from the multiple decrement life tables discussed in Section 7.3.3.

7.3.4 Contributory causes of mortality

In Scotland, up to 10 causes of morbidity or mortality can be recorded alongside an underlying cause of death on each death record. The underlying cause of death is normally included in this list of 10 causes; however, there are some situations in which coding guidelines indicate that if two contributory causes appear together, a third underlying cause should be listed (National Records of Scotland, n.d.). Codes other than the underlying cause of death which appear in this section are described by National Records of Scotland as: "'contributory factors', diseases or injuries, none of which was the underlying cause of the death, but each of which will have contributed in some way to the occurrence of the death (e.g. by hastening it)." (National Records of Scotland, n.d.). In analysis of the number of contributory causes of mortality recorded alongside each death, the underlying cause was discounted from the list of causes in each mortality record leaving only contributory causes. The number of contributory causes recorded in Scotland in each year 2001 to 2019 is shown in Figure 7.1. A large increase in the recorded number of contributory causes can be observed in both sexes from 2015 onwards[1]. This is the result of a new mortality recording system introduced by the Scottish Government in this year which is discussed in Section 7.4.1 below. To examine diversity in contributory causes of mortality the share of deathsat age x to which each contributory

[1]There is a notable increase in the number of contributory causes recorded among men from 2013 to 2014. The act referenced above came into force in May 2015 meaning that it is unlikely to be responsible for this increase. The reason for this increase is unclear. It is notable that the same increase is not observed in deaths among females in Scotland suggesting that the increase may not be associated with a systemic change.

cause was attributed to was multiplied by the life table (d_x) (Discussed in Section 3.3.1) to produce life table distributions of contributory causes. This means that, in effect, deaths for which only the underlying cause of death was recorded were not included in analysis.

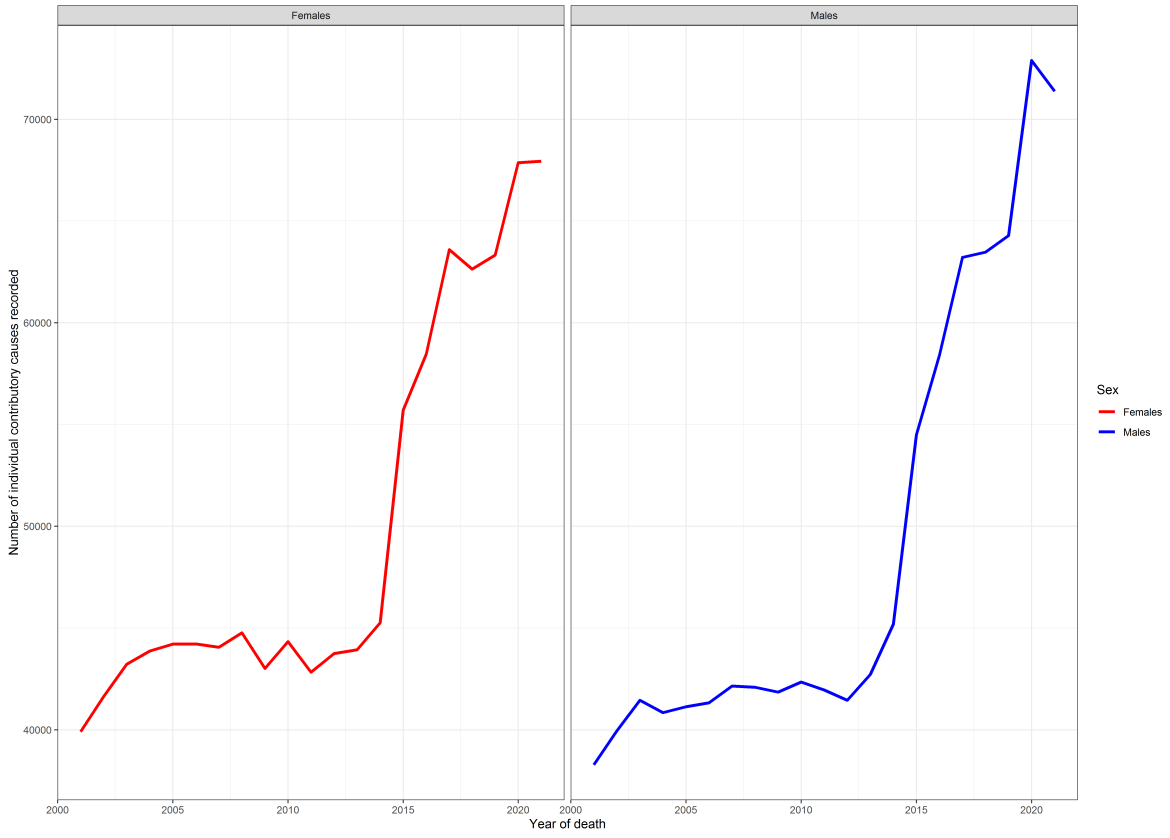


Figure 7.1: The total number of contributory causes of morbidity and mortality recorded on death certificates in Scotland in men and women separately.

Normalised alpha diversity, as described in section 7.3.3, was calculated in contributory causes extracted as discussed above at q values of 1, 2, and ∞. Diversity was calculated with years as subcommunities for males and females separately in each year 2001 to 2021. Diversity in contributory causes was also calculated in this way for the Scottish population divided by age into the three age ranges described in Chapter 5 deaths among those aged: 0 to 39, 40 to 79 and 80+. Trends in the diversity of contributory causes in these groups is shown in Appendix A.19.

Normalised alpha diversity was also calculated separately for the contributory causes of mortality recorded alongside leading underlying causes of mortality. Leading causes were defined as those which had been among the five most prevalent by ASMR across Scotland in either sex in 2020 or 2021 or had an ASMR of greater than 45 per 100,000 in either year in either sex. These causes were: acute

myocardial infarction (ICD-10 code I21), Alzheimer's disease (ICD-10 code G30), COVID-19 (ICD-10 code U07), malignant neoplasm of the bronchus and lung (ICD-10 code C34), other chronic obstructive pulmonary disease (ICD-10 code J44), chronic ischaemic heart disease (ICD-10 code I25), and malignant neoplasm of the prostate (ICD-10 code C61). The diversity of contributory causes recorded alongside these underlying causes was calculated in each year from 2001 to 2021, except for COVID-19 mortality, which occurred only in 2020 and 2021 (COVID-19 mortality was defined as described in Chapter 6).

7.3.5 Statistical analysis

Pearson's correlation coefficient was calculated to assess the relationship between normalised alpha diversity in contributory causes and normalised alpha diversity in underlying causes. Diversity was compared at the same q-value (underlying cause diversity at $q = 1$ was compared to contributory cause diversity at $q = 1$ and so on). The relationship between these measures was examined over the years 2001 to 2021 in each q value and each sex separately.

7.3.6 Software

Analysis in this chapter was performed in *R* version 4.1.3 (R Core Team, 2022). Data manipulation and processing was performed using version 1.3.1 of the *tidyverse* package, plots were created using version 3.3.5 of the *ggplot2* package alongside version 1.1-2 of the package *RColorBrewer* (Neuwirth, 2014; Wickham, 2016; Wickham et al., 2019). Calculation of diversity was performed using version 2.0 of the *rdiversity* package (Mitchell et al., 2020).

7.4 Results

The results section of this chapter is structured as follows. First a summary of the number of contributory causes recorded on death certificates in Scotland from 2001 to 2021 is presented (Section 7.4.1). Then trends in the diversity of contributory causes of mortality in Scotland are examined and are compared to trends in the diversity of underlying causes of mortality (Section 7.4.2). Finally, the number and diversity of contributory causes recorded alongside leading underlying causes of mortality are examined and compared to those recorded alongside COVID-19 in 2020 and 2021 (Section 7.4.3).

7.4.1 Contributory causes of morbidity and mortality in Scotland

In Scotland, up to 10 contributory causes of morbidity and mortality can be listed on the death certificate alongside the underlying cause of death. In most death certificates, one of these 10 causes is the underlying cause. Contributory causes are not recorded for every death. Table 7.1 shows summary data on the contributory causes of morbidity and mortality recorded on death certificates in Scotland. Between 0 and 10 contributory causes were recorded on death certificates in this period. On average, a greater number of causes was recorded in deaths among younger and older age groups than those aged 40 to 79, though differences were small.

Table 7.1: Summary statistics of the observed distribution of contributory causes among males and females across Scotland and within set age ranges across the years 2001 to 2021.

Sex	Mean number of contributory causes	Standard deviation	Lower quartile	Upper quartile
Deaths across all ages				
Females	1.7	1.4	1	3
Males	1.8	1.5	1	3
Deaths among those aged 0 to 39				
Females	1.8	1.9	0	3
Males	2.1	1.9	1	3
Deaths among those aged 40 to 79				
Females	1.6	1.5	0	2
Males	1.7	1.5	1	3
Deaths among those aged 80+				
Females	1.8	1.4	1	3
Males	1.9	1.4	1	3

From 2001 to 2014, the median number of contributory causes listed on death certificates alongside the underlying cause was 1. This increased to a median of 2 contributory causes per death certificate in 2015 in both sexes. Trends in the mean number of contributory causes are displayed in Figure 7.2A and show a similar increase around these years.

The increase in the average number of contributory causes recorded on each death certificate does not coincide with any change in mortality coding practice (i.e., a change in the guidelines for how each cause should be recorded) described by the data holders National Records of Scotland. It does however, come at around the same time as the introduction of new arrangements on certification of death which began on the 13th May 2015 having been introduced in the Certification of Death Act (2011). This Act did not, per se, aim to increase the scope of causes recorded on death certificates or to change coding practices. However, it did legislate to increase scrutiny of the accuracy of death certificates and to change the forms on which information was recorded. In doing so, the aim of the Act was to increase the accuracy of these records. It seems that as a result, an increased number of causes of mortality tended to be recorded on death certificates.

7.4.2 Diversity in contributory causes of mortality in Scotland

When examining the diversity of contributory causes of mortality in Scotland, a general trend for increasing diversity can be observed until around 2015-2016, as shown in Figure 7.2B. Following this, the diversity of contributory causes of mortality fell under all measures of diversity in females and fell or stagnated under diversity at $q = 1$ and $q = 2$ in males (upper panels of Figure 7.2B). The only case where a change in the direction of the trend was not observed was in males at $q = \infty$ (lower panel of Figure 7.2B). The change in trends observed here is striking as it might be expected that alongside the increase in the average number of causes recorded following 2015, the diversity of contributory causes commonly recorded on death certificates would increase. However, this is clearly not what occurred here. Despite the trends discussed above in both sexes and across diversity measures, the contributory causes of mortality was more diverse in 2001 than 2021. This indicates that a larger variety of ICD-10 3 character codes describing diseases and conditions were commonly recorded on death certificates. This is despite the average number of causes recorded for each individual remaining fairly constant before and after the increase in 2015 noted above.

In Appendix B.1.4 an alternative measure of diversity, related to measures of evenness, is used to calculate trends in the diversity of contributory causes of mortality in Scotland. This measure explicitly accounts of zero values in the distribution of mortality causes, meaning it is sensitive to causes which do not occur in every year of the study period. Using this measure diversity in causes of mortality is shown to fall from 2001 to 2010, after which trends are similar to those under normalised alpha diversity as discussed above with an increasing trend to 2015 followed by a reduction in diversity. Overall, this means that using this alternative measure, diversity in contributory causes of mortality fell from 2001 to 2019 in Scotland in both sexes.

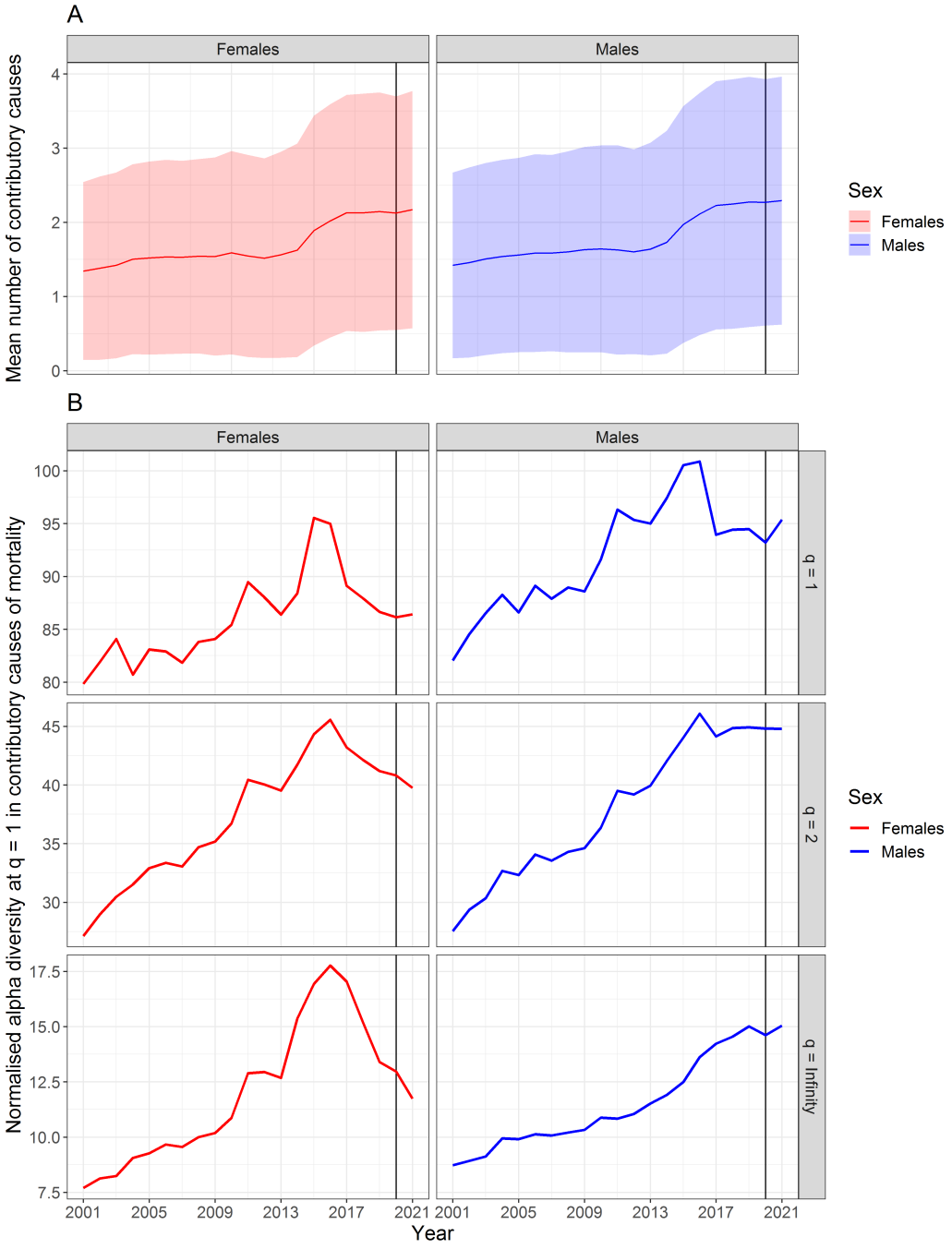


Figure 7.2: A) The mean number of contributory causes recorded on each death certificate in Scotland from 2001 to 2021 in males and females separately. Confidence intervals indicate the standard deviation from the mean. B) Normalised alpha diversity at $q = 1$, $q = 2$, and $q = \infty$ of contributory causes of mortality in Scotland in males and females in Scotland from 2001 to 2021. A vertical line indicates the year 2020 when the COVID-19 pandemic began.

Despite no discernible change in the average number of causes recorded for each death, among females the most common contributory causes became, over time, more likely to appear on death certificates after 2015. This is indicated by trends in diversity at $q = \infty$. This reduction in diversity is associated with increases in the prevalence of the ICD-10 code R68 which denotes "Other general symptoms and signs" and includes conditions such as "Hypothermia, not associated with low environmental temperature" and "clubbing of fingers" as well as "Other specified general symptoms and signs". The significance of the increase in the records on which this code appears is unclear due the general nature of the conditions which it codes for. This code appeared on 3.6% of all death certificates among females in 2016 increasing to 7.7% in 2021. It appeared most frequently alongside degenerative diseases such as dementia and Alzheimer's disease. Together, these causes (ICD-10 codes: F01, F03, G30) were the underlying cause on 23.8% of the death records where code R68 appeared among women in 2021. In deaths among males, chronic ischaemic heart disease (ICD-10 code I25) was the most common contributory cause of mortality from 2004 onwards; before this point "Pneumonia, organism unspecified" (ICD-10 code J18) was the most common.

In Appendix B.2.4 I perform a simple redistribution of garbage codes to valid causes of mortality, described in Section 4.5.4. This redistribution is performed to illustrate how trends in the diversity of mortality causes would appear accounting for garbage codes. Appendix B.2.4 shows that with garbage codes redistributed, trends in the diversity of contributory causes of mortality in Scotland are similar to those reported here up until 2016 where with garbage codes redistributed, diversity in causes of mortality continues to increase across measures and in both sexes. This indicates that despite the trends noted above, the diversity of valid causes of mortality recorded in Scotland increased throughout the study period.

Trends in the diversity of contributory causes of morbidity and mortality at different ages in Scotland are presented in Appendix A.19. Deaths among those aged 40 to 79 tended to be recorded alongside the most diverse set of contributory causes. Trends in deaths among those aged 40 to 79 and 80+ are mostly similar to those across all ages shown in Figure 7.2, though the change in trends following 2015, discussed above, is only seen in deaths among those aged 80+. In deaths among those aged 0 to 40 diversity in contributory causes fell from 2001 to 2019.

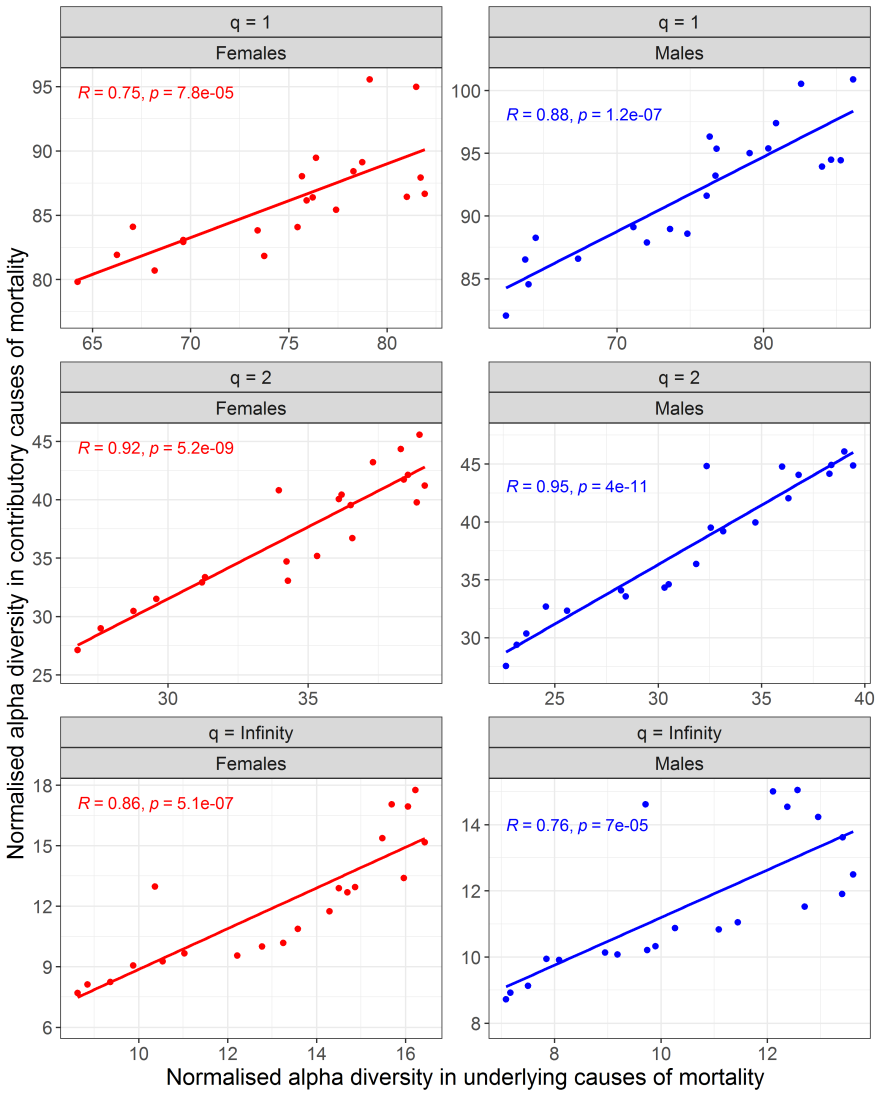


Figure 7.3: The normalised alpha diversity in contributory causes of mortality plotted against the normalised alpha diversity of underlying causes at $q = 1$, $q = 2$, and $q = \infty$ in deaths among males and females in Scotland from 2001 to 2019. Pearson's correlation coefficient and p-value are displayed in each facet for the relationship between underlying cause diversity and contributory cause diversity.

Previous chapters of this thesis have demonstrated that diversity in underlying causes of mortality in Scotland increased from 2001 to 2021, though with a change in trends in the 2010s, similar to the trends in contributory causes noted here. In the late 2010s, diversity in underlying causes is shown to have stalled or fallen, with reductions most notable at $q = \infty$. Figure 7.3 plots diversity in underlying causes of mortality against diversity in contributory causes of mortality, with measurements compared at each q value (i.e. diversity in underlying mortality causes at $q = 1$ is compared to diversity in contributory causes at $q = 1$) [2] . In both

[2]Diversity in underlying causes in 2020 and 2021 is measured here including deaths attributed to COVID-19; for further discussion see Chapter 6.

sexes, and across measures of diversity, a strong relationship is observed between diversity in contributory causes and diversity in underlying causes, as measured by Pearson's correlation coefficient. This observation means that variation in the range of underlying causes recorded on death certificates each year has increased alongside variation in the range of contributory causes. The relationship between underlying cause diversity and contributory cause diversity is found to be weakest among men under diversity at $q = \infty$ (lower panels of Figure 7.3). In this case, while the diversity of underlying causes of mortality fell in the late 2010s, diversity in contributory causes continued to increase.

7.4.3 Contributory causes of morbidity and mortality in the case of COVID-19 mortality

7.4.3.1 The number of contributory causes recorded alongside COVID-19

In the previous section, the average number of contributory causes of mortality recorded across all underlying causes of mortality from 2001 to 2021 is discussed. Figure 7.4 shows the distribution of the number of contributory causes recorded in 2020 and 2021 alongside all underlying causes and in seven of the leading underlying causes of mortality in these years. These causes were any that were among the five most prevalent by ASMR in either year in either sex or had a cause-specific ASMR of greater than 45 per 100,000 population in either year.

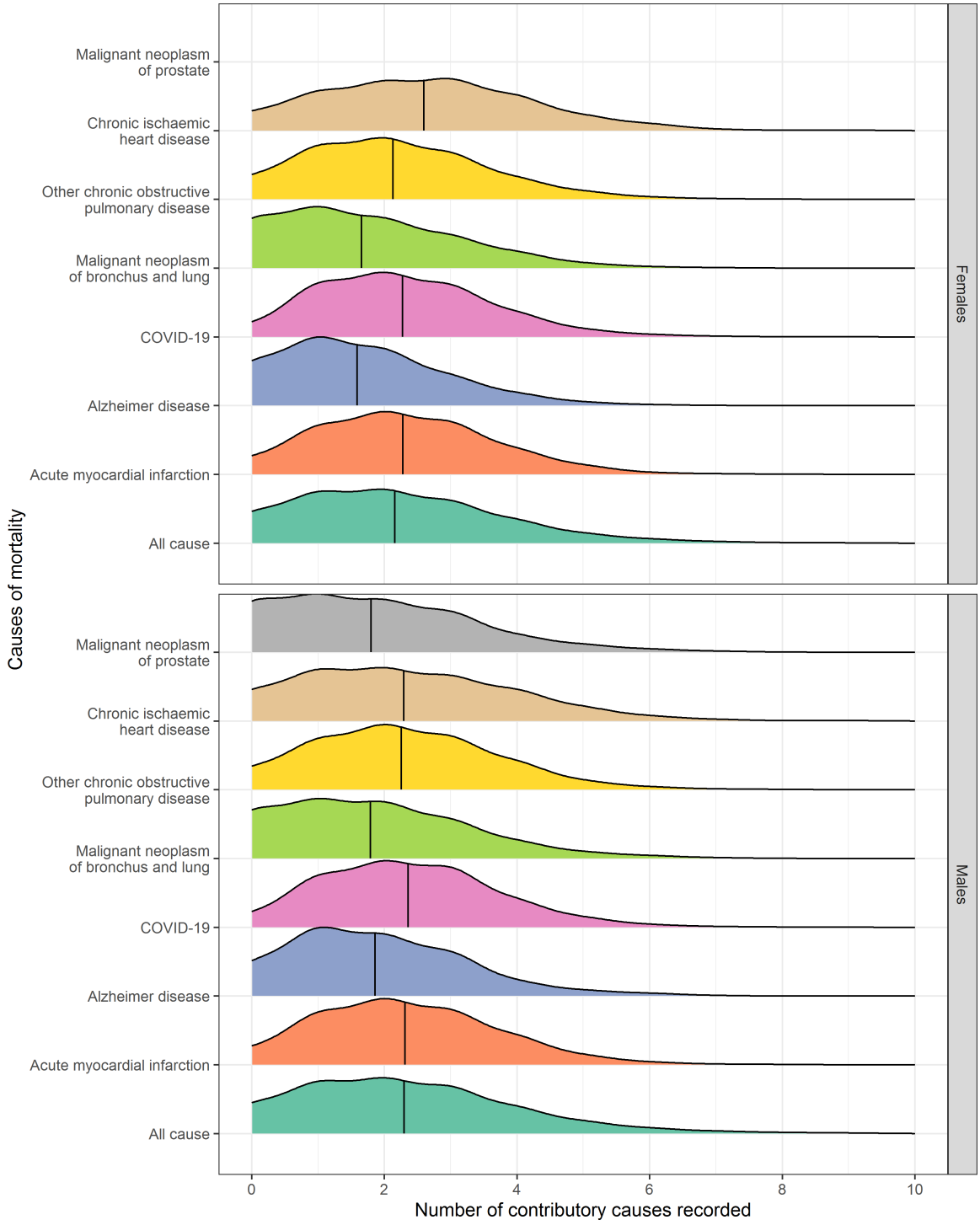


Figure 7.4: Ridge-plot illustrating the number of contributory causes of mortality recorded alongside all underlying causes and leading individual underlying causes of mortality in 2020 and 2021. Vertical black lines indicate the mean number of contributory causes recorded in each case.

Across all causes, and in each of the underlying causes shown in Figure 7.4, most records contained between 1 and 3 contributory causes. It was most common for deaths to be recorded without a contributory cause in deaths where Alzheimer's disease and malignant neoplasm of the bronchus and lung were the underlying cause. Compared to the other common causes of mortality discussed here, deaths where COVID-19 was the recorded cause were the least likely to have no contributory causes recorded alongside them. In other words, almost all individuals who were recorded as having died from COVID-19, also had other contributory causes listed.

7.4.3.2 Diversity in the contributory causes recorded alongside COVID-19 compared to other leading causes

Diversity in contributory causes of morbidity and mortality was measured in the causes recorded alongside each of the leading underlying causes of death discussed in Section 7.4.3.1. In Figure 7.5, the diversity of these contributory causes is plotted from 2001 to 2021. Compared to the diversity of contributory causes recorded across all underlying causes shown in Figure 7.2, less diverse contributory causes were recorded alongside each of the underlying causes shown here. This is what might be expected as the range of contributory causes recorded alongside each underlying cause is likely to differ. Therefore, when all causes are combined the overall distribution can be expected to be more diverse.

Figure 7.5: The normalised alpha diversity at q = 1, q = 2, and q = ∞ of the contributory causes of mortality recorded alongside leading underlying causes among male and females separately in Scotland from 2001 to 2021.

In most of the leading causes examined in this section, the diversity of contributory causes changed little across the study period. The diversity of causes recorded alongside Alzheimer's disease increased most notably (blue lines in Figure 7.2). Increases were also observed in the diversity of causes recorded alongside acute myocardial infarction at q values 2 and ∞ (orange lines in Figure 7.2). This indicates that in this cause, the most common contributory causes were recorded alongside fewer deaths in 2021 than in 2001, and instead, overall rarer causes were recorded.

A notable increase in the diversity of contributory causes recorded alongside "other chronic obstructive pulmonary disease" (COPD) is observed between 2010 and 2016 (yellow lines in Figure 7.2). The increase at the beginning of this period corresponds to a change in mortality coding practices introduced by the Scottish

Government in 2010 and the reduction in 2017 corresponds to another change in coding systems (National Records of Scotland, 2011; National Records of Scotland & Scottish Government, 2017). Neither of these changes were reported by NRS to directly affect deaths coded as being attributed to COPD by the National Records of Scotland. However, the fact that changes here occur in the same years that these coding practice changes came into effect suggests they were associated with these updates.

Greater diversity in contributory mortality causes is found for deaths associated with COVID-19 than for other leading mortality causes in 2020 and 2021. This was more prominent at higher values of q (the pink lines are further from the other lines). This indicates that the contributory causes associated with COVID-19 were less dominated by the most common causes of morbidity and mortality than in other underlying causes. Despite this, under diversity at $q = 1$, the contributory causes recorded alongside COVID-19 were similar to other leading causes such as COPD and chronic ischaemic heart disease. While the specific contributory causes recorded alongside each underlying cause of death are likely to differ, this indicates a similar degree of variation in the range of secondary causes found alongside each leading underlying causes of death.

7.4.3.3 The causes of morbidity and mortality recorded most commonly alongside COVID-19

The ten most common contributory causes of morbidity and mortality recorded alongside COVID-19 in men and women in Scotland are shown separately for 2020 and 2021 in Figure 7.6. Cardiovascular disease, COPD, non-insulin-dependent diabetes and degenerative diseases such as dementia and Alzheimer disease feature prominently across sexes. Considerable differences in these distributions are evident between 2020 and 2021, likely reflecting changes in the demographics of those affected by COVID-19. In both sexes, a reduction can be observed in the proportion of COVID-19 deaths alongside which degenerative diseases were recorded. Likely associated with this, the median age of deaths attributed to COVID-19 fell in both males and females from 2020 to 2021 from 85 to 81 in females and 81 to 77 in males.

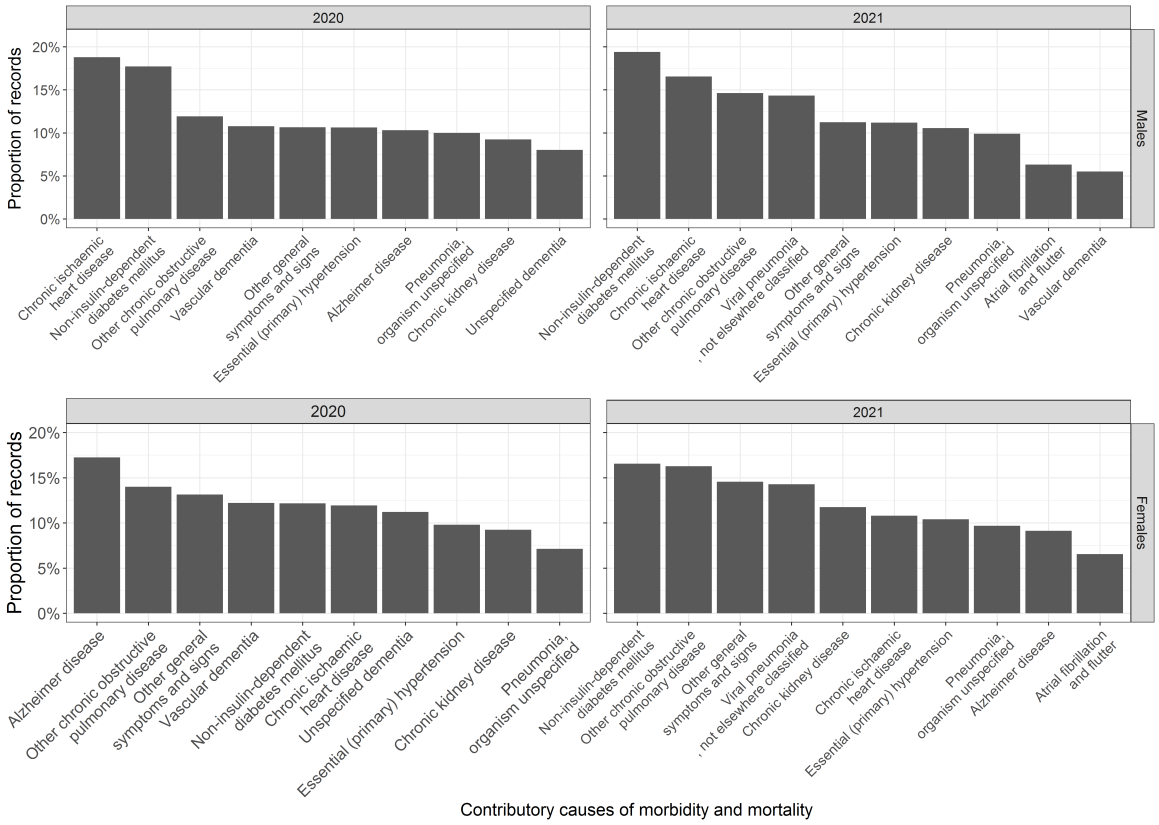


Figure 7.6: The proportion of the mortality records where COVID-19 was the underlying cause of death on which each contributory cause of morbidity and mortality was recorded (males and females separately in Scotland in 2020 and 2021).

In 2021, the most common contributory cause of morbidity or mortality to be recorded alongside COVID-19 in both males and females was non-insulin-dependent diabetes mellitus. In males, this was the result of only a small increase in prevalence from 2020 to 2021 (17.7% to 19.4%) while in females it increased from being recorded alongside 12.1% of COVID-19 deaths in 2020 to 16.5% in 2021. Contributory causes may appear alongside COVID-19 on death certificates through three primary mechanisms: either they were underlying causes of ill health present before COVID-19 infection; they were causes of ill health which occurred due to COVID-19 infection; or they were causes of ill health which were acquired alongside COVID-19 infection (for example while in hospital receiving treatment). It is not possible to determine which of these mechanisms is responsible for the prevalence of each of the contributory causes reported in Figure 7.6 however, it is important to consider that it is possible not all were present before COVID-19 infection.

7.5 Discussion

The first section of results presented in this chapter aimed to present analysis of temporal trends in the diversity of contributory causes of morbidity and mortality recorded on death certificates in Scotland from 2001 to 2021. This was followed in the second section by an examination of the contributory causes recorded alongside COVID-19 deaths and a comparison with the diversity of causes recorded alongside other leading causes. The discussion section of this chapter is structured as follows. First I summarise the findings of this chapter (Section 7.5.1) , then I reflect on the interpretation of these findings (Section 7.5.2). Next, I present a comparison to existing studies (Section 7.5.3), followed by a discussion of the strengths and limitations of this research (Section 7.5.4). Finally, I discuss the implications of the research presented in this chapter (Section 7.5.5).

7.5.1 Principal findings

In this section, the key findings of this chapter are presented, structured by the research question posed in Section 7.2 which they address.

7.5.1.1 What are the national trends in the diversity of contributory causes of morbidity and mortality in Scotland?

The diversity of contributory causes of morbidity and mortality in both men and women in Scotland was higher in 2021 than in 2001. This indicates that, over time, there was greater variation in the diseases and conditions, as denoted by ICD-10 3-character codes, that were recorded on death certificates. This occurred despite little change in the average number of causes recorded for each death, other than an increase in the mid 2010s which is suggested to have been related to the implementation of the Certification of Death Act (2011) (Scottish Government, 2011). Much of the increase in diversity occurred from 2001 to 2015 while the average number of causes changed little. This holds true under the observed distribution of mortality causes and with garbage codes redistributed to valid causes of mortality. This suggests that the contributory causes of mortality have truly become

more diverse over the period rather than any increase simply having been associated with more causes being recorded for each individual. The increase in diversity in contributory causes of mortality may however, be driven by an increase in the precision of diagnosis and mortality coding.

Despite the overall increase in the diversity of contributory causes noted in this chapter, a change in trends is observed in the years following 2015-2016 in both sexes. In females, this presents as a reduction in diversity across measures whereas among males, reductions are noted at $q = 1$ and $q = 2$ but not at $q = \infty$. This implies that following 2015, variation in the range of causes recorded in both sexes decreased and that among females the most common contributory cause was found on an increasing number of death certificates. In contrast, in males, the proportion of death certificates which contained the most common contributory causes continued to fall. It should be noted that when garbage codes are redistributed, diversity in contributory causes of mortality is shown to continue to increase following 2016 in both males and females. This indicates that in fact the share of death certificates on which the most common contributory cause appears has continued to fall across the study period. Appendix B.1.4 shows that when diversity in contributory causes of mortality is measured using a measure of diversity which explicitly accounts for zero values in the distributions of mortality causes a reduction in diversity is observed from 2001 to 2010. In the 2010s trends in diversity under this measure are shown to be similar to under normalised alpha diversity as discussed in this chapter. This means that under this measure diversity in contributory of mortality causes fell over the study period.

7.5.1.2 What is the relationship between the diversity of contributory causes and the diversity of underlying mortality causes in Scotland?

Increases in the diversity of contributory causes were found to be well correlated with increases in the diversity of underlying causes of mortality described elsewhere in this thesis. This implies that, as variation in the underlying causes of mortality faced by population increased, a more varied set of contributory causes were recorded. This implies that overall the causes of death recorded on each death certificate, both underlying and contributory, are becoming more complex

over time and likely increasingly distinctive for each individual. While this increase in diversity may be associated with an increases in the underlying variation in causes of ill health in Scotland it should be noted that it may also may be caused by an increase in diagnostic and mortality coding precision.

7.5.1.3 Were the contributory causes of mortality recorded alongside COVID-19 more or less diverse than those recorded alongside other leading causes of mortality in 2020 and 2021?

Deaths attributed to COVID-19 are shown to have been the least likely of the leading causes of mortality in Scotland in 2020/2021 to have been recorded without an accompanying contributory cause. Further to this, the diversity of contributory causes of mortality recorded alongside COVID-19 was shown to be higher than the other leading causes of mortality in these years. This was especially prominent at more conservative measures of diversity. This means that each non-COVID-19 underlying cause assessed in this chapter had a single contributory cause which was recorded as a factor in a large proportion of deaths whereas for COVID-19, no single cause was found on more than 20% of death certificates. Overall this indicates that in deaths attributed to COVID-19 a wider variety of co-morbidities were recorded compared to other common causes.

The evidence discussed above indicates that the contributory causes recorded alongside COVID-19 deaths were more diverse than those for other leading causes. The most common causes to have appeared alongside COVID-19 are also examined in this chapter. These were predominated by diseases of the cardiovascular system, COPD, degenerative diseases and diabetes. From 2020 to 2021, a reduction in the prevalence of degenerative diseases and an increase in the prevalence of diabetes is observed in both sexes. The reduction in prevalence of degenerative diseases may be due to a change in policy regarding old age care homes in Scotland from 2020 to 2021 which aimed to further protect these facilities from COVID-19 (Burton et al., 2021).

7.5.2 Interpretation

The increased diversity in contributory causes of mortality noted in this chapter indicates greater variation in the contributory causes which appeared alongside each underlying cause, meaning that the combinations of co- and multi-morbidities in the population are increasing in variety. However, over the same period, increases in the diversity of underlying causes are observed. It can be assumed that, with some overlap, each underlying cause of mortality has a set of contributory causes which are commonly recorded alongside it. Therefore, the increase in diversity in contributory causes may simply be a result of increasing variation in the distribution of underlying causes in the population.

In this chapter, reductions in the diversity of contributory causes of mortality at $q = \infty$ are observed among women in the years following 2016. These reductions were associated with increasing prevalence of the code R68 on death certificates. This code has been included among lists of "garbage codes" a term used in the literature to describe codes which are not useful for public health research (Johnson et al., 2021; Ellingsen et al., 2022). This highlights an issue of measuring diversity across all three-character ICD-10 codes as some represent broader and, arguably, less useful definitions. It may not be of value in the measurement of population health to use a measure sensitive to changes in these codes. However, it is notable that that this cause has been used so often in recent years. It appeared more than twice as often on death certificates among females in 2021 (7.7%) than in 2016 (3.6%), while in 2001 it appeared on less than 0.05% of records. The doubling of the records on which this cause appeared followed the introduction of procedures outlined in the Certification of Death Act (2011) (Scottish Government, 2011). It is unclear why an uninformative ICD-10 code would have become so common following the introduction of a policy designed to increase the accuracy of the recording of causes of death. The increase in prevalence of ICD-10 code R68, a "garbage code", on death certificates may be an indicator that this goal is not being achieved. As is observed in Appendix B.2.4; diversity in contributory causes of mortality continues to increase in Scotland when contributory causes of mortality are redistributed to valid causes of mortality. This suggests that garbage codes are masking some trends here and means that variation in the valid contributory causes of mortality faced by the Scottish population has continued to increase across the study period.

7. Diversity in contributory causes of mortality

Those whose deaths were attributed to COVID-19 in 2020 and 2021 were more likely to have had a contributory cause of morbidity and mortality recorded on their death certificate and these causes are shown in this chapter to have been more diverse than those recorded alongside other leading causes of mortality. This means that there was greater variation in the diseases and conditions recorded alongside COVID-19 than alongside other causes of death. Together, the fact that a contributory cause was recorded more often for these deaths and the higher diversity of these contributory causes, might suggest that those who died due to COVID-19 were, overall, less healthy and more likely to have an underlying health condition than in other causes of mortality. It has been proposed by sources in the literature that a number of conditions are potential risk factors for COVID-19 (Alkundi et al., 2020; Atkins et al., 2020; de Jong et al., 2021). These conditions include non-insulin-dependent diabetes mellitus, which is shown to be the most common contributory cause in COVID-19 deaths in both males and females in 2021.

Another possible explanation for the increased diversity in contributory causes in COVID-19 deaths is an increased desire for accuracy among the individuals recording causes of death. In other words, physicians and coroners may have been more likely to be thorough in recording conditions associated with COVID-19 because of its status as a novel cause of mortality. The regular scrutiny of death certificates introduced by the Certification of Death Act (2011), which involves routine review of death certificates, may make this less likely as this process is intended to increase accuracy across causes.

Finally, the higher diversity in contributory causes of mortality recorded alongside COVID-19 compared to other leading causes of mortality may be associated with the demographics of the sections of the population which were susceptible to COVID-19 mortality. The mortality rate associated with COVID-19 was high across a wide range of population groups in Scotland including for example across groups with different racial backgrounds and across deprivation gradients (Amele et al., 2023; Albani et al., 2022). The sets of morbidities which are common among these groups will differ markedly (Bhopal et al., 2011). By, in effect, drawing from across these disparate population groups with their distinct morbidity profiles, deaths attributed to COVID-19 would necessarily accumulate a diverse set of contributory causes. This effect may be amplified by the fact that COVID-19 caused more deaths among older people and across population groups, older individuals are likely to face more co- and multi-morbidities (Salive, 2013).

7.5.3 Comparison to existing studies

The only previous study of diversity in mortality causes in an MCOD context is presented by Trias-Llimós and Permanyer (2023). They use methods which assess the underlying cause of death and contributory causes together and find that increases in diversity in the USA were more pronounced with contributory causes taken into account. Their study did not assess the diversity of contributory causes without also considering underlying causes and assessed diversity using coarse-grained causes. The advantages and disadvantages of using coarse-grained diversity are discussed in Chapter 2. The difference between the methods for the calculation of diversity used in this previous study and those used here presents a more difficult comparison.

The MCOD framing used by Trias-Llimós and Permanyer (2023) highlights differences in the entire set of mortality causes faced by each individual whereas this study generalises this, examining the range of contributory causes across the population. Both approaches have merit, the approach adopted by Trias-Llimós and Permanyer (2023) specifically makes it possible to make inferences about the variety of multimorbidities which occur in the population through their focus on sets of mortality causes, whereas the methods used in this chapter provide insight into the uncertainty in the range of contributory causes each individual was likely to face regardless of which causes appear together. The increases in diversity observed in this chapter indicate increased uncertainty in the individual co-morbid conditions each individual may develop. On the other hand, the findings of Trias-Llimós and Permanyer (2023) indicate increased uncertainty in the set of multimorbidities an individual might face.

In Chapter 4, I discuss the fact that variation in causes of mortality is not a topic of great concern in the sphere of public health. Variation in contributory causes of mortality and other indicators of ill-health are, if anything, even less of a factor in public health decision making. Increasing attention is being paid to changes in the distribution of causes of ill-health, however, and alongside the research of Trias-Llimós and Permanyer (2023) discussed above, Dyson et al. (2021) have presented a study using the Gini coefficient to examine the cause distribution of DALYs in regions around the globe. While these studies aim to examine different aspects of health, with the first focusing on MCOD and the second on the causes of morbidity and ill-health during life, they are both concerned with assessing variation in

health outcomes other than the underlying cause of death. The results of these studies and of this chapter are comparable, indicating increased variety in causes of contributory causes of morbidity, mortality, and ill-health during the 21st century. This has the potential to impact decision-making in the resourcing of health care and public health initiatives as more conditions are occurring more commonly and must be addressed. These risks remain little-discussed in the public health policy literature though evidence of this chapter and previous work suggest that the study of variation in causes of mortality and ill-health may aid in understanding areas and populations where these risks manifest most strongly (Trias-Llimós & Permanyer, 2023; Dyson et al., 2021).

7.5.4 Strengths and limitations

The results presented in this chapter suggest that measurements of diversity in contributory causes may be biased by recorded causes which appear commonly but have little clinical significance. An example of this is noted previously in the effect of the high prevalence of cause R68 on diversity in contributory causes in women in the late 2010s. This issue is difficult to address, previous studies of diversity in mortality causes attempt to solve it through grouping causes broadly by ICD-10 chapter. I show through a simple redistribution method in Appendix A.19 that the trends in diversity in contributory causes of mortality reported in this chapter are also present up to 2016 with with garbage codes accounted for and redistributed to valid codes. Following 2016, with garbage codes redistributed, diversity in contributory causes continues to increase while in the observed distribution discussed in this chapter a reduction in diversity is observed. This confirms that the presence of garbage codes has impacted the trends in diversity and as is discussed above means that variation in the valid contributory causes of mortality recorded in Scotland has continued to increase across the study period.

Some in the literature suggest that a source of uncertainty in the study of mortality causes is the accuracy of their recording (O'Malley et al., 2005; Koetsier et al., 2012). Despite standard national coding practices some regional differences are though to occur (Lanska & Peterson, 1995). Further variation has been proposed between the individual medical practitioners who record mortality causes (Lu et al., 2000; Anderson, 2011; Danilova, 2016). These factors have the potential to bias the analysis presented in this chapter and may be more acute in the case

of contributory causes where a greater degree of individual judgement between practitioners may be involved in determining which causes merit inclusion. There is likely no way to avoid this potential bias in research such as is presented in this chapter but it should be considered as a factor.

In this chapter, normalised alpha diversity was calculated over a distribution of contributory causes of mortality extracted from life tables. Through this process, deaths which had no contributory cause recorded alongside them were in effect ignored. This presents a limitation of this approach as it means that a case which occurs with some regularity, a death being recorded with only the underlying cause, is not accounted for. In future research it would be valuable to compare the results reported in this chapter against those produced using a measure which explicitly accounts for these cases.

As noted in Section 7.4.1 and shown in Appendix B.1.4, measuring diversity in contributory causes of mortality causes using a measure of diversity which explicitly accounts for 0 can produce different results to those reported in this chapter. In Appendix B.1.4 I only analyse overall trends in diversity in contributory causes of mortality in Scotland. Given the disparate trends found in this analysis the use of an alternative measure of mortality causes may affect other conclusions drawn in this chapter. When measuring diversity it is important to consider the characteristics of the measure of used. In future analysis, it may be considered advantageous to use a measure which treats zero values differently to those used in this thesis, in which case the alternative measure of diversity discussed in Appendix B.1.4 may be of use.

7.5.5 Research implications

This chapter presents this first analysis of diversity in the contributory causes of mortality in isolation. The increasing diversity of contributory causes found in this chapter adds to the evidence, presented in this thesis, that the conditions, diseases and causes of ill-health facing the Scottish population became more varied from 2001 to 2021. Confirmation that diversity in contributory causes of mortality has increased alongside diversity in underlying causes in Scotland in these years has implications at the societal level. Treatment and prevention of causes of morbidity and mortality are necessary no matter where on the death certificate a cause of

ill-health appears. Therefore, at the population level, the cost to healthcare and public health systems associated with the increased range of contributory causes is likely to be just as, if not more, significant than those required to address the increasing range of underlying causes (Carreras et al., 2013).

Chapter 8

Discussion

8.1 Chapter overview

Chapter 8 brings this thesis to a close. I begin this chapter with a summary, in Section 8.2, of the key findings and conclusions of this thesis. Following this, in Section 8.3, I interpret, discuss and contextualise the implications of the findings of this thesis. I reflect on the strengths and limitations of the research presented in this thesis in Section 8.4. In Section 8.5 I give suggestions for future research directions for the field of diversity in mortality causes and offer two short case studies on possible avenues for methodological development. Finally, I end this chapter with a conclusion summarising this thesis as a whole (Section 8.6).

8.2 Summary of key findings

In summary, this thesis has shown that diversity in the underlying and contributory causes of mortality recorded in Scotland had increased in 2021 relative to 2001; though trends over the study period have fluctuated somewhat. These increases are shown to have occurred more rapidly in males when rare causes of mortality are accounted for but at a similar rate when considering the most common causes. Further, trends are shown to be weaker when zero values in the distribution of mortality causes are considered. Overall, I suggest that increasing diversity in mortality causes, while indicating positive trends at the level of

the most common causes, may present a challenge for public health and health care systems. Diversification indicates that diagnostic uncertainty may increase as physicians encounter a greater variety of causes with potentially overlapping symptoms. In addition, healthcare and public health systems must spread limited resources further in order to make gains in population health while economies of scale lose their power. In later years of the study period, especially following 2015, I show differing trends in diversity in causes of mortality as increases slow in females and reverse in males. This is potentially worrying as it implies adverse trends in the prevalence and rate of mortality associated with the most common causes. A reduction over time in diversity in causes of mortality indicates other challenges for public health systems. It indicates an increased proportion of mortality due to a smaller number of common causes and means the burden of these causes on healthcare systems may increase. This increased health care and public health burden has the potential to be especially challenging as the very fact that these already prominent causes are becoming proportionally more common indicates that public health and medical systems are struggling to address them.

I have shown that the causes of mortality which have driven increasing diversity in underlying causes of mortality differ considerably between different subpopulations in Scotland by age, area-level deprivation and urban-rural class. However despite this, trends in diversity are similar across these subpopulations. This means that despite differing disease burdens, changes in the patterns of mortality occurred in similar ways. Across subpopulations, diversification means an increased variety of causes must be accounted for at the individual and population level meaning the adoption of a more holistic approach in public health and medical care. However, this more holistic approach cannot become too general and there must be a focus on specialisms in different diseases and conditions in different subpopulations.

I present evidence that higher life expectancy and lower lifespan diversity over time are linked to a greater diversity in causes of mortality. However, I also demonstrate that subpopulations with higher lifespan diversity cross-sectionally exhibit a higher diversity in causes of mortality, under less conservative mesaures of diversity. These findings support existing hypotheses about the relationship between lifespan diversity and diversity in causes of mortality, although my conclusions differ slightly from previous research. Specifically, I propose that longer lifespans and greater diversity in mortality causes have both been driven by the same patterns of mortality, rather than by an older population with more homogeneous ages at death causing an increase in diversity of mortality causes. I argue that reduc-

tions in the rate of mortality of the most common causes of death, especially at younger ages, led to fewer premature deaths; an increase in life expectancy; and a reduction in lifespan diversity, as well as an increase in the diversity of mortality causes. No matter the reason for these trends, the increase in diversity of causes of mortality alongside a population living to older, more homogeneous, ages has complex implications. On a positive note, it implies that, while the population faced increasingly varied causes of death, total inequalities in health outcomes across the population slightly reduced. However, any potential gains to efficiency in healthcare and public health caused by deaths at more homogeneous ages may be counteracted by individuals facing a wider variety of causes of mortality.

Finally, in this thesis I have shown that despite myriad proposed cause-specific impacts of the COVID-19 pandemic, diversity in causes of mortality in Scotland with COVID-19 excluded from analysis in the years 2020 and 2021 was within the expected ranges of counterfactual projections. This is shown to be true in deaths across different age groups and furthermore, the contribution to diversity of each ICD-10 Chapter is shown to have changed little in 2020 and 2021 compared to previous years. I conclude that, overall, variation in causes of mortality was affected little by the pandemic and that trends in variation in mortality causes were relatively, and remarkably, robust to its influences.

8.3 Interpretation and research implications

In this section, I will discuss the findings of the research conducted in this thesis and interpret their implications within the context of public and population health literature. This section covers five themes which emerge from the research in this thesis:

The first theme addresses the interpretation of changes in the diversity of mortality causes and their implications for population and public health (Section 8.3.1). In this context, I also discuss how the use of fine-grained cause methods, as opposed to coarse-grained cause methods, can influence the study of diversity of mortality causes. Finally, within this theme, I examine the impact of the increasing prevalence of external causes of mortality on the study of diversity in mortality causes in Scotland.

The second theme focuses on the implications of the research presented in this thesis for public health specifically in Scotland, both during the study period (2001 to 2019) and in the future (Section 8.3.2). I conclude this theme with a discussion of how diversity in causes of mortality may factor into public health discourse around future burdens on the UK health system.

The third theme explores the relationship between diversity in causes of mortality and lifespan diversity as it relates to epidemiological transition theory (Section 8.3.3), drawing on the findings of Chapter 5.

The fourth theme discusses subpopulation-level differences in diversity in mortality causes (Section 8.3.4), examining why such differences may manifest.

The fifth and final theme (Section 8.3.5) centres on the COVID-19 pandemic and how the findings of this thesis may improve our understanding of its impact.

8.3.1 Implications of changes in diversity in mortality causes for population and public health

Measuring diversity in causes of mortality introduces an interesting question in analysis of population health. Diversity measures are fundamentally sensitive to the proportion of deaths due to a particular cause, rather than to absolute numbers (or rates), therefore: is it better for the rate of mortality of a disease to decrease or for it to cause a smaller proportion of deaths? The obvious answer to this might be both; however, for one cause of death to reduce in proportion, at least one other must, by definition, increase. Given this fact it could be argued that it is beneficial to achieve improvements in health by the reduction of all-cause mortality rates through equal proportional reductions in the rate of mortality associated with all causes of death. For example a 10% reduction in all-cause ASMR might be achieved through a 10% reductions in the ASMR of each individual cause of death. If this were to occur, diversity in mortality causes would remain relatively constant as the distribution of mortality causes shrinks across the board but the proportion of individuals who die of each cause does not change. This is, of course, not an outcome which is likely to occur in reality. Furthermore, it is not necessarily a desirable outcome. It is likely more beneficial for population health and society

in general to reduce the mortality rate and proportional prevalence of causes of mortality with the highest societal cost (whether that be financial in medical and public health settings or through causing a greater degree of suffering and morbidity). This likely means reducing mortality due to common causes of mortality; for example, heart disease and lung cancer the most common causes of death in Scotland have significant societal costs (Lawson, 2013; Jäkel et al., 2013; Liao et al., 2008). It may therefore be the case that the greatest improvements to population health might come alongside increasing diversity in mortality causes.

The consequences of changes in the diversity of causes of mortality are closely linked to uncertainty at the individual-level and resourcing at the population-level. Increasing diversity in mortality causes, as I have observed in Scotland when measured using normalised alpha diversity, can, generally, be understood as an increase in uncertainty in the causes of mortality each individual is likely to face and increased variation in causes of mortality across the population as a whole. Such trends are often driven by reductions in the prevalence of the most common cause or causes of death. These causes become less likely to be the reason an individual has died and consequently, other causes become *relatively* more common. Many of the improvements to overall health outcomes across the world can be traced to reductions in mortality associated with the most common causes of death (Omran, 1971; Santosa et al., 2014). There are a number of explanations for these reductions in mortality. These causes attract a great deal of attention from researchers, policymakers and public health officials, as well as resources from funding bodies (Carter & Nguyen, 2012; Stark & Shah, 2017). The combined impact of this attention has been linked to reductions in mortality rates attributed to the common causes of death, particularly diseases of the cardiovascular system. However, in addition improvements in health in developed nations in the latter half of the 20th Century have been linked to a variety of other influences including shifting employment arrangements and improving living conditions as well as changes in other social determinants in health (Marmot, 2017; Walsh et al., 2010b). Evolving patterns of tobacco smoking and obesity in Scotland and other developed nations have also been associated with changing mortality rates due to common causes of death such as cancers and diseases of the cardiovascular system (Hotchkiss et al., 2014).

At the population level, increasing diversity in causes of mortality means that in order to make improvements in health practitioners must address a broader variety of diseases and conditions. This wider purview is described by Bergeron-Boucher et al. (2020) as a more holistic approach, meaning that the entire distribution of causes must be considered. This more holistic approach may, however, be substan-

tially more difficult and more costly to implement, because previously effective economies of scale in health treatment are likely to be lost. As a result, in addition to the need to distribute resources across a wider range of causes, a loss of efficiency will cause these resources to have less of an effect. As discussed in Section 2.4.2, prevention strategies in public health are likely to be far less impacted by increasing diversity than more direct treatments. This is primarily because interventions which aim to prevent the occurrence of ill-health have the advantage of being wide-ranging and can affect the up-stream causes of a large number of diseases and conditions at once. In Section 2.4.2 I use the example of campaigns to reduce the tobacco smoking which, in Scotland and around the world, have reduced the prevalence of a risk factor which has been linked to: heart disease, a number of cancers, and strokes. Reducing the prevalence of smoking in the population has played a large part in reducing the mortality risks associated with each of these causes of mortality (Critchley & Capewell, 2003). In this way prevention and upstream public health interventions may be insulated from increased fragmentation in causes of ill-health. The utility of interventions which focus on prevention also means that the holistic approach described above should not come at the expense of existing campaigns. Were a shift in focus to mean that causes of mortality and ill-health which have become relatively rare in our society through the successes of public heath campaigns become more common it could have grave consequences for population health. It is therefore, important that the impact of increased diversity in causes of mortality on population health be considered in the context of overarching population trends.

At the individual level, increased diversity in causes has been proposed to have the effect of increasing uncertainty in diagnosis as a wider variety of diseases and conditions become common and physicians become more likely to come across conditions which were previously rare (Bergeron-Boucher et al., 2020). This increased diagnostic uncertainty may cause an increased rate of misdiagnosis which has the potential to result in a reduced quality of care for individual patients. However, there may be reason to question this interpretation. One explanation for increasing diversity in mortality causes may be increasing diagnostic accuracy. As discussed in Chapter 4, if diagnostic accuracy is causing diversity in mortality causes to increase, it is unlikely that increased diversity in mortality causes will make diagnosis more difficult. Furthermore, increasingly diagnostic science relies on computer assistance which is likely less vulnerable the impact of more diverse causes of mortality on diagnostic uncertainty (Dilsizian & Siegel, 2013). Despite this it may be possible that diversification in mortality causes could impact certainty in physician led diagnosis.

In a situation where all-cause mortality rates and life expectancies are improving alongside increasing diversity in mortality causes, I suggest the ramifications for policy and practice are mixed. On the positive side, increased diversity in mortality causes likely indicates a smaller proportion deaths have occurred due to the most common cause and that the rates of mortality associated with these causes have fallen faster than all-cause rates. This suggests that either public health campaigns or medical interventions/treatments have been effective in their aims of reducing mortality related to these causes. I show that increases in diversity in the 2000s were driven by reducing rates of acute myocardial infarction. Reducing heart disease mortality is a long-standing public health goal in Scotland, and campaigns with this aim were, to an extent, successful in this period (Critchley & Capewell, 2003; Critchley et al., 2003; Hotchkiss et al., 2014). However, as discussed in previous paragraphs; increased diversity in causes of mortality has some negative implications for healthcare and health improvement even if it comes alongside general improvements in other mortality measures. These potential individual-level and population-level consequences, indicate increasing pressure on medical care and heath promotion systems as well as an increased burden of uncertainty on medical practitioners.

I suggest that the reductions over time in the diversity of mortality causes observed after 2016 among males in Scotland indicate potential challenges for public health and healthcare systems, whether overall mortality rates have increased or decreased. This is because falling diversity in mortality causes likely indicates that the most common causes of death are becoming, proportionally, more common. There are three likely outcomes indicated by falling diversity in mortality causes, each of which is associated with the most common causes increasing in prevalence: 1) mortality rates associated with the most common causes have increased while those associated with other causes have plateaued or fallen, 2) a new and highly prevalent cause of mortality has arisen, or 3) mortality rates associated with the most common cause or causes have changed little but improvements have been made across the rest of the distribution of causes (therefore, the most common cause has become proportionally more common). The first of these outcomes is observed in males in Scotland following 2016, among whom mortality associated with acute myocardial infarction has increased while the mortality rates of other causes of death have not changed to the same extent. The second outcome is observed during the COVID-19 pandemic in Scotland, where the large number of COVID-19 deaths caused a reduction in the diversity of mortality causes. It is alternatively possible that falling diversity in mortality causes could be caused by

a simple reduction in the number of causes which are recorded in a population. However, this is unlikely to occur without being associated with either the first or third outcome described above, or a change in policy or practice around mortality recording or healthcare.

A reduction in diversity in mortality causes is, therefore, likely to be associated with worrying trends in the prevalence of the most common causes of mortality: the most common existing causes are becoming more common, a new common cause has appeared, or the most common causes are lagging improvements in other causes. In each of these cases, the greatest gains in population health are likely to be made by addressing the most common causes. While other causes should not be ignored, the more holistic approach needed to address diversification in causes of death may be less beneficial to direct improvements in overall population health.

To simplify the arguments of this section regarding population-level effects, changes in the diversity of mortality causes might be thought of from a health economics perspective. Increasing diversity in causes means resources for healthcare and health promotion must be spread wider to make gains in health. In contrast, falling diversity likely means that greater improvements are possible by a focus on the most common causes. However, to understand the implications of changes in diversity in mortality causes it is necessary to examine the overall context of population health in which such changes have taken place.

8.3.1.1 Measuring diversity in causes of mortality at the fine-grained level

The potential consequences of increasingly diverse causes of mortality on population health and health care systems have been touched upon in research and commentary from a number of perspectives, which will be discussed further in upcoming sections. Despite this, formal measurement of variation in causes of morbidity and mortality in the literature is, as I have discussed throughout this thesis, limited. A key methodological difference separates the work of this thesis with most studies carried out in the 21st century. This disagreement stems from the definition of a cause of mortality. In this thesis I have presented evidence of diversity in mortality causes measured across fine-grained ICD-10 three character codes. This stands in contrast to other studies of diversity in mortality causes in recent

years which have examined coarse-grained causes, generally grouped together by ICD-10 Chapter (Bergeron-Boucher et al., 2020; Trias-Llimós & Permanyer, 2023). I have discussed the relative merits of these approaches previously (Section 2.4.3) however, it is a point which warrants further discussion here.

A key argument which I raise in Section 2.4.3 for the use of fine-grained cause methods is the impact of differential degrees of variation within groups of causes. That is to say that, for example, the diversity of the fine-grained causes of mortality within ICD-10 Chapter II: Neoplasms is significantly greater than that of the fine-grained causes within ICD-10 Chapter VI: Diseases of the nervous system (This can be observed in analysis of monthly diversity within ICD-10 Chapters in Chapter 6). In grouping causes at the level of Chapters this variation is lost.

The evidence I have presented in this thesis provides further support for the use of fine-grained cause methods. I have shown that, within ICD-10 Chapters, fine-grained causes of mortality have exhibited varied trends in their contribution to overall variation in the distribution of mortality causes. Perhaps a key example of this is among cancers, contained within ICD-10 Chapter II. Bergeron-Boucher et al. (2020) present evidence of changes in the proportion of deaths caused by each group of causes addressed in their study. They suggest that in the UK from 1994 to 2017 "neoplasms" (all of ICD-10 Chapter II) were among the groups of causes which saw the largest increase in the share of deaths they caused. However, I have shown that underlying this, in Scotland, the dynamics of change in the prevalance of individual cancers was more complex. From 2001 to 2019, lung cancer, the most common cancer and one of the most common causes of death overall in both sexes, reduced in prevalence. This meant it contributed to increasing diversity, as a reduction in dominance in the most common causes increases the evenness of the distribution of mortality causes. However, concurrently a number of other relatively less common cancers have increased in prevalence, also contributing to increased diversity. Overall, while strategies to reduce lung cancer mortality rates have been successful a variety of other forms of cancer have become, proportionally, more likely to occur as a cause of mortality. Not only does this imply diversification of causes *within* Chapter II but it means screening and treatment planning for cancers must take the holistic approach discussed previously. The implications of this diversity in mortality causes within-chapters or within-groups are clearly important and should be accounted for in studies of variation in mortality causes.

There are drawbacks to the use of fine-grained cause diversity, some of which I have discussed in Chapter 7. By including all fine-grained causes of mortality, garbage codes (for example, R68: "Other general symptoms and signs", discussed in Chapter 7) may have an impact on the measurement of diversity. Although analysis in Appendix B.2 suggests that in most cases the results presented in this thesis are robust to the effect of garbage codes. Further, it is possible that multiple codes describing similar causes of mortality which may be considered broadly homogeneous by medical professionals are treated as completely separate. Therefore, while I have advocated for the use of fine-grained cause methods in the calculation of diversity it may be most advantageous to adopt a compromise position as discussed in Section 2.4.3. This could mean causes grouped using a systematic rubric for the classification of diseases at a more granular level that ICD-10 Chapter but at a higher level than the fine-grained causes examined in this thesis. Another possibility would be the use of similarity-sensitive measures of diversity an approach discussed later, in Section 8.5.

8.3.1.2 External causes of morbidity and mortality

The implications of diversification in mortality causes I discuss in this thesis - fragmentation in causes of ill-health and increased diagnostic uncertainty - mostly take a focus on medical aspects of health. They consider the prevention, diagnosis, and treatment of diseases and conditions in this medical context (Bergeron-Boucher et al., 2020; Trias-Llimós & Permanyer, 2023). However, one of the key drivers of diversification noted in this thesis, especially in more deprived areas and urban areas, are causes within ICD-10 Chapter XX: External causes of morbidity and mortality. Violent deaths; causes of mortality associated with alcohol and drug abuse; and suicides have increased in prevalence and mortality rate and have been key to the increasing contribution of Chapter XX to diversity in mortality causes. The implications of diversity described above do not necessarily apply to these external causes of mortality in the same way: their increased prevalence is not likely to lead to increased diagnostic uncertainty as they often occur outside of medical institutions and present differently to various other causes of ill-health (Augarde et al., 2022). Despite this, external causes of death impose a significant burden on the healthcare systems in Scotland, and this burden may intensify if the prevalence and mortality rate of external causes continue to rise.

Drug overdoses, suicide attempts and violent actions such as knife injuries can require significant medical interventions. Much of the focus on reducing mortality due to these causes is through preventative interventions as well as through preventative treatment such as rehabilitation and psychological interventions (Rehder et al., 2021). Efforts spread beyond traditional public health spheres with prevention campaigns including the actions of the police, civil society groups, and other organisations. This means that while the increases in diversity driven by external causes might not be expected to have a significant effect on diagnostic uncertainty, their potential impact on population level resource costs are substantial.

In Scotland, and across the world, increases in the rate and prevalence of deaths of despair, and external causes of morbidity and mortality more widely, have been associated with rising economic instability (Guimarães, 2021; Beseran et al., 2022; King et al., 2022). In Scotland this has been linked to economic isolation associated with neo-liberal de-industrialisation policies, especially in the 2000s, and austerity policies in the 2010s (Allik et al., 2020; Walsh et al., 2010b; DeVerteuil, 2022).

Both upstream and downstream solutions to rising mortality due to external causes have been proposed. Many downstream programmes aiming to, for example, make drug use safer and reduce street violence are in place in Scotland and across the UK (Leyland & Dundas, 2010; Matheson et al., 2014). The observed effects of these interventions have been mixed (McAuley et al., 2012). This type of intervention is likely particularly vulnerable to cost-cutting exercises during periods of austerity, as well as as a result of increased diversity in mortality causes: as resources must be spread wider, fewer resources may be available to individual bespoke programmes.

Upstream proposals, for example wholesale revisions of drug legality and policing, although they are more difficult to achieve, may more adequately address rising diversity as they have the potential to reduce mortality from a range of causes of death at once (Atkinson et al., 2019; Dalgarno et al., 2021). Simultaneously, the most radical, obvious, and widely agreed upon solution to reduce mortality from causes whose rise has been linked to economic instability, including deaths of despair, is to reduce economic instability (Walsh et al., 2020a). This is a purported aim of many governing political parties around the world, although evidence for their success is mixed (Asteriou & Price, 2001). Despite the adverse health outcomes observed in Scotland in the late 2010s having come alongside falling diversity in mortality causes, it is still important to consider interventions which

have wide-ranging effects in terms of the causes they address, even those which are aspirational such as improving socioeconomic positions. I note in Chapter 2 that public health interventions targeted at prevention are likely to be least affected by increasing diversity in mortality causes. The interventions discussed in this paragraph are key examples of how prevention campaigns may also be valuable in combating increased diversity in causes of mortality. By addressing the root cause of wide range of causes of ill-health, interventions which improve socioeconomic conditions or other social determinants in health can address increasing the increasing variety of causes of mortality observed in this thesis. This is in addition to other improvements to health outcomes achieved through such interventions.

8.3.2 Implications of the findings in this thesis for population health in Scotland

In the 2000s, opposing trends are observed in lifespan diversity and diversity in causes of mortality in Scotland. Lifespan diversity in the population fell, only slightly in females, but to a greater degree in males, while diversity in causes of mortality, measured using normalised alpha diversity, increased. As discussed in Chapter 5, this has potentially contradictory implications. Reduced lifespan diversity suggests greater equality in age at death, potentially indicating progress towards more equitable health outcomes. In practical terms, healthcare at the time of death can be provided for individuals at more similar, and mostly older, ages. It means that it may be beneficial to provide greater funding in social care services, old age care homes, and geriatric care in hospital. However, counteracting this, I have shown increasing diversity in causes of mortality, both underlying and contributory, which indicates that a greater variety of diseases and conditions must be catered for. In this thesis I have shown that from 2001 to 2010 the drivers of diversification in causes of mortality were falling rates and prevalences of cardiovascular diseases alongside increasing prevalence of a range of cancers, external causes and degenerative diseases. Risk factors of these varied causes must be addressed in the population and treatment for them must be provided which will probably stretch already limited resources.

Life expectancies continued to increase and mortality rates continued to decrease in Scotland in the early years of the 2010s, although these increases were considerably slower than in previous years (National Records of Scotland, 2022a). The results of this thesis have shown that lifespan diversity fell and inequalities between subpopulations in this measure narrowed until around 2015, indicating overall improvements in health outcomes and continued improvements in health equity. However, from the mid-2010s onwards, improvements in many measures of population health have stalled or reversed in Scotland (Ramsay et al., 2020). The reasons for these stalling improvements are not fully understood; however, in the literature, they have been strongly linked to austerity policies introduced by the government of the United Kingdom in the early 2010s (Walsh et al., 2022). These policies have particularly been linked to increases in the rate of deaths of despair as well as other causes which are associated with deprivation and economic instability (Richardson et al., 2020). Among other effects, these policies are thought to have reduced the standard of living of large sections of the population and made accessing health-improving influences more difficult (McKendrick et al., 2016; Garthwaite & Bambra, 2017; Holland, 2017; Macdonald & Morgan, 2021).

Alongside stalling improvements in population health, I find in Chapter 6 that from 2015 onwards, diversity in causes of mortality fell in males across Scotland. In Chapter 4, I observe that following 2015 the prevalence and mortality rate of acute myocardial infarction increased, following years of decline. It has been suggested, in the literature, that this increase in rate is associated with the effect of austerity policies (Walsh et al., 2020b). Among females, increases in diversity in mortality causes slowed across measures of diversity after 2009. Lung cancer and degenerative diseases are found to have been the most common causes of mortality in females in the latter half of the 2010s. Mortality rates associated with these causes are shown in Chapter 6 to have remained relatively constant in this period. This period of relative stagnation with regard to the most common causes led to the trend of slower increases in mortality cause diversity among females. In both males and females, these effects later in the study period are observed more strongly at more conservative measures of diversity (q values of 2 and ∞). As suggested in Section 8.3.1, this implies a greater focus is needed on the most common causes in Scotland. I show in Chapter 6 that reductions in diversity among men following 2016 were less pronounced at less conservative measures. Similarly increases to diversity in mortality causes among women were stronger under less conservative measures of diversity. This indicates that further

diversification at the level of relatively rare causes may be occurring alongside, and slightly counteracting, adverse trends in the most common causes. Further evidence of this is the increasing additive value of cancers, deaths of despair and other external causes of morbidity and mortality (ICD-10 Chapter XX) in this period.

Together, these trends in mortality cause diversity mean a greater variety of relatively uncommon causes of mortality has been observed at the same time as adverse (static or increasing) trends at the level of the most common causes. To address this, renewed efforts to reduce mortality associated with the most common causes of mortality are needed but must come alongside a recognition of the widening variety of causes which the population faces. These recommendations and the work of this thesis are predicated on the study of causes of mortality. I do not assess the impact of each cause of mortality on society more generally whether this be through the burden each cause exerts on health and social care systems or through the impact of each cause on the length and quality of life among the public. It is generally accepted in the literature that in order to maximise population health these factors must be taken into account (Young, 2004; Arah, 2009; Jones et al., 2019). Therefore, while increasing diversity in mortality causes may pose challenges for public health and healthcare systems; it is not necessarily true that diversification itself must be addressed to improve population health.

The trends in diversity in causes of mortality discussed to this point were measured under normalised alpha diversity under the Reeve et al. (2016) framework. In Appendix B.1, I examine alternative measures of diversity finding that under uniformity - a measure related to measures of evenness which explicitly accounts for zero values in the distribution of mortality causes - there is no clear trend in diversity from 2001 to 2015, thereafter diversity increases up to the year 2019. Under this measure diversity in mortality causes in Scotland is found to have increased in 2019 compared to 2001 but the near continuous increases from 2001 to 2015 found under the measures discussed above are not observed. The use of this alternative measure of diversity is discussed further in Section 8.4.1.1. The findings discussed in this section should be considered in the context of this alternative measure of diversity under which trends are found to differ somewhat.

Alongside the potential for undesirable consequences arising from increasing diversity of mortality causes, in this thesis I show increasing lifespan diversity in the latter years of the 2010s across Scotland in both sexes. Furthermore, I find increasing inequalities in this measure between subpopulations after more than

a decade of improvements. This continues long-term increases in deprivation-related inequalities in variation in lifespan in Scotland which had stalled or slightly reversed in the 2000s (Seaman et al., 2016a; Seaman et al., 2019). It means greater inequality in age at death across Scotland, between subpopulations and within more deprived areas and urban areas.

Alongside diversity in underlying causes of mortality, in this thesis I have presented trends in the diversity of contributory causes recorded in Scotland. I show that diversity in contributory causes of mortality has a strong relationship with diversity in underlying causes of death in Scotland. Previous research has not made these comparisons directly. Concurrent increases in these measures have the potential to compound on each of the implications of trends in the diversity of causes of mortality discussed in this chapter. An increased diversity in underlying causes of mortality suggests that the diseases and conditions which must be treated and prevented in a population have become more varied; additional to this, the contributory causes of morbidity and mortality recorded on a death certificate must also be treated through medical care or addressed by public health interventions. Increasing prevalence of co- and mutli-morbidities is raised as a concerning trend in public and population health literatures (Grundy & Stuchbury, 2022; Finucane et al., 2021). My evidence suggests that not only must health care systems in Scotland prepare to mitigate an increased rate of multimorbidity. They must acknowledge that the recorded contributory causes of mortality have become more varied over time and therefore, a more varied set of multimotbidities may need to be treated. The analysis I present in Chapter 7 does not explicitly measure the diversity of multimorbidities; that is, I do not explore how commonly each set of multimorbidities appears together [1]. Despite this, my evidence points to an increase in variation in the conditions and causes of ill-health at the time of death across the population of Scotland.

The cost of ongoing treatment and care for chronic diseases and morbidities in the developed world is significant (Carreras et al., 2013). The contributory causes of mortality recorded on death certificates give an indication of morbidities but do not directly measure them as they only assess conditions which have hastened or been a factor in an individuals death. Despite this, if the wider variety of contributory causes is indicative of greater variety in the burden of morbidity in the population, there may be even greater need for medical care in future than trends

[1] This analysis has been performed by Trias-Llimos and Permanyer (2023) showing an increase in diversity in the sets of multimorbidities recorded at the time of death in the USA over the 21st century.

in the underlying cause of death imply (Brown et al., 2001; Lehnert et al., 2013). Furthermore, it is possible that the increase in diversity in contributory causes noted here was, at least in part, been driven by increasing diagnostic precision. Even with these caveats, the results presented here are an empirical indicator that the diseases, conditions, injuries and morbidities recorded at the time of death in the Scottish population have become increasingly diverse, above and beyond increasing diversity in underlying causes of death.

8.3.2.1 Variation in mortality causes and the burden on the health-care system

I have shown, through study of diversity in mortality causes, that the causes of mortality and ill-health faced by the Scottish population have become more varied from 2001 to 2019. An increasingly varied burden on the healthcare system in the UK has been discussed in the public health literature as a potential future challenge. Thanks to improved living conditions and health care systems individuals are living to longer ages and, crucially, are living for longer with chronic - but treatable - conditions (Salive, 2013; McKee et al., 2009; Kingston et al., 2018).

Mckee et al. (2021) suggest that increased multi-morbidities in the UK require a "holistic approach delivered by multidisciplinary teams". The findings and recommendations of this thesis support this conclusion. Increased diversity in causes of mortality, both underlying and contributory, should be addressed through a more holistic approach by health care and public health professionals. Advantages and disadvantages are offered by the specific health care context within the UK for addressing the increasingly varied causes of ill-health and multimorbidity. The UK has a strong generalist primary care approach which is well-suited to addressing increased variation in the disease burden, however, Mckee et al. (2021) note that: "the disadvantage, in a country that has somewhat fewer medical specialists than many others (despite growth in specialist posts in recent years), is that it might be difficult to obtain specialist expertise when needed". A number of studies have suggested that in future multi-morbidities are likely to become increasingly common and the research of this thesis suggests these multi-morbidities may fur-

ther have increasingly varied causes (Foreman et al., 2018; Kingston et al., 2018). Measuring diversity in causes of mortality and ill-health provides a way to assess variation in the burden of disease on health care systems and may be of value to future discussion of public health policy.

After 2015, among males in Scotland, an increase in the prevalence (proportion of deaths) associated with acute myocardial infarction, the most common cause, has occurred. As I have noted this has potentially contrary implications to increases in diversity in previous years. The generalist approach of the UK health care system may be invaluable in combating such changes in trends in mortality cause diversity. This system offers flexibility to treat both the most common causes and relatively rare causes through a generalist first point of contact (Kringos et al., 2013). However, due to funding and staff cuts, and especially following the COVID-19 pandemic, access to primary care in the UK has become increasingly restricted (Douglas et al., 2020; McKee et al., 2021). In future more specialist health care resources may be needed in the UK to address the increasing variety of causes of ill-health and mortality. This must, however, be accompanied by a competent and well-funded primary care service able to respond concurrently to potential resurgences in the most common causes.

8.3.3 Lifespan diversity, diversity in causes of mortality and the epidemiological transition

In Chapter 5, I examine and discuss the relationship between lifespan diversity and diversity in causes of mortality. I show that over time this relationship was negative and increasing diversity in causes of mortality was associated with falling lifespan diversity. I suggest that both falling lifespan diversity and increasing diversity in causes of mortality were driven by a reduced prevalence of the most common causes of mortality, which reduced in rate most acutely at younger ages. Therefore, the dynamics of the relationship between lifespan diversity and diversity in causes of mortality was closely linked to the prevailing mortality patterns in Scotland in the study period. It is conceivable, therefore, that this relationship has differed in the past and might change again in the future as the population undergoes epidemiological transition.

The theory of epidemiological transition, described in detail in Section 2.2.4, aims to explain changes in mortality patterns and population structures over time. Improving living standards; medical care; and public health systems drive these changes. Many studies have examined and theorised the link between the epidemiological transition and variation in lifespans. The early stages of the epidemiological transition were associated with steady and significant reductions in variation in lifespans from the late 1800s to around the 1950s (Edwards, 2010; Shkolnikov et al., 2003; Wilmoth & Horiuchi, 1999). These significant reductions in lifespan variation lead the field of demography to, broadly, adopt theories regarding progressions in human mortality which postulated that "mortality compression", the concentration of deaths around the modal age at death, would continue indefinitely until the upper limit of human lifespan was met (Permanyer & Scholl, 2019). This is also known as the "rectangularisation hypothesis" which suggests that as mortality compression occurs and more individuals reach ages closer to the upper limit, survival curves[2] would gradually become rectangular (Fries, 1980). At the point of rectangularisation all deaths occur at the modal age, meaning the probability of survival is 100% at all other ages and that there is no variation in lifespans. Progress towards rectangularisation was observed for much of the 20th century, though following 1950 stalls in reductions in lifespan variation began to be observed (Permanyer & Scholl, 2019). Some aspects of the rectangularisation hypothesis are contested, especially as no upper limit to human lifespan has been met and life expectancy has increased near-continuously in most nations (Robine, 2001). Despite this it has been fairly consistently shown across the world that increasing life expectancies are associated with falling inequalities in lifespan (Permanyer & Scholl, 2019).

Historic changes in lifespan variation are underpinned by the dynamics of mortality associated with the epidemiological transition (Robine, 2001). The first stage of epidemiological transition is associated with high infectious disease mortality and especially high infant mortality (Omran, 1971). Since infectious diseases can strike people of any age, deaths were common at all ages, but especially among young children. As a result, there was high variation in lifespans during this period, despite low life expectancies. As the epidemiological transition progresses through it's second and third stages infectious disease mortality reduces and non-communicable diseases become more common. Therefore, deaths among the young are prevented and become less common. However, the leading causes of death - heart attacks, other cardiovascular diseases, and strokes - commonly occur at a wide range of ages from late-middle age to old age (Santosa et al.,

[2]Survival curves are graphs of survival probability against age.

2014; Acosta et al., 2022). However, advances in medical care and public health have largely reduced premature mortality from such conditions meaning a greater proportion of the population live to advanced ages, continuing the reductions in lifespan variation. In the following stages of the transition, mortality from cardiovascular and cerebrovascular diseases decreases, while the prevalence of degenerative diseases, which tend to occur at advanced ages, increases, further compressing mortality (Robine, 2001; Omran, 1998).

No study has examined diversity in mortality causes over a wide enough period to assess changes associated with epidemiological transition. However, based on the available knowledge, we can speculate on how the epidemiological transition may have affected variation in the distribution of causes of mortality. At the peak of each stage of the transition, diversity in causes of mortality was likely relatively low, as the most common causes of mortality accounted for a large proportion of deaths. Using coarse-grained methods, we would expect low diversity in the first stage of transition, as the majority of deaths were due to infectious diseases. However, over time fewer people died due to infectious diseases. This likely resulted in diversification in causes of mortality as deaths shifted from infectious diseases to non-communicable diseases, which were previously less common. However, over time non-communicable diseases began to be dominant with an ever greater proportion of deaths occurring due to heart attacks and strokes. This would have presented as a reduction in diversity, as proportionally, a small subset of causes of mortality were becoming more common and causing a large proportion of deaths. In the next stage of the transition, we have evidence from this thesis and earlier research of a diversification of causes of mortality, as the proportion of deaths caused by cardiovascular diseases decreased. During this period (late 20th and early 21st centuries), deaths were redistributed to other causes especially degenerative diseases and cancers, increasing diversity. This hypothetical timeline is, of course, simplified but is illustrative of what might be expected. An interesting future research direction could be an exploration of diversity in causes of mortality over the epidemiological transition, although differences in the classification of disease over time make this challenging.

Over time, therefore, while the epidemiological transition has mostly brought about near-continuous reductions in lifespan variation it is likely the different stages of transition have been associated with waves of high, and low, diversity in mortality causes. The relationship between diversity in causes of mortality and lifespan diversity I report in this thesis might therefore be expected to differ at different points of transition. Most studies of diversity in mortality causes have found

increasing diversity associated with falling cardiovascular mortality rates. This is characteristic of the transition from the third to the fourth stage of Omran's (1998) updated theory of transition. Going further back in history, the progression from the second stage of transition to the third occurred in many developed nations in the late 1800s and early 1900s. In this period and up to 1950 cardiovascular diseases became the most common causes of death across the developed world increasing in prevalence (Thom et al., 1985). The conclusions of this section would suggest that in these years as variation in lifespan diversity fell over time diversity in mortality causes would also have fallen. The population was dying at more homogeneous ages and from less varied causes of death. Therefore the conclusions drawn in this thesis regarding the relationship between lifespan diversity and diversity in causes of mortality may not hold in future as patterns of mortality continue to evolve.

8.3.4 Diversity in causes of mortality across subpopulations in Scotland

Disparities in diversity in mortality causes between urban-rural subpopulations examined in this study are not found to have a clear pattern across measures. While more urban areas are observed to have greater diversity in causes under diversity at $q = 1$, at more conservative measures of diversity, less variance between these subpopulations is found and no clear tendency is observed. This evidence indicates that, while the single most common cause of mortality may differ across the urban-rural gradient, in each year the most common cause of death is responsible for a similar share of deaths across urban-rural classes. The effects of living in urban or rural areas (such as exposure to pollution and greater or lesser access to healthcare or green space among other factors) have been proposed to have wide ranging and varied impacts on health (Hartley, 2004; Levin, 2003). It is possible that some combination of these effects contributes to the patterns in diversity noted here.

When analysing populations by SIMD income deprivation quintiles, diversity in causes of mortality tended to be higher in more deprived areas when rare causes are weighted more strongly. However, this was reversed at more conservative measures meaning that in the most deprived areas, deaths were more dominated by the most common causes. This has similar implications to the temporal trends

in diversity in mortality causes across Scotland described in Section 8.3.2. In these deprived areas, this thesis has shown there is greater uncertainty in the mortality causes each individual might face but, concurrently, a greater proportion of individuals face the most common causes of death than in less deprived areas.

In this thesis I have pointed to various cause-specific differences in the burden of mortality faced by different subpopulations in Scotland. Many of these differences have been discussed broadly in the literature and many studies have examined the reasons why they exist. Deaths of despair are more common in more deprived areas and urban areas, leading, combined with other influences, to greater diversity at $q = 1$. It has been suggested that higher rates of deaths of despair, often attributed to self-destructive behaviour such as suicide and substance abuse, among more income-deprived groups is linked to marginalisation politically, economically, and in terms of community support (Allik et al., 2020). Marginalisation in each of these different spheres is proposed to make it more likely for individuals to engage in self-destructive behaviours and other behaviours associated with poor health. Interestingly, some of these factors, especially social isolation, have been linked to higher rates of deaths of despair among rural populations in the USA (DeVerteuil, 2022). Deaths of despair are shown in this thesis have a greater effect on the diversity of mortality causes, and be more prevalent, among urban populations than their more rural counterparts in Scotland. This international difference may be due to the interplay of socioeconomic and socio-geographic factors which differs across nations. That is to say that in some nations, such as the USA, deprivation is more common in rural areas than it is in Scotland.

The prevalence of the most common causes of mortality, such as heart disease and lung cancer, is shown to be higher in more deprived areas. As previously discussed, the prevalence of the most common causes is likely responsible for much of the observed changes and differences in diversity. Higher rates of smoking and obesity have been suggested as a cause of higher rates of cardiovascular disease in more deprived areas. Although evidence has suggested that even controlling for these factors cardiovascular disease risk is greater in more deprived areas (Yen & Kaplan, 1999; Sundquist et al., 2004). Worse access to health promoting influences, such as safe areas to exercise and access to healthy food, compared to less deprived areas have been associated with these increased risks (Fleury et al., 2000). In Scotland, deprived areas have a greater density of fast-food establishments and some of these areas have been shown to have greater density of alcohol outlets (Ellaway et al., 2010; Macdonald et al., 2018). Greater exposure to both of

these influences has been linked to poorer health outcomes in these areas. Higher prevalence of lung cancer in more deprived areas in England has been linked to higher rates of tobacco smoking (Riaz et al., 2011; Payne et al., 2022). Furthermore, in Scotland, social class has been shown to have an effect on lung cancer even when the effects of smoking are accounted for (Hart et al., 2001).

The previous research described above explains some of the reasons for the observed differences in distribution of mortality causes across subpopulations in Scotland. It does not, however, fully explain the similar rates of diversification observed between subpopulations in Scotland over the study period. In the 2000s, despite the causes of disease differing in prevalence across these groups, reductions in the prevalence of the most common cause and increases in variation in relatively rare causes occurred at similar rates. I have observed that reductions in the mortality rate and prevalence of acute myocardial infarction, alongside other diseases of the circulatory system, were key to increasing diversity, especially at $q = \infty$, in the 2000s. While the magnitude of these reductions in terms of mortality rate were different, they resulted in similar changes in the proportion of deaths which were attributed to acute myocardial infarction. This finding, in terms of mortality rate have been reported previously in the literature, percentage yearly reductions in the rate of chronic heart disease (CHD) (using a definition encompassing ICD-10 codes I10 to I25) in the 2000s have been shown to be greater in less deprived areas in this period, causing relative inequalities in CHD mortality to remain relatively constant (Hotchkiss et al., 2011; Hotchkiss et al., 2014). Hotchkiss et al. (2014) show that improved and increased medical treatments were responsible for around half of the improvements to heart disease mortality and that these advances were implemented relatively equally across the population. The fact that improvements to health procedures and reductions in medical risk factors were felt relatively equally across populations may be linked to the similar increases in diversity during the 2000s.

The study of diversity in mortality causes suggests that improvements to health were brought about in similar ways across subpopulations in Scotland. Reductions in the mortality rate of the most common causes were more rapid than reductions in the all-cause rate, leading to a redistribution in mortality causes and greater diversity in causes of mortality. Diversity in causes of mortality was examined in subpopulations in this thesis in order to assess where the consequences of diversification may be felt most strongly. The findings of this thesis show that diversification occurred at similar rates across all subpopulations. Trends across subpopulations are, for the most part, also shown to be similar when considering measures which

explicitly account for zero values in the distribution of mortality and when garbage codes are redistributed to valid cause of mortality as discussed in Appendix B.1. However, while I suggest much of this diversification was driven by similar rates of reduction in the prevalence of common causes, the causes of mortality which saw consequent increases in prevalence were different across subpopulations. Therefore, while an increased variation in causes has been faced across the population the potential interventions needed to address this differs among different groups.

Following 2015, reductions in the diversity of mortality causes are observed with relatively consistency in deaths among males across Scottish subpopulations. This is particularly evident under diversity $q = \infty$ which indicates that the proportion of deaths attributed to acute myocardial infarction has increased at a similar rate across groups. This is remarkable because many of the proposed reasons for faltering improvements to health outcomes in Scotland after 2015 are thought to more harshly affect more income deprived areas. In deaths among females, slowing improvements to diversity at more conservative measures are shown to be more pronounced in more deprived areas. The most common three-character ICD-10 cause of mortality among females following 2010 was lung cancer. Lung cancer mortality rates are known to be high and improvements to these rates are known to be slow in more deprived areas compared to less deprived neighbourhoods (Hart et al., 2001; Brown et al., 2019).

8.3.5 The COVID-19 pandemic

In this thesis, diversity in mortality causes in Scotland in 2020 and 2021 with COVID-19 mortality excluded from analysis, was shown to be within projected counterfactual trends. This potentially indicates that the COVID-19 pandemic has limited impact on variation in the distribution of mortality causes when COVID-19 deaths are excluded. Nonetheless, the effects of the pandemic on population health may be felt beyond 2021. As discussed in Chapter 6, it has been suggested that a range of causes of ill-health may become more common due to knock-on effects of the pandemic. In Scotland and the UK, cancers have been especially prominent in discussions of future consequences of the COVID-19 pandemic. Despite the efforts of public health and health care systems screening services for cancers were curtailed during the most stringent NPIs, such as lock-downs, and in other periods of the pandemic. As I suggest in Chapter 6, measuring the diversity of causes within ICD-10 Chapter II which includes all cancers may be a useful tool for future research.

When COVID-19 deaths are included in analysis, I find that diversity in causes of mortality was reduced compared to previous years under diversity at $q = 1$ and $q = 2$. This is what would be expected given the addition of a new highly prevalent cause of mortality to the distribution of causes of death. With the introduction of successful vaccines and improved treatments the mortality rate associated with COVID-19 has fallen dramatically, both in Scotland and around the world in 2021 and 2022 (Public Health Scotland, 2023). Despite this, at the time of writing, in 2023, COVID-19 still circulates and risks becoming an endemic virus worldwide with little evidence of willingness or ability to eradicate the disease (Steere-Williams, 2022). Therefore, deaths due to COVID-19 will continue to occur. Future studies which assess variation in causes of mortality must, therefore, address this when examining trends over time. Notional 2022 mortality rates produced by the Scottish Government suggest that COVID-19 was the sixth most common cause in Scotland in this year. In future therefore, COVID-19 may cease to reduce diversity and become a net contributor to variation in causes of mortality. This increase in diversity will reflect that compared to 2018 or 2019, in 2022 the variety, and number, of causes of death that the population of Scotland faced has increased. This is, in fact, a ready and obvious demonstration of the implications of increased diversity discussed throughout this thesis. In 2022, and presumably in the years to come, alongside all of the causes of mortality previously faced by the population, COVID-19 mortality must also be treated. This will, and demonstrably has, stretched healthcare and public health budgets (Cho & Kurpierz, 2020).

I show in chapter 7 that diversity in contributory causes was greater in deaths attributed to the novel coronavirus than other leading causes. This indicates a wide range of diseases with possible risk associations with COVID-19 (Atkins et al., 2020). Given that COVID-19 may remain endemic in societies (Steere-Williams, 2022; Adepoju, 2022), causes of morbidity and mortality which are exacerbated by COVID-19 may occur more often in the distribution of underlying and contributory mortality causes. Therefore, the prevalence of diseases for which COVID-19 is known to increase the risk of death may be higher than expected in coming years compared to previous trends. The effect this may have on the distribution of mortality causes, and population health more generally, is unclear.

8.3.5.1 Short-term population health consequences of the COVID-19 pandemic

The impact of the COVID-19 pandemic has been a hotly discussed issue in the literature and the media. This discussion has frequently focused on the concentration of deaths among vulnerable or frail individuals who may have had underlying chronic health conditions and would have faced a high risk of death during the pandemic, even if the COVID-19 virus had not been present (Cha & Jin, 2020; Drury et al., 2020; Youd & Moore, 2020). It is not the case that only vulnerable individuals or those with underlying health conditions died following COVID-19 infection, with evidence that mortality occurred in otherwise healthy individuals (Merchant et al., 2021). However, the majority of deaths occurred among elderly individuals and in those with underlying health conditions (Banerjee et al., 2020). There has been some discussion regarding whether this would, after the pandemic, lead to a lower burden of mortality because of the death of a fraction of the population who were, previously, at a high risk of death. This argument contends that deaths which would have occurred in 2020, 2021, or within a short period following these years, were effectively brought forward therefore, fewer deaths might be expected to occur in 2022 and later years (Ashkenazi & Rapaport, 2020). This view is not universally supported and has been discredited by some in the literature (Sridhar, 2020). Claims of deaths being "brought forward" or isolated to the elderly or frail have especially been challenged following their use to promote reasoning sceptical of NPIs during the pandemic (Darian-Smith, 2021).

In fact, early evidence from government agencies suggests that in 2022 the rate of mortality in the UK was in excess of what might be expected given pre-pandemic trends (Office for National Statistics, 2022). By 2022, the vast majority of UK residents had been vaccinated against COVID-19, COVID-19 related NPIs had been lifted and the mortality rate due to COVID-19 had fallen considerably (Hussain, 2022; Raleigh, 2022). It is suggested that COVID-19 mortality may contribute to some of the increase in excess deaths but, crucially, COVID-19 mortality rates are not considered high enough to explain all excess mortality (Raleigh, 2022). Furthermore, analysis has suggested that, in England and Wales, mortality rates of the 10 leading causes of death were within expected ranges in 2022 given previous years (Raleigh, 2022). Potential reasons for the increase in mortality in 2022 include a summer heatwave in the UK, but many point to the impact of the pandemic, particularly on the healthcare system. The NHS has faced effective budget cuts in the 2010s due to the austerity policies of successive UK Governments. This left

the NHS ill-prepared for the pressures of the pandemic (Mellish et al., 2020). In the aftermath waiting lists for procedures and tests have increased dramatically and across the UK healthcare capacity has been stretched thin. This has been described as a crisis widely in the media and has also been tentatively linked to the rise in excess deaths (Sridhar, 2022; Raleigh, 2022).

The UK is not isolated in experiencing high mortality in 2022 compared to the years before the COVID-19 pandemic. Statistical reporting agencies have documented excess mortality of >10%[3] in a number of EU nations (Eurostat, 2023) as well as in the USA and Australia (OECD.stat, 2023) in most weeks (OECD) or months (Eurostat) in 2022. It is important to approach these remarkable and continuous excess mortality findings with a degree of caution. Excess mortality is a metric used to indicate a difference in mortality rates from a standard average. The excess mortality figures presented in this section are derived from the average rates recorded between 2015 and 2019. If mortality trends were static, statistically speaking, we could anticipate mortality rates higher than this average about half of the time. Nevertheless, the high excess mortality in 2022 in the UK and across other nations has been greeted with concern (Sridhar, 2022).

In Chapter 6 I discuss the reasons for the limited impact of the COVID-19 pandemic on diversity in causes of mortality. I suggest two primary reasons: the dominance of the most common causes, which did not dramatically change in mortality rate in 2020 and 2021 compared to previous years, and changes in prevalence among relatively less common causes effectively cancelling each other out. This overall signifies that there was not an overarching pattern in the impact of the pandemic on the distribution of mortality causes. That is to say that, with COVID-19 excluded from analysis, the proportion of deaths attributed to the most common causes did not increase or decrease dramatically and relatively rare causes did not, as a whole, become more or less varied.

It remains unclear how the increased mortality rates observed in 2022 will impact the distribution of mortality causes. By examining the diversity of mortality causes in 2022 in Scotland, as well as across the UK and internationally, we can assess whether high mortality rates were driven by increases in the prevalence of

[3]An excess mortality of 10% indicates that the number of deaths recorded was 10% higher than an average based on the years 2015-2019.

the most common causes or by more rare causes occurring more often. Moreover, through examining the contribution of causes of mortality to diversity in 2022 it may be possible to assess the contested claim that deaths among the elderly and those with chronic may have been "brought forward" during the pandemic.

8.4 Strengths and limitations

8.4.1 The measures of diversity used in this thesis

In this thesis, diversity in causes of mortality is examined using measures of diversity proposed in the Reeve et al. (2016) framework described in Chapter 2. This thesis presents the first use of these diversity measures in the study of diversity in mortality causes. While the use of this framework has become increasingly widespread in the ecological literature it is not, yet, a standard in the field. It might be argued therefore, that other, more established, measures should be applied to the study of diversity in mortality causes.

The measures described in the Reeve et al. (2016) framework were chosen because of three primary features which, when combined, offer advantages over the enormous range of other measures of diversity in the literature. First, they describe diversity with meaningful units using the "effective number of types" formulation which, unlike other diversity measures, is mathematically intuitive and produces values which are directly comparable across populations. Secondly, this framework offers the ability to use the viewpoint parameter, q, to vary the importance of the prevalence of types to assess different aspects of variation in a distribution. This makes it possible to examine changes in the distribution at the level of rarer types as well as the most common types using measures which are analogous in their mathematical formulation and performance. This made it possible for me to examine trends in variation in the least and most common causes of mortality and identify diverging trends such as in the latter years of the 2010s in Scotland. Finally, the framework is flexible in its application, allowing for the range of research presented in this thesis using a number of measures: naïve-similarity subcommunity normalised alpha diversity in causes of mortality and lifespans; similarity-sensitive subcommunity normalised alpha lifespan diversity; and the measure of similarity-sensitive metacommunity gamma diversity

used in the creation of a similarity matrix for analysis of lifespan diversity. The Reeve et al. (2016) framework describes several other measures which have potentially promising applications in the study of diversity in mortality causes or age at death, some of which are discussed in Section 8.5. These measures use the same notation and can be used in conjunction to address a wide array of questions relating to variation in distributions both across populations and between subpopulations. While other systems for measurement of diversity provide some of the benefits described here, the Reeve et al. (2016) framework is effectively unique in combining them in a unified framework.

8.4.1.1 Alternative measures of diversity

In Appendix B.1, I explore the use of alternative measures of diversity in the study of diversity in mortality causes. One of these measures, uniformity, is distinct in that it explicitly includes zero values in the distribution of mortality causes in the calculation of diversity. A zero value occurs in the distribution of mortality causes if a cause of mortality occurs in the data set but does not occur in every year. In any year where this cause does not occur, it is assigned a zero value in the distribution of mortality. Uniformity explicitly accounts for and is sensitive to these zero values. Due to the use of fine-grained causes zero values occur regularly in the distributions of mortality examined in this thesis. Zero values would occur less often were coarse-grained causes to be analysed. Most measures of diversity used in the literature are not sensitive to zero values, instead measuring variation in the causes of mortality which occur in each year (or to be more exact, in measuring variation in the distribution of types in each community) in isolation (Shannon, 1948; Reeve et al., 2016; Simpson, 1949; Berger & Parker, 1970; Mitchell, 2019). As discussed in Appendix B.1 this measure bears relation to measures of evenness. Under this alternative measure of diversity, trends in diversity in underlying causes of mortality across Scotland are shown to differ slightly from those reported in the body of this thesis, with no clear trend in diversity up to 2015, followed by a marked increase in diversity in causes of mortality. This means that overall, diversity under this measure was higher in 2019 than in 2001 but that the trend for near-continuous increases up to 2015 found in Chapter 4 of this thesis is not observed. Similar differences in trend are observed in analysis of diversity in underlying causes of mortality in Scottish subpopulations. Greater differences are observed when examining diversity in contributory causes of mortality where a negative trend is found under uniformity from 2001 to 2010. In Chapter 7 I find that

diversity in contributory mortality causes increased in this period using normalised alpha diversity. Following 2010 trends under both uniformity and the measures used in Chapter 7 are shown to be similar with increases up to 2015 followed by reductions in diversity in females and a plateau in males.

Measures of diversity which do not account for zero values are the conventional standard in measuring diversity in mortality causes. Indeed none of the studies of mortality cause diversity found in the literature review described in Section 2.4 use measures which account for zero values. However, some may consider it advantageous to utilise measures which do account for zero values such as uniformity. I have shown, in Appendix B.1, that in most cases, the overall patterns in diversity in the years 2001 to 2019 are comparable to those discussed in the body of the thesis. However, the distinct trends observed under this alternative measure of diversity should be taken into account when considering the results of this thesis.

8.4.2 The use of ICD-10 three-character codes

In using all of the three-character codes assigned to deaths within Scotland, this study includes codes which are uninformative with regard to the true cause of death. These codes are often described as 'garbage codes' in the literature and are regarded as not being useful for population health research (Johnson et al., 2021; Ellingsen et al., 2022). Including these codes in analysis is potentially problematic because if they are associated with changes in diversity in mortality causes it would not be meaningful or useful in a population health context. In Appendix B.2, I perform a simple redistribution, assigning deaths attributed to garbage codes to valid causes of mortality, and calculate diversity under this altered distribution of mortality causes to examine the sensitivity of the results in this thesis to garbage codes. I perform this analysis for underlying causes of mortality in Scotland at all ages, and in deaths within twenty year age ranges and in deaths at all ages in Scottish subpopulation. I also perform this analysis for contributory causes of mortality in deaths at all ages across Scotland. Under this analysis the trends in underlying mortality cause diversity across Scotland and in Scottish subpopulations noted in Chapter 4 are found to be upheld. This suggests the conclusions of this thesis related to these trends are robust to the impact of garbage codes inn the distribution of mortality causes. When examining diversity in contributory causes of mortality, redistribution of deaths assigned to garbage codes is found to have limited impact on trends from 2001 to 2015. However, following 2015 diversity in contributory

causes of mortality with garbage codes redistributed continues to increase up to 2021, whereas in the observed distribution a reduction in diversity is observed following 2015 in females and diversity plateaus in males. Therefore, under this alternative analysis diversity in contributory causes of mortality increased across the study period indicating an increase in variation in the valid causes of mortality recorded in Scotland. The impact of garbage codes should be considered when assessing the conclusions of this thesis and it is likely advantageous in future research to explicitly account for garbage codes in analysis of diversity in mortality causes.

8.4.3 Accuracy in diagnosis

The accuracy with which the cause of each death is recorded is of key importance to the measurement of diversity in mortality causes. Should misclassification of causes be occurring on a large scale, it has the potential to hugely impact any conclusions drawn from research into the distribution of mortality causes as presented in this thesis. Studies have shown that there is considerable risk that the mortality recording in medical settings is not completely accurate (D'Amico et al., 1999; Zellweger et al., 2019). Legislation such as the Certification of Death Act (2011) in Scotland has aimed to ensure accuracy in diagnosis but there is still a risk that the causes of death examined in this thesis have been misclassified. It is difficult to envisage a method to address this misclassification without wholesale audits of mortality records.

8.4.4 The Scottish index of multiple deprivations

The Scottish Index of Multiple Deprivations (SIMD) is one of the most widely used measures of socioeconomic deprivation in research based in Scotland. While not without its detractors and limitations, the SIMD is generally considered to be the most advanced method for assessing area-level deprivation in Scotland (Clelland & Hill, 2019). The SIMD is an area-level measure of deprivation, which uses index scores calculated for small areas to assign a value for deprivation to geographic districts (Clelland & Hill, 2019). The use of area-level deprivation measures has been criticised in the literature because, by design, they are not indicators of the level of deprivation each individual faces. This means that some individuals who

would not be classified as facing deprivation but who live in relatively more deprived areas are designated as facing higher relative deprivation because of the area they live in. By the same mechanism some individuals who face material deprivation will be included in "less deprived" areas as they have neighbours in a better socioeconomic position (Lupton & Tunstall, 2003). The Scottish Government advice on the use of SIMD measures highlights this: 'not everyone living in a deprived area is deprived, and not all deprived people live in deprived areas' (The Scottish Government, 2017). Significant variation has been shown in the proportion of deprived individuals captured by SIMD across Scottish regions (Clelland & Hill, 2019). It is therefore possible that a degree of ecological fallacy is introduced by the use of these area-level measures in this thesis. Nevertheless, the use of area-level indicators is widespread in the literature, and most sources agree that despite the limitations discussed above, they are appropriate for the study of health inequalities.

One approach to overcome ecological fallacy would be the use of individual-level indicators of socioeconomic position. Unfortunately, it was not possible in the scope of this thesis to link any such indicators, for example educational attainment or social class, to mortality data. Even if this were possible, these approaches are not without their own limitations (Świgost, 2017; Kashem et al., 2019; Ingleby et al., 2020). A leading limitation of these indicators, and strength of area-level indicators, is that independent of an ones personal circumstances, the area in which one lives has a large impact on an individual's health (Dankwa-Mullan & Pérez-Stable, 2016).

8.4.5 Life tables

The calculation of life tables introduces certain demographic changes in the distribution of deaths within a population. This is discussed in Chapter 5 where the relative advantages, and disadvantages, of using life tables while assessing variation in age at death are examined. For example, in Scotland, the available population data ends with an open-ended age class at 90+. In order to calculate life tables, it is necessary to model mortality at ages older than 90. The modelling process has the effect of artificially extending the distribution of deaths to older ages. The approach for modelling these older age mortality rates used in this thesis is con-

sistent with standard practice (Seaman, 2017; Wilmoth et al., 2022). Life tables are generally promoted for use in population health analysis because they produce theoretically comparable study populations and account for biases in population structure.

In Results Chapters 4, 5 and 7, diversity in mortality causes is calculated using multiple decrement life tables while in Chapter 6, observed mortality data is used. Chapter 6 used different methods because diversity in causes of mortality was calculated in each month rather than across years. The number of deaths in each month was considered to be too small to reliably calculate life tables. Methods have been proposed in the literature to further model mortality rates and produce estimated life tables. However, in Chapter 3 (Methods), both life table and observed data methods for the calculation off diversity in mortality causes are shown to produce qualitatively similar results. Therefore, observed data were used in Chapter 6. For diversity in causes of mortality, the use of life tables, is found to have little effect on the calculation of diversity, suggesting that their use may be an unnecessary complication and source of uncertainty.

8.5 Future research directions

This section begins with a discussion of future research directions for the field of study of diversity in mortality causes. This is followed by a section of short case studies on specific methodologies which may be worth exploring further in future research.

This thesis has presented research on diversity in mortality causes in Scotland. Novel methods have been applied to this study which have potential to be explored in a wider range of contexts to further understanding of variation in the distribution of mortality causes. The work in this thesis adds to the evidence base regarding diversity in mortality causes within low-mortality, developed nations. All but one of the previous studies of diversity in mortality causes found in Chapter 2 has addressed the distributions of causes in either European nations, the USA, Japan, Australia or a combination of these countries. Ascaso-Terrén et al. (1988) is the only previous study to have examined diversity outside of this group by addressing deaths in Chile, Thailand, Egypt, Israel and Cuba alongside European nations and

the USA. The prevailing causes of mortality faced by populations vary a great deal across the world especially in regions where people face higher mortality rates and poorer health outcomes (Bollyky et al., 2019). Further exploration of the diversity of mortality causes among populations outside of the low-mortality, mostly Western nations that research in this area has been focused on could provide valuable information on how resources should be best focused to address poor health outcomes in these populations. If diversity in mortality causes in these nations is found to be lower than in the previously studied populations then a focus on the reducing the prevalence of most common mortality causes might provide the greatest benefit. Whereas should diversity in causes of mortality prove to be similar or even higher among these populations then a more holistic approach, focusing on a wide range of causes, should be adopted to provide the greatest benefit. Further to these potential direct insights, study of variation in causes of mortality in populations across the world may help inform theories of epidemiological transition in populations where traditional models have been found to be insufficient (Santosa et al., 2014).

Some of the most recent advances in the study of variation in lifespans has come from the examination of variation in healthy lifespans, meaning that the length of time each individual spends in a state of full health is considered rather only the length of their life (Permanyer et al., 2022). Extending analysis of diversity to causes of disease and morbidity affecting populations during life is a logical parallel here. This would provide similar insights to those gained from the study of diversity in mortality causes. Increased diversity in lifetime diseases and morbidities would indicate greater individual-level uncertainty in the conditions one might face and mean healthcare and public health resources must be spread to treat and prevent a greater variety of causes. However, a framing which extended to the study of causes of poor health during life could help to address the increasing acknowledgement in the literature of the importance of health across the life course (Mitchell, 2005). This framing could further address concerns regarding the ability of clinical measures of health including those based on the measurement of mortality to tackle ill-health during life (Halfon & Hochstein, 2002). Dyson et al. (2021) have presented research related to the study of variation in causes of disease burden measuring the Gini coefficient of the distribution of cause-specific DALYs in various regions across the world extracted from the Global Burden of Disease study. The authors find reductions in the concentration of DALYs around the leading causes from 1990 to 2018. The methods employed by Dyson et al. (2021) are informative with regard to concentration among the leading causes and reranking of causes

(i.e. causes moving higher or lower on a list ranked by prevalence). However, the application of diversity measures to this data can expand the scope of study into variation in the distribution of causes of ill-health. A possible approach to address this is to calculate diversity in the distribution of DALYs by cause.

Beta diversity is defined as the diversity between subcommunities within a metacommunity. That is to say, the differences (or similarities) in the causes of mortality faced by different groups rather than the variation within those groups. Previously, Trias-Llimós and Permanyer (2023) have used what is effectively a beta diversity measure to compare the contribution of changes in diversity of causes within subpopulations of the USA grouped by educational attainment (alpha diversity) to changes in the diversity between these groups. They suggest that increased diversity in causes across the US population was driven more by changes between educational attainment groups than within them. These methods have potential in the study of inequalities between groups of the population making it possible to examine whether inter- or intra-group variation was a greater driver of changes across the population. It is also possible to apply these beta diversity measures to the distribution of ages at death to conduct similar analysis.

The Reeve et al. (2016) framework under which diversity is calculated in this thesis presents four different measures of subcommunity beta diversity. The derivation of these measures is explained in detail by Reeve et al. (2016) and in the thesis of Mitchell (2019). These measures assess different aspects of between-group diversity. A particularly promising measure in the study of diversity in mortality causes is representativeness (normalised beta diversity (reversed)), which describes how typical the distribution of mortality causes within a subcommunity (e.g. a council area, region, or deprivation quintile) is of the distribution of mortality causes across rest of the metacommunity (e.g. Scotland). It is maximised if the distribution of causes in a subcommunity is identical (in terms of the causes present in the distribution and the prevalence of each cause) to that of the metacommunity. Conversely, it is minimised when the causes of death within a subcommunity are not found in any other subcommunity which makes up the metacommunity. Representativeness measure may be valuable where little prior knowledge exists on the relative health of subpopulations to identify groups which have particularly unusual distributions of mortality causes and patterns of health.

8.5.1 Brief case studies of future research possibilities

This section presents short reports detailing potential future directions for research into diversity in mortality causes. These reports present work carried out during the completion of this thesis that is in addition to the research questions addressed in this work. The first, Section 8.5.1.1, presents a possible route to apply the use of similarity-sensitive measures in analysis of diversity in mortality causes. In Section 8.5.1.2 I discuss the issues presented by the study of diversity in mortality causes (or indeed lifespan diversity) among small populations and present a possible method for overcoming these issues. The methods utilised and the findings of these case studies of are explained in detail in Appendix E.

8.5.1.1 Similarity-sensitive diversity in causes of mortality

8.5.1.1.1 Background Throughout this thesis diversity in causes of mortality has been measured using naïve-similarity measures. These measures treat each cause of mortality as entirely distinct. In Chapter 5 similarity-sensitive measures were used in the study of lifespan diversity. These measures have not previously been applied to diversity in causes of mortality. Causes might be considered more closely related if they are within the same ICD-10 Chapter or if they are associated with the same risk factors. In acknowledging similarity between causes it is possible to bridge the gap between fine-grained cause diversity, as used in this thesis, and course-grained cause diversity which has been used more widely in studies of diversity in mortality causes in the 21st century. Previously issues related to the use of fine-grained ICD-10 codes in the calculation of diversity have been discussed. These stem from the fact that these codes do not necessarily indicate “equal” causes of mortality, some may have wider definitions than others for example. Some studies of diversity in mortality causes group causes together to overcome this however, in doing this variation between fine-grained causes is lost. These limitations are addressed in Chapter 2. Using similarity-sensitive measures might make it possible to address issues related to fine-grained cause diversity without entirely losing the detail of variation at the level of fine-grained causes of mortality.

The selection of a suitable framework for addressing similarity between causes of mortality requires further research. The properties of similarity-sensitive measures in the context of diversity in mortality causes should be examined and these measures validated before they are used in earnest in the study of population health. Appendix E presents a possible avenue for the application of similarity with similarity informed by the structure of ICD-10 categorisation.

8.5.1.1.2 Implications for future research Trends in the similarity-sensitive diversity of mortality causes are found to be similar to those in naive-similarity diversity presented previously in this thesis. The most notable deviation is that changes in trends at diversity in mortality causes at q - ∞ in the 2010s are more pronounced under similarity-sensitive measures. The use of similarity-sensitive measures of diversity in analysis of mortality causes has the potential to solve several of the issues raised by the dichotomy between fine-grained cause diversity and coarse-grained cause diversity discussed in Chapter 2. By producing a more meaningful similarity matrix, through a more systematic process of linking causes than the ICD-10 categorisation used here, future research can account for the fact that fine-grained causes are not necessarily descriptive on the same level. However, unlike grouping causes this method would preserve variation in prevalence between fine-grained causes. Producing a similarity matrix which more methodically addresses the relations between each of the more than 1,700 ICD-10 three-character codes was beyond the scope of this thesis. These methods may however, be a valuable tool for future research.

Another reason that similarity-sensitive methods were not employed in this thesis and a key drawback to the use of similarity-sensitive measures is that it complicates any study of the causes which have driven changes in diversity. While under naïve-similarity diversity each cause of mortality is treated separately and its contribution to diversity can be understood in isolation, similarity-sensitive measures mean that the contribution of each cause to diversity is dependent on its relation to other causes as well as its prevalence. These concerns are not insurmountable but require further study and the adaptation of methods (such as additive value analysis as introduced in this thesis) to suit similarity sensitive methods.

8.5.1.2 Diversity in the study of population health in small populations

8.5.1.2.1 Background The measurement of diversity in any system can be biased by the size of the study population as discussed in Section 3.3.3.2. This limits the direct measurement of diversity in mortality causes among small populations. It may be desirable to examine diversity among small populations to investigate place-based determinants of health or to assess differences between small subsets of the population. Various influences of area in which one lives have been linked to health including access to green space, the influence of the built environment, access to healthcare, and air pollution (Campbell et al., 2000; Dankwa-Mullan & Pérez-Stable, 2016). Addressing how these factors affect variation in the distribution of mortality causes would be made possible through analysis of mortality at the level of individual communities and small populations. While deaths and populations have been assigned to datazones in Scotland in this thesis these datazones were then aggregated into larger groups for analysis.

The measurement of diversity in mortality causes (and/or ages at death) within small populations can aid in research related to a number of other research gaps. One key area mentioned in this thesis is the comparative effects of the urban-rural gradient and socioeconomic deprivation. To examine this it would be ideal to measure diversity in mortality causes at the level of small-area geographies, for example, census districts or other administrative geographies. Socio-geographic and socioeconomic factors vary across these small communities making study of the effects of various influences on diversity in mortality causes possible.

Having the ability to reliably measure diversity in mortality causes within small-area geographies would also enable study of geographic variation. Understanding how diversity in mortality causes varies geographically could potentially help inform public health planning. Areas with higher diversity in causes might benefit from more public health spending on a wider variety of causes whereas in areas with low diversity in mortality causes greater gains in health might be made by focussing on the most common causes.

In order to study diversity in the causes of mortality among small populations I suggest the use of "random resampling" methods. This relies on examining the distribution of causes of death within small populations as a part of a larger population. Here, as an example, I examine the diversity of mortality causes within postcode sectors in Scotland, these are administrative geographies with average populations of around 5,000 individuals (Brown et al., 2014). An algorithm is used to reassign the cause of each death within each postcode sector. For each individual a cause is sampled, with replacement, from the distribution of causes in individuals of the same age and sex across Scotland with likelihood based on the prevalence of each cause in this distribution. Diversity in the sampled causes of mortality for each postcode sector can then be calculated. This process is then repeated over a set number of iterations and an average of the diversity in each postcode sector is calculated. This produces an estimate of how diverse the mortality causes of each postcode sector would be expected to be given the number of deaths and the sex and age of those who died.

Appendix E demonstrates an example of the random resampling method described above. The distributions of mortality causes among deaths in the years 2001 to 2005 in Scottish census postcode sectors are examined. The algorithm described in the previous paragraph is applied to these distributions and the resampled "expected diversity in mortality causes" is compared to the observed diversity in mortality causes in deaths within these small-area geographies.

8.5.1.2.2 Implications for future research Notable variation from expected diversity in mortality causes is observed in many postcode sectors. Further a potential tendency for more deprived areas to have less diverse causes of mortality than would be expected given the age and sex structure of the area. Measuring the diversity of mortality causes in small-area geographies such as postcode sectors or within other small subsections of population is complicated by small numbers of deaths. However, examining diversity at this level may be beneficial to answer research questions related to place-based or localised health determinants or to examine spatial variation in mortality cause diversity at small scales. In this section I have proposed a potential method for future research in this area allowing for analysis of diversity among small subpopulations. Further development of the methods described here might include addressing economic deprivation or other determinants of health in the resampling process. This would involve choosing the cause of mortality for each individual from the distribution of causes faced by those at the same age and sex who died in areas with similar levels of deprivation.

Analysis of whether an area was more, or less, diverse than expected given both the age and sex structure of the population and the relative level of deprivation in the area would then be possible. Under the methods described in this section addressing diversity in causes of mortality at small populations is made possible.

8.6 Conclusions

Measuring diversity in causes of mortality offers insight into the distribution of mortality causes and patterns of mortality which are not possible through all-cause or cause-specific analysis. In this thesis I have proposed novel methods for measuring diversity in mortality causes and lifespan diversity which may be of benefit to the field in future allowing for more robust comparison between populations, more comprehensive analysis of the contribution of fine-grained causes of death to overall diversity and clearer interpretation of results. Using these methods, I have examined the diversity in causes of mortality across Scotland and in subpopulations. Furthermore, I have explored the relationship between diversity in causes of mortality and lifespan diversity. Finally, I have examined the impact of the COVID-19 pandemic on variation in causes of mortality.

In conclusion, I have shown that the measurement of diversity in causes of mortality is a useful tool for analysis of patterns of mortality and population health. I have presented evidence that diversity in mortality causes was higher in 2019 than 2001 across Scotland and across subpopulations. Diversity in mortality causes is shown to be patterned complexly across subpopulations although, in most groups, trends over time in diversity were similar. In more deprived areas and urban areas diversity in mortality causes was higher when rare causes are weighted heavily but in more deprived areas the most common causes were more prevalent. I have demonstrated that improvements in population health, such as increasing life expectancies and falling lifespan diversity, in the 2000s were associated with a diversification in causes of mortality. I suggest that each of these trends were driven by falling mortality rates and prevalences of the most common causes of mortality. However, I have noted worrying trends in population health in Scotland since 2015. Alongside stalling life expectancy and increasing lifespan diversity diversification in causes of mortality has stalled in females and reversed in males. This was evident before the COVID-19 pandemic and disruption due to the pandemic may have the effect of worsening these trends in coming years. However, I have shown that

variation in causes of mortality with COVID-19 deaths excluded was consistent with previous trends during the COVID-19 pandemic. Measuring diversity in causes of mortality, using the methods proposed in this thesis, can be an essential tool in studying such disruptions in population health and may provide valuable insights for policy.

Bibliography

Aburto, J. M., Alvarez, J.-A., Villavicencio, F., & Vaupel, J. W. (2019). The threshold age of keyfitz' entropy. https://doi.org/arXiv:1901.07963

Aburto, J. M., Basellini, U., Baudisch, A., & Villavicencio, F. (2021). Drewnowski's index to measure lifespan variation: Revisiting the gini coefficient of the life table. https://doi.org/10.48550/arXiv.2111.11256

Aburto, J. M., & van Raalte, A. (2018). Lifespan dispersion in times of life expectancy fluctuation: The case of central and eastern europe. *Demography*, *55*(6), 2071–2096. https://doi.org/10.1007/s13524-018-0729-9

Acosta, E., Mehta, N., Myrskylä, M., & Ebeling, M. (2022). Cardiovascular mortality gap between the united states and other high life expectancy countries in 2000–2016. *The Journals of Gerontology: Series B*, *77*, S148–S157. https://doi.org/10.1093/geronb/gbac032

Adepoju, P. (2022). Africa prepares for endemic COVID-19. *Nature Medicine*, *28*(5), 876–878. https://doi.org/10.1038/d41591-022-00040-0

Ahmad, O. B., Boschi-Pinto, C., Lopez, A. D., Murray, C. J., Lozano, R., & Inoue, M. (2001). *Age standardization of rates: A new WHO standard*. World Health Organization.

Albani, V., Welsh, C. E., Brown, H., Matthews, F. E., & Bambra, C. (2022). Explaining the deprivation gap in COVID-19 mortality rates: A decomposition analysis of geographical inequalities in england. *Social Science & Medicine*, *311*, 115319. https://doi.org/10.1016/j.socscimed.2022.115319

Albert, C., & Davia, M. A. (2011). Education is a key determinant of health in europe: A comparative analysis of 11 countries. *Health Promotion International*, *26*(2), 163–170. https://doi.org/10.1093/heapro/daq059

Alboukadel Kassambara. (2020). *Package 'ggpubr'* (Version 0.4.0). Retrieved August 15, 2022, from https://cran.r-project.org/web/packages/ggpubr/ggpubr.pdf

Alfonso Viguria, U., & Casamitjana, N. (2021). Early interventions and impact of COVID-19 in spain. *International Journal of Environmental Research and Public Health*, *18*(8), 4026. https://doi.org/10.3390/ijerph18084026

Alkundi, A., Mahmoud, I., Musa, A., Naveed, S., & Alshawwaf, M. (2020). Clinical characteristics and outcomes of COVID-19 hospitalized patients with diabetes in the united kingdom: A retrospective single centre study. *Diabetes Research and Clinical Practice*, *165*, 108263. https://doi.org/10.1016/j.diabres.2020.108263

Allik, M., Brown, D., Dundas, R., & Leyland, A. H. (2020). Deaths of despair: Cause-specific mortality and socioeconomic inequalities in cause-specific mortality among young men in scotland. *International Journal for Equity in Health*. https://doi.org/10.1186/s12939-020-01329-7

Alvarez-Uria, G., Gandra, S., & Laxminarayan, R. (2016). Poverty and prevalence of antimicrobial resistance in invasive isolates. *International Journal of Infectious Diseases*, *52*, 59–61. https://doi.org/10.1016/j.ijid.2016.09.026

Alzahrani, S. I., Aljamaan, I. A., & Al-Fakih, E. A. (2020). Forecasting the spread of the COVID-19 pandemic in saudi arabia using ARIMA prediction model under current public health interventions. *Journal of Infection and Public Health*, *13*(7), 914–919. https://doi.org/10.1016/j.jiph.2020.06.001

Amele, S., Kibuchi, E., McCabe, R., Pearce, A., Henery, P., Hainey, K., Fagbamigbe, A. F., Kurdi, A., McCowan, C., Simpson, C. R., Dibben, C., Buchanan, D., Demou, E., Almaghrabi, F., Anghelescu, G., Taylor, H., Tibble, H., Rudan, I., Nazroo, J., … Katikireddi, S. V. (2023). Ethnic inequalities in positive SARS-CoV-2 tests, infection prognosis, COVID-19 hospitalisations and deaths: Analysis of 2 years of a record linked national cohort study in scotland [Publisher: BMJ Publishing Group Ltd Section: Original research]. *J Epidemiol Community Health*, *77*(10).

Anderson, R. N. Coding and classifying causes of death: Trends and international differences. In: In *International handbook of adult mortality*. International Handbooks of Population. Dordrecht: Springer Netherlands, 2011, pp. 467–489. ISBN: 978-90-481-9996-9.

Andersson, G., & Philipov, D. (2002). Life-table representations of family dynamics in sweden, hungary, and 14 other FFS countries: A project of descriptions of demographic behavior. *Demographic Research*, *7*, 67–144.

Andreev, E. M. (2019). Reflections on demographic theories. *Population and Economics*, *3*(2), 1–9. https://doi.org/10.3897/popecon.3.e37965

Andrews, N., Tessier, E., Stowe, J., Gower, C., Kirsebom, F., Simmons, R., Gallagher, E., Chand, M., Brown, K., Ladhani, S. N., Ramsay, M., & Bernal, J. L. (2021). Vaccine effectiveness and duration of protection of comirnaty, vaxzevria and spikevax against mild and severe COVID-19 in the UK. https://doi.org/10.1101/2021.09.15.21263583

Arah, O. A. (2009). On the relationship between individual and population health. *Medicine, Health Care and Philosophy*, *12*(3), 235–244. https://doi.org/10.1007/s11019-008-9173-8

Arthi, V., & Parman, J. (2021). Disease, downturns, and wellbeing: Economic history and the long-run impacts of COVID-19. *Explorations in Economic History*, *79*, 101381. https://doi.org/10.1016/j.eeh.2020.101381

Artzrouni, M. Mathematical demography (K. Kempf-Leonard, Ed.). In: *Encyclopedia of social measurement* (K. Kempf-Leonard, Ed.). Ed. by Kempf-Leonard, K. New York: Elsevier, 2005, pp. 641–651. ISBN: 978-0-12-369398-3. https://doi.org/10.1016/B0-12-369398-5/00360-1.

Asah, F. N., Nielsen, P., & Sæbø, J. I. Challenges for health indicators in developing countries: Misconceptions and lack of population data (J. Choudrie, M. S. Islam, F. Wahid, J. M. Bass & J. E. Priyatma, Eds.) [Series Title: IFIP Advances in Information and Communication Technology]. In: *Information and communication technologies for development* (J. Choudrie, M. S. Islam, F. Wahid, J. M. Bass & J. E. Priyatma, Eds.). Ed. by Choudrie, J., Islam, M. S., Wahid, F., Bass, J. M., & Priyatma, J. E. Vol. 504. Series Title: IFIP Advances in Information and Communication Technology. Cham: Springer International Publishing, 2017, pp. 593–604. ISBN: 978-3-319-59110-0 978-3-319-59111-7. https://doi.org/10.1007/978-3-319-59111-7_48.

Ascaso Terrén, C., Canela Soler, J., & Sentís Vílalta, J. (1988). Una aproximación ecológica a las causas de mortalidad en 14 países: el Índice de diversidad de shannon. *Gaceta Sanitaria*, *2*(8), 224–229. https://doi.org/10.1016/S0213-9111(88)70931-0

Ashkenazi, I., & Rapaport, C. (2020). Saving lives versus saving dollars: The acceptable loss for coronavirus disease 2019*. *Critical Care Medicine*, *48*(8), 1243–1244. https://doi.org/10.1097/CCM.0000000000004452

Asteriou, D., & Price, S. (2001). Political instability and economic growth: UK time series evidence. *Scottish Journal of Political Economy*, *48*(4), 383–399. https://doi.org/10.1111/1467-9485.00205

Atkins, J. L., Masoli, J. A. H., Delgado, J., Pilling, L. C., Kuo, C.-L., Kuchel, G. A., & Melzer, D. (2020). Preexisting comorbidities predicting COVID-19 and mortality in the UK biobank community cohort. *The Journals of Gerontology: Series A*, *75*(11), 2224–2230. https://doi.org/10.1093/gerona/glaa183

Atkinson, A. M., McAuley, A., Trayner, K. M. A., & Sumnall, H. R. (2019). 'we are still obsessed by this idea of abstinence': A critical analysis of UK news media representations of proposals to introduce drug consumption rooms in glasgow, UK. *International Journal of Drug Policy*, *68*, 62–74. https://doi.org/10.1016/j.drugpo.2019.03.010

Augarde, E., Gunnell, D., Mars, B., & Hickman, M. (2022). An ecological study of temporal trends in 'deaths of despair' in england and wales. *Social Psychiatry and Psychiatric Epidemiology*, *57*(6), 1135–1144. https://doi.org/10.1007/s00127-022-02251-9

Ball, S., Banerjee, A., Berry, C., Boyle, J. R., Bray, B., Bradlow, W., Chaudhry, A., Crawley, R., Danesh, J., Denniston, A., Falter, F., Figueroa, J. D., Hall, C., Hemingway, H., Jefferson, E., Johnson, T., King, G., Lee, K. K., McKean, P., … Wyatt, M. (2020). Monitoring indirect impact of COVID-19 pandemic on services for cardiovascular diseases in the UK. *Heart*, *106*(24), 1890–1897. https://doi.org/10.1136/heartjnl-2020-317870

Banerjee, A., Chen, S., Pasea, L., Lai, A. G., Katsoulis, M., Denaxas, S., Nafilyan, V., Williams, B., Wong, W. K., Bakhai, A., Khunti, K., Pillay, D., Noursadeghi, M., Wu, H., Pareek, N., Bromage, D., McDonagh, T. A., Byrne, J., Teo, J. T. H., … Hemingway, H. (2021). Excess deaths in people with cardiovascular diseases during the COVID-19 pandemic. *European Journal of Preventive Cardiology*, *28*(14), 1599–1609. https://doi.org/10.1093/eurjpc/zwaa155

Banerjee, A., Pasea, L., Harris, S., Gonzalez-Izquierdo, A., Torralbo, A., Shallcross, L., Noursadeghi, M., Pillay, D., Sebire, N., Holmes, C., Pagel, C., Wong, W. K., Langenberg, C., Williams, B., Denaxas, S., & Hemingway, H. (2020). Estimating excess 1-year mortality associated with the COVID-19 pandemic according to underlying conditions and age: A population-based cohort study. *The Lancet*, *395*(10238), 1715–1725. https://doi.org/10.1016/S0140-6736(20)30854-0

Barnett, A. G., De Looper, M., & Fraser, J. F. (2008). The seasonality in heart failure deaths and total cardiovascular deaths. *Australian and New Zealand Journal of Public Health*, *32*(5), 408–413. https://doi.org/10.1111/j.1753-6405.2008.00270.x

Batty, G. D., Gale, C. R., Kivimäki, M., & Bell, S. (2019). Assessment of relative utility of underlying vs contributory causes of death. *JAMA Network Open*, *2*(7), e198024. https://doi.org/10.1001/jamanetworkopen.2019.8024

Batty, G. D., Gale, C. R., Kivimäki, M., Deary, I. J., & Bell, S. (2020). Comparison of risk factor associations in UK biobank against representative, general population based studies with conventional response rates: Prospective cohort study and individual participant meta-analysis. *The BMJ*, *368*. https://doi.org/10.1136/bmj.m131

Baum, D. A. (2009). Species as ranked taxa. *Systematic Biology*, *58*(1), 74–86. https://doi.org/10.1093/sysbio/syp011

Baxter, M. A., Murphy, J., Cameron, D., Jordan, J., Crearie, C., Lilley, C., Sadozye, A., Maclean, M., Hall, P., Phillips, A., Greger, A., Madeleine, J., & Petty, R. D. (2021). The impact of COVID-19 on systemic anticancer treatment delivery in scotland. *British Journal of Cancer*, *124*(8), 1353–1356. https://doi.org/10.1038/s41416-021-01262-8

Beltrán-Sánchez, H., Preston, S. H., & Canudas-Romo, V. (2008). An integrated approach to cause-of-death analysis: Cause-deleted life tables and decompositions of life expectancy. *Demographic research*, *19*, 1323. https://doi.org/10.4054/DemRes.2008.19.35

Berger, W. H., & Parker, F. L. (1970). Diversity of planktonic foraminifera in deep-sea sediments. *Science*, *168*(3937). https://doi.org/10.1126/science.168.3937.1345

Bergeron-Boucher, M.-P., Aburto, J. M., & van Raalte, A. (2020). Diversification in causes of death in low-mortality countries: Emerging patterns and implications. *BMJ Global Health*, *5*. https://doi.org/10.1136/bmjgh-2020-002414

Bergeron-Boucher, M.-P., Bourbeau, R., & Légaré, J. (2017). Changes in cause-specific mortality among the elderly in canada, 1979–2011. *Canadian Studies in Population*, *43*(3), 215. https://doi.org/10.25336/p69w3w

Bernal, J. L., Andrews, N., Gower, C., Stowe, J., Robertson, C., Tessier, E., Simmons, R., Cottrell, S., Roberts, R., O'Doherty, M., Brown, K., Cameron, C., Stockton, D., McMenamin, J., & Ramsay, M. (2021). Early effectiveness of COVID-19 vaccination with BNT162b2 mRNA vaccine and ChAdOx1 adenovirus vector vaccine on symptomatic disease, hospitalisations and mortality in older adults in england [Pages: 2021.03.01.21252652]. https://doi.org/10.1101/2021.03.01.21252652

Beseran, E., Pericàs, J. M., Cash-Gibson, L., Ventura-Cots, M., Porter, K. M. P., & Benach, J. (2022). Deaths of despair: A scoping review on the social determinants of drug overdose, alcohol-related liver disease and suicide. *International Journal of Environmental Research and Public Health*, *19*(19), 12395. https://doi.org/10.3390/ijerph191912395

Bhatnagar, P., Wickramasinghe, K., Williams, J., Rayner, M., & Townsend, N. (2015). The epidemiology of cardiovascular disease in the UK 2014. *Heart*, *101*(15). https://doi.org/10.1136/heartjnl-2015-307516

Bhise, V., Rajan, S. S., Sittig, D. F., Morgan, R. O., Chaudhary, P., & Singh, H. (2018). Defining and measuring diagnostic uncertainty in medicine: A systematic review. *Journal of General Internal Medicine*, *33*(1), 103–115. https://doi.org/10.1007/s11606-017-4164-1

Bhopal, R., Fischbacher, C., Povey, C., Chalmers, J., Mueller, G., Steiner, M., Brown, H., Brewster, D. H., & Bansal, N. (2011). Cohort profile: Scottish health and ethnicity linkage study of 4.65 million people exploring ethnic variations in disease in scotland. *International Journal of Epidemiology*, (5), 1168–1175.

Bird, S. M., McAuley, A., Perry, S., & Hunter, C. (2016). Effectiveness of scotland's national naloxone programme for reducing opioid-related deaths: A before (2006–10) versus after (2011–13) comparison. *Addiction*, *111*(5), 883–891. https://doi.org/10.1111/add.13265

Bishop, K., Balogun, S., Eynstone-Hinkins, J., Moran, L., Martin, M., Banks, E., Rao, C., & Joshy, G. (2022). Analysis of multiple causes of death: A review of methods and practices. https://doi.org/10.1101/2022.08.01.22278086

Bollyky, T. J., Templin, T., Cohen, M., Schoder, D., Dieleman, J. L., & Wigley, S. (2019). The relationships between democratic experience, adult health, and cause-specific mortality in 170 countries between 1980 and 2016: An observational analysis. *The Lancet*, *393*(10181), 1628–1640. https://doi.org/10.1016/S0140-6736(19)30235-1

Boniface, S., Scannell, J. W., & Marlow, S. (2017). Evidence for the effectiveness of minimum pricing of alcohol: A systematic review and assessment using the bradford hill criteria for causality. *BMJ Open*, *7*(5), e013497. https://doi.org/10.1136/bmjopen-2016-013497

Brackin, W, Brashnyk, P, Brown, F., Churchill, B., Cohen, T., Cohn, K, Delisle, D., DeSouki, G, Friedman, J, & Gilfillan, R. (2015). *What is population health?* Retrieved November 9, 2022, from https://onlinepublichealth.gwu.edu/resources/what-is-population-health/

Bradford, D. R. R., Allik, M, McMahon, A. D., & Brown, D. (2022). Assessing the risk of endogeneity bias in health and mortality inequalities research using composite measures of multiple deprivation which include health-related indicators: A case study using the scottish index of multiple deprivation and population health and mortality data. https://doi.org/https://doi.org/10.1101/2022.12.01.22282961

Bromiley, P. A., Thacker, N. A., & Bouhova-Thacker, E. (2004). Shannon entropy, renyi entropy, and information, 9.

Brown, D., Allik, M., Dundas, R., & Leyland, A. H. (2019). All-cause and cause-specific mortality in scotland 1981–2011 by age, sex and deprivation: A population-based study. *European Journal of Public Health*, *29*(4), 647–655. https://doi.org/10.1093/eurpub/ckz010

Brown, D., Allik, M., Dundas, R., & Leyland, A. H. (2014). *Carstairs scores for scottish postcode sectors, datazones and output areas from the 2011 census*. University of Glasgow. Retrieved November 30, 2022, from https://eprints.gla.ac.uk/99555/

Brown, M. L., Lipscomb, J., & Snyder, C. (2001). The burden of illness of cancer: Economic cost and quality of life. *Annual Review of Public Health*, *22*(1), 91–113. https://doi.org/10.1146/annurev.publhealth.22.1.91

Brown, M., MacArthur, J., McKechanie, A., Hayes, M., & Fletcher, J. (2010). Equality and access to general health care for people with learning disabilities: Reality or rhetoric? *Journal of Research in Nursing*, *15*(4), 351–361. https://doi.org/10.1177/1744987110370019

Brønnum-Hansen, H., Östergren, O., Tarkiainen, L., Hermansen, ¤., Martikainen, P., van der Wel, K. A., & Lundberg, O. (2021). Changes in life expectancy and lifespan variability by income quartiles in four nordic countries: A study based on nationwide register data. *BMJ open*, *11*(6), e048192. https://doi.org/10.1136/bmjopen-2020-048192

Bucks, D., Baylis, A., Dougall, D., & Robertson, R. (2018). *A vision for population health: Towards a healthier future*. The King's Fund.

Burton, J. K., Reid, M., Gribben, C., Caldwell, D., Clark, D. N., Hanlon, P., Quinn, T. J., Fischbacher, C., Knight, P., Guthrie, B., & McAllister, D. A. (2021). Impact of COVID-19 on care-home mortality and life expectancy in scotland. *Age and Ageing*, *50*(4), 1029–1037. https://doi.org/10.1093/ageing/afab080

Campbell, C., Sommerfield, T., Clark, G. R. C., Porteous, L., Milne, A. M., Millar, R., Syme, T., & Thomson, C. S. (2021). COVID-19 and cancer screening in scotland: A national and coordinated approach to minimising harm. *Preventive Medicine*, *151*, 106606. https://doi.org/10.1016/j.ypmed.2021.106606

Campbell, N. C., Elliott, A. M., Sharp, L., Ritchie, L. D., Cassidy, J., & Little, J. (2000). Rural factors and survival from cancer: Analysis of scottish cancer registrations [Number: 11 Publisher: Nature Publishing Group]. *British Journal of Cancer*, *82*(11), 1863–1866. https://doi.org/10.1054/bjoc.1999.1079

Carreras, M., Ibern, P., Coderch, J., Sánchez, I., & Inoriza, J. M. (2013). Estimating lifetime healthcare costs with morbidity data. *BMC Health Services Research*, *13*(1), 440. https://doi.org/10.1186/1472-6963-13-440

Carstairs, V., & Morris, R. (1989). Deprivation: Explaining differences in mortality between scotland and england and wales. *British Medical Journal*. https://doi.org/10.1136/bmj.299.6704.886

Carter, A. J., & Nguyen, C. N. (2012). A comparison of cancer burden and research spending reveals discrepancies in the distribution of research funding. *BMC Public Health*, *12*(1), 526. https://doi.org/10.1186/1471-2458-12-526

Casant, J., & Helbich, M. (2022). Inequalities of suicide mortality across urban and rural areas: A literature review. *International Journal of Environmental Research and Public Health*, *19*(5), 2669. https://doi.org/10.3390/ijerph19052669

Case, A., & Deaton, A. (2015). Rising morbidity and mortality in midlife among white non-hispanic americans in the 21st century. *Proceedings of the National Academy of Sciences*, *112*(49), 15078–15083. https://doi.org/10.1073/pnas.1518393112

Caswell, H., & Van Daalen, S. (2021). Healthy longevity from incidence-based models: More kinds of health than stars in the sky. *Demographic Research*, *45*. https://doi.org/10.4054/DemRes.2021.45.13

Caswell, H., & Zarulli, V. (2018). Matrix methods in health demography: A new approach to the stochastic analysis of healthy longevity and DALYs. *Population Health Metrics*, *16*(1), 8. https://doi.org/10.1186/s12963-018-0165-5

Cha, S., & Jin, Y. (2020). Have inequalities in all-cause and cause-specific child mortality between countries declined across the world? *International Journal for Equity in Health*, *19*(1), 1–13. https://doi.org/10.1186/s12939-019-1102-3

Chadeau-Hyam, M., Bodinier, B., Elliott, J., Whitaker, M. D., Tzoulaki, I., Vermeulen, R., Kelly-Irving, M., Delpierre, C., & Elliott, P. (2020). Risk factors for positive and negative COVID-19 tests: A cautious and in-depth analysis of UK biobank data. *International Journal of Epidemiology*, *49*(5), 1454–1467. https://doi.org/10.1093/ije/dyaa134

Chao, A., Chiu, C.-H., & Jost, L. (2014a). Unifying species diversity, phylogenetic diversity, functional diversity, and related similarity and differentiation measures through hill numbers. *Annual Review of Ecology, Evolution, and Systematics*, *45*(1), 297–324. https://doi.org/10.1146/annurev-ecolsys-120213-091540

Chao, A., Gotelli, N. J., Hsieh, T. C., Sander, E. L., Ma, K. H., Colwell, R. K., & Ellison, A. M. (2014b). Rarefaction and extrapolation with hill numbers: A framework for sampling and estimation in species diversity studies. *Ecological Monographs*, *84*(1), 45–67. https://doi.org/10.1890/13-0133.1

Cheema, K. M., Dicks, E., Pearson, J., & Samani, N. J. (2022). Long-term trends in the epidemiology of cardiovascular diseases in the UK: Insights from the british heart foundation statistical compendium. *Cardiovascular Research*, *118*(10), 2267–2280. https://doi.org/10.1093/cvr/cvac053

Chen, J., & McGeorge, R. (2020). Spillover effects of the COVID-19 pandemic could drive long-term health consequences for non-COVID-19 patients. *Health Affairs Forefront*. https://doi.org/10.1377/forefront.20201020.566558

Cheung, S. L. K., Robine, J.-M., Tu, E. J.-C., & Caselli, G. (2005). Three dimensions of the survival curve: Horizontalization, verticalization, and longevity extension. *Demography*, *42*(2), 243–258. https://doi.org/10.1353/dem.2005.0012

Chi, H. (1988). Life-table analysis incorporating both sexes and variable development rates among individuals. *Environmental Entomology*, *17*(1), 26–34. https://doi.org/10.1093/ee/17.1.26

Cho, C. H., & Kurpierz, J. (2020). Stretching the public purse: Budgetary responses to COVID-19 in canada. *Journal of Public Budgeting, Accounting & Financial Management*, *32*(5), 771–783. https://doi.org/10.1108/JPBAFM-05-2020-0070

Clelland, D., & Hill, C. (2019). Deprivation, policy and rurality: The limitations and applications of area-based deprivation indices in scotland. *Local Economy*, *34*(1), 33–50. https://doi.org/10.1177/0269094219827893

Coale, A. J., Demeny, P., & Vaughan, B. (2013). *Regional model life tables and stable populations: Studies in population*. Elsevier.

Corburn, J. (2017). Urban place and health equity: Critical issues and practices. *International Journal of Environmental Research and Public Health*, *14*(2), 117. https://doi.org/10.3390/ijerph14020117

Corrao, G., Rea, F., & Blangiardo, G. C. (2021). Lessons from COVID-19 mortality data across countries. *Journal of Hypertension*, *39*(5), 856–860. https://doi.org/10.1097/HJH.0000000000002833

Critchley, J. A., Capewell, S., & Unal, B. (2003). Life-years gained from coronary heart disease mortality reduction in scotland: Prevention or treatment? *Journal of Clinical Epidemiology*, *56*(6), 583–590. https://doi.org/10.1016/S0895-4356(03)00059-3

Critchley, J. A., & Capewell, S. (2003). Mortality risk reduction associated with smoking cessation in patients with coronary heart disease: A systematic review. *JAMA*, *290*(1), 86–97. https://doi.org/10.1001/jama.290.1.86

Dalgarno, P., O'Rawe, S., & Hammersley, R. (2021). Illegal drugs in the UK: Is it time for considered legalisation to improve public health? *Drug Science, Policy and Law*, *7*. https://doi.org/10.1177/20503245211005351

D'Amico, M., Agozzino, E., Biagino, A., Simonetti, A., & Marinelli, P. (1999). Ill-defined and multiple causes on death certificates – a study of misclassification in mortality statistics. *European Journal of Epidemiology*, *15*(2), 141–148. https://doi.org/10.1023/A:1007570405888

Danilova, I. A. (2016). Problems of the quality of cause-specific mortality statistics at old ages. *Advances in Gerontology*, *6*(1), 1–5. https://doi.org/10.1134/S2079057016010021

Dankwa-Mullan, I., & Pérez-Stable, E. J. (2016). Addressing health disparities is a place-based issue. *American Journal of Public Health*, *106*(4), 637–639. https://doi.org/10.2105/AJPH.2016.303077

Darian-Smith, E. (2021). Dying for the economy: Disposable people and economies of death in the global north. *State Crime Journal*, *10*, 61–79. https://doi.org/10.13169/statecrime.10.1.0061

de Bruin, A. M., van Rossum, A. C., Visser, M. C., & Koole, G. M. (2007). Modeling the emergency cardiac in-patient flow: An application of queuing theory. *Health Care Management Science*, *10*(2), 125–137. https://doi.org/10.1007/s10729-007-9009-8

de Jong, M., Woodward, M., & Peters, S. A. (2021). Diabetes and COVID-19–related mortality in women and men in the UK biobank: Comparisons with influenza/pneumonia and coronary heart disease. *Diabetes Care*, *44*(2), e22–e24. https://doi.org/10.2337/dc20-2378

De Livera, A. M., Hyndman, R. J., & Snyder, R. D. (2011). Forecasting time series with complex seasonal patterns using exponential smoothing. *Journal of the American Statistical Association*, *106*(496), 1513–1527. https://doi.org/10.1198/jasa.2011.tm09771

De Ramos, I. P., Auchincloss, A. H., & Bilal, U. (2022). Exploring inequalities in life expectancy and lifespan variation by race/ethnicity and urbanicity in the united states: 1990 to 2019. *SSM - Population Health*, *19*, 101230. https://doi.org/10.1016/j.ssmph.2022.101230

Declich, S., & Carter, A. O. (1994). Public health surveillance: Historical origins, methods and evaluation. *Bulletin of the World Health Organization*, *72*(2), 285–304.

DeVerteuil, G. (2022). Deaths of despair and the social geographies of health denial [_eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/gec3.12607]. *Geography Compass*, *16*(2), e12607. https://doi.org/10.1111/gec3.12607

Dilsizian, S. E., & Siegel, E. L. (2013). Artificial intelligence in medicine and cardiac imaging: Harnessing big data and advanced computing to provide personalized medical diagnosis and treatment. *Current Cardiology Reports*, *16*(1), 441. https://doi.org/10.1007/s11886-013-0441-8

Docherty, K. F., Butt, J. H., Boer, R. A. d., Dewan, P., Køber, L., Maggioni, A. P., McMurray, J. J. V., Solomon, S. D., & Jhund, P. S. (2020). Excess deaths during the covid-19 pandemic: An international comparison. https://doi.org/10.1101/2020.04.21.20073114

Douglas, M., Katikireddi, S. V., Taulbut, M., McKee, M., & McCartney, G. (2020). Mitigating the wider health effects of covid-19 pandemic response. *BMJ*, *369*, m1557. https://doi.org/10.1136/bmj.m1557

Drury, J., Reicher, S., & Stott, C. (2020). COVID-19 in context: Why do people die in emergencies? it's probably not because of collective psychology. *British Journal of Social Psychology*, *59*(3), 686–693. https://doi.org/10.1111/bjso.12393

Dutey-Magni, P. F., Williams, H., Jhass, A., Rait, G., Lorencatto, F., Hemingway, H., Hayward, A., & Shallcross, L. (2021). COVID-19 infection and attributable mortality in UK care homes: Cohort study using active surveillance and electronic records (march–june 2020). *Age and Ageing*, *50*(4), 1019–1028. https://doi.org/10.1093/ageing/afab060

Dyson, H., Van Gestel, R., & van Doorslaer, E. (2021). The relative importance and stability of disease burden causes over time: Summarizing regional trends on disease burden for 290 causes over 28 years. *Population Health Metrics*, *19*(1), 30. https://doi.org/10.1186/s12963-021-00257-0

Désesquelles, A., Demuru, E., Salvatore, M. A., Pappagallo, M., Frova, L., Meslé, F., & Egidi, V. (2014). Mortality from alzheimer's disease, parkinson's disease, and dementias in france and italy: A comparison using the multiple cause-of-death approach. *Journal of Aging and Health*, *26*(2), 283–315. https://doi.org/10.1177/0898264313514443

Earle, M. (2010, October 25). *Certification of death (scotland) bill*. SPICe. https://archive2021.parliament.scot/ResearchBriefingsAndFactsheets/S3/SB_10-83.pdf

Edwards, R. (2010). *Trends in world inequality in life span since 1970* (w16088). National Bureau of Economic Research. Cambridge, MA. https://doi.org/10.3386/w16088

Ellaway, A., Macdonald, L., Forsyth, A., & Macintyre, S. (2010). The socio-spatial distribution of alcohol outlets in glasgow city. *Health & Place*, *16*(1), 167–172. https://doi.org/10.1016/j.healthplace.2009.08.007

Ellingsen, C. L., Alfsen, G. C., Ebbing, M., Pedersen, A. G., Sulo, G., Vollset, S. E., & Braut, G. S. (2022). Garbage codes in the norwegian cause of death registry 1996–2019. *BMC Public Health*, *22*(1), 1301. https://doi.org/10.1186/s12889-022-13693-w

Elliott, J., Bodinier, B., Whitaker, M., Delpierre, C., Vermeulen, R., Tzoulaki, I., Elliott, P., & Chadeau-Hyam, M. (2021). COVID-19 mortality in the UK biobank cohort: Revisiting and evaluating risk factors. *European Journal of Epidemiology*, *36*(3), 299–309. https://doi.org/10.1007/s10654-021-00722-y

Etches, V., Frank, J., Ruggiero, E. D., & Manuel, D. (2006). MEASURING POPULATION HEALTH: A review of indicators. *Annual Review of Public Health*, *27*(1), 29–55. https://doi.org/10.1146/annurev.publhealth.27.021405.102141

Eurostat. (2013). *Revised european standard population*.

Eurostat. (2023). *Excess mortality - statistics*. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Excess_mortality_-_statistics

Evans, T. (2001). *Challenging inequities in health: From ethics to action*. Oxford University Press.

Exworthy, M., Blane, D., & Marmot, M. (2003). Tackling health inequalities in the united kingdom: The progress and pitfalls of policy. *Health Services Research*, *38*(6), 1905–1922. https://doi.org/10.1111/j.1475-6773.2003.00208.x

Farmer, J., Hinds, K., Richards, H., & Godden, D. (2005). Urban versus rural populations' views of health care in scotland [Publisher: SAGE Publications]. *Journal of Health Services Research & Policy*, *10*(4), 212–219. https://doi.org/10.1258/135581905774414240

Farmer, J., Philip, L., King, G., Farrington, J., & MacLeod, M. (2010). Territorial tensions: Misaligned management and community perspectives on health services for older people in remote rural areas. *Health & Place*, *16*(2), 275–283. https://doi.org/10.1016/j.healthplace.2009.10.010

Fedeli, U., Zoppini, G., Goldoni, C. A., Avossa, F., Mastrangelo, G., & Saugo, M. (2015). Multiple causes of death analysis of chronic diseases: The example of diabetes. *Population Health Metrics*, *13*(1), 21. https://doi.org/10.1186/s12963-015-0056-y

Fenton, L., Minton, J., Ramsay, J., Kaye-Bardgett, M., Fischbacher, C., Wyper, G. M. A., & McCartney, G. (2019a). Recent adverse mortality trends in scotland: Comparison with other high-income countries [Publisher: British Medical Journal Publishing Group Section: Epidemiology]. *BMJ Open*, *9*(10), e029936. https://doi.org/10.1136/bmjopen-2019-029936

Fenton, L., Wyper, G. M., McCartney, G., & Minton, J. (2019b). Socioeconomic inequality in recent adverse all-cause mortality trends in scotland. *Journal of Epidemiology and Community Health*. https://doi.org/10.1136/jech-2019-212300

Ferrara, P., & Albano, L. (2022). Advances in population-based healthcare research: From measures to evidence. *International Journal of Environmental Research and Public Health*, *19*(20), 13122. https://doi.org/10.3390/ijerph192013122

Figueroa, J. D., Brennan, P., Theodoratou, E., Poon, M., Purshouse, K., Din, F., Jin, K., Mesa-Eguiagaray, I., Dunlop, M., Hall, P., Cameron, D., Wild, S., & Sudlow, C. (2020a). Trends in excess cancer and cardiovascular deaths in scotland during the COVID-19 pandemic 30 december – 20 april suggest underestimation of COVID-19 related deaths. https://doi.org/10.1101/2020.05.02.20086231

Figueroa, J. D., Brennan, P. M., Theodoratou, E., Poon, M. T. C., Purshouse, K., Din, F. V. N., Jin, K., Mesa-Eguiagaray, I., Dunlop, M. G., Hall, P. S., Cameron, D., Wild, S. H., & Sudlow, C. L. M. (2020b). Distinguishing between direct and indirect consequences of covid-19. *BMJ*, *369*, m2377. https://doi.org/10.1136/bmj.m2377

Figueroa, J. D., Gray, E., Pashayan, N., Deandrea, S., Karch, A., Vale, D. B., Elder, K., Procopio, P., van Ravesteyn, N. T., Mutabi, M., Canfell, K., & Nickson, C. (2021). The impact of the covid-19 pandemic on breast cancer early detection and screening. *Preventive Medicine*, *151*, 106585. https://doi.org/10.1016/j.ypmed.2021.106585

Finucane, A. M., Bone, A. E., Etkind, S., Carr, D., Meade, R., Munoz-Arroyo, R., Moine, S., Iyayi-Igbinovia, A., Evans, C. J., Higginson, I. J., & Murray, S. A. (2021). How many people will need palliative care in scotland by 2040? a mixed-method study of projected palliative care need and recommendations for service delivery. *BMJ Open*, *11*(2), e041317. https://doi.org/10.1136/bmjopen-2020-041317

Fleury, J., Keller, C., & Murdaugh, C. (2000). Social and contextual etiology of coronary heart disease in women. *Journal of Women's Health & Gender-Based Medicine*, *9*(9), 967–978. https://doi.org/10.1089/15246090050199991

Foreman, K. J., Marquez, N., Dolgert, A., Fukutaki, K., Fullman, N., McGaughey, M., Pletcher, M. A., Smith, A. E., Tang, K., Yuan, C.-W., Brown, J. C., Friedman, J., He, J., Heuton, K. R., Holmberg, M., Patel, D. J., Reidy, P., Carter, A., Cercy, K., … Murray, C. J. L. (2018). Forecasting life expectancy, years of life lost, and all-cause and cause-specific mortality for 250 causes of death: Reference and alternative scenarios for 2016–40 for 195 countries and territories. *The Lancet*, *392*(10159), 2052–2090. https://doi.org/10.1016/S0140-6736(18)31694-5

Fries, J. F. (1980). Aging, natural death, and the compression of morbidity. *The New England Journal of Medicine*, *303*(3), 130–135. https://doi.org/10.1056/NEJM198007173030304

Fu, B., Huang, F., & Maller, R. (2021). Modelling australian life tables with advanced ages – a report prepared for the australian government actuary. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4020309

Garthwaite, K., & Bambra, C. (2017). "how the other half live": Lay perspectives on health inequalities in an age of austerity. *Social Science & Medicine*, *187*, 268–275. https://doi.org/10.1016/j.socscimed.2017.05.021

Gemmell, I., McLoone, P., Boddy, F. A., Dickinson, G. J., & Watt, G. C. (2000). Seasonal variation in mortality in scotland. *International Journal of Epidemiology*. https://doi.org/10.1093/ije/29.2.274

Global Burden of Disease Collaborative Network. (2020). Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: A systematic analysis for the global burden of disease study 2019. *Lancet (London, England)*, *396*(10258), 1204–1222. https://doi.org/10.1016/S0140-6736(20)30925-9

Glover, M., Buxton, M., Guthrie, S., Hanney, S., Pollitt, A., & Grant, J. (2014). Estimating the returns to UK publicly funded cancer-related research in terms of the net value of improved health outcomes. *BMC Medicine*, *12*(1). https://doi.org/10.1186/1741-7015-12-99

Grebenik, E. (1991). Demographic research in britain 1936–1986. *Population Studies*, *45*(1), 3–30. https://doi.org/10.1080/0032472031000145876

Gross, C. P., Anderson, G. F., & Powe, N. R. (1999). The relation between funding by the national institutes of health and the burden of disease. *New England Journal of Medicine*, *340*(24), 1881–1887. https://doi.org/10.1056/NEJM199906173402406

Grundy, E. M., & Stuchbury, R. (2022). Multimorbidity as assessed by reporting of multiple causes of death: Variations by period, sociodemographic characteristics and place of death among older decedents in england and wales, 2001–2017. *Journal of Epidemiology and Community Health*, *76*(8), 699–706. https://doi.org/10.1136/jech-2021-217846

Guimarães, R. M. (2021). American dream or nightmare? deaths from despair and the future of capitalism. *Trabalho, Educação e Saúde*, *19*. https://doi.org/10.1590/1981-7746-sol00327

Hadjidemetriou, G. M., Sasidharan, M., Kouyialis, G., & Parlikad, A. K. (2020). The impact of government measures and human mobility trend on COVID-19 related deaths in the UK. *Transportation Research Interdisciplinary Perspectives*, *6*, 100167. https://doi.org/10.1016/j.trip.2020.100167

Halfon, N., & Hochstein, M. (2002). Life course health development: An integrated framework for developing health, policy, and research. *Milbank Quarterly*, *80*(3). https://doi.org/10.1111/1468-0009.00019

Hart, C. L., Hole, D. J., Gillis, C. R., Smith, G. D., Watt, G. C., & Hawthorne, V. M. (2001). Social class differences in lung cancer mortality: Risk factor explanations using two scottish cohort studies. *International Journal of Epidemiology*, *30*(2), 268–274. https://doi.org/10.1093/ije/30.2.268

Hartley, D. (2004). Rural health disparities, population health, and rural culture. *American Journal of Public Health*, *94*(10), 1675–1678. https://doi.org/10.2105/AJPH.94.10.1675

Health, T. L. P. (2020). Education: A neglected social determinant of health. *The Lancet. Public Health*, *5*(7), e361. https://doi.org/10.1016/S2468-2667(20)30144-4

Helleringer, S., & Queiroz, B. L. (2022). Commentary: Measuring excess mortality due to the COVID-19 pandemic: Progress and persistent challenges. *International Journal of Epidemiology*, *51*(1), 85–87. https://doi.org/10.1093/ije/dyab260

Hiam, L., Minton, J., & McKee, M. (2021). What can lifespan variation reveal that life expectancy hides? comparison of five high-income countries. *Journal of the Royal Society of Medicine*, *114*(8), 389–399. https://doi.org/10.1177/01410768211011742

Hill, M. O. (1973). Diversity and evenness: A unifying notation and its consequences. *Ecology*, *54*, 427–432. https://doi.org/10.2307/1934352

Hoem, J. M. (1977). A markov chain model of working life tables. *Scandinavian Actuarial Journal*, *1977*(1), 1–20. https://doi.org/10.1080/03461238.1977.10405621

Holland, W. W. (2017). Health improvement in a period of austerity: A public health perspective. *Global Policy*, *8*, 30–37. https://doi.org/10.1111/1758-5899.12216

Horwood, C. M., Hakendorf, P., Thompson, C. H., Horwood, C. M., Hakendorf, P., & Thompson, C. H. (2018). Comparison of specialist and generalist care. *Australian Health Review*, *42*(5). https://doi.org/10.1071/AH17197

Hosseinpoor, A. R., Bergen, N., Barros, A. J. D., Wong, K. L. M., Boerma, T., & Victora, C. G. (2016). Monitoring subnational regional inequalities in health: Measurement approaches and challenges. *International Journal for Equity in Health*, *15*(1), 18. https://doi.org/10.1186/s12939-016-0307-y

Hosseinpoor, A. R., Williams, J. S., Amin, A., Carvalho, I. A. d., Beard, J., Boerma, T., Kowal, P., Naidoo, N., & Chatterji, S. (2012). Social determinants of self-reported health in women and men: Understanding the role of gender in population health. *PLOS ONE*, *7*(4), e34799. https://doi.org/10.1371/journal.pone.0034799

Hotchkiss, J. W., Davies, C., Gray, L., Bromley, C., Capewell, S., & Leyland, A. H. (2011). Trends in adult cardiovascular disease risk factors and their socio-economic patterning in the scottish population 1995–2008: Cross-sectional surveys. *BMJ Open*, *1*(1), e000176. https://doi.org/10.1136/bmjopen-2011-000176

Hotchkiss, J. W., Davies, C. A., Dundas, R., Hawkins, N., Jhund, P. S., Scholes, S., Bajekal, M., O'Flaherty, M., Critchley, J., Leyland, A. H., & Capewell, S. (2014). Explaining trends in scottish coronary heart disease mortality between 2000 and 2010 using IMPACTSEC model: Retrospective analysis using routine data. *BMJ*, *348*, g1088. https://doi.org/10.1136/bmj.g1088

Hougaard, P. (1984). Life table methods for heterogeneous populations: Distributions describing the heterogeneity. *Biometrika*, *71*(1), 75–83. https://doi.org/10.1093/biomet/71.1.75

Hsieh, J. J. (1991). A general theory of life table construction and a precise abridged life table method. *Biometrical Journal*, *33*(2), 143–162. https://doi.org/10.1002/bimj.4710330204

Huang, D., Yang, S., & Liu, T. (2020a). Life expectancy in chinese cities: Spatially varied role of socioeconomic development, population structure, and natural conditions. *International Journal of Environmental Research and Public Health*, *17*(18), 6597. https://doi.org/10.3390/ijerph17186597

Huang, F., Maller, R., & Ning, X. (2020b). Modelling life tables with advanced ages: An extreme value theory approach. *Insurance: Mathematics and Economics*, *93*, 95–115. https://doi.org/10.1016/j.insmatheco.2020.04.004

Human Mortality Database. (n.d.). Human mortality database. *Max Planck Institute for Demographic Research (Germany), University of California, Berkeley (USA), and French Institute for Demographic Studies (France)*.

Hunter, P. R., Izsák, J., & Nehaul, L. (2003). Changes of neoplasm concentration with geographical co-ordinates. *Health and Place*, *9*(4), 305–313. https://doi.org/10.1016/S1353-8292(02)00063-1

Hussain, Z. (2022). UK health officials analyse recent rise in excess deaths [Publisher: British Medical Journal Publishing Group Section: News]. *BMJ*, *378*, o2085. https://doi.org/10.1136/bmj.o2085

Hyndman, R., Athanasopoulos, G., Bergmeir, C., Caceres, G., Chhay, L., O'Hara-Wild, M., Petropoulos, F., Razbash, S., Wang, E., & Yasmeen, F. (2022). *Forecast: Forecasting functions for time series and linear models*. https://pkg.robjhyndman.com/forecast/

Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and practice (2nd ed)*. OTexts: Melbourne, Australia. https://otexts.com/fpp2/index.html

Hyndman, R. J., & Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for r. *Journal of Statistical Software*, *26*(3), 1–22. https://doi.org/10.18637/jss.v027.i03

Imai, N., Gaythorpe, K. A., Abbott, S., Bhatia, S., van Elsland, S., Prem, K., Liu, Y., & Ferguson, N. M. (2020). Adoption and impact of non-pharmaceutical interventions for COVID-19. *Wellcome Open Res*, *5*, 59.

Ingleby, F. C., Belot, A., Atherton, I., Baker, M., Elliss-Brookes, L., & Woods, L. M. (2020). Assessment of the concordance between individual-level and area-level measures of socio-economic deprivation in a cancer patient cohort in england and wales. *BMJ Open*, *10*(11), e041714. https://doi.org/10.1136/bmjopen-2020-041714

Izsák, J. (1982). Comparative study of diversity indices on mortality data. *Biológia*, *30*.

Izsák, J., & Juhás-Nagy, P. (1984). Diversity studies on mortality data. *Annales Universitatis Scientiarum Budapestinensis*.

Izsák, J. (1986). Measuring the secular changes of the concentration of death causes. *Genus*, *42*(3), 197–208.

Izsák, J. (1988). Secular changes of the concentration of neoplasm death causes in the male population of some countries. *Genus*, *44*(3), 119–130.

Izsák, J. (1993a). Comparative analysis of death cause diversity curves in various countries. *Genus*, *49*(1), 67–77.

Izsák, J. (1993b). Measuring epidemiological diversity and concentration: A short review. *Abstracta Botanica*, *17*, 245–250.

Izsák, J., & György, K. (1991). Concentration analysis of diagnostic groups in a geronto-epidemiological survey. *Statisztikai Szeml*.

Jansen, L., Hakulinen, T., & Brenner, H. (2012). Standard errors of non-standardised and age-standardised relative survival of cancer patients. *British Journal of Cancer*, *106*(3), 569–574. https://doi.org/10.1038/bjc.2011.560

John, R., & Innocent, P. (2005). Modeling uncertainty in clinical diagnosis using fuzzy logic. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *35*(6), 1340–1350. https://doi.org/10.1109/TSMCB.2005.855588

Johnson, S. C., Cunningham, M., Dippenaar, I. N., Sharara, F., Wool, E. E., Agesa, K. M., Han, C., Miller-Petrie, M. K., Wilson, S., Fuller, J. E., Balassyano, S., Bertolacci, G. J., Davis Weaver, N., GBD Cause of Death Collaborators, Arabloo, J., Badawi, A., Bhagavathula, A. S., Burkart, K., Cámera, L. A., … Naghavi, M. (2021). Public health utility of cause of death data: Applying empirical algorithms to improve data quality. *BMC Medical Informatics and Decision Making*, *21*(1), 175. https://doi.org/10.1186/s12911-021-01501-1

Jones, N. L., Gilman, S. E., Cheng, T. L., Drury, S. S., Hill, C. V., & Geronimus, A. T. (2019). Life course approaches to the causes of health disparities. *American Journal of Public Health*, *109*, S48–S55. https://doi.org/10.2105/AJPH.2018.304738

Jost, L. (2006). Entropy and diversity. *Oikos*. https://doi.org/10.1111/j.2006.0030-1299.14714.x

Jost, L. (2007). Partitioning diversity into independent alpha and beta components. *Ecology*, *88*(10), 2427–24. https://doi.org/10.1890/06-1736.1

Jost, L. (2010). The relation between evenness and diversity. *Diversity*, *2*(2), 207–232. https://doi.org/10.3390/d2020207

Jurasinski, G., Retzer, V., & Beierkuhnlein, C. (2009). Inventory, differentiation, and proportional diversity: A consistent terminology for quantifying species diversity. *Oecologia*, *159*(1), 15–26. https://doi.org/10.1007/s00442-008-1190-z

Jäkel, A., Plested, M., Dharamshi, K., Modha, R., Bridge, S., & Johns, A. (2013). A systematic review of economic evaluations in second and later lines of therapy for the treatment of non-small cell lung cancer. *Applied Health Economics and Health Policy*, *11*(1). https://doi.org/10.1007/s40258-012-0001-1

Kannisto, V. (2000). Measuring the compression of mortality. *Demographic Research*, *3*(6). https://doi.org/10.4054/DemRes.2000.3.6

Kashem, T., Al Sayah, F., Tawiah, A., Ohinmaa, A., & Johnson, J. A. (2019). The relationship between individual-level deprivation and health-related quality of life. *Health and Quality of Life Outcomes*, *17*(1), 176. https://doi.org/10.1186/s12955-019-1243-5

Katikireddi, S. V., Cezard, G., Bhopal, R. S., Williams, L., Douglas, A., Millard, A., Steiner, M., Buchanan, D., Sheikh, A., & Gruer, L. (2018). Assessment of health care, hospital admissions, and mortality by ethnicity: Population-based cohort study of health-system performance in scotland. *The Lancet Public Health*, *3*(5), e226–e236. https://doi.org/10.1016/S2468-2667(18)30068-9

Keaver, L., Pérez-Ferrer, C., Jaccard, A., & Webber, L. (2020). Future trends in social inequalities in obesity in england, wales and scotland. *Journal of Public Health (United Kingdom)*, *42*(1). https://doi.org/10.1093/pubmed/fdz022

Kelly, A., & Panush, R. S. (2017). Diagnostic uncertainty and epistemologic humility. *Clinical Rheumatology*, *36*(6), 1211–1214. https://doi.org/10.1007/s10067-017-3631-8

Kerasidou, A., & Kingori, P. (2019). Austerity measures and the transforming role of a&e professionals in a weakening welfare system [Publisher: Public Library of Science]. *PLOS ONE*, *14*(2), e0212314. https://doi.org/10.1371/journal.pone.0212314

Keyfitz, N., & Caswell, H. (2005). *Applied mathematical demography* (3rd). Springer New York, NY.

Kindig, D., & Stoddart, G. (2003). What is population health? *American Journal of Public Health*, *93*(3), 380–383. Retrieved November 8, 2022, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1447747/

Kindig, D. A. (2007). Understanding population health terminology. *The Milbank Quarterly*, *85*(1), 139–161. https://doi.org/10.1111/j.1468-0009.2007.00479.x

King, A. C., Ahn, D. F., Atienza, A. A., & Kraemer, H. C. (2008). Exploring refinements in targeted behavioral medicine intervention to advance public health. *Annals of Behavioral Medicine*, *35*(3), 251–260. https://doi.org/10.1007/s12160-008-9032-0

King, L., Scheiring, G., & Nosrati, E. (2022). Deaths of despair in comparative perspective. *Annual Review of Sociology*, *48*(1), 299–317. https://doi.org/10.1146/annurev-soc-030320-031757

Kingston, A., Robinson, L., Booth, H., Knapp, M., & Jagger, C. (2018). Projections of multi-morbidity in the older population in england to 2035: Estimates from the population ageing and care simulation (PACSim) model. *Age and Ageing*, *47*(3), 374–380. https://doi.org/10.1093/ageing/afx201

Koetsier, A., Peek, N., & De Keizer, N. Identifying types and causes of errors in mortality data in a clinical registry using multiple information systems. In: In *Studies in health technology and informatics*. 2012. ISBN: 978-1-61499-100-7. https://doi.org/10.3233/978-1-61499-101-4-771.

Kraft, A. N., Thatcher, E. J., & Zenk, S. N. (2020). Neighborhood food environment and health outcomes in u.s. low-socioeconomic status, racial/ethnic minority, and rural populations: A systematic review. *Journal of Health Care for the Poor and Underserved*, *31*(3), 1078–1114. https://doi.org/10.1353/hpu.2020.0083

Krieger, N. (2012). Who and what is a "population"? historical debates, current controversies, and implications for understanding "population health" and rectifying health inequities. *The Milbank Quarterly*, *90*(4), 634–681. https://doi.org/10.1111/j.1468-0009.2012.00678.x

Kringos, D., Boerma, W., Bourgueil, Y., Cartier, T., Dedeu, T., Hasvold, T., Hutchinson, A., Lember, M., Oleszczyk, M., Pavlic, D. R., Svab, I., Tedeschi, P., Wilm, S., Wilson, A., Windak, A., Zee, J. V. d., & Groenewegen, P. (2013). The strength of primary care in europe: An international comparative study [Publisher: British Journal of General Practice Section: Research]. *British Journal of General Practice*, *63*(616), e742–e750. https://doi.org/10.3399/bjgp13X674422

Kuh, D, Ben-Shlomo, Y, Lynch, J, Hallqvist, J, & Power, C. (2003). Life course epidemiology. *Journal of Epidemiology and Community Health*, *57*(10), 778–783. https://doi.org/10.1136/jech.57.10.778

Kuh, D., Cooper, R., Hardy, R., Richards, M., & Ben-Shlomo, Y. (2013). *A life course approach to healthy ageing*. OUP Oxford.

Kuk, J. L., & Ardern, C. I. (2010). Age and sex differences in the clustering of metabolic syndrome factors: Association with mortality risk. *Diabetes Care*, *33*(11), 2457–2461. https://doi.org/10.2337/dc10-0942

Lai, F., Macmillan, J., Daudelin, D. H., & Kent, D. M. (2006). The potential of training to increase acceptance and use of computerized decision support systems for medical diagnosis [Publisher: SAGE Publications Inc]. *Human Factors*, *48*(1), 95–108. https://doi.org/10.1518/001872006776412306

Lanska, D. J., & Peterson, P. M. (1995). Geographic variation in reporting of stroke deaths to underlying or contributing causes in the united states. *Stroke*, *26*(11), 1999–2003. https://doi.org/10.1161/01.STR.26.11.1999

Lawson, K. D. (2013). *The scottish cardiovascular disease policy model* (Doctoral dissertation). University of Glasgow. Retrieved November 12, 2023, from https://eleanor.lib.gla.ac.uk/record=b3001608

Lee, S., Shafe, A. C., & Cowie, M. R. (2011). UK stroke incidence, mortality and cardiovascular risk management 1999-2008: Time-trend analysis from the general practice research database. *BMJ Open*, *1*(2). https://doi.org/10.1136/bmjopen-2011-000269

Lehnert, T., Sonntag, D., Konnopka, A., Riedel-Heller, S., & König, H.-H. (2013). Economic costs of overweight and obesity. *Best Practice & Research Clinical Endocrinology & Metabolism*, *27*(2), 105–115. https://doi.org/10.1016/j.beem.2013.01.002

Leinster, T., & Cobbold, C. A. (2012). Measuring diversity: The importance of species similarity. *Ecology*. https://doi.org/10.1890/10-2402.1

Levi, F., Lucchini, F., Negri, E., & Vecchia, C. L. (2001). The fall in breast cancer mortality in europe. *European Journal of Cancer*, *37*(11), 1409–1412. https://doi.org/10.1016/S0959-8049(01)00144-7

Levin, K. A. (2003). Urban-rural differences in self-reported limiting long-term illness in scotland. *Journal of Public Health*, *25*(4), 295–302. https://doi.org/10.1093/pubmed/fdg084

Levin, K. A., & Leyland, A. H. (2005). Urban/rural inequalities in suicide in scotland, 1981–1999. *Social Science & Medicine*, *60*(12), 2877–2890. https://doi.org/10.1016/j.socscimed.2004.11.025

Levin, K. A., & Leyland, A. H. (2006). A comparison of health inequalities in urban and rural scotland. *Social Science and Medicine*, *62*(6). https://doi.org/10.1016/j.socscimed.2005.08.045

Leyland, A. H., & Dundas, R. (2010). The social patterning of deaths due to assault in scotland, 1980–2005: Population-based study. *Journal of Epidemiology & Community Health*, *64*(5), 432–439. https://doi.org/10.1136/jech.2009.095018

Liao, L., Allen, L. A., & Whellan, D. J. (2008). Economic burden of heart failure in the elderly. *PharmacoEconomics*, *26*(6). https://doi.org/10.2165/00019053-200826060-00001

Liddell, C., Morris, C., Gray, B., Czerwinska, A., & Thomas, B. (2016). Excess winter mortality associated with alzheimer's disease and related dementias in the UK: A case for energy justice. *Energy Research & Social Science*, *11*, 256–262. https://doi.org/10.1016/j.erss.2015.11.007

Long, K. A. (1993). THE CONCEPT OF HEALTH: Rural perspectives. *Nursing Clinics of North America*, *28*(1), 123–130. https://doi.org/10.1016/S0029-6465(22)02842-0

Long, M. J., Ament, R. P., Dreachslin, J. L., & Kobrinski, E. J. (1985). A reconsideration of economies of scale in the health care field. *Health Policy*, *5*(1), 25–44. https://doi.org/10.1016/0168-8510(85)90064-8

Lopez, A. D., & Adair, T. (2019). Is the long-term decline in cardiovascular-disease mortality in high-income countries over? evidence from national vital statistics net. *International Journal of Epidemiology*, *48*(6), 1815–1823. https://doi.org/10.1093/ije/dyz143

Lu, T.-H., Lee, M.-C., & Chou, M.-C. (2000). Accuracy of cause-of-death coding in taiwan: Types of miscoding and effects on mortality statistics. *International Journal of Epidemiology*, *29*(2), 336–343. https://doi.org/10.1093/ije/29.2.336

Luengo-Fernandez, R., Leal, J., & Gray, A. M. (2012). UK research expenditure on dementia, heart disease, stroke and cancer: Are levels of spending related to disease burden? *European Journal of Neurology*, *19*(1). https://doi.org/10.1111/j.1468-1331.2011.03500.x

Luengo-Fernández, R., Leal, J., Gray, A., Petersen, S., & Rayner, M. (2006). Cost of cardiovascular diseases in the united kingdom. *Heart*, *92*(10). https://doi.org/10.1136/hrt.2005.072173

Lupton, R., & Tunstall, R. (2003). *Is targeting deprived areas an effective means to reach poor people? an assessment of one rationale for area-based funding programmes* (CASE 070). Centre for Analysis of Social Exclusion, LSE.

Macdonald, K., & Morgan, H. M. (2021). The impact of austerity on disabled, elderly and immigrants in the united kingdom: A literature review. *Disability & Society*, *36*(7), 1125–1147. https://doi.org/10.1080/09687599.2020.1779036

Macdonald, L., Olsen, J. R., Shortt, N. K., & Ellaway, A. (2018). Do 'environmental bads' such as alcohol, fast food, tobacco, and gambling outlets cluster and co-locate in more deprived areas in glasgow city, scotland? *Health & Place*, *51*, 224–231. https://doi.org/10.1016/j.healthplace.2018.04.008

Mackenbach, J. P. (1994). The epidemiologic transition theory. *Journal of Epidemiology and Community Health*, *48*(4), 329–331.

Mackenbach, J. P. (2002). Income inequality and population health: Evidence favouring a negative correlation between income inequality and life expectancy has disappeared. *BMJ*, *324*(7328), 1–2. https://doi.org/10.1136/bmj.324.7328.1

Macran, S., Weatherly, H., & Kind, P. (2003). Measuring population health: A comparison of three generic health status measures. *Medical Care*, *41*(2), 218–231.

Magurran, A. E. (2021). Measuring biological diversity. *Current Biology*, *31*(19), R1174–R1177. https://doi.org/10.1016/j.cub.2021.07.049

Malagón, T., Yong, J. H. E., Tope, P., Miller Jr., W. H., Franco, E. L., & McGill Task Force on the Impact of COVID-19 on Cancer Control and Care. (2022). Predicted long-term impact of COVID-19 pandemic-related care delays on cancer mortality in canada. *International Journal of Cancer*, *150*(8), 1244–1254. https://doi.org/10.1002/ijc.33884

Mansfield, K. E., Mathur, R., Tazare, J., Henderson, A. D., Mulick, A. R., Carreira, H., Matthews, A. A., Bidulka, P., Gayle, A., Forbes, H., Cook, S., Wong, A. Y. S., Strongman, H., Wing, K., Warren-Gash, C., Cadogan, S. L., Smeeth, L., Hayes, J. F., Quint, J. K., … Langan, S. M. (2021). Indirect acute effects of the COVID-19 pandemic on physical and mental health in the UK: A population-based study. *The Lancet Digital Health*, *3*(4), e217–e230. https://doi.org/10.1016/S2589-7500(21)00017-0

Marmot, M. (2017). The health gap: Doctors and the social determinants of health. *Scandinavian Journal of Public Health*, *45*(7). https://doi.org/10.1177/1403494817717448

Masisi, L., Nelwamondo, V., & Marwala, T. (2008). The use of entropy to measure structural diversity. https://doi.org/10.48550/arXiv.0810.3525

Matheson, C., Pflanz-Sinclair, C., Aucott, L., Wilson, P., Watson, R., Malloy, S., Dickie, E., & McAuley, A. (2014). Reducing drug related deaths: A pre - implementation assessment of knowledge, barriers and enablers for naloxone distribution through general practice. *BMC Family Practice*, *15*(1), 12. https://doi.org/10.1186/1471-2296-15-12

Maynard, A. (2017). Shrinking the state: The fate of the NHS and social care. *Journal of the Royal Society of Medicine*, *110*(2), 49–51. https://doi.org/10.1177/0141076816686923

Mays, G. P., & Smith, S. A. (2011). Evidence links increases in public health spending to declines in preventable deaths. *Health Affairs*, *30*(8). https://doi.org/10.1377/hlthaff.2011.0196

McAuley, A., Best, D., Taylor, A., Hunter, C., & Robertson, R. (2012). From evidence to policy: The scottish national naloxone programme. *Drugs: Education, Prevention and Policy*, *19*(4), 309–319. https://doi.org/10.3109/09687637.2012.682232

McCartney, G., Popham, F., McMaster, R., & Cumbers, A. (2019). Defining health and health inequalities. *Public Health*, *172*, 22–30. https://doi.org/10.1016/j.puhe.2019.03.023

McCartney, G., Walsh, D., Fenton, L., & Devine, R. (2022). *Resetting the course for population health: Evidence and recommendations to address stalled mortality improvements in scotland and the rest of the UK* (Research Reports or Papers). Glasgow Centre for Population Health.

McCartney, G., Collins, C., Walsh, D., & Batty, D. (2011). *Accounting for scotland's excess mortality: Towards a synthesis*. Glasgow Centre for Population Health. Retrieved December 19, 2022, from https://www.gcph.co.uk/publications/238_accounting_for_scotlands_excess_mortality_towards_a_synthesis

McCartney, G., Walsh, D., Whyte, B., & Collins, C. (2012). 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. https://doi.org/10.1093/eurpub/ckr136

McIsaac, S. J., & Wilkinson, R. G. (1997). Income distribution and cause-specific mortality. *European Journal of Public Health*, *7*(1), 45–53. https://doi.org/10.1093/eurpub/7.1.45

McKee, M., Dunnell, K., Anderson, M., Brayne, C., Charlesworth, A., Johnston-Webber, C., Knapp, M., McGuire, A., Newton, J. N., Taylor, D., & Watt, R. G. (2021). The changing health needs of the UK population. *The Lancet*, *397*(10288), 1979–1991. https://doi.org/10.1016/S0140-6736(21)00229-4

McKee, M., Suhrcke, M., Nolte, E., Lessof, S., Figueras, J., Duran, A., & Menabde, N. (2009). Health systems, health, and wealth: A european perspective. *The Lancet*, *373*(9660), 349–351. https://doi.org/10.1016/S0140-6736(09)60098-2

McKendrick, J. H., Asenova, D., MacRae, C., Reynolds, R., Egan, J., Hastings, A., Mooney, G., & Sinclair, S. (2016). Conceptualising austerity in scotland as a risk shift: Ideas and implications. *Scottish Affairs*, *25*(4), 451–478. https://doi.org/10.3366/scot.2016.0152

McKenzie, K., Murray, A., & Booth, T. (2013). Do urban environments increase the risk of anxiety, depression and psychosis? an epidemiological study. *Journal of Affective Disorders*, *150*(3), 1019–1024. https://doi.org/10.1016/j.jad.2013.05.032

McLoone, P., & Boddy, F. A. (1994). Deprivation and mortality in scotland, 1981 and 1991. *BMJ*, *309*(6967), 1465–1470. https://doi.org/10.1136/bmj.309.6967.1465

McMonagle, C., Brown, D., Reeve, R., & Mancy, R. (2022). Trends in the diversity of mortality causes and age-standardised mortality rates among subpopulations within scotland, 2001–2019. *SSM - Population Health*, *19*, 101192. https://doi.org/10.1016/j.ssmph.2022.101192

Meier, P. S. (2010). Polarized drinking patterns and alcohol deregulation: Trends in alcohol consumption, harms and policy: United kingdom 1990-2010. *NAT Nordisk alkohol & narkotikatidskrift*, *27*(5).

Mellish, T. I., Luzmore, N. J., & Shahbaz, A. A. (2020). Why were the UK and USA unprepared for the COVID-19 pandemic? the systemic weaknesses of neoliberalism: A comparison between the UK, USA, germany, and south korea. *Journal of Global Faultlines*, *7*, 9–45. https://doi.org/10.13169/jglobfaul.7.1.0009

Merchant, H. A., Kow, C. S., & Hasan, S. S. (2021). COVID-19 first anniversary review of cases, hospitalization, and mortality in the UK. *Expert Review of Respiratory Medicine*, *15*(8), 973–978. https://doi.org/10.1080/17476348.2021.1890035

Miall, N., Fergie, G., & Pearce, A. (2022). *Health inequalities in scotland: Trends in deaths, health and wellbeing, health behaviours, and health services since 2000*. University of Glasgow. https://doi.org/10.36399/gla.pubs.282637

Mishel, M. H. (1988). Uncertainty in illness. *Image: the Journal of Nursing Scholarship*, *20*(4), 225–232. https://doi.org/10.1111/j.1547-5069.1988.tb00082.x

Missov, T. I., Németh, L., & Dańko, M. J. (2016). How much can we trust life tables? sensitivity of mortality measures to right-censoring treatment. *Palgrave Communications*, *2*(1), 1–10. https://doi.org/10.1057/palcomms.2015.49

Mitchell, R. (2005). Commentary: The decline of death—how do we measure and interpret changes in self-reported health across cultures and time? *International Journal of Epidemiology*, *34*(2), 306–308. https://doi.org/10.1093/ije/dyh361

Mitchell, S. (2019). *The measurement, dynamics, and interpretation of biological diversity* (Doctoral dissertation). University of Glasgow. 10.5525/gla.thesis.41212

Mitchell, S., Reeve, R., & White, T. (2020). *Rdiversity*. https://github.com/boydorr/rdiversity

Mitratza, M., Kunst, A. E., & Kardaun, J. W. P. F. (2019). Detecting mortality trends in the netherlands across 625 causes of death. *International Journal of Environmental Research and Public Health*, *16*(21), 4150. https://doi.org/10.3390/ijerph16214150

Mittra, J., Mastroeni, M., Haddow, G., Wield, D., & Barlow, E. (2019). Re-imagining healthcare and medical research systems in post-devolution scotland. *Sociological Research Online*, *24*(1). https://doi.org/10.1177/1360780418823221

Moore, G., Audrey, S., Barker, M., Bond, L., Bonell, C., Cooper, C., Hardeman, W., Moore, L., O'Cathain, A., Tinati, T., Wight, D., & Baird, J. (2014). Process evaluation in complex public health intervention studies: The need for guidance. *J Epidemiol Community Health*, *68*(2), 101–102. https://doi.org/10.1136/jech-2013-202869

Moreno, C. E., & Rodríguez, P. (2010). A consistent terminology for quantifying species diversity? *Oecologia*, *163*(2), 279–282. https://doi.org/10.1007/s00442-010-1591-7

Moser, K., Shkolnikov, V. M., & Leon, D. A. (2005). World mortality 1950–2000: Divergence replaces convergence from the late 1980s. *Bulletin of the World Health Organization*. https://doi.org/10.1007/978-1-4020-6174-5_1

Muggeo, V. M. R. (2008). Segmented: An r package to fit regression models with broken-line relationships. *R News*, *8*(1), 20–25. https://cran.r-project.org/doc/Rnews/

Murphy, M. (2021). Recent mortality in britain: A review of trends and explanations. *Age and ageing*, *50*(3), 676–683. https://doi.org/10.1093/ageing/afab016

Murray, C. J. L., Salomon, J. A., & Mathers, C. (2000). A critical examination of summary measures of population health. *Bulletin of the World Health Organization*, *78*, 981–994. https://doi.org/10.1590/S0042-96862000000800008

Muszyńska-Spielauer, M., & Luy, M. (2022). Well-being adjusted health expectancy: A new summary measure of population health. *European Journal of Population*. https://doi.org/10.1007/s10680-022-09628-1

Naghavi, M., Wang, H., Lozano, R., Davis, A., Liang, X., Zhou, M., Vollset, S. E., Abbasoglu Ozgoren, A., Abdalla, S., Abd-Allah, F., Abdel Aziz, M. I., Abera, S. F., Aboyans, V., Abraham, B., Abraham, J. P., Abuabara, K. E., Abubakar, I., Abu-Raddad, L. J., Abu-Rmeileh, N. M., … Temesgen, A. M. (2015). Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: A systematic analysis for the global burden of disease study 2013. *The Lancet*, *385*(9963). https://doi.org/10.1016/S0140-6736(14)61682-2

National Records of Scotland. (n.d.). *Cause of death text and codes*. Retrieved November 18, 2022, from https://www.nrscotland.gov.uk/files/statistics/vital-events/ve-deaths-cause-of-death-text-codes.pdf

National Records of Scotland. (2011). *Changes to the coding of causes of death between 2010 and 2011*. Retrieved April 28, 2021, from https://www.nrscotland.gov.uk/files/statistics/vital-events/changes-to-coding-of-causes-of-death-between-2010-2011.pdf

National Records of Scotland. (2021). *Small area mid-year population estimates NHS open data*. https://www.nrscotland.gov.uk/statistics-and-data/statistics/statistics-by-theme/population/population-estimates/mid-year-population-estimates

National Records of Scotland. (2022a). *Life expectancy in scotland 2019-2021*. Retrieved March 14, 2023, from https://www.nrscotland.gov.uk/files/statistics/life-expectancy-in-scotland/19-21/life-expectancy-19-21-report.pdf

National Records of Scotland. (2022b). *Mid-2021 small area population estimates, scotland*. Retrieved December 20, 2022, from https://www.nrscotland.gov.uk/files//statistics/population-estimates/sape-2021/sape-21-report.pdf

National Records of Scotland & Scottish Government. (2017). The impact of the implementation of IRIS software for ICD-10 cause of death coding on mortality statistics in scotland. Retrieved September 28, 2021, from https://www.nrscotland.gov.uk/files/statistics/vital-events/impact-of-implementation-iris-for-icd.pdf

National Records of Scotland Web Team. (2013, May 31). *Death certificates and coding the causes of death*. Retrieved January 31, 2023, from https://www.nrscotland.gov.uk/statistics-and-data/statistics/statistics-by-theme/vital-events/deaths/deaths-background-information/death-certificates-and-coding-the-causes-of-death

Nazroo, J. Y. (2003). The structuring of ethnic inequalities in health: Economic position, racial discrimination, and racism. *American Journal of Public Health*, *93*(2), 277–284. https://doi.org/10.2105/AJPH.93.2.277

Neuwirth, E. (2014). *RColorBrewer: ColorBrewer palettes*. https://CRAN.R-project.org/package=RColorBrewer

Newton, J. N., Briggs, A. D. M., Murray, C. J. L., Dicker, D., Foreman, K. J., Wang, H., Naghavi, M., Forouzanfar, M. H., Ohno, S. L., Barber, R. M., Vos, T., Stanaway, J. D., Schmidt, J. C., Hughes, A. J., Fay, D. F. J., Ecob, R., Gresser, C., McKee, M., Rutter, H., … Davis, A. C. J. (2015). Changes in health in england, with analysis by english regions and areas of deprivation, 1990–2013: A systematic analysis for the global burden of disease study 2013. *The Lancet*, *386*(10010), 2257–2274. https://doi.org/10.1016/S0140-6736(15)00195-6

Nigri, A., Barbi, E., & Levantesi, S. (2021). The relationship between longevity and lifespan variation. *Statistical Methods and Applications*. https://doi.org/10.1007/s10260-021-00584-4

Norman, P., Boyle, P., Exeter, D., Feng, Z., & Popham, F. (2011). Rising premature mortality in the UK's persistently deprived areas: Only a scottish phenomenon? *Social Science & Medicine*, *73*(11), 1575–1584. https://doi.org/10.1016/j.socscimed.2011.09.034

Noss, R. F. (1990). Indicators for monitoring biodiversity: A hierarchical approach. *Conservation Biology*. https://doi.org/10.1111/j.1523-1739.1990.tb00309.x

Numbers, K., & Brodaty, H. (2021). The effects of the COVID-19 pandemic on people with dementia. *Nature Reviews Neurology*, *17*(2), 69–70. https://doi.org/10.1038/s41582-020-00450-z

Németh, L., & Missov, T. I. (2018). Adequate life-expectancy reconstruction for adult human mortality data. *PLOS ONE*, *13*(6), e0198485. https://doi.org/10.1371/journal.pone.0198485

OECD.stat. (2023). *Mortality (by week) : Excess deaths by week, 2020-2023*. https://stats.oecd.org/index.aspx?queryid=104676

Office for National Statistics. (2022). *Monthly mortality analysis, england and wales - office for national statistics - december 2022*. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/deaths/bulletins/monthlymortalityanalysisenglandandwales/december2022

O'Malley, K. J., Cook, K. F., Price, M. D., Wildes, K. R., Hurdle, J. F., & Ashton, C. M. (2005). Measuring diagnoses: ICD code accuracy. *Health Services Research*, *40*(5), 1620–1639. https://doi.org/10.1111/j.1475-6773.2005.00444.x

Omran, A. R. (1971). The epidemiologic transition: A theory of the epidemiology of population change. *The Milbank Quarterly*, *83*(4), 731–757. https://doi.org/10.1111/j.1468-0009.2005.00398.x

Omran, A. R. (1998). The epidemiologic transition theory revisited thirty years later. *World Health Statistics Quarterly*, *53*, 21. https://doi.org/https://apps.who.int/iris/handle/10665/330604

Paplikar, A., Rajagopalan, J., & Alladi, S. (2022). Care for dementia patients and caregivers amid COVID-19 pandemic. *Cerebral Circulation - Cognition and Behavior*, *3*, 100040. https://doi.org/10.1016/j.cccb.2022.100040

Pascariu, M. D. (2022). *MortalityLaws: Parametric mortality models, life tables and HMD*. https://CRAN.R-project.org/package=MortalityLaws

Pateman, T. (2011). Rural and urban areas: Comparing lives using rural/urban classifications. *Regional Trends*, *43*(1), 11–86. https://doi.org/10.1057/rt.2011.2

Patton, G. C., Coffey, C., Sawyer, S. M., Viner, R. M., Haller, D. M., Bose, K., Vos, T., Ferguson, J., & Mathers, C. D. (2009). Global patterns of mortality in young people: A systematic analysis of population health data. *The Lancet*, *374*(9693), 881–892. https://doi.org/10.1016/S0140-6736(09)60741-8

Payne, N. W. S., Brown, K. F., Delon, C., Kotrotsios, Y., Soerjomataram, I., & Shelton, J. (2022). Socio-economic deprivation and cancer incidence in england: Quantifying the role of smoking. *PLOS ONE*, *17*(9), e0272202. https://doi.org/10.1371/journal.pone.0272202

Permanyer, I., & Scholl, N. (2019). Global trends in lifespan inequality: 1950-2015 [Publisher: Public Library of Science]. *PLOS ONE*, *14*(5), e0215742. https://doi.org/10.1371/journal.pone.0215742

Permanyer, I., Spijker, J., & Blanes, A. (2022). On the measurement of healthy lifespan inequality. *Population Health Metrics*, *20*(1), 1. https://doi.org/10.1186/s12963-021-00279-8

Permanyer, I., Spijker, J., Blanes, A., & Renteria, E. (2018). Longevity and lifespan variation by educational attainment in spain: 1960–2015. *Demography*, *55*(6). https://doi.org/10.1007/s13524-018-0718-z

Petchey, O. L., & Gaston, K. J. (2006). Functional diversity: Back to basics and looking forward. *Ecology Letters*, *9*(6), 741–758. https://doi.org/10.1111/j.1461-0248.2006.00924.x

Petrou, S., & Kupek, E. (2019). Epidemiological trends and risk factors for tobacco, alcohol and drug use among adolescents in scotland, 2002-13. *Journal of Public Health (United Kingdom)*, *41*(1). https://doi.org/10.1093/pubmed/fdy006

Phin, N. (2021). Public health scotland - the first year: Successes and lessons. *Journal of the Royal College of Physicians of Edinburgh*, *51*(1), 34–39. https://doi.org/10.4997/jrcpe.2021.239

Piffaretti, C., Moreno-Betancur, M., Lamarche-Vadel, A., & Rey, G. (2016). Quantifying cause-related mortality by weighting multiple causes of death. *Bulletin of the World Health Organization*, *94*(12), 870–879. https://doi.org/10.2471/BLT.16.172189

Pilar, M. R., Eyler, A. A., Moreland-Russell, S., & Brownson, R. C. (2020). Actual causes of death in relation to media, policy, and funding attention: Examining public health priorities. *Frontiers in Public Health*, *8*.

Plana-Ripoll, O., Dreier, J. W., Momen, N. C., Prior, A., Weye, N., Mortensen, P. B., Pedersen, C. B., Iburg, M., Christensen, M. K., Munk, T., Laursen, Agerbo, E., Pedersen, M. G., Brandt, Frohn, M., Geels, C., Christensen, J. H., & Mcgrath, J. J. (2022). Analysis of mortality metrics associated with a comprehensive range of disorders in denmark , 2000 to 2018 : A population-based cohort study. *PLoS medicine*, 1–19. https://doi.org/10.1371/journal.pmed.1004023

Preston, S. H., Heuveline, P., & Guillot, M. (2001). *Demography: Measuring and modeling population processes*. Blackwell Publishers.

Prowle, M., & Harradine, D. (2014). Austerity and financial governance: A UK case study of the national health service. *International Journal of Public Sector Management*, *27*(3), 212–224. https://doi.org/10.1108/IJPSM-02-2013-0028

Public Health Scotland. (n.d.-a). *COVID-19 daily dashboard*. Retrieved October 17, 2022, from https://public.tableau.com/views/COVID-19DailyDashboard_15960160643010/

Public Health Scotland. (n.d.-b). COVID-19: ICD-10 analytical guidance, 20. Retrieved October 27, 2022, from https://www.isdscotland.org/Products-and-Services/Terminology-Services/docs/COVID-19-ICD-10-Analytical-Guidance-V1.3.pdf

Public Health Scotland. (n.d.-c). *Daily COVID-19 Cases in Scotland - Scottish Health and Social Care Open Data*. Retrieved October 17, 2022, from https://www.opendata.nhs.scot/tr/dataset/covid-19-in-scotland

Public Health Scotland. (2022a). *Cancer mortality in scotland - annual update to 2021 - cancer mortality - publications - public health scotland*. Retrieved November 12, 2023, from https://publichealthscotland.scot/publications/cancer-mortality/cancer-mortality-in-scotland-annual-update-to-2021/view-the-cancer-mortality-data-files/#Lung%20and%20mesothelioma/respiratory

Public Health Scotland. (2022b). *Scottish heart disease statistics*. https://publichealthscotl
scot/media/11607/2022-01-25-heart-disease-report.pdf

Public Health Scotland. (2023). *Overview of the COVID-19 data dashboard - COVID-19 daily cases in scotland dashboard - COVID-19 data and intelligence - COVID-19 - conditions and diseases - our areas of work - public health scotland*. Retrieved January 6, 2023, from https://publichealthscotland.scot/our-areas-of-work/conditions-and-diseases/covid-19/covid-19-data-and-intelligence/covid-19-daily-cases-in-scotland-dashboard/overview-of-the-covid-19-data-dashboard/

Pugh, C. M., Watson, A., Bell, R. H., Brasel, K. J., Jackson, G. P., Weber, S. M., & Kao, L. S. (2009). Surgical education in the internet era. *Journal of Surgical Research*, *156*(2), 177–182. https://doi.org/10.1016/j.jss.2009.03.021

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Raleigh, V. (2022). What is driving excess deaths in england and wales? [Publisher: British Medical Journal Publishing Group Section: Opinion]. *BMJ*, *379*, o2524. https://doi.org/10.1136/bmj.o2524

Ramezanian, M., Haghdoost, A. A., Mehrolhassani, M. H., Abolhallaje, M., Dehnavieh, R., Najafi, B., & Fazaeli, A. A. (2019). Forecasting health expenditures in iran using the ARIMA model (2016-2020). *Medical Journal of the Islamic Republic of Iran*, *33*, 25. https://doi.org/10.34171/mjiri.33.25

Ramsay, J., Minton, J., Fischbacher, C., Fenton, L., Kaye-Bardgett, M., Wyper, G. M., Richardson, E., & McCartney, G. (2020). How have changes in death by cause and age group contributed to the recent stalling of life expectancy gains in scotland? comparative decomposition analysis of mortality data, 2000-2002 to 2015-2017. *BMJ Open*, *10*(10), e036529. https://doi.org/10.1136/bmjopen-2019-036529

Rao, C. R. (1982a). Diversity and dissimilarity coefficients: A unified approach. *Theoretical Population Biology*, *21*(1), 24–43. https://doi.org/10.1016/0040-5809(82)90004-1

Rao, C. R. (1982b). Diversity: Its measurement, decomposition, apportionment and analysis. *Sankhyā: The Indian Journal of Statistics, Series A (1961-2002)*, *44*(1), 1–22.

Rao, C. (2020). Medical certification of cause of death for COVID-19. *Bulletin of the World Health Organization*, *98*(5), 298–298A. https://doi.org/10.2471/BLT.20.257600

Rashed, T., & Jürgens, C., eds. *Remote sensing of urban and suburban areas*. Vol. 10. Remote Sensing and Digital Image Processing. Dordrecht: Springer Netherlands, 2010. ISBN: 978-1-4020-4371-0 978-1-4020-4385-7. https://doi.org/10.1007/978-1-4020-4385-7.

Rattka, M., Dreyhaupt, J., Winsauer, C., Stuhler, L., Baumhardt, M., Thiessen, K., Rottbauer, W., & Imhof, A. (2021). Effect of the COVID-19 pandemic on mortality of patients with STEMI: A systematic review and meta-analysis. *Heart*, *107*(6), 482–487. https://doi.org/10.1136/heartjnl-2020-318360

Redelings, M. D., Wise, M., & Sorvillo, F. (2007). Using multiple cause-of-death data to investigate associations and causality between conditions listed on the death certificate. *American Journal of Epidemiology*, *166*(1), 104–108. https://doi.org/10.1093/aje/kwm037

Reeve, R., Leinster, T., Cobbold, C. A., Thompson, J., Brummitt, N., Mitchell, S. N., & Matthews, L. (2016). How to partition diversity. *arXiv*, 1404.6520. https://doi.org/10.1111/j.1462-2920.2009.01900.x

Rehder, K., Lusk, J., & Chen, J. I. (2021). Deaths of despair: Conceptual and clinical implications. *Cognitive and Behavioral Practice*, *28*(1), 40–52. https://doi.org/10.1016/j.cbpra.2019.10.002

Reidpath, D. D., & Allotey, P. (2003). Infant mortality rate as an indicator of population health. *Journal of Epidemiology & Community Health*, *57*(5), 344–346. https://doi.org/10.1136/jech.57.5.344

Reilly, J., Crawford, D., & O'Boyle, D. (2020). *Care home review a rapid review of factors relevant to the management of COVID-19 in the care home environment in scotland*. Scottish Government. https://www.gov.scot/publications/root-cause-analysis-care-home-outbreaks/

Revelle, W. (2022). *Psych: Procedures for psychological, psychometric, and personality research*. Northwestern University. https://CRAN.R-project.org/package=psych

Riaz, S. P., Horton, M., Kang, J., Mak, V., Lüchtenborg, M., & Møller, H. (2011). Lung cancer incidence and survival in england: An analysis by socioeconomic deprivation and urbanization. *Journal of Thoracic Oncology*, *6*(12), 2005–2010. https://doi.org/10.1097/JTO.0b013e31822b02db

Richardson, E., Fenton, L., Parkinson, J., Pulford, A., Taulbut, M., McCartney, G., & Robinson, M. (2020). The effect of income-based policies on mortality inequalities in scotland: A modelling study. *The Lancet Public Health*, *5*(3), e150–e156. https://doi.org/10.1016/S2468-2667(20)30011-6

Robine, J.-M. (2001). Redefining the stages of the epidemiological transition by a study of the dispersion of life spans: The case of france. *Population: An English Selection*, *13*(1), 173–193.

Robine, J.-M., Crimmins, E. M., Horiuchi, S., & Yi, Z., eds. *Human longevity, individual life duration, and the growth of the oldest-old population*. Dordrecht: Springer Netherlands, 2007. ISBN: 978-1-4020-4846-3 978-1-4020-4848-7.

Robinson, M., Mackay, D., Giles, L., Lewsey, J., Richardson, E., & Beeston, C. (2021). Evaluating the impact of minimum unit pricing (MUP) on off-trade alcohol sales in scotland: An interrupted time–series study. *Addiction*, *116*(10). https://doi.org/10.1111/add.15478

Routledge, R. D. (1979). Diversity indices: Which ones are admissible? *Journal of Theoretical Biology*, *76*(4), 503–515. https://doi.org/10.1016/0022-5193(79)90015-8

Roy, B., Riley, C., Sears, L., & Rula, E. Y. (2018). Collective well-being to improve population health outcomes: An actionable conceptual model and review of the literature. *American Journal of Health Promotion*, *32*(8), 1800–1813. https://doi.org/10.1177/0890117118791993

Rushworth, G. F., Diack, L., MacRobbie, A., Munoz, S.-A., Pfleger, S., & Stewart, D. (2015). Access to medicines in remote and rural areas: A survey of residents in the scottish highlands & western isles. *Public Health*, *129*(3), 244–251. https://doi.org/10.1016/j.puhe.2015.01.005

Rényi, A. On measures of entropy and information. In: In *Proceedings of the fourth berkeley symposium on mathematical statistics and probability. 1*. Berkeley, California, USA, 1961.

Salive, M. E. (2013). Multimorbidity in older adults. *Epidemiologic Reviews*, *35*(1), 75–83. https://doi.org/10.1093/epirev/mxs009

Santosa, A., Wall, S., Fottrell, E., Högberg, U., & Byass, P. (2014). The development and experience of epidemiological transition theory over four decades: A systematic review. *Global Health Action*, *7*, 23574. https://doi.org/10.3402/gha.v7.23574

Sasson, I. (2016). Trends in life expectancy and lifespan variation by educational attainment: United states, 1990–2010. *Demography*, *53*(2). https://doi.org/10.1007/s13524-015-0453-7

Savarese, G., & D'Amario, D. Sex differences in heart failure. In: In *Sex-specific analysis of cardiovascular function*. Vol. 1065. Cham: Springer International Publishing, 2018, pp. 529–544. ISBN: 978-3-319-77931-7.

Schaffer, A. L., Dobbins, T. A., & Pearson, S.-A. (2021). Interrupted time series analysis using autoregressive integrated moving average (ARIMA) models: A guide for evaluating large-scale health interventions. *BMC Medical Research Methodology*, *21*(1), 58. https://doi.org/10.1186/s12874-021-01235-8

Certification of Death Act (2011) (2011).

Scottish Government. (2014). *Heart disease improvement plan*. Retrieved February 6, 2021, from https://www.gov.scot/publications/heart-disease-improvement-plan/

Scottish Government. (2020). *SIMD 2020 technical notes*. Retrieved February 1, 2023, from https://www.gov.scot/binaries/content/documents/govscot/publications/statistics/2020/09/simd-2020-technical-notes/documents/simd-2020-technical-notes/simd-2020-technical-notes/govscot%3Adocument/SIMD%2B2020%2Btechnical%2Bnotes.pdf

Seaman, R., Leyland, A. H., & Popham, F. (2016a). How have trends in lifespan variation changed since 1950? a comparative study of 17 western european countries. *European Journal of Public Health*, *26*(2), 360–362. https://doi.org/10.1093/eurpub/ckv185

Seaman, R., Mitchell, R., Dundas, R., Leyland, A., & Popham, F. (2015). How much of the difference in life expectancy between scottish cities does deprivation explain? *BMC Public Health*, *15*(1), 1057. https://doi.org/10.1186/s12889-015-2358-1

Seaman, R. (2017). *Total mortality inequality in scotland: The case for measuring lifespan variation* (Doctoral dissertation). University of Glasgow. https://theses.gla.ac.uk/8318/1/2016SeamanPhD.pdf

Seaman, R., Leyland, A. H., & Popham, F. (2016b). Increasing inequality in age of death at shared levels of life expectancy: A comparative study of scotland and england and wales. *SSM - Population Health*, *2*, 724–731. https://doi.org/10.1016/j.ssmph.2016.10.001

Seaman, R., Riffe, T., Leyland, A. H., Popham, F., & van Raalte, A. (2019). The increasing lifespan variation gradient by area-level deprivation: A decomposition analysis of scotland 1981–2011. *Social Science and Medicine*, *230*. https://doi.org/10.1016/j.socscimed.2019.04.008

Sears, L. E., Agrawal, S., Sidney, J. A., Castle, P. H., Rula, E. Y., Coberley, C. R., Witters, D., Pope, J. E., & Harter, J. K. (2014). The well-being 5: Development and validation of a diagnostic instrument to improve population well-being. *Population Health Management*, *17*(6), 357–365. https://doi.org/10.1089/pop.2013.0119

Seretakis, D., Lagiou, P., Lipworth, L., Signorello, L. B., Rothman, K. J., & Trichopoulos, D. (1997). Changing seasonality of mortality from coronary heart disease. *JAMA*, *278*(12), 1012–1014. https://doi.org/10.1001/jama.1997.03550120072036

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*(3). https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Shkolnikov, V. M., Andreev, E. M., & Begun, A. (2003). Gini coefficient as a life table function: Computation from discrete data, decomposition of differences and empirical examples. *Demographic Research*, *8*(11), 305–358. https://doi.org/10.4054/DemRes.2003.8.11

Shkolnikov, V. M., & Andreev, E. M. (2010, May). *Spreadsheet for calculation of life-table dispersion measures* (TR-2010-001). Max Planck Institute for Demographic Research. Rostock. Retrieved August 7, 2022, from https://www.demogr.mpg.de/en/publications_databases_6118/publications_1904/mpidr_technical_reports/spreadsheet_for_calculation_of_life_table_dispersion_measures_3837

Silles, M. A. (2009). The causal effect of education on health: Evidence from the united kingdom. *Economics of Education Review*, *28*(1), 122–128. https://doi.org/10.1016/j.econedurev.2008.02.003

Simpkin, A. L., & Schwartzstein, R. M. (2016). Tolerating uncertainty — the next medical revolution? *New England Journal of Medicine*, *375*(18), 1713–1715. https://doi.org/10.1056/NEJMp1606402

Simpson, E. H. (1949). Measurement of diversity. *Nature*. https://doi.org/10.1038/163688a0

Singh, G. K., Azuine, R. E., Siahpush, M., & Kogan, M. D. (2013). All-cause and cause-specific mortality among US youth: Socioeconomic and rural–urban disparities and international patterns. *Journal of Urban Health*, *90*(3), 388–405. https://doi.org/10.1007/s11524-012-9744-0

Smith, G. D., Bartley, M, & Blane, D. (1990). The black report on socioeconomic inequalities in health 10 years on. *BMJ : British Medical Journal*, *301*(6748), 373–377.

Smyth, H., Gorey, S., O'Keeffe, H., Beirne, J., Kelly, S., Clifford, C., Kerr, H., Mulroy, M., & Ahern, T. (2022). Generalist vs specialist acute medical admissions - what is the impact of moving towards acute medical subspecialty admissions on efficacy of care provision? [Publisher: Elsevier]. *European Journal of Internal Medicine*, *98*, 47–52. https://doi.org/10.1016/j.ejim.2021.12.007

Sridhar, D. (2020). Staying sane in our new COVID-19 world. *The Lancet*, *396*(10262), 1552. https://doi.org/10.1016/S0140-6736(20)32345-X

Sridhar, D. (2022). What's behind the mystery of thousands of excess deaths this summer? *The Guardian*. https://www.theguardian.com/commentisfree/2022/sep/13/whats-behind-the-mystery-of-thousands-of-excess-deaths-this-summer

Stark, D. E., & Shah, N. H. (2017). Funding and publication of research on gun violence and other leading causes of death. *JAMA*, *317*(1), 84–85. https://doi.org/10.1001/jama.2016.16215

Steere-Williams, J. (2022). Endemic fatalism and why it will not resolve COVID-19. *Public Health*, *206*, 29–30. https://doi.org/10.1016/j.puhe.2022.02.011

Steingart, R. M., Packer, M., Hamm, P., Coglianese, M. E., Gersh, B., Geltman, E. M., Sollano, J., Katz, S., Moyé, L., Basta, L. L., Lewis, S. J., Gottlieb, S. S., Bernstein, V., McEwan, P., Jacobson, K., Brown, E. J., Kukin, M. L., Kantrowitz, N. E., & Pfeffer, M. A. (1991). Sex differences in the management of coronary artery disease. *New England Journal of Medicine*, *325*(4), 226–230. https://doi.org/10.1056/NEJM199107253250402

Sud, A., Torr, B., Jones, M. E., Broggio, J., Scott, S., Loveday, C., Garrett, A., Gronthoud, F., Nicol, D. L., Jhanji, S., Boyce, S. A., Williams, M., Riboli, E., Muller, D. C., Kipps, E., Larkin, J., Navani, N., Swanton, C., Lyratzopoulos, G., … Turnbull, C. (2020). Effect of delays in the 2-week-wait cancer referral pathway during the COVID-19 pandemic on cancer survival in the UK: A modelling study [Publisher: Elsevier]. *The Lancet Oncology*, *21*(8). https://doi.org/10.1016/S1470-2045(20)30392-2

Sudharsanan, N., Aburto, J. M., Riffe, T., & van Raalte, A. (2022). Large variation in the epidemiological transition across countries: Is it still valuable as a mortality theory? *International Journal of Epidemiology*, *51*(4), 1057–1061. https://doi.org/10.1093/ije/dyac107

Sundquist, K., Malmström, M., & Johansson, S.-E. (2004). Neighbourhood deprivation and incidence of coronary heart disease: A multilevel study of 2.6 million women and men in sweden. *Journal of Epidemiology & Community Health*, *58*(1), 71–77. https://doi.org/10.1136/jech.58.1.71

Sutton, M. (2002). Vertical and horizontal aspects of socio-economic inequity in general practitioner contacts in scotland. *Health Economics*, *11*(6), 537–549. https://doi.org/10.1002/hec.752

Taiyun Wei, Viliam Simko, Michael Levy, Yihui Xie, Yan Jin, Jeff Zemla, Moritz Freidank, Jun Cai & Tomas Protivinsky. (2021). *Package 'corrplot'* (Version 0.92). Retrieved August 15, 2022, from https://cran.r-project.org/web/packages/corrplot/corrplot.pdf

Tapia Granados, J. A., & Ionides, E. L. (2017). Population health and the economy: Mortality and the great recession in europe. *Health Economics*, *26*(12), e219–e235. https://doi.org/10.1002/hec.3495

Taylor, D. (2016). The pharmaceutical industry and the future of drug development. *Pharmaceuticals in the Environment*, *2016*. https://doi.org/10.1039/9781782622345-00001

Telfair, J., & Shelton, T. L. (2012). Educational attainment as a social determinant of health. *North Carolina Medical Journal*, *73*(5), 358–365. https://doi.org/10.18043/ncm.73.5.358

The Health Foundation. (2019). *Mortality and life expectancy trends in the UK*. Retrieved March 27, 2023, from https://www.health.org.uk/publications/reports/mortality-and-life-expectancy-trends-in-the-uk

The Health Foundation. (2023). *General practice tracker*. https://www.health.org.uk/news-and-comment/charts-and-infographics/general-practice-tracker

The Scottish Government. (2017). *Scottish index of multiple deprivation (SIMD) 2016*. https://doi.org/10.2307/j.ctt1t892cc.53

Thom, T. J., Epstein, F. H., Feldman, J. J., & Leaverton, P. E. (1985). Trends in total mortality and mortality from heart disease in 26 countries from 1950 to 1978. *International Journal of Epidemiology*, *14*(4), 510–520. https://doi.org/10.1093/ije/14.4.510

Toch-Marquardt, M., Menvielle, G., Eikemo, T. A., Kulhánová, I., Kulik, M. C., Bopp, M., Esnaola, S., Jasilionis, D., Mäki, N., Martikainen, P., Regidor, E., Lundberg, O., Mackenbach, J. P., & Consortium, f. t. E.-G.-S. (2014). Occupational class inequalities in all-cause and cause-specific mortality among middle-aged men in 14 european populations during the early 2000s. *PLOS ONE*, *9*(9), e108072. https://doi.org/10.1371/journal.pone.0108072

Trias-Llimós, S., & Permanyer, I. (2023). Cause-of-death diversity from a multiple-cause perspective in the united states. *Demography*, 10410415. https://doi.org/10.1215/00703370-10410415

Tucker, C. M., & Cadotte, M. W. (2013). Unifying measures of biodiversity: Understanding when richness and phylogenetic diversity should be congruent. *Diversity and Distributions*, *19*(7), 845–854. https://doi.org/10.1111/ddi.12087

Tucker, R., Moffatt, F., & Timmons, S. (2022). Austerity on the frontline- a preliminary study of physiotherapists working in the national health service in the UK. *Physiotherapy Theory and Practice*, *38*(8), 1037–1049. https://doi.org/10.1080/09593985.2020.1812139

Turnock, B. (2012). *Public health* [e]. Jones & Bartlett Publishers.

Valles, S. A. (2018). *Philosophy of population health: Philosophy for a new public health era* (1st ed.). Routledge. https://doi.org/10.4324/9781315163307

van Raalte, A. A., & Caswell, H. (2013). Perturbation analysis of indices of lifespan variability. *Demography*, *50*(5), 1615–1640. https://doi.org/10.1007/s13524-013-0223-3

van Raalte, A. A., Martikainen, P., & Myrskylä, M. (2013). Lifespan variation by occupational class: Compression or stagnation over time? *Demography*, *51*(1), 73–95. https://doi.org/10.1007/s13524-013-0253-x

Van Raalte, A. A., Sasson, I., & Martikainen, P. (2018). The case for monitoring life-span inequality. *Science*, *362*(6418). https://doi.org/10.1126/science.aau5811

Van Rossum, T., Ferretti, P., Maistrenko, O. M., & Bork, P. (2020). Diversity within species: Interpreting strains in microbiomes. *Nature Reviews Microbiology*, *18*(9), 491–506. https://doi.org/10.1038/s41579-020-0368-1

Vaupel, J., Thatcher, R., & Kannisto, V. (1998). *The force of mortality at ages 80 to 120*. Odense University Press.

Vaupel, J. W., Zhang, Z., & Raalte, A. A. v. (2011). Life expectancy and disparity: An international comparison of life table data. *BMJ Open*, *1*(1), e000128. https://doi.org/10.1136/bmjopen-2011-000128

Verheij, R., Maas, J., & Groenewegen, P. (2008). Urban—rural health differences and the availability of green space. *European Urban and Regional Studies*, *15*(4), 307–316. https://doi.org/10.1177/0969776408095107

Vilar-Compte, D., Camacho-Ortiz, A., & Ponce-de León, S. (2017). Infection control in limited resources countries: Challenges and priorities. *Current Infectious Disease Reports*, *19*(5), 20. https://doi.org/10.1007/s11908-017-0572-y

Villegas, A. M., & Hiabu, M. (2021, October 11). *Estimating and modelling mortality rates in the absence of population data*. Retrieved October 7, 2023, from https://www.actuaries.org/IAA/Documents/SECTIONS/Sections%20Colloquium%202021/20211005_JointSectionColloquium2021_AVillegas_Slides.pdf

Walker, B. H. (2021). *Urban-rural differences in lifespan variation in the united states* (Doctoral dissertation). Mississippi State University.

Walkowiak, M. P., & Walkowiak, D. (2022). Underestimation in reporting excess COVID-19 death data in poland during the first three pandemic waves. *International Journal of Environmental Research and Public Health*, *19*(6), 3692. https://doi.org/10.3390/ijerph19063692

Wallace, M. (2022). Mortality advantage reversed: The causes of death driving all-cause mortality differentials between immigrants, the descendants of immigrants and ancestral natives in sweden, 1997–2016. *European Journal of Population*. https://doi.org/10.1007/s10680-022-09637-0

Wallace, R. B., Remington, P. L., & Wadland, W. C. (2021). The meaning of population health: Whose population is this anyway? *American Journal of Preventive Medicine*, *60*(2), 303–305. https://doi.org/10.1016/j.amepre.2020.09.012

Walsh, D., Bendel, N., Jones, R., & Hanlon, P. (2010a). It's not 'just deprivation': Why do equally deprived UK cities experience different health outcomes? *Public Health*, *124*(9), 487–495. https://doi.org/10.1016/j.puhe.2010.02.006

Walsh, D., McCartney, G., Collins, C., Taulbut, M., & Batty, G. D. (2017). History, politics and vulnerability: Explaining excess mortality in scotland and glasgow. *Public Health*, *151*, 1–12. https://doi.org/10.1016/j.puhe.2017.05.016

Walsh, D., Dundas, R., McCartney, G., Gibson, M., & Seaman, R. (2022). Bearing the burden of austerity: How do changing mortality rates in the UK compare between men and women? *J Epidemiol Community Health*, *76*(12), 1027–1033. https://doi.org/10.1136/jech-2022-219645

Walsh, D., Lowther, M., McCartney, G., & Reid, K. (2020a). Can scotland achieve its aim of narrowing health inequalities in a post-pandemic world? *Public Health in Practice*, *1*, 100042. https://doi.org/10.1016/j.puhip.2020.100042

Walsh, D., McCartney, G., Minton, J., Parkinson, J., Shipton, D., & Whyte, B. (2020b). Changing mortality trends in countries and cities of the UK: A population-based trend analysis. *BMJ Open*, *10*(11), e038135. https://doi.org/10.1136/bmjopen-2020-038135

Walsh, D., Taulbut, M., & Hanlon, P. (2010b). 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. https://doi.org/10.1093/eurpub/ckp063

Warnecke, R. B., Oh, A., Breen, N., Gehlert, S., Paskett, E., Tucker, K. L., Lurie, N., Rebbeck, T., Goodwin, J., Flack, J., Srinivasan, S., Kerner, J., Heurtin-Roberts, S., Abeles, R., Tyson, F. L., Patmios, G., & Hiatt, R. A. (2008). Approaching health disparities from a population perspective: The national institutes of health centers for population health and health disparities. *American Journal of Public Health*, *98*(9), 1608–1615. https://doi.org/10.2105/AJPH.2006.102525

Whyte, B., & Ajetunmobi, T. (2012). Still "the sick man of europe"? scottish mortality in a european context 1950-2010. https://www.gcph.co.uk/publications/391_still_the_sick_man_of_europe

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

Wilmoth, J., Andreev, K., Jdanov, D., Glei, D., Boe, C., Bubenheim, M., Philipov, D., Shkolnikov, V., & Vachon, P. (2022). Methods protocol for the human mortality database.

Wilmoth, J. R., & Horiuchi, S. (1999). Rectangularization revisited: Variability of age at death within human populations. *Demography*, *36*(4), 475–495. https://doi.org/10.2307/2648085

Wilson, L., Bhatnagar, P., & Townsend, N. (2017). Comparing trends in mortality from cardiovascular disease and cancer in the united kingdom, 1983–2013: Joinpoint regression analysis. *Population Health Metrics*, *15*(1), 23. https://doi.org/10.1186/s12963-017-0141-5

Wingard, D. L., Cohn, B. A., Kaplan, G. A., Cirillo, P. M., & Cohen, R. D. (1989). Sex differentials in morbidity and mortality risks examined by age and cause in the same cohort. *American Journal of Epidemiology*, *130*(3), 601–610. https://doi.org/10.1093/oxfordjournals.aje.a115374

Wishart, G. C., Azzato, E. M., Greenberg, D. C., Rashbass, J., Kearins, O., Lawrence, G., Caldas, C., & Pharoah, P. D. (2010). PREDICT: A new UK prognostic model that predicts survival following surgery for invasive breast cancer. *Breast Cancer Research*, *12*(1). https://doi.org/10.1186/bcr2464

World Health Organization. (2019). *International statistical classification of diseases and related health problems (11th ed.)*. https://icd.who.int/

Wraw, C., Minton, J., Mitchell, R., Wyper, G. M., Campbell, C., & McCartney, G. (2020). Can changes in spending on health and social care explain the recent mortality trends in scotland? a protocol for an observational study. *BMJ Open*. https://doi.org/10.1136/bmjopen-2019-036025

Wu, J., Mafham, M., Mamas, M. A., Rashid, M., Kontopantelis, E., Deanfield, J. E., de Belder, M. A., & Gale, C. P. (2021). Place and underlying cause of death during the COVID-19 pandemic: Retrospective cohort study of 3.5 million deaths in england and wales, 2014 to 2020. *Mayo Clinic Proceedings*, *96*(4), 952–963. https://doi.org/10.1016/j.mayocp.2021.02.007

Wu, W., Hu, J., & Zhang, J. Prognostics of machine health condition using an improved ARIMA-based prediction method [ISSN: 2158-2297]. In: In *2007 2nd IEEE conference on industrial electronics and applications*. 2007 2nd IEEE Conference on Industrial Electronics and Applications. ISSN: 2158-2297. 2007, 1062–1067. https://doi.org/10.1109/ICIEA.2007.4318571.

Wyper, G. M., Grant, I., Fletcher, E., McCartney, G., & Stockton, D. L. (2019). The impact of worldwide, national and subnational severity distributions in burden of disease studies: A case study of cancers in scotland. *PLoS ONE*, *14*(8), e0221026. https://doi.org/10.1371/journal.pone.0221026

Yen, I. H., & Kaplan, G. A. (1999). Neighborhood social environment and risk of death: Multilevel evidence from the alameda county study. *American Journal of Epidemiology*, *149*(10), 898–907. https://doi.org/10.1093/oxfordjournals.aje.a009733

Youd, E., & Moore, L. (2020). COVID-19 autopsy in people who died in community settings: The first series [Publisher: BMJ Publishing Group Section: Short report]. *Journal of Clinical Pathology*, *73*(12), 840–844. https://doi.org/10.1136/jclinpath-2020-206710

Young, T. K. (2004). *Population health: Concepts and methods* [Google-Books-ID: Sd3f_zd2npoC]. Oxford University Press.

Yurkov, A. M., Kemler, M., & Begerow, D. (2011). Species accumulation curves and incidence-based species richness estimators to appraise the diversity of cultivable yeasts from beech forest soils. *PLOS ONE*, *6*(8), e23671. https://doi.org/10.1371/journal.pone.0023671

Zeileis, A., & Grothendieck, G. (2005). Zoo: S3 infrastructure for regular and irregular time series. *Journal of Statistical Software*, *14*(6), 1–27. https://doi.org/10.18637/jss.v014.i06

Zellweger, U., Junker, C., Bopp, M., Egger, M., Spoerri, A., Zwahlen, M., Puhan, M., Röösli, M., Oris, M., & Bochud, M. (2019). Cause of death coding in switzerland: Evaluation based on a nationwide individual linkage of mortality and hospital in-patient records. *Population Health Metrics*, *17*(2). https://doi.org/10.1186/s12963-019-0182-z

Świgost, A. (2017). Approaches towards social deprivation: Reviewing measurement methods. *Bulletin of Geography. Socio-economic Series*, (38), 131–141. https://doi.org/10.1515/bog-2017-0039

Appendix A

Supplementary tables and figures

A.1 Nations examined in international studies of diversity in causes of mortality.

Paper	Nations studied
An ecological approach to causes of death in 14 countries: the Shannon index of diversity, *Ascaso Terrén, Canela Soler, Sentís Vílalta*, 1988 (Only available in Spanish)	Egypt, Chile, Cuba, USA, Israel, Japan, Thailand, France, Italy, Portugal, Sweden, Yugoslavia, Austria, Spain
Secular Changes of the Concentration of Neoplasm Death Causes in the Male Population of Some Countries, *Izsák*, 1988	England and Wales, Norway, Finland, Hungary, Japan
Comparative Analysis of Death Cause Diversity Curves in Various Countries, *Izsák*, 1993	England and Wales, Norway, Finland, Hungary, Japan
Diversification in causes of death in low-mortality countries: emerging patterns and implications, *Bergeron-Boucher, Aburto, van Raalte*, 2020	Australia, Austria, Belgium, Canada, Denmark, Finland, France, germany, Japan, Netherlands, Spain, Sweden, Switzerland, United Kingdom, USA

A.2 The number of causes recorded in each ICD-10 Chapter in men and women across Scotland and the percentage of all-cause ASMR associated with causes within that Chapter in 2001 and 2019.

Chapter	Number of individual causes		Percntage of total ASMR	
	2001	2019	2001	2019
Females				
I: Infections	0.96%	1.20%	29	25
II: Neoplasms	25.60%	28.03%	97	95
III: Blood	0.24%	0.31%	17	17
IV: Metabolic	1.53%	2.20%	22	23
IX: Circulatory	40.30%	24.02%	57	56
V: Mental	4.84%	8.75%	15	17
VI: Nervous	2.27%	7.86%	29	35
VIII: Ear	0.00%	0.00%	1	1
X: Respiratory	11.85%	11.76%	29	32
XI: Digestive	5.19%	5.78%	45	44
XII: Skin	0.21%	0.44%	10	12
XIII: Musculoskeletal	0.91%	0.88%	31	32
XIV: Genitourinary	1.87%	1.86%	27	24
XV: Pregnancy	0.01%	0.01%	1	2
XVI: Perinatal	0.24%	0.13%	22	12
XVII: Congenital	0.25%	0.22%	24	23
XVIII: NEC	0.75%	1.76%	14	17
XX: External Causes	2.96%	4.79%	82	79
VII: Eye	0.00%	0.00%	0	1
Males				
I: Infections	0.98%	1.17%	29	32
II: Neoplasms	27.49%	29.14%	91	88
III: Blood	0.21%	0.24%	16	17
IV: Metabolic	1.47%	2.37%	21	19
IX: Circulatory	39.97%	26.90%	50	55
V: Mental	3.57%	6.66%	12	14
VI: Nervous	1.97%	6.52%	30	33
VIII: Ear	0.00%	0.00%	1	1

X: Respiratory	12.36%	11.24%	38	39
XI: Digestive	4.92%	4.89%	42	48
XII: Skin	0.10%	0.24%	8	10
XIII: Musculoskeletal	0.34%	0.50%	25	29
XIV: Genitourinary	1.84%	1.58%	21	27
XVI: Perinatal	0.19%	0.16%	20	18
XVII: Congenital	0.20%	0.23%	26	28
XVIII: NEC	0.41%	0.86%	11	11
XX: External Causes	3.98%	7.28%	127	118
Individual causes associated with >0.25% of total ASMR in 2001 or 2019				
Females				
I: Infections	0.58%	0.59%	1	1
II: Neoplasms	22.52%	24.30%	19	21
IV: Metabolic	1.00%	1.31%	2	1
IX: Circulatory	38.35%	22.30%	18	20
V: Mental	4.65%	8.51%	3	3
VI: Nervous	1.47%	6.88%	3	4
X: Respiratory	10.88%	10.97%	5	7
XI: Digestive	3.50%	3.67%	8	8
XIII: Musculoskeletal	0.28%	0.00%	1	0
XIV: Genitourinary	1.41%	1.42%	3	2
XVIII: NEC	0.51%	1.62%	1	3
XX: External Causes	1.16%	3.30%	1	3
Males				
I: Infections	0.56%	0.58%	1	1
II: Neoplasms	24.72%	25.94%	19	22
IV: Metabolic	1.01%	1.77%	2	2
IX: Circulatory	38.29%	24.94%	17	17
V: Mental	3.35%	6.46%	4	3
VI: Nervous	1.10%	5.38%	2	4
X: Respiratory	11.64%	10.09%	6	6
XI: Digestive	3.15%	2.88%	6	6
XIV: Genitourinary	1.38%	1.14%	3	2
XX: External Causes	1.34%	5.06%	2	4
XVIII: NEC	0.00%	0.54%	0	2

A.3 The 5 causes in Chapter XX with the highest additive value to diversity at q = 1 in men in 2001 and 2019 in the most and least income deprived quintiles in Scotland.

ICD-10 Code	ICD-10 Descpription	Additive Value
	2001	
	Most deprived	
X70	Intentional self-harm by hanging, strangulation and suffocation	0.016
W19	Unspecified fall	0.011
X99	Assault by sharp object	0.009
Y12	Poisoning by and exposure to narcotics and psychodysleptics [hallucinogens], not elsewhere classified, undetermined intent	0.008
X00	Exposure to uncontrolled fire in building or structure	0.007
	Least deprived	
W19	Unspecified fall	0.015
X70	Intentional self-harm by hanging, strangulation and suffocation	0.012
W10	Fall on and from stairs and steps	0.006
V47	Car occupant injured in collision with fixed or stationary object	0.005
V43	Car occupant injured in collision with car, pick-up truck or van	0.005
	2019	
	Most deprived	
W19	Unspecified fall	0.064
X70	Intentional self-harm by hanging, strangulation and suffocation	0.061
X41	Accidental poisoning by and exposure to antiepileptic, sedative-hypnotic, antiparkinsonism and psychotropic drugs, not elsewhere classified	0.054
X59	Exposure to unspecified factor	0.006
X99	Assault by sharp object	0.005
	Least deprived	
W19	Unspecified fall	0.014
X70	Intentional self-harm by hanging, strangulation and suffocation	0.012
X42	Accidental poisoning by and exposure to narcotics and psychodysleptics [hallucinogens], not elsewhere classified	0.010

Y84	Other medical procedures as the cause of abnormal reaction of the patient, or of later complication, without mention of misadventure at the time of the procedure	0.005
Y83	Surgical operation and other surgical procedures as the cause of abnormal reaction of the patient, or of later complication, without mention of misadventure at the time of the procedure	0.005

A.4 Trends in diversity in mortality causes in deaths among those aged 0 to 19 and 20 to 39 across Scotland with mortality records pooled across 4 years.

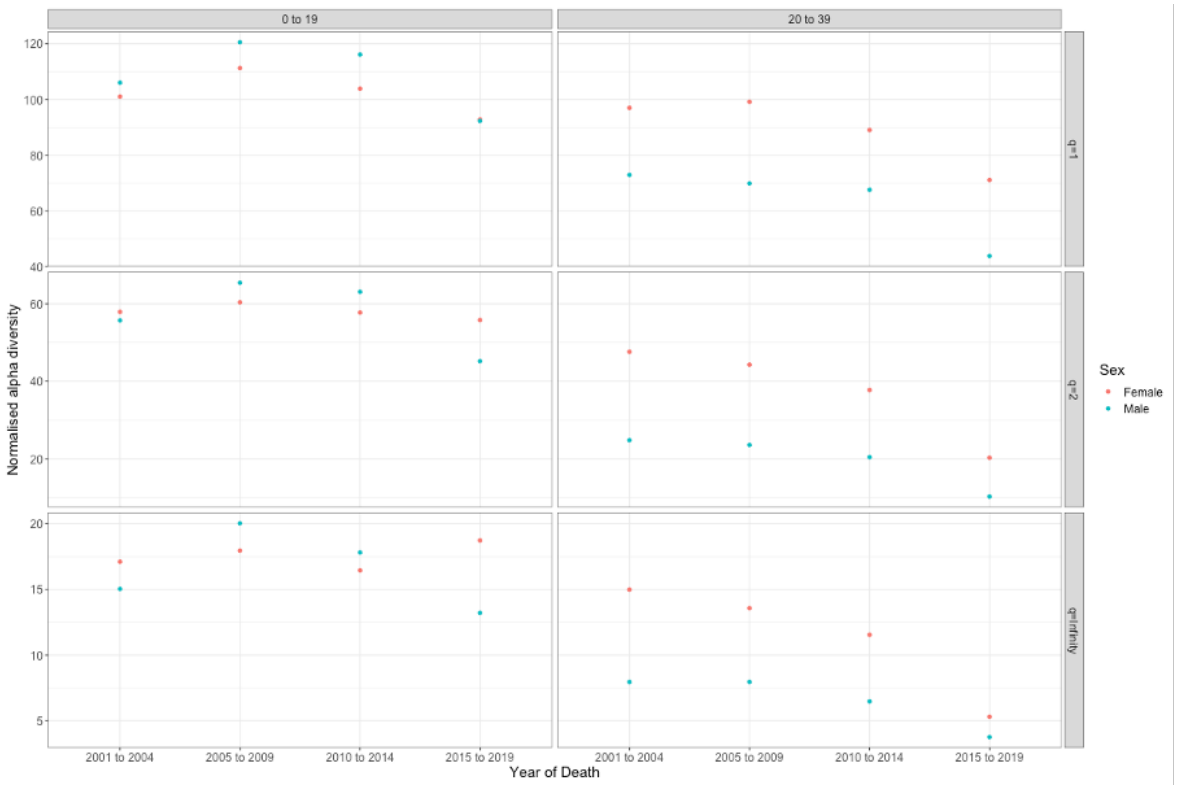
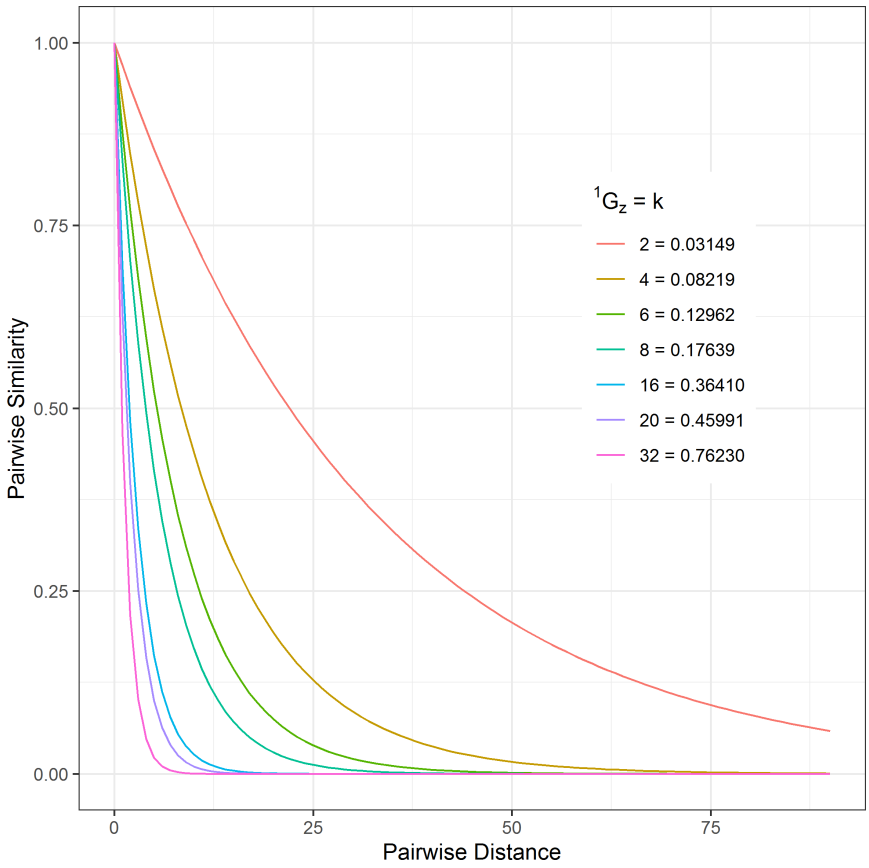
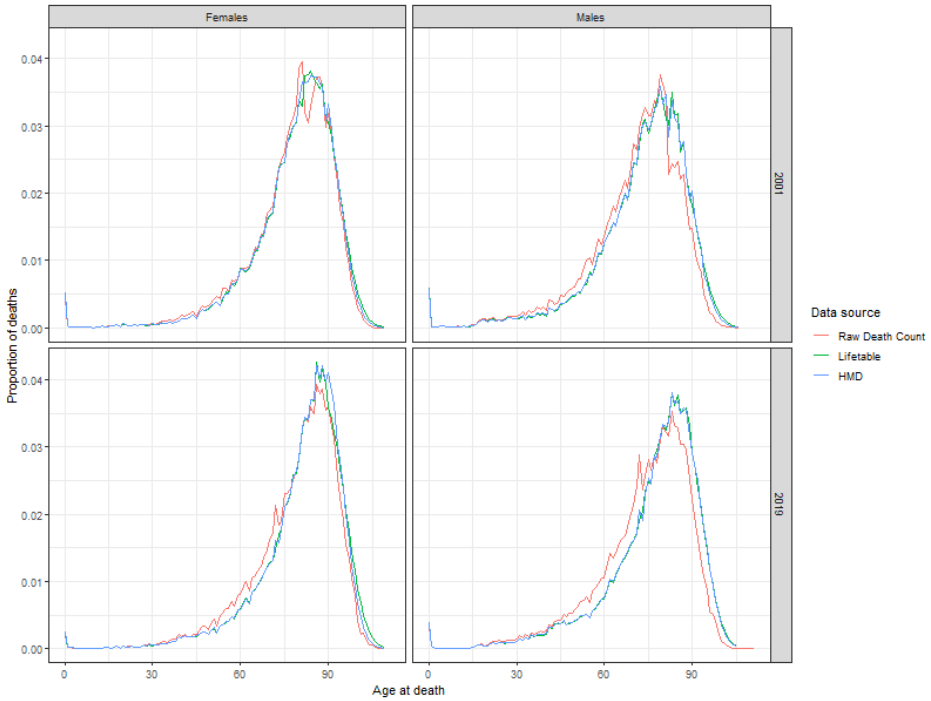


Figure A.1: Trends in diversity in mortality causes in deaths among those aged 0 to 19 and 20 to 39 across Scotland with mortality records pooled across 4 years.

A.5 Pairwise distance in age at death against pairwise similarity at various values of k.



A.6 Distributions of age at death in the Scottish population from life tables calculated from mortality data, publicly available life tables obtained from the Human Mortality Database and observed raw death counts in males and females in 2001 and 2019.



A.7 Correlation between similarity-sensitive normalised alpha diversity at q = 1 in age at death and life expectancy over the years 2001 to 2019 in Scotland and Scottish Subpopulations in males and females separately.

Diversity in causes of mortality q-value	Correlation coefficient	P-value
Scotland		
Females	-0.85	4.07E-06
Males	-0.94	3.66E-09
Scottish Government Urban-Rural Classes		
	Females	
Large Urban Areas	-0.73	4.42E-04
Other Urban Areas	-0.71	6.70E-04
Accessible Small Towns	-0.81	2.95E-05
Remote Small Towns	-0.57	0.01
Accessible Rural Areas	-0.72	5.07E-04
Remote Rural Areas	-0.54	0.02
	Males	
Large Urban Areas	-0.87	1.31E-06
Other Urban Areas	-0.86	3.04E-06
Accessible Small Towns	-0.92	3.30E-08
Remote Small Towns	-0.66	0
Accessible Rural Areas	-0.85	3.59E-06
Remote Rural Areas	-0.81	2.82E-05
SIMD Income Deprivation Quintiles		
	Females	
Most Deprived	-0.65	0
2	-0.7	9.36E-04
3	-0.82	1.67E-05
4	-0.72	4.75E-04
Least Deprived	-0.89	3.32E-07
	Males	
Most Deprived	-0.86	3.08E-06
2	-0.79	5.49E-05

3	-0.91	4.74E-08
4	-0.92	1.78E-08
Least Deprived	-0.96	1.64E-10

A.8 Correlation between normalised alpha diversity in causes of mortality and life expectancy over the years 2001 to 2019 in Scotland and Scottish Subpopulations in males and females separately.

Diversity in causes of mortality q-value	Correlation coefficient	P-value
Scotland		
Females	-0.85	4.07E-06
Males	-0.94	3.66E-09
Scottish Government Urban-Rural Classes		
Females		
Large Urban Areas	-0.73	4.42E-04
Other Urban Areas	-0.71	6.70E-04
Accessible Small Towns	-0.81	2.95E-05
Remote Small Towns	-0.57	0.01
Accessible Rural Areas	-0.72	5.07E-04
Remote Rural Areas	-0.54	0.02
Males		
Large Urban Areas	-0.87	1.31E-06
Other Urban Areas	-0.86	3.04E-06
Accessible Small Towns	-0.92	3.30E-08
Remote Small Towns	-0.66	0
Accessible Rural Areas	-0.85	3.59E-06
Remote Rural Areas	-0.81	2.82E-05
SIMD Income Deprivation Quintiles		
Females		
Most Deprived	-0.65	0
2	-0.7	9.36E-04
3	-0.82	1.67E-05

4	-0.72	4.75E-04
Least Deprived	-0.89	3.32E-07
Males		
Most Deprived	-0.86	3.08E-06
2	-0.79	5.49E-05
3	-0.91	4.74E-08
4	-0.92	1.78E-08
Least Deprived	-0.96	1.64E-10

A.9 Correlation between similarity-sensitive normalised alpha diversity in age at death at q = 1 and normalised alpha diversity in causes of mortality over the years 2001 to 2019 in Scotland and Scottish Subpopulations in males and females separately.

Diversity in causes of mortality q-value	Correlation coefficient	P-value
All ages		
Scotland		
Females		
1	-0.76	1.85E-04
2	-0.8	4.51E-05
Inf	-0.86	2.45E-06
Males		
1	-0.87	1.49E-06
2	-0.89	3.96E-07
Inf	-0.94	2.58E-09
Scottish Government Urban-Rural Classes		
Females		
Large Urban Areas		
1	-0.62	4.61E-03
2	-0.71	7.40E-04

Inf	-0.66	1.93E-03
Males		
1	-0.77	1.07E-04
2	-0.77	1.08E-04
Inf	-0.79	5.44E-05
Other Urban Areas		
Females		
1	-0.51	0.02701521
2	-0.56	0.01295917
Inf	-0.63	0.00377924
Males		
1	-0.74	2.98E-04
2	-0.76	1.68E-04
Inf	-0.83	1.06E-05
Accessible Small Towns		
Females		
1	-0.68	0.00139861
2	-0.69	1.00E-03
Inf	-0.71	7.07E-04
Males		
1	-0.82	1.64E-05
2	-0.88	5.22E-07
Inf	-0.89	2.55E-07
Remote Small Towns		
Females		
1	-0.17	4.76E-01
2	-0.29	2.35E-01
Inf	-0.15	5.33E-01
Males		
1	-0.43	6.82E-02
2	-0.56	1.35E-02
Inf	-0.61	5.30E-03
Accessible Rural Areas		
Females		
1	-0.74	3.13E-04
2	-0.78	7.02E-05
Inf	-0.85	3.36E-06
Males		

1	-0.69	1.03E-03
2	-0.73	3.74E-04
Inf	-0.77	1.17E-04
Remote Rural Areas		
Females		
1	-0.31	1.94E-01
2	-0.36	1.35E-01
Inf	-0.43	6.38E-02
Males		
1	-0.56	1.22E-02
2	-0.64	3.20E-03
Inf	-0.69	1.15E-03
SIMD Income Deprivation Quintiles		
Most Deprived		
Females		
1	-0.25	2.96E-01
2	-0.27	2.65E-01
Inf	-0.12	6.21E-01
Males		
1	-0.6	6.91E-03
2	-0.59	7.56E-03
Inf	-0.57	1.15E-02
Quintile 2		
Females		
1	-0.42	7.47E-02
2	-0.52	2.30E-02
Inf	-0.6	6.33E-03
Males		
1	-0.5	2.88E-02
2	-0.56	1.26E-02
Inf	-0.7	8.15E-04
Quintile 3		
Females		
1	-0.72	5.33E-04
2	-0.75	2.28E-04
Inf	-0.75	1.99E-04
Males		
1	-0.84	7.55E-06

2	-0.88	8.48E-07
Inf	-0.92	3.30E-08
Quintile 4		
Females		
1	-0.58	8.53E-03
2	-0.61	5.66E-03
Inf	-0.58	9.85E-03
Males		
1	-0.82	2.09E-05
2	-0.87	1.14E-06
Inf	-0.88	6.57E-07
Least deprived		
Females		
1	-0.76	1.37E-04
2	-0.78	9.07E-05
Inf	-0.74	3.34E-04
Males		
1	-0.94	2.37E-09
2	-0.94	1.90E-09
Inf	-0.91	6.40E-08

A.10 Cross-sectional correlation between normalised alpha diversity in causes of mortality and life expectancy in the years 2001 to 2019 between Scottish Subpopulations in males and females separately.

Year	Diversity in Mortality causes q-value	Correlation coefficient	P-value
Scottish Government Urban-Rural Classifications			
Females			
2001	1	-0.6379081	0.17292878
2002	1	7.91E-04	0.99881394
2003	1	-0.7525032	0.08430183

2004	1	-0.2338277	0.65565071
2005	1	-0.6426064	0.16877033
2006	1	-0.3674921	0.47357678
2007	1	-0.3044716	0.55740532
2008	1	-0.6346114	0.17587191
2009	1	-0.4599179	0.35876504
2010	1	-0.4450909	0.37645119
2011	1	-0.2655676	0.61101332
2012	1	-0.5515342	0.25658427
2013	1	-0.4044069	0.42645898
2014	1	-0.6560113	0.15714057
2015	1	-0.2868489	0.58152796
2016	1	-0.383159	0.45338746
2017	1	-0.4012106	0.43047557
2018	1	-0.1471217	0.78090966
2019	1	-0.0421215	0.93685506
2001	2	-0.4842091	0.3304498
2002	2	0.48355392	0.33120247
2003	2	-0.3163184	0.54134737
2004	2	0.83839105	0.03706577
2005	2	0.73473159	0.09621788
2006	2	0.60216178	0.2059289
2007	2	0.10569699	0.84204493
2008	2	0.06381518	0.90440717
2009	2	0.44340844	0.37847684
2010	2	0.70008974	0.12143136
2011	2	0.53715323	0.27176353
2012	2	0.68468718	0.13345872
2013	2	0.61266789	0.19598427
2014	2	0.31435108	0.54400493
2015	2	0.52793328	0.28167116
2016	2	0.57699485	0.23055517
2017	2	0.92457259	0.00831938
2018	2	0.86569178	0.02584667
2019	2	0.80822297	0.05164102
2001	Inf	-0.2853636	0.58357356
2002	Inf	0.04005104	0.93995556
2003	Inf	-0.2588426	0.62040732
2004	Inf	0.13374135	0.80058408
2005	Inf	0.39835151	0.43407872

2006	Inf	0.78412914	0.06487053
2007	Inf	0.03344271	0.94985464
2008	Inf	-0.0273605	0.95896955
2009	Inf	-0.1798089	0.73319334
2010	Inf	0.09694227	0.85504212
2011	Inf	0.36160414	0.48123502
2012	Inf	0.34805448	0.49900027
2013	Inf	0.69074264	0.12867147
2014	Inf	0.46673396	0.35073586
2015	Inf	-0.1850288	0.72562414
2016	Inf	0.71409031	0.11093078
2017	Inf	0.74761629	0.0875082
2018	Inf	0.16674624	0.75219878
2019	Inf	0.7795116	0.06756316
		Males	
2001	1	-0.7889429	0.06211685
2002	1	-0.9383847	0.0055777
2003	1	-0.7196746	0.10685919
2004	1	-0.7778493	0.06854471
2005	1	-0.2361719	0.65232871
2006	1	-0.6732675	0.14269117
2007	1	-0.2774832	0.59445794
2008	1	-0.6905476	0.12882451
2009	1	-0.6103844	0.19812863
2010	1	-0.7262539	0.10214853
2011	1	-0.8086534	0.05141737
2012	1	-0.889736	0.01756691
2013	1	-0.7786435	0.06807497
2014	1	-0.670062	0.14533027
2015	1	0.03693332	0.94462522
2016	1	-0.1354036	0.79813592
2017	1	-0.201637	0.70164347
2018	1	-0.4248103	0.40111597
2019	1	-0.5805297	0.22702897
2001	2	-0.6425708	0.16880162
2002	2	-0.3092721	0.55088261
2003	2	-0.3677787	0.47320507
2004	2	-0.3404245	0.50908897
2005	2	0.44705344	0.37409317
2006	2	-0.5923507	0.21539578

2007	2	0.75141093	0.08501382
2008	2	-0.304669	0.55713662
2009	2	-0.646201	0.16561743
2010	2	-0.6913087	0.12822787
2011	2	-0.4186808	0.4086748
2012	2	-0.6270837	0.18266976
2013	2	-0.0383598	0.94248855
2014	2	-0.6803765	0.13691252
2015	2	0.26721619	0.60871593
2016	2	0.3253794	0.52915515
2017	2	-0.1134832	0.83050588
2018	2	-0.5945519	0.21325685
2019	2	-0.1858808	0.72438999
2001	Inf	-0.6893602	0.12975768
2002	Inf	-0.4612269	0.35721813
2003	Inf	-0.5941601	0.21363688
2004	Inf	-0.349501	0.49709448
2005	Inf	0.11886837	0.82253722
2006	Inf	-0.9911912	1.16E-04
2007	Inf	0.19860441	0.70601024
2008	Inf	-0.8096163	0.05091859
2009	Inf	-0.4051731	0.42549804
2010	Inf	-0.4195334	0.40762061
2011	Inf	-0.1826266	0.72910562
2012	Inf	-0.727612	0.10118791
2013	Inf	0.03581427	0.94630156
2014	Inf	-0.637565	0.17323409
2015	Inf	0.39386169	0.43975677
2016	Inf	-0.4250981	0.40076226
2017	Inf	-0.3406168	0.50883392
2018	Inf	-0.618353	0.19068739
2019	Inf	-0.0525529	0.92124324
SIMD Quintiles			
Females			
2001	1	-0.9056057	0.0343167
2002	1	-0.8046148	0.10058074
2003	1	-0.6581946	0.22718271
2004	1	0.16487624	0.79102809
2005	1	-0.7704568	0.12737408
2006	1	-0.8275552	0.08370348

2007	1	-0.7066556	0.18209649
2008	1	-0.6181635	0.26641587
2009	1	-0.4659002	0.42901728
2010	1	0.07986049	0.89842665
2011	1	-0.4692276	0.42527237
2012	1	0.054278	0.93092505
2013	1	0.80545313	0.09994756
2014	1	-0.8577308	0.06302429
2015	1	-0.9056108	0.03431397
2016	1	-0.6095859	0.27503709
2017	1	-0.7377856	0.15468655
2018	1	-0.8917764	0.04203763
2019	1	-0.7385159	0.15405928
2001	2	0.87712828	0.0507389
2002	2	0.88681393	0.04492717
2003	2	0.72461067	0.16612941
2004	2	0.9573238	0.01051508
2005	2	0.66226724	0.22328815
2006	2	0.97371537	0.00509527
2007	2	0.88191882	0.04783647
2008	2	0.90271802	0.03588736
2009	2	0.95937965	0.00976754
2010	2	0.94718412	0.01445472
2011	2	0.75708154	0.13836622
2012	2	0.91624245	0.02873009
2013	2	0.9772596	0.00410246
2014	2	0.98086908	0.00316729
2015	2	0.82888802	0.08275249
2016	2	0.82266717	0.08721962
2017	2	0.69727854	0.19059957
2018	2	0.39179146	0.51422904
2019	2	0.46469856	0.43037153
2001	Inf	0.73804607	0.15446276
2002	Inf	0.9495703	0.0134912
2003	Inf	0.87181085	0.054023
2004	Inf	0.72928006	0.16204685
2005	Inf	-0.0416389	0.94699904
2006	Inf	0.5795864	0.30574559
2007	Inf	0.94261371	0.01635957
2008	Inf	0.72718895	0.16387152

2009	Inf	0.80320806	0.10164603
2010	Inf	0.86920271	0.05565733
2011	Inf	0.76279671	0.13363639
2012	Inf	0.92714992	0.023344
2013	Inf	0.88953658	0.04333427
2014	Inf	0.72352923	0.16707912
2015	Inf	0.84957282	0.06843414
2016	Inf	0.87833439	0.05000308
2017	Inf	0.79135285	0.11075682
2018	Inf	0.31039235	0.61123708
2019	Inf	0.4713931	0.42283909
		Males	
2001	1	-0.8936163	0.04098192
2002	1	-0.899453	0.03769006
2003	1	-0.3051216	0.61762228
2004	1	-0.9596052	0.00968665
2005	1	-0.38526	0.52189186
2006	1	-0.9091932	0.03239708
2007	1	-0.7261608	0.16477084
2008	1	-0.4528308	0.44379795
2009	1	-0.6619983	0.22354472
2010	1	-0.9163787	0.02866064
2011	1	-0.8894452	0.04338745
2012	1	-0.2354567	0.70300083
2013	1	-0.5449497	0.34221113
2014	1	-0.9880203	0.00157115
2015	1	-0.0747477	0.90491692
2016	1	-0.9251998	0.02428031
2017	1	-0.2191517	0.72321726
2018	1	-0.8803563	0.04877723
2019	1	-0.9868	0.00181691
2001	2	0.76849359	0.12897053
2002	2	0.52315794	0.36566815
2003	2	0.56827792	0.31753635
2004	2	-0.0706682	0.9100974
2005	2	0.80029511	0.10386267
2006	2	0.72217271	0.1682726
2007	2	0.94045881	0.01728395
2008	2	0.91972689	0.02697053
2009	2	0.79443035	0.1083691

2010	2	0.91787621	0.02790062
2011	2	0.84304037	0.07286468
2012	2	0.96314722	0.00844547
2013	2	0.97586694	0.00448409
2014	2	0.89958328	0.03761758
2015	2	0.95813717	0.01021717
2016	2	0.95939903	0.00976058
2017	2	0.96702139	0.00715356
2018	2	0.95688658	0.01067638
2019	2	0.88433206	0.04639499
2001	Inf	-0.7160514	0.17368849
2002	Inf	-0.2075335	0.73766937
2003	Inf	-0.3701863	0.53966219
2004	Inf	-0.835828	0.07785524
2005	Inf	-0.3644308	0.54647804
2006	Inf	-0.9887587	0.00142832
2007	Inf	0.38980928	0.51655213
2008	Inf	-0.2939234	0.63122561
2009	Inf	0.44606596	0.45149223
2010	Inf	0.72126015	0.16907686
2011	Inf	0.52401894	0.36473417
2012	Inf	0.94362042	0.01593343
2013	Inf	0.86739091	0.05680164
2014	Inf	0.98381842	0.00246495
2015	Inf	0.82593549	0.08486368
2016	Inf	0.54448919	0.3427028
2017	Inf	0.71531448	0.17434381
2018	Inf	0.32745843	0.59064289
2019	Inf	0.10094124	0.87169622

A.11 Cross-sectional correlation between similarity-sensitive normalised alpha diversity at q = 1 and normalised alpha diversity in causes of mortality in the years 2001 to 2019 between Scottish Subpopulations in males and females separately.

Diversity in causes of mortality q-value	Correlation coefficient	P-value
Scottish Government Urban-Rural Classifications		
Females		
	2001	
1	0.81	0.05
2	0.67	0.15
Inf	0.22	0.67
	2002	
1	0.56	0.25
2	-0.49	0.33
Inf	-0.42	0.41
	2003	
1	0.57	0.24
2	-0.06	0.91
Inf	-0.19	0.72
	2004	
1	0.31	0.55
2	-0.72	0.11
Inf	-0.19	0.72
	2005	
1	0.93	0.01
2	-0.44	0.38
Inf	-0.19	0.72
	2006	
1	0.76	0.08
2	0.09	0.87
Inf	-0.46	0.36
	2007	
1	0.7	0.12
2	0.09	0.87
Inf	-0.16	0.77
	2008	
1	0.8	0.06
2	-0.09	0.87
Inf	-0.15	0.78
	2009	

1	0.4	0.43
2	-0.66	0.16
Inf	0.03	0.95
	2010	
1	0.89	0.02
2	-0.32	0.54
Inf	0.13	0.81
	2011	
1	0.62	0.19
2	-0.58	0.23
Inf	-0.69	0.13
	2012	
1	0.72	0.1
2	-0.57	0.24
Inf	-0.84	0.04
	2013	
1	0.89	0.02
2	-0.2	0.71
Inf	-0.75	0.09
	2014	
1	0.62	0.19
2	-0.27	0.6
Inf	-0.51	0.31
	2015	
1	0.85	0.03
2	0.17	0.75
Inf	-0.16	0.77
	2016	
1	0.18	0.74
2	-0.71	0.11
Inf	-0.96	0
	2017	
1	0.63	0.18
2	-0.89	0.02
Inf	-0.59	0.22
	2018	
1	0.39	0.44
2	-0.82	0.04

Inf	-0.56	0.24
	2019	
1	0.49	0.33
2	-0.48	0.33
Inf	-0.41	0.42
	Males	
	2001	
1	0.99	0
2	0.59	0.21
Inf	0.62	0.19
	2002	
1	0.89	0.02
2	0.18	0.74
Inf	0.38	0.46
	2003	
1	0.75	0.09
2	0.5	0.31
Inf	0.57	0.23
	2004	
1	0.84	0.04
2	0.51	0.3
Inf	0.49	0.32
	2005	
1	0.06	0.92
2	-0.37	0.48
Inf	-0.22	0.68
	2006	
1	0.69	0.13
2	0.78	0.07
Inf	0.95	0
	2007	
1	0.65	0.16
2	-0.39	0.45
Inf	0.28	0.59
	2008	
1	0.75	0.09
2	0.33	0.52
Inf	0.79	0.06

	2009	
1	0.59	0.22
2	0.58	0.23
Inf	0.27	0.61
	2010	
1	0.78	0.07
2	0.44	0.38
Inf	0.08	0.88
	2011	
1	0.84	0.04
2	-0.03	0.95
Inf	-0.1	0.84
	2012	
1	0.79	0.06
2	0.08	0.89
Inf	0.23	0.66
	2013	
1	0.91	0.01
2	-0.11	0.83
Inf	-0.27	0.61
	2014	
1	0.82	0.05
2	0.68	0.13
Inf	0.37	0.48
	2015	
1	0.32	0.54
2	-0.07	0.9
Inf	0.03	0.95
	2016	
1	0.3	0.56
2	-0.11	0.84
Inf	0.5	0.31
	2017	
1	0.26	0.62
2	0.18	0.74
Inf	0.4	0.44
	2018	
1	0.52	0.29

2	0.75	0.09
Inf	0.52	0.29
	2019	
1	0.7	0.12
2	0.14	0.8
Inf	0.01	0.99
	SIMD Income Deprivation Quintile	
	Females	
	2001	
1	0.84	0.08
2	-0.75	0.14
Inf	-0.56	0.32
	2002	
1	0.75	0.15
2	-0.85	0.07
Inf	-0.85	0.07
	2003	
1	0.55	0.33
2	-0.68	0.2
Inf	-0.78	0.12
	2004	
1	-0.31	0.61
2	-0.93	0.02
Inf	-0.72	0.17
	2005	
1	0.68	0.21
2	-0.67	0.22
Inf	0.06	0.92
	2006	
1	0.77	0.13
2	-0.98	0
Inf	-0.46	0.43
	2007	
1	0.44	0.46
2	-0.99	0
Inf	-0.87	0.05
	2008	
1	0.51	0.38

2	-0.94	0.02
Inf	-0.81	0.09
	2009	
1	0.34	0.58
2	-0.99	0
Inf	-0.88	0.05
	2010	
1	-0.02	0.98
2	-0.91	0.03
Inf	-0.89	0.04
	2011	
1	0.47	0.43
2	-0.75	0.14
Inf	-0.78	0.12
	2012	
1	-0.09	0.88
2	-0.94	0.02
Inf	-0.96	0.01
	2013	
1	-0.77	0.13
2	-0.96	0.01
Inf	-0.86	0.06
	2014	
1	0.89	0.05
2	-0.97	0.01
Inf	-0.71	0.18
	2015	
1	0.87	0.06
2	-0.88	0.05
Inf	-0.89	0.05
	2016	
1	0.56	0.33
2	-0.87	0.06
Inf	-0.92	0.02
	2017	
1	0.64	0.24
2	-0.78	0.12
Inf	-0.87	0.06

	2018	
1	0.86	0.06
2	-0.45	0.44
Inf	-0.39	0.52
	2019	
1	0.69	0.19
2	-0.53	0.36
Inf	-0.58	0.3
	Males	
	2001	
1	0.87	0.05
2	-0.7	0.19
Inf	0.79	0.11
	2002	
1	0.89	0.04
2	-0.41	0.49
Inf	0.34	0.58
	2003	
1	0.37	0.54
2	-0.49	0.4
Inf	0.45	0.45
	2004	
1	0.92	0.03
2	0.06	0.92
Inf	0.83	0.08
	2005	
1	0.33	0.59
2	-0.78	0.12
Inf	0.38	0.53
	2006	
1	0.87	0.05
2	-0.76	0.14
Inf	0.99	0
	2007	
1	0.7	0.19
2	-0.9	0.04
Inf	-0.32	0.6
	2008	

1	0.42	0.49
2	-0.91	0.03
Inf	0.33	0.59
	2009	
1	0.65	0.23
2	-0.74	0.15
Inf	-0.37	0.54
	2010	
1	0.9	0.04
2	-0.9	0.04
Inf	-0.72	0.17
	2011	
1	0.86	0.06
2	-0.86	0.06
Inf	-0.62	0.26
	2012	
1	0.25	0.68
2	-0.95	0.01
Inf	-0.92	0.02
	2013	
1	0.59	0.29
2	-0.95	0.01
Inf	-0.84	0.08
	2014	
1	0.98	0
2	-0.85	0.07
Inf	-0.99	0
	2015	
1	0.1	0.87
2	-0.93	0.02
Inf	-0.82	0.09
	2016	
1	0.93	0.02
2	-0.97	0.01
Inf	-0.6	0.29
	2017	
1	0.24	0.7
2	-0.97	0.01

Inf	-0.76	0.14
	2018	
1	0.88	0.05
2	-0.96	0.01
Inf	-0.35	0.57
	2019	
1	0.98	0
2	-0.86	0.06
Inf	-0.05	0.93

A.12 P-values of correlations between various indices of variation in age at death.

	e†	Life table entropy	Standard Deviation	Coefficient of Variance	Inter-Quartile Range	AID	Gini Coefficient	Lifetable diversity q =1	Lifetable diversity q =2	Raw death count diversity q =1	Raw death count diversity q =2
e†	0	1.09E-33	3.76E-08	1.31E-34	2.72E-34	2.33E-38	6.13E-38	4.95E-43	7.29E-47	6.35E-20	2.23E-17
Life table entropy	2.54E-35	0	4.42E-08	1.35E-44	2.51E-29	2.16E-27	1.12E-40	6.22E-29	8.21E-31	1.05E-17	2.58E-15
Standard Deviation	4.34E-09	1.64E-08	0	4.42E-08	4.42E-08	2.02E-08	4.42E-08	2.76E-08	3.76E-08	3.76E-08	4.42E-08
Coefficient of Variance	2.67E-36	2.18E-46	7.97E-09	0	6.01E-29	1.14E-29	9.98E-46	2.12E-30	5.28E-31	1.57E-18	5.24E-16
Inter-Quartile Range	5.92E-36	7.61E-31	7.50E-09	1.88E-30	0	4.57E-32	2.16E-31	2.72E-34	1.69E-38	8.26E-20	9.75E-18
AID	4.23E-40	7.19E-29	1.84E-09	3.34E-31	1.14E-33	0	1.06E-32	4.72E-50	2.39E-38	1.13E-23	1.68E-20
Gini Coefficient	1.18E-39	1.89E-42	1.17E-08	1.56E-47	5.53E-33	2.52E-34	0	3.98E-34	1.19E-34	8.26E-20	3.55E-17
Lifetable diversity q =1	8.25E-45	2.01E-30	2.76E-09	6.07E-32	5.91E-36	7.15E-52	9.05E-36	0	9.98E-46	1.41E-22	1.04E-19
Lifetable diversity q =2	1.12E-48	2.28E-32	4.18E-09	1.43E-32	3.01E-40	4.42E-40	2.37E-36	1.57E-47	0	3.00E-20	6.00E-18
Raw death count diversity q =1	2.65E-21	6.57E-19	4.87E-09	7.86E-20	3.59E-21	3.89E-25	3.61E-21	5.04E-24	1.20E-21	0	5.01E-38
Raw death count diversity q =2	1.48E-18	2.15E-16	7.37E-09	4.03E-17	5.74E-19	6.47E-22	2.53E-18	4.96E-21	3.33E-19	9.46E-40	0

A.13 Parameters of ARIMA models used to forecast monthly diversity in mortality causes. Models were run separately in each sex for each value of q, across all ages and causes, at twenty-year age ranges across all causes and across all causes in selected ICD-10 Chapters.

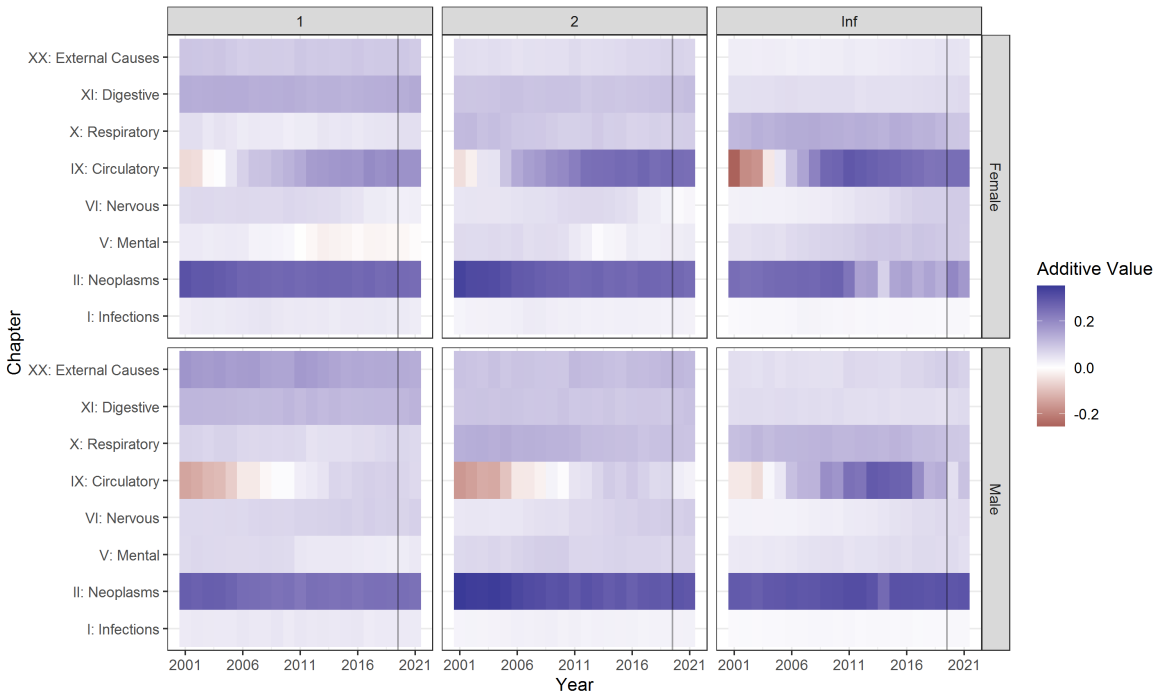
Diversity q-value	ARIMA parameters	Ljung-Box test p-value	Notes
All-ages, all-cause			
Females			
1	(1,0,1)(1,1,1)	0.669	
2	(1,0,1)(0,1,1)	0.6857	
Inf	(0,1,1)(0,1,1)	0.3438	
Males			
1	(0,1,1)(0,1,1)	0.55589	
2	(0,1,1)(0,1,1)	0.5999	
Inf	(0,1,1)(0,1,1)	0.7794	
Twenty-year age ranges, all-cause			
Females, 0 to 19			
1	(2,0,0)(2,1,1)	0.2002	
2	(3,0,3)(2,1,1)	0.2397	
Inf	(0,0,0)(1,1,1)	0.8169	
Males, 0 to 19			
1	(0,1,1)(0,0,0)	0.6132	
2	(0,1,1)(0,0,0)	0.6084	
Inf	(2,0,2)(1,1,2)	0.1353	
Females, 20 to 39			
1	(0,1,1)(0,1,1)	0.5695	
2	(0,1,1)(0,1,1)	0.4099	
Inf	(0,1,1)(0,0,0)	0.477	
Males, 20 to 39			
1	(0,1,1)(0,1,1)	0.2553	
2	(1,0,0)(1,1,2)	0.5115	
Inf	(2,0,0)(1,1,2)	0.6878	

Females, 40 to 59			
1	(0,1,1)(0,1,1)	0.4224	
2	(0,1,2)(0,0,0)	0.2136	
Inf	(0,1,1)(0,1,1)	0.2332	
Males, 40 to 59			
1	(1,1,1)(0,1,1)	0.4168	
2	(0,1,2)(0,0,0)	0.2799	
Inf	(1,1,1)(0,0,0)	0.268	
Females, 60 to 79			
1	(0,1,1)(0,1,1)	0.2028	
2	(0,0,0)(2,1,0)	0.1109	
Inf	(1,1,3)(0,0,0)	0.03033	Model suggests correlated residuals under both seasonal and non-seasonal conditions
Males, 60 to 79			
1	(0,1,1)(0,1,1)	0.7957	
2	(0,1,1)(0,1,1)	0.9807	
Inf	(0,1,1)(0,1,1)	0.7655	
Females, 80+			
1	(2,0,0)(1,1,2)	0.319	Model fit using data from 2010-2019
2	(2,0,0)(1,1,1)	0.707	Model fit using data from 2010-2020
Inf	(0,1,1)(0,1,2)	0.8996	Model fit using data from 2010-2021
Males, 80+			
1	(0,1,1)(0,1,1)	0.2425	
2	(0,1,1)(0,1,1)	0.7744	
Inf	(0,1,1)(0,1,1)	0.7655	
All-ages, ICD-10 Chapter			
Females, I: Certain infectious and parasitic diseases			
1	(0,1,1)(0,1,1)	0.2461	
2	(0,1,1)(0,0,0)	0.8536	
Inf	(0,1,1)(0,0,0)	0.9636	
Males, I: Certain infectious and parasitic diseases			
1	(0,1,1)(0,0,0)	0.1178	

2	(0,1,1)(0,0,0)	0.6106	
Inf	(1,1,3)(0,0,0)	0.9014	
Females, II: Neoplasms			
1	(0,1,1)(0,1,1)	0.07779	Model fit using data from 2010-2019
2	(1,1,1)(0,0,0)	0.8639	Model fit using data from 2010-2019
Inf	(1,1,1)(0,0,0)	0.6328	
Males, II: Neoplasms			
1	(0,1,1)(0,1,1)	0.9387	
2	(1,1,3)(0,0,0)	0.7285	
Inf	(1,1,3)(0,0,0)	0.7768	
Females, V: Mental and behavioural disorders			
1	(0,1,1)(0,1,1)	0.7441	
2	(0,1,1)(0,1,1)	0.8015	
Inf	(0,1,1)(0,1,1)	0.7303	
Males, V: Mental and behavioural disorders			
1	(0,1,1)(0,1,1)	0.7441	
2	(0,1,1)(0,1,1)	0.7957	
Inf	(0,1,1)(0,1,1)	0.5161	
Females, VI: Diseases of the nervous system			
1	(4,1,0)(0,0,0)	0.5126	Model fit using data from 2010-2019
2	(2,1,2)(0,0,0)	0.08591	Model fit using data from 2010-2019
Inf	(0,1,1)(0,0,0)	0.03096	Model suggests correlated residuals under both seasonal and non-seasonal conditions; Model fit using data from 2010-2019
Males, VI: Diseases of the nervous system			
1	(0,1,1)(0,0,0)	0.3635	Model fit using data from 2010-2019
2	(4,1,3)(0,0,0)	0.09975	Model fit using data from 2010-2019

Inf	(0,1,1)(0,0,0)	0.3052	Model fit using data from 2010-2019
Females, IX: Diseases of the circulatory system			
1	(0,0,1)(0,1,1)	0.2346	
2	(1,1,1)(0,0,0)	0.2951	
Inf	(0,1,1)(0,1,1)	0.6099	
Males, IX: Diseases of the circulatory system			
1	(0,1,1)(0,0,0)	0.5867	
2	(0,1,1)(0,0,0)	0.3868	
Inf	(0,1,1)(0,0,0)	0.4267	
Females, X: Diseases of the respiratory system			
1	(0,1,1)(0,1,1)	0.2346	
2	(0,1,1)(0,1,1)	0.05796	
Inf	(0,1,1)(0,0,0)	0.09614	
Males, X: Diseases of the respiratory system			
1	(2,1,1)(0,0,0)	0.09687	
2	(2,1,1)(0,0,0)	0.8247	
Inf	(0,1,1)(0,1,1)	0.4373	
Females, XX: External causes of morbidity and mortality			
1	(2,1,1)(0,0,0)	0.6301	
2	(2,1,1,)(0,0,0)	0.6301	
Inf	(0,1,1)(0,1,1)	0.8027	
Males, XX: External causes of morbidity and mortality			
1	(0,1,1)(0,1,1)	0.073	
2	(0,1,1)(0,1,1)	0.1163	
Inf	(0,1,1)(0,1,1)	0.5676	

A.14 The additive value of selected ICD-10 chapters under diversity at q = 1, q = 2, and q= ∞ in men and women in Scotland separately in each year 2001-2020.



A.15 Cause-specific age-standardised mortality rate in the years 2016 to 2021 in males and females within those aged 0 to 19, 20 to 39, 40 to 59, 60 to 79 and 80+ separately. Causes shown are those which were within the five most common in any year from 2016-2019.

ICD-10 Code	ICD-10 description	Age-standardised mortality rate (per 100,000)					
		2016	2017	2018	2019	2020	2021
	Deaths among those aged 0 to 19						
	Females						
P36	Bacterial sepsis of newborn	1.3	1.4	1	0.5	0.7	1.3
Q91	Edwards syndrome and Patau syndrome	1.2	0.3	0.7	0.4	0.5	0.4
R95	Sudden infant death syndrome	1.2	0.7	1.4	1.6	0.9	1.9

V43	Car occupant injured in collision with car, pick-up truck or van	0.9	0.2	0	0	0.4	0.4
X70	Intentional self-harm by hanging, strangulation and suffocation	0.9	0.9	1.7	2	0.9	0.9
P07	Disorders related to short gestation and low birth weight, not elsewhere classified	0.8	1.4	1.4	0.9	0.9	2.6
C71	Malignant neoplasm of brain	0.4	0.9	0.7	0.7	0.9	0.4
G80	Cerebral palsy	0.2	0.7	0.4	0.4	0	0
R99	Other ill-defined and unspecified causes of mortality	0.3	0.7	1.4	0.4	0.9	0.4
P77	Necrotizing enterocolitis of fetus and newborn	0.8	0.5	1.2	0.4	0.4	0.6
P02	Fetus and newborn affected by complications of placenta, cord and membranes	0.5	0.2	0.3	1.1	0.5	0.9
P52	Intracranial nontraumatic haemorrhage of fetus and newborn	0.3	0.5	0.3	0	0.9	0.6
P27	Chronic respiratory disease originating in the perinatal period	0.7	0.3	0	0.2	0.2	1.1
	Males						
R95	Sudden infant death syndrome	1.7	2.1	3.1	2.2	1.4	0.7
P07	Disorders related to short gestation and low birth weight, not elsewhere classified	1.6	1.8	1.1	2	1.7	0.9
X70	Intentional self-harm by hanging, strangulation and suffocation	1.5	1.2	4	3.7	3.7	2.3
G80	Cerebral palsy	1.2	0.7	0.5	0.9	0.9	0.8
G40	Epilepsy	1.2	0.4	0.3	0.7	1	0.3
P77	Necrotizing enterocolitis of fetus and newborn	0.6	1.9	0.5	0.8	0.5	0.7
W75	Accidental suffocation and strangulation in bed	0.2	0.8	0.2	0.7	0.2	0.4

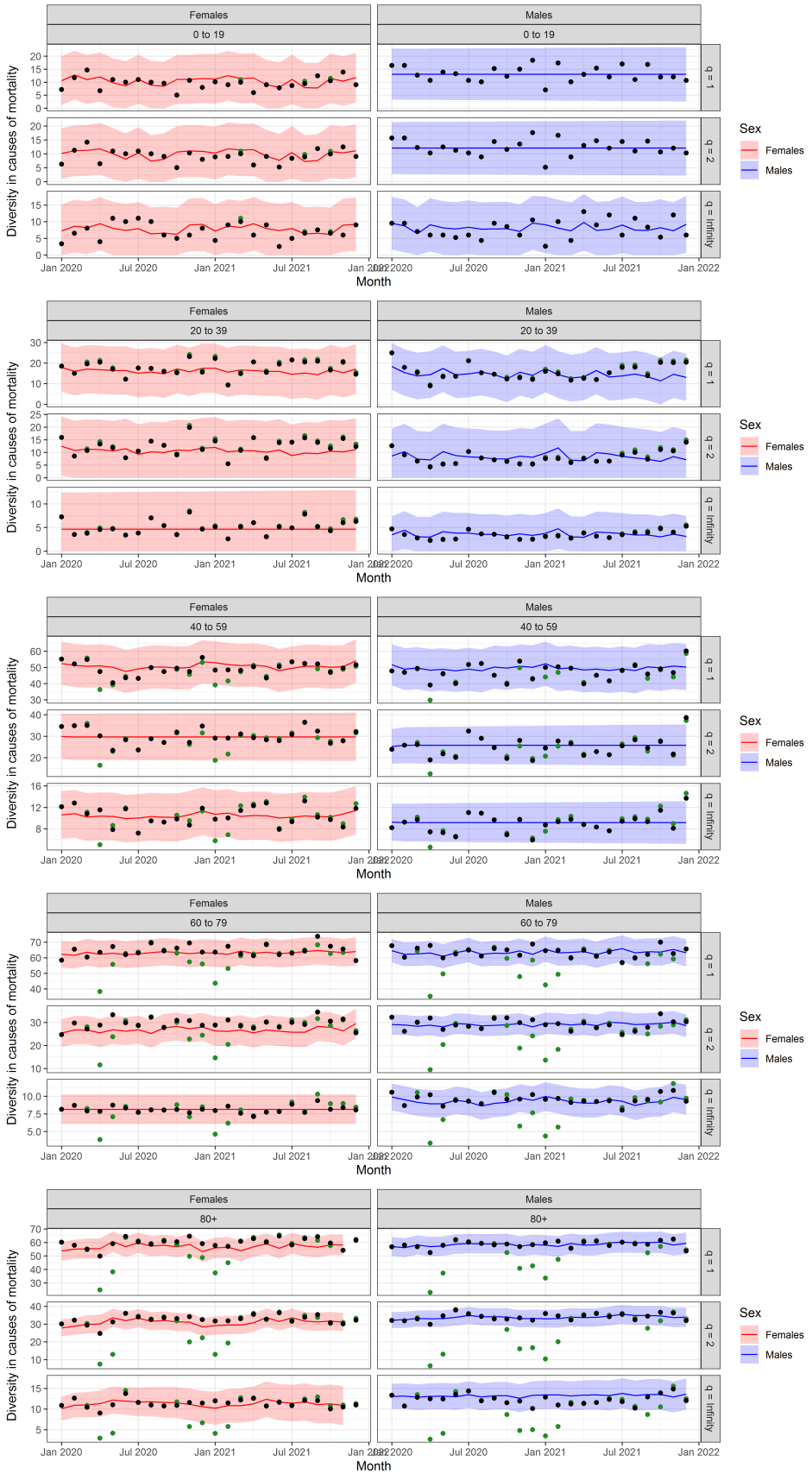
C71	Malignant neoplasm of brain	0.9	0.7	1.5	0.5	0.7	1
X42	Accidental poisoning by and exposure to narcotics and psychodysleptics [hallucinogens], not elsewhere classified	0.3	0.4	1.4	1.8	1.6	1.1
X41	Accidental poisoning by and exposure to antiepileptic, sedative-hypnotic, antiparkinsonism and psychotropic drugs, not elsewhere classified	0.7	0.7	1.2	0.7	1.2	0.5
P91	Other disturbances of cerebral status of newborn	0.6	0.3	0.3	1.2	0.9	0.5
P36	Bacterial sepsis of newborn	0.2	0.2	0.7	1	0.5	1.1
Q24	Other congenital malformations of heart	0.3	0.5	0.8	0.7	1	1
Deaths among those aged 20 to 39							
Females							
X42	Accidental poisoning by and exposure to narcotics and psychodysleptics [hallucinogens], not elsewhere classified	10	13.7		15.4	13.7	13
C50	Malignant neoplasm of breast	4.4	3.2		2.2	3	3.4
X70	Intentional self-harm by hanging, strangulation and suffocation	4.2	3.8		5.3	5.4	5.5
K70	Alcoholic liver disease	2.6	1.7		1.8	2.9	1.8
C53	Malignant neoplasm of cervix uteri	2.2	2.2		2	1.6	2.1
X41	Accidental poisoning by and exposure to antiepileptic, sedative-hypnotic, antiparkinsonism and psychotropic drugs, not elsewhere classified	2.2	1.8		1.7	2.5	2.2
R99	Other ill-defined and unspecified causes of mortality	1.1	1.3		2.6	1.1	0.7
U07	COVID-19	0	0		0	1.3	2.4
Males							

X42	Accidental poisoning by and exposure to narcotics and psychodysleptics [hallucinogens], not elsewhere classified	31.9	34.6	38.8	47.4	34.4
X70	Intentional self-harm by hanging, strangulation and suffocation	15.4	18.5	22.9	21.5	16.7
X41	Accidental poisoning by and exposure to antiepileptic, sedative-hypnotic, antiparkinsonism and psychotropic drugs, not elsewhere classified	4.9	4.7	4.9	7.1	8.5
K70	Alcoholic liver disease	3.5	4.1	3.1	2.4	3.1
R99	Other ill-defined and unspecified causes of mortality	3.1	2.8	1.9	1.6	2.9
I25	Chronic ischaemic heart disease	2.5	3.3	2.1	1.5	2.1
F10	Mental and behavioural disorders due to use of alcohol	1.5	1.4	2.3	1.7	1.2
G40	Epilepsy	1	1.8	1.4	2.6	1.7
X99	Assault by sharp object	0.9	2.3	1.6	2.5	1.3
U07	COVID-19	0	0	0	1.4	4.8
	Deaths among those aged 40 to 59					
	Females					
C50	Malignant neoplasm of breast	29.8	26.4	29.8	27.1	27.7
C34	Malignant neoplasm of bronchus and lung	24.2	23.2	20.7	20.5	20.7
K70	Alcoholic liver disease	15.7	17	16.9	14.9	17.9
X42	Accidental poisoning by and exposure to narcotics and psychodysleptics [hallucinogens], not elsewhere classified	12.3	16.4	21.5	21.5	21
J44	Other chronic obstructive pulmonary disease	9.9	10.2	10.5	8.8	11.7
U07	COVID-19	0	0	0	15.5	24.1
	Male					
I25	Chronic ischaemic heart disease	40.2	36.3	37.2	39.5	46.1

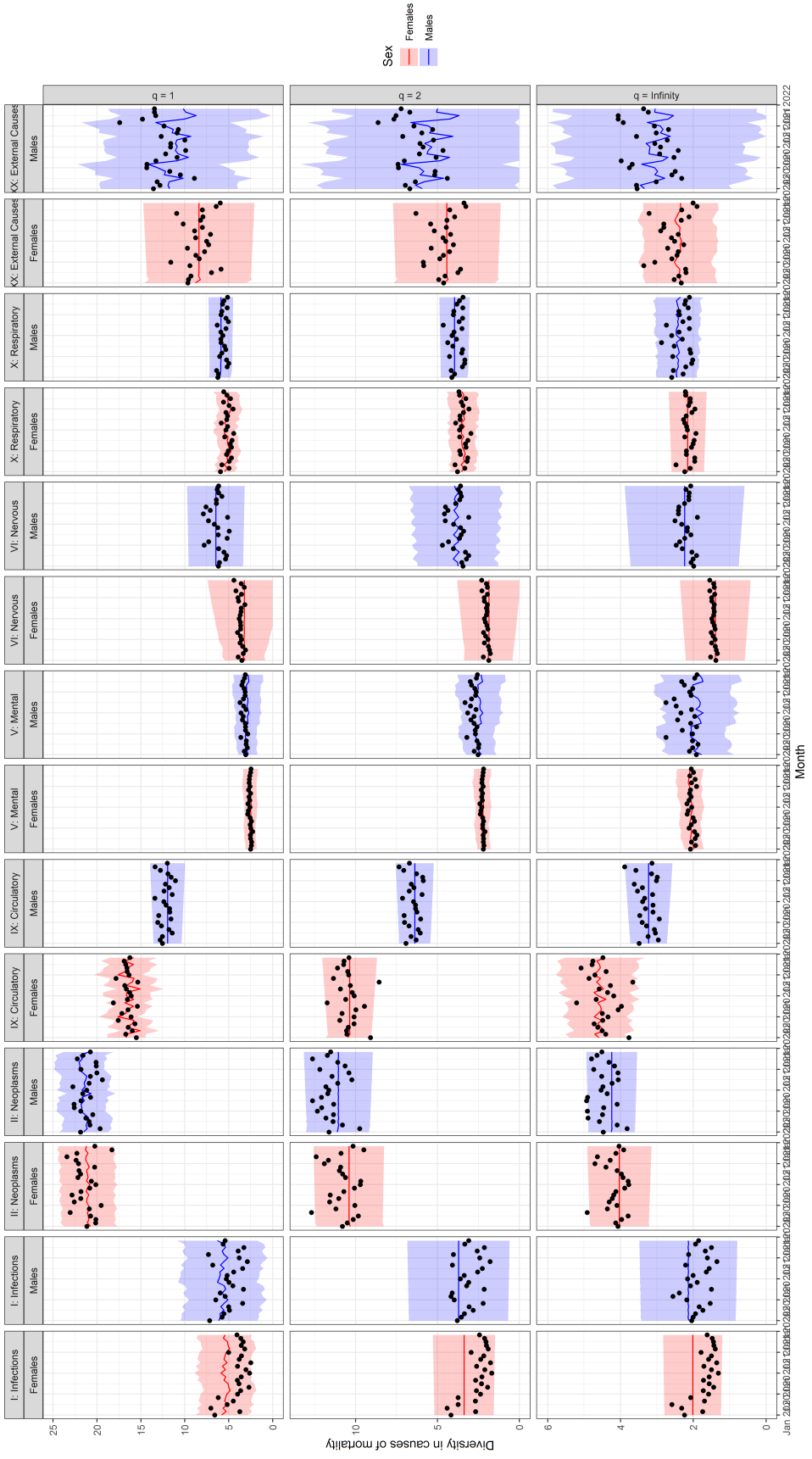
I21	Acute myocardial infarction	32.9	35.6	33.3	39.1	29.1
K70	Alcoholic liver disease	32.3	31	24.1	31.6	32.5
X42	Accidental poisoning by and exposure to narcotics and psychodysleptics [hallucinogens], not elsewhere classified	31.6	36.5	59.6	66.5	54.6
C34	Malignant neoplasm of bronchus and lung	28.9	29.4	23.9	24.8	24
U07	COVID-19	0	0	0	26.3	32.2
	Deaths among those aged 60 to 79					
	Females					
C34	Malignant neoplasm of bronchus and lung	205.4	219.9	215	197.9	198.4
J44	Other chronic obstructive pulmonary disease	155.9	153.2	139.8	111.5	113.7
I21	Acute myocardial infarction	97.9	105.5	98.7	104.3	93.5
C50	Malignant neoplasm of breast	72.7	70.9	67	66.4	69.7
I25	Chronic ischaemic heart disease	70.1	52	47.7	44.6	53.8
U07	COVID-19	0	0	0	135.1	126.3
	Males					
C34	Malignant neoplasm of bronchus and lung	259.2	257.5	250.5	224.8	224.9
I21	Acute myocardial infarction	207.4	229.1	227.6	242	232.1
I25	Chronic ischaemic heart disease	162.9	156.1	124.8	134.7	152.3
J44	Other chronic obstructive pulmonary disease	147.3	145.2	135	114.1	115.3
C61	Malignant neoplasm of prostate	83.2	88.1	86.4	90.9	83.1
U07	COVID-19	0	0	0	240.2	208.9
	Deaths among those aged 80+					
	Females					
F03	Unspecified dementia	794.2	847.1	698.1	656.4	634
G30	Alzheimer disease	726.3	897.7	898.3	932.2	887.5
F01	Vascular dementia	601.1	729.3	721.7	662.7	607.6
I21	Acute myocardial infarction	595.8	620.1	584.1	588.8	597.1
J18	Pneumonia, organism unspecified	518.5	500.6	400.8	301	285.2

J44	Other chronic obstructive pulmonary disease	482	499.4	435.8	353.1	303.9
U07	COVID-19	0	0	0	1283.2	751.3
		Males				
I21	Acute myocardial infarction	941.3	921.9	991.1	958.6	971.1
I25	Chronic ischaemic heart disease	717	739.6	635.1	630.4	597.4
F01	Vascular dementia	658.8	746.7	741.6	675.5	570
J44	Other chronic obstructive pulmonary disease	644.6	627.7	526.7	462.9	433.8
C34	Malignant neoplasm of bronchus and lung	626.5	571.2	540.7	530.9	534.7
J18	Pneumonia, organism unspecified	625.4	646.1	546.1	469.2	398.3
G30	Alzheimer disease	547.7	643.9	703.8	624.4	625.5
C61	Malignant neoplasm of prostate	529.7	612.6	586.8	558.7	600.5
U07	COVID-19	0	0	0	1781.8	1097.5

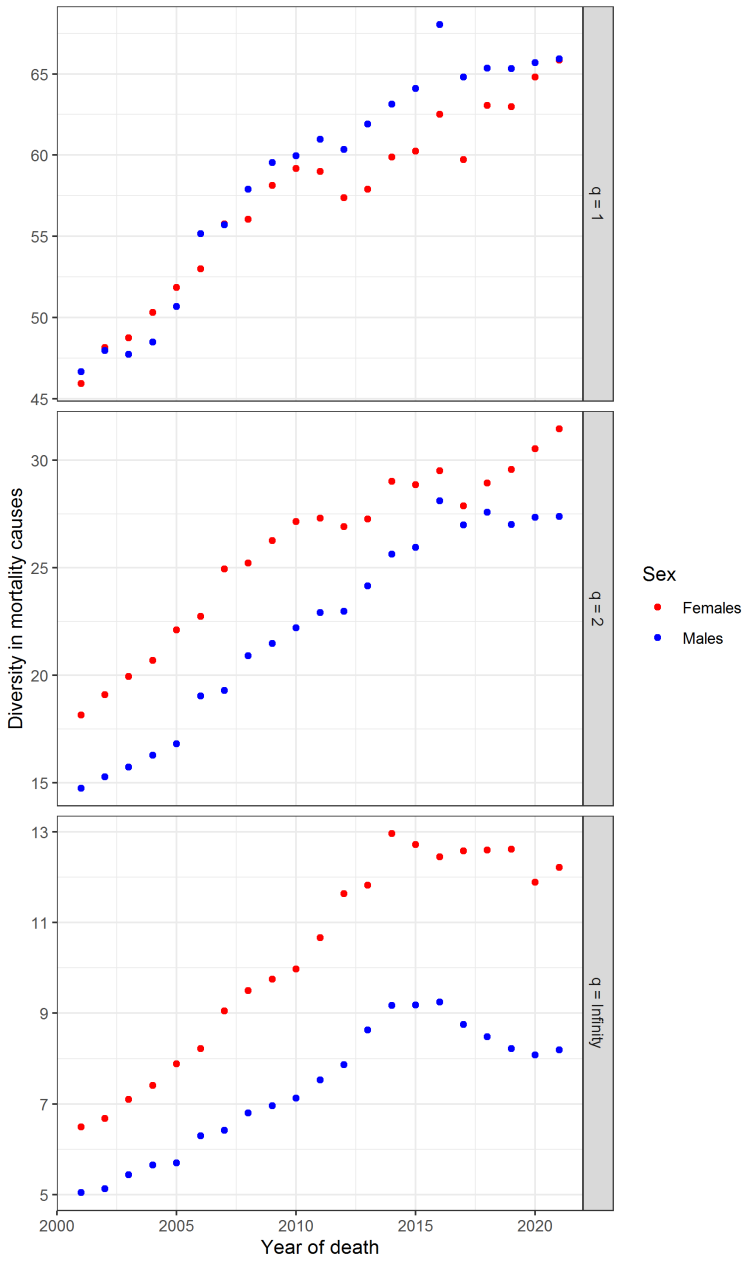
A.16 Diversity in mortality causes at q = 1, q = 2 and q = ∞ in each month in 2020 and 2021 in deaths among twenty-year age groups in males and females, with COVID-19 mortality included (green) and excluded (black). Coloured lines represent the central estimate of ARIMA model forecasts with ribbons indicating 99% confidence intervals.


Females
Males
0 to 19
20 to 39
40 to 59
60 to 79
80+
q = 1
q = 2
q = infinity
Diversity in causes of mortality
Month
Jan 2020
Jul 2020
Jan 2021
Jul 2021
Jan 2022
Sex
Females
Males

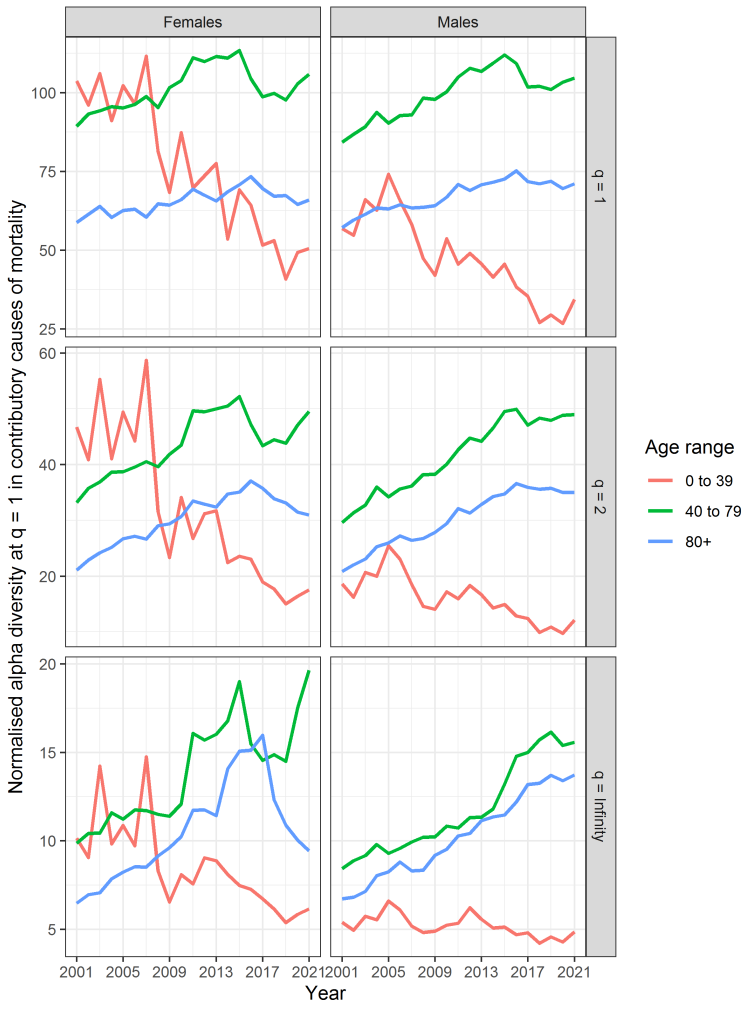
A.17 Diversity in mortality causes at q = 1, q = 2 and q = ∞ in each month in 2020 and 2021 in deaths in males and females separately in selected ICD-10 Chapters. Coloured lines represent the central estimate of ARIMA model forecasts with ribbons indicating 99% confidence intervals.


Sex
Females
Males
q = 1
q = 2
q = Infinity
I: Infections
II: Neoplasms
IX: Circulatory
V: Mental
VI: Nervous
X: Respiratory
XX: External Causes
Females
Males
Diversity in causes of mortality
Month

A.18 Normalised alpha diversity at $q = 1$, $q = 2$, and $q = \infty$ in causes of mortality within ICD-10 Chapter II: Neoplasms, in deaths among males and females at all ages across Scotland in each year 2001 to 2019.



A.19 Normalised alpha diversity at $q = 1$, $q = 2$, and $q = \infty$ of contributory causes of mortality in Scotland in males and females in deaths at ages 0 to 39, 40 to 79 and 80+ separately in Scotland from 2001 to 2021.



Appendix B

Alternative methods and measures for the calculation of diversity

Here, different methods for the measurement of diversity in causes of mortality are explored, testing the sensitivity of the conclusions drawn in this thesis to the use of alternative methods. First, I examine alternative measures of diversity, including some which treat the distribution of mortality causes differently to the measures used in this thesis. Secondly, I redistribute deaths which were attributed to garbage codes to meaningful causes of death and recalculate diversity in causes of mortality with garbage codes redistributed.

B.1 Alternative measures of diversity in causes of mortality

To examine the impact of alternative measures of diversity on the results reported in this thesis, in this section I compare four measures of diversity to those used in this thesis. These measures are: Shannon's Index, Shannon's Entropy and a measure referred to here as uniformity. These measures are used to calculate diversity in causes of mortality using distributions of mortality causes extracted from life tables. The calculation of these measures is described below

B.1.1 Methods

The measures of diversity used in this section are detailed below. First, I measure diversity in causes of mortality in males and females at all ages across Scotland under Shannon's index, using the *vegan* package in R. The formula for this measure is shown in Equation B.1, where *H'* is the Shannon index, and p_i is the proportion of deaths attributed to cause *i*. This measure can take values between 0, where all deaths are attributed to the same cause, and the natural logarithm of *N* (where *N* is the total number of causes of deaths in the population) when an equal share of deaths are attributed to every cause in the dataset (Shannon, 1948).

$$H' = -\sum_{i}^{N} p_i \log p_i \tag{B.1}$$

Secondly, I measure Shannon's entropy in causes of mortality in males and females at all ages across Scotland using methods described in another study of diversity in causes of mortality by Bergeron-Boucher et al. (2020). Normalised alpha diversity at q = i as used in the main body of this thesis is equivalent to the exponent of Shannon's Entropy as discussed in Chapter 3. The formula for Shannon's entropy is shown below in Equation B.2 where $\hat{H}$ is Shannon entropy, p_i is the proportion of deaths attributed to cause *i* and *N* is the total number of causes of mortality across all years in the study population. This measure can take values between 0, when all deaths are attributed to a single cause of mortality and 1 when all deaths are equally spread across causes of mortality.

$$\hat{H} = -\sum_{i=1}^{N} \frac{p_i \log(p_i)}{\log(N)} \tag{B.2}$$

Finally, I assess the distribution of mortality causes through a method described here as *uniformity* and defined by Equations B.3 and B.4 where p_i is the share of deaths attributed to each cause of mortality. In this section *N* is treated in two ways; to calculate *uniformity N* is the total number of causes of mortality recorded in any year in the population of interest during the study period. While to calculate *uniformity (excluding zero values) N* is treated as the total number of mortality causes recorded in a given year within a given population. Uniformity is more

closely related to measures of evenness than to traditional measures of diversity. Evenness measures aim to examine only how evenly distributed individuals are across the types in a distribution (that is how evenly distributed causes of mortality are within a population). Such measures do not account for "richness" meaning the number of causes of mortality recorded in each year. Richness and evenness are discussed by some in the field as being two components of diversity and many diversity measures, including those used in the body of this thesis aim to assess both components (Jost, 2010) . Unlike the measures of diversity used in the main body of this thesis, uniformity aims to account for zeros in the distribution of mortality causes. That is causes of mortality which do not occur in every year. Uniformity (excluding zero values) is measured here to aid in demonstrating the properties of this measure. This measure can take values between $1-(N-1)/N$ and 1; values of $1-(N-1)/N$ occur when all deaths are assigned to one cause of mortality in a year and, as above, values of 1 are reached if causes of mortality are equally distributed across causes.

$$uniformity = 1 - concentration \tag{B.3}$$

where

$$concentration = 1/2 \sum_{i}^{N} |p_i - 1/N| \tag{B.4}$$

In this section, diversity in causes of mortality was calculated in distributions of mortality cause extracted from multiple decrement life tables as described in Chapters 3, 4 and 7. First, these measures were used to calculate diversity in underlying causes of mortality in deaths across all ages in Scotland and Scottish subpopulations in the years 2001 to 2019. Then, they were used to calculated diversity in underlying causes of mortality in deaths within twenty-year age ranges in Scotland in the years 2001 to 2019. Finally, each of the measures described above was used to calculate diversity in the contributory causes of mortality recorded in deaths at all ages across Scotland in each year 2001 to 2019.

B.1.2 Analysis of alternative measures of diversity in mortality causes in deaths across all ages across Scotland and in Scottish subpopulations

Figure B.1 shows trends in the diversity of all underlying causes of mortality recorded in Scotland from 2001 to 2019 under each of the three measures of diversity described above and under normalised alpha diversity at $q = 1$, $q = 2$, and $q = \infty$ using Reeve et al. (2016) framework as used in the body of this thesis. As can be observed, trends under uniformity (excluding zero values), Shannon Index, and Shannon Entropy are shown to be similar to those reported in this thesis, particularly to diversity at $q = 1$. The results presented in this thesis are therefore upheld under these alternative measures. When considering a measure more closely related to evenness, trends are shown to differ somewhat from those reported in this thesis. In deaths among females, the trend in uniformity from 2001 to 2014 is effectively flat indicating no change under this measure while in deaths among males a slightly positive trend is observed. From 2015 onwards in both sexes, uniformity is observed to increase markedly. The reason for this sudden increase is unclear though given its timing it may be associated with the Recording of Death Act (2011), discussed in Chapter 3, which took effect in that year meaning it may be associated with improved recording accuracy. Overall, uniformity in causes of mortality was higher in 2019 than in 2001, indicating an increase in evenness similar to the increase in diversity reported throughout this thesis. However, despite this there is not clearly a trend of increasing uniformity as is observed under other measures of diversity.

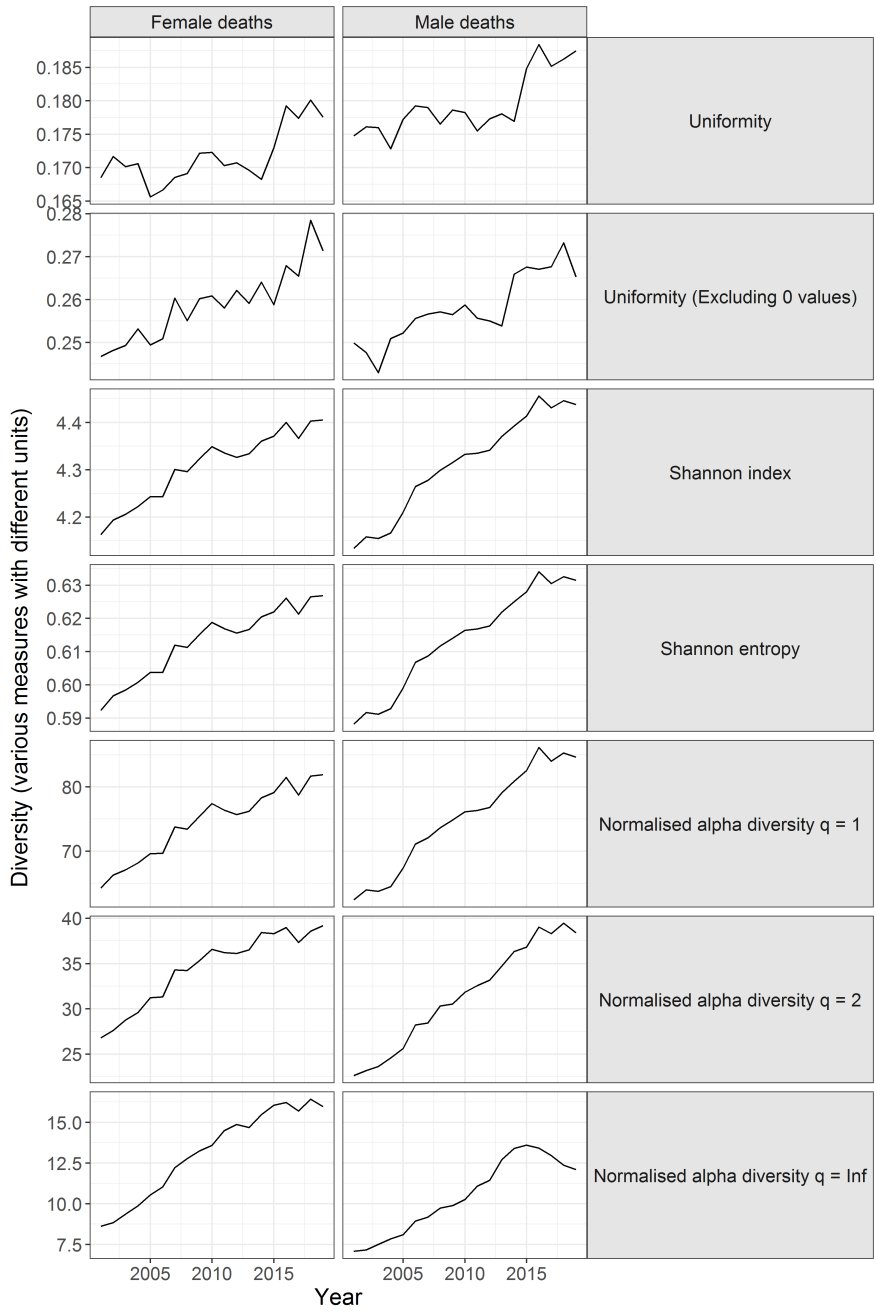


Figure B.1: Diversity in causes of mortality, calculated - from distributions of mortality causes extracted from multiple-decrement life tables - under: uniformity, uniformity (excluding zero values), Shannon's Index, under Shannon Entropy, and under the measures used in the main body of this thesis. Trends in the diversity of mortality causes from 2001 to 2019 are shown in males and females in Scotland separately.

B.1.2.1 Scottish Subpopulations

In Figure B.2, diversity in mortality causes (across all causes of mortality) is measured in deaths within SIMD income deprivation quintiles using the alternative measures of diversity discussed above. Trends in the diversity of mortality causes measured using uniformity (excluding zero values), Shannon's Entropy and Shannon's Index are shown to be relatively similar to trends under the measures of diversity used in the body of this thesis. Namely, a positive trend is found in most quintiles across the study period and in all quintiles diversity is found to be higher in in 2019 than in 2001. When considering uniformity, an increase is still observed and from 2001 to 2019 in both sexes. However, as observed in deaths across Scotland, trends in uniformity are mostly flat from 2001 to around 2015. As discussed above this indicates that when examining measures more related to evenness and taking account of zero values in the distribution of mortality the trends noted in Chapter 4 and throughout this thesis may not be observed.

Figure B.2: Diversity in causes of mortality, calculated - from distributions of mortality causes extracted from multiple-decrement life tables - under: uniformity, uniformity (excluding zero values), Shannon's Index, under Shannon Entropy, and under the measures used in the main body of this thesis. Trends in the diversity of mortality causes from 2001 to 2019 are shown in males and females in SIMD income deprivation quintiles separately.

Figure B.3 shows analysis of alternative measures of diversity in the measurement of mortality cause diversity in Scottish Government urban-rural classes. As discussed for SIMD quintiles above, the trends reported in Chapter 4 are observed to be consistent when diversity is measured using uniformity (excluding zero values) Shannon's Index and Shannon's Entropy. Similarly to above, trends diverge slightly when examining uniformity. In rural areas and small towns, a relatively flat trend is observed under uniformity in deaths among both males and females. In urban areas, uniformity changes little from 2001 to 2015 then increases in the final years of the study period leading to an increase in diversity in 2019 compared to 2001. Uniformity is also noted to be markedly higher in urban areas than other areas of Scotland, this is also observed when diversity is measured under q = 1 in normalised alpha diversity but to a lesser extent.

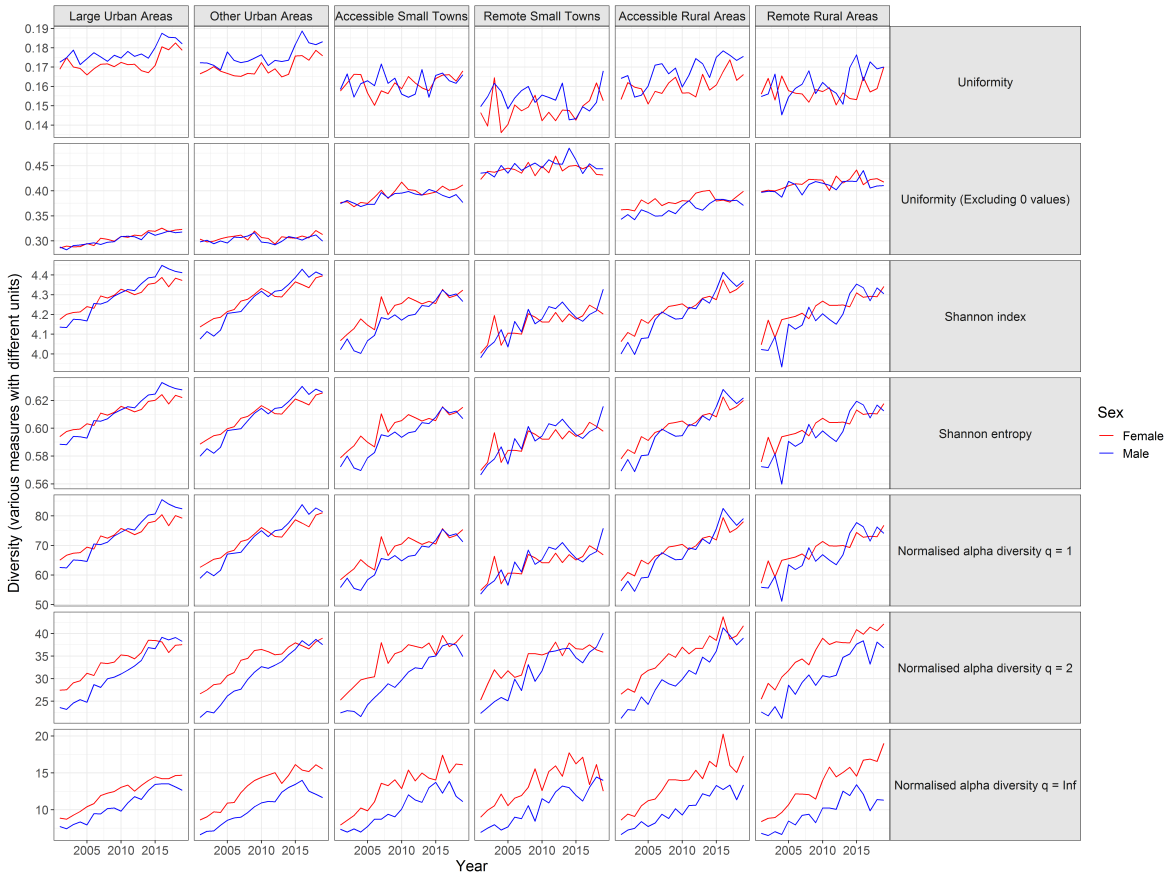


Figure B.3: Diversity in causes of mortality, calculated - from distributions of mortality causes extracted from multiple-decrement life tables - under: uniformity, uniformity (excluding zero values), Shannon's Index, under Shannon Entropy, and under the measures used in the main body of this thesis.Trends in the diversity of mortality causes from 2001 to 2019 are shown in males and females in SIMD income deprivation quintiles separately.

B.1.3 Analysis of alternative measures of diversity in mortality causes in deaths separated by age

Here, the analysis described in the previous section is repeated for deaths in Scotland within twenty-year age ranges, with diversity measured in distributions of mortality causes extracted from multiple decrement life tables. Figure B.4 shows that under Shannon's index and Shannon's entropy, diversity in mortality causes follows similar trends to those found when diversity is measured using the Reeve et al. (2016) framework. Under uniformity (excluding zero values) no clear trend is found among deaths in each age range. Uniformity is shown to to fall from 2001 to 2019 in death among those aged 0 to 19 and to rise over this period in deaths among those aged 60 to 79 and 80+. These trends correspond to those reported in

this thesis, under Shannon's index and under Shannon's entropy. In deaths among those aged 20 to 39 a very slight reduction over time is observed when assessing the distribution of mortality causes using uniformity, while in diversity measured under Shannon's index, Shannon's entropy and the measures used in this thesis show a more marked reduction. In deaths among those aged 40 to 59, a slight increase in diversity is observed from 2001 to 2019 in the measures used in this thesis. However, under uniformity no clear trend is observed over this period. Therefore, for deaths in most age ranges the trends reported in this thesis are upheld under all measures described in this section. However, for deaths in those aged 20 to 39 and 40 to 59, when zero values are accounted for, using uniformity, trends differ slightly.

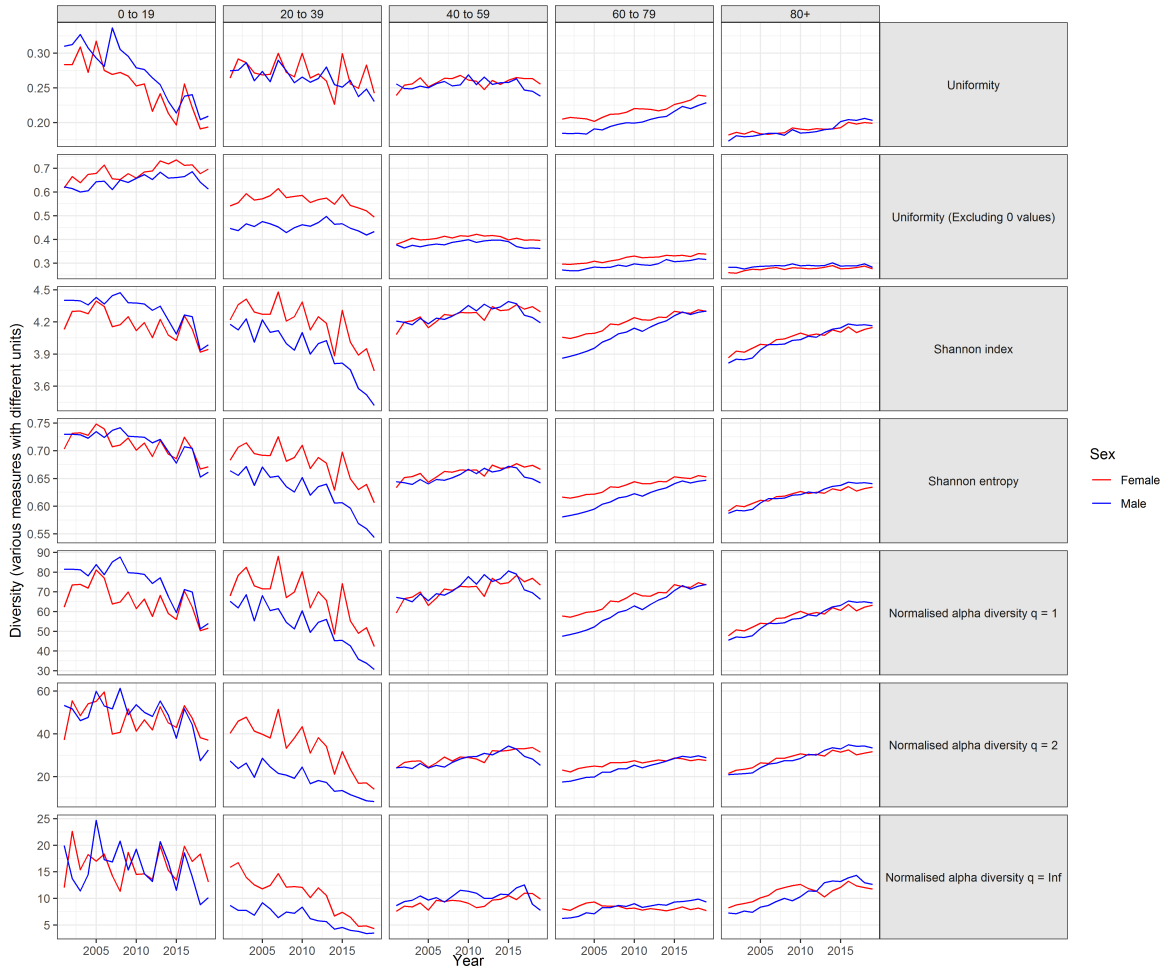


Figure B.4: Diversity in causes of mortality, calculated from distributions of mortality causes extracted from multiple-decrement life tables, calculated under: uniformity, uniformity (excluding zero values), Shannon's Index, under Shannon Entropy, and under the measures used in the main body of this thesis. Trends in the diversity of mortality causes from 2001 to 2019 are shown for deaths within twenty-year age ranges in males and females in Scotland separately.

B.1.4 Alternative measures of diversity in contributory causes of mortality

Finally in this section, diversity is measured in contributory causes of mortality in Scotland. Contributory causes of mortality were extracted from multiple decrement life tables as described in Chapter 7. Shannon's entropy, Shannon's index and uniformity were then calculated as described above for the distribution of contributory mortality causes across Scotland in deaths among males and females separately in each year 2001 to 2021. Figure B.5 compares trends in diversity under these alternative measures to trends measured under the Reeve et al. (2016) framework. Trends under Shannon's index and Shannon's entropy are shown to be qualitatively identical to those under the measures used in Chapter 7. Diversity increases from 2001 to 2015 in both sexes, after which diversity falls in females and to falls slightly then plateaus in males. These patterns are discussed in Chapter 7. Using measures more related to evenness, under both uniformity and uniformity (excluding zero values), a reduction is observed from 2001 to around 2010 in females, with no clear trend in males. After 2010 which, trends in uniformity and uniformity (excluding zero values) are found to be relatively similar to those under the other measures discussed above. That is to say there is an increase from 2010 to 2015 in both sexes followed by a reduction in females and sharp reduction followed by a plateau in males. This indicates a distinct difference in the observed trends in diversity in mortality causes in the 2000s dependant on the measure used. Using measures related to evenness reductions are found while under more traditional diversity measures increases over time are observed. The difference in trends under this alternative measure should be considered when examining the trends and conclusions noted in Chapter 7. This consideration is discussed in the chapter itself.

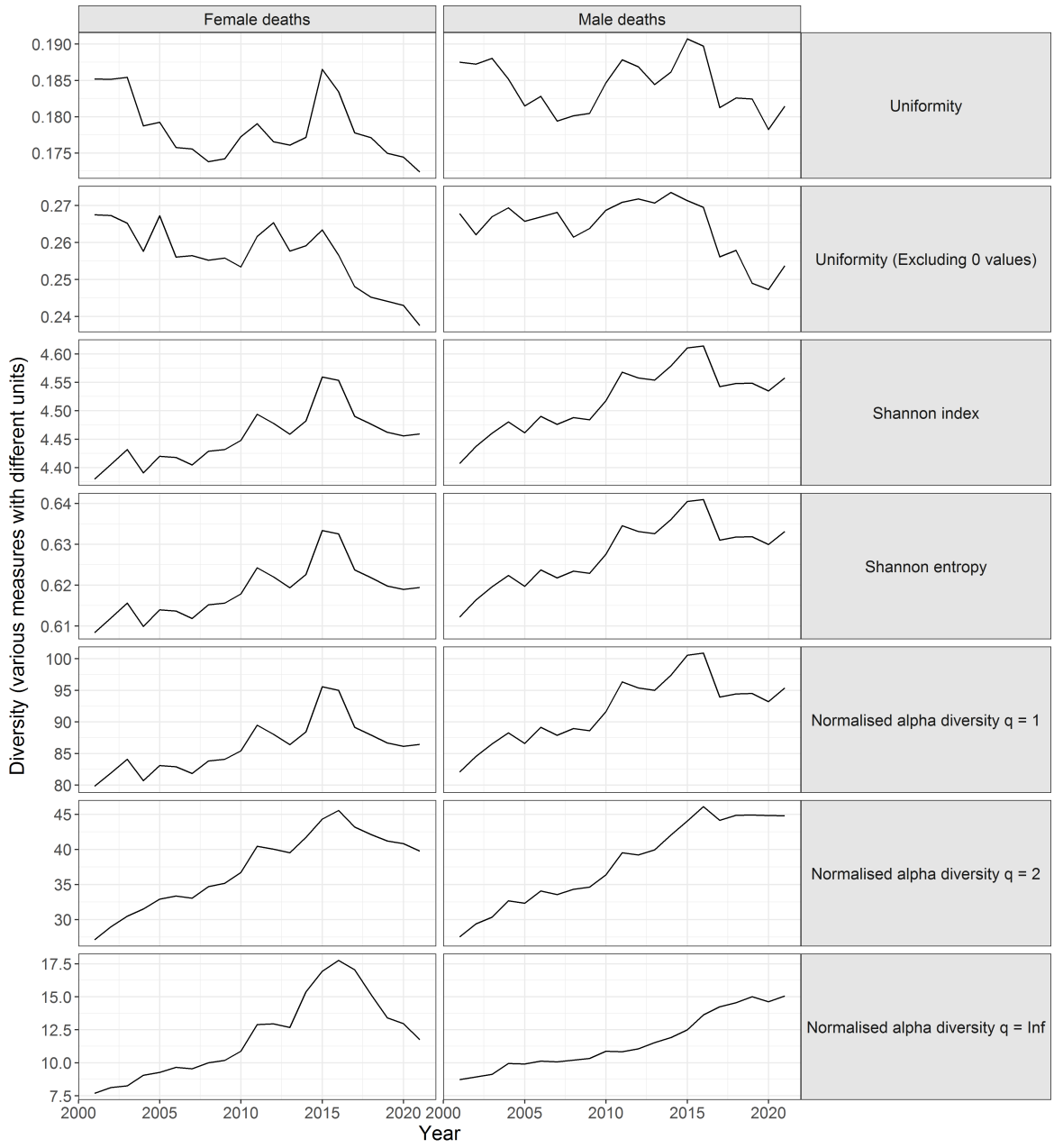


Figure B.5: Diversity in contributory causes of mortality, calculated from distributions of mortality causes extracted from multiple-decrement life tables, calculated under the measures used in this thesis, under Shannon's Index, Shannon Entropy and uniformity. Trends in the diversity of mortality causes from 2001 to 2019 are shown for deaths in males and females in Scotland separately.

B.2 Redistribution of garbage codes

So-called "garbage codes" are ICD-10 codes which are not deemed useful or relevant in the study of population health (Ellingsen et al., 2022). Generally this is because the cause of mortality they signify is vague or uninformative. Deaths assigned to these causes are, in some analysis of population health data, redistributed to different causes of mortality with the aim of providing more meaningful analysis (Ellingsen et al., 2022). In this thesis garbage codes were retained in the distribution of mortality causes. To assess the sensitivity of the results of this thesis to garbage codes, in this section I perform a simple redistribution of garbage codes to valid ICD-10 three character codes and compare trends in diversity to those calculated from the observed distribution of mortality causes. First, I calculate diversity in underlying mortality causes with garbage codes redistributed in deaths at all ages across Scotland, in SIMD income deprivation quintiles and Scottish Government urban-rural classes. Then, diversity in underlying mortality causes with garbage codes redistributed is calculated for deaths across Scotland within twenty-year age ranges. Finally, diversity in contributory causes of mortality is calculated with garbage codes redistributed in deaths across Scotland.

B.2.1 Methods

Here, I redistribute deaths assigned to garbage codes to other causes of mortality to examine whether trends in the diversity of mortality causes calculated under this alternative distribution are different to those reported in this thesis. For this analysis garbage codes were any codes described as GBD Level 1, 2, or 3 garbage codes[1] according to Ellingsen et al. (2022). Redistribution of deaths assigned to garbage codes was performed proportionally, using the distribution of mortality causes extracted from multiple decrement life-tables. Deaths attributed to each garbage code were reassigned to valid codes which occurred in the same sex, at the same age, in the same year and within the same ICD-10 Chapter for Chapters I-XVII and XIX-XXII (All causes within Chapter XVIII can be considered garbage codes and so any deaths attributed to these causes were reassigned proportionally across all causes). For example, all deaths which were attributed to garbage codes within Chapter I in males aged 65 in 2001 were given new valid codes with proportion equal to the share of deaths attributed to each valid code in Chapter I in males aged 65 in 2001.

[1]The GBD system has one further level of garbage code however, these garbage codes are described as having "limited policy implications" and so are excluded here.

B.2.2 Diversity in underlying causes of mortality with garbage codes redistributed in deaths across Scotland and in Scottish subpopulations.

The proportion of deaths attributed to garbage codes in Scotland and in the Scottish subpopulations examined in this study in each year 2001 to 2019 is shown in Figure B.6. In most cases the proportion of deaths assigned to garbage codes was consistently between 10% and 15%. However, in male deaths within the most deprived SIMD quintile, an increase in the proportion of deaths assigned to garbage codes can be observed in the latter years of the study period. This increase is driven mostly by an increase in the proportion of deaths assigned to cause X42: "Accidental poisoning by and exposure to narcotics and psychodysleptics [hallucinogens], not elsewhere classified". The increase in prevalence of this cause has not been widely reported but is likely linked to the overall increase in deaths attributed to causes within ICD-10 Chapter XX: External causes in the 2010s which has been more widely discussed in the literature (Allik et al., 2020) and is referenced throughout this thesis .

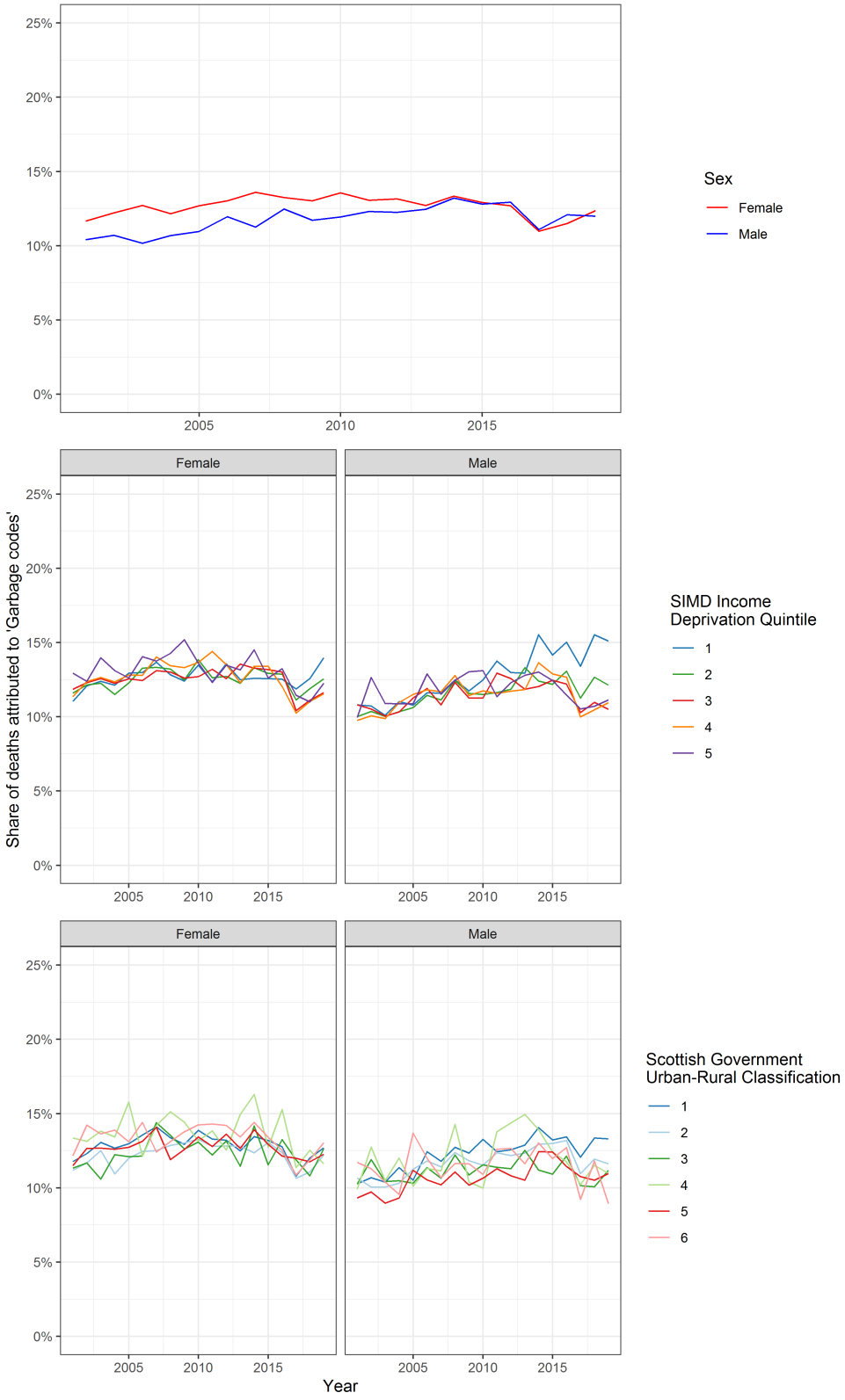


Figure B.6: The share of deaths attributed to garbage codes among males and females of all ages in the distribution of mortality causes extracted from multiple decrement life tables in each year 2001 to 2019. Trends in the share of deaths attributed to garbage codes are shown separately for Scotland as a whole, SIMD income deprivation quintiles and Scottish Government urban-rural classes.

B. Alternative methods and measures for the calculation of diversity

Once garbage codes were redistributed to valid codes, diversity in causes of mortality was calculated under the Reeve et al. (2016) framework as described in Chapter 3. Diversity in this redistributed set of causes of mortality was calculated in each year from 2001 to 2019 in deaths at all ages among males and females across Scotland. Figures B.7 shows trends in diversity in mortality causes calculated from the observed distribution of mortality causes and from the redistributed mortality causes described above.

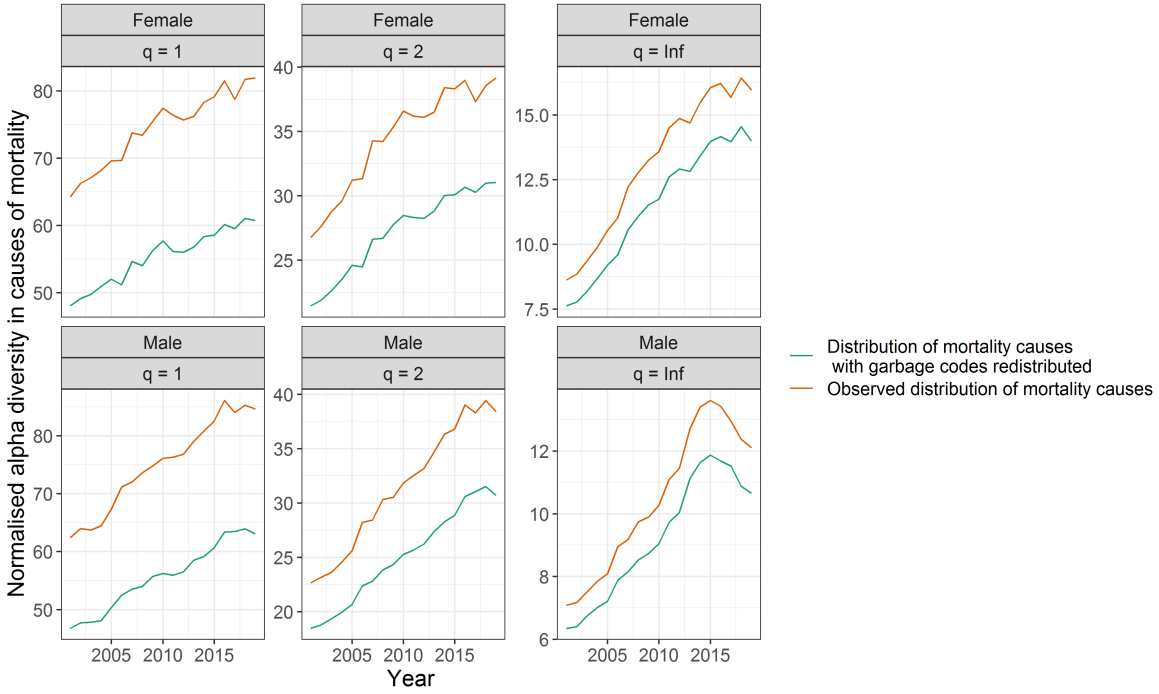


Figure B.7: Comparison of trends over time in diversity in mortality causes at q = 1, q = 2 and q = ∞ calculated from the observed distribution of mortality causes (red lines) and from the distribution of mortality causes with garbage codes redistributed to valid codes (blue lines). Diversity is calculated in both cases in the distribution of mortality causes extracted from multiple decrement life tables for males and females across the population of Scotland as a whole in each year 2001 to 2019.

It is clear from Figure B.7 that the redistribution of causes has an effect on the measurement of diversity in each year of the study period. However, it does not affected overall trends and rather causes an intercept shift, decreasing diversity in almost every year. To show this more clearly, diversity in causes of mortality with garbage codes redistributed is plotted against diversity in mortality causes without redistribution in Figure B.8. Together, these figures show that while the redistribution of garbage codes does have an effect on the measurement of diversity in mortality causes, there is no significant impact on the trends in diversity in causes of mortality reported in this thesis. Figures B.9 and B.10 recreate Figure

B.7, showing trends in diversity in redistributed mortality causes in SIMD income deprivation quintiles and Scottish Government urban-rural classes showing that the redistribution of garbage codes does not qualitatively impact trends in these subpopulations.

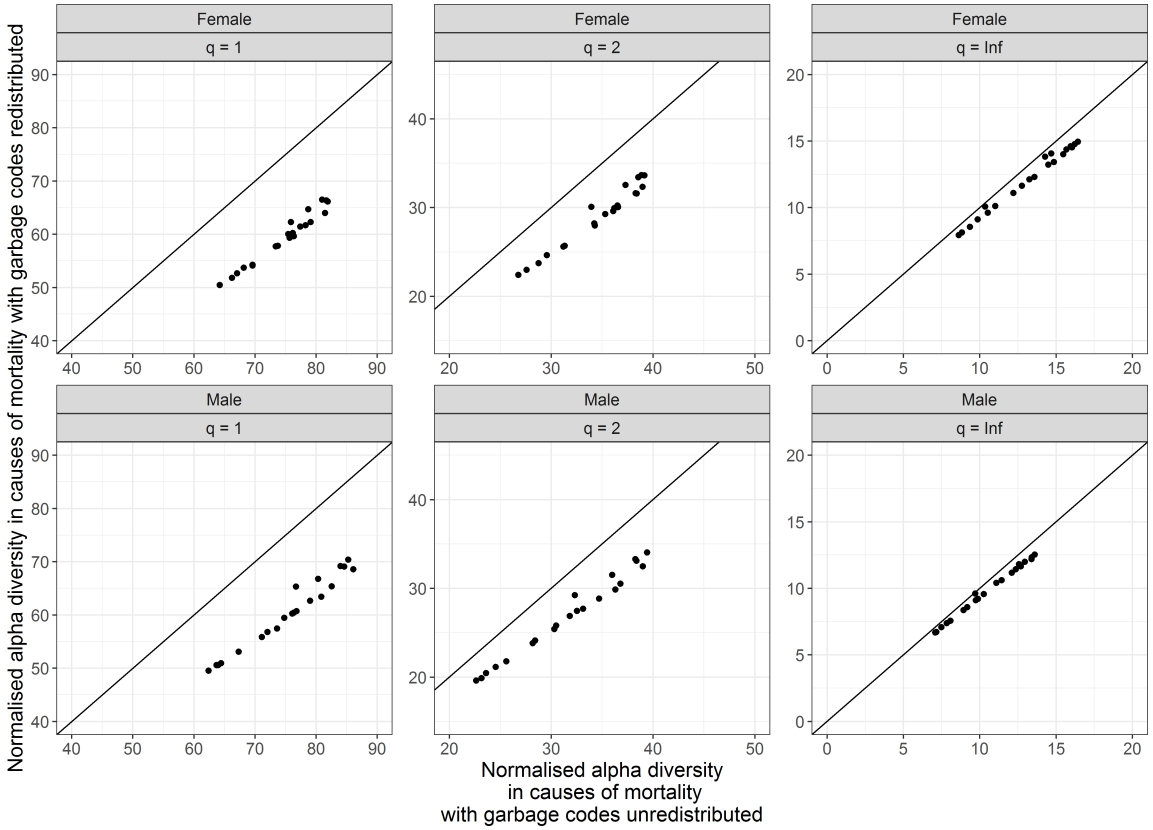


Figure B.8: Direct comparison of diversity in mortality causes at q = 1, q = 2 and q = ∞ calculated from the observed distribution of mortality causes and from the distribution of mortality causes with garbage codes redistributed to valid codes in each year 2001 to 2019 with a diagonal line indicating a 1:1 relationship. Diversity is calculated in both cases in the distribution of mortality causes extracted from multiple decrement life tables in each year 2001 to 2019.

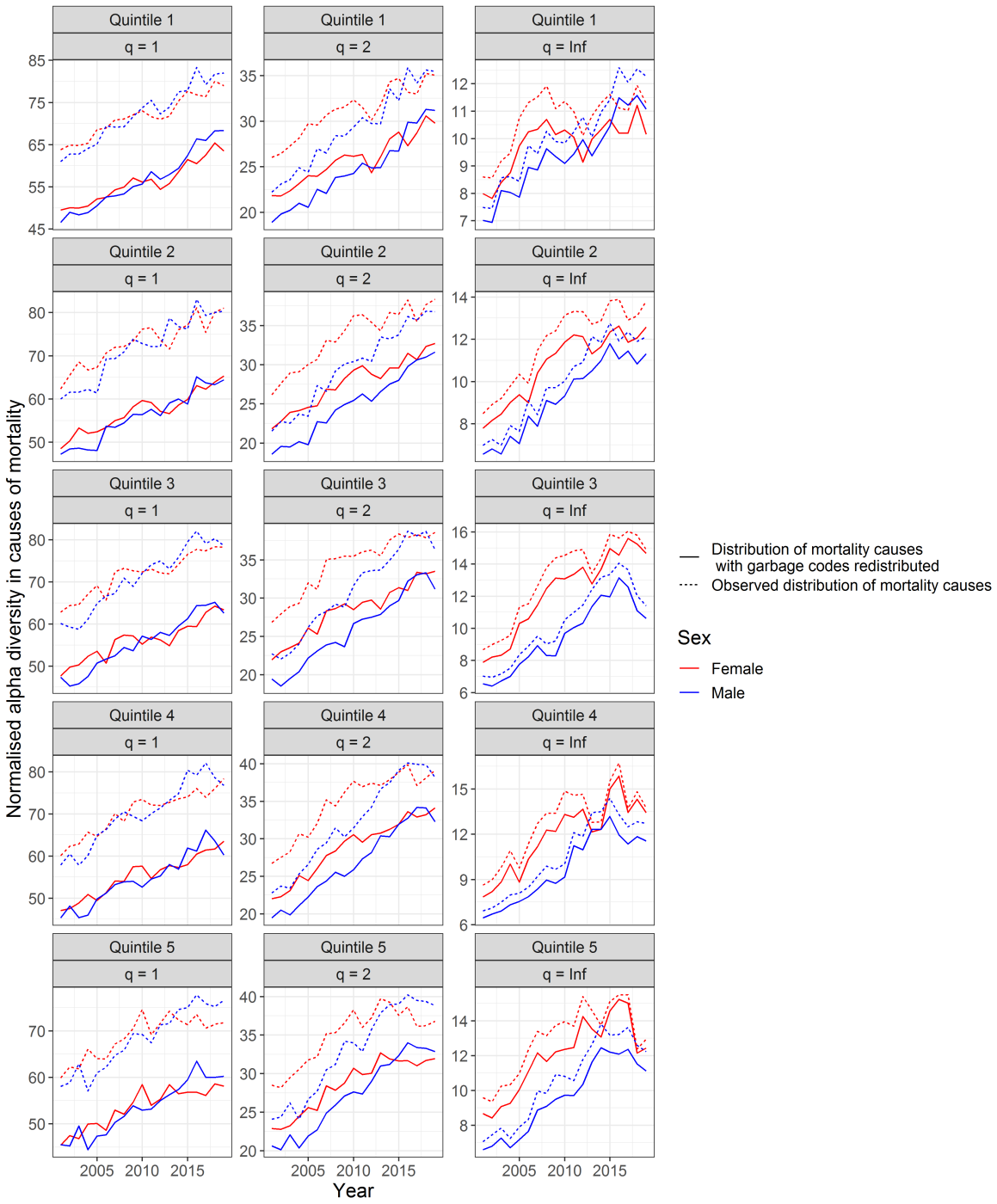


Figure B.9: Comparison of trends over time in diversity in mortality causes at q = 1, q = 2 and q = ∞ calculated from the observed distribution of mortality causes (solid lines) and from the distribution of mortality causes with garbage codes redistributed to valid codes (dotted lines). Diversity is calculated in both cases in the distribution of mortality causes extracted from multiple decrement life tables for males and females within SIMD income deprivation quintiles in each year 2001 to 2019.

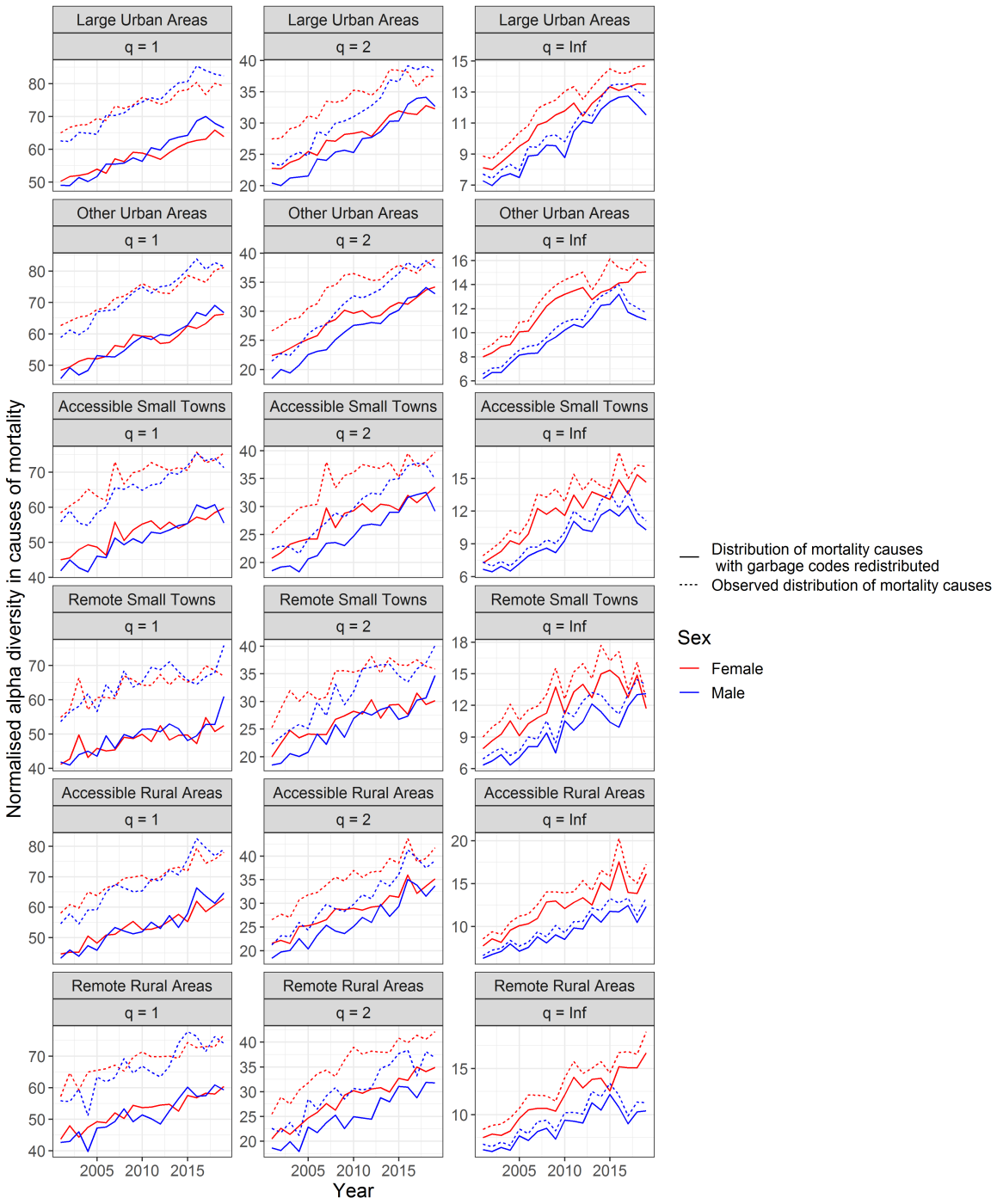


Figure B.10: Comparison of trends over time in diversity in mortality causes at q = 1, q = 2 and q = ∞ calculated from the observed distribution of mortality causes (solid lines) and from the distribution of mortality causes with garbage codes redistributed to valid codes (dotted lines). Diversity is calculated in both cases in the distribution of mortality causes extracted from multiple decrement life tables for males and females within Scottish Government urban-rural classes in each year 2001 to 2019.

B.2.3 Diversity in underlying causes of mortality with garbage codes redistributed in deaths within twenty-year age ranges

In this section, I examine the impact of redistributing garbage codes on the measurement of diversity in mortality causes in deaths across Scotland separated by age. Figure B.11 shows the share of deaths which were attributed to garbage codes in deaths within twenty-year age ranges in Scotland in each year from 2001 to 2019. This figure shows that deaths among those aged 20 to 39 were most likely to be attributed to a garbage code with as many as 45% of deaths in this age-range attributed to garbage codes. Furthermore, while the share of deaths attributed to garbage codes has changed little during the study period in deaths at most age ranges it has increased markedly in deaths among those aged 20 to 39. This is primarily due to the high prevalence of cause X42: "Accidental poisoning by and exposure to narcotics and psychodysleptics [hallucinogens], not elsewhere classified" which was the most common cause of death among males this age group for much of the 2010s.

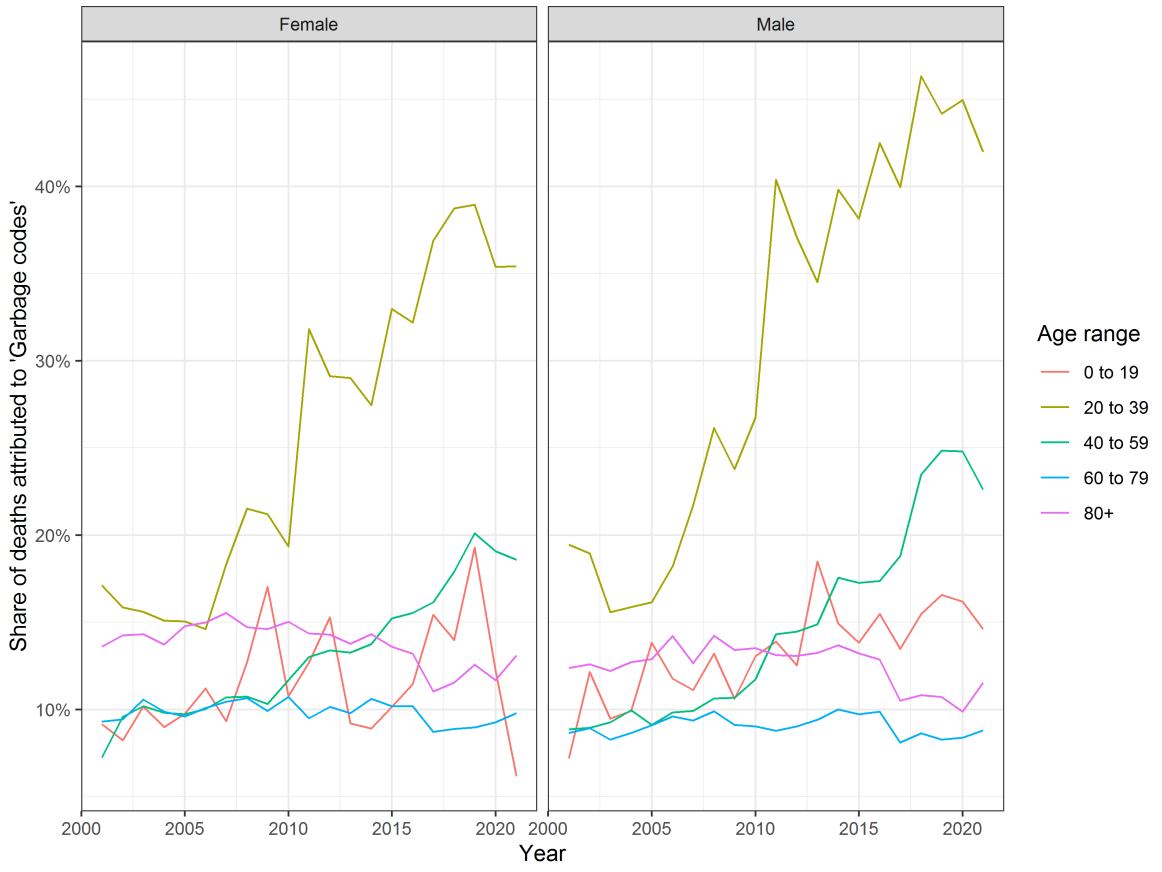


Figure B.11: The share of deaths attributed to garbage codes in the distribution of mortality causes extracted from multiple decrement life tables among males and females within twenty-year age ranges in each year 2001 to 2019.

B. Alternative methods and measures for the calculation of diversity

Deaths attributed to garbage codes were redistributed as described above. Following this redistribution, diversity in causes of mortality was calculated under the Reeve et al. (2016) as described in Chapters 3 and 4 for deaths among males and females within each twenty-year age range separately in each year 2001 to 2019. Diversity calculated from the redistributed distribution of mortality causes is plotted alongside diversity calculated from the observed distribution of mortality causes in Figure B.12.

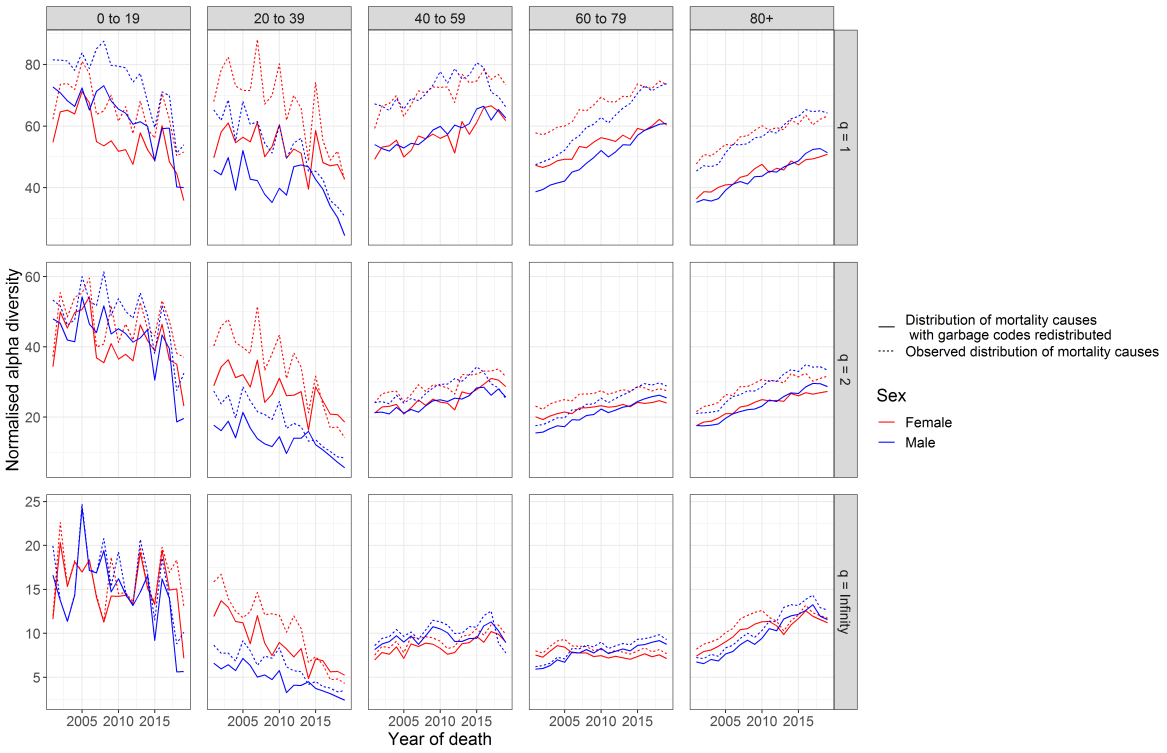


Figure B.12: Comparison of trends over time in diversity in mortality causes at q = 1, q = 2 and q = ∞ calculated from the observed distribution of mortality (solid lines) and from the distribution of mortality causes with garbage codes redistributed to valid codes (dotted lines). Diversity is calculated in both cases in the distribution of mortality causes extracted from multiple decrement life tables for males and females within twenty-year age ranges in each year 2001 to 2019.

Despite the large share of deaths attributed to garbage codes at some ages, Figure B.12 shows that trends in diversity in mortality causes are remarkably consistent whether calculated from observed distributions of mortality causes or redistributed mortality causes. There are likely two main, and interconnected, reasons for this: first, the distribution of mortality causes in Scotland is highly dominated by a relatively small number of causes as noted in Chapters 4 and 6; because of this distributional changes can have a limited effect on overall diversity. Secondly, the method used for redistribution involved reassigning garbage code deaths to causes of mortality proportionally meaning that much of the distribution of mortality was

maintained despite the redistribution. Together, these reasons help to explain why the effect of redistribution is most salient under diversity at $q = 1$ which is least sensitive to changes in the most common causes and more sensitive to changes in the distribution of relatively rare causes. In other terms, the redistribution of causes does cause there to be fewer rare causes of mortality (because most garbage codes are relatively rare) but does not have a large effect on the dominance of the most common causes. To use an example, among males aged 20 to 39 in 2019 the two most common causes were X42, described above, and X70 "Intentional self-harm by hanging, strangulation and suffocation" which together caused 46% of deaths. Following redistribution the most common causes in this age group were X70 and X99 "Assault by sharp object", together allocated as the cause of 44% of deaths, both of these causes are within ICD-10 Chapter XX: External Causes of Morbidity and Mortality. Cause X42 is also within Chapter XX meaning many of the deaths attributed to X42 in the observed data were reassigned to X70 and X99. In other words, the proportional redistribution caused a large share of deaths assigned to a common cause to be attributed to other common causes, reducing diversity slightly. While considerable change occurred in the distribution of mortality causes, it did not lead to a large change in the dominance of the most common causes. This causes diversity in mortality causes to change little, particularly at more conservative measures. Although other methods of redistribution might lead to different impacts on diversity, the fact that qualitative patterns in trends in diversity remain after the redistribution approach tested here helps to demonstrate the overall robustness of the results in this thesis.

B.2.4 Redistribution of contributory causes of mortality

Finally, in this section I present analysis of diversity in contributory causes of mortality in Scotland with garbage codes redistributed to valid ICD-10 three character codes. Contributory causes of mortality were defined and extracted from multiple decrement life tables as described in Chapter 7. The redistributive algorithm described above was applied to this distribution of mortality causes reassigning garbage codes to valid codes proportionally by exact age at death, sex, and ICD-10 Chapter. Diversity in the distribution of contributory causes of mortality was then calculated as described in Chapter 7. Figure B.13 shows diversity in contributory mortality causes calculated from the distribution of mortality causes with garbage codes redistributed, alongside diversity in contributory mortality causes in the observed distribution of mortality causes in each year 2001 to 2019. This

figure shows that from 2001 to 2015, the results reported in Chapter 7 are upheld when garbage codes are redistributed and that diversity in contributory causes of mortality increased in both sexes and across measures of diversity. Following 2015 trends differ, diversity in the observed distribution of contributory causes falls in females under diversity at q = 1, q = 2 and q = ∞ and plateaus under diversity at q = 1 and q = 2 in males. When diversity is measured in contributory causes of mortality with garbage codes redistributed an initial reduction is noted from 2015 to 2017 under diversity at q = 1 and q = 2 in both sexes after which diversity continues to increase. Under diversity at q = ∞ no reduction in diversity after 2015 in either sex is observed though increases slow and plateau after 2019.

This analysis suggests that diversity in valid contributory causes of mortality continued its general trend of increase across the study period from 2001 to 2019 and did not fall following 2015. This has an impact on some of the conclusions discussed in Chapter 7; and should be considered when analysing the results of this Chapter.

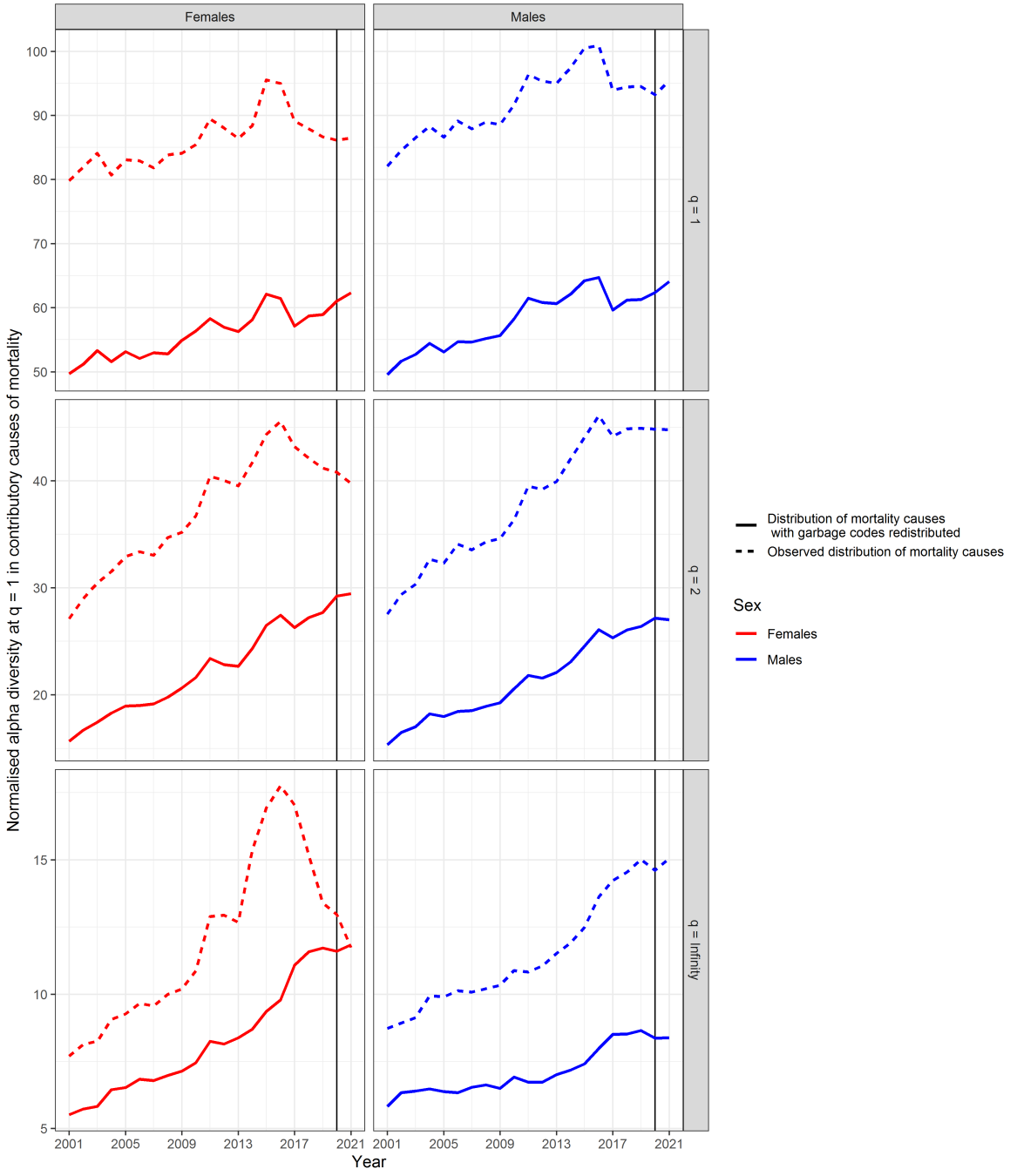


Figure B.13: Comparison of trends over time in diversity in contributory causes of mortality at q = 1, q = 2 and q = ∞ calculated from the observed distribution of mortality causes (solid lines) and from the distribution of mortality causes with garbage codes redistributed to valid codes (dotted lines). Diversity is calculated in both cases in the distribution of contributory causes of mortality extracted from multiple decrement life tables for males and females separately in each year 2001 to 2019.

Appendix C

Further analysis of diversity in mortality causes within subpopulations among twenty-year age ranges

C.1 Diversity trends

The overall tends in diversity in causes of mortality across Scotland by age range observed within Chapter 4 hold within both SIMD income deprivation quintiles and Scottish government urban-rural classifications. Diversity in mortality causes fell in those younger than 40 and increased in most cases in older age ranges. At older age ranges less variation in diversity between groups is noted in both sets of subpopulations.

C.1.1 SIMD income deprivation quintiles

In men, those in the most deprived areas tended to face the least diversity in mortality causes at q values >1 across the age ranges examined here. In contrast, for diversity at q = 1, this pattern is only seen among those aged 60 to 79 and 80+ with no clear pattern at younger ages.

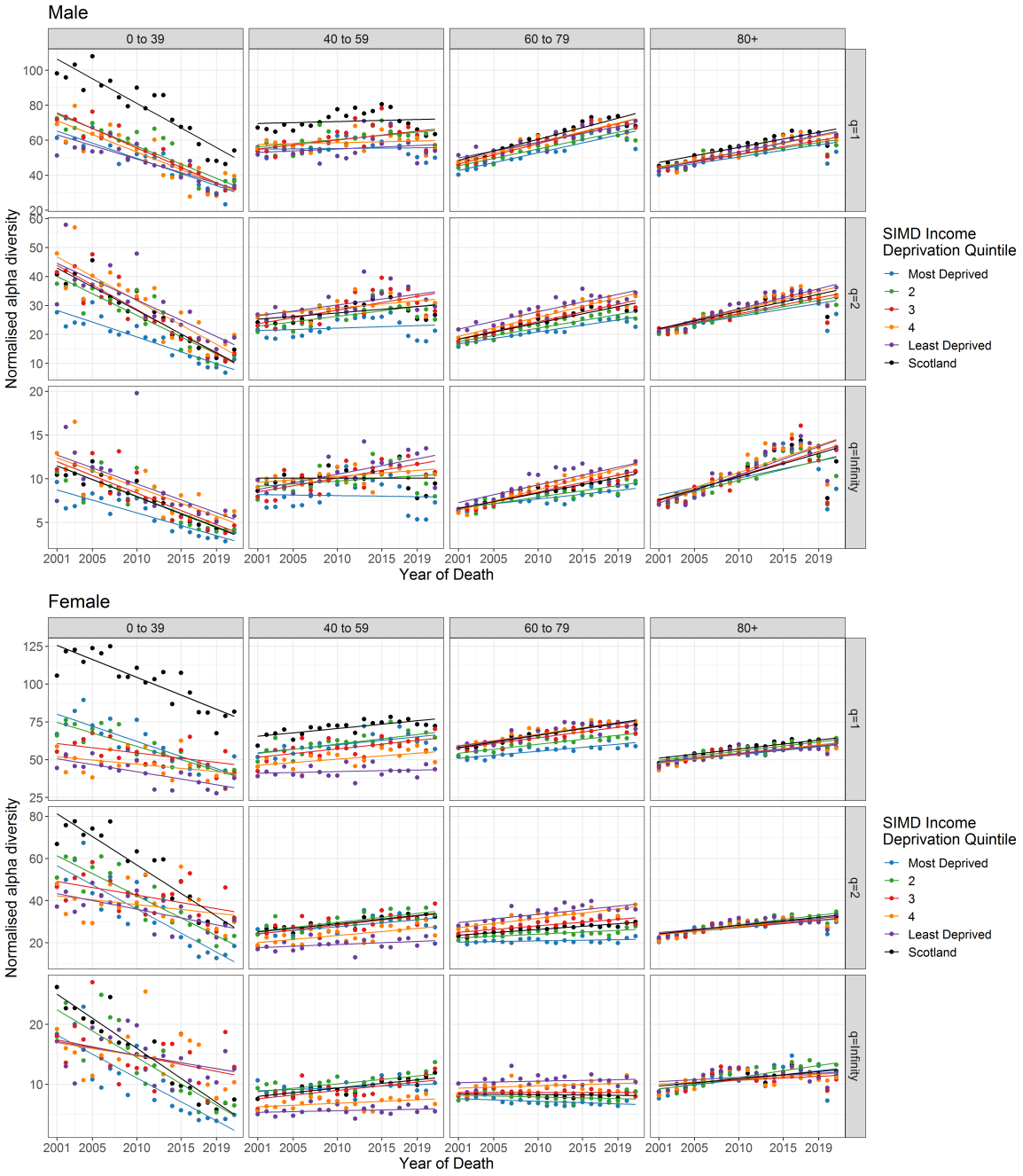


Figure C.1: Trends in diversity in mortality causes from 2001 to 2019 in deaths across twenty-year age ranges in SIMD income deprivation quintiles.

In women, more variation in trends across ages can be observed. Diversity in causes of mortality in deaths among the youngest age group tended to be higher in more deprived areas under each diversity measure in 2001; however diversity fell most quickly in these areas. Under the more conservative measures, this led to the most deprived areas facing the least diverse causes by 2019.

In deaths among women aged 40 to 59 the least diverse causes were found in the least deprived areas across measures and diversity changed little across the study period in these areas while increasing elsewhere. This is reversed in deaths at ages 60 to 79 with those in more deprived areas facing the least diverse causes of mortality. In this age range diversity in mortality causes is shown to fall at $q = \infty$ in deaths across Scotland. Figure C.1 shows this was driven by trends in more deprived areas with little change in diversity in the median quintile and less deprived areas. Finally, in deaths among women aged 80+ trends are less clear at $q = 1$ and $q= 2$ however, diversity increased most quickly in more deprived areas at $q = \infty$.

Diversity in mortality causes is found to be higher in urban areas across the age ranges shown in Figure 13 in both sexes under diversity at $q = 1$. In men under diversity at $q = 2$ and $q = \square$, this is also observed in deaths among those aged 80+ however, in other age ranges patterns are less clear with similar trends across the populations of urban-rural classification.

In women the trend for higher diversity in urban areas continues under more conservative measures of diversity among deaths in those aged 0 to 39 and 40 to 59. However, in older ages this reverses and women in urban areas tended to face less diverse causes of mortality. For the most part trends in diversity were consistent across subpopulations in women however diversity fell most quickly in urban areas among those aged 0 to 39 and the previously noted reductions in diversity among those aged 60 to 79 were only observed in urban and small town areas.

C.1.2 Scottish Government urban-rural classifications

Diversity in mortality causes is found to be higher in urban areas across the age ranges shown in Figure C.2 in both sexes under diversity at $q = 1$. In men under diversity at $q = 2$ and $q = \infty$, this is also observed in deaths among those aged 80+ however, in other age ranges patterns are less clear with similar trends across the populations of urban-rural classification.

In women the trend for higher diversity in urban areas continues under more conservative measures of diversity among deaths in those aged 0 to 39 and 40 to 59. However, in older ages this reverses and women in urban areas tended to face less diverse causes of mortality. For the most part trends in diversity were consistent across subpopulations in women however diversity fell most quickly in urban areas among those aged 0 to 39 and the previously noted reductions in diversity among those aged 60 to 79 were only observed in urban and small town areas.

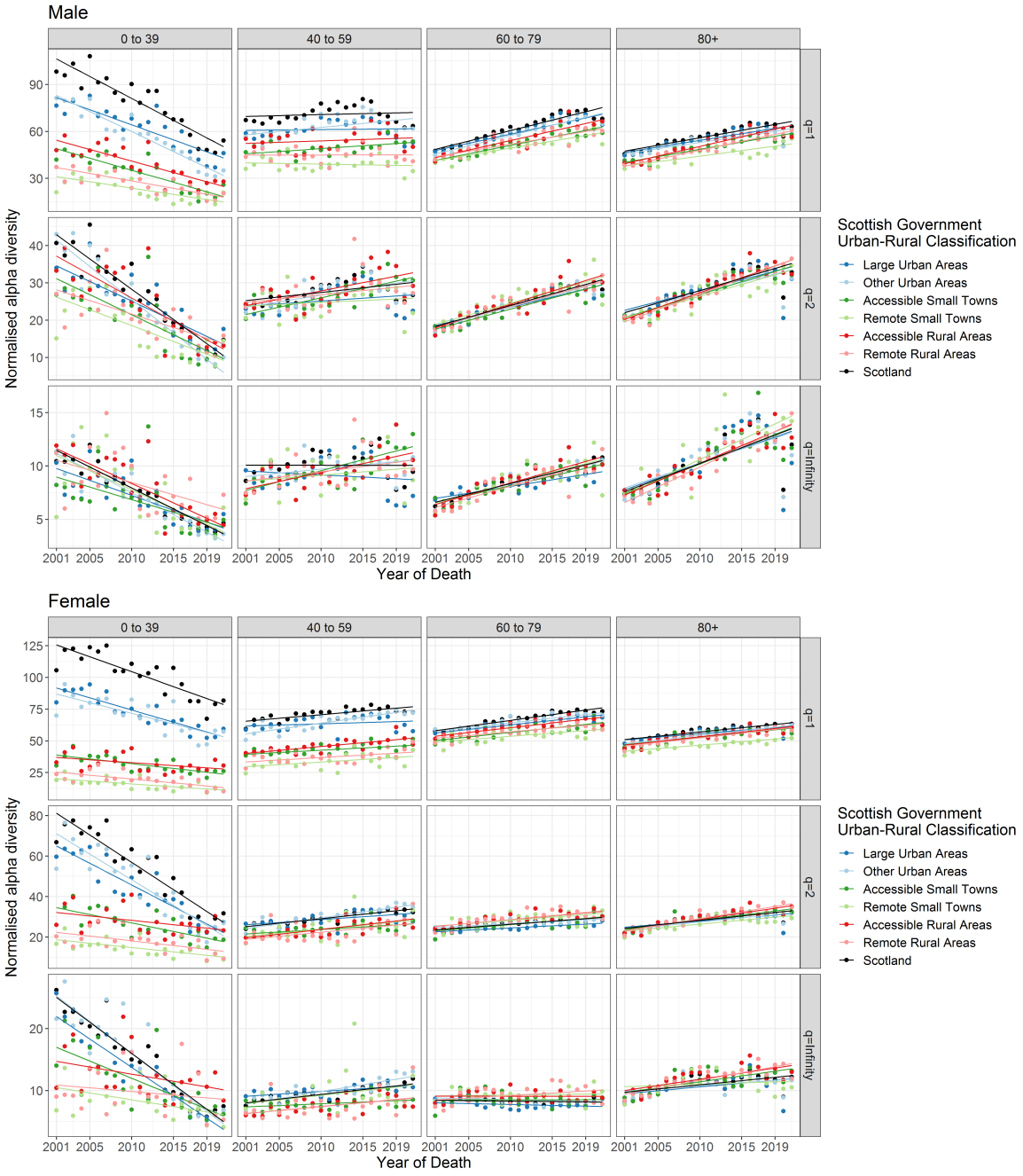


Figure C.2: Trends in diversity in mortality causes from 2001 to 2019 in deaths across twenty-year age ranges in Scottish Government urban-rural classifications.

C.1.3 Trends among the youngest age class

Issues related to the measurement of diversity with small sample sizes are discussed in Section 3.3.3.2. In most cases throughout this thesis sample sizes are considered to be large enough for a reliable measurement of diversity in causes of mortality to be made. However, within subpopulations divided by age at death, as presented within this appendix, some subgroups present with particularly small sample sizes with the potential to bias measurements of diversity. This is particularly a concern among the youngest age class where the fewest deaths occur. Figure XX presents the diversity of causes of mortality in deaths among those aged 0-39 in SIMD income deprivation quintiles and Scottish Government urban-rural classifications as discussed in previous sections with deaths grouped across four-year periods to increase sample sizes. Trends and subpopulation tendencies are observed to differ little from those discussed previously. This indicates that the conclusions drawn in previous sections are robust to the potential biasing effects of the smaller sample sizes in these subpopulations

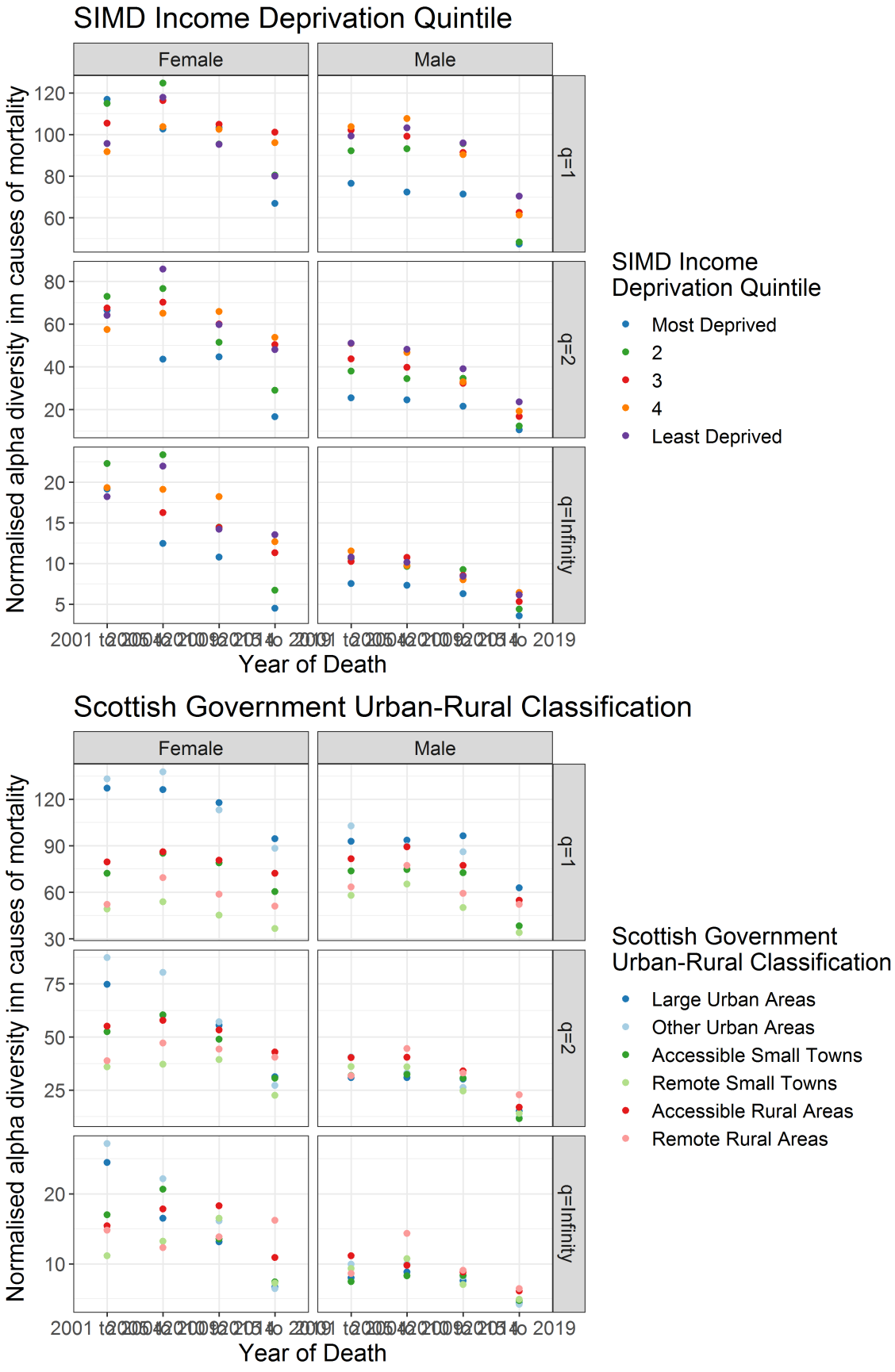


Figure C.3: Trends in diversity in mortality causes from 2001 to 2019 in deaths among those aged 0-39 in SIMD income deprivation quintiles and Scottish Government urban-rural classifications with deaths pooled across four year periods.

C.2 Additive value

C.2.1 SIMD income deprivaiton quintiles

Inequalities in the distribution of cause-specific mortality by income deprivation are clear in Figure C.4. It shows that Chapter XX is more prominent in mortality at younger ages among more deprived areas. Previous research has clearly shown strong deprivation-related effects on mortality in Chapter XX with specifically deaths of despair linked to more deprived areas and this is clearly evident here.

Trends observed in cancers (Chapter II) at ages 40 to 59 in women across Scotland are more prominent in the least deprived areas where all-cause diversity stagnated across the study period. The specific cancers which were most prominent in each quintile differs with breast and ovarian cancers (ICD-10 codes C50 and 57) most common in the least deprived areas and lung cancer (C34) more dominant in more deprived areas.

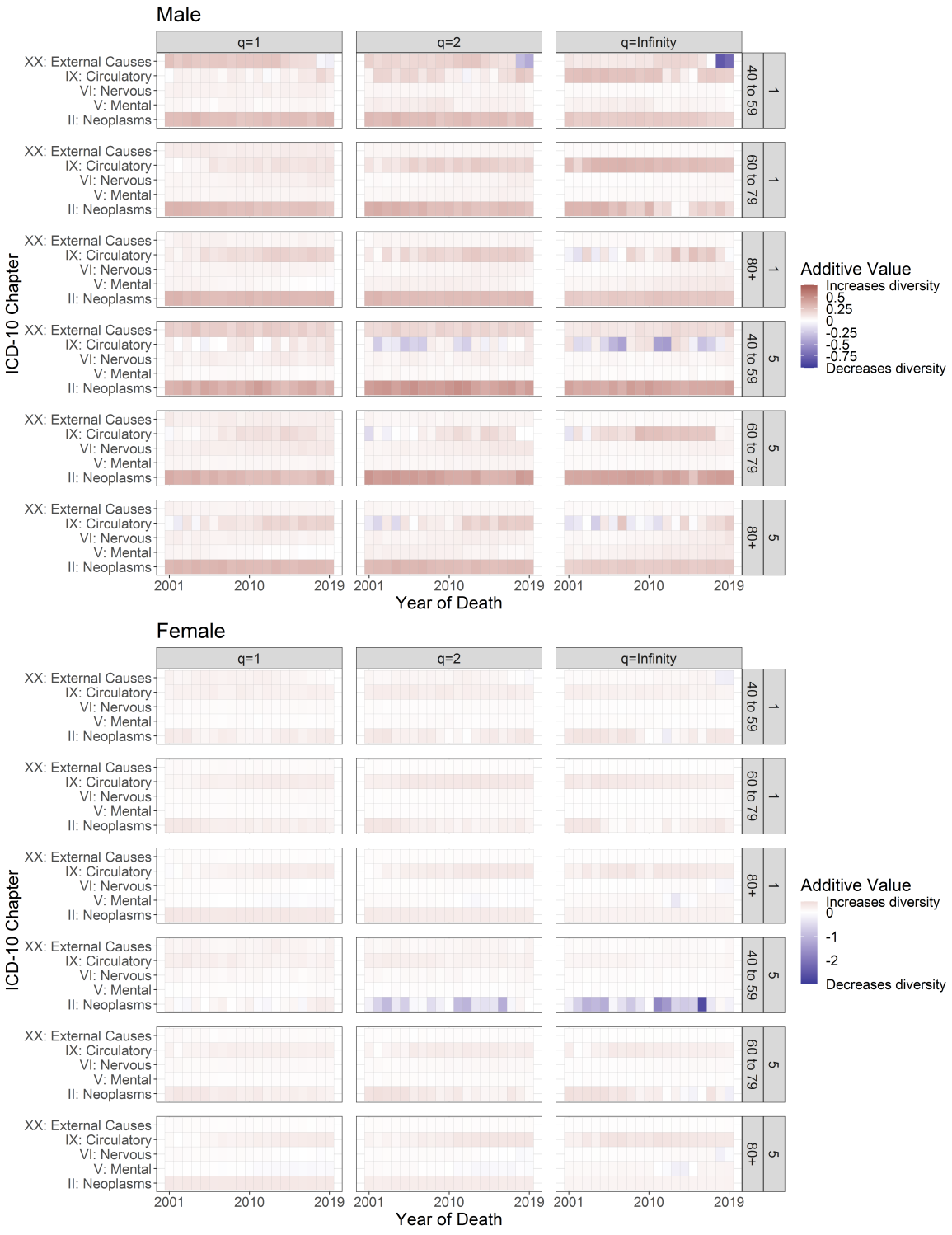


Figure C.4: Additive value to diversity in mortality cause of selected ICD-10 Chapters in deaths among men and women in twenty-year age ranges in the most and least deprived SMD income deprivation quintiles.

Figure C.1 shows less variation between quintiles at older ages and this is evident here with similar profiles of additive value across ICD-10 Chapters among those aged 60 to 79 and 80+.

C.2.2 Scottish Government urban-rural classifications

Figure C.5 shows that in deaths among those aged 0 to 39 Chapter XX: External Causes of Morbidity and Mortality was much more dominant in the distribution of mortality causes in urban areas than in small towns or rural areas. The increasing dominance of causes within this chapter likely explains reductions in diversity at q = ∞ occurring more rapidly in urban areas as for the most part profiles of additive value are observed to be similar across subpopulations.

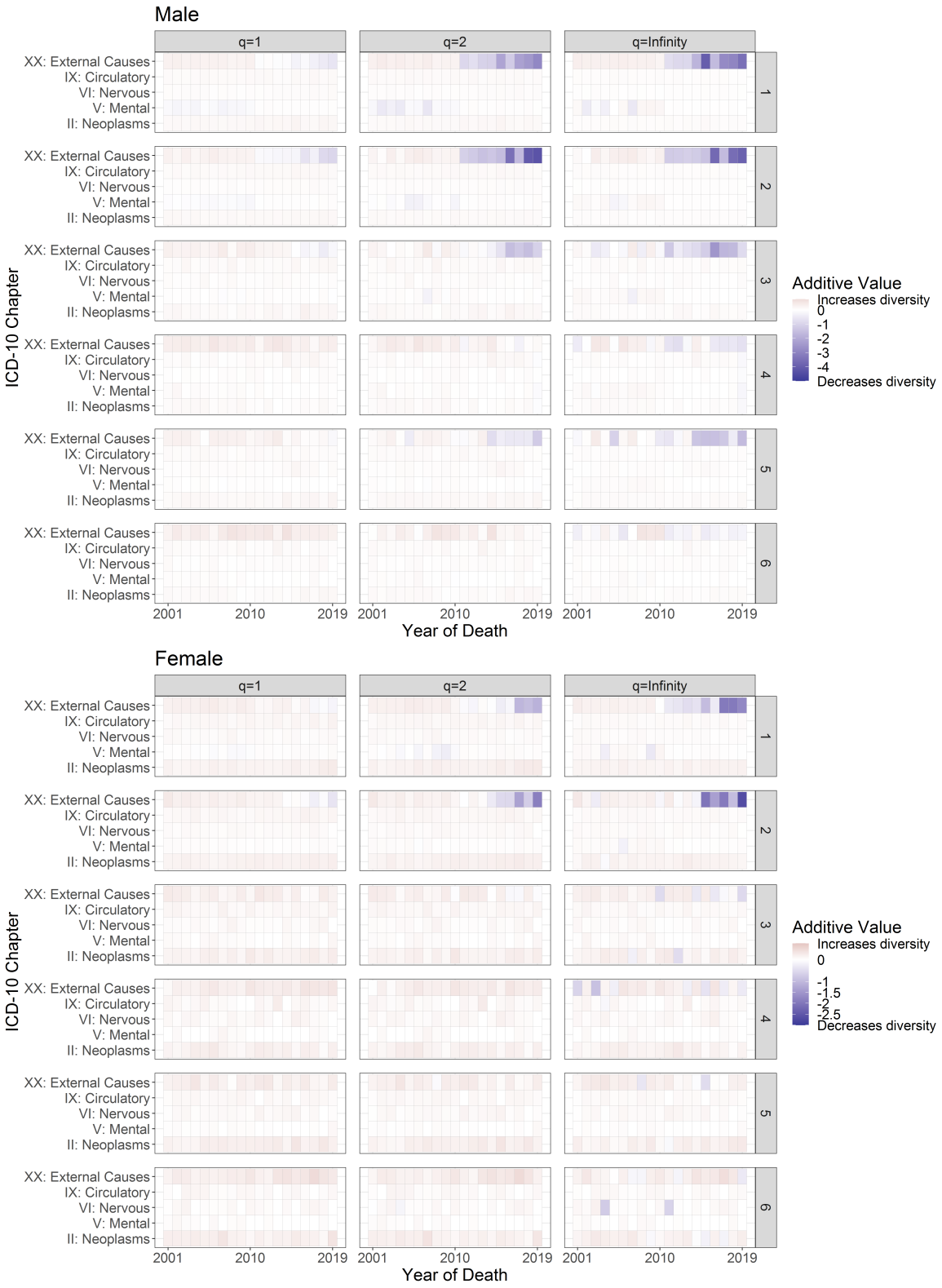


Figure C.5: Additive value to diversity in mortality cause of selected ICD-10 Chapters in deaths among males and females aged 0 to 39 in the population of Scottish Government urban-rural classifications.

C.2. Additive value

The additive value of Chapter II: Neoplasms in women aged 40 to 59 is found to be higher in more rural areas as shown in Figure C.6. Similarly, the additive value of Chapter IX: Diseases of the circulatory system is found to be higher in these areas among men in this age range. These trends are also noted among the populations of the least deprived areas in Scotland.

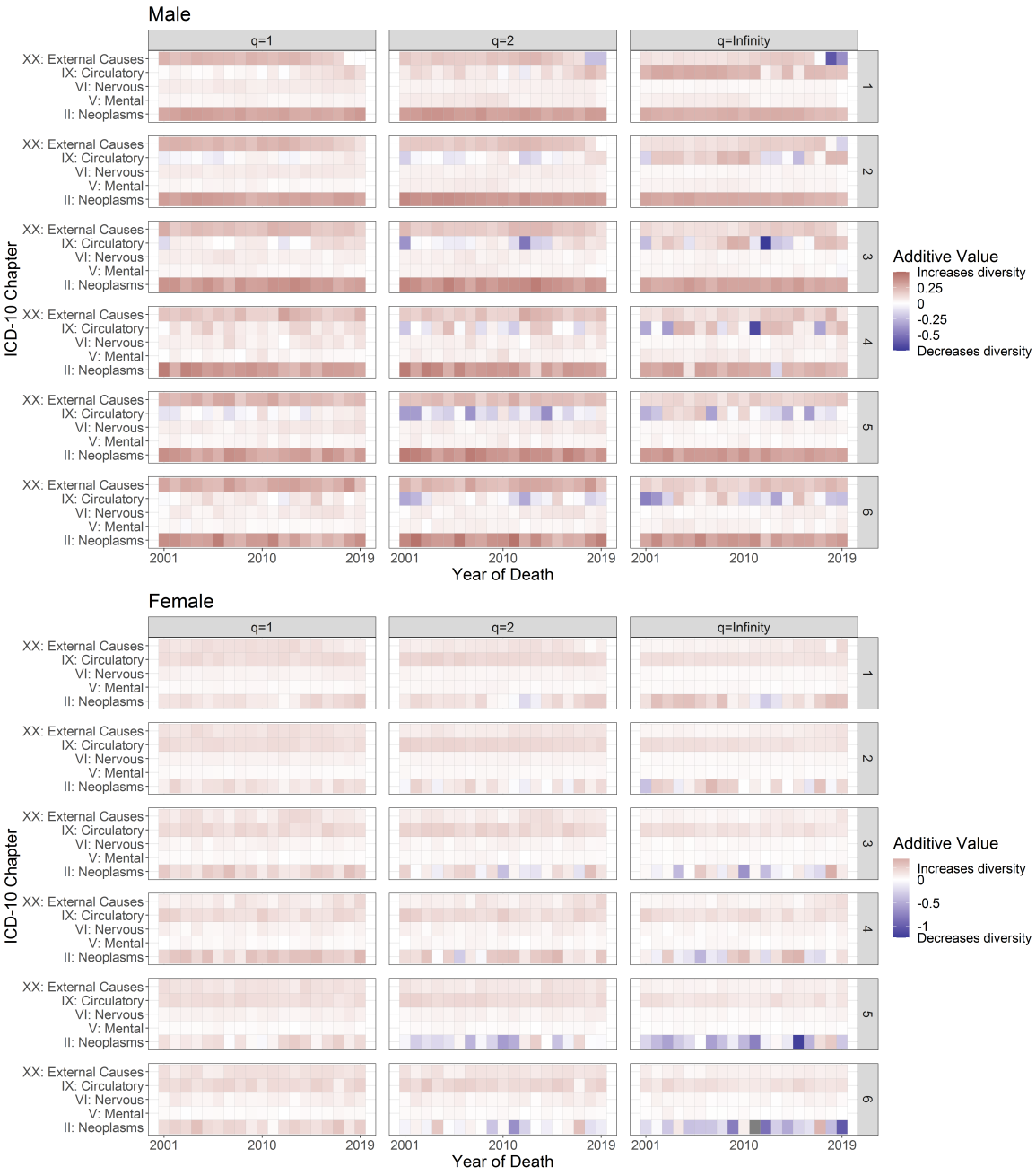


Figure C.6: Additive value to diversity in mortality cause of selected ICD-10 Chapters in deaths among males and females aged 40 to 59 in the population of Scottish Government urban-rural classifications.

C.2. Additive value

The trends in additive value among older age groups are, like in SIMD income deprivation quintiles, similar across urban-rural classifications reflecting the small differences in diversity observed in these ages.

Appendix D

The relationship between lifespan diversity and life expectancy

Figure D.1 shows that lifespan diversity was negatively correlated with life expectancy in Scotland and in most subpopulations over the years 2001 to 2019. Meaning that as the population died at an older age, the variation in age at death was smaller. This is indicative of greater equality in lifespan within populations as individuals lived longer. This association is fairly consistent across both SIMD income deprivation quintiles (upper panels of Figure D.1) and urban-rural classifications (lower panels of Figure D.1).

The trends of increasing or stagnating lifespan diversity noted in the years following 2010 occurred alongside faltering improvements to life expectancy across Scotland as well as within a number of the subpopulations examined here. This indicates that reductions in within-group inequality in lifespan slowed at the same time as increasing average age at death in these groups.

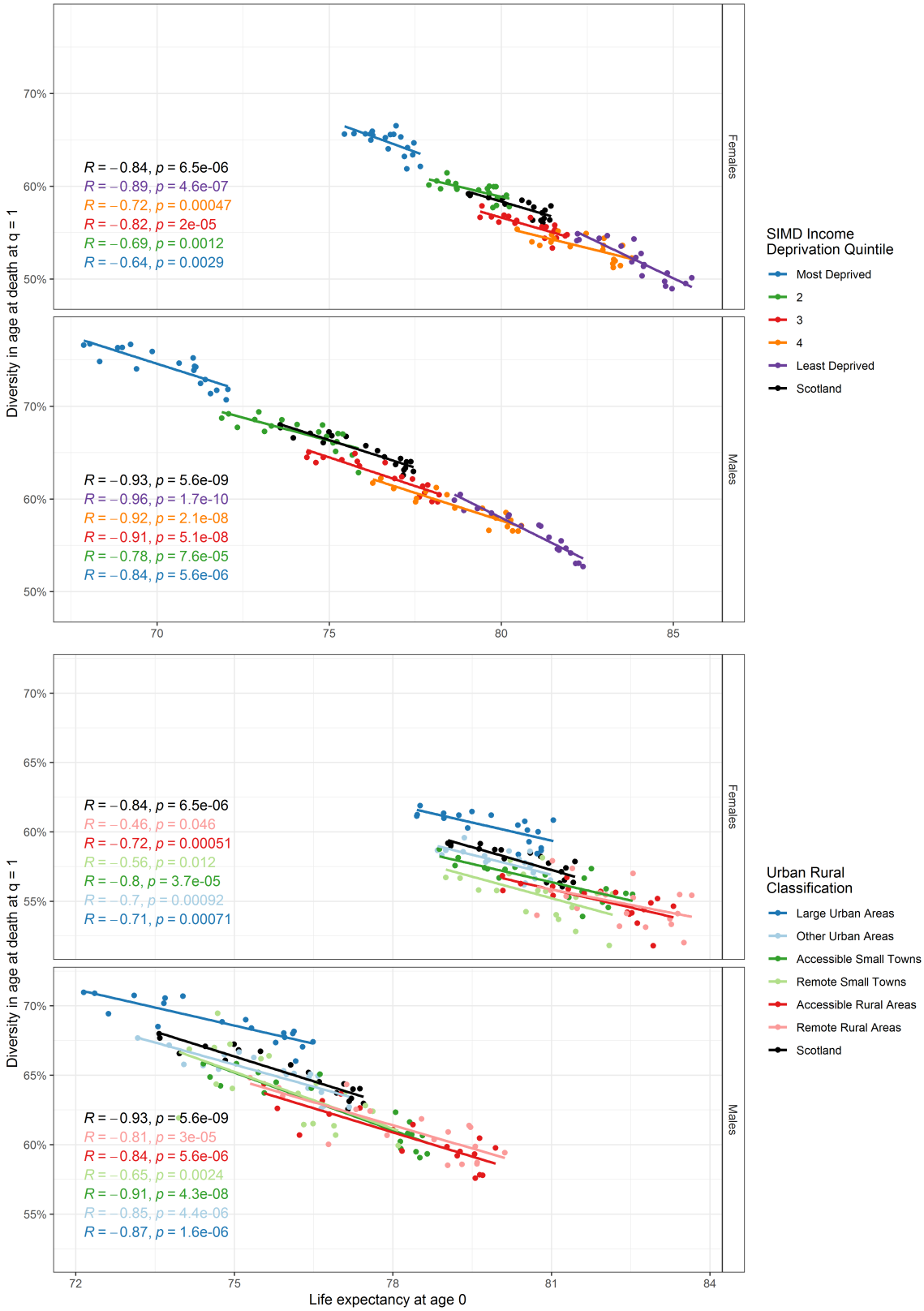


Figure D.1: The relationship between normalised alpha lifespan diversity at q=1 and life expectancy at age 0 in Scotland in the years 2001 to 2019 in males and females, across the country as a whole and in SIMD income deprivation quintiles and Scottish Government urban/rural classifications. Pearson's correlation coefficient alongside p-values are presented, colour coded, for each relationship.

Alongside trends over time, a general tendency for reduced lifespan diversity can be observed among subpopulations with higher life expectancies. This indicates that in areas where the population lived longer there was less variation in ages at death. This is especially clear between SIMD income deprivation quintiles between which greater differences in both measures are observed than between urban-rural subpopulations.

Appendix E

Future research case study methods and results

E.1 Similarity-sensitive diversity in causes of mortality: Case study methods

Mortality and population data were used in the calculation of is life tables for the male and female population of Scotland in the years 2001 to 2019 as described in Chapter 4 of this thesis. Naïve-similarity normalised alpha diversity in causes of mortality was calculated using mortality cause distributions extracted from these life tables as described in Chapter 4.

The formation of similarity matrices for the calculation of diversity is explained in detail in Chapter 5. In the case of ages at death a formula for similarity between pairs of ages based on the distance between these ages is used. This is possible because ages at death present a continuous set of types where differences between types are mathematically relevant (i.e. a difference of 1 year in age can be treated as the same between ages 0 and age 1 as between age 99 and age 100). In the case of causes of mortality a formulaic approach such as this is not applicable. Instead in this case study similarity is based on the structure of ICD-10 classification. Two layers of stratification are used by the ICD-10 system above three-character codes; these are: Blocks which are groups of three-character ICD-10 codes describing similar diseases or conditions (e.g. C00-C14 Malignant neoplasms of lip, oral cavity

and pharynx, which groups codes classifying neoplasms in the mouth and throat); and Chapters, the highest level of categorisation in the ICD-10 structure grouping diseases and conditions by disease type or body system or example Chapter II: Neoplasms or Chapter IX: Diseases of the Circulatory System.

A matrix was constructed with rows and columns representing each ICD-10 three-character code present in the Scottish mortality data used in this study. As a default, causes are set to have 0 similarity to each other, without further adaptation this represents naïve similarity with all causes treated as distinct. All causes within the same Chapter are set to have a similarity of C to each other and all causes within the same Block are set to have a similarity of B to each other. The process of selecting values for C and B to take is one the key areas which requires further research. Here, in a similar process to that described in Chapter 5, values were chosen by calculating the similarity-sensitive metacommunity gamma diversity at q=1 ($({}^1G^Z)$) of the distribution of mortality causes in across the years 2001-2019 in both sexes combined. This measure calculates the effective number of causes (or groups of causes) in the distribution. In this case B was set to be half of C in all cases indicating that codes in the same block are twice as similar to each other as codes in the same chapter. With this constraint, similarity matrices were created by gradually reducing C from a value of 1 and calculating $({}^1G^Z)$ under each matrix with the aim of producing a value of 19 under this measure. This number was chosen to align with the proposed number of groups of causes used by Bergeron-Boucher et al. (2020). This resulted in a value for C of 0.4695 (and therefore B was (0.23475).

Using the similarity matrix described above similarity-sensitive normalised alpha diversity in causes of mortality was calculated for deaths among males and females in Scotland in each year from 2001-2019.

E.1.0.0.1 Similarity-sensitive diversity in causes of mortality: Case study results In Figure E.1 trends from 2001 to 2019 in the similarity-sensitive diversity in causes of mortality among males and females in Scotland are shown. Within each sex no notable exceptions are observed to trends in naïve-similarity diversity as observed previously in this thesis. Differences in diversity between sexes are smaller under similarity-sensitive measures than in naïve-similarity measures at q = 1 and q = 2.

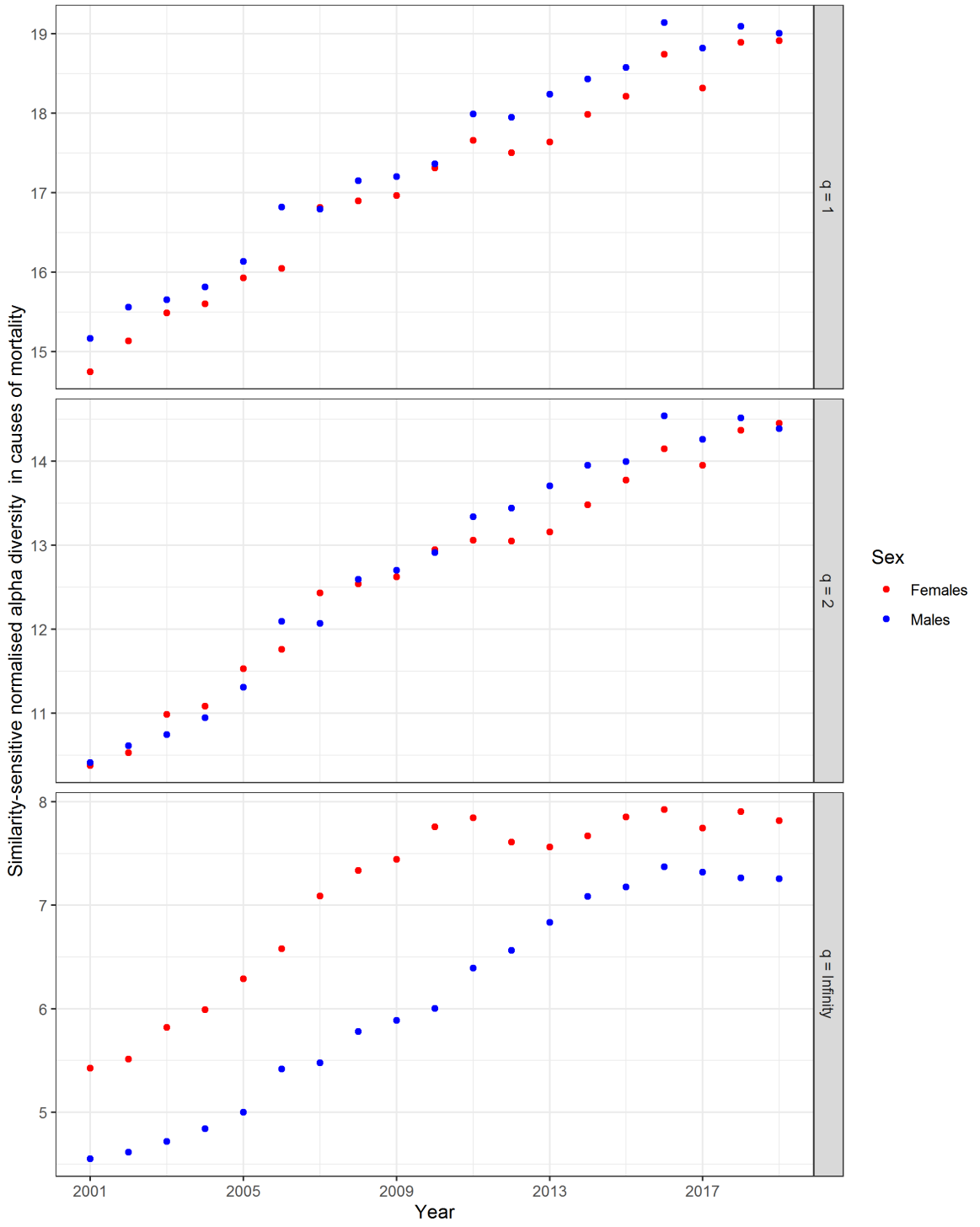


Figure E.1: Similarity-sensitive normalised alpha diversity at $q = 1$, $q = 2$, and $q = \infty$ in causes of mortality in deaths among males and females separately in Scotland in each year 2001 to 2019.

Similarity-sensitive diversity is compared to naïve-similarity diversity in Figure E.2. Strong correlation between measures can be observed in both sexes under diversity at q = 1 and q = 2. Under diversity at q = ∞, slight deviation can be observed. These datapoints represent the late years of the 2010s when reductions in naïve similarity diversity at this measure were more pronounced than under similarity-sensitive diversity.

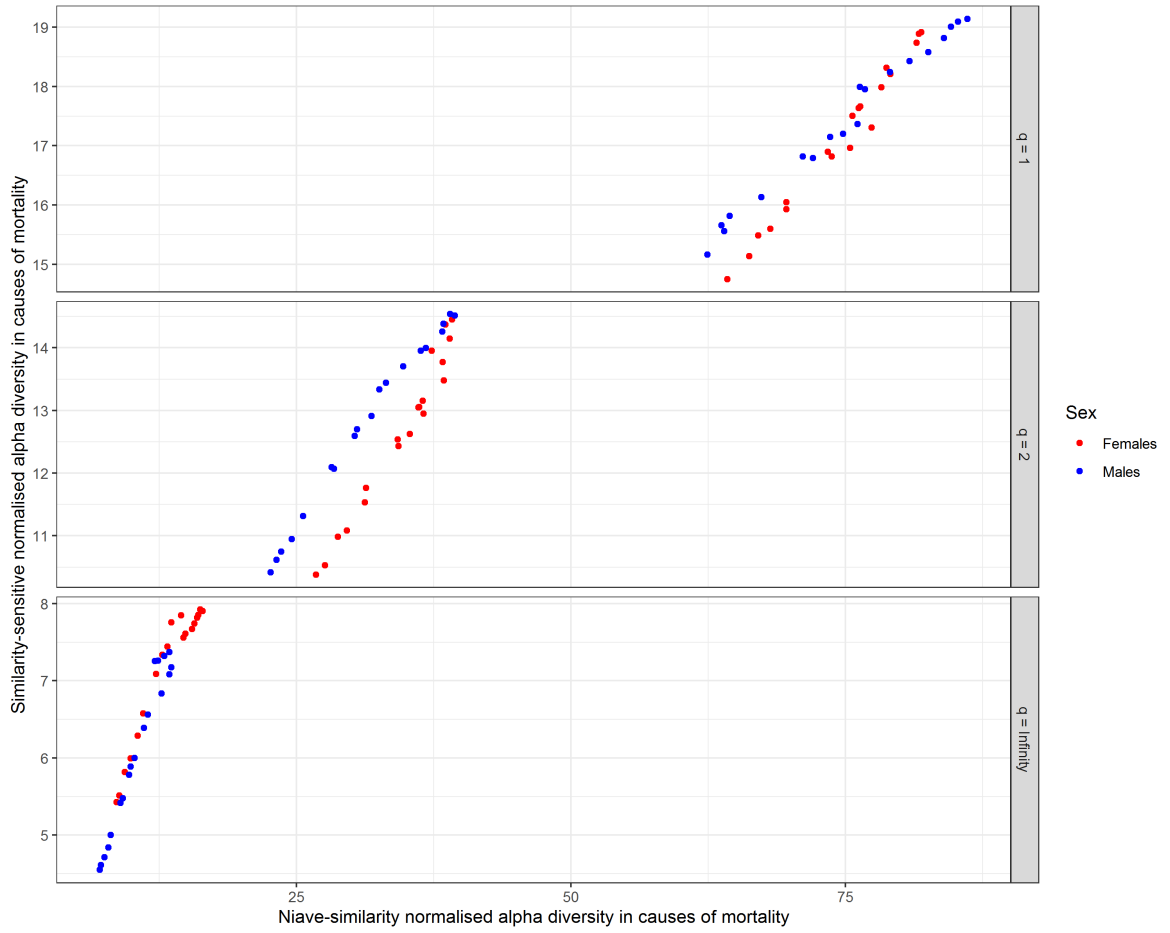


Figure E.2: Similarity-sensitive diversity in causes of mortality plotted against naïve-similarity diversity in causes of mortality at values of q equal to 1, 2, and ∞ for deaths among males and females in Scotland from 2001 to 2019.

E.2 Diversity in the study of population health in small populations: case study methods

Mortality data as described in Section 3.2.1 for the years 2001 to 2005 was used for this case study. A census postcode sector was assigned each mortality record based on the postcode associated with the record. Census postcode sectors are administrative geographies created for the dissemination of census data. The deaths within each postcode sector in Scotland in the years 2001 to 2005 were combined for each sex separately and normalised alpha diversity (as described previously in this thesis) at q = 1 was calculated in the distribution of mortality causes in each postcode sector for males and females separately.

The 2001 Carstairs score, a measure of area-level deprivation, for each postcode sector was obtained from Public Health Scotland publicly available data. This area level deprivation data was used indicatively to demonstrate the methods described in this section. Carstairs scores are standardised with high positive values indicating more deprived areas and more negative values indicating less deprived areas (Brown et al., 2014).

The algorithm described in the previous section was applied to the mortality records within each postcode sector for males and females separately. The distribution of mortality causes within each postcode sector was resampled and normalised alpha diversity at q = 1 was calculated for men and women separately in each postcode sector. This was repeated 1000 times and the mean of the values of diversity produced was extracted.

E.2.0.0.1 Diversity in the study of population health in small populations: Case study results The normalised alpha diversity at $q = 1$ of the causes of mortality recorded in the years 2001 to 2005 in each Scottish census postcode sector is plotted against the mean resampled diversity value for males and females separately in Figure E.3. A diagonal line is plotted, points which fall on this line represent postcode sectors where diversity in mortality causes was exactly what would be expected given the population size and the ages of those who died. Points to the left of the diagonal line indicate postcode sectors where causes of mortality were less diverse that would be expect and those on the right represent areas where causes of mortality were more diverse than would be expected.

Points are coloured by 2001 Carstairs score, a measure of deprivation, as an example of how these methods might be utilised. A possible tendency for areas with greater deprivation to have less diverse mortality causes than would be expected can be observed.

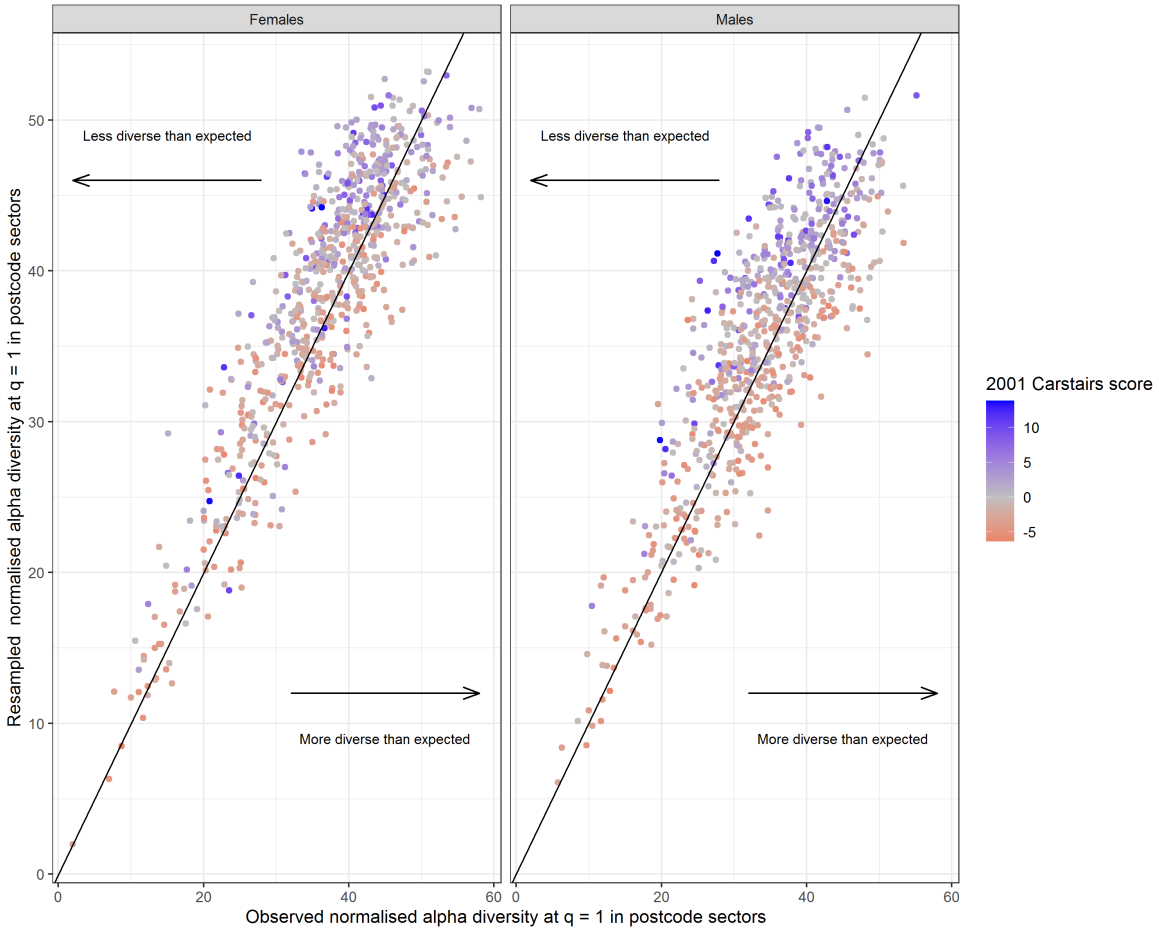


Figure E.3: Normalised alpha diversity at $q = 1$ in causes of mortality calculated in deaths recorded within individual census postcode sectors in Scotland in the years 2001 to 2005 plotted against the resampled diversity of each census postcode sector. Points are coloured by 2001 Carstairs Deprivation score, blue positive values indicate more deprived areas with red points indicating less deprived areas.