Modelling mortality heterogeneity using health trajectories and multimorbidity

Michelle Kundai Vhudzijena

A thesis in fulfilment of the requirements for the degree of

Doctor of Philosophy

School of Risk and Actuarial Studies

UNSW Business School

October 2023

PLEASE TYPE

The University of New South Wales

Thesis/Dissertation Sheet

Surname or Family name**: Vhudzijena**	
First name: **Michelle Kundai**	Other name/s:
Abbreviation for degree as given in the University calendar: **PhD**	
School: **School of Risk and Actuarial Studies**	Faculty: **UNSW Business School**
Title: **Modelling mortality heterogeneity using health trajectories and multimorbidity**	

Abstract 350 words maximum: (PLEASE TYPE)

Mortality heterogeneity is a generally well understood area of longevity risk that remains relatively unexplored in the actuarial pricing of longevity linked products. However, with increasing amounts of longitudinal individual level data, there exists an extraordinary opportunity to derive more nuanced and realistic mortality risk profiles that can improve the design and demand of annuities and other longevity linked products.

Deriving mortality risk profiles using the clustering of health trajectories and unsupervised machine learning algorithms is seldomly investigated in the literature. This thesis applies a three dimensional k–means clustering algorithm to joint trajectories of self reported health and body mass index to develop mortality risk profiles. We are able to determine distinct mortality risk profiles from the clusters that exhibit significant differences in life expectancy and annuity prices for both males and females at varying ages.

Disregarding health status in longevity linked products has been shown to cause adverse selection from individuals with chronic conditions due to inaccurate pricing of mortality and morbidity risks. However, we are unaware of work in the actuarial literature that shows the impact of risk factors on health status. Therefore, the second project explores the effectiveness of utilising hidden markov models with covariates to demonstrate mortality heterogeneity. We find that the clusters generated by the hidden markov models have a better fit to empirical data than models without clustering.

It is important to address the link between multimorbidity and the pricing of health and longevity linked products in the actuarial literature. The last project in this thesis seeks to find the best way to incorporate multimorbidity in the pricing of long term care products. We compare two different ways of incorporating multimorbidity in multiple state models. We find that our proposed five state multimorbidity and functional disability model is able to capture the dynamics of health over time more accurately than the three state health and functional disability model with a multimorbidity predictor. The results from the later model weakly suggest morbidity expansion when in effect there is very strong evidence of morbidity expansion. This inadvertently leads to the gross mispricing of life annuities, long term care and lifecare annuities.

Declaration relating to disposition of project thesis/dissertation

I hereby grant to the University of New South Wales or its agents the right to archive and to make available my thesis or dissertation in whole or in part in the University libraries in all forms of media, now or here after known, subject to the provisions of the Copyright Act 1968. I retain all property rights, such as patent rights. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

I also authorise University Microfilms to use the 350 word abstract of my thesis in Dissertation Abstracts International (this is applicable to doctoral theses only).

……

Signature

……..…….

Witness

……..……..……

Date

The University recognises that there may be exceptional circumstances requiring restrictions on copying or conditions on use. Requests for restriction for a period of up to 2 years must be made in writing. Requests for a longer period of restriction may be considered in exceptional circumstances and require the approval of the Dean of Graduate Research.

FOR OFFICE USE ONLY

Date of completion of requirements for Award:

Declarations

ORIGINALITY STATEMENT

☑ I hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at UNSW or any other educational institution, except where due acknowledgement is made in the thesis. Any contribution made to the research by others, with whom I have worked at UNSW or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.

COPYRIGHT STATEMENT

☑ I hereby grant the University of New South Wales or its agents a non-exclusive licence to archive and to make available (including to members of the public) my thesis or dissertation in whole or part in the University libraries in all forms of media, now or here after known. I acknowledge that I retain all intellectual property rights which subsist in my thesis or dissertation, such as copyright and patent rights, subject to applicable law. I also retain the right to use all or part of my thesis or dissertation in future works (such as articles or books).

For any substantial portions of copyright material used in this thesis, written permission for use has been obtained, or the copyright material is removed from the final public version of the thesis.

AUTHENTICITY STATEMENT

☑ I certify that the Library deposit digital copy is a direct equivalent of the final officially approved version of my thesis.

Inclusion of Publications Statement

UNSW is supportive of candidates publishing their research results during their candidature as detailed in the UNSW Thesis Examination Procedure.

Publications can be used in the candidate's thesis in lieu of a Chapter provided:

- The candidate contributed **greater than 50%** of the content in the publication and are the "primary author", i.e. they were responsible primarily for the planning, execution and preparation of the work for publication.
- The candidate has obtained approval to include the publication in their thesis in lieu of a Chapter from their Supervisor and Postgraduate Coordinator.
- The publication is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in the thesis.

☑ The candidate has declared that **their thesis contains no publications, either published or submitted for publication.**

Candidate's Declaration

I declare that I have complied with the Thesis Examination Procedure.

Acknowledgements

Undertaking this PhD has been one of the most challenging things I have ever done to date. It was particularly demanding given that I started this journey with an eight month old who did not believe in sleeping before midnight. It has continuously tested my intelligence, patience and resilience. However, all this has been worthwhile since I have greatly improved my research skills and hopefully advanced knowledge in the management of longevity risk – a very underresearched area. Needless to say, I am very encouraged by my own tenacity and resilience to the end.

I would like to thank my supervisors Professor Michael Sherris, Associate Professor Andrés Villegas and Associate Professor Jonathan Ziveyi. From helping me narrow down a very broad research agenda to accommodating numerous missed deadlines. I presented at multiple prestigious conferences that they helped me identify and they always encouraged me throughout the process. I could not have been led by a better team.

I am also grateful to the School of Risk and Actuarial Studies. The research class run by Professor Hazel Bateman was very instructive and provided constructive feedback on my work. I received valuable commentary from fellow students and academic staff for all three projects. I would also like thank the school manager, Maree Withers and all the administrative team for their assistance with funding and administration.

In addition, I thank members of my PhD review panel: Associate Professor Anthony Asher, Associate Professor Katja Hanewald, Associate Professor Yang Shen and Associate Professor Jae Kyung Woo. They have consistently provided me with great advice and offered excellent support during this four year journey.

I am very grateful for my examiners, Professor Jing Ai and Professor Vladimir Canudas–Romo. Their reports demonstrate their enormous attention to detail and provide me with tools to improve my problem formulation and the communication of my research ideas. I

appreciate all the time they took to review my thesis and their encouragement to work on converting my chapters into publications.

I also acknowledge the financial support I received from the Australian Research Council Centre of Excellence in Population Ageing Research (CEPAR) and the University of New South Wales. The Scientia PhD scholarship covered my tuition and provided a living stipend during the course of my PhD. The career development support offered by Scientia is unparalleled and helped me create a robust career plan.

My friends, Doreen Kabuche and Salvatory Kessy offered extensive support with debugging R, Stata and Matlab. We commiserated with each other over not getting enough time on the computational cluster Katana and were intrigued by the sheer amount of work we were doing.

I am also hugely indebted to my mother, Margaret Vhudzijena, who constantly encouraged me to keep working after putting my daughter Alexandra to bed, no matter the time of day. I would not have completed my thesis without my dearest husband Adolf Gasseler. He spent so many hours helping me rehearse my presentations and driving me to campus so that I could write without distractions. Lastly, I thank God and all my ancestors for making me who I am.

Abstract

Mortality heterogeneity is a generally well understood area of longevity risk that remains relatively unexplored in the actuarial pricing of longevity linked products. However, with increasing amounts of longitudinal individual level data, there exists an extraordinary opportunity to derive more nuanced and realistic mortality risk profiles that can improve the design and demand of annuities and other longevity linked products.

Deriving mortality risk profiles using the clustering of health trajectories and unsupervised machine learning algorithms is seldomly investigated in the literature. The first project in this thesis applies a three dimensional k–means clustering algorithm to joint trajectories of self reported health and body mass index to develop mortality risk profiles. We are able to determine distinct mortality risk profiles from the clusters that exhibit significant differences in life expectancy and annuity prices for both males and females at varying ages.

Disregarding health status in longevity linked products has been shown to cause adverse selection from individuals with chronic conditions due to inaccurate pricing of mortality and morbidity risks. However, we are unaware of work in the actuarial literature that shows the impact of risk factors on health status. Therefore, the second project explores the effectiveness of utilising hidden markov models with covariates to demonstrate mortality heterogeneity. We find that the clusters generated by the hidden Markov models have a better fit to empirical data than models without clustering.

It is important to address the link between multimorbidity and the pricing of health and longevity linked products in the actuarial literature. The last project in this thesis seeks to find the best way to incorporate multimorbidity in the pricing of long term care products. We compare two different ways of incorporating multimorbidity in multiple state models. We find that our proposed five state multimorbidity and functional disability model is

able to capture the dynamics of health over time more accurately than the three state health and functional disability model with a multimorbidity predictor. The results from the later model weakly suggest morbidity expansion when in effect there is very strong evidence of morbidity expansion. This inadvertently leads to the gross mispricing of life annuities, longterm care and lifecare annuities.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ADLs	Activities of Daily Living
ANOVA	Analysis of Variance
APC	Age Period Cohort
ARIMA	Autoregressive Integrated Moving Average
BIC	Bayesian Information Criterion
BMI	Body Mass Index
CAE	Common Age Effect
CBD	Cairns–Blake–Dowd
CEPAR	Center of Excellence in Population Ageing Research
CI	Confidence Interval
CIRS	Cumulative Illness Rating Scale
IADLs	Instrumental Activities of Daily Living
ICED	Index of Coexisting Disease
CVD	Cardiovascular Disease
GLMs	Generalised Linear Models
GLMM	Generalised Linear Mixed Models
GDP	Gross Domestic Product
GP	Gaussian–Process
HIV	Human Immunodeficiency Virus
HLE	Healthy Life Expectancy
HMM	Hidden Markov Models
HR	Hazard Ratio
HRS	Health and Retirement Study
HSD	Honest Significant Difference
ICD	International Classification of Diseases

IMD	Index of Multiple Deprivation
ICD–10	International Classification of Diseases 10^{th} Revision
ICD–11	International Classification of Diseases 11^{th} Revision
KL	Kullback–Leibler
LTSM	Long Short-Term Memory
OECD	Organisation for Economic Co–operation and Development
OR	Odds ratio
NLPCA	Non–Linear Principal Component Analysis
QIC	Quasilikelihood Information Criterion
ReLU	Rectified Linear Unit
SII	Slope Index of Inequality
SOA	Society of Actuaries
SNPs	Single Nucleotide Polymorphisms
TLE	Total Life Expectancy
US	United States
USA	United States of America
UK	United Kingdom
UNSW	University of New South Wales
VAR	Vector Autoregression
VECMs	Vector Error Correction Models
WHO	World Health Organisation

Chapter 1

Introduction

1.1 Motivation

Total pension assets for the 22 largest economies stood at US$47.861 trillion in December 2022 with an average ratio of pension assets to gross domestic product of 62% (Hall et al., 2023). Despite the sheer scale of the pensions industry, longevity risk remains poorly managed globally. One of the main reasons for this phenomenon is that markets for annuities which provide regular and constant income during retirement are thin globally except in countries where there is mandatory annuitisation (Brugiavini, 1993; Cannon & Tonks, 2008). There are a variety of reasons that have been suggested to understand why there is poor uptake of annuity products from both demand and supply side perspectives. These include mental accounting, behavioural biases, environmental limitations and non–contributory age pensions that mimic annuities (MacDonald et al., 2013; O'Meara et al., 2015; Thaler, 1999). Of interest to this thesis is that the pricing of annuities is sub–optimal and this negatively affects demand (J. R. Brown, 2009; Cappelletti et al., 2013; Steinorth, 2012). This thesis extends the research on how to make annuities fairer so that they meet the needs of most retirees instead of a select few.

To improve product design, more attention has to be given to underwritten annuities simply because they take into account an individual's circumstances rather than what is deemed to be the average individual's mortality experience from aggregate population level dynamics. The rate for standard annuity products is mainly determined by the mortality assumption which is an estimate of an individual's probability of death. When

actuaries estimate this probability, they usually use mortality tables that rely on historical crude death rates that have been smoothed through various standard graduation techniques (Booth & Tickle, 2008; Tabeau et al., 2002). At the end of the day, standard annuities are based on "average" estimates of an individual's probability to die in a given year based on their gender. Combined with adverse selection from individuals perceived to be of higher risk because of higher longevity, annuities become more expensive than they should be and this results in an inefficient market (Meyricke & Sherris, 2013).

If we consider newer, more robust and popular extrapolative methods in mortality modelling and forecasting such as Lee–Carter and Cairns-Blake–Dowd and their variants, we are still unaware of what really happens at the individual level (Cairns et al., 2006, 2009; Lee & Carter, 1992). For example, the original Lee & Carter (1992) forecasts mortality rates by applying singular value decomposition to a matrix of age–specific mortality rates that is centered by subtracting the mean logarithm at a specific age and ultimately derives β_x and κ_t

$$\log(\mu_{x,t}) = \alpha_x + \beta_x \kappa_t, \tag{1.1}$$

where $\log(\mu_{x,t})$ is the log force of mortality at age x and calendar year t, β_x is a vector that measures how the log change of mortality varies with time at age x and κ_t is the time index in calendar year t. As reflected in Equation (1.1), this is clearly not a structural model as there are no covariates and it ultimately does not lead to a good understanding of the underlying factors for mortality and/or its associated rates. Resultantly, many researchers find the continued use of extrapolative methods that do not appreciate the laws of cause and effect troubling (Gutterman & Vanderhoof, 1998).

However, it is not without reason that extrapolative methods remain dominant in mortality modelling and forecasting. The main reason is that an explicit relationship between mortality and its determinants is a complex relationship to model because the variables are interconnected (Tuljapurkar & Boe, 1998). Despite the lack of complete understanding of how mortality is connected with its determinants, explanatory methods have been used to forecast mortality rates and have performed better than extrapolative methods in some cases (Girosi & King, 2008; Murray & Lopez, 1997c, 1997b; Murray & Lopez, 1997a). This success can be attributed to attempts to directly link mortality to its determinants.

This inadvertently motivates significant literature on incorporating socio–demographic information into mortality models and can potentially increase the demand for longevity linked products by using the mortality heterogeneity to design better products that match the characteristics of the different risk profiles (Alai et al., 2018; R. L. Brown & McDaid, 2003; Cairns et al., 2019; Madrigal et al., 2011; Meyricke & Sherris, 2013; Sherris & Zhou, 2014; Su & Sherris, 2012; Tuljapurkar & Boe, 1998; Wen et al., 2021; M. Xu et al., 2019). Therefore in this thesis, we extend this literature by examining overlooked concepts such as health trajectories and multimorbidity which are sources of mortality heterogeneity. The health trajectories and multimorbidity status are derived from individual level data. Identifying different risk profiles in a heterogeneous population allows us to better quantify mortality or morbidity risks particularly when we integrate machine learning and longitudinal individual–level data.

Recently, there has been huge progress in the availability of data that allow more individualised product design and better computational capacity to analyse data and make more informed decisions. Some of these data are available at population level through more integrated data collection from government departments. Machine learning can be very useful in analysing huge datasets mainly because it provides easily implementable data driven solutions that can find patterns in data that can take decades of research using standard statistical techniques (Goodfellow et al., 2016). There are two classes of machine learning methods: supervised learning and unsupervised learning (Hastie et al., 2009; James et al., 2013). The output is known in supervised learning and this guides the learning process, whereas there is no output in unsupervised learning. The algorithms try to find patterns in the data. Longitudinal data, that is, data with repeated measurements over time provide researchers with the opportunity to analyse within–individual change compared to between–individual changes from cross–sectional data (Diggle, 2002; Fitzmaurice et al., 2011). Combining insights from both longitudinal data and machine learning has great potential to produce a better understanding of the mortality of retirees and thus should lead to better designed products for retirees.

We are unaware of the combined use of trajectories of the determinants of mortality and clustering techniques in the literature as a means of determining mortality risk profiles. Hence, our first objective in Chapter 3 is to consider the use of a k–means clustering technique on longitudinal individual level data to determine mortality risk profiles. Since

we would like to identify and characterise these risk profiles, we ask what clusters emerge when we jointly model body mass index (BMI) and self reported health trajectories of older adults? Once we have established these clusters, we then ask the following secondary research questions. What is the relationship between mortality and cluster allocation? How does cluster allocation impact mortality amongst older adults when controlling for socio–economic variables? What are the pricing implications for annuities in each cluster?

We follow up the previous study with an inquiry on how to link health status with trajectories of various predictors of mortality and morbidity in the actuarial pricing of health and longevity linked products in Chapter 4. While health status has been shown to impact mortality and the pricing of longevity linked products, there is very little evidence of linking health status with covariates such as BMI, smoking status or income. Hidden Markov models (HMMs) are known to be capable of modelling unobservable states using a known output (Ghassempour et al., 2014; Rabiner & Juang, 1986). Therefore, the second objective of this thesis is to determine the effectiveness of using Hidden Markov models (HMMs) with covariates to demonstrate mortality heterogeneity. Our specific research questions are as follows. To what extent do the clusters developed from the multivariate-time series clustering of health trajectories using HMMs with covariates, provide well developed risk profiles that exhibit mortality heterogeneity? Does clustering provide a better fit to empirical data when estimating transition rates and life expectancy in a multi–state model of health status and functional disability while controlling for age and gender?

Lastly, despite the fact that multimorbidity is known to have an impact on mortality and quality of life particularly amongst people of older ages, it is usually omitted in the actuarial pricing of health and longevity linked products. In Chapter 5 we compare two main strategies of incorporating multimorbidity in multiple state models. The first strategy includes modelling multimorbidity as a predictor that affects transitions amongst states in a simple three state model of functional disability. The second approach is to model multimorbidity as a state in a five state model of multimorbidity and functional disability. Our research questions are as follows. What is the impact of multimorbidity on transition rates in a three state model of health status and functional disability that controls for age and gender? To what extent does a five state model of multimorbidity and functional disability capture differences in mortality and functional disability risks?

What are the pricing and life expectancy implications of using the two different methods of capturing the effect of multimorbidity on transition rates in multiple state health models?

1.2 Thesis contributions

In this section we summarise the main contributions of this thesis using empirical data from the United States Health and Retirement Study (HRS).

In Chapter 3 we extend the literature on body mass index trajectories of older adults by identifying trajectories of joint BMI and self reported health status (Zajacova et al., 2015; H. Zheng et al., 2013). We identify three clusters: normal, stable BMI and declining very good health (A), normal, stable BMI and declining fair health (B) and high, increasing BMI and declining good health (C). One–Way Anova tests show that the clusters are unique across different socio–economic characteristics and pairwise tests demonstrate that the differences between clusters are statistically significant even after adjusting for multiple testing.

The estimated predicted probabilities of death are consistently highest in the normal, stable BMI and declining fair health cluster at different ages for both males and females. Hence, they have the shortest longevity prospects. Individuals in Cluster A have the lowest probabilities of death at any age for males and females. We also find that education, total–non housing wealth and total household income have no significant impact on mortality when controlling for socio–economic variables and other risk factors in models including and excluding cluster allocation. Our second major contribution from this project is methodological since we demonstrate how to apply an existing objective data driven technique of determining risk profiles to mortality modelling literature.

In Chapter 4 we extend the literature on multistate models of health and functional disability that incorporate health status amongst older ages (Fong et al., 2015; Z. Li et al., 2017; Sherris & Wei, 2021). We find that there is significant mortality heterogeneity amongst people who become chronically ill during their lifetimes. Individuals in the cluster with highest education, wealth and income tend to have longer and healthier life expectancies than those with lesser means. These results encourage practitioners to develop products that are more tailored to the mortality and morbidity experience

of specific groups. Methodologically, models with clustering outperform models without which shows that the HMMs are better at capturing mortality heterogeneity due to a better fit to the empirical data.

One significant contribution we make in Chapter 5 extends the literature on whether there is mortality morbidity or compression over a certain duration (Kingston et al., 2018; Tetzlaff et al., 2017). Our results suggest that the choice of model has a great impact on the conclusions of whether there is morbidity compression or expansion. Results from the three state model of functional disability and health weakly suggest morbidity expansion while results from the five state model of multimorbidity and functional disability illustrate that gains in life expectancy are being lost to increasing multimorbidity which supports morbidity expansion. These varying conclusions carry over to pricing and can result in serious mispricing of life annuities, lifecare annuities and long term care products.

1.3 Thesis outline

The rest of this thesis is structured as follows. The literature review is provided in Chapter 2 and it covers the principal literature in mortality modelling, heterogeneity and applications of machine learning to mortality modelling. Chapters 3, 4 and 5 answer specific research questions and each of these chapters has its own discussion and conclusion. In Chapter 3 we demonstrate how to apply a three dimensional k-means clustering algorithm to create mortality risk profiles. Chapter 4 illustrates how to use hidden Markov models with covariates and k–medoid clustering to place individuals with varying level of mortality risks into different groups. We demonstrate the results using three state models of health and functional disability. Chapter 5 compares two methods of integrating multimorbidity in multiple state models. We conclude in Chapter 6 where we summarise and discuss the overall contributions of this thesis, its limitations and suggested future work.

Chapter 2

Literature review

2.1 Introduction

We begin by reviewing the literature in the following areas: mortality modelling, heterogeneity, determinants of mortality, longitudinal data, machine learning, cause of death mortality modelling and multimorbidity. Section 2.2 provides a general overview mortality modelling approaches and discusses the importance of explanatory models given the increase in availability of individual level longitudinal data. Sections 2.3 and 2.4 introduce the concepts of mortality heterogeneity and determinants of mortality which are the main areas of interest in this thesis. We also examine the machine learning literature and its applications to mortality modelling. This literature is important as techniques from this field are applied in Chapters 3 and 4 to investigate mortality heterogeneity. Lastly, we analyse the literature on cause of death modelling in Section 2.6. The literature on cause of death modelling is crucial because it introduces us to the concept of multimorbidity which we further explore in Chapter 5.

Chapter 3 explores the use of k–means clustering to cluster single and joint trajectories of body mass index and self–reported health from individual level data. The literature associated with this chapter is provided in Sections 2.2, 2.3, 2.4 and 2.5. Chapter 4 investigates the use of hidden markov models with covariates and k–medoids clustering to cluster health trajectories. The literature review sections relevant to this chapter are Sections 2.2, 2.3, 2.4 and 2.5. Hidden markov models provide a statistically robust way to determine meaningful distance measures between health trajectories. Chapter

5 provides an actuarial lens on multimorbidity and long term care. The sections of literature review associated with the chapter are Sections 2.2, 2.3, 2.4 and 2.6. We notice that while actuaries realise the importance of cause of death data, they have overlooked multimorbidity as the precursor to cause of death modelling and forecasting. In section 2.7 we provide a summary of our review and attend to the gaps we identify in our literature review. These gaps lead to the works in Chapter 3, Chapter 4 and Chapter 5.

2.2 Mortality modelling

Mortality modelling methods can generally be split into three groups: expectation, extrapolation and explanation (Booth & Tickle, 2008; Haberman & Renshaw, 2011; Haberman & Sibbett, 1995; Pitacco, 2004; Tabeau, 2001; Tuljapurkar & Boe, 1998; Wong-Fupuy & Haberman, 2004). Expectation methods are usually based on the subjective opinions of experts. Extrapolative methods use past experience to generate forecasts. Explanatory methods rely on a structural or cause of death approach using covariates beyond age and gender. It is not unusual for extrapolative models to have some explanatory and/or expectation based features and vice versa.

Stochastic extrapolative models provide age specific mortality indicators over time and their associated ranges of uncertainty instead of point forecasts (Booth & Tickle, 2008; Wong-Fupuy & Haberman, 2004). There are a variety of extrapolative techniques such as principal components based (Brouhns et al., 2002; De Jong & Tickle, 2006; Lee & Carter, 1992; N. Li & Lee, 2005; Renshaw & Haberman, 2003a), regression–based (Renshaw & Haberman, 2003b; Sithole et al., 2000), Bayesian methods (Cairns et al., 2006, 2009; Raftery et al., 2013; Wiśniowski et al., 2015), continuous time affine models (Blackburn & Sherris, 2013; Y. Xu et al., 2019), parametric models (Heligman & Pollard, 1980; Siler, 2007) and non–parametric models (Hyndman et al., 2013; Hyndman & Ullah, 2007). The different methods can be assessed under a variety of criteria that can be qualitative (biological reasonableness, plausibility of uncertainty and robustness of forecasts) or quantitative (parsimony, model fit, et cetera) (Cairns et al., 2009; Cairns et al., 2011; Dowd et al., 2010a, 2010b).

While there is general consensus on the age pattern of mortality at young and adult

ages, there are varying views on what happens at older ages. Oeppen & Vaupel (2002) show that life expectancy has steadily increased over the past 150 years at a rate of 0.25 years each year even though others argue that mortality gains in older ages are slowing down amongst older ages (Horiuchi & Wilmoth, 1998). Some of the reasons for different longevity outcomes are the choice of mortality indicator used in forecasting, the assumptions made, and the level of expert judgement adopted (Bergeron-Boucher et al., 2019; Stoeldraijer et al., 2013; Wong-Fupuy & Haberman, 2004). Regardless of the extensive work on forecasting mortality rates, forecasts have improved but tend to be inaccurate and this can have a catastrophic impact on pension funds and life insurers (Basellini et al., 2022).

The main disadvantage of extrapolative methods such as Lee–Carter(LC) is an over reliance on historical experience without an appreciation of advances in medicine, climate change, antibacterial resistance, pandemics, et cetera (Gutterman & Vanderhoof, 1998). The type of data that would make explanatory methods become the gold standard of mortality forecasting is currently unavailable, but as more individual level data at population level become available there is great potential in this area. This information needs to be fed into mortality models not only to improve forecasts but also to understand the structural ways in which mortality and its determinants interact. Many experts argue that there is no explicit manner to model the relationship between mortality and covariates. In fact, the inter-relatedness of determinants of mortality makes explanatory methods difficult to model (Edwards & Tuljapurkar, 2005; Lin & Liu, 2007; Tuljapurkar & Boe, 1998). However, this has not stopped researchers from explicitly forecasting or modelling mortality using cause of death and socio–economic data.

For example, Girosi & King (2008) apply a Bayesian approach based on partial pooling of expected mortality to forecast all–cause mortality using gross domestic product and tobacco use as exogenous covariates in a linear regression setting. Other researchers consider the global burden of disease approach, where estimates of cause specific mortality rates for a specific age group and gender are found by regressing gross domestic product per person, human capital, smoking intensity and time (Murray & Lopez, 1997a; Murray & Lopez, 1997b, 1997c). This establishes 25 year mortality forecasts for 9 cause of death disease clusters for optimistic, pessimistic and baseline scenarios. Their results show a decrease in communicable diseases and Human Immunodeficiency Virus (HIV) related

deaths. Heart disease, depression and road traffic deaths are expected to account for the highest disease burden in the future and smoking related deaths also increase. Their results are relevant mainly because of the incorporation of socio–economic variables into mortality modelling. This regression equation approach for global burden of diseases has been updated annually and is now a study with 194 countries and its results are published annually in a special issue of renowned medical journal, the Lancet. The specific impact of a cause of death on life expectancy and the disease burden through years of life lost due to certain injury or illness is documented for each country from 1990. The most recent issue provides global age–specific fertility, healthy average life expectancy and mortality forecasts with 95% uncertainty intervals (Abbafati et al., 2020).

The influence of explanatory models on public policy and growing interest in socio-economic information to explain mortality differentials has probably motivated the increasing literature on extrapolative methods integrating information on more determinants of mortality into mortality models. Mortality differentials are usually shown by differences in life expectancy for different socio–economic groups. In many of the recent studies in actuarial literature, population mortality models use socio–economic data from deprivation indices in the UK (Cairns et al., 2019; Villegas & Haberman, 2014; Wen et al., 2021). Not many studies have used more granular data in the form of individual–level data that has the potential to better explain the heterogeneity in a population (Madrigal et al., 2011; Meyricke & Sherris, 2013; M. Xu et al., 2019).

2.3 Heterogeneity

Mortality heterogeneity is the non-uniformity of each individual's susceptibility to death in a population. Pitacco (2019) provides a detailed review of mortality heterogeneity for an actuarial audience. The individual risk factors that explain mortality heterogeneity can be split into two groups; observable and unobservable risk factors. Observable risk factors include mortality determinants such as age, sex, income, postcode and these are commonly used as rating factors. Unobservable risk factors are defined as "frailty", a term introduced by Vaupel et al. (1979).

2.3.1 Frailty

According to Vaupel et al. (1979), frailty $\mu(x, t, z)$ for an individual aged x at time t can be defined mathematically as follows:

$$\mu(x, t, z) = z \cdot \mu(x, t, 1), \tag{2.1}$$

where z is the level of frailty and $\mu(x, t, 1)$ is the standard force of mortality which can be described by an appropriate law of mortality such as Gompertz or Gompertz–Makeham (Gompertz, 1825; Makeham, 1867). Frailty is positive and lower values imply lower chances of dying. There are a variety of probability density functions for frailty. Su & Sherris (2012) consider Gamma and log Normal distributions of frailty and find that the Gamma distribution has a better fit to Australian data. The Gamma distribution is ideal for modelling frailty due to its flexibility and tractability (Vaupel et al., 1979). A Gompertz-Gamma model of fixed frailty is the same as Perks Law of Mortality which implies that mortality curve tends to flatten at older ages (Perks, 1932; Pitacco, 2019; Thatcher et al., 1998).

While assuming that frailty is constant makes calculations much easier, Markov processes can be used to model variable frailty over an individuals lifetime (Le Bras, 1976; Pitacco, 2019). Yashin et al. (1994) show that while statistical models for fixed frailty and variable frailty are conceptually different, they produce identical mortality patterns. Liu & Lin (2012) extend their previous work Lin & Liu (2007) on modelling ageing as a finite-state continuous-time Markov process with a single absorbing state by adding a stochastic Gamma time change process to capture systematic mortality risk. This enhancement is motivated by subordinated Markov processes in finance to model stock prices (Daal & Madan, 2005; Madan et al., 1998; Madan & Milne, 1991). One key contribution of the model is its dual ability to model both systematic mortality risk and heterogeneity through physiological age.

2.3.2 Mortality models using population level data and other covariates beyond age and gender

Due to the impact of mortality heterogeneity on life insurance, pension funds and social security systems, most stochastic mortality models are now being adapted so that they can capture the heterogeneity amongst populations. For extrapolative models, there are a variety of ways to capture more socio–economic information, health status, demographic code, et cetera. This section discusses some of the research in this area using population level data.

In Su & Sherris (2012), the authors adapt Lin & Liu (2007) Markov ageing model to Australian population data by adding a Gompertzian factor to the mortality rate at older ages that is health dependent. Their model is more parsimonious than the LC model since only 12 parameter estimates are required compared to over 100 for the latter. Specifically, one would need 3 parameters at each age for forecasts from age 65 to 100 for the LC model. They compare the Markov ageing model to the frailty model and find that the Markov ageing model tends to flatten at older ages whereas mortality increases linearly with age for the frailty model. Since it is unclear what happens to mortality rates at older ages, practitioners have to choose a mortality model that exhibits reasonable behaviour at older ages for that particular population. The Markov ageing model results in higher priced annuities than the frailty model for individuals with similar health conditions for both whole life annuities and deferred life annuities. This means that the Markov ageing model has a more conservative mortality assumption than the frailty model. The different rates for annuities for various physiological ages and frailty factors supports the notion that annuities should be offered at prices that show appreciation for the mortality heterogeneity in a population. Individuals at a higher physiological age will generally be offered cheaper annuities.

Sherris & Zhou (2014) develop a stochastic Markov Ageing model that captures both systematic mortality risk and heterogeneity using Australian data. This model is compared with Vaupel's fixed individual frailty model and the Le Bras' variable individual frailty which both measure heterogeneity but do not measure systematic mortality risk (Le Bras, 1976; Vaupel et al., 1979). For the Markov ageing model, health status is used to divide the Australian population into 5 distinct groups based on the severity of the

health conditions, ranked from best to worst health. Severity is defined as the probability of death from a specific cause of death. The authors also assume that health conditions are independent and that for individuals who die from multiple causes, their mortality rate is from the condition with the highest mortality overall. At age 65, the Markov ageing model is more pessimistic about future expected life times than the frailty models. This results in much lower premiums than the frailty models. Moreover, the Markov ageing model has the lowest standard deviation for an annuity pool with $1,000$ individuals with best health implying less heterogeneity than the frailty models as the individuals age. Despite the adverse selection that comes from individuals with the best health, annuity providers lower their risk by catering only to these individuals than exposing themselves to larger risk from a pool of individuals with mixed health. Systematic mortality risk increases with pool size and this risk is not diversifiable. Frailty models fail to capture this element as they only show an increase in idiosyncratic risk that can be diversified as pool size increases. In summary, Sherris & Zhou (2014) demonstrate that the model risk is very significant and motivates models that fully capture both systematic mortality risk and heterogeneity.

Olivieri & Pitacco (2016) address the lack of rigour in identifying separate risk groups for life annuities by proposing the use of frailty models to separate different risk profiles rather than applying adjustments to population rates with no statistical basis (Le Bras, 1976; Vaupel et al., 1979). The authors adopt a fixed frailty model for Italian population data where the force of mortality is Gompertzian and the distribution of the frailty is gamma which implies that mortality decelerates at older ages (Gompertz, 1825; Perks, 1932). Risk groups are then defined by then specifying frailty levels for each group. Without considering systematic risk, Olivieri & Pitacco (2016) show that the liabilities of larger annuity portfolios with greater heterogeneity exhibit lower volatility than homogeneous portfolios.

Villegas & Haberman (2014) develop a relative mortality age–period–cohort (APC) forecasting model that measures current levels of socio–economic mortality differentials and forecasts trends within different socio–economic groups in England based on a multiple deprivation index. The deprivation index is a composite score of one of 32,844 areas with an average population of 1,500 individuals in England based the area's income, employment, education, health, crime, living environment and barriers to housing and services.

The relative approach model uses the national population as a reference mortality set that is forecasted alongside subpopulations exhibiting varying socio–economic deprivation. In their research, they collate the different deprivation areas into 5 quintiles ranked from least deprived to most deprived. Their results show that the forecasted mortality rates within each quintile are coherent and consistent with the overall England and Wales forecasts during the same forecasting period. Their results show a general increase in the mortality differentials between the most deprived and least deprived groups with younger age groups (50–64) experiencing up to 4 times the mortality of the least deprived socio–economic group. Importantly, the results suggest that socio–economic factors show greater variability of annuity rates than variability due to gender.

Cairns et al. (2019) develop an affluence index to proxy socio–economic differentials in mortality for Danish males and females. For any individual, the affluence index in a specific year and a certain age is the sum of one's net wealth in the year and the product of capitalisation factor for retirement income and taxable income in the year before. Individuals are then allocated to 10 groups based on the values of the affluence and allowed to make transitions until a lockdown age (in line with the retirement age) where no changes are made to the socio–economic groups. Age specific mortality rates are then forecast in each group using a multi-population gravity model with parameters estimated through a Bayesian approach (Cairns et al., 2006; Plat, 2009). While the model generates forecasts that are coherent and consistent with national mortality forecasts for Danish males, it fails to produce similar results for females. This implies that while affluence can be an effective proxy of socio–economic mortality differentials for males, an index that takes into account more socio–economic variables such as the Index of Multideprivation in the UK will generate more coherent and consistent forecasts for both males and females as shown by Villegas & Haberman (2014) that avoid crossovers between socio–economic groupings and maintain the hierarchy chosen initially.

Wen et al. (2021) consider 12 age–period multi–population mortality models to determine which model provides the best fit when socio-economic data is accounted for based on the Index of Multideprivation in England (Cairns et al., 2006; Kleinow, 2015; Lee & Carter, 1992; Plat, 2009). Across the deciles from the most deprived to the least deprived, the Common Age Effect (CAE) model with an age effect that is shared by all the groups has the best performance as shown by the lowest Bayesian Information Criterion (BIC) for

both males and females (Kleinow, 2015; Wen et al., 2021). It is important to note that second best performing model, the Plat model with a common age affect across all groups, has significant patterns in its residuals for some groups which are removed by adding a common cohort feature across all groups that improves its fit. This adjusted Plat model with a cohort effect has better performance than the CAE. These results are similar to those shown by Villegas & Haberman (2014) in which age–period–cohort models also had the lowest BIC.

2.3.3 Mortality models using individual level data

Survival models and generalised linear models (GLMs) are the main methods of modelling mortality using individual level data. A lot of work in the actuarial literature is focussed on post retirement mortality because the main beneficiaries are retirees. Madrigal et al. (2011) use GLMs based on cross–sectional data from the longevity analytics firm, Club Vita. The baseline exposed to risk is approximately 2.9 million with 90,000 deaths from 91 pension schemes. Despite the richness of the data set, they focus on the effects of age, gender, geo–demographic codes and measures of affluence through pension amount and last known salary. Variables for measures of affluence and geo–demographic codes from ACORN (a post code and household segmentation tool) are clustered using Ward's method and recursive partitioning (Ward, 1963). Their results show that salary at exit and geo–demographic codes are strong predictors of male mortality. Pension amount is an inconsistent predictor of mortality due to the fact that individuals can contribute to a variety of schemes making sole reliance on this amount not ideal. However, pension amount is a relevant predictor for females signifying the importance of an individual approach for determining pension liabilities. Overall, the results show that there is heterogeneity in the mortality experience of pensioners that can lead to differences in life expectancy of up to 10 years and a reduction in annuity values of up to 60%.

Meyricke & Sherris (2013) use a generalised linear mixed model (GLMM) and individual level data to model both underwriting factors and frailty. Their results show that common underwriting factors such as education, age, gender and health history do not fully capture the heterogeneity within a population. Heterogeneity remains significant after underwriting since there is not much improvement in the reduction in variance of frailty.

Incorporating frailty in a mortality model is particularly significant for individuals from high risk groups as seen by a higher variation in annuity values for different levels of frailty especially when compared to low and medium risk groups. This investigation does not explore the impact of time varying–covariates on mortality despite the use of a GLMM. As such, a more recent work of M. Xu et al. (2019) looks at both mortality heterogeneity and time trends using longitudinal mortality models. This work extends the work using generalised linear model framework by adding the impact of time trends (Madrigal et al., 2011; Meyricke & Sherris, 2013). Their results show that mortality improvements are mainly from the simultaneous decline in cancer, cardio-vascular disease and hypertension (M. Xu et al., 2019). Population level inferences are made using marginal models. An attempt to use the LC model to model mortality trends with individual level data that is split into different groups is unsuccessful because the differences in mortality improvements for the different subpopulations are not statistically significant. Recently, using United Kingdom (UK) electronic health records, Kulinskaya et al. (2021) apply landmark analysis and a Gompertz survival model to estimate the survival probabilities and life expectancies of 648 risk profiles for males and females in each deprivation quantile.

2.4 Determinants of mortality

In this subsection, we discuss the determinants of mortality. The seminal work of R. L. Brown & McDaid (2003) and Tuljapurkar & Boe (1998) provides a comprehensive treatment of the determinants of mortality and applications to actuarial studies based on the information available at the time. Thereby, this subsection mostly updates previous results with new information from recently published articles. We consider age, gender and socioeconomic determinants of mortality (education, income and marital status). We also discuss health behaviours, obesity, environmental risks and race and ethnicity.

2.4.1 Age

Generally, mortality laws are used to show how the log force of mortality evolves over time. This usually presents itself in a one factor parameterisation function which shows high mortality rates in the early stages of childhood, an exponential increase in mortality

during middle age and deceleration of mortality rates during older ages. Gender differences in mortality have long been the norm with females living longer than males based on life expectancy at birth. However, there are also differences of mortality and gender within the patterns of age structure. In most countries the gender gap in mortality is narrowing (Glei & Horiuchi, 2007; Trovato & Lalu, 1998). It is generally appreciated that while mortality differentials exist before senescent mortality, this mortality differential stops having an impact at older ages (Vaillant & Mukamal, 2001). Age interacts with many variables in different and complicated ways and these interactions should be considered when pricing annuities (R. L. Brown & McDaid, 2003).

2.4.2 Gender

Gender has long been used in the pricing of insurance and annuities with females living longer than males across different times and countries. This is despite the mortality–morbidity paradox. That is, at any point in time, women are more frail yet it is men who have a higher probability of death. An understanding of how gender interacts with other risk factors both biologically and behaviourally is not common knowledge (R. L. Brown & McDaid, 2003). Internationally, the gender gap in mortality is narrowing. Recent work shows that $86 - 89\%$ of the gender age gap is attributed to gender specific preferences and health behaviours on longevity using utility functions (Oksuzyan et al., 2008; Rieker & Bird, 2005; Schünemann et al., 2017). These results are of particular relevance to insurance companies and annuity providers that use unisex rates in pricing.

2.4.3 Education

Generally, higher educated individuals have lower mortality rates than those with lower education levels. Using US longitudinal data, Preston & Elo (1995) find increasing educational differentials in mortality amongst white males. In contrast, white females aged $25-64$ experience narrowing educational differentials in mortality, while the older females $(65 - 74)$ have not changed much. Their analysis is based on a comparison of the age standardised death rate across three indices of inequality: slope index of inequality (SII), SII/death rate and an index of dissimilarity. Educational differentials are significant for both cause-specific mortality and all-cause mortality (Huisman et al., 2005; Kulhánová

et al., 2014; Mackenbach et al., 2008). In a comparison of US data from 1990 and 2008, Olshansky et al. (2012) find that differences in life expectancy at birth continue to widen. In 2008, highly educated males are expected to outlive their less educated peers by 14.2 years while for females the difference is 10.3 years.

2.4.4 Health behaviours

Males have higher mortality from major causes of death and gender differences in health behaviours drive many causes such as heart disease, lung cancer, motor vehicle deaths, suicide, etc (Waldron, 2012). Males smoke and drink heavily more than females and this partially explains sex mortality ratios (Waldron, 1991). Additionally, males tend to engage in riskier sexual behaviour making them more likely to be infected by sexually transmitted diseases such as Human Immunodeficiency Virus (HIV) (Smith, 1991). Diets have improved and exercise has increased for both males and females since the 1980*s* but it is unclear whether there are any gender differences on mortality based on diet or exercise (Waldron, 2012). Recently, other health behaviours such as sitting, and sleep duration are getting more attention in how they predict all–cause mortality. In a systematic review of prospective studies with over $590,000$ participants, Chau et al. (2013) find that sitting for more than 10 hours each day increases the risk of death by 34% after controlling for exercise. Long sleep durations of greater than 8 hours of sleep each day and short durations of sleep of less than 7 hours have been found to increase mortality by 30% and 12% respectively across countries using data from meta analysis of 16 studies with more than 1 million participants (Akiko et al., 2004; Cappuccio et al., 2010; Gallicchio & Kalesan, 2009).

2.4.5 Income

Higher income has been associated with lower mortality rates in multiple studies (Deaton & Paxson, 1999; Knox & Tomlin, 1997). In an analysis of over $500,000$ records of pensioner data, Madrigal et al. (2011) find that last known salary has better predictive power for mortality than the amount of pension. In a study of 1.4 billion tax returns from the USA covering the period between 1999 and 2014, Chetty et al. (2016) study the relationship between income and mortality using race and ethnicity adjusted life expectancy at

age 40. Chetty et al. (2016) show that the gap between the richest 1% and poorest 1% was 14.6 years during the period under investigation. The richest men gained 2.34 years in life expectancy whilst the poorest only gained 0.32 years from 1999 to 2014, highlighting the stark income mortality differentials. In the bottom income quartile, the authors find that factors such as the proportion of the immigrant population, median house values, proportion of college educated individuals and population density in counties were positively correlated with life expectancy in contrast to economic indicators such as inequality, levels of unemployment and access to healthcare.

2.4.6 Marital status

Married people exhibit lower mortality than unmarried individuals. However, social ties through social networks (family, friends and the community) are also associated with lower mortality rates (Berkman, 1984; Rogers, 1996). Income interacts with marital status in such a manner that married individuals with higher incomes have lower odds of dying compared to couples with lower incomes (Rogers, 1995). H. C. Wang et al. (2016) use Taiwanese population data from 1975–2011 to show that married individuals have longer life expectancy than single and widowed people, and the impact is stronger for males than females. This is further evidence for using marital status as a rating factor in life insurance and annuities. For Taiwan, discounts based on marital status for life insurance products are higher than those from using smoking and obesity as rating factors. However, as discussed in Rogers (1995), those who are married are more likely to have higher incomes and higher levels of education.

2.4.7 Obesity

Body mass index (BMI) is calculated as weight in kilograms divided by height in meters squared. Higher body mass index (BMI) is associated with higher mortality. However, central obesity (accumulation of fat in the abdominal area) in individuals with normal BMI is associated with an even higher risk of mortality (Cerhan et al., 2014; Coutinho et al., 2013; Sahakyan et al., 2015). Using Cox-regression analysis for a large sample of over 650,000 Caucasians with a median follow–up (50^{th} percentile of time to observe event of interest) of 9 years, Cerhan et al. (2014) find that for all BMIs, having a large

waist circumference increases one's risk of death. A high waist circumference is likely to decrease life expectancy in females and males by 3 years and 5 years respectively. Having a high waist to hip ratio doubles the risk of death after controlling for BMI (Sahakyan et al., 2015).

2.4.8 Race and ethnicity

Race and/or ethnicity cannot be used to price annuities. However, mortality differentials amongst different racial groups in countries across the world are known to exist and persist. Studies investigating black–white mortality differentials in the USA indicate that allostatic burden (repeated chronic stress that can manifest through stress hormones and biomarkers) can partially explain why Blacks experience higher mortality than Caucasians even after controlling for socio–economic status and health behaviours (Duru et al., 2012; Geronimus et al., 2006). Hill et al. (2007) find that the life expectancy at birth between Aboriginal Australians and non–indigeneous is 13.1 years for males and 12.4 years for females. The authors argue that the results show excess mortality for Aboriginals when compared to other indigenous people such as the Maori in New Zealand and American Indians in the USA.

2.4.9 Environmental risks

The World Health Organisation (WHO) defines environmental risks as "all the physical, chemical and biological factors external to a person, and all related behaviours, but excluding those natural environments that cannot reasonably be modified" (Neira & Prüss-Ustün, 2016). Using a burden of disease approach, Neira & Prüss-Ustün (2016) find that environmental risk factors account for $23\%(95\%CI : 13\% - 34\%)$ for all deaths worldwide. The authors demonstrate that the majority of deaths related to environmental effects are in Africa, South East Asia and non–OECD countries in Europe. Children under 5 and adults over age 50 are most susceptible to dying of diseases caused by environmental factors. Gender differences in environmental risk factors are negligible.

2.5 Machine learning

In this section, we provide a general overview of machine learning and how machine learning has been applied in mortality modelling and health.

Machine learning can be split into two categories, supervised and unsupervised (Hastie et al., 2009; James et al., 2013). With supervised learning, the output is known. Whereas in unsupervised learning, the algorithms try to find patterns in the data and there is no known output. Variables and outputs are known as predictors and responses respectively. The training set is data through which algorithms predict the output. Predictors can be quantitative or qualitative. Qualitative input or output can be categorical or discrete. Quantitative input or output is numerical. Inspired by the developments in machine learning, statistical learning is becoming more popular as data sets become larger and the range of statistical techniques to analyse increases. Classification problems are instances in which a qualitative output is predicted. Regression problems are methods for which the output is numerical. Some of the most popular techniques used in classification and regression problems are described in the sections that follow.

2.5.1 Machine learning and mortality modelling

Demographic and actuarial experts have generally not appreciated many machine learning techniques due to a lack of understanding of how these techniques work. The most widely used techniques such as the Lee–Carter (LC) and the Cairns–Blake–Dowd (CBD) models are highly interpretable amongst these audiences and easy to implement (Cairns et al., 2006, 2009; Lee & Carter, 1992). However, machine learning techniques have the potential to exhibit superior predictive performance and the ability to capture non–linearities in mortality data.

One of the first applications of machine learning to mortality modelling is through the use of a Gaussian–Process (GP) to graduate crude rates and ascertain improvement rates (Rasmussen, 2004; Wu, 2016). Some advantages of GPs are that they are data–driven methods that handle uncertainty extremely well since they are based on the Bayesian paradigm. One of the main findings by Ludkovski et al. (2018) is that mortality amongst ages $55-70$ has been stable and chances are that it has worsened and this is contrary to

the opinion that all ages have been making significant gains as reported by the Society of Actuaries (SOA) in 2015 and this makes the associated improvement factors inaccurate (Retirement Plans Experience Committee, 2015).

There are multiple variants of the LC models and comparing the different extensions is not trivial because each variant has its own strengths and weaknesses (Booth et al., 2002; Lee & Miller, 2001; Wilmoth, 1993). Deprez et al. (2017) show how to use a boosting regression tree machine to compare the performance of different stochastic mortality models. In their paper, they back–test different versions of the LC model to show that age–period–cohort model (Renshaw & Haberman, 2006) has a better fit than the age–period model (Lee & Carter, 1992) through changing the feature space of gender, age, year to include a cohort effect. In addition, they also test the Renshaw–Haberman model with an identical boosting regression tree machine and find that the model captures cohort effects better than the LC but remains susceptible to shocks such as the 1918 influenza epidemic. Lastly, they apply a boosting Poisson regression tree to cause-of-death mortality data and manage to show how the conditional probabilities of dying change over time (Alai et al., 2015; Deprez et al., 2017). These results are significant because machine learning techniques are used to find interactions between features that the original models fail to consider and the evolution in mortality rates.

Levantesi & Pizzorusso (2019) extend the work by Deprez et al. (2017) by adding the Plat model for comparison and using different tree based machine learning estimators (Plat, 2009). The Plat model merges LC and the CBD models. Hainaut (2018) applies a neural network to predict mortality rates using US, UK and French data. The use of a neural network is motivated by non–linear principal component analysis (NLPCA) from chemical engineering literature (Hastie & Stuetzle, 1989; Kramer, 1991). Specifically, Hainaut (2018) summarises the term structure in the LC model using various neural network architectures. The best performing neural network is a (3–2–3) network, that is, 3 and 2 neurons in the input/output and intermediate layers. While the author makes a timely contribution to the application of machine learning techniques, this work has been further improved in a number of applications which are discussed below.

The main contribution from Richman & Wüthrich (2019) is identifying the specific functional form of the LC model using representation learning (Bengio et al., 2013). This allows one to compare a variety of the numerous LC extensions in a multi–population

context instead of just 2 or 3 models as done in Hainaut (2018). Furthermore, multi–population models produce more coherent mortality forecasts due to shared improvements in socio–economic conditions and overall health status (Kleinow, 2015; N. Li & Lee, 2005) than single population mortality models. Richman & Wüthrich (2019) extend Hainaut (2018) by considering gender differences and derive a one step process for predicting mortality rates. They consider the effects of both a random walk with and without drift, with the best performing neural network (DEEP5) still outperforming other LC models. The DEEP5 network has 5 intermediate layers with a rectified linear unit (ReLU) activation layer. All networks are fit via a backward–propagation algorithm. This is different from the genetic algorithm used by (Hainaut, 2018). The input layer has age, gender, country and year. While the results are impressive as demonstrated by the lowest mean square error in 51 out of 76 countries, the authors did not consider ensembling, cohort effects and adding more features. Moreover, the results are point forecasts, using probability density forecasts would show uncertainty (Bishop, 1994)

To avoid the problematic "black box" approach of machine learning techniques, Nigri et al. (2019) use a Long Short-Term Memory (LTSM) neural network to capture the evolution of time trend κ_t in the LC model. The main advantage of this technique is that it retains the interpretability of the LC models while improving the prediction accuracy. Demographers and actuaries are more likely to accept this technique as it preserves the interpretive capacity instead of the black-box methods of (Hainaut, 2018; Richman & Wüthrich, 2019). Furthermore, the LC model is usually modelled with an Autoregressive Integrated Moving Average (ARIMA) process which fails to capture non–linearities in the time index due to a constant variance assumption. Using mean absolute error and and root mean square error, the LSTM consistently outperform the best ARIMA models in seven countries for both males and females. Lastly, the log–force mortality rates from a LTSM neural network fits the data better than that of an ARIMA using Australian data. The authors also argue that the forecasted mortality curve is more realistic than that determined by the ARIMA. However, there is no general consensus on the shape of the curve at later ages in life (Gavrilov & Gavrilova, 2019; Gompertz, 1825; Horiuchi & Wilmoth, 1998).

2.5.2 Predictive models in health diagnostics and biomedicine

Weng et al. (2019) compare a multivariate Cox regression and two machine learning techniques to predict all cause mortality in 10 years and found deep learning models had better predictive accuracy. In their study, they use longitudinal data from the UK Biobank (over $500,000$ participants) with 60 predictor variables including age, Townsend deprivation index, diet and medical history. This study is important because it is one of the first to demonstrate the application of neural networks and random forests to longitudinal data. The authors highlight that while machine learning algorithms can identify the variables that are most important in predicting mortality, it is still unclear how exactly these variables contribute. In addition, they fail to provide a method of picking the most important variables since there is no consensus amongst the different methods. Deep learning models identify smoking, age and prior diagnosis of cancer as most important, while the Cox model ranks age, prior diagnosis of cancer and gender as having the strongest Cox regression coefficients.

The above work is extended in Zhao et al. (2019), whereby genetic data of 204 single nucleotide polymorphisms (SNPs) are added to socio-economic variables to improve prediction accuracy. SNPs are the most common genetic variations amongst people. The results show that age remains the most important predictor followed by two SNPS, gender and electronic health record length. The best performing algorithms are gradient boosting trees and convolutional neural networks. However, the authors failed to identify the predictors with the greatest impact on mortality from the convolutional neural network because of its black box approach. It is also interesting to note that longitudinal data show how BMI changes over time is more predictive of cardiovascular mortality than one single value from cross–sectional data.

Statistical and machine learning techniques have been used extensively to predict cardiovascular mortality, the presence of single diseases such as diabetic retinopathy and to identify genes that predict resistance to tuberculosis under different treatments (Arcadu et al., 2019; Chen et al., 2019; Zhao et al., 2019). The deep convolutional neural network that is used to predict progression of diabetic retinopathy, analyses images from a single visit and identifies specific locations in the eye that need to be inspected in contrast to the fovea or optic nerve in routine appointments with a doctor. Most of the data used

in these studies are cross–sectional data, with application of machine learning techniques to longitudinal data appearing in the past few years in medical literature. It remains uncertain whether genetic data can be used for pricing as it is different from health behaviours that people can change. Macdonald (2014) provides a review of the effects of adverse selection on insurance, the privacy concerns and how genetic data can be used in actuarial models. Genetic testing is becoming much cheaper, with the sequencing of the entire human genome continuously decreasing over time. It is highly likely that genetic data could be used as rating factors in the future (Bélisle-Pipon et al., 2019; Tiller et al., 2019; Zhao et al., 2023).

2.6 Cause of death mortality modelling

In this section, we discuss cause of death mortality modelling. This literature is important because it leads us to the literature on old age multimorbidity from which cause of death modelling is derived. The specific literature on multimorbidity is more thoroughly investigated in Chapter 5.

Mortality forecasts from the decomposition of cause of death data are difficult to implement and do not usually perform better than forecasts based on overall mortality (Booth & Tickle, 2008; Wilmoth, 1995). Substantial expert judgement is needed to know whether a specific cause of death has a cohort effect and consequently apply an appropriate model with cohort effects. Even when the appropriate model is used, (say an age–period–cohort model for a cause of death with a period effect) the results are no better than LC model forecasts (Cesare & Murphy, 2009). Despite this, countries still use cause of death mortality forecasts to aid policy development around disease elimination and to identify populations at high risk of specific diseases. Most of the research in the actuarial literature is focussed on how to account for interdependence amongst different causes of diseases and modelling cohort effects (Alai et al., 2015; Arnold (-Gaille) & Sherris, 2013, 2015; H. Li & Lu, 2019; Lyu et al., 2020; M. Zheng & Klein, 1995).

Vector Autoregression (VAR) and Vector Error Correction Models (VECMs) are used to model cause dependence in cause of death mortality forecasts (Arnold (-Gaille) & Sherris, 2013; Hamilton, 1994; Lütkepohl, 2005). Using cause of death data for Swiss females,

Arnold (-Gaille) & Sherris (2013) demostrate that VECMs outperform ARIMA models and are able to capture long-term trends and stationary relationships between variables. Lyu et al. (2020) develop a multipopulation extension of cause of death forecasting that leverages on the multipopulation extensions of LC (N. Li & Lee, 2005). The authors model dependence with a 2 step-beta convergence test that performs much better than N. Li & Lee (2005) and Lee & Carter (1992) due to its ability to capture both "intercause coherence" and "international coherence" (Barro, 1991; D'Albis et al., 2012; Lyu et al., 2020). However, international comparisons of cause specific mortality rates in developed countries have shown that it is not prudent to assume that different countries will have similar mortality trajectories when they have different historical cause-of-death experience (Arnold (-Gaille) & Sherris, 2013, 2015). While some countries have similar trends, there are differences in long run stochastic trends in most countries (Arnold (-Gaille) & Sherris, 2015; Tuljapurkar et al., 2000) .

Alai et al. (2018) analyse the impact on life expectancy of a hypothetical cause of death mortality reduction by socioeconomic circumstances using a logistic multinomial regression that captured dependence among causes and different cohorts and find that different socioeconomic groups are affected differently by different causes of death. As such, specific causes may need to be targeted to reduce socioeconomic inequalities. H. Li & Lu (2019) extend the literature on long term trends and within-cohort dependance by using a hierarchical Archimedan copula to model within-cohort dependance and estimate net mortality intensities (Arnold (-Gaille) & Sherris, 2013, 2015; M. Zheng & Klein, 1995). The LC model is then used to forecast mortality rates for each cohort and cause of disease for US males. Their results show that the high positive dependence between cardiovascular diseases (CVD) and cancer might explain why there have been improvements in mortality from CVD but only marginal results with cancer despite a War on Cancer in the US.

Other authors have used Compositional Data Analysis (CoDa) to forecast mortality rates and apply this to cause of death data. For example, Bergeron-Boucher et al. (2017) establish coherent forecasts of life expectancy at birth for 15 Western European countries in 2050. The CoDa coherent forecasts outperform LC coherent forecasts despite having wider prediction intervals than LC forecasts. Performance is based mean absolute error.

2.6.1 Multiple cause of death and mortality modelling

International classification of diseases (ICD) forms from the World Health Organisation have consistently provided multiple cause of death data. However, cause of death analysis has been limited to the underlying cause of death.

Few investigations have considered forecasting mortality using multiple cause of death data. The World Health Assembly has approved a new ICD–11 that provides more information on individuals at population level including detailed pharmaceutical benefits, postal code, genotype and socioeconomic risk factors. We know that people do not normally die from one cause of death, that is, the underlying cause of death is only one part of a bigger story. There are other intermediate causes and terminal causes that contribute to deaths. This is particularly important amongst older people since they tend to acquire multiple comorbidities which are usually chronic in nature and it is difficult to establish the underlying cause of death. However, the underlying cause of death is ideal for identifying the point of public health intervention and making international comparisons.

One could argue that the data available are quite rich such that over reliance on single causes of death might lead to an oversimplification of the causes of death because it misses out the impact of other causes. Recent research by Moreno-Betancur et al. (2017) proposes that weights are assigned to each cause of death using a Cox-regression framework for any disease mentioned on an ICD–10 form. This technique is innovative because mortality can be "conceptualised as a mixture of disease processes" (Moreno-Betancur et al., 2017). More weight is put on the underlying cause. Their results show that the impact of socioeconomic inequalities on diseases has been underestimated on diseases with low exposure. However, the choice of weights is still subjective, there is no use of temporal information and other statistical approaches such as a Bayesian approach and machine learning can be explored to improve the methodology.

2.7 Conclusion

Deriving appropriate risk profiles is a critical aspect in the design and pricing of actuarially fair health and longevity risk products. Since markets for annuities are thin globally except in countries with mandatory annuitisation, it is very useful to explore mortality modelling

methods that can increase the demand for annuities by covering a broader population. Given access to individual level longitudinal data, actuaries can realise an opportunity to adopt statistically robust ways of determining mortality and morbidity risk profiles.

In the actuarial literature, risk profiles have been determined using deprivation indices, looking at single determinants of mortality such as education, or on a subjective basis such as high and low education, good health and bad health. This thesis seeks to combine our knowledge of the determinants of mortality heterogeneity and statistical learning techniques to derive better defined risk profiles.

The use of health trajectories and unsupervised machine learning algorithms has not been explored in the literature as a means of developing mortality risk profiles. Therefore, we apply a k–means clustering algorithm for longitudinal data to joint trajectories of body mass index and self reported health to develop mortality risk profiles in Chapter 3. This fills this gap by demonstrating that different clusters have different longevity prospects which directly impacts the pricing of annuities. In the scenario where health status is included in actuarial models as a source of mortality heterogeneity, no one has explored the notion of linking health status with covariates using hidden Markov models in pricing. Hence, in Chapter 4 we use hidden Markov models with covariates to place individuals into distinct risk profiles.

Multimorbidity has not been explored in the actuarial literature as a source of mortality heterogeneity. Yet, multimorbidity has a direct impact on cause of death mortality modelling and forecasting. To the best of our knowledge, there is no link in the literature between multimorbidity and the pricing of long term care products. In the epidemiology literature, multistate models are used to calculate mortality rates, incidence and prevalence rates of multimorbidity for both men and women. However, in most of these models they do not consider functional disability or recovery from multimorbidity which have significant impacts on the pricing of long term care products. In Chapter 5 we demonstrate how best to incorporate multimorbidity in actuarial pricing of annuities, long term care and life care annuities. We do this by comparing different multistate models of functional disability with recovery.

Chapter 3

Mortality heterogeneity and the k–means clustering of body mass index and self reported health trajectories

Earlier versions of this chapter were presented at the following conferences and events:

- Fourth Insurance Data Science Conference, Università Cattolica del Sacro Cuore, Milan, Italy. "Mortality Heterogeneity and Clustering using Joint Body Mass Index and Self-Reported Health Trajectories", 17 June 2022.
- CEPAR 29^{th} Colloquium on Pensions and Retirement Research, University of New South Wales, Sydney, Australia, "Mortality Heterogeneity and Clustering Using Joint Body Mass Index and Self–Reported Health Trajectories", 1 December 2021.
- Centre of Actuarial Studies Seminar Series, Department of Economics, University of Melbourne, Melbourne, Australia. "Mortality Heterogeneity and Clustering Using Joint Body Mass Index and Self–Reported Health Trajectories", 22 October 2021.
- United as One: 24^{th} Congress on Insurance: Mathematics and Economics, Virtual Conference. "Clustering and Mortality Heterogeneity using Joint Body Mass Index and Self–Reported Health Trajectories", 9 July, 2021.

3.1 Introduction

Due to the impact of mortality heterogeneity on life insurance, pension funds and social security systems, there is increasing interest in capturing heterogeneity using deprivation indices or single variables such as education or income in stochastic mortality models (Cairns et al., 2019; Villegas & Haberman, 2014; Wen et al., 2021). However, clustering techniques have yet to be applied to longitudinal individual level data to determine homogenous risk profiles.

As such, our main goal in this paper is to extend the literature on "actuarial precision" by developing more objective determination of risk profiles (Kulinskaya et al., 2020, 2021). We carry out this using k–means longitudinal clustering to create groups of individuals with similar risk profiles based on observable risk factors. Machine learning techniques are used to cluster the data and reduce the level of judgement used in the model. We also determine the pricing implications on annuities for the different clusters. Therefore, we seek to answer the following questions:

- What clusters emerge based on the joint modelling of body mass index and self reported health trajectories?
- Is there an association between cluster membership and each risk factor?
- Is there an association between mortality and cluster allocation?
- Does cluster allocation impact mortality when controlling for socio–economic variables?
- Can the clusters impact mortality while controlling for other risk factors and the predictors BMI and self–reported health are excluded?
- What are the pricing implications for annuities in each cluster?

We identify three clusters: normal, stable BMI and declining very good health (A), normal, stable BMI and declining fair health (B), high, increasing BMI and declining good health (C). One–Way Anova tests show that the clusters are unique across different socio–economic characteristics and pairwise tests show that the differences between clusters are statistically significant even after adjusting for multiple testing. The estimated predicted

probabilities of death are consistently highest in the normal,stable BMI and declining fair health cluster at different ages for both males and females. Hence, they have the shortest longevity prospects. Individuals in Cluster A have the lowest probabilities of death at any age for males and females. We also find that in the presence of other socio–economic variables and risk factors, the role of education, total–non housing wealth and total household income is mediated through other socio–economic variables and risk factors.

The rest of the chapter is organized as follows: Section 3.2 briefly describes the data. Section 3.3 describes the methods used to cluster the longitudinal data and estimate marginal models for the different clusters. Section 3.4 provides the results of clustering, marginal models and pricing implications. Section 3.5 is a discussion of the results and we conclude in Section 3.6.

3.2 Data

The data are extracted from the United States Health and Retirement Study (HRS) which is a nationally representative survey that follows Americans aged 50 and older. The HRS is sponsored by the National Institute on Aging (grant number NIA U01AG009740) and is conducted by the University of Michigan. It has 13 waves of data collected biennially from 1992 to 2016. The first wave has individuals born in the period $1931 - 1941$. We are interested in this cohort since members have been observed for the longest duration. Actuarial studies in the literature that use longitudinal data have relied on this data set (Meyricke & Sherris, 2013; M. Xu et al., 2019). Additionally, incorporating many cohorts would make direct comparisons to literature that uses the initial cohort difficult. Moreover, we are more likely to get variable trajectories of reasonable length when we use the oldest cohort unlike younger cohorts. However, we will consider different cohorts in future work as we did not test this assumption. As is standard survey practice, the HRS omits individuals in institutions such as prisons or aged care facilities. However, if a member transitions from a regular household to an institution, they continue to be interviewed despite their new residency status. This means that transitions to institutions and subsequent transitions are adequately captured in the data.

The HRS data provides information on a variety of variables covering demographics, fam-

ily structure, health, education, financial information, employment history, social security and pensions. Some of the data have restricted access such as genetic data, social security and geocodes. In this study, we focus on known determinants of mortality that can be used as rating factors that are available from the RAND[1] file and are not restricted. The RAND file provides cleaned and processed variables. Where a variable has missing information, the RAND file provides the reasons for missingness. Therefore, for each individual there is a record of person identifier, mortality status, interview date, age at interview, gender, marital status, education, doctor diagnosed health conditions, self reported body mass index, drinking behaviour, smoking status, self reported health, value of primary residence, total non–housing wealth and total household income.

3.3 Methodology

3.3.1 K–means clustering for longitudinal data

In this chapter we use k–means clustering for longitudinal data, a non–parametric data driven partitioning method technique (Genolini et al., 2013; Genolini et al., 2015). With this method one can model the evolution of variable trajectories over time singularly or jointly and thereby cluster individuals with similar trajectories. For an individual i at time j, a single variable p trajectory is defined as

$$y_{1.p} = y_{ijp}, y_{i2p}, \ldots, y_{inp}, \tag{3.1}$$

where $i = 1, ..., N$ and $j = 1, ..., n$.

In the case of joint trajectories, the matrix that defines the joint trajectory for an individual i is

$$y_i = \begin{pmatrix} y_{i1a} & y_{i2a} & \cdots & y_{ina} \\ y_{i1b} & y_{i2b} & \cdots & y_{inb} \\ \vdots & \vdots & \ddots & \vdots \\ y_{i1p} & y_{i2p} & y_{i3p} & y_{inp} \end{pmatrix}, \tag{3.2}$$

[1]Publicly available HRS dataset: https://hrs.isr.umich.edu/data-products

where a, b and p are single trajectories of different variables.

The sequence

$$y_{ij} = \begin{pmatrix} y_{ija} \\ y_{ijb} \\ \dots \\ y_{ijp} \end{pmatrix} \tag{3.3}$$

represents the i^{th} individual's state at occasion j. The distance $d(y_1, y_2)$ between two joint trajectories y_1 and y_2 is given by calculating the distance between the fixed columns j

$$d(y_1, y_2) = \|(d_1.(y_{11.}, y_{21.}), d_2.(y_{12.}, y_{22.}), \dots, d_j.(y_{1j.}, y_{2j.}))\|. \tag{3.4}$$

We use the Euclidean distance,

$$d(y_1, y_2) = \sqrt[p]{\sum_{j=1}^{n} (d_j.(y_{1j}, y_{2j}))^p}. \tag{3.5}$$

Data are normalised so that all variables are on a comparable scale. The number of clusters chosen is based on concordance of multiple criteria and expert judgement (Caliñski & Harabasz, 1974; Kryszczuk & Hurley, 2010; Ray & Turi, 1999).

While there are multiple variables available to cluster the data, the most meaningful results from our exploratory analysis were from body mass index (a continuous variable) and self reported health (an ordinal variable). Therefore, in this thesis we focus on the clustering results from these two variables. Binary variables such as doctor diagnosed conditions and health behaviour are not amenable k–means clustering. Nominal variables such as marital status do not provide meaningful distances between categories. The clustering of wealth and income variables did not provide well separated risk groups. Hence, we use single and joint trajectories for body mass index and self-reported health to cluster the data. Furthermore, single variable trajectories of BMI and self reported health trajectories are used in the epidemiology literature to create different mortality and morbidity risk profiles using a variety of techniques that exclude k–means clustering (Cheng et al., 2021; Zajacova et al., 2015; H. Zheng et al., 2013). Our work also extends

this literature.

3.3.2 Marginal models

3.3.2.1 Model specification

Marginal models are an extension of generalised linear models (Nelder & Wedderburn, 1972) to longitudinal data. Marginal models allow one to make inferences about population means over time but do not account for random effects. Similar to standard linear regression models, the systematic component of a generalised linear model η_i linearly combines the covariates X_i and unknown regression coefficients β_i. This implies that the mean response can be expressed as a simple weighted sum of the regression parameter, β

$$\eta_i = \beta_1 X_{i1} + \beta_2 X_{i1} + \ldots + \beta_p X_{ip}.$$

The mean of each response $\mu_{ij} = E(Y_{ij}|X_{i1}, \ldots, X_{in_i}) = E(Y_{ij}|X_{ij}) = Pr(Y_{ij} = 1|X_{ij})$ is assumed to depend on the covariates through a known link function where $Y_{ij} = (Y'_{i1}, Y'_{i1}, \ldots, Y'_{in_i})$. The prime symbol denotes the transpose of the matrix. The link function applies a suitable transformation of the mean response and then links the covariates to the transformed mean of the distribution of responses. The transformed mean of the distribution of responses $g(\mu_{ij})$ is defined as follows:

$$g(\mu_{ij}) = \eta_{ij} = X'_{ij}\beta,$$

where the prime symbol denotes the transpose of the matrix. While we can use either the logit link function or the probit link function, we have chosen the logit link since it has been shown to have a better fit in an earlier study using the HRS data (M. Xu et al., 2019). The link function ensures that the predicted probabilities are within the range 0 to 1 for a binary response. Therefore we have the following log logit equation

$$\log\left(\frac{\mu_{ij}}{1-\mu_{ij}}\right) = \eta_{ij} = X'_{ij}\beta.$$

The conditional variance of each Y_{ij}, given the covariates, is assumed to depend on the

mean according to

$$Var(Y_{ij}|X_{ij}) = \phi\upsilon(\mu_{ij}),$$

where $\phi\upsilon(\mu_{ij})$ is a known "variance function" (that is, a known function of the mean, υ_{ij}) and ϕ is a scale parameter that may be known or may need to be estimated. Since the response variable is a Bernoulii distribution, we set $\upsilon(\mu_{ij}) = \mu_{ij}(1 - \mu_{ij})$ and $\phi = 1$.

For balanced longitudinal designs, a separate scale parameter ϕ could be estimated at each occasion; alternatively, the scale parameter could depend on times of measurement, with $\phi(t_{ij})$, being some parametric function of t_{ij}.

The within-subject association among the vector of repeated responses, given the covariates, is assumed to be a function of an additional set of parameters α_{jk}, whereby, α_{jk} is assumed to have an unstructured log odds ratio pattern

$$\alpha_{jk} = Corr(Y_{ij}, Y_{ik}).$$

3.3.2.2 Model estimation

We use the generalised estimating equations (GEE) approach to estimate β (Zeger & Liang, 1986). This arises from minimising the function

$$\sum_{i=1}^{N}(y_i - \mu_i(\beta))' V_i^{-1}(y_i - \mu_i(\beta)) \tag{3.6}$$

with respect to β and V_i. The corresponding covariance matrix is constructed as the product of standard deviations and correlations

$$V_i = A_i^{\frac{1}{2}} Corr(Y_i) A_i^{\frac{1}{2}}, \tag{3.7}$$

where A_i is a diagonal matrix with $Var(Y_{ij}|X_{ij}) = \phi\upsilon(\mu_{ij}))$ along the diagonal and $A_i^{\frac{1}{2}}$ is a diagonal matrix with the standard deviations,$\sqrt{\phi\upsilon\mu_{ij}}$.

Note that

$$\mu_{ij} = \mu_{ij}(\beta) = g^{-1}(X'_{ij}\beta). \tag{3.8}$$

If the minimum exists, then it must solve the following generalised estimations equations

$$\sum_{i=1}^{N} D'_i V_i^{-1}(y_i - \mu_i) = 0, \tag{3.9}$$

where the derivative matrix $D_i = \frac{\partial \mu_i}{\partial \beta}$. The pairwise within-subject association is estimated by

$$\hat{\alpha}_{jk} = \frac{1}{N} \sum e_{ij} e_{ik}. \tag{3.10}$$

We consider the following three models:

- Model 1: All variables excluding cluster allocation
- Model 2: All variables including cluster allocation
- Model 3: All variables including cluster allocation excluding BMI and self reported health.

3.3.3 Implications for annuity pricing

The conditional probability that individual i dies in period j given that the individual is alive at the start of the period is

$$p(t_{ij}) = Pr[t_{ij} = t | t_{ij} \geqslant t, X_{ij}].$$

The log-odds for a logit link are

$$\log \frac{p(t_{ij})}{1 - p(t_{ij})} = E(Y_{ij}|X_{ij}) = (X'_{ij}\beta).$$

Therefore,

$$p(t_{ij}) = \frac{1}{1 + \exp -(X'_{ij}\hat{\beta})}.$$

Since each time interval j is 2 years, the probability of death in the preceding 2 years is

$$ {}_2q_{x-2} = q_{x-2} + (1 - q_{x-2})q_{x-1} \approx q_{x-2} + q_{x-1}. \tag{3.11} $$

This means that

$$ q_{x+n} =_2 q_{x+n-1} - q_{x+n-1}. \tag{3.12} $$

If $q_{x+1} = cq_x$ and $q_{x+2} = cq_{x+1}$ then

$$ {}_2q_x \approx (1 + c)q_x \tag{3.13} $$

and

$$ {}_2q_{x+1} \approx c(1 + c)q_x. \tag{3.14} $$

Therefore, the estimated annual probability of death at age x is

$$ q_x \approx \frac{{}_2q_x}{1 + \frac{{}_2q_{x+1}}{{}_2q_x}}. \tag{3.15} $$

We project annual probabilities of death after ages $65, 70, 75, 80$ and 85 using the topping out technique by Haberman & Renshaw (2009) which is an adaptation of the Coale & Kisker (1990) method that assumes that the force of mortality $\mu_{x+j,t+j}$ increases with age. Therefore,

$$ \begin{aligned} u_j = \log(\mu_{x+j,t+j}) &= a + bj + cj(j + 1) \\ q_{x+j,t+j} &= 1 - \exp(-\mu_{x+j,t+j}) \end{aligned} \tag{3.16} $$

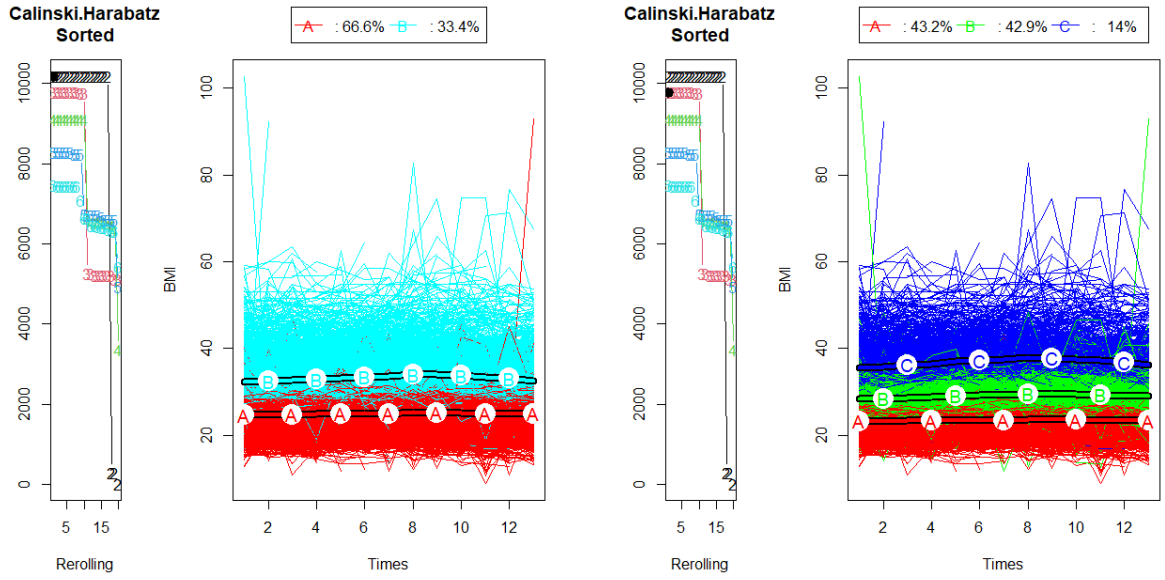
for $j = -1, 0, \ldots, \omega - x$ where $t = 2016, \omega = 110, x = 65, 70, \ldots, 85$ and $\mu_{110,2016+j} = 0.7$ and $q_{x+j,t+j}$ is derived from Equation (3.15). All our analysis is performed in R (R Core Team, 2022).

3.4 Results

3.4.1 Clustering results

3.4.1.1 Body mass index trajectories

On the right of Figure 3.1(a) are the two clusters and the mean body mass index trajectories of the 9,815 subjects. On the left, the black dot represents the partition which is stable after re-rolling 17 times. There are 66.6% individuals in cluster A and 33.4% individuals in Cluster B. Cluster A (red trajectories) has steady low BMIs that are associated with normal and underweight individuals. Cluster B (blue trajectories) has high and increasing BMIs which taper off slightly after eight waves. Some individuals have BMI's greater than 50 or less than 15 which might suggest that some of the data are entered incorrectly. However, excluding individuals with extreme values of BMI did not change cluster allocation so we did not discard individuals with extreme values.



(a) Mean body mass index trajectories with two clusters

(b) Mean body mass index trajectories with three clusters

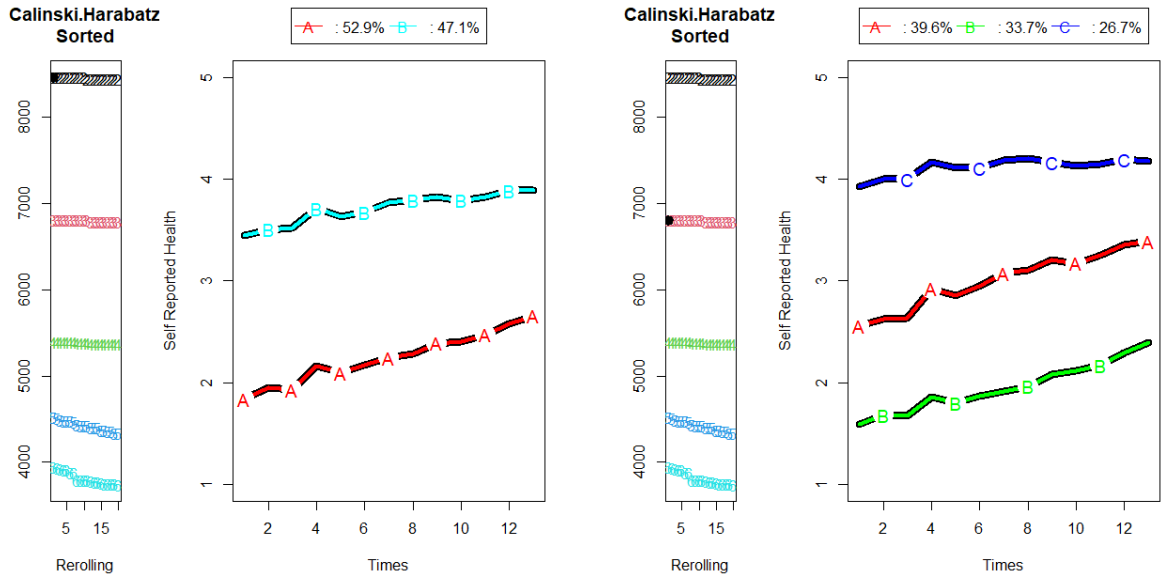
Figure 3.1: Single variable mean body mass index trajectories using Calinski Harabatz criterion

Figure 3.1(b) shows three mean body mass trajectories. Individuals are assigned to the following partitions: 43.2% in cluster A, 42.9% in Cluster B and 14% in Cluster C. Cluster B and Cluster C have similar BMI trajectories that are high and increasing. However,

individuals in Cluster B are in the overweight category and those in Cluster C fall into the World Health Organisation (WHO) obesity classes I, II and III.

3.4.1.2 Self reported health trajectories

Figure 3.2(a) shows 2 trajectories for self reported health. In the HRS, Individuals assess their health on a scale of 1–5 where 1 is excellent, 2 is very good, 3 is good, 4 is fair and 5 is poor. There are 52.9% individuals in cluster A and 47.1% individuals in Cluster B. Trajectory A shows individuals with very good health but whose health deteriorates over time towards good health. Trajectory B presents individuals initially in good health but whose health worsens over time toward fair health. Figure 3.2(b) shows 3 trajectories for self reported health. There are 39.6% individuals in Cluster A, 33.7% individuals in Cluster B and 26.7% in Cluster C. Trajectory A presents individuals initially in good health but whose health worsens over time toward fair health. Trajectory B (in green) shows individuals with very good health but whose health deteriorates over time towards good health. Trajectory C shows individuals whose fair health state is constant over time.

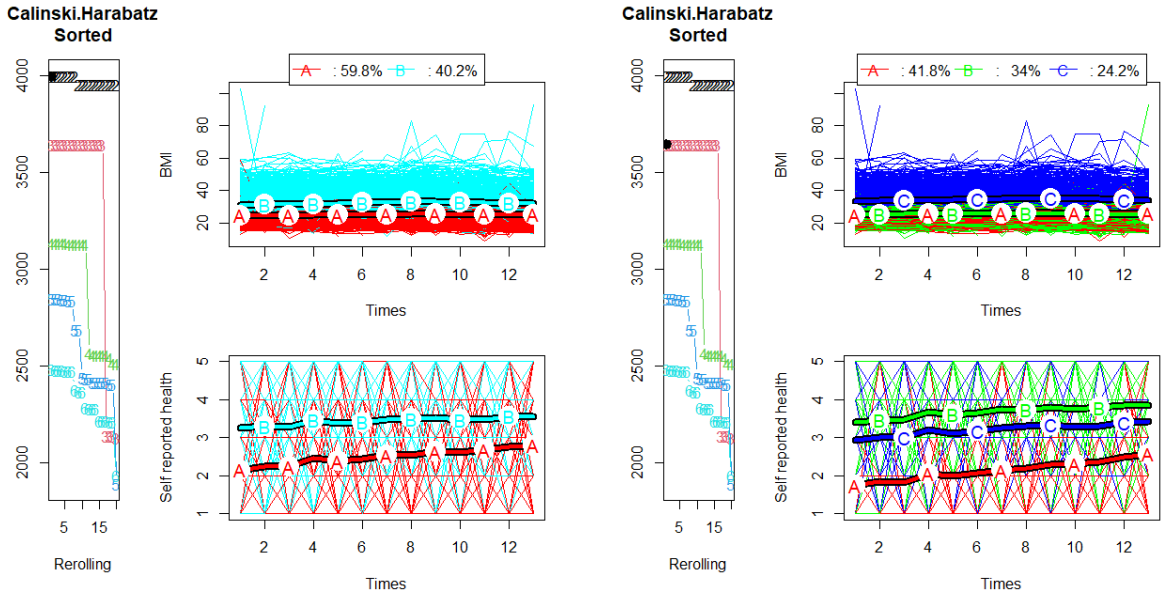


(a) Mean self reported health trajectories with 2 clusters **(b)** Mean self reported health with 3 clusters

Figure 3.2: Single variable mean self reported health trajectories using Calinski Harabatz criterion

3.4.1.3 Joint body mass index and self reported health trajectories

Figure 3.3 shows the joint trajectories for body mass index and self reported health with 2 clusters. There are 59.8% subjects in Cluster A and 40.2% in Cluster B. The partitioning shown by BMI results is similar to Figure 3.1. However, the population in Cluster A has reduced from 66.6% to 59.8% while Cluster B's size has increased to 40.2% from 33.4%. For self reported health, the segmentation is similar to results in Figure 3.2. The population in Cluster A has increased from 52.9% to 59.8% while Cluster B's size has decreased to 40.2%.



(a) Mean joint trajectories for body mass index and self reported health trajectories with 2 clusters

(b) Mean joint trajectories for body mass index and self reported health trajectories with 3 clusters

Figure 3.3: Mean joint trajectories for body mass index and self reported health trajectories

Figure 3.3(b) shows the joint trajectories for body mass index and self reported health with three clusters. There are 41.8% individuals in Cluster A, 34% in Cluster B and 24.2% in Cluster C. Unlike Figure 3.3(a) this segmentation is able to split individuals who have steady normal BMIs by their self reported health. We can clearly see that decisions based on an assumption that of steady and normal weight are incorrect as they mask the effects of deteriorating health. Remarkably, individuals in Cluster B have the worst health status that is higher than obese individuals in Cluster C.

3.4.1.4 Cluster summaries

Table 3.1 provides summaries of the characteristic features of the different clusters and the overall dataset in 1992 based on the segmentation from joint body mass index and self reported health trajectories in Figure 3.3(b). All the individuals have roughly the same average age of 55 years and were born in 1936. Females represent 53% of the overall population and there are more females than males in each cluster. However, Cluster C has the highest proportion of females than any other cluster. Body mass index (BMI) is calculated as weight in kilograms divided by height in meters squared (kg/m^2). We classify BMI using the World Health Organisation recommendations on nutritional status. The majority (40.95%) are overweight, that is a 25.0 <BMI< 29.9 and only 1.4% are underweight with a BMI of less than 18.5. Roughly a third of the sample have a normal weight 18.5 <BMI< 24.9. Almost a quarter of the group are obese which is further classified into three groups: Obesity Class I is 30.0 <BMI< 34.9, Obesity Class II is 35.0 <BMI< 39.9 and Obesity Class III is BMI> 40. The majority of individuals in Cluster A have a normal weight (48.05%) with only (5.68%) being obese. Most of the individuals in Cluster B are overweight (49.55%) and a significant proportion (39.20%) have normal weight. More than three quarters (77.34%) of Cluster C are in obesity classes I,II and III.

The average value of the primary residence in Cluster A is 80% higher than Cluster B. Average total non–housing wealth in Cluster A is at least double the values in Clusters B and C. Both the total household income in Clusters B and C are lower than the overall average. However, the average value of primary residence, household income and total household income in Cluster C is at least 12% higher than the same values in Cluster B. Cluster A has the highest proportion (48.17%) of individuals with “some college” and “college and above education”. Clusters B and C have similar levels of individuals who did not attend college. Approximately 68% have at least graduated from high school. However, the proportion of people who left high school without graduating are over represented in Cluster B. Cluster A has higher than average rates of people who are married. Individuals who have been divorced, never married, partnered, separated or widowed are higher than average in Clusters B and C.

Overall, Clusters B and C have higher than average rates of any disease. People

in Cluster C have the highest rates of high blood pressure, diabetes and arthritis $(48.63\%, 16.48\%, 43.30\%)$, respectively. Individuals in Cluster A have the lowest rates of any disease. Individuals in Cluster B and Cluster C have 6 times and 7 times the proportion of diabetes in Cluster A. Cluster B has double the average rates of lung disease (10.23%) and stroke (4.68%), respectively. Psychiatric problems (11.73%) are most common in Cluster B. Cluster B has the highest level of current smokers (39.17%) and Cluster A has the highest number of people who have ever drunk alcohol (68.99%). Cluster A has the highest levels of people in excellent or very good health (84.33%). Clusters B and C have (40.97%) and (29.21%) in fair or poor health, respectively. This implies that Cluster B has the worst health status overall and these rates are higher than the overall average rates for fair and poor health.

Table 3.1: Baseline summary statistics for Wave 1 in 1992 based on the segmentation from joint body mass index and self reported health trajectories in Figure 3.3(b)

Description	**Cluster A**[a]	**Cluster B**[b]	**Cluster C**[c]	**Overall**[d]
Clusters				
Cluster	41.82%	33.97%	24.21%	100%
Socio–demographic				
Age in years	55.86	56.37	55.77	56.01
Year of birth	1936.33	1935.83	1936.42	1936.18
Wealth and income				
Value of primary residence	$111,313	$61,314	$69,416	$84,187
Total non-housing wealth	$224,732	$97,040.25	$109,163	$153,381
Total household income	$60,894	$33,748	$39,819	$46,571
Gender				
Female	52.67%	50.69%	57.32%	53.12%
Male	47.33%	49.31%	42.68%	46.88%
Education				
College and above	25.29%	9.57%	11.41%	16.59%
GED[*]	4.02%	5.82%	6.02%	5.11%
High School Graduate	34.84%	29.51%	33.29%	32.65%
Lt High-school[†]	12.98%	39.62%	32.37%	26.72%
Some college	22.87%	15.48%	16.92%	18.92%
Marital status				
Divorced	9.67%	12.27%	10.90%	10.85%
Married	79.39%	67.43%	70.12%	73.08%
Married, spouse absent	0.58%	0.45%	0.34%	0.48%
Never married	2.34%	4.89%	4.17%	3.65%

Table 3.1: Baseline summary statistics for Wave 1 in 1992 based on the segmentation from joint body mass index and self reported health trajectories in Figure 3.3(b) *(continued)*

Description	Cluster A[a]	Cluster B[b]	Cluster C[c]	Overall[d]
Partnered	2.24%	2.76%	2.69%	2.53%
Separated	1.71%	4.71%	3.41%	3.14%
Widowed	4.07%	7.50%	8.38%	6.28%
Doctor diagnosed health conditions				
High blood pressure	20.95%	40.46%	48.70%	34.29%
Diabetes	2.34%	13.08%	16.50%	9.41%
Cancer	3.17%	6.60%	4.67%	4.70%
Lung disease	1.34%	10.23%	4.97%	5.24%
Heart problems	3.70%	17.04%	12.29%	10.31%
Stroke	0.56%	4.68%	2.57%	2.45%
Psychiatric problems	2.10%	11.73%	7.95%	6.79%
Arthritis	21.58%	40.94%	43.35%	33.43%
Health behaviour				
Ever drinks alcohol	68.99%	54.29%	53.91%	60.35%
Ever smoked	59.00%	70.55%	61.49%	63.53%
Smokes now	21.34%	39.17%	21.30%	27.39%
Self reported health				
Excellent	43.02%	3.51%	10.31%	21.68%
Fair	1.29%	27.11%	19.28%	14.42%
Good	14.20%	38.54%	36.15%	27.78%
Poor	0.19%	16.98%	9.97%	8.26%
Very good	41.29%	13.86%	24.28%	27.86%
Body mass index				
Normal weight	48.04%	39.20%	1.05%	33.66%
Obesity class I	5.51%	7.98%	49.03%	16.88%
Obesity class II	0.15%	0.39%	18.90%	4.77%
Obesity class III	0.02%	0.09%	9.51%	2.34%
Overweight	45.21%	49.55%	21.51%	40.95%
Underweight	1.07%	2.79%	0.00%	1.40%

[a] Cluster A: normal, stable BMI and declining very good health;

[b] Cluster B: normal, stable BMI and declining fair health;

[c] Cluster C: high, increasing BMI and declining good health;

[d] Overall is the entire dataset without clustering;

[*] GED means graduated high school by taking a General Education Development Test;

[†] Lt-High school means left high school without graduation.

3.4.1.5 Dependence between cluster membership and predictors

To test if there is an association between cluster membership and each socioeconomic characteristic, we conduct One–Way Analysis of Variance (ANOVA) tests with Holm corrections ($p < 0.05$) to adjust for multiple testing (Fisher, 1919; Holm, 1979). Table 3.2 shows the results for group comparisons for the continuous variables. For statistically significant variables, we carry out pairwise comparisons with Tukey's Honest Significant Difference (HSD) test and a 95% family wise confidence level to show the actual differences between clusters (Tukey, 1949). There is strong evidence that the mean age ($F = 76.47, p < 0.0001$), value of primary residence ($F = 1375.36, p < 0.0001$), total non–housing wealth ($F = 579.47, p < 0.0001$) and total household income ($F = 6.74, p = 0.0012$) are different across each cluster.

Table 3.3 shows the results for overall group and pairwise comparisons for categorical variables using Chi-squared (χ^2) tests with Holm corrections ($p < 0.05$) to adjust for multiple testing (Pearson, 1900). Since the adjusted p–values are all less than 0.05 for all the predictors, we reject the null hypothesis that there is no dependence between cluster membership and each predictor. We conclude that there is very strong evidence ($p < 0.0001$) that cluster membership is associated with gender, education, marital status, drinking behaviour, smoking behaviour and all the doctor diagnosed conditions. Consequently, pairwise tests specify which clusters are different from each other for all predictors. The null hypothesis is that the proportions of each predictor are the same in each cluster. We strongly reject the null hypothesis ($p < 0.0001$) for pairwise comparisons of all the variables except drinking behaviour ($p = 0.2682$).

Table 3.2: Comparison tests for numerical variables: age, value of primary residence, total non–housing wealth and total household income using one-way ANOVA test and Tukey's HSD test

Predictor	**Term**	**F value**	**Pr(>F)†**	**Adjusted p–value**	**B–A[1]**	**C–A[2]**	**C–B[3]**
Age		76.4724	0.0000	0.0000			
	Difference in means				-0.7765	-0.4625	0.3140
	Lower bound 95% CI				-0.9262	-0.6237	0.1405
	Upper bound 95% CI				-0.6267	-0.3012	0.4875
	Adjusted p-value				0.0000	0.0000	0.0001
Value of primary residence		1,375.3582	0.0000	0.0000			
	Difference in means				-85,885.1285	-73,760.0160	12,125.1125
	Lower bound 95% CI				-90,096.9118	-78,295.9778	7,243.9864
	Upper bound 95% CI				-81,673.3451	-69,224.0541	17,006.2386
	Adjusted p-value				0.0000	0.0000	0.0000
Total non-housing wealth		579.4693	0.0000	0.0000			
	Difference in means				-241,345.0243	-208,593.3787	32,751.6456
	Lower bound 95% CI				-259,616.6825	-228,271.3960	11,576.2288
	Upper bound 95% CI				-223,073.3662	-188,915.3615	53,927.0624
	Adjusted p-value				0.0000	0.0000	0.0008
Total household income		6.7363	0.0012	0.0012			
	Difference in means				-23,450.4145	-36,939.1088	-13,488.6943
	Lower bound 95% CI				-46,506.5967	-61,769.9122	-40,208.9986
	Upper bound 95% CI				-394.2324	-12,108.3055	13,231.6100
	Adjusted p-value				0.0451	0.0014	0.4632

[1] Tukey's HSD test for pairwise comparison between Cluster B and Cluster A with a 95% family wise confidence level;

[2] Tukey's HSD test for pairwise comparison between Cluster C and Cluster A with a 95% family wise confidence level

[3] Tukey's HSD test for pairwise comparison between Cluster C and Cluster B with a 95% family wise confidence level;

† One–Way Anova test to compare the effect of cluster membership on predictors with Holm corrections ($p<0.05$).

Table 3.3: Comparison tests for categorial variables: gender, education, marital status, doctor diagnosed conditions and health behaviour using χ^2 tests with Holm corrections (p<0.05)

Predictor	**Term**	χ^2	**p–value**	**Adjusted p–value**	**B–A[1]**	**C–A[2]**	**C–B[3]**
Gender		169.7636	<0.0001	<0.0001			
	Raw p-value				0.0391	0.0000	0.0000
	Adjusted p-value				0.0391	0.0000	0.0000
Education		47,330.3578	<0.0001	<0.0001			
	Raw p-value				0.0000	0.0000	0.0000
	Adjusted p-value				0.0000	0.0000	0.0000
Marital Status		1,381.3050	<0.0001	<0.0001			
	Raw p-value				0.0000	0.0000	0.0000
	Adjusted p-value				0.0000	0.0000	0.0000
High blood pressure		3,930.0947	<0.0001	<0.0001			
	Raw p-value				0.0000	0.0000	0.0000
	Adjusted p-value				0.0000	0.0000	0.0000
Diabetes		5,850.3772	<0.0001	<0.0001			
	Raw p-value				0.0000	0.0000	0.0000
	Adjusted p-value				0.0000	0.0000	0.0000
Cancer		193.6758	<0.0001	<0.0001			
	Raw p-value				0.0000	0.0106	0.0000
	Adjusted p-value				0.0000	0.0106	0.0000
Lung disease		2,695.0048	<0.0001	<0.0001			
	Raw p-value				0.0000	0.0000	0.0000
	Adjusted p-value				0.0000	0.0000	0.0000
Heart problems		2,860.0503	<0.0001	<0.0001			
	Raw p-value				0.0000	0.0000	0.0000
	Adjusted p-value				0.0000	0.0000	0.0000

Table 3.3: Comparison tests for categorial variables: gender, education, marital status, doctor diagnosed conditions and health behaviour using χ^2 tests with Holm corrections (p<0.05) *(continued)*

Predictor	**Term**	χ^2	**p–value**	**Adjusted p–value**	**B–A[1]**	**C–A[2]**	**C–B[3]**
Stroke		1,388.0545	<0.0001	<0.0001			
	Raw p-value				0.0000	0.0000	0.0000
	Adjusted p-value				0.0000	0.0000	0.0000
Psychiatric problems		3,088.0434	<0.0001	<0.0001			
	Raw p-value				0.0000	0.0000	0.0000
	Adjusted p-value				0.0000	0.0000	0.0000
Arthritis		3,536.8922	<0.0001	<0.0001			
	Raw p-value				0.0000	0.0000	0.0000
	Adjusted p-value				0.0000	0.0000	0.0000
Ever drinks alcohol		2,726.0009	<0.0001	<0.0001			
	Raw p-value				0.0000	0.0000	0.2682
	Adjusted p-value				0.0000	0.0000	0.2682
Ever smoked		641.3706	<0.0001	<0.0001			
	Raw p-value				0.0000	0.0000	0.0000
	Adjusted p-value				0.0000	0.0000	0.0000
Smokes now		2,031.3510	<0.0001	<0.0001			
	Raw p-value				0.0000		0.0000
	Adjusted p-value				0.0000		0.0000

[1] χ^2 test for pairwise comparison between Cluster B and Cluster A;

[2] χ^2 test for pairwise comparison between Cluster C and Cluster A;

[3] χ^2 test for pairwise comparison between Cluster C and Cluster B.

3.4.2 Mortality experience in each cluster

Table 3.4 presents the mortality experience of each cluster using training data. Cluster B has the worst mortality rate (57.08%) which is higher than the overall mortality rate (38.82%). Clusters A (21.49%) has the lowest mortality rate. Cluster C (42.74%) has a mortality rate lower than the overall mortality rate but higher than Cluster A's.

Table 3.4: Crude mortality rates from 1992–2016 using training data

Cluster	Number Alive	Percentage	Number Dead	Mortality rate
A	3,266	41.59%	702	21.49%
B	2,691	34.27%	1,536	57.08%
C	1,895	24.13%	810	42.74%
Overall	7,852	100.00%	3,048	38.82%

[a] Cluster A: normal, stable BMI and declining very good health;
[b] Cluster B: normal, stable BMI and declining fair health;
[c] Cluster C: high, increasing BMI and declining good health;
[*] Overall is the entire dataset without clustering.

3.4.3 Marginal model results

Firstly, we fit a marginal model with mortality as the response and clusters as the predictor. We reject the null hypothesis that the proportion of deaths in each cluster is the same and find very strong evidence ($\chi^2(2) = 798, p < 0.0001$) of an association between mortality and cluster allocation. Pairwise, we have very strong evidence of differences in mortality amongst clusters ($p < 0.0001$). The mortality odds are $0.303(0.274, 0.334)$ times lower in Cluster A than in Cluster B. They are $0.461(0.412, 0.516)$ times lower in Cluster A than in Cluster C, and $1.525(1.391, 1.671)$ times higher in Cluster B than C.

Table 3.5 presents the regression coefficients for Model 1 which includes all the predictors without cluster allocation. The goal of this baseline analysis is to identify the variables impact mortality and quantify their impact on mortality in terms of odds ratios.

Table 3.5: Odds ratios for mortality fitted using Model 1

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Beta						
Intercept	0.007	0.006	0.009	1 686.9720	0.0000	***
Education (ref: College and above)						
Lt High-school†	0.940	0.816	1.083	0.7293	0.3931	
GED‡	0.920	0.756	1.119	0.6972	0.4037	
High School Graduate	1.072	0.939	1.224	1.0572	0.3038	
Some college	1.058	0.919	1.218	0.6204	0.4309	
Gender (ref: Female)						
Male	1.516	1.397	1.646	99.4247	0.0000	***
Sociodemographic						
Age	1.671	1.601	1.744	550.3427	0.0000	***
Marital status (ref: Married)						
Married spouse absent	1.868	1.478	2.360	27.4343	0.0000	***
Partnered	1.363	1.099	1.690	7.9640	0.0048	**
Separated	1.258	0.990	1.598	3.5100	0.0610	.
Divorced	1.165	1.032	1.315	6.0808	0.0137	*
Separated/divorced	1.394	1.013	1.918	4.1519	0.0416	*
Widowed	1.070	0.964	1.189	1.6150	0.2038	
Never married	1.419	1.174	1.714	13.1255	0.0003	***
Doctor diagnosed health conditions						
Psychiatric problems	1.155	1.052	1.267	9.1438	0.0025	**
High blood pressure	1.185	1.088	1.290	15.2774	0.0001	***
Diabetes	1.431	1.316	1.556	70.5618	0.0000	***
Cancer	1.722	1.573	1.885	139.2812	0.0000	***
Lung disease	1.428	1.303	1.565	57.8264	0.0000	***
Heart problems	1.316	1.214	1.428	44.0040	0.0000	***
Stroke	1.403	1.266	1.553	42.0635	0.0000	***
Arthritis	0.911	0.837	0.992	4.5754	0.0324	*
Health behaviour						
Ever drinks alcohol	0.816	0.752	0.885	24.0349	0.0000	***
Ever smoked	1.398	1.276	1.532	51.5381	0.0000	***
Smokes now	1.472	1.331	1.627	56.5806	0.0000	***
Body mass index (ref: Normal weight)						
Underweight	1.941	1.650	2.283	64.1075	0.0000	***
Overweight	0.716	0.655	0.784	53.0161	0.0000	***
Obesity class I	0.615	0.550	0.687	72.9836	0.0000	***

Table 3.5: Odds ratios for mortality fitted using Model 1 *(continued)*

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Obesity class II	0.710	0.606	0.831	18.0449	0.0000	***
Obesity class III	0.846	0.698	1.026	2.8777	0.0898	.
Self reported health (ref: Excellent)						
Very good	1.202	0.960	1.504	2.5681	0.1090	
Good	1.672	1.346	2.077	21.5397	0.0000	***
Fair	2.960	2.377	3.687	93.8567	0.0000	***
Poor	5.322	4.243	6.675	209.1711	0.0000	***
Wealth and income						
Value of primary residence	0.822	0.762	0.887	25.5420	0.0000	***
Total non-housing wealth	1.008	0.963	1.055	0.1230	0.7258	
Total household income	0.748	0.309	1.814	0.4122	0.5209	

Note: Significance levels

*** p<0.001; ** p<0.01; * p<0.05; . p<0.1; p<1.

† GED means graduated high school by taking a General Education Development Test;

‡ Lt-High school means left high school without graduation.

All the variables have the expected impact on mortality except for education, total non–housing wealth, total household income and body mass index. There is not enough evidence to conclude that having left high school decreases mortality risk by 6% ($OR = 0.94(0.82, 1.09), p = 0.3931$) or graduating with a GED decreases your mortality risk by 8% ($OR = 0.92(0.756, 1.119), p = 0.4037$) when compared to an average individual with an education level of college and above. The mortality odds of being underweight or overweight when compared to the average patient who has a normal weight are 96% higher ($CI = (1.650, 2.283), p < 0.0001$) and 28% lower ($CI = (0.655, 0.784), p < 0.0001$). This indicates that being underweight presents a greater mortality risk than being overweight. An obese individual (Class I, II or III) has lower mortality odds $39\%, 29\%, 15\%$ than the average individual with normal weight, respectively. These results are consistent with the literature on BMI and mortality.

Amongst doctor diagnosed health conditions: cancer, diabetes, lung disease and stroke have the highest impact on mortality. In particular, we have strong evidence that the odds of an average individual dying if the individual has cancer is 1.7 times ($CI = (1.57, 1.89), p < 0.0001$) the mortality odds of an average individual who does not have

cancer while holding all other variables constant. Having arthritis is associated with a decrease in an average person's mortality odds by 0.9 times ($(CI = (0.84, 0.99), p < 0.0001$) when compared to the mortality odds of an average individual who does not have arthritis while holding all other variables constant. The odds of an average individual dying if they are married and their spouse is absent is 1.8 times ($CI = (1.57, 1.88), p < 0.0001$) when compared to the mortality odds of an average individual who is married while holding all other variables constant. The odds of an average individual dying if they have poor health status is 4.6 times $CI = (4.24, 6.68), p < 0.0001$ when compared to the mortality odds of an average individual who has excellent health all other things being equal.

Table 3.6: Odds ratios for mortality fitted using Model

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Beta						
Intercept	0.007	0.005	0.009	1 682.2840	0.0000	***
Education (ref: College and above)						
Lt High-school[†]	0.916	0.795	1.055	1.4753	0.2245	
GED[‡]	0.896	0.735	1.092	1.1876	0.2758	
High School Graduate	1.061	0.929	1.212	0.7591	0.3836	
Some college	1.050	0.911	1.209	0.4522	0.5013	
Gender (ref: Female)						
Male	1.515	1.395	1.644	98.2199	0.0000	***
Clusters (ref: Cluster A[a])						
Cluster B[b]	1.238	1.101	1.393	12.7302	0.0004	***
Cluster C[c]	1.475	1.264	1.721	24.4355	0.0000	***
Sociodemographic						
Age	1.696	1.623	1.772	559.2876	0.0000	***
Marital status (ref: Married)						
Married spouse absent	1.860	1.472	2.351	26.9545	0.0000	***
Partnered	1.373	1.108	1.702	8.3837	0.0038	**
Separated	1.255	0.988	1.594	3.4544	0.0631	.
Divorced	1.163	1.030	1.314	5.9339	0.0149	*
Separated/divorced	1.383	1.005	1.903	3.9619	0.0465	*
Widowed	1.063	0.957	1.182	1.3151	0.2515	
Never married	1.391	1.150	1.681	11.6154	0.0007	***
Doctor diagnosed health conditions						

Table 3.6: Odds ratios for mortality fitted using Model *(continued)*

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Psychiatric problems	1.155	1.052	1.268	9.1004	0.0026	**
High blood pressure	1.177	1.081	1.282	14.0126	0.0002	***
Diabetes	1.401	1.288	1.524	61.8989	0.0000	***
Cancer	1.725	1.576	1.889	139.0086	0.0000	***
Lung disease	1.423	1.298	1.560	56.3127	0.0000	***
Heart problems	1.304	1.202	1.415	40.8241	0.0000	***
Stroke	1.392	1.257	1.543	39.9951	0.0000	***
Arthritis	0.899	0.826	0.980	5.9114	0.0150	*
Health behaviour						
Ever drinks alcohol	0.819	0.755	0.889	22.9715	0.0000	***
Ever smoked	1.400	1.277	1.534	51.6080	0.0000	***
Smokes now	1.462	1.322	1.618	54.2892	0.0000	***
Body mass index (ref: Normal weight)						
Underweight	1.932	1.642	2.272	63.1679	0.0000	***
Overweight	0.704	0.643	0.771	57.3787	0.0000	***
Obesity class I	0.531	0.462	0.611	79.1340	0.0000	***
Obesity class II	0.575	0.473	0.698	31.2048	0.0000	***
Obesity class III	0.690	0.551	0.865	10.3462	0.0013	**
Self reported health (ref: Excellent)						
Very good	1.173	0.937	1.468	1.9409	0.1636	
Good	1.515	1.214	1.890	13.5548	0.0002	***
Fair	2.562	2.034	3.227	63.9354	0.0000	***
Poor	4.599	3.624	5.835	157.7008	0.0000	***
Wealth and income						
Value of primary residence	0.829	0.769	0.895	23.4959	0.0000	***
Total non-housing wealth	1.008	0.964	1.054	0.1247	0.7240	
Total household income	0.764	0.321	1.818	0.3708	0.5426	

Note: Significance levels

*** p<0.001; ** p<0.01; * p<0.05; . p<0.1; p<1.

[a] Cluster A: normal, stable BMI and declining very good health;

[b] Cluster B: normal, stable BMI and declining fair health;

[c] Cluster C: high, increasing BMI and declining good health;

[†] GED means graduated high school by taking a General Education Development Test;

[‡] Lt-High school means left high school without graduation.

Table 3.6 presents the regression coefficients for Model 2 which includes all the important

variables from Table 3.5 and adds cluster membership. The goal of this analysis is to quantify the impact of cluster allocation on mortality. There is strong evidence that the odds of an average individual dying if they are in Cluster B are 1.238 times higher ($(CI = 1.10, 1.39), p = 0.0004$) when compared to the mortality odds of an average individual who is in Cluster A, holding all other variables constant. The odds of an average individual dying if they are in Cluster C are 1.48 times higher ($(CI = 1.26, 1.72), p < 0.0001$) when compared to the mortality odds of an average individual who is in Cluster A holding all other variables constant. All the other variables have generally similar impact as observed in Table 3.5 with minor changes in the regression coefficients and confidence intervals.

Table 3.7: Odds ratios for mortality fitted using Model 3

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Beta						
Intercept	0.007	0.006	0.009	3 521.5266	0.0000	***
Education (ref: College and above)						
Lt High-school[†]	1.023	0.888	1.179	0.1008	0.2245	
GED[‡]	0.921	0.752	1.127	0.6429	0.2758	
High School Graduate	1.070	0.936	1.222	0.9832	0.3836	
Some college	1.049	0.910	1.209	0.4387	0.5013	
Gender (ref: Female)						
Male	1.486	1.367	1.615	86.6297	0.0000	***
Clusters (ref: Cluster A)						
Cluster B[b]	1.985	1.791	2.200	170.9773	0.0004	***
Cluster C[c]	1.496	1.338	1.671	50.4198	0.0000	***
Sociodemographic						
Age	1.720	1.645	1.798	569.7416	0.0000	***
Marital status (ref: Married)						
Married spouse absent	2.211	1.729	2.827	39.9437	0.0000	***
Partnered	1.403	1.132	1.738	9.5908	0.0038	**
Separated	1.391	1.078	1.794	6.4316	0.0631	*
Divorced	1.249	1.104	1.413	12.5026	0.0149	***
Separated/divorced	1.493	1.075	2.072	5.7276	0.0465	*
Widowed	1.097	0.986	1.220	2.9034	0.2515	.
Never married	1.409	1.168	1.700	12.8578	0.0007	***
Doctor diagnosed health conditions						
Psychiatric problems	1.327	1.207	1.459	34.2215	0.0026	***
High blood pressure	1.157	1.063	1.260	11.3773	0.0002	***

Table 3.7: Odds ratios for mortality fitted using Model 3 *(continued)*

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Diabetes	1.430	1.316	1.554	70.9857	0.0000	***
Cancer	1.944	1.772	2.133	196.7051	0.0000	***
Lung disease	1.697	1.544	1.864	121.6194	0.0000	***
Heart problems	1.451	1.337	1.576	78.8145	0.0000	***
Stroke	1.607	1.448	1.784	79.4370	0.0000	***
Arthritis	0.923	0.848	1.004	3.4720	0.0150	.
Health behaviour						
Ever drinks alcohol	0.741	0.683	0.804	51.9473	0.0000	***
Ever smoked	1.396	1.273	1.530	50.3977	0.0000	***
Smokes now	1.616	1.461	1.788	86.9698	0.0000	***
Wealth and income						
Value of primary residence	0.818	0.756	0.886	24.4136	0.0000	***
Total non-housing wealth	1.018	0.970	1.069	0.5154	0.0000	
Total household income	0.413	0.127	1.346	2.1521	0.0000	

Note: Significance levels

*** p<0.001; ** p<0.01; * p<0.05; . p<0.1; p<1.

[a] Cluster A: normal, stable BMI and declining very good health;

[b] Cluster B: normal, stable BMI and declining fair health;

[c] Cluster C: high, increasing BMI and declining good health;

[†] GED means graduated high school by taking a General Education Development Test;

[‡] Lt-High school means left high school without graduation.

Table 3.7 presents the regression coefficients for Model 3 which includes all the variables from Table 3.6. The goal of this analysis is to determine the impact of cluster allocation impact when the clustering variables: body mass index and self reported health are removed. There is strong evidence that the odds of an average individual dying if they are in Cluster B is 1.99 times ($CI = (1.79, 2.20), p = 0.004$) when compared to the mortality odds of an average individual who is in Cluster A while holding all other variables constant. The odds of an average individual dying if they are in Cluster C is 1.50 times ($CI = (1.34, 1.67), p < 0.0001$) when compared to the mortality odds of an average individual who is in Cluster A while holding all other variables constant. Contrary to the results in Table 3.6, these results are more in line with the raw mortality experience shown in Table 3.4.

3.4.3.1 Model performance

Table 3.8 presents the marginal model results for all three models. We use Quasilikelihood Information Criterion (QIC) to determine the model with the best fit (Pan, 2001). Comparing Model 2 and Model 1 shows that clustering impacts mortality when controlling for socio–economic variables since it has the lowest QIC. This shows that clustering does provide additional information in the prediction of mortality. Omitting the variables used to cluster the data results in the worst model performance as shown by the high QIC in model 3 when compared to model 2. Hence, the joint trajectories of BMI and self reported health cannot accurately predict mortality in the absence of the actual BMI and self reported health variables. This proves that just relying on clusters results in the worst performance overall even though it more parsimonious.

Table 3.8: Comparison of goodness of fit using quasilikelihood information criterion

Model	QIC	Parameters
Model 1	22,003.67	37
Model 2	**21,983.42**	**39**
Model 3	22,617.05	30

Table 3.9: Odds ratios for mortality fitted using Model 2 while excluding variables with $p < 0.2$

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Beta						
Intercept	0.007	0.006	0.009	1 944.1840	0.0000	***
Gender (ref: Female)						
Male	1.506	1.387	1.635	95.3176	0.0000	***
Clusters (ref: Cluster A)						
Cluster B†	1.219	1.084	1.370	10.9564	0.0009	***
Cluster C‡	1.451	1.244	1.692	22.4239	0.0000	***
Sociodemographic						
Age	1.696	1.624	1.772	558.7873	0.0000	***
Marital status (ref: Married)						
Married spouse absent	1.851	1.463	2.340	26.4019	0.0000	***
Partnered	1.357	1.096	1.681	7.8200	0.0052	**
Separated	1.234	0.971	1.568	2.9464	0.0861	.

Table 3.9: Odds ratios for mortality fitted using Model 2 while excluding variables with $p < 0.2$ *(continued)*

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Divorced	1.172	1.039	1.322	6.6491	0.0099	**
Separated/divorced	1.379	1.002	1.899	3.8829	0.0488	*
Widowed	1.061	0.956	1.178	1.2382	0.2658	
Never married	1.385	1.144	1.676	11.1677	0.0008	***
Doctor diagnosed health conditions						
Psychiatric problems	1.146	1.043	1.259	8.0724	0.0045	**
High blood pressure	1.175	1.079	1.280	13.6446	0.0002	***
Diabetes	1.396	1.284	1.518	60.8503	0.0000	***
Cancer	1.735	1.585	1.900	141.7971	0.0000	***
Lung disease	1.415	1.291	1.552	54.7508	0.0000	***
Heart problems	1.314	1.211	1.425	43.2840	0.0000	***
Stroke	1.403	1.266	1.555	41.6205	0.0000	***
Arthritis	0.897	0.824	0.978	6.1579	0.0131	*
Health behaviour						
Ever drinks alcohol	0.825	0.761	0.895	21.6319	0.0000	***
Ever smoked	1.395	1.273	1.529	50.5611	0.0000	***
Smokes now	1.456	1.316	1.610	53.0883	0.0000	***
Body mass index (ref: Normal weight)						
Underweight	1.942	1.650	2.284	64.0333	0.0000	***
Overweight	0.705	0.644	0.772	57.0354	0.0000	***
Obesity class I	0.533	0.464	0.613	78.4823	0.0000	***
Obesity class II	0.578	0.476	0.701	30.6895	0.0000	***
Obesity class III	0.692	0.552	0.868	10.1722	0.0014	**
Self reported health (ref: Excellent)						
Very good	1.180	0.943	1.476	2.0969	0.1476	
Good	1.524	1.222	1.900	14.0369	0.0002	***
Fair	2.552	2.029	3.210	64.0277	0.0000	***
Poor	4.547	3.590	5.761	157.5238	0.0000	***
Wealth and income						
Value of primary residence	0.837	0.780	0.898	24.4468	0.0000	***

Note: Significance levels

*** p<0.001; ** p<0.01; * p<0.05; . p<0.1; p<1.

Table 3.9: Odds ratios for mortality fitted using Model 2 while excluding variables with $p < 0.2$ *(continued)*

		95% CI			
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**

[a] Cluster A: normal, stable BMI and declining very good health;

[b] Cluster B: normal, stable BMI and declining fair health;

[c] Cluster C: high, increasing BMI and declining good health;

[†] GED means graduated high school by taking a General Education Development Test;

[‡] Lt-High school means left high school without graduation.

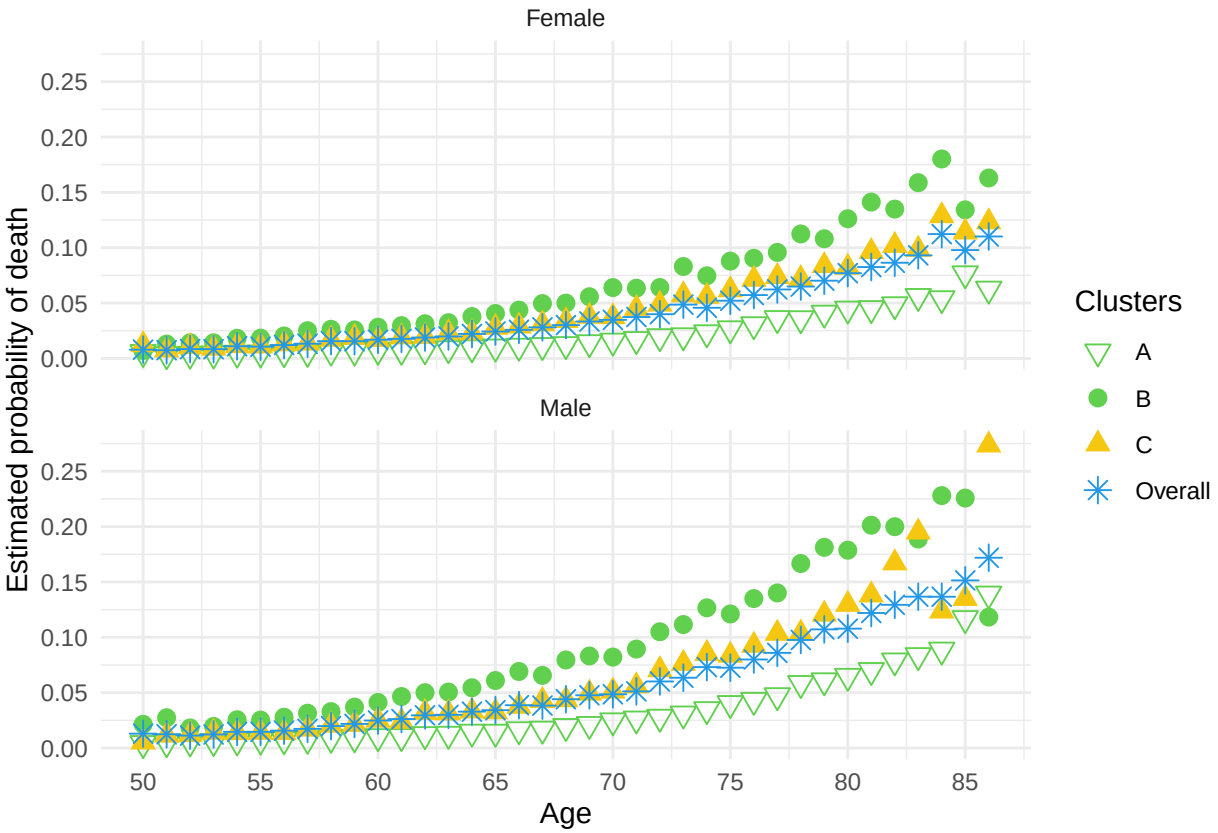


Figure 3.4: Predicted probability of death for males and females at different ages on training data

Figure 3.4 and Figure 3.5 show the predicted probabilities of death based for males and females on the training data and testing data, respectively. We use marginal Model 2 to fit the data but exclude variables with $p > 0.2$. Both training and test data exhibit the same behaviour where individuals in Cluster A have the lowest mortality rates and Cluster B has the highest mortality rates. The mortality rates of individuals in Cluster C lie in between Cluster A and Cluster B. The overall data represents males and females without

specifying the cluster to which the individual belongs. Males have higher mortality rates than females. The data are noisy after age 80 for both males and females.

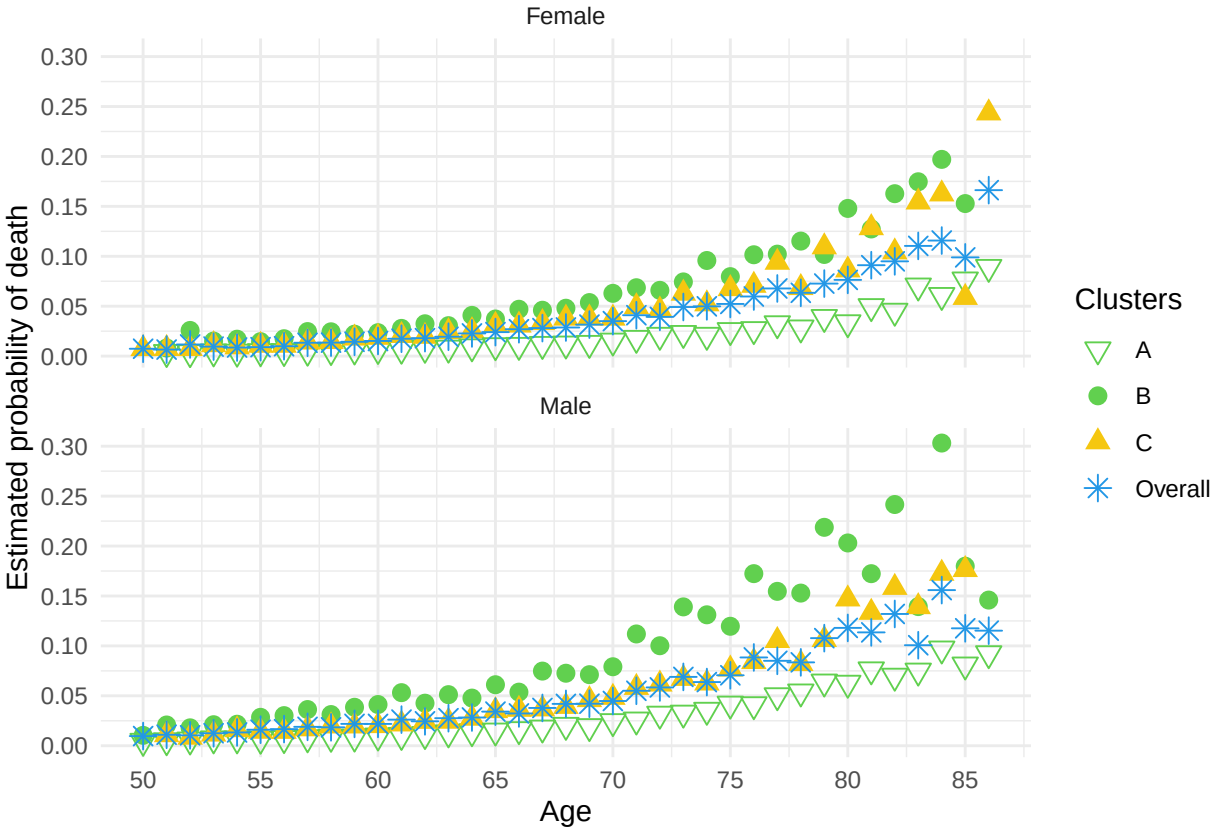


Figure 3.5: Predicted probability of death at different ages for males and females on test data

3.4.4 Pricing implications for each cluster

Table 3.10 shows the prices for whole life immediate annuities at 5 year intervals from ages 65 to 85 for males and females in the different clusters and their relative differences to the overall rate. If prices are charged at actuarially fair rates, individuals from ages 65 to 80 in Cluster A would pay between 14% and 36% above overall rates. However, individuals in Clusters B would pay between 22% and and 28% lower than the overall rates. Individuals in Cluster C would pay between 8% and 10% lower than the overall rate in Cluster C for ages 65 to 75 but 4% higher at age 80. Similar rates are observed for females with Cluster A having the most expensive annuities and the cheapest annuities in Cluster B. Generally, annuities for males are cheaper than those for females.

Table 3.10: Comparison of cluster specific prices for underwritten annuities for males and females at ages 65, 70, 75, 80 and 85 at 3% with sample with no cluster allocation

		Cluster A		Cluster B		Cluster C		
		Normal, stable BMI and very good health		Normal, stable BMI and fair health		High, increasing BMI and good health		
Age	Sex	Price	Change[a]	Price	Change[b]	Price	Change[c]	Overall[*]
65	Male	13.86	23.31%	8.83	-21.44%	10.88	-3.20%	11.24
70	Male	13.96	19.73%	8.52	-26.93%	10.40	-10.81%	11.66
75	Male	11.61	34.53%	6.42	-25.61%	7.96	-7.76%	8.63
80	Male	9.22	40.33%	4.88	-25.72%	6.80	3.50%	6.57
85	Male	5.63	6.03%	13.62	156.50%	2.63	-50.47%	5.31
65	Female	17.27	18.29%	12.08	-17.26%	14.75	1.03%	14.60
70	Female	14.99	19.92%	12.26	-1.92%	9.34	-25.28%	12.50
75	Female	11.65	16.15%	9.36	-6.68%	8.34	-16.85%	10.03
80	Female	11.91	36.27%	6.16	-29.52%	7.00	-19.91%	8.74
85	Female	12.17	77.41%	5.02	-26.82%	6.87	0.15%	6.86

[a] Difference in price of whole annuity immediate betweeen Cluster A and overall group at same age;
[b] Difference in price of whole annuity immediate betweeen Cluster B and overall group at same age;
[c] Difference in price of whole annuity immediate betweeen Cluster C and overall group at same age;
[*] Price of whole annuity immediate calculated without specifying cluster allocation

3.5 Discussion

The goal of this study is to investigate mortality heterogeneity using distinct risk profiles determined from the clustering of longitudinal data using BMI and self reported health trajectories. This is the first study in the actuarial literature using k–means clustering of longitudinal data. Our results from single variable clustering prove to be not interesting as they corroborate findings already in the epidemiological literature showing three trajectories for BMI: low and steady, medium and increasing, and high and steady. For self reported health, we find two trajectories: good health and deteriorating, plus poor health and deteriorating. However, clustering both BMI and self reported health offers quite interesting and highly nuanced trajectories which have not yet been observed elsewhere in the literature. We find this particular segmentation highly relevant for the insurance and annuities industries as it allows us to study BMI in conjunction with health status. The clusters that emerge from this analysis are different from those determined using a non–parametric hierarchical clustering of BMI trajectories estimated with Principal Analysis through Conditional Expectation using the same dataset (Zajacova et al., 2015).

The main difference is that the authors find a decreasing BMI trajectory for both males and females containing around 15% of individuals. Their stable BMI trajectory contains a much larger proportion of individuals for both males (69%) and females (81%) which further validates the result that relying on BMI trajectories is myopic and fails to fully capture the evolution of health status over time. Our results show that this stable trajectory can be further split into 2: one with deteriorating health from fair to poor and another with very good health which deteriorates to good health. The high and increasing

BMI trajectory is similar in both studies except that we find a higher proportion (24%) compared to males 16% and females 8% in Zajacova et al. (2015). We also find that dichotomising gender did not change the trajectories significantly so our clustering focuses on the overall dataset.

Our results suggest that relying on joint trajectories of BMI and self reported health to price annuities will lead to fairer pricing. This is particularly important for individuals who have poor health and normal BMI. If BMI alone was used to price annuities, these individuals would pay more whereas they are much more likely to live for shorter periods. Furthermore, education, total–housing wealth, total household income are shown to be not statistically significant when predicting mortality in Tables 3.5, 3.6 and 3.7. This suggests that their effects on mortality are already accounted for through the predictors: marital status, doctor diagnosed health conditions, health behaviour, body mass index and self–reported health.

One limitation of this study is the use of body mass index as a measure of obesity. While a high BMI is associated with high mortality, central obesity (accumulation of fat in the abdominal area) in individuals with normal BMI is associated with an even higher risk of mortality (Cerhan et al., 2014; Coutinho et al., 2013; Sahakyan et al., 2015). Using Cox-regression analysis for a large sample of over 650,000 Caucasians with a median follow up of 9 years, Cerhan et al. (2014) find that for all BMIs, having a large waist circumference increases one's risk of death. A high waist circumference is likely to decrease life expectancy in females and males by 3 years and 5 years respectively. Having a high waist to hip ratio doubles the risk of death after controlling for BMI (Sahakyan et al., 2015). This means that other measures of obesity might lead to more accurate estimates of probabilities of death at each age.

In addition, the marginal models we fit did not show the impact of time on mortality. As such, future research could show the impact of time on mortality and estimate cohort probabilities of death with age specific trends instead of using period life tables. We could also use multiple imputation by chained equations to replace missing data instead of last one carried forward to reduce bias in the regression estimates (Huque et al., 2018; Sterne et al., 2009; White et al., 2011). In terms of clustering techniques, we could also investigate more robust clustering methods to deal with categorical predictors such as hidden Markov models (Ghassempour et al., 2014). Further research on clustering using

other variables or more robust clustering techniques can be explored in the future. It would also be interesting to forecast trajectories of obesity measures or self reported health and determine the impact on life expectancy and healthy life expectancy of the population. We could also use quintiles from a logistic regresssion to determine risk profiles instead of clustering individual level data and also consider survival analysis.

3.6 Conclusion

Our results demonstrate the importance of using individual level data to identify individuals with different mortality risk profiles. These findings add to the increasing literature on using data that is specific to an individual and not population level data. Longitudinal individual level data is readily accessible and quantifies mortality heterogeneity by providing insights that are difficult to obtain from data based on a single time point. The joint modelling of BMI and self reported health show that individuals with normal BMI and declining fair health trajectories have the worst mortality outcomes than those with either a normal BMI and declining very good health or those who are obese and have declining good health. We find strong evidence of an association between cluster membership and each socio–economic variable and other risk factor. We also find strong evidence of an association between mortality and cluster allocation. The clusters continue to impact mortality when controlling for other risk factors. The clusters are poor predictors of mortality in the absence of the predictors BMI and self reported health whilst controlling for socio–economic variables and other risk factors. We find that there are significant differences in the pricing of annuities due to differences in body mass index and health status for both males and females. These results raise awareness for more rigorous determination of mortality risk profiles particularly when based on BMI, health status, education, income and wealth. Our results are relevant to policymakers, insurers, annuity providers and pension funds. From a business perspective, precise calculation of mortality risk can have a direct impact on the amount of reserves and profits attributable to shareholders.

3.7 Disclosure statement

No potential conflict of interest is reported by the authors.

3.8 Funding

This research is supported by the Australian Research Council Centre of Excellence in Population Ageing Research (project number CE170100005). Michelle Vhudzijena also acknowledges funding from the University of New South Wales Scientia PhD Scholarship Scheme from 2019–2023.

3.9 Acknowledgements

The authors acknowledge the technical assistance of Dr Gordana Popovic from Stats Central at the University of New South Wales, Sydney, Australia.

Appendix

3.A Testing for interactions between clusters and predictors using training data

We were also interested in determining whether there was an interaction between a predictor and the clusters since the predictors had significant effects on the mortality odds. Section 3.A presents the results of all the tests we carried out. Table 3.11 reports the results from testing whether the effect of self reported health on mortality risk varies by cluster whilst controlling for significant variables reported in Model 2. We find no evidence that there is an interaction between self reported health and cluster membership.

Table 3.11: Interaction of self reported health and clusters on mortality odds fitted using Model 2 which excludes variables with $p < 0.2$

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Beta						
Intercept	0.007	0.006	0.009	1715.788	0.000	***
Gender (ref: Female)						
Male	1.506	1.387	1.635	95.623	0.000	***
Sociodemographic						
Age	1.686	1.614	1.762	546.358	0.000	***
Marital status (ref: Married)						
Married spouse absent	1.836	1.454	2.320	25.995	0.000	***
Partnered	1.353	1.091	1.678	7.584	0.006	**
Separated	1.245	0.980	1.582	3.223	0.073	.
Divorced	1.173	1.040	1.323	6.757	0.009	**
Separated/divorced	1.380	1.003	1.899	3.922	0.048	*
Widowed	1.061	0.956	1.177	1.243	0.265	
Never married	1.389	1.149	1.679	11.501	0.001	***
Doctor diagnosed health conditions						
Psychiatric problems	1.151	1.048	1.264	8.694	0.003	**
High blood pressure	1.177	1.081	1.282	13.976	0.000	***
Diabetes	1.408	1.295	1.531	64.329	0.000	***
Cancer	1.730	1.580	1.894	141.141	0.000	***
Lung disease	1.426	1.302	1.563	57.792	0.000	***

Table 3.11: Interaction of self reported health and clusters on mortality odds fitted using Model 2 which excludes variables with $p < 0.2$ *(continued)*

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Heart problems	1.316	1.214	1.428	44.091	0.000	***
Stroke	1.404	1.267	1.555	42.051	0.000	***
Arthritis	0.905	0.831	0.986	5.250	0.022	*
Health behaviour						
Ever drinks alcohol	0.827	0.763	0.897	21.155	0.000	***
Ever smoked	1.391	1.270	1.525	49.859	0.000	***
Smokes now	1.452	1.313	1.606	52.792	0.000	***
Body mass index (ref: Normal weight)						
Underweight	1.944	1.653	2.287	64.398	0.000	***
Overweight	0.707	0.646	0.774	56.144	0.000	***
Obesity class I	0.537	0.468	0.618	76.619	0.000	***
Obesity class II	0.581	0.478	0.705	30.207	0.000	***
Obesity class III	0.689	0.548	0.865	10.291	0.001	**
Wealth and income						
Value of primary residence	0.836	0.779	0.897	24.725	0.000	***
Self reported health (ref: Excellent)						
Very good	1.106	0.862	1.419	0.624	0.430	
Good	1.290	0.996	1.670	3.723	0.054	.
Fair	3.324	2.511	4.399	70.490	0.000	***
Poor	6.343	4.513	8.916	113.094	0.000	***
Clusters (ref: Cluster A[a])						
Cluster B[b]	0.645	0.210	1.979	0.588	0.443	
Cluster C[c]	1.547	0.840	2.852	1.957	0.162	
Self reported health × Clusters						
Very good × Cluster B	2.502	0.786	7.964	2.409	0.121	
Good × Cluster B	2.370	0.762	7.373	2.221	0.136	
Fair × Cluster B	1.408	0.452	4.387	0.348	0.555	
Poor × Cluster B	1.254	0.395	3.979	0.148	0.701	
Very good × Cluster C	0.983	0.506	1.907	0.003	0.958	
Good × Cluster C	1.139	0.604	2.145	0.161	0.688	
Fair × Cluster C	0.633	0.334	1.200	1.962	0.161	
Poor × Cluster C	0.687	0.352	1.343	1.203	0.273	

Note: Significance levels

*** p<0.001; ** p<0.01; * p<0.05; . p<0.1; p<1.

Table 3.11: Interaction of self reported health and clusters on mortality odds fitted using Model 2 which excludes variables with $p < 0.2$ *(continued)*

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	

[a] Cluster A: normal, stable BMI and declining very good health;

[b] Cluster B: normal, stable BMI and declining fair health;

[c] Cluster C: high, increasing BMI and declining good health;

Table 3.12 reports the results from testing whether the effect of gender on mortality odds varies by cluster whilst controlling for significant variables reported in Model 2. We find that there is no interaction between gender and cluster membership.

Table 3.12: Interaction of gender and clusters on mortality odds fitted using Model 2 which excludes variables with $p < 0.2$

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Beta						
Intercept	0.007	0.005	0.009	1781.17	0.000	***
Sociodemographic						
Age	1.696	1.624	1.772	560.21	0.000	***
Marital status (ref: Married)						
Married spouse absent	1.849	1.462	2.337	26.39	0.000	***
Partnered	1.355	1.094	1.679	7.77	0.005	**
Separated	1.228	0.966	1.561	2.82	0.093	.
Divorced	1.169	1.036	1.319	6.43	0.011	*
Separated/divorced	1.373	0.998	1.890	3.78	0.052	.
Widowed	1.059	0.954	1.175	1.16	0.282	
Never married	1.383	1.143	1.674	11.10	0.001	***
Doctor diagnosed health conditions						
Psychiatric problems	1.144	1.041	1.256	7.83	0.005	**
High blood pressure	1.175	1.079	1.280	13.65	0.000	***
Diabetes	1.396	1.284	1.518	61.06	0.000	***
Cancer	1.732	1.582	1.896	140.73	0.000	***
Lung disease	1.416	1.292	1.553	55.07	0.000	***
Heart problems	1.314	1.212	1.426	43.45	0.000	***
Stroke	1.403	1.266	1.555	41.73	0.000	***
Arthritis	0.898	0.825	0.978	6.06	0.014	*
Health behaviour						

Table 3.12: Interaction of gender and clusters on mortality odds fitted using Model 2 which excludes variables with $p < 0.2$ *(continued)*

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Ever drinks alcohol	0.827	0.763	0.897	21.05	0.000	***
Ever smoked	1.394	1.272	1.528	50.33	0.000	***
Smokes now	1.457	1.318	1.612	53.62	0.000	***
Body mass index (ref: Normal weight)						
Underweight	1.953	1.659	2.299	64.82	0.000	***
Overweight	0.702	0.641	0.769	58.14	0.000	***
Obesity class I	0.532	0.463	0.611	79.18	0.000	***
Obesity class II	0.575	0.473	0.698	31.34	0.000	***
Obesity class III	0.683	0.544	0.858	10.76	0.001	**
Self reported health (ref: Excellent)						
Very good	1.182	0.945	1.479	2.14	0.143	
Good	1.526	1.224	1.903	14.11	0.000	***
Fair	2.552	2.029	3.211	63.99	0.000	***
Poor	4.550	3.591	5.766	157.41	0.000	***
Wealth and income						
Value of primary residence	0.837	0.780	0.898	24.54	0.000	***
Gender (ref: Female)						
Male	1.670	1.431	1.948	42.52	0.000	***
Clusters (ref: Cluster A[a])						
Cluster B[b]	1.311	1.123	1.530	11.78	0.001	***
Cluster C[c]	1.575	1.303	1.905	21.95	0.000	***
Gender × Cluster						
Male × Cluster B	0.877	0.729	1.056	1.92	0.166	
Male × Cluster C	0.862	0.700	1.061	1.96	0.161	

Note: Significance levels

*** p<0.001; ** p<0.01; * p<0.05; . p<0.1; p<1.

[a] Cluster A: normal, stable BMI and declining very good health;

[b] Cluster B: normal, stable BMI and declining fair health;

[c] Cluster C: high, increasing BMI and declining good health;

Table 3.13 reports the results from testing whether the effect of marital status on mortality odds varies by cluster whilst controlling for significant variables reported in Model 2. We find that being married and having an absent spouse increases mortality odds ($OR = 0.48(0.27, 0.86), p = 0.013$) in Cluster B than Cluster A. A similar effect is observed with

being divorced ($OR = 0.74(0.56, 0.98), p = 0.036$).

Table 3.13: Interaction of marital status and clusters on mortality odds fitted using Model 2 which excludes variables with $p < 0.2$

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Beta						
Intercept	0.007	0.006	0.009	1859.109	0.000	***
Gender (ref: Female)						
Male	1.508	1.389	1.637	95.662	0.000	***
Sociodemographic						
Age	1.698	1.625	1.774	558.686	0.000	***
Doctor diagnosed health conditions						
Psychiatric problems	1.147	1.044	1.260	8.203	0.004	**
High blood pressure	1.174	1.077	1.279	13.422	0.000	***
Diabetes	1.398	1.286	1.520	61.561	0.000	***
Cancer	1.740	1.589	1.906	142.890	0.000	***
Lung disease	1.420	1.295	1.556	55.638	0.000	***
Heart problems	1.320	1.217	1.432	44.847	0.000	***
Stroke	1.406	1.268	1.558	42.181	0.000	***
Arthritis	0.897	0.824	0.978	6.157	0.013	*
Health behaviour						
Ever drinks alcohol	0.824	0.760	0.894	21.800	0.000	***
Ever smoked	1.395	1.273	1.530	50.524	0.000	***
Smokes now	1.460	1.320	1.616	53.911	0.000	***
Body mass index (ref: Normal weight)						
Underweight	1.959	1.665	2.304	65.910	0.000	***
Overweight	0.706	0.645	0.773	56.425	0.000	***
Obesity class I	0.532	0.463	0.611	79.141	0.000	***
Obesity class II	0.577	0.475	0.700	30.932	0.000	***
Obesity class III	0.693	0.552	0.868	10.136	0.002	**
Self reported health (ref: Excellent)						
Very good	1.180	0.944	1.476	2.111	0.146	
Good	1.514	1.214	1.887	13.592	0.000	***
Fair	2.527	2.010	3.179	62.812	0.000	***
Poor	4.508	3.559	5.711	155.912	0.000	***
Wealth and income						
Value of primary residence	0.836	0.779	0.898	24.504	0.000	***
Marital status (ref: Married)						
Married spouse absent	2.994	1.872	4.788	20.956	0.000	***

Table 3.13: Interaction of marital status and clusters on mortality odds fitted using Model 2 which excludes variables with $p < 0.2$ *(continued)*

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Partnered	1.082	0.686	1.708	0.116	0.734	
Separated	0.616	0.223	1.700	0.877	0.349	
Divorced	1.456	1.155	1.835	10.109	0.002	**
Separated/divorced	1.442	0.609	3.416	0.693	0.405	
Widowed	1.028	0.840	1.259	0.073	0.787	
Never married	1.125	0.642	1.973	0.170	0.680	
Clusters (ref: Cluster A[a])						
Cluster B[b]	1.249	1.082	1.441	9.185	0.002	**
Cluster C[c]	1.469	1.229	1.755	17.903	0.000	***
Marital status × Cluster						
Married spouse absent × Cluster B	0.483	0.273	0.857	6.200	0.013	*
Partnered × Cluster B	1.406	0.808	2.447	1.455	0.228	
Separated × Cluster B	2.089	0.727	6.003	1.871	0.171	
Divorced × Cluster B	0.742	0.561	0.981	4.387	0.036	*
Separated/divorced × Cluster B	0.773	0.294	2.035	0.271	0.603	
Widowed × Cluster B	1.059	0.836	1.341	0.225	0.635	
Never married × Cluster B	1.204	0.652	2.223	0.353	0.552	
Married spouse absent × Cluster C	0.614	0.336	1.120	2.526	0.112	
Partnered ×Cluster C	1.266	0.696	2.303	0.597	0.440	
Separated ×Cluster C	2.210	0.720	6.788	1.920	0.166	
Divorced × Cluster C	0.777	0.563	1.073	2.348	0.125	
Separated/divorced × Cluster C	1.374	0.506	3.733	0.389	0.533	
Widowed × Cluster C	1.004	0.769	1.311	0.001	0.975	
Never married × Cluster C	1.440	0.746	2.779	1.181	0.277	

Note: Significance levels

*** p<0.001; ** p<0.01; * p<0.05; . p<0.1; p<1.

[a] Cluster A: normal, stable BMI and declining very good health;

[b] Cluster B: normal, stable BMI and declining fair health;

[c] Cluster C: high, increasing BMI and declining good health;

Table 3.14 reports the results from testing whether the effect of high blood pressure on mortality odds varies by cluster whilst controlling for significant variables reported in Model 2. We find that there is no interaction between high blood pressure and cluster membership.

Table 3.14: Interaction of high blood pressure and clusters on mortality odds fitted using Model 2 which excludes variables with $p < 0.2$

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Beta						
Intercept	0.007	0.006	0.009	1888.926	0.000	***
Gender (ref: Female)						
Male	1.506	1.387	1.635	95.215	0.000	***
Sociodemographic						
Age	1.696	1.623	1.772	558.397	0.000	***
Marital status (ref: Married)						
Married spouse absent	1.851	1.463	2.341	26.351	0.000	***
Partnered	1.358	1.096	1.682	7.831	0.005	**
Separated	1.234	0.971	1.569	2.954	0.086	.
Divorced	1.172	1.039	1.322	6.637	0.010	**
Separated/divorced	1.380	1.002	1.900	3.894	0.048	*
Widowed	1.061	0.956	1.177	1.226	0.268	
Never married	1.383	1.143	1.674	11.066	0.001	***
Doctor diagnosed health conditions						
Psychiatric problems	1.146	1.043	1.258	8.016	0.005	**
Diabetes	1.397	1.284	1.519	60.959	0.000	***
Cancer	1.736	1.585	1.901	141.683	0.000	***
Lung disease	1.416	1.291	1.552	54.768	0.000	***
Heart problems	1.313	1.211	1.425	43.181	0.000	***
Stroke	1.403	1.265	1.555	41.450	0.000	***
Arthritis	0.898	0.824	0.978	6.116	0.013	*
Health behaviour						
Ever drinks alcohol	0.825	0.761	0.895	21.602	0.000	***
Ever smoked	1.395	1.273	1.529	50.621	0.000	***
Smokes now	1.455	1.315	1.609	52.920	0.000	***
Body mass index (ref: Normal weight)						
Underweight	1.944	1.653	2.287	64.425	0.000	***
Overweight	0.704	0.643	0.771	57.263	0.000	***
Obesity class I	0.532	0.463	0.612	78.774	0.000	***
Obesity class II	0.577	0.476	0.701	30.814	0.000	***
Obesity class III	0.693	0.553	0.869	10.076	0.002	**
Self reported health (ref: Excellent)						
Very good	1.180	0.942	1.478	2.062	0.151	
Good	1.525	1.220	1.906	13.699	0.000	***

Table 3.14: Interaction of high blood pressure and clusters on mortality odds fitted using Model 2 which excludes variables with $p < 0.2$ *(continued)*

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Fair	2.555	2.026	3.221	62.862	0.000	***
Poor	4.553	3.586	5.781	154.720	0.000	***
Wealth and income						
Value of primary residence	0.837	0.780	0.898	24.434	0.000	***
Doctor diagnosed health conditions						
High blood pressure	1.178	1.010	1.373	4.373	0.036	*
Clusters (ref: Cluster A[a])						
Cluster B[b]	1.207	1.024	1.422	5.033	0.025	*
Cluster C[c]	1.504	1.198	1.888	12.402	0.000	***
High blood pressure × Cluster						
High blood pressure × Cluster B	1.013	0.840	1.223	0.019	0.890	
High blood pressure × Cluster C	0.955	0.755	1.208	0.148	0.700	

Note: Significance levels

*** p<0.001; ** p<0.01; * p<0.05; . p<0.1; p<1.

[a] Cluster A: normal, stable BMI and declining very good health;

[b] Cluster B: normal, stable BMI and declining fair health;

[c] Cluster C: high, increasing BMI and declining good health;

Table 3.15 reports the results from testing whether the effect of current smoking status on mortality odds varies by cluster whilst controlling for significant variables reported in Model 2. We find that there is an interaction between current smoking status and cluster membership.

Table 3.15: Interaction of current smoking status and clusters on mortality odds fitted using Model 2 which excludes variables with $p < 0.2$

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Beta						
Intercept	0.007	0.005	0.008	1936.67	0.000	***
Gender (ref: Female)						
Male	1.512	1.393	1.642	96.99	0.000	***
Sociodemographic						
Age	1.692	1.619	1.768	553.22	0.000	***

Table 3.15: Interaction of current smoking status and clusters on mortality odds fitted using Model 2 which excludes variables with $p < 0.2$ *(continued)*

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Marital status (ref: Married)						
Married spouse absent	1.847	1.461	2.335	26.32	0.000	***
Partnered	1.345	1.085	1.667	7.34	0.007	**
Separated	1.239	0.974	1.576	3.04	0.081	.
Divorced	1.177	1.044	1.328	7.04	0.008	**
Separated/divorced	1.392	1.012	1.915	4.12	0.042	*
Widowed	1.058	0.953	1.174	1.11	0.292	
Never married	1.384	1.144	1.676	11.15	0.001	***
Doctor diagnosed health conditions						
Psychiatric problems	1.146	1.043	1.259	8.07	0.004	**
High blood pressure	1.181	1.084	1.287	14.42	0.000	***
Diabetes	1.396	1.284	1.518	60.93	0.000	***
Cancer	1.742	1.591	1.908	144.06	0.000	***
Lung disease	1.421	1.296	1.557	56.22	0.000	***
Heart problems	1.315	1.212	1.427	43.41	0.000	***
Stroke	1.404	1.266	1.556	41.75	0.000	***
Arthritis	0.901	0.827	0.982	5.67	0.017	*
Health behaviour						
Ever drinks alcohol	0.828	0.764	0.898	20.81	0.000	***
Ever smoked	1.385	1.263	1.519	47.89	0.000	***
Body mass index (ref: Normal weight)						
Underweight	1.945	1.652	2.289	64.05	0.000	***
Overweight	0.707	0.645	0.774	55.81	0.000	***
Obesity class I	0.534	0.465	0.614	77.81	0.000	***
Obesity class II	0.580	0.478	0.705	29.97	0.000	***
Obesity class III	0.695	0.554	0.872	9.89	0.002	**
Self reported health (ref: Excellent)						
Very good	1.187	0.949	1.485	2.25	0.133	
Good	1.532	1.229	1.909	14.40	0.000	***
Fair	2.561	2.036	3.221	64.57	0.000	***
Poor	4.578	3.614	5.799	158.96	0.000	***
Wealth and income						
Value of primary residence	0.840	0.783	0.901	23.77	0.000	***
Health behaviour						
Smokes now	1.928	1.594	2.332	45.74	0.000	***
Clusters (ref: Cluster A[a])						

Table 3.15: Interaction of current smoking status and clusters on mortality odds fitted using Model 2 which excludes variables with $p < 0.2$ *(continued)*

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Cluster B[b]	1.336	1.175	1.520	19.45	0.000	***
Cluster C[c]	1.517	1.288	1.787	24.98	0.000	***
Smokes now × Cluster						
Smokes now × Cluster B	0.668	0.537	0.831	13.16	0.000	***
Smokes now × Cluster C	0.797	0.609	1.043	2.74	0.098	.

Note: Significance levels

*** p<0.001; ** p<0.01; * p<0.05; . p<0.1; p<1.

[a] Cluster A: normal, stable BMI and declining very good health;

[b] Cluster B: normal, stable BMI and declining fair health;

[c] Cluster C: high, increasing BMI and declining good health;

Table 3.16 reports the results from testing whether the effect of stroke on mortality odds varies by cluster whilst controlling for significant variables reported in Model 2. We find no evidence of an interaction between stroke and cluster membership.

Table 3.16: Interaction of stroke and clusters on mortality odds fitted using Model 2 which excludes variables with $p < 0.2$

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Beta						
Intercept	0.007	0.006	0.009	1945.137	0.000	***
Gender (ref: Female)						
Male	1.506	1.388	1.636	95.647	0.000	***
Sociodemographic						
Age	1.696	1.623	1.771	557.800	0.000	***
Marital status (ref: Married)						
Married spouse absent	1.860	1.473	2.348	27.222	0.000	***
Partnered	1.360	1.098	1.684	7.929	0.005	**
Separated	1.241	0.977	1.576	3.139	0.076	.
Divorced	1.173	1.040	1.324	6.755	0.009	**
Separated/divorced	1.383	1.004	1.903	3.949	0.047	*
Widowed	1.062	0.957	1.179	1.297	0.255	
Never married	1.387	1.146	1.677	11.329	0.001	***

Table 3.16: Interaction of stroke and clusters on mortality odds fitted using Model 2 which excludes variables with $p < 0.2$ *(continued)*

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Doctor diagnosed health conditions						
Psychiatric problems	1.146	1.043	1.259	8.081	0.004	**
High blood pressure	1.177	1.080	1.282	13.900	0.000	***
Diabetes	1.396	1.284	1.518	60.963	0.000	***
Cancer	1.737	1.587	1.902	142.653	0.000	***
Lung disease	1.414	1.290	1.551	54.720	0.000	***
Heart problems	1.312	1.210	1.423	42.907	0.000	***
Arthritis	0.897	0.824	0.977	6.190	0.013	*
Health behaviour						
Ever drinks alcohol	0.825	0.761	0.894	21.715	0.000	***
Ever smoked	1.397	1.274	1.531	50.959	0.000	***
Smokes now	1.454	1.314	1.608	52.712	0.000	***
Body mass index (ref: Normal weight)						
Underweight	1.941	1.651	2.283	64.397	0.000	***
Overweight	0.705	0.643	0.771	57.234	0.000	***
Obesity class I	0.535	0.465	0.615	77.843	0.000	***
Obesity class II	0.579	0.477	0.704	30.346	0.000	***
Obesity class III	0.693	0.552	0.870	10.015	0.002	**
Self reported health (ref: Excellent)						
Very good	1.178	0.941	1.474	2.044	0.153	
Good	1.514	1.214	1.890	13.488	0.000	***
Fair	2.532	2.011	3.188	62.523	0.000	***
Poor	4.517	3.563	5.725	155.307	0.000	***
Wealth and income						
Value of primary residence	0.837	0.780	0.898	24.497	0.000	***
Doctor diagnosed health conditions						
Stroke	1.501	1.176	1.916	10.623	0.001	**
Clusters (ref: Cluster A[a])						
Cluster B[b]	1.246	1.101	1.410	12.159	0.000	***
Cluster C[c]	1.437	1.223	1.688	19.426	0.000	***
Stroke × Cluster						
Stroke × Cluster B	0.879	0.666	1.161	0.827	0.363	
Stroke × Cluster C	1.019	0.757	1.372	0.015	0.901	

Note: Significance levels

*** p<0.001; ** p<0.01; * p<0.05; . p<0.1; p<1.

Table 3.16: Interaction of stroke and clusters on mortality odds fitted using Model 2 which excludes variables with $p < 0.2$ *(continued)*

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	

[a] Cluster A: normal, stable BMI and declining very good health;

[b] Cluster B: normal, stable BMI and declining fair health;

[c] Cluster C: high, increasing BMI and declining good health;

Table 3.17 reports the results from testing whether the effect of lung disease on mortality odds varies by cluster whilst controlling for significant variables reported in Model 2. We find some evidence of an interaction between lung disease and cluster membership.

Table 3.17: Interaction of lung disease and clusters on mortality odds fitted using Model 2 which excludes variables with $p < 0.2$

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Beta						
Intercept	0.007	0.006	0.009	1947.65	0.000	***
Gender (ref: Female)						
Male	1.508	1.389	1.637	96.17	0.000	***
Sociodemographic						
Age	1.694	1.621	1.770	555.98	0.000	***
Marital status (ref: Married)						
Married spouse absent	1.850	1.464	2.339	26.50	0.000	***
Partnered	1.354	1.093	1.679	7.67	0.006	**
Separated	1.232	0.969	1.565	2.91	0.088	.
Divorced	1.172	1.039	1.321	6.64	0.010	**
Separated/divorced	1.379	1.002	1.898	3.89	0.049	*
Widowed	1.060	0.955	1.177	1.21	0.272	
Never married	1.383	1.143	1.674	11.13	0.001	***
Doctor diagnosed health conditions						
Psychiatric problems	1.148	1.045	1.260	8.29	0.004	**
High blood pressure	1.172	1.076	1.277	13.23	0.000	***
Diabetes	1.396	1.284	1.517	61.01	0.000	***
Cancer	1.737	1.586	1.902	142.75	0.000	***
Heart problems	1.318	1.215	1.429	44.30	0.000	***
Stroke	1.401	1.264	1.552	41.44	0.000	***
Arthritis	0.897	0.824	0.977	6.18	0.013	*

Table 3.17: Interaction of lung disease and clusters on mortality odds fitted using Model 2 which excludes variables with $p < 0.2$ *(continued)*

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Health behaviour						
Ever drinks alcohol	0.825	0.761	0.894	21.74	0.000	***
Ever smoked	1.391	1.269	1.525	49.70	0.000	***
Smokes now	1.456	1.316	1.610	53.45	0.000	***
Body mass index (ref: Normal weight)						
Underweight	1.934	1.644	2.276	63.12	0.000	***
Overweight	0.705	0.644	0.772	57.12	0.000	***
Obesity class I	0.533	0.464	0.613	78.45	0.000	***
Obesity class II	0.577	0.476	0.701	30.72	0.000	***
Obesity class III	0.691	0.550	0.867	10.15	0.001	**
Self reported health (ref: Excellent)						
Very good	1.168	0.933	1.461	1.83	0.176	
Good	1.492	1.196	1.861	12.54	0.000	***
Fair	2.499	1.986	3.145	60.95	0.000	***
Poor	4.477	3.533	5.673	154.05	0.000	***
Wealth and income						
Value of primary residence	0.837	0.780	0.898	24.47	0.000	***
Doctor diagnosed health conditions						
Lung disease	1.719	1.385	2.133	24.20	0.000	***
Clusters (ref: Cluster A[a])						
Cluster B[b]	1.274	1.123	1.445	14.15	0.000	***
Cluster C[c]	1.489	1.268	1.748	23.55	0.000	***
Lung disease × Cluster						
Lung disease × Cluster B	0.780	0.612	0.994	4.04	0.044	*
Lung disease × Cluster C	0.831	0.633	1.089	1.80	0.180	

Note: Significance levels

*** p<0.001; ** p<0.01; * p<0.05; . p<0.1; p<1.

[a] Cluster A: normal, stable BMI and declining very good health;

[b] Cluster B: normal, stable BMI and declining fair health;

[c] Cluster C: high, increasing BMI and declining good health;

Table 3.17 reports the results from testing whether the effect of cancer on mortality odds varies by cluster whilst controlling for significant variables reported in Model 2. The effect of cancer on mortality odds is the same in each cluster.

Table 3.18: Interaction of cancer and clusters on mortality odds fitted using Model 2 which excludes variables with $p < 0.2$

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Beta						
Intercept	0.007	0.006	0.009	1926.191	0.000	***
Gender (ref: Female)						
Male	1.505	1.387	1.634	95.052	0.000	***
Sociodemographic						
Age	1.696	1.623	1.772	558.464	0.000	***
Marital status (ref: Married)						
Married spouse absent	1.852	1.464	2.341	26.499	0.000	***
Partnered	1.355	1.094	1.679	7.758	0.005	**
Separated	1.233	0.970	1.568	2.938	0.086	.
Divorced	1.172	1.039	1.323	6.669	0.010	**
Separated/divorced	1.378	1.001	1.898	3.865	0.049	*
Widowed	1.061	0.956	1.178	1.253	0.263	
Never married	1.385	1.145	1.676	11.223	0.001	***
Doctor diagnosed health conditions						
Psychiatric problems	1.146	1.043	1.258	8.046	0.005	**
High blood pressure	1.175	1.078	1.280	13.565	0.000	***
Diabetes	1.396	1.284	1.518	60.843	0.000	***
Lung disease	1.416	1.292	1.552	55.046	0.000	***
Heart problems	1.314	1.211	1.425	43.379	0.000	***
Stroke	1.403	1.266	1.555	41.739	0.000	***
Arthritis	0.898	0.824	0.978	6.119	0.013	*
Health behaviour						
Ever drinks alcohol	0.825	0.761	0.895	21.670	0.000	***
Ever smoked	1.395	1.272	1.528	50.505	0.000	***
Smokes now	1.456	1.316	1.610	53.220	0.000	***
Body mass index (ref: Normal weight)						
Underweight	1.939	1.648	2.281	63.669	0.000	***
Overweight	0.705	0.644	0.772	57.124	0.000	***
Obesity class I	0.533	0.464	0.613	78.606	0.000	***
Obesity class II	0.578	0.476	0.701	30.717	0.000	***
Obesity class III	0.692	0.551	0.867	10.222	0.001	**
Self reported health (ref: Excellent)						
Very good	1.176	0.940	1.471	2.002	0.157	
Good	1.515	1.215	1.888	13.635	0.000	***

Table 3.18: Interaction of cancer and clusters on mortality odds fitted using Model 2 which excludes variables with $p < 0.2$ *(continued)*

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Fair	2.535	2.016	3.189	63.322	0.000	***
Poor	4.522	3.570	5.726	156.719	0.000	***
Wealth and income						
Value of primary residence	0.837	0.780	0.898	24.509	0.000	***
Doctor diagnosed health conditions						
Cancer	1.826	1.538	2.169	47.070	0.000	***
Clusters (ref: Cluster A[a])						
Cluster B[b]	1.240	1.091	1.410	10.859	0.001	***
Cluster C[c]	1.480	1.259	1.738	22.657	0.000	***
Cancer × Cluster						
Cancer × Cluster B	0.938	0.759	1.159	0.352	0.553	
Cancer × Cluster C	0.929	0.727	1.188	0.345	0.557	

Note: Significance levels

*** p<0.001; ** p<0.01; * p<0.05; . p<0.1; p<1.

[a] Cluster A: normal, stable BMI and declining very good health;

[b] Cluster B: normal, stable BMI and declining fair health;

[c] Cluster C: high, increasing BMI and declining good health;

Table 3.19 reports the results from testing whether the effect of heart problems on mortality odds varies by cluster whilst controlling for significant variables reported in Model 2. There is some weak evidence of an interaction between heart problems and cluster membership.

Table 3.19: Interaction of heart problems and clusters on mortality odds fitted using Model 2 which excludes variables with $p < 0.2$

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Beta						
Intercept	0.007	0.006	0.009	1933.679	0.000	***
Gender (ref: Female)						
Male	1.509	1.390	1.639	96.239	0.000	***
Sociodemographic						
Age	1.696	1.623	1.772	557.163	0.000	***

Table 3.19: Interaction of heart problems and clusters on mortality odds fitted using Model 2 which excludes variables with $p < 0.2$ *(continued)*

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Marital status (ref: Married)						
Married spouse absent	1.839	1.455	2.324	25.981	0.000	***
Partnered	1.358	1.097	1.681	7.893	0.005	**
Separated	1.230	0.969	1.561	2.891	0.089	.
Divorced	1.171	1.038	1.321	6.625	0.010	*
Separated/divorced	1.380	1.002	1.899	3.898	0.048	*
Widowed	1.062	0.957	1.179	1.290	0.256	
Never married	1.383	1.143	1.674	11.090	0.001	***
Doctor diagnosed health conditions						
Psychiatric problems	1.142	1.040	1.255	7.701	0.005	**
High blood pressure	1.180	1.083	1.286	14.367	0.000	***
Diabetes	1.391	1.279	1.513	59.135	0.000	***
Cancer	1.735	1.584	1.900	141.790	0.000	***
Lung disease	1.420	1.296	1.557	55.834	0.000	***
Stroke	1.401	1.264	1.553	41.458	0.000	***
Arthritis	0.899	0.825	0.979	5.984	0.014	*
Health behaviour						
Ever drinks alcohol	0.824	0.760	0.893	21.965	0.000	***
Ever smoked	1.393	1.271	1.527	50.080	0.000	***
Smokes now	1.452	1.312	1.606	52.436	0.000	***
Body mass index (ref: Normal weight)						
Underweight	1.937	1.647	2.277	63.982	0.000	***
Overweight	0.706	0.644	0.772	57.105	0.000	***
Obesity class I	0.536	0.467	0.616	77.014	0.000	***
Obesity class II	0.576	0.474	0.700	30.711	0.000	***
Obesity class III	0.690	0.549	0.867	10.180	0.001	**
Self reported health (ref: Excellent)						
Very good	1.184	0.945	1.482	2.155	0.142	
Good	1.525	1.220	1.906	13.738	0.000	***
Fair	2.539	2.013	3.203	61.826	0.000	***
Poor	4.521	3.562	5.739	153.674	0.000	***
Wealth and income						
Value of primary residence	0.838	0.780	0.899	24.358	0.000	***
Doctor diagnosed health conditions						
Heart problems	1.259	1.058	1.498	6.733	0.010	**
Clusters (ref: Cluster A[a])						

Table 3.19: Interaction of heart problems and clusters on mortality odds fitted using Model 2 which excludes variables with $p < 0.2$ *(continued)*

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Cluster B[b]	1.255	1.098	1.435	11.037	0.001	***
Cluster C[c]	1.319	1.109	1.570	9.734	0.002	**
Heart problems × Cluster						
Heart problems × Cluster B	0.964	0.788	1.180	0.125	0.724	
Heart problems × Cluster C	1.242	0.993	1.552	3.612	0.057	.

Note: Significance levels

*** p<0.001; ** p<0.01; * p<0.05; . p<0.1; p<1.

[a] Cluster A: normal, stable BMI and declining very good health;

[b] Cluster B: normal, stable BMI and declining fair health;

[c] Cluster C: high, increasing BMI and declining good health;

Table 3.20 reports the results from testing whether the effect of age on mortality odds varies by cluster whilst controlling for significant variables reported in Model 2. There is some weak evidence of an interaction between age and cluster membership.

Table 3.20: Interaction of age and clusters on mortality odds fitted using Model 2 which excludes variables with $p < 0.2$

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Beta						
Intercept	0.007	0.006	0.009	1937.44	0.000	***
Gender (ref: Female)						
Male	1.504	1.386	1.633	95.04	0.000	***
Marital status (ref: Married)						
Married spouse absent	1.850	1.466	2.336	26.77	0.000	***
Partnered	1.359	1.097	1.683	7.89	0.005	**
Separated	1.228	0.967	1.560	2.83	0.092	.
Divorced	1.172	1.039	1.322	6.70	0.010	**
Separated/divorced	1.371	0.995	1.889	3.72	0.054	.
Widowed	1.060	0.955	1.176	1.20	0.273	
Never married	1.384	1.144	1.673	11.23	0.001	***
Doctor diagnosed health conditions						
Psychiatric problems	1.147	1.045	1.260	8.31	0.004	**
High blood pressure	1.173	1.077	1.278	13.40	0.000	***

Table 3.20: Interaction of age and clusters on mortality odds fitted using Model 2 which excludes variables with $p < 0.2$ *(continued)*

		95% CI				
Term	**Odds ratio**	**Lower**	**Upper**	**Wald**	**p–value**	
Diabetes	1.393	1.281	1.515	60.40	0.000	***
Cancer	1.731	1.582	1.895	141.75	0.000	***
Lung disease	1.415	1.292	1.551	55.33	0.000	***
Heart problems	1.313	1.211	1.424	43.45	0.000	***
Stroke	1.405	1.268	1.556	42.37	0.000	***
Arthritis	0.897	0.824	0.977	6.29	0.012	*
Health behaviour						
Ever drinks alcohol	0.825	0.761	0.894	21.85	0.000	***
Ever smoked	1.396	1.274	1.530	51.00	0.000	***
Smokes now	1.447	1.307	1.601	51.09	0.000	***
Body mass index (ref: Normal weight)						
Underweight	1.933	1.645	2.272	63.82	0.000	***
Overweight	0.704	0.643	0.771	57.70	0.000	***
Obesity class I	0.534	0.464	0.613	78.22	0.000	***
Obesity class II	0.578	0.476	0.703	30.40	0.000	***
Obesity class III	0.695	0.554	0.872	9.91	0.002	**
Self reported health (ref: Excellent)						
Very good	1.158	0.924	1.452	1.62	0.203	
Good	1.480	1.184	1.851	11.82	0.001	***
Fair	2.481	1.968	3.128	59.09	0.000	***
Poor	4.415	3.478	5.604	148.78	0.000	***
Wealth and income						
Value of primary residence	0.837	0.780	0.898	24.63	0.000	***
Sociodemographic						
Age	1.776	1.641	1.922	202.87	0.000	***
Clusters (ref: Cluster A[a])						
Cluster B[b]	1.291	1.128	1.477	13.82	0.000	***
Cluster C[c]	1.488	1.252	1.767	20.44	0.000	***
Age × Cluster						
Age × Cluster B	0.923	0.843	1.010	3.01	0.083	.
Age × Cluster C	0.979	0.883	1.086	0.16	0.689	

Note: Significance levels

*** p<0.001; ** p<0.01; * p<0.05; . p<0.1; p<1.

[a] Cluster A: normal, stable BMI and declining very good health;

[b] Cluster B: normal, stable BMI and declining fair health;

[c] Cluster C: high, increasing BMI and declining good health;

Chapter 4

Modelling mortality and functional disability risks using hidden Markov models with covariates

Earlier versions of this chapter were presented at the following conferences and events:

- Seventeenth International Longevity Risk and Capital Markets Solutions Conference, Waterloo, Canada. "Modelling mortality risks using hidden Markov models with covariates", 13 September 2022.
- CEPAR Longevity Risk Workshop, Retirement Income: Risks and Solutions, University of New South Wales, Sydney, Australia. "Modelling mortality and functional disability risks using hidden Markov models with covariates", 28 November 2022.

4.1 Introduction

There is a large body of literature on the pricing of long term care products using multistate models. For example, Renshaw & Haberman (1995) demonstrate how to graduate transition rates in a three state Markov Chain Monte Carlo model with recovery from sickness in a generalised linear model (GLM) framework using United Kingdom (UK) data (Forfar et al., 1988; Nelder & Wedderburn, 1972). An extension of this work is in Rickayzen & Walsh (2002), where a multistate Markov chain model is used to estimate the

number and trend of people with disabilities in the UK but the authors do not consider the financial implications nor the uncertainty. Using United States (US) longitudinal data, Pritchard (2006) shows that not accounting for recovery from illness in a multiple state model tends to increase costs of long–term care. Fong et al. (2015) develop a multistate functional disability model that has sex and age specific transition probabilities in a three state model with recovery from disabled to nondisabled for individuals aged 50 to 100.

Hanewald et al. (2019) extend the work in Fong et al. (2015) by developing a generalised linear model that allows for age effects, time trends and age–time interactions in a multi-state Markov model using data from the Chinese Longitudinal Healthy Longevity Survey. However, there is no recovery from the functionally disabled state to the non–disabled state (healthy). Kogure et al. (2019) propose integrating Lee–Carter with additional parameters on mortality differentials in health status using a Bayesian approach. This is because long–term care models based on longitudinal data (Fong et al., 2015; Z. Li et al., 2017; Shao et al., 2017; Sherris & Wei, 2021) only cover short durations and there are gaps between survey intervals which would result in loss of information. However, their work fails to model transitions between states.

Sherris & Wei (2021) develop a five state Markov model that includes an independent health state to capture differences in mortality risk due to the presence of a chronic condition such as diabetes and heart disease. Their results show that disregarding health status can cause adverse selection from individuals with chronic conditions due to inaccurate pricing of mortality and disability risks. However, while this is a remarkable extension of the multistate Markov models with systematic trend and uncertainty; it fails to make a more holistic use of the full breadth and depth of longitudinal data by neglecting the impact of individual risk factors such as body mass index, education, wealth and income on health status. These variables are known to have a significant impact on mortality and morbidity as discussed in Chapter 2.4. As such, we incorporate these predictors by fitting hidden Markov models (HMMs) with health status as the response and covariates of body mass index (BMI) and self reported health. We then cluster the multivariate time series using HMMs to create an indicator variable that captures the relationship between health status and covariates and illustrates the heterogeneity of mortality risk amongst individuals. Our motivation is that, there are variations in risk amongst people who are in ill health and their particular dynamics need to be accounted for in longevity products

for fairer and better management of mortality and functional disability risks.

Hidden Markov models are doubly embedded stochastic processes that are used to model an output based on an assumption that the system being modelled is Markovian in nature and has unobservable states. The parameters of the hidden states are determined from the observed output. Emission probabilities are probabilities from the hidden state to the observed state while transition probabilities describe the Markov process from one hidden state to another. HMMs are commonly used to analyse time series and can be used to cluster trajectories since they offer a model based clustering method that is embedded in a probabilistic framework.

Therefore, in this paper we cluster health trajectories to place individuals with similar mortality/morbidity risk profiles in the same groups to demonstrate heterogeneity. We verify if this clustering improves the estimation of transition rates among the different states of a three state Markov chain model with systematic trend and uncertainty. This leads us to the following research questions:

1. To what extent do the clusters developed from the multivariate-time series clustering of health trajectories using HMMs with covariates, provide well developed risk profiles that exhibit mortality heterogeneity?

2. Does clustering provide a better fit to empirical data when estimating transition rates and life expectancy in a multistate model of health status and functional disability while controlling for age and gender?

The rest of the chapter is structured as follows: Section 4.2 briefly describes the data and the steps taken to clean the dataset. Section 4.3 describes the methods used to estimate the hidden Markov models, cluster the trajectories and estimate multistate Markov models. Section 4.4 provides the results of the HMMs, clustering and multistate models. Section 4.5 is a discussion of the results and we conclude in Section 4.6.

4.2 Data

The data are from the United States Health and Retirement Study (HRS) which is conducted by the University of Michigan. We use waves 4 to 13 to fit the multistate models.

The year 1998 is the first interview year in the multistate model data. Earlier waves have inconsistencies in survey questions pertaining to activities of daily living. This handling of the data is similar to previous studies using this dataset (Fong et al., 2015; Z. Li et al., 2017; Shao et al., 2017; Sherris & Wei, 2021). Waves 1 to 13 are used to estimate the HMMs and cluster the trajectories as they have more information on health status and the covariates. We exclude individuals who provide inappropriate responses, who fail to respond in any wave and those who do not respond at least twice to the survey. We also limit our sample to individuals who acquire a chronic condition. Our final sample contains $3,940$ participants and $60,855$ observations. All individuals are at least aged 45 and are observed from 1992–2016. This is departure from previous studies because we want to fit HMMs with a multinomial logistic regression to each individual who acquires a chronic disease during the observation period. Table 4.1 summarises the variables used in our analysis.

Table 4.1: Descriptions of variables extracted from the Health and Retirement Study

Variable	**Description**
Socio–demographic	
HHIDPN	Person specific identifier in each household
RAGENDER	Gender
RAEDUC	Highest level of education
RMSTAT	Marital status
HACOHORT	Cohort
RABMONTH	Birth month
RABYEAR	Birth year
RABDATE	Date of birth
RADMONTH	Month of death
RADYEAR	Year of death
RADDATE	Date of death
Interview	
RxIWEND	End date of interview
RxIWENDY	End year of interview
RxIWSTAT	Interview status

Table 4.1: Descriptions of variables extracted from the Health and Retirement Study *(continued)*

Variable	Description
Activities of daily living	
RxWALKRA	Some difficulty walking across room
RxDRESSA	Some difficulty dressing
RxEATA	Some difficulty eating
RxBEDA	Some difficulty getting in and out of bed
RxTOILTA	Some difficulty using the toilet
RxBATHA	Some difficulty bathing or showering
Doctor diagnosed health conditions	
RxDIABE	Ever had diabetes
RxLUNGE	Ever had lung disease
RxSTROKE	Ever had a stroke
RxHEARTE	Ever had heart problems
Health behaviour	
RxBMI	Self reported body mass index
RxSMOKEN	Current smoking status
RxDRINK	Ever drinks any alcohol
RxSHLT	Self–reported health
Wealth and income	
HAHOUS	Value of primary residence
HATOTN	Total non–housing wealth
HITOT	Total household income

Note: x indicates the wave of the HRS dataset

4.3 Methodology

4.3.1 Approach

We follow the procedure outlined in Ghassempour et al. (2014) to develop clusters using hidden Markov models. We define T_i as the health trajectories for $i = 1, \ldots, N$ where N is the number of individuals. Each T_i is a $t \times d$ matrix where d are the columns relating to the the multivariate time series of the response and covariates (health status, BMI, self reported health) and t are the rows relating to the number of waves, $t_{max} = 13$. Similar to Sherris & Wei (2021), we define health status as the presence of at least one of 4 chronic doctor diagnosed conditions (heart problems, diabetes, lung disease and stroke). This is different from at Ghassempour et al. (2014), who consider the presence of three doctor diagnosed conditions in their analysis. The reason for this departure is that payments in a health and functional disability product are triggered by the onset of only one disease.

4.3.2 Hidden Markov Models

In this section we provide some background to hidden Markov models. Figure 4.1 shows a schematic of a hidden Markov model with q_T hidden states and observed output O_T.

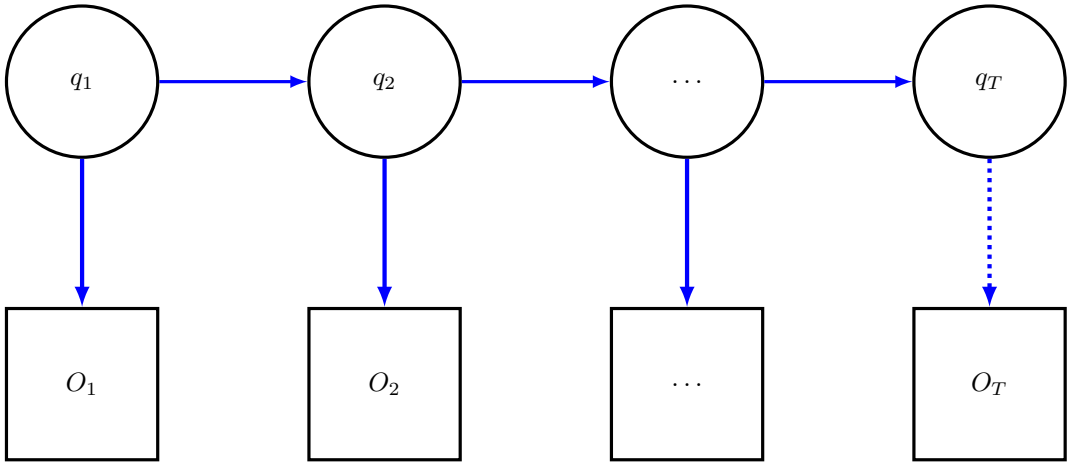


Figure 4.1: A graphical representation of a Hidden Markov Model

4.3.2.1 Principles

As described in Rabiner & Juang (1986), we assume that a Hidden Markov Model (HMM) has $Q = q_1, q_2, \ldots, q_N$ finite N hidden (unobserved) states and $V = v_1, v_2, \ldots, v_M$ discrete

set M of possible observations. The HMM is specified as $\lambda = (A, B, \pi)$ where $A = \{a_{ij}\}$ is the matrix of state transition probabilities, a_{ij} is the probability from state q_i at time t_i to state q_j at time $t_i + 1$, $B = \{b_j(k)\}$ is the matrix of emission probabilities, $b_j(k)$ is the observation symbol probability in state j and $\pi = \pi_i$, is the initial state distribution, π is the probability of q_i at $t = 1$. Matrices A and B are unknown. Emission probabilities are probabilities from the hidden state to the observed state while transition probabilities describe the Markov process from one hidden state to another. The observation sequence $O = (O_1, O_2, \ldots, O_t)$ represents the outcome generated by the HMM. To solve a HMM we need solutions to the following problems: evaluation, estimation and training.

4.3.2.2 The evaluation problem

The evaluation problem seeks a solution to compute $\Pr(O|\lambda)$ given an observation sequence. For a fixed state sequence $I = i_1, i_2, \ldots, i_t$,

$$\Pr(O|I, \lambda) = b_{i1}(O_1)b_{i2}(O_2)\ldots b_{it}(O_T). \tag{4.1}$$

While

$$\Pr(I|\lambda) = \pi_{1_1} a_{i_1 i_2} \ldots a_{i_{T-1} i_T}. \tag{4.2}$$

The joint probability of O and I is $\Pr(O, I|\lambda) = \Pr(O|I, \lambda)\Pr(I, \lambda)$. Therefore

$$\begin{aligned} \Pr(O|\lambda) &= \sum_{all\ I} \Pr(O|I, \lambda)\Pr(I|\lambda) \\ &= \sum_{i_1, i_2, \ldots, i_T} \pi_{i1} b_{i_1}(O_1) a_{i_1 i_2} b_{i_2}(O_2) \ldots a_{i_{T-1} i_T} b_{iT} O_T. \end{aligned} \tag{4.3}$$

This means we need to make $2(T-1)N^T$ calculations to determine $\Pr(O|\lambda)$ directly. Since this is not feasible we use the forward–backward procedure to solve for $\Pr(O|\lambda)$ (Baum & Eagon, 1967). The forward variable $\alpha_t(i)$, is the probability of the partial observation sequence until time t and state q_i given the model λ,

$$\alpha_t(i) = \Pr(O_1, O_2, \ldots, O_t, i_t = q_i|\lambda). \tag{4.4}$$

We solve for $\alpha_t(i)$ in three steps: initialisation, induction and termination. First, initialise the forward probabilities with the joint probability of state q_i and initial observation O_1,

$$\alpha_t(i) = \pi_i b_i O_1 \quad 1 \leq i \leq N. \tag{4.5}$$

Secondly, induce $\alpha_{t+1}(j)$ recursively using the initial conditions

$$\alpha_{t+1}(j) = \left[\sum_{i,j=1}^{N} \alpha_t(i) a_{ij}\right] b_j(O_{t+1}). \tag{4.6}$$

Lastly, compute

$$\Pr(O|\lambda) = \sum_{i=1}^{N} \alpha_T(i) \tag{4.7}$$

to terminate the forward procedure.

4.3.2.3 The estimation problem

The second question we have to solve is what is the optimal hidden state sequence $I = i_1, i_2, \ldots, i_T$ given an observation sequence. That is, calculate the probability of being in state q_i at time t given the observation sequence $O = O_1, O_2, \ldots, O_T$ and the model λ, that is

$$\begin{aligned} \gamma_t(i) &= \Pr(i_t = q_i|O, \lambda) \\ &= \Pr(O_1, O_2, \ldots, O_t, i_t = q_i|\lambda) \Pr(O_{t+1}, O_{t+2}, \ldots, O_T, i_t = q_i|\lambda) \\ &= \frac{\alpha_t(i)\beta_t(i)}{\Pr O|\lambda}. \end{aligned} \tag{4.8}$$

From Equation (4.8), we define the backward variable $\beta_t(i)$ as

$$\beta_t(i) = \Pr(O_{t+1}O_{t+2}\ldots O_T|i_t = q_i, \lambda), \tag{4.9}$$

the probability of the partial observation sequence from $t+1$ to the end T given the state q_i at time t and the model λ. We solve for $\beta_t(i)$ inductively.

Firstly, initialise the procedure by setting

$$\beta_t(i) = 1, 1 \leq i \leq N. \tag{4.10}$$

Secondly, compute

$$\beta_t(i) = \sum_{j}^{N} a_{ij} b_j(O_{t+1}(j)) \tag{4.11}$$

for $t = T-1, T-2, \ldots, 1$.

4.3.2.4 The training problem

The third problem we have to solve is finding the parameters of $\lambda = (A, B, \pi)$ that maximise $Pr(O|\lambda)$. So we define the latent variable ξ_{ij}, the probability of a path being in state q_i at time t and making a transition to state q_j at time t_{i+1} given the observation sequence O and the model λ as follows

$$\xi_t(i,j) = \Pr(i_t = q_i, i_{t+1} = q_j | O, \lambda). \tag{4.12}$$

This can be written as

$$\begin{aligned} \xi_t(i,j) &= \frac{\alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}{\Pr(O|\lambda)} \\ &= \frac{\alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)}{\sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_t(i) a_{ij} b_j(O_{t+1}) \beta_{t+1}(j)} \end{aligned} \tag{4.13}$$

where $\Pr(O|\lambda)$ is a normalisation factor. Equation (4.13) is related to Equation (4.8) by summing $\xi_t(i,j)$ over j

$$\gamma_t(i) = \sum_{j=1}^{N} \xi_t(i,j). \tag{4.14}$$

The expected number of transitions made from state q_i is $\sum_{t=1}^{T-1} \gamma_t(i)$ where $\gamma_t(i)$ is a latent variable. The expected number of transitions made from state q_i to state q_j is $\sum_{t=1}^{T-1} \xi_t(i,j)$. We then use the Baum–Welch algorithm to update the parameters of the HMM until convergence. The re–estimated values of λ^* are defined as follows:

$$\pi_i^* = \gamma_1(i), \quad 1 \leq i \leq N, \tag{4.15}$$

$$a_{ij}^* = \frac{\sum_{t=1}^{T-1} \xi_t(i,j)}{\sum_{t=1}^{T-1} \gamma_t(i)} \tag{4.16}$$

and

$$b_j^*(k) = \frac{\sum_{t=1}^{T} 1_{O_{t=k}} \gamma_t(j)}{\sum_{t=1}^{T} \gamma_t(j)}. \tag{4.17}$$

Figure 4.2 shows the forward, emission and backward probabilities needed to calculate $\xi_t(i,j)$ in one step of training.

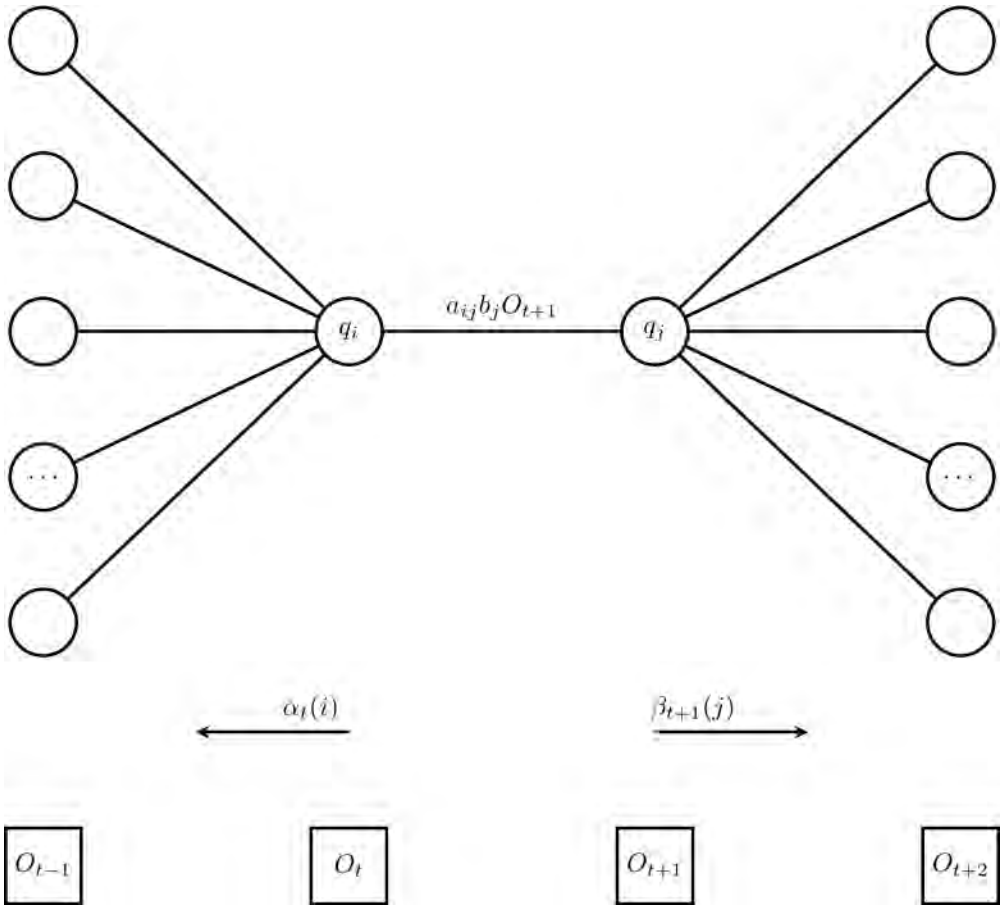


Figure 4.2: Forward, emission and backward probabilities required to calculate $\xi_t(i,j)$

Note that

$$\max_Q Q(\lambda, \lambda^*) \Rightarrow \Pr(O|\lambda^*) > \Pr(O|\lambda). \tag{4.18}$$

4.3.2.5 Distance

HMMs allow us to define a meaningful distance measure between trajectories through the use of probability densities that define trajectories. We use the Kullback–Leibler (KL) Divergence to calculate the distance between trajectories (Kullback & Leibler, 1951). The KL distance between probability densities, $\Pr(T|\lambda_i)$ and $\Pr(T|\lambda_j)$ is the distance between the trajectories λ_i and λ_j and is defined as follows:

$$D_{KL}(\Pr(T|\lambda_i) \parallel \Pr(T|\lambda_j)) = \int \Pr(T|\lambda_i) \log \frac{\Pr(T|\lambda_i)}{\Pr(T|\lambda_j)}\, dT, \tag{4.19}$$

which leads us to

$$D_{KL}(\Pr(T|\lambda_i) \parallel \Pr(T|\lambda_j)) \approx \frac{1}{n}\sum_{\alpha=1}^{N} \log \frac{\Pr(T_\alpha|\lambda_i)}{\Pr(T_\alpha|\lambda_j)}. \tag{4.20}$$

If we assume that the KL distance from Equation (4.19) is based on one data point so that $\Pr(T|\lambda_i)$ is concentrated on the observed trajectory T_i, then

$$D_{KL}(\Pr(T|\lambda_i) \parallel \Pr(T|\lambda_j)) \approx \log \frac{\Pr(T_i|\lambda_i)}{\Pr(T_i|\lambda_j)}. \tag{4.21}$$

This distance measure is asymmetric and we can make it symmetric as follows

$$D_{KL_{sym}}(\Pr(T|\lambda_i) \parallel \Pr(T|\lambda_j)) \approx \frac{1}{2}\Big(D_{KL}(\Pr(T|\lambda_i) \parallel \Pr(T|\lambda_j)) + D_{KL}(\Pr(T|\lambda_j) \parallel \Pr(T|\lambda_i))\Big). \tag{4.22}$$

This means that the KL distance is

$$D_{KL_{sym}}(\Pr(T|\lambda_i) \parallel \Pr(T|\lambda_j)) \approx \frac{1}{2}\left(\log \frac{\Pr(T_i|\lambda_i)}{\Pr(T_i|\lambda_j)} + \log \frac{\Pr(T_j|\lambda_j)}{\Pr(T_j|\lambda_i)}\right), \tag{4.23}$$

(see García-García et al., 2009).

For simplicity, we can combine the main principles from Equation (4.23) with Equation (4.20) to calculate the KL distance. Firstly, we define an N-dimensional vector $\tilde{\Pr}(\lambda)$ which is a probability mass function that can sufficiently describe the $\Pr(T|\lambda)$ over the set of N observed trajectories T_i,

$$\begin{aligned} \Pr(T|\lambda) \to \tilde{\Pr}(\lambda) &= \frac{1}{Z_\lambda}\Big\{\Pr(T_1|\lambda), \Pr(T_2|\lambda), \ldots, \Pr(T_N|\lambda)\Big\} \\ &= \Big\{\tilde{\Pr}(T_1|\lambda), \tilde{\Pr}(T_2|\lambda), \ldots, \tilde{\Pr}(T_N|\lambda)\Big\}, \end{aligned} \tag{4.24}$$

where Z is a normalising factor $Z_\lambda = \sum_{i=1}^{N} \Pr(T_i|\lambda)$ and $\tilde{\Pr}(T_i|\lambda) = \frac{1}{Z}\Pr(T_i|\lambda)$. Then the distance between HMM models λ_i and λ_j is

$$D_{KL}(\lambda_i, \lambda_j) = D_{KL}(\Pr_{\lambda_i} \parallel \Pr_{\lambda_j)} \approx \sum_{i=1}^{N} \tilde{\Pr}(T_i|\lambda_i) \log \frac{\tilde{\Pr}(T_i|\lambda_i)}{\tilde{\Pr}(T_i|\lambda_j}, \tag{4.25}$$

and the distance between λ_j and λ_i is

$$D_{KL}(\lambda_j, \lambda_i) = D_{KL}(\Pr_{\lambda_j} \parallel \Pr_{\lambda_i)} \approx \sum_{j=1}^{N} \tilde{\Pr}(T_j|\lambda_j) \log \frac{\tilde{\Pr}(T_j|\lambda_j)}{\tilde{\Pr}(T_j|\lambda_i}. \tag{4.26}$$

We use the average of Equation (4.25) and Equation (4.26) to calculate the KL divergence.

4.3.3 Clustering of HMMs using k–medoids clustering

K–medoids clustering is a partitioning method (Kaufman & Rousseeuw, 1987). It works by ensuring that the dissimilarity of objects and the nearest medoid is minimal. Dissimilarity measures how far away two objects are from each other. One has to specify the number of medoids to initialise the algorithm. We want to find the medoids $m_1, m_2, \ldots, m_k \subset \{1, 2, \ldots, N\}$ that minimises the objective function

$$\sum_{i=1}^{N} \min_{t,=1,2,\ldots,k} d(i, m_t). \tag{4.27}$$

An object is assigned to the cluster C_i corresponding to the nearest medoid

$$d(i, m_{C_i}) \leq d(i, m_w) \quad \forall w = (1, \ldots, k). \tag{4.28}$$

Object i is put into cluster v_i when medoid m_{v_i} is nearer to i than any other medoid m_w. A cluster index is then used to assess the quality of clustering. There are two steps in a k–medoids clustering, a building phase and then a swapping phase.

4.3.3.1 Algorithm

As described in Struyf et al. (1997) in the build phase, we create initial medoids where

$$m_1 \quad \text{is the smallest} \sum_{i=1}^{N} d(i, m_t),$$

and $m_2, \ldots, m_k$ decreases the objective function. In the swap phase, we take all pairs of (i, j) where $i \in \{m_1, m2, \ldots, m_k\}$ and $j \notin \{m_1, m_2, \ldots, m_k\}$ and swap i with j will to minimise the objective function. We repeat until we reach convergence.

4.3.3.2 Sihlouette index and plot

To interpret the cluster analysis we use silhouettes to visualise the results and a sihlouette index to quantify the quality of segmentation (Rousseeuw, 1987). We define the silhouette index $s(i)$ as follows:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad -1 \leq s(i) \leq 1, \tag{4.29}$$

where $a(i)$ is the average dissimilarity of i to all other objects of cluster A (intracluster distance)

$$a(i) = \frac{1}{|A| - 1} \sum_{j \in A, j \notin i} d(i, j), \tag{4.30}$$

and $b(i)$ is the smallest average dissimilarity of i to all objects of C, that is, intercluster distance

$$b(i) = \min_{C \neq A} d(i, C). \tag{4.31}$$

The average dissimilarity of i to all objects of any cluster C

$$d(i, C) = \frac{1}{|C|} \sum_{j \in C} d(i, j). \tag{4.32}$$

If $s(i) = 1$ this means that the i is well clustered while $s(i) = -1$ means that the i is far away from A than B. A $s(i) = 0$ means that the object is equally distant from A and B.

4.3.3.3 Dunn index

The Dunn index measures the level of separation and compactness of a partition (Dunn, 1973, 1974). If

$$\begin{aligned} C &= C_1, C_2, \ldots, C_p, \\ C_i \cap C_j &= \varnothing \quad \text{and} \\ C_i &= \varnothing, \end{aligned} \tag{4.33}$$

then the Dunn index (D) is

$$D(p, C) = \min_{1 \leq i \leq p} \left\{ \min_{1 \leq j \leq p, i \neq j} \left\{ \frac{\delta(C_i, C_j)}{\max_{1 \leq k \leq p} \Delta(C_k)} \right\} \right\}, \tag{4.34}$$

where the intercluster distance is $\delta(C_i, C_j)$ and the intracluster distance is $\Delta(C_k)$.

4.3.3.4 Davies–Bouldin index

The Davies–Bouldin index (DB) measures the similarity of clusters (Davies & Bouldin, 1979). The smallest DB value indicates the recommended number of clusters as it minimises the average similarity of clusters. Therefore,

$$DB(p, C) = \frac{1}{p} \sum_{i=1}^{p} \max_{i \neq j} \left\{ \frac{\Delta(C_i) + \Delta(C_j)}{\delta(C_i, C_j)} \right\}, \tag{4.35}$$

where $\delta(C_i, C_j)$ is the intercluster distance, $\Delta(C_i)$ and $\Delta(C_j)$ are the intracluster distances of clusters C_i and C_j, respectively.

4.3.4 Extending the three state functional disability model

Following the proportional hazard specification in Z. Li et al. (2017), we model the transition intensity type s of type $s = 1, \ldots, S$ for an individual k for $k = 1, \ldots, K$ at time t years with

$$\lambda_{k,s}(t) = \exp(\beta_s + \gamma_s' w_k(t) + \alpha_s \psi(t)) H_{k,s}(t), \tag{4.36}$$

where β_s is the time invariant baseline log–intensity for transition type s, $w_k(t)$ is a vector of the observed predictors for each individual k, $\psi(t)$ is frailty which is a stochastic latent

process, γ_s is a vector measuring the sensitivity of $\lambda_{k,s}(t)$ with respect to $w_k(t)$, α_s is a scalar measuring the sensitivity of $\lambda_{k,s}(t)$ with respect to $\psi(t)$ and the baseline hazard function for duration dependence $H_{k,s}(t) = 1$ (Koopman et al., 2008a). Figure 4.3 shows the three state functional disability model that we will use in this analysis.

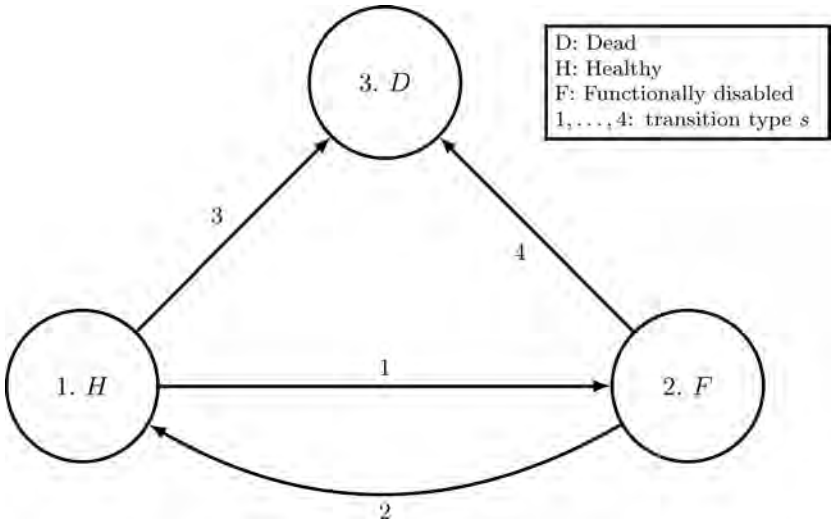


Figure 4.3: Three state functional disability model

Similar to Z. Li et al. (2017) we consider three models: a static model, a static model with a linear time trend, and a frailty model with time trend. However, we compare all the models under 2 cases, one with clustering and another without. The conversion of the calendar year to time is presented in Table 4.2.

For the static model, the transition rate $\lambda_{k,s}(t)$ is assumed to be dependent on age and sex only:

$$\begin{aligned} \ln \lambda_{k,s} &= \beta_s + \gamma_s^{age} x_k(t) + \gamma_s^{female} F_k \\ \ln \lambda_{k,s} &= \beta_s + \gamma_s^{age} x_k(t) + \gamma_s^{female} F_k + \gamma_s^{cluster} C_k \end{aligned} \tag{4.37}$$

where β_s is the time invariant baseline log–intensity for transition type s, $x_k(t)$ is the k^{th} individual's age at time t, F_k is the binary variable indicating the gender for the individual k, C_k is the categorical variable indicating the cluster for the individual k, γ_s^{age} measures the sensitivity of $\ln \lambda_{k,s}(t)$ with respect to age and $\gamma_s^{cluster}$ measures the sensitivity of $\ln \lambda_{k,s}(t)$ with respect to cluster α_s is a scalar measuring the sensitivity of $\ln \lambda_{k,s}(t)$ with respect to sex.

For the model with systematic trend,

$$\begin{aligned}\ln \lambda_{k,s} &= \beta_s + \gamma_s^{age} x_k(t) + \gamma_s^{female} F_k + \phi_s^{time} t \\ \ln \lambda_{k,s} &= \beta_s + \gamma_s^{age} x_k(t) + \gamma_s^{female} F_k + \phi_s^{time} t + \gamma_s^{cluster} C_k,\end{aligned} \tag{4.38}$$

where ϕ_s measures the sensitivity of $\ln \lambda_{k,s}(t)$ with respect to the time trend t.

For the model with systematic trend and uncertainty,

$$\begin{aligned}\ln \lambda_{k,s} &= \beta_s + \gamma_s^{age} x_k(t) + \gamma_s^{female} F_k + \phi_s^{time} t + \alpha_s \psi_i \\ \ln \lambda_{k,s} &= \beta_s + \gamma_s^{age} x_k(t) + \gamma_s^{female} F_k + + \phi_s^{time} t + \alpha_s \psi_i + \gamma_s^{cluster} C_k\end{aligned} \tag{4.39}$$

where α_s measures the sensitivity of $\ln \lambda_{k,s}(t)$ with respect to the latent factor ψ that is modelled as a random walk,

$$\psi_j = \psi_{j-1} + \epsilon_j, \quad \epsilon_j \sim N(0, \sigma^2), \quad \psi_0 = 0 \quad \text{and} \quad \sigma^2 = t_j - t_{j-1}. \tag{4.40}$$

We estimate the transition rates using code developed in Fu et al. (2021) that is available at Functional Disability China US. We make some adjustments to allow the incorporation of clusters.

Table 4.2: Conversion of calendar year to time for HRS

Year	Time
1998–1999	1
2000–2001	3
2002–2003	5
2004–2005	7
2006–2007	9
2008–2009	11
2010–2011	13
2012–2013	15
2014–2015	17
2016–2017	19

4.4 Results

We now provide the results. Table 4.3 shows the cluster quality assessment across 3 indices where the response is the presence of any one of the four chronic diseases and BMI as the covariate. Table 4.4 shows the cluster quality assessment across 3 indices where the response is the presence of any one of the four chronic diseases and the covariates are BMI and self–reported health. Optimal cluster solution minimises the Davies Bouldin Index (a measure of similiarity of the clusters) and maximises both the Sihlouette and Dunn indices. The Sihlouette index quantifies the quality of segmentation while the Dunn index is a measure of compactness and separation of a cluster. However, there is no consensus amongst the different criteria as shown in the Table 4.3 and 4.4. The following results are based on opting for a 3 cluster solution from Table 4.4.

Table 4.3: Cluster quality assessment with BMI as the only covariate

Clusters	Sihlouette	Dunn	Davies Bouldin
2	0.744	0.019	0.394
3	0.519	0.019	0.962
4	0.646	0.051	0.560
5	0.655	0.026	0.590
6	0.591	0.024	0.721
7	0.558	0.016	0.797
8	0.416	0.003	0.893
9	0.426	0.003	0.819
10	0.423	0.006	0.835

Table 4.4: Cluster quality assessment with covariates of BMI and self reported health status

Clusters	Sihlouette	Dunn	Davies Bouldin
2	0.352	0.015	0.935
3	0.449	0.017	1.005
4	0.617	0.014	0.653
5	0.644	0.012	0.686
6	0.703	0.018	0.570
7	0.746	0.033	0.517
8	0.763	0.023	0.489
9	0.784	0.018	0.408
10	0.675	0.018	0.542

4.4.1 Cluster profiles

We summarise the characteristics of the individuals in each cluster using information from the first wave in Table 4.5. Age wise, individuals in Cluster 2 are slightly younger than all the other clusters. Individuals in Cluster 3 own homes with the largest values,

have the highest household income and have the highest proportion of individuals who graduated from college. All clusters have more males than females. Cluster 1 has the lowest proportion of people in excellent or very good health and the highest proportion of people who are obese. Individuals in Cluster 2 generally have mean proportions or values close to the Overall (model with no clusters).

Table 4.5: Baseline summary statistics for wave 1 in 1992 using three clusters estimated from hidden Markov models with covariates of BMI and self reported health

Description	Cluster 1	Cluster 2	Cluster 3	Overall[a]
Socio–demographic				
Year of birth	1934	1935	1934	1935
Wealth and income				
Value of primary residence	$69,258	$71,320	$71,911	$71,074
Total non-housing wealth	$106,968	$120,069	$109,048	$115,193
Total household income	$40,469	$37,842	$41,771	$39,205
Education				
Lt High-school[b]	33.01%	35.94%	29.37%	33.93%
GED[c]	8.74%	5.41%	6.35%	6.23%
High School Graduate	24.27%	32.33%	30.56%	30.45%
Some college	22.33%	15.64%	18.65%	17.54%
College and above	11.65%	10.68%	15.08%	11.84%
Gender				
Male	55.83%	55.64%	57.54%	56.10%
Female	44.17%	44.36%	42.46%	43.90%
Body mass index				
Underweight	0.49%	0.90%	0.79%	0.80%
Normal weight	24.76%	25.56%	26.59%	25.65%
Overweight	47.57%	50.68%	54.76%	51.02%
Obese	27.18%	22.86%	17.86%	22.53%
Health behaviour				
Ever drinks	58.25%	61.05%	59.52%	60.20%
Marital status				
Married	72.82%	73.38%	78.57%	74.44%
Married, spouse absent	0.49%	0.45%	0.00%	0.36%
Partnered	4.37%	3.46%	3.57%	3.65%

Table 4.5: Baseline summary statistics for wave 1 in 1992 using three clusters estimated from hidden Markov models with covariates of BMI and self reported health *(continued)*

Description	Cluster 1	Cluster 2	Cluster 3	Overall[a]
Separated	1.46%	2.41%	2.38%	2.23%
Divorced	10.19%	10.98%	10.71%	10.77%
Widowed	5.83%	6.47%	3.97%	5.79%
Never married	4.85%	2.86%	0.79%	2.76%
Self reported health				
Excellent	9.22%	14.44%	27.38%	16.38%
Very good	24.27%	28.72%	31.75%	28.58%
Good	49.51%	32.63%	25.40%	34.11%
Fair	13.11%	16.84%	13.10%	15.32%
Poor	3.88%	7.37%	2.38%	5.61%
Smoking status				
Current smoker	41.75%	44.81%	40.08%	43.19%

[a] Overall represents the model with no clusters;

[b] Lt-High school means left high school without graduation

[c] GED means graduated high school by taking a General Education Development Test.

4.4.2 Multistate model results

Table 4.6 presents the estimated parameters for the static model specified in Equation (4.37). Disability rates (H→F) increase with age and females are more likely to become disabled than males. Recovery rates (F→H) decrease with age and females are more likely to recover. Mortality rates (H→D,F→D) increase with age and females tend to live longer than males. All these findings corroborate results from earlier studies (Fong et al., 2015; Z. Li et al., 2017; Shao et al., 2017; Sherris & Wei, 2021). However, the size of the coefficients for functional disability and mortality for the functionally disabled are smaller than those reported in Fu et al. (2021). This could be due to the size of our sample. Table 4.7 reports the estimated parameters for the static model with clustering. We find that the functional disability rates from the healthy state decrease with clustering. By contrast, while the transition rates from the functionally disabled to the healthy state vary with clustering, this effect is not significant.

Table 4.8 reports the estimated parameters for the trend model specified in Equation (4.38). Functional disability rates and mortality rates increase with time. This is in contrast to Fu et al. (2021), where disability rates and mortality rates decrease with time. However, this effect of time is very significant ($p < 0.01$), highlighting that individuals who acquire chronic diseases tend to deteriorate with time. Another anomaly is that females are less likely to recover from disability than males as shown by the negative coefficient. Nevertheless, this effect is not significant.

Table 4.9 shows the estimated parameters for the trend model with clustering. We note that individuals in Cluster 3 are more likely to recover from disability than those in Cluster 1 while individuals in Cluster 2 are less likely to recover from disability than those in Cluster 1. Belonging to Clusters 2 and 3 results in lower chance of becoming functionally disabled when compared to Cluster 1. Individuals in Cluster 2 have higher mortality rates than individuals in Cluster 1 while individuals in Cluster 3 have lower mortality rates from the functionally disabled state than those in Cluster 1.

Table 4.10 shows the estimated parameters for the frailty model specified in Equation (4.39). Functional disability and mortality rates have significant systematic uncertainty. Fu et al. (2021) show than only the functional disability rates have significant uncertainty and the size of the stochastic frailty factor is close to zero for all other transitions. This is likely due to the notion that the risk profiles of the samples used to estimate the models are very different. Table 4.11 shows the estimated parameters for the frailty model with clustering. All parameters exhibit similar behaviour to the trend model with clustering. However, there is greater uncertainty in disability, mortality and recovery rates.

Table 4.6: Static model without clustering: estimated parameters with standard errors in parentheses

Transition	H→F		F→H		H→D		F→D	
s	1		2		3		4	
$\hat{\beta}_s$	-3.6100	***	-2.0366	***	-3.4748	***	-2.2359	***
	(0.0603)		(0.0879)		(0.0587)		(0.0756)	
$\hat{\gamma}_s^{\text{age}}$	0.5081	***	-0.3927	***	0.6377	***	0.4537	***
	(0.0311)		(0.0442)		(0.0318)		(0.0332)	
$\hat{\gamma}_s^{\text{female}}$	0.3147	***	0.0226		-0.2470	***	-0.2952	***
	(0.0567)		(0.0977)		(0.0548)		(0.0616)	
Log likelihood	-14,361							

Note: Age covariate is calculated using age last birthday.

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 4.7: Static model with clustering: estimated parameters with standard errors in parentheses

Transition	H→F		F→H		H→D		F→D	
s	1		2		3		4	
$\hat{\beta}_s$	-3.2831	***	-2.0241	***	-3.6716	***	-2.2016	***
	(0.1070)		(0.1687)		(0.1324)		(0.1197)	
$\hat{\gamma}_s^{\text{age}}$	0.5088	***	-0.3896	***	0.6327	***	0.4497	***
	(0.0312)		(0.0444)		(0.0319)		(0.0333)	
$\hat{\gamma}_s^{\text{female}}$	0.3022	***	0.0193		-0.2282	***	-0.2942	***
	(0.0569)		(0.0977)		(0.0550)		(0.0617)	
$\hat{\gamma}_s^{\text{cluster2}}$	-0.2905	***	-0.1123		0.4760	***	0.0563	
	(0.1065)		(0.1795)		(0.1302)		(0.1120)	
$\hat{\gamma}_s^{\text{cluster3}}$	-0.3677	***	0.0198		0.0950		-0.0705	
	(0.0965)		(0.1610)		(0.1254)		(0.1053)	
Log likelihood	-14,330							

Note: Age covariate is calculated using age last birthday.

‡ Reference level for Cluster 2 and Cluster 3 is Cluster 1;

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 4.8: Trend model without clustering: estimated parameters with standard errors in parentheses

Transition	H→F		F→H		H→D		F→D	
s	1		2		3		4	
$\hat{\beta}_s$	-3.7414	***	-1.4331	***	-5.1971	***	-3.1831	***
	(0.0709)		(0.1065)		(0.0880)		(0.1015)	
$\hat{\gamma}_s^{\text{age}}$	0.4776	***	-0.2905	***	0.3114	***	0.3580	***
	(0.0322)		(0.0448)		(0.0328)		(0.0342)	
$\hat{\gamma}_s^{\text{female}}$	0.3266	***	-0.0301		-0.1302	**	-0.2115	***
	(0.0567)		(0.0983)		(0.0546)		(0.0619)	
$\hat{\gamma}_s^{\text{time}}$	0.2352	***	-0.9571	***	2.3323	***	1.0764	***
	(0.0643)		(0.1117)		(0.0674)		(0.0681)	
Log likelihood	-13,525							

Note: Age covariate is calculated using age last birthday.

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 4.9: Trend model with clustering: estimated parameters with standard errors in parentheses

Transition	H→F		F→H		H→D		F→D	
s	1		2		3		4	
$\hat{\beta}_s$	-3.4226	***	-1.4370	***	-5.5091	***	-3.1672	***
	(0.1135)		(0.1778)		(0.1479)		(0.1369)	
$\hat{\gamma}_s^{\text{age}}$	0.4773	***	-0.2712	***	0.3140	***	0.3418	***
	(0.0322)		(0.0455)		(0.0326)		(0.0343)	
$\hat{\gamma}_s^{\text{female}}$	0.3158	***	-0.0433		-0.1035	*	-0.2039	***
	(0.0570)		(0.0984)		(0.0547)		(0.0620)	
$\hat{\gamma}_s^{\text{cluster2}}$	-0.2697	**	-0.1973		0.6893	***	0.1878	*
	(0.1066)		(0.1799)		(0.1300)		(0.1121)	
$\hat{\gamma}_s^{\text{cluster3}}$	-0.3712	***	0.1230		0.0998		-0.1989	*
	(0.0965)		(0.1614)		(0.1253)		(0.1054)	
$\hat{\gamma}_s^{\text{time}}$	0.2433	***	-1.0182	***	2.3758	***	1.1561	***
	(0.0646)		(0.1141)		(0.0672)		(0.0692)	
Log likelihood	-13,450							

Note: Age covariate is calculated using age last birthday.

‡ Reference level for Cluster 2 and Cluster 3 is Cluster 1;

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 4.10: Frailty model without clustering: estimated parameters with standard errors in parentheses

Transition	H→F		F→H		H→D		F→D	
s	1		2		3		4	
$\hat{\beta}_s$	-3.6783	***	-1.3683	***	-5.2663	***	-3.2166	***
	(0.0749)		(0.1157)		(0.0858)		(0.1000)	
$\hat{\gamma}_s^{\text{age}}$	0.4774	***	-0.2907	***	0.3153	***	0.3494	***
	(0.0322)		(0.0449)		(0.0328)		(0.0341)	
$\hat{\gamma}_s^{\text{female}}$	0.3276	***	-0.0318		-0.1353	**	-0.2057	***
	(0.0567)		(0.0983)		(0.0547)		(0.0619)	
$\hat{\gamma}_s^{\text{time}}$	0.1268		-1.0671	***	2.4808	***	1.1786	***
	(0.0775)		(0.1368)		(0.0672)		(0.0696)	
$\hat{\alpha}_s$	-0.0741	**	-0.0745		0.1799	***	0.1190	***
	(0.0288)		(0.0521)		(0.0246)		(0.0263)	
Log likelihood	-13,485							

Note: Age covariate is calculated using age last birthday.

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

Table 4.11: Frailty model with clustering: estimated parameters with standard errors in parentheses

Transition	H→F		F→H		H→D		F→D	
s	1		2		3		4	
$\hat{\beta}_s$	-3.1092	***	-1.1362	***	-5.5344	***	-3.2034	***
	(0.1314)		(0.2213)		(0.1444)		(0.1343)	
$\hat{\gamma}_s^{\text{age}}$	0.4748	***	-0.2706	***	0.3209	***	0.3311	***
	(0.0322)		(0.0455)		(0.0326)		(0.0341)	
$\hat{\gamma}_s^{\text{female}}$	0.3159	***	-0.0424		-0.1083	**	-0.1984	***
	(0.0570)		(0.0984)		(0.0547)		(0.0620)	
$\hat{\gamma}_s^{\text{cluster2}}$	-0.2660	**	-0.1985		0.6799	***	0.1899	*
	(0.1066)		(0.1799)		(0.1301)		(0.1121)	
$\hat{\gamma}_s^{\text{cluster3}}$	-0.3684	***	0.1175		0.0935		-0.1824	*
	(0.0965)		(0.1615)		(0.1253)		(0.1055)	
$\hat{\gamma}_s^{\text{time}}$	-0.0183		-1.2712	***	2.2852	***	1.1142	***
	(0.0888)		(0.1655)		(0.0624)		(0.0658)	
$\hat{\alpha}_s$	-0.1827	***	-0.1652	**	0.1953	***	0.1321	***
	(0.0362)		(0.0700)		(0.0215)		(0.0229)	
Log likelihood	-13,382							

Note: Age covariate is calculated using age last birthday.

‡ Reference level for Cluster 2 and Cluster 3 is Cluster 1;

$^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

4.4.3 Model performance

Table 4.12 show the AIC and BIC for the static, trend and frailty models. We note that trend and frailty models with clustering have a better fit than those without clustering as shown by the smaller values of AIC and BIC. The frailty model with clustering has the best fit. We cannot say which of the static models performs the best since the AIC and BIC are not in consensus. However, the static model has the worst fit amongst all of the models. The model selection criteria are calculated as follows:

$$AIC = 2\kappa - 2\ln\hat{L}, \qquad BIC = \kappa \ln n - 2\ln\hat{L},$$

where κ is the number of estimated parameters in the model, $\hat{L}$ is the maximum likelihood of the model and n is the number of observations used to fit the model.

Table 4.12: Model selection using Akaike information criterion (AIC) and Bayesian information criterion (AIC)

	No Clustering		**Clustering**	
Model	**AIC**	**BIC**	**AIC**	**BIC**
Static	28,745	28,850	28,700	28,874
Trend	27,082	27,222	26,947	27,156
Frailty	27,010	27,184	26,819	27,063

Table 4.13 presents the results of the likelihood tests. The null model is the model without the clustering variable. All p–values are close to zero and this suggests that there is very strong evidence to reject the null models in favor of the alternative models. The likelihood tests also confirm that the static model does not provide a good fit to the empirical data. The frailty model with clustering outperforms all the other models.

Table 4.13: Likelihood tests for models with and without clustering

Null model	**Alternative model**	**p–Value**	**Chi–square**	**Symbol**[1]
Static	Static with Clustering	2.8842e-10	61.1	***
Trend	Trend with Clustering	0.0000e+00	150.9	***
Frailty	Frailty with Clustering	0.0000e+00	206.4	***
Static with Clustering	Trend with Clustering	0.0000e+00	1760.6	***
Trend with Clustering	Frailty with Clustering	0.0000e+00	136.2	***

[1] *Note:* $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$.

4.4.4 Future lifetime statistics

Table 4.14 shows the future lifetime statistics for healthy 65 year old individuals across different clusters and the model with no clusters using simulated results from the esti-

mated static parameters in Tables 4.6 and 4.7. The overall mean life expectancy for males and females is 14.62 years and 13.17 years, respectively. This is lower than the values (20.17 years and 17.69 years) calculated in Fu et al. (2021), showing that the sample in this analysis is negatively skewed towards individuals with lower life expectancy. This makes sense since we focus on people with chronic diseases and these individuals are likely to have a lower life expectancy overall. However, even amongst these individuals with various health issues, we observe heterogeneity in life expectancy, healthy life expectancy, the onset of disability and the proportion of years spent in a healthy state. Cluster 3 has the most favourable future lifetime statistics and Cluster 1 has the worst mortality experience. People in Cluster 2 have lower life expectancy and healthier life expectancy than the overall model. The mean life expectancy for the model with no clusters for healthy 75 year olds is 9.81 years and 8.68 years for males and females, respectively.

Table 4.15 presents the future lifetime statistics for healthy 65 year old individuals starting in 1998 across different clusters and the model with no clusters using simulated results from the estimated trend models in Tables 4.8 and 4.9. The overall mean life expectancy for males and females is 12.33 years and 11.94 years, respectively. These values are less than those reported using the static model because the static model fails to capture the mortality deterioration in the sample. Regardless, we notice significant heterogeneity in the longevity and morbidity experience across clusters. Cluster 3 outperforms all the clusters and the model with no clusters in life expectancy and healthy life expectancy. Cluster 2 has the lowest life expectancy for both males and females while Cluster 1 has the lowest HLE/TLE ratio (Healthy Life Expectancy/Total Life Expectancy). Individuals in Cluster 1 spend more time disabled than others because they tend not to recover from disability when compared to other clusters. They also tend have an earlier disability onset. Females spend more years disabled than males across clusters. Table 4.18 shows results analogous to the results for 65–year olds but for 75–year olds.

Table 4.14: Static model with clustering: Future lifetime statistics for healthy 65–year old males and females across clusters

	Cluster 1		Cluster 2		Cluster 3		Overall[a]	
	Female	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**
Total future lifetime								
Mean	12.823	11.362	13.952	12.451	15.146	13.578	14.624	13.174
Standard error	0.078	0.072	0.083	0.077	0.088	0.082	0.086	0.080
Standard deviation	7.844	7.214	8.314	7.728	8.761	8.228	8.558	8.044
Healthy future lifetime								
Mean	9.881	9.630	10.963	10.684	12.136	11.809	11.636	11.379
Standard error	0.066	0.066	0.071	0.071	0.076	0.076	0.074	0.074
Standard deviation	6.622	6.586	7.121	7.109	7.636	7.638	7.413	7.441
Disabled future lifetime								
Mean	2.942	1.732	2.989	1.768	3.010	1.769	2.988	1.795
Standard error	0.043	0.032	0.043	0.032	0.044	0.032	0.043	0.032
Standard deviation	4.264	3.148	4.340	3.191	4.380	3.185	4.342	3.222
Healthy future lifetime over total future lifetime								
Mean	0.804	0.867	0.816	0.875	0.827	0.883	0.822	0.879
Standard error	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
Standard deviation	0.253	0.223	0.242	0.212	0.231	0.201	0.236	0.206
Age at onset of disability conditional on becoming disabled								
Mean	74.446	74.353	75.300	75.233	76.285	76.403	75.831	75.952
Standard error	0.086	0.099	0.092	0.108	0.100	0.117	0.096	0.113
Standard deviation	6.274	6.193	6.724	6.730	7.257	7.320	7.010	7.101

Note: Simulation results for 10,000 individuals with maximal age of 110 years.

[a] Overall represents the model with no clusters.

Table 4.15: Trend model with clustering: Future lifetime statistics for healthy 65–year old males and females across clusters in 1998

	Cluster 1		Cluster 2		Cluster 3		Overall[a]	
	Female	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**
Total future lifetime								
Mean	12.503	12.272	10.881	10.553	12.837	12.515	12.328	11.938
Standard error	0.044	0.043	0.042	0.040	0.044	0.043	0.044	0.043
Standard deviation	4.412	4.328	4.151	4.017	4.428	4.307	4.425	4.314
Healthy future lifetime								
Mean	9.818	10.455	9.063	9.354	10.710	11.092	10.224	10.546
Standard error	0.047	0.047	0.042	0.042	0.046	0.045	0.045	0.045
Standard deviation	4.656	4.660	4.219	4.148	4.609	4.530	4.536	4.462
Disabled future lifetime								
Mean	2.685	1.817	1.818	1.199	2.127	1.423	2.104	1.393
Standard error	0.035	0.029	0.030	0.024	0.034	0.027	0.033	0.027
Standard deviation	3.484	2.860	3.002	2.407	3.359	2.732	3.307	2.673
Healthy future lifetime over total future lifetime								
Mean	0.792	0.851	0.846	0.892	0.844	0.890	0.841	0.888
Standard error	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
Standard deviation	0.254	0.226	0.240	0.208	0.233	0.202	0.236	0.205
Age at onset of disability conditional on becoming disabled								
Mean	73.116	73.350	72.531	72.622	73.506	73.657	73.298	73.403
Standard error	0.058	0.066	0.062	0.072	0.068	0.078	0.066	0.076
Standard deviation	4.390	4.444	3.959	3.994	4.507	4.510	4.428	4.393

Note: Simulation results for 10,000 individuals with maximal age of 110 years.

[a] Overall represents the model with no clusters.

Table 4.16 shows the future lifetime statistics for healthy 65 year old individuals starting in 1998 across different clusters and the model with no clusters using simulated results from the estimated frailty parameters in Tables 4.10 and 4.11. The overall mean life expectancy for males and females is 12.11 years and 11.71 years, respectively. These values are slightly lower than the values from the 1998 trend model. As observed earlier, Cluster 3 has the highest life expectancy, healthy life expectancy and age at onset of disability. Cluster 1 has the worst mortality experience and greatest number of years spent in disability. Males spend most of their future lifetime healthy while females spend a lesser proportion healthy. This is true across clusters with Cluster 2 having a longevity and morbidity experience quite similar to the overall cluster. We observe

similar results when we simulate healthy 75–year olds. The results are shown in Table 4.19.

Table 4.16: Frailty model with clustering: Future lifetime statistics for healthy 65–year old males and females across clusters in 1998

	Cluster 1		Cluster 2		Cluster 3		Overall[a]	
	Female	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**
Total future lifetime								
Mean	12.803	12.555	11.217	10.878	13.248	12.889	12.110	11.713
Standard error	0.002	0.001	0.001	0.001	0.002	0.002	0.001	0.001
Standard deviation	4.655	4.520	4.352	4.206	4.678	4.526	4.282	4.144
Healthy future lifetime								
Mean	9.468	10.168	8.844	9.236	10.483	10.952	10.000	10.302
Standard error	0.001	0.002	0.001	0.001	0.002	0.001	0.001	0.001
Standard deviation	4.569	4.607	4.220	4.196	4.574	4.549	4.396	4.312
Disabled future lifetime								
Mean	3.335	2.387	2.374	1.641	2.765	1.936	2.110	1.411
Standard error	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Standard deviation	3.719	3.124	3.297	2.690	3.611	2.981	3.211	2.606
Healthy future lifetime over total future lifetime								
Mean	0.758	0.821	0.814	0.865	0.814	0.864	0.841	0.888
Standard error	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Standard deviation	0.255	0.230	0.248	0.217	0.235	0.208	0.234	0.204
Age at onset of disability conditional on becoming disabled								
Mean	72.484	72.884	72.076	72.279	73.077	73.333	72.962	73.049
Standard error	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002
Standard deviation	4.265	4.354	3.902	3.918	4.427	4.463	4.252	4.219

Note: Simulation results for 10,000 individuals with maximal age of 110 years.

[a] Overall represents the model with no clusters.

Table 4.17: Comparison of future lifetime statistics for healthy 65–year old males and females across clusters using trend and frailty models

	Cluster 1[a]		Cluster 2[b]		Cluster 3[c]	
Change	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**
Trend in 1998						
Total future lifetime	101.42%	102.79%	88.26%	88.40%	104.13%	104.83%
Healthy future lifetime	96.03%	99.14%	88.64%	88.70%	104.76%	105.19%
Disabled future lifetime	127.60%	130.47%	86.40%	86.06%	101.10%	102.15%
Age at onset of disability	99.75%	99.93%	98.95%	98.94%	100.28%	100.35%
Frailty in 1998						
Total future lifetime	105.72%	107.19%	92.63%	92.87%	109.40%	110.03%
Healthy future lifetime	94.68%	98.70%	88.44%	89.65%	104.83%	106.31%
Disabled future lifetime	158.09%	169.24%	112.53%	116.35%	131.08%	137.27%
Age at onset of disability	99.35%	99.78%	98.79%	98.95%	100.16%	100.39%

Note: Age at onset of disability is conditional on becoming disabled.

[a] Percentage difference in future lifetime betweeen Cluster 1 and model with no clusters of the same sex;

[b] Percentage difference in future lifetime betweeen Cluster 2 and model with no clusters of the same sex;

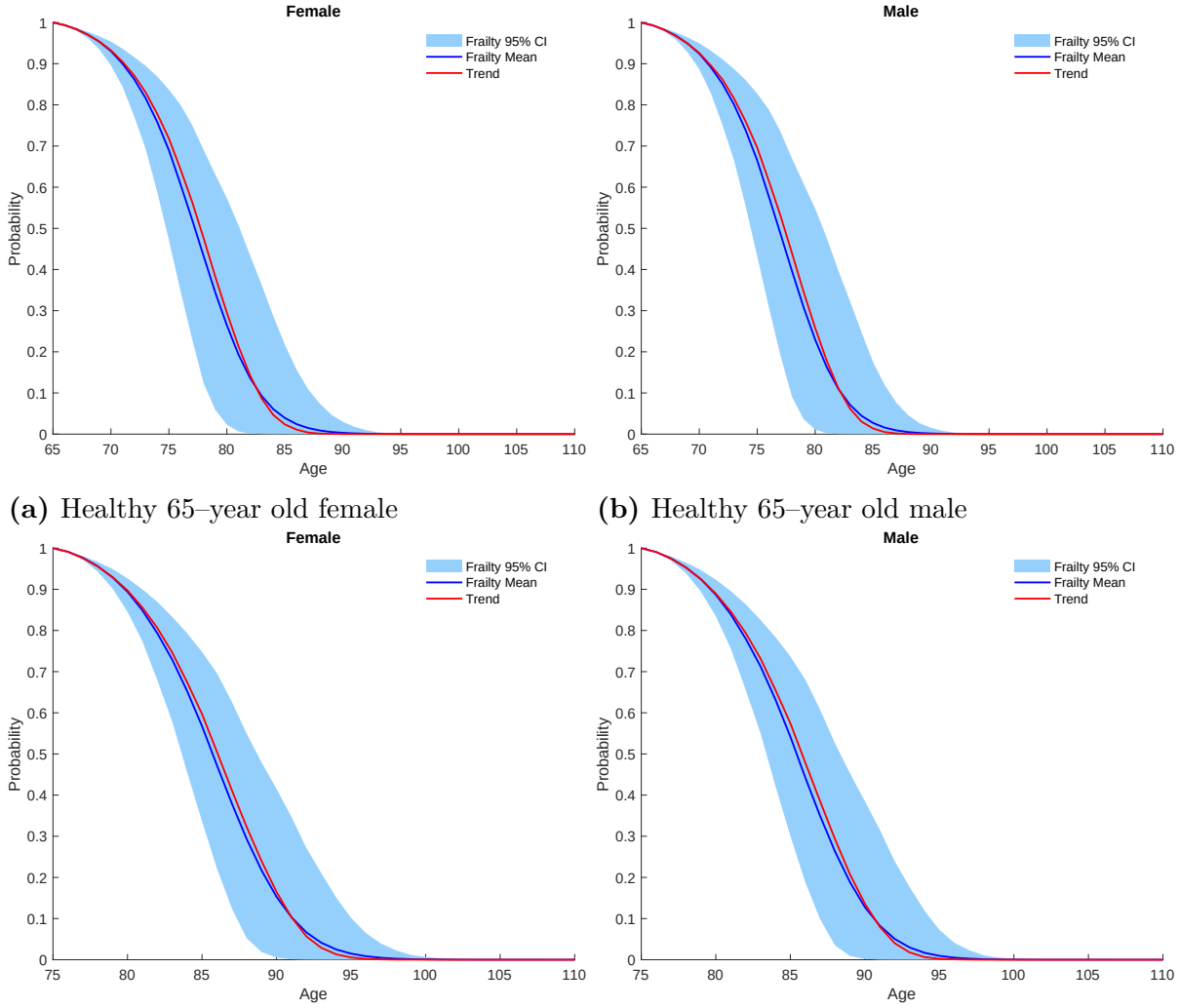
[c] Percentage difference in future lifetime betweeen Cluster 3 and model with no clusters of the same sex.

Table 4.17 shows comparisons of future lifetime statistics across the three clusters. All the clusters are compared to the relevant model with no clusters. We find differences in mean life expectancy and healthy life expectancy. Individuals in Cluster 3 have the highest total future lifetime, healthy future lifetime and age at onset of disability conditional on being disabled for trend and frailty models. Individuals in Cluster 2 have the lowest total future lifetime and healthy future lifetime. They also experience disability earlier than all other groups. Cluster 1 generally experiences better life expectancy than Cluster 2 but worse experience than Cluster 3. In terms of disability, Cluster 1 spends the largest amount of time disabled while Cluster 3 spends the least. The results in Table 4.20 are similar to those in Table 4.17 except that they are for healthy individuals aged 75.

4.4.5 Survival curves of trend and frailty models

Figures 4.4 and 4.5 illustrate the survival curves of the trend and frailty models for the models excluding clusters and including clusters, respectively. As expected, healthy

females in 1998 have higher survival probabilities and more uncertainty than males at ages 65 and 75. Amongst all the clusters we note that there is great overlap between the male and female survival confidence intervals even though females slightly outlive males. Figure 4.6 presents the probability of being disabled for healthy males and females across clusters at age 65. Females are generally more likely to be disabled than males. However, we note that individuals in Cluster 1 are more likely to become disabled earlier than individuals in Clusters 2 and 3. The bell curve is due to the fact that as time increases, mortality deterioration exceeds the impact of disability. Note that the sample is includes very sick individuals who have relatively low longevity prospects than average.



(a) Healthy 65–year old female

(b) Healthy 65–year old male

(c) Healthy 75–year old female

(d) Healthy 75–year old male

Figure 4.4: Survival curves of trend and frailty models for healthy 65 and 75 year old males and females for the case with no clusters in 1998

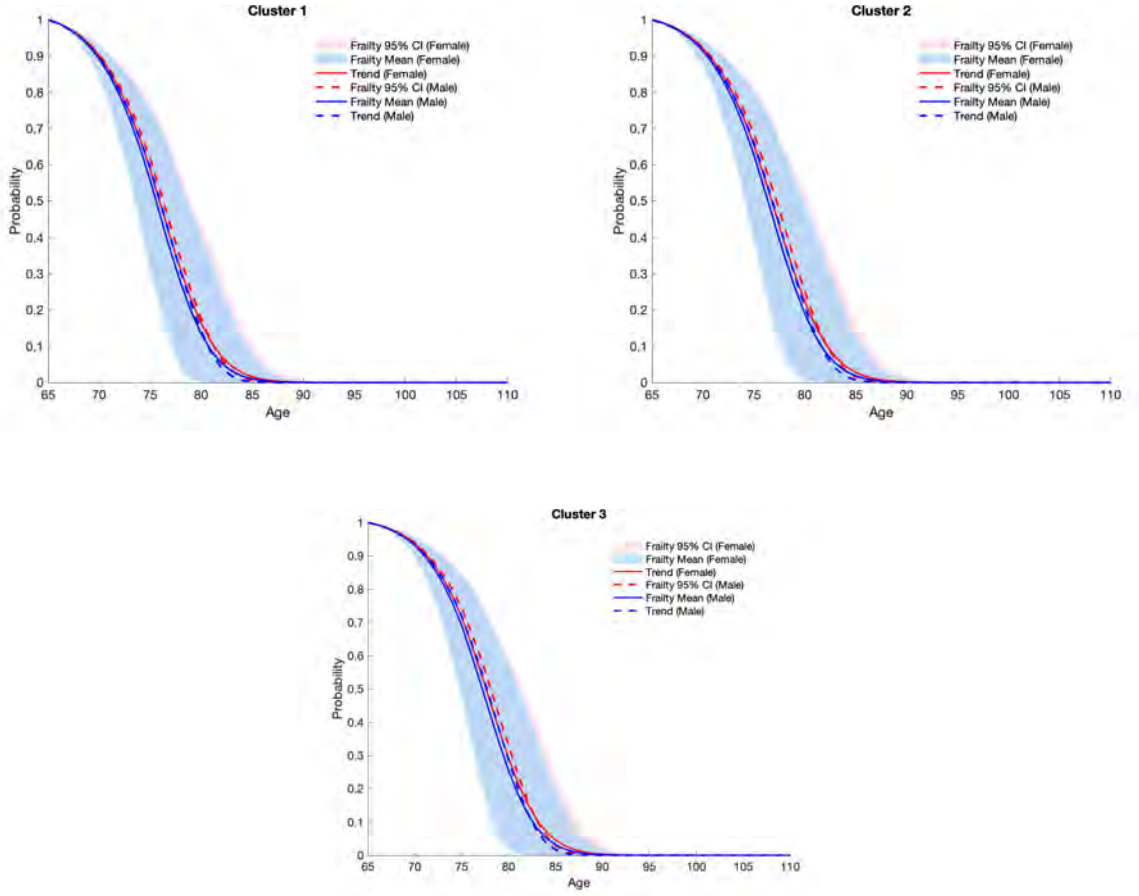


Figure 4.5: Survival curves of trend and frailty models across clusters in 1998 for healthy 65 year old males and females

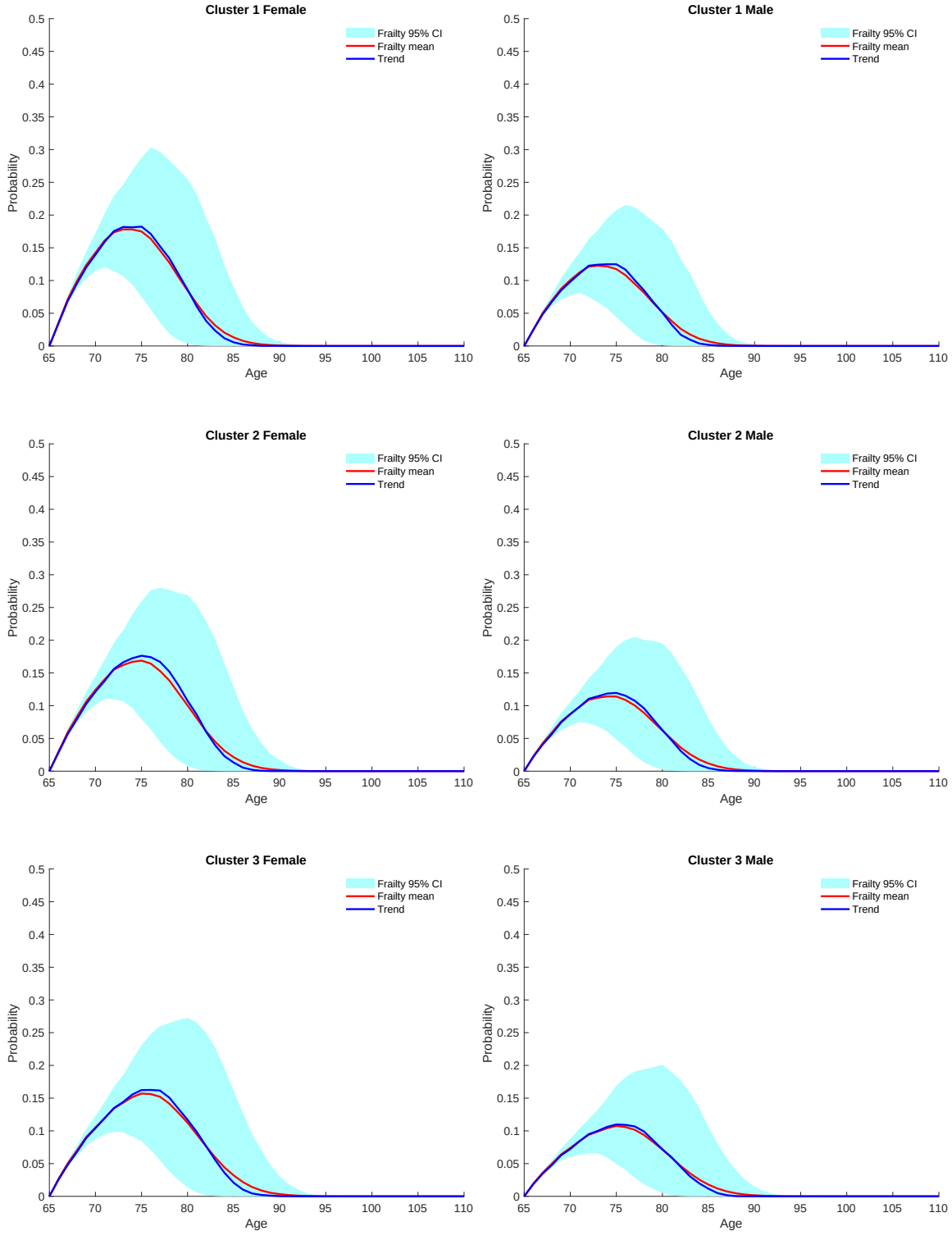


Figure 4.6: Probability of being disabled for healthy males and females aged 65 using trend and frailty models across clusters in 1998

4.5 Discussion

4.5.1 Summary of key findings

The main aim of this work is to develop better risk profiling techniques using longitudinal individual level data from other determinants of mortality and morbidity beyond age and gender. We introduce hidden Markov models with covariates as a novel way of developing mortality/morbidity risk profiles through the construction of health trajectories for individuals over their lifetime. This allows us to leverage a machine learning clustering technique with a statistically robust multistate Markov model. Secondly, we evaluate whether clustering improves the results from a traditional three state functional disability Markov model fitted using the United States Health and Retirement Study.

We find that models with clusters provide a better fit to the empirical data. However, the static model provides a poor fit and this is likely due to its inability to capture the deterioration in mortality and morbidity over time. Trend and frailty models are able to capture this effect with the frailty model with clustering outperforming all models. We also find evidence of three clusters with varying mortality and disability experience. Cluster 2 has the lowest longevity prospects while Cluster 3 has the best mortality experience. This is likely due to higher levels of generally good health and lower obesity. Individuals in Clusters 1 and 2 have higher than average obesity levels and lower rates of education at college level. Cluster 2 also has the highest levels of current smokers and people who drink alcohol, which might contribute to the high mortality levels. All these negative health behaviours were reported to increase mortality risk in Chapter 2.4. It is important to note that the combined use of hidden Markov models and k–medoids clustering allows us to delineate differences in education, wealth and income amongst clusters even though these variables were not used in estimating the HMMs.

4.5.2 Contributions

Other studies in the actuarial literature mostly focus on age effects (Fong et al., 2015), age and time interactions (Aro et al., 2014; Hanewald et al., 2019) or health status (Sherris & Wei, 2021). However, there is not much work linking health status with its driving factors such as BMI, income or self reported health. It is important to incorporate

health status as a function of its drivers because there are varying levels of mortality and functional disability risks associated with these determinants of mortality and morbidity. These differences need to be accurately priced to enable better management and design of longevity risk and critical illness products. In this chapter, we extend the literature by linking health status with BMI and self reported health.

We use hidden Markov models and k–medoids clustering to determine clusters that demonstrate varying life and healthy life expectancy. Hidden Markov models are used because they allow us to calculate a meaningful distance measure (Kullback–Leibler Divergence) between trajectories. HMMs can also handle covariates that are either categorical or continuous in nature. K–medoids clustering is a simple partitioning algorithm that is more interpretable than k–means clustering because its medoids are representative of individuals that exist in the clusters.

Additionally, we are unaware of studies that specifically focus on people who acquire at least one chronic illness throughout their lifetime. Focussing on these individuals allows us to report varying levels of mortality risk and other socioeconomic characteristics that are specific to those who acquire chronic conditions. Isolating these effects would be masked or dampened where we to study a larger population. Other three state Markov studies using both Chinese and US data do not highlight these differences (Fu et al., 2021; Z. Li et al., 2017). Although Q. Wang et al. (2022) use a combined neural network and GLM to incorporate other predictors beyond age, gender and time such as marital status and current smoking behaviour. They do not consider clustering which can show patterns amongst different predictors instead of relying on single variables.

4.5.3 Limitations and future work

While hidden Markov models are very robust, it is quite challenging to use them when the sample size is very large. Even with our relative small sample size, it takes more than 12 hours to compute the likelihood of the dataset with more than 17 million observations in the likelihood matrix. More time is also needed to calculate the distance matrix and perform the clustering. Hence, multiple solutions can be considered to deal with the high dimensionality and the increasing computational power needed to estimate HMMs such as penalised HMMs, hidden Markov neural networks and recursive multinomial approxi-

mations (Rimella, 2021; Städler & Mukherjee, 2013). Future work will consider some of these strategies.

One limitation of this study is that when we compare Chapters 3 and 4, we note that k–means clustering for longitudinal data is more adept at handling missing values than HMMs. With HMMs, one cannot compute the likelihood when there are missing values unless they choose to ignore the missing data. However, more robust HMMs that can handle missing data could be explored but were not available due to ease of implementation and time constraints at the time of completion of this study (Chassan & Concordet, 2023; Pandolfi et al., 2023).

We have only considered the impact of BMI and self reported health in this chapter. However, HMMs have the capability to incorporate more covariates whether the variables are categorical or continuous in nature. In the future we will consider other predictors and determine whether this has an effect on the clusters produced.

4.6 Conclusion

Multivariate trajectories of BMI and self–reported health estimated using hidden Markov models are able to capture mortality differences amongst different segments. The more the covariates used the better the segmentation. We find that clustering provides a better fit to empirical data than results not based on clustering. Frailty models had the best performance while static models had the worst fit. We use likelihood tests and compare AIC and BIC to confirm the differences in models. One limitation is the small sample size used in this analysis which affects the estimation of transition rates. Future work will look at pricing implications and a larger sample size.

Appendix

4.A Simulated future lifetime statistics for healthy 75–year olds using trend and frailty models

Table 4.18: Trend: Future lifetime statistics for healthy 75–year old males and females in 1998 across clusters

	Cluster 1		Cluster 2		Cluster 3		Overall[a]	
	Female	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**
Total future lifetime								
Mean	10.782	10.633	9.464	9.196	11.336	11.070	10.782	10.482
Standard error	0.042	0.041	0.039	0.038	0.043	0.041	0.042	0.041
Standard deviation	4.212	4.112	3.909	3.800	4.278	4.143	4.212	4.095
Healthy future lifetime								
Mean	7.795	8.497	7.439	7.845	8.747	9.307	8.342	8.817
Standard error	0.043	0.044	0.040	0.040	0.043	0.044	0.043	0.043
Standard deviation	4.281	4.366	3.952	3.948	4.343	4.346	4.273	4.254
Disabled future lifetime								
Mean	2.987	2.136	2.025	1.351	2.589	1.763	2.441	1.666
Standard error	0.034	0.029	0.029	0.024	0.034	0.028	0.033	0.027
Standard deviation	3.359	2.857	2.934	2.385	3.417	2.801	3.259	2.682
Healthy future lifetime over total future lifetime								
Mean	0.729	0.797	0.799	0.858	0.784	0.845	0.785	0.845
Standard error	0.003	0.002	0.003	0.002	0.003	0.002	0.003	0.002
Standard deviation	0.278	0.255	0.269	0.237	0.264	0.234	0.266	0.236
Age at onset of disability conditional on becoming disabled								
Mean	81.895	82.264	81.478	81.539	82.433	82.659	82.138	82.346
Standard error	0.048	0.054	0.051	0.058	0.055	0.062	0.054	0.060
Standard deviation	3.943	4.030	3.597	3.553	4.095	4.098	3.957	3.920

Note: Simulation results for 10,000 individuals with maximal age of 110 years.

[a] Overall represents the model with no clusters.

Table 4.19: Frailty: Future lifetime statistics for healthy 75–year old males and females in 1998 across clusters

	Cluster 1		Cluster 2		Cluster 3		Overall[a]	
	Female	**Male**	**Female**	**Male**	**Female**	**Male**	**Female**	**Male**
Total future lifetime								
Mean	10.959	10.781	9.704	9.429	11.604	11.316	10.622	10.308
Standard error	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Standard deviation	4.398	4.283	4.100	3.964	4.477	4.326	4.092	3.959
Healthy future lifetime								
Mean	7.347	8.111	7.089	7.581	8.421	9.024	8.163	8.622
Standard error	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Standard deviation	4.141	4.287	3.895	3.932	4.280	4.315	4.156	4.131
Disabled future lifetime								
Mean	3.612	2.670	2.615	1.847	3.183	2.292	2.459	1.686
Standard error	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001
Standard deviation	3.606	3.097	3.207	2.675	3.640	3.072	3.217	2.650
Healthy future lifetime over total future lifetime								
Mean	0.690	0.762	0.757	0.820	0.750	0.814	0.784	0.845
Standard error	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Standard deviation	0.280	0.261	0.277	0.250	0.267	0.241	0.266	0.236
Age at onset of disability conditional on becoming disabled								
Mean	81.242	81.665	81.035	81.260	81.835	82.174	81.835	82.036
Standard error	0.001	0.002	0.002	0.002	0.002	0.002	0.002	0.002
Standard deviation	3.693	3.840	3.466	3.510	3.915	3.980	3.817	3.811

Note: Simulation results for 10,000 individuals with maximal age of 110 years.

[a] Overall represents the model with no clusters.

Table 4.20: Comparison of future lifetime statistics across clusters for healthy 75–year old males and females using trend and frailty models

	Cluster 1[a]		Cluster 2[b]		Cluster 3[c]	
	Female	**Male**	**Female**	**Male**	**Female**	**Male**
Trend in 1998						
Total future lifetime	100.00%	101.44%	87.77%	87.73%	105.14%	105.61%
Healthy future lifetime	93.44%	96.37%	89.17%	88.98%	104.86%	105.56%
Disabled future lifetime	122.39%	128.27%	82.99%	81.14%	106.08%	105.85%
Age at onset of disability	99.70%	99.90%	99.20%	99.02%	100.36%	100.38%
Frailty in 1998						
Total future lifetime	103.17%	104.59%	91.35%	91.47%	109.24%	109.79%
Healthy future lifetime	90.00%	94.08%	86.84%	87.93%	103.16%	104.67%
Disabled future lifetime	146.87%	158.39%	106.33%	109.59%	129.44%	135.98%
Age at onset of disability	99.28%	99.55%	99.02%	99.05%	100.00%	100.17%

Note: Age at onset of disability is conditional on becoming disabled.

[a] Percentage difference in future lifetime betweeen Cluster 1 and model with no clusters of the same sex;

[b] Percentage difference in future lifetime betweeen Cluster 2 and model with no clusters of the same sex;

[c] Percentage difference in future lifetime betweeen Cluster 3 and model with no clusters of the same sex.

4.B Simulated survival curves of trend and frailty models for healthy 75–year olds

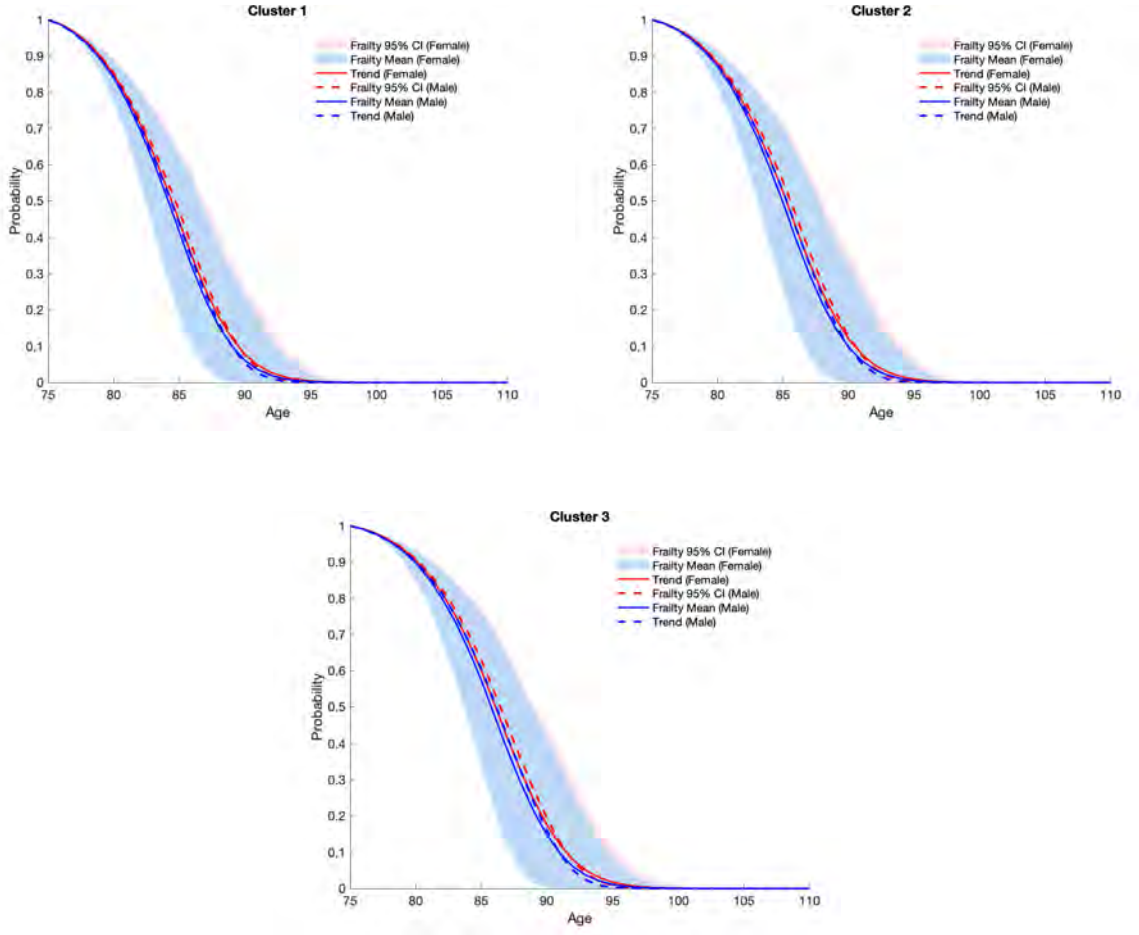


Figure 4.7: Survival curves of trend and frailty model across clusters in 1998 for healthy 75–year old males and females

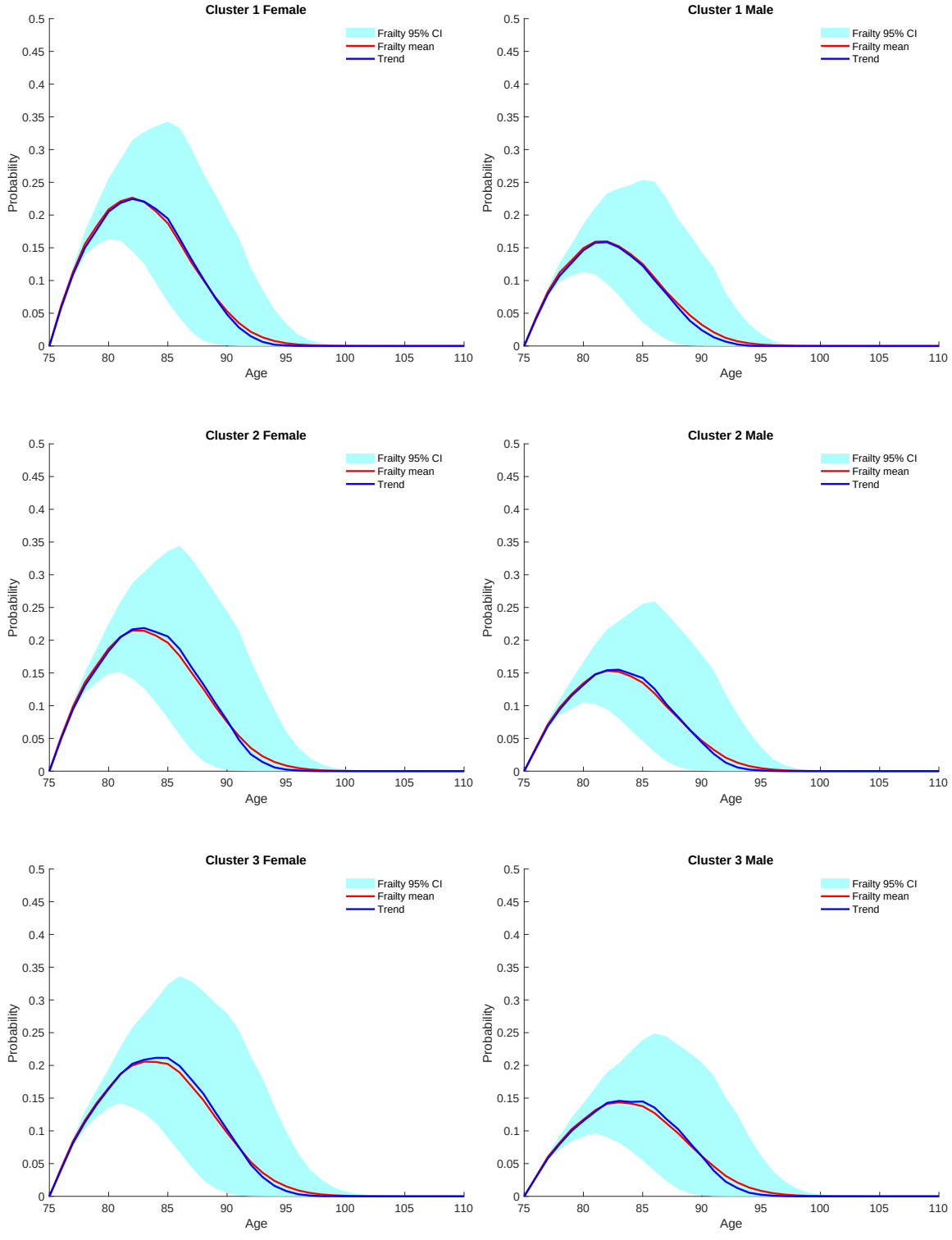


Figure 4.8: Probability of being disabled for healthy males and females aged 75 using trend and frailty models across clusters in 1998

Chapter 5

Multimorbidity and functional disability: Implications for life annuities and long–term care insurance

Earlier versions of this chapter were presented at the following conferences:

- Perspectives on Actuarial Risks in Talks of Young Researchers, Valencia, Spain. "An actuarial lens on multimorbidity and long–term care", 1 February 2023.
- 2023 International Congress of Actuaries, Sydney, Australia. "An actuarial lens on multimorbidity and long–term care", 30 May 2023.

5.1 Introduction

5.1.1 Background

Multimorbidity is commonly described as the presence of more than one chronic condition at a particular time. It is an under researched area in the actuarial literature considering the high prevalence of multiple diseases amongst older individuals. The majority of existing research on multimorbidity amongst older adults focusses on quantifying prevalence,

the incidence of morbidity and health care costs; describing the common comorbidities and their pathology; identifying risk factors and determining the impact of multimorbidity on mortality. The backdrop of an ever increasing ageing population worldwide and significant costs of chronic conditions on disease burdens globally has garnered more attention towards multimorbidity.

5.1.1.1 Definitions of multimorbidity

There are various ways to define multimorbidity. In a systematic review by Marengoni et al. (2011), the authors find that the customary way to define multimorbidity is as the occurrence of 2 or 3 diseases in any individual. However, the pitfalls of this definition are the disregard of the nature of the disease(s) and its potential impact on disability. This leads to a second definition of multimorbidity as the occurrence of diseases, cognitive impairment and/or functional disability. While this definition is exhaustive, it conflates functional disability with multimorbidity which provides challenges in an insurance setting since these two risks are different. Valderas et al. (2009) introduce a chronological aspect to defining multimorbidity, where either the duration or the sequence of morbidity is scrutinised. Adding chronology is particularly useful when dealing with longitudinal data since duration of multimorbidity or its sequence can be easily determined. Lastly, an indexation approach to multimorbidity is common amongst physicians. This entails the process where each additional disease is analysed in reference to the initial index disease an individual acquired. In a systematic review of 12 such indices, De Groot et al. (2003) find that the Charlson Index, Cumulative Illness Rating Scale (CIRS), the Index of Coexisting Disease (ICED) and the Kaplan Index are the most reliable and valid at estimating multimorbidity in randomised control trials and prognostic studies (Charlson et al., 1987; Greenfield et al., 1993; Imamura et al., 1997). It is important to note that most research articles in the epidemiological literature refrain from using an index linked definition of multimorbidity because of the lack of clinical expertise needed to select an index disease and/or the data needed to apply such a definition.

Besides the definition, another crucial aspect in defining multimorbidity is to specify the pool of conditions from which the number of diseases is ascertained. Fortin et al. (2012) find that studies that consider between 4 and 7 conditions have more variability in the

estimates of prevalence of multimorbidity. They recommend considering a group of 12 or more diseases to counteract this effect.

For actuarial purposes, it seems that defining multimorbidity as the presence of more than 2 chronic diseases with or without cognitive impairment in an individual is useful in the design of longevity and health linked insurance products. Definitions that combine chronic diseases and functional disability are not ideal for the pricing or reserving of functional disability risks accurately.

5.1.1.2 Risk factors

Generally, being older, female and having a lower socio–economic status is positively associated with multimorbidity. Van den Akker et al. (1998) find very strong evidence of an association between prevalence of multimorbidity and various risk factors. They observe this by running Chi–squared tests with multimorbidity as the response variable with separate and independent variables of sex, age, living arrangements, health insurance and education. The data are from a sample of over $60,000$ individuals from the south of the Netherlands.

Through multiple logistic regression, Van Den Akker et al. (2000) find that increasing age, low education and having more than 2 conditions at the beginning of the study are associated with differences between individuals with new multimorbidity and those with new monomorbidity. The authors use a nested–case control study of general practitioner records from Netherlands. Sex was not a significant predictor of multimorbidity.

Using the Heidelberg cohort of the European Prospective Investigation, Nagel et al. (2008) find that low education attainment is positively associated with multimorbidity in both men (Odds Ratio (OR)=1.43; 95% Confidence Interval (CI)=$[1.28, 1.61]$) and women ($OR = 1.33$; $95\% CI = [1.18, 1.57]$). They define multimorbidity as the presence of 2 or more chronic diseases. The authors also find that BMI plays a stronger intermediate role than smoking in the relationship between education and multimorbidity. Other studies also reach similar conclusions on the relationship between education and multimorbidity (Schäfer et al., 2012; Schiøtz et al., 2017). There is very strong evidence of an association between education, occupation and literacy with transitions from a healthy state to multimorbid state (Dugravot et al., 2020).

Rocca et al. (2014) find that multimorbidity varies by race/ethnicity using county data from Minnesota, USA. Amongst individuals aged 65+, multimorbidity (≥ 2 chronic conditions) is more common amongst Caucasians than Asians and Blacks. However the dataset undersamples Blacks and Asians. Using UK data Siah et al. (2022) also find that multimorbidity is positively associated with deprivation.

5.1.1.3 Projections of prevalence of multimorbidity

Kingston et al. (2018) estimate future prevalence, morbidity onset and life expectancy from 2015 to 2035 using a discrete time dynamic microsimulation model fit on English data. Multimorbidity is defined as the presence of 2 or more diseases with and without impairment. They find that prevalence rates of diabetes (+179.4%), cancer (+118.1%) and respiratory disease (+101.5%) will increase by more than 100% over the 20 year period for individuals aged 65+. More than half of the older population will have multimorbidity regardless of the definition of multimorbidity used. Interestingly, the authors note that there is an overall gain in life expectancy at age 65 from 2015 to 2035. This gain consists of a reduction in life expectancy for people with less than 2 chronic diseases and an increase in the life expectancy of individuals with 2 or more diseases. Moreover, the authors suggest that there will be morbidity expansion whereby people will spend more time multimorbid even though they are living longer (Gruenberg, 1977). Kingston et al. (2018) do not include a time trend in their modelling.

In a study of more than 1 million English people, Chan et al. (2019) find significant differences in the onset of multimorbidity and life expectancy amongst 5 quintiles based on deprivation even after controlling for age, sex and smoking status. They use a five state model of multimorbidity with the following states: Healthy, 1 Disease, 2 Diseases, More than 2 diseases and Dead. Multimorbidity is defined as the presence of more than 2 diseases from a list 30 chronic diseases that can be managed with medication but have no cure. The authors do not allow recovery from any of the illness states even though this has been shown to be significant in some studies. Individuals who have other diseases that are not chronic in nature are labeled healthy. The results show that while the most deprived spend the shortest times in a multimorbid state, they experience the highest mortality rates and their onset of multimorbidity is earlier than all other groups. These

results are insightful and extend Kingston et al. (2018) who project multimorbidity in England up to 2035.

5.1.1.4 Multimorbidity and functional disability

The literature on functional disability and morbidity is sparse. However, Koller et al. (2014) examine the impact of multimorbidity on long term care using a Cox–regression framework. They use administrative data from a German statutory health insurance company and select more than $123,000$ individuals who are at least 65 years old in 2004. Multimorbidity is defined as the presence of more than 3 chronic conditions from a list of 46 ICD–10 codes in 3 out of 4 quarters of the year 2004. Kaplan–Meier curves show that the multimorbid group had a higher probability of making a claim on a long term care insurance policy than the non–morbid group over a five year period. Even after controlling for age, sex and the interaction between age and sex, multi–morbid individuals had a 41% higher risk of needing long term care than those without multimorbidity (Hazard Ratio (HR) $= 1.41; 95\% CI = [1.36, 1.47]$). One weakness of this analysis is that morbid and non–morbid groups are determined at the beginning of the study. This means that the dynamics of people whose morbidity status changes over the 5 year period is lost. A strength of this study is that it shows that individuals in multimorbid disease clusters (based on the most common diseases) are associated with higher risks of long term care than individuals who are not multimorbid.

5.1.1.5 Multimorbidity and mortality

Out of four methods of survival analysis: Cox proportional hazards, regularised Cox models, accelerated failure time models, and a neural network survival model; Siah et al. (2022) find that the standard Cox model has similar performance to all other models based on its Concordance Index on test data (0.81) (Faraggi & Simon, 1995; Harrell et al., 1982; Simon et al., 2011; Wei, 1992; Wilson et al., 1962). The response is the five–year overall survival rate with various predictors including the most common pairs and triplets of diseases amongst the aged (65-80 years) and elderly groups (80 years +). The data are a sample of approximately $390,000$ patients from across the UK covering the period from $2010 - 2012$. Multimorbidity (having more than four chronic conditions) is the fifth

highest risk factor ($HR = 2.44; 95\%CI = [2.22, 2.69]$) after pancreas cancer, oesophageal cancer, lung cancer and liver cancer. Parkinson's disease is the lowest ranked of the top ten risk factors. All hazard ratios are corrected for multiple testing using a Benjamini–Hochberg adjustment rate with a 5% false discovery rate (Benjamini, 2010; Benjamini & Hochberg, 1995).

Dugravot et al. (2020) study whether socio–economic inequalities in mortality appear before or after frailty, disability and multimorbidity using 6,425 residents from the Whitehall II study aged 50+ on 31 August 2017. One interesting outcome is that the authors find that multimorbidity had the strongest association with mortality ($HR = 4.12; 95\%CI = [3.41, 4.98]$) compared to disability ($HR = 2.38; 95\%CI = [1.93, 2.93]$) and frailty ($HR = 1.34; 95\%CI = [1.34, 2.22]$) using separate 2–state mortality models. The impact on mortality is determined using a Cox proportional hazard model with a Weibull distribution (Carroll, 2003).

5.1.2 Literature gap

Overall, we are unaware of a link in the literature between multimorbidity and the pricing of long term care products. Sherris & Wei (2021) propose a five state health status and disability model but do not consider multimorbidity. Commonly used three state models of health and functional disability do not distinguish between multimorbid and healthy individuals who are not multimorbid (Fong et al., 2015; Hanewald et al., 2019; Z. Li et al., 2017; Shao et al., 2017; Q. Wang et al., 2022). Multistate models are used to calculate mortality rates, incidence and prevalence rates of multimorbidity for both men and women (Chan et al., 2019; Kingston et al., 2018). However, they do not consider functional disability or recovery from multimorbidity.

5.1.3 Research objectives

The overall aim of this paper is to determine how best to incorporate multimorbidity in multistate models of functional disability since these models are used in practice to price or reserve for long–term care products. In this paper we consider how to incorporate multimorbidity in the pricing of health and longevity linked products using multistate

models of functional disability with recovery. We compare two methods as seen in the literature review, one where multimorbidity is treated as a covariate in a regression model and another where multimorbidity is treated as a state. We do this by firstly extending a three state health and functional disability model by adding multimorbidity as a predictor. Secondly, we propose a five state model of multimorbidity and functional disability. Our last objective is to determine the pricing implications of the methods used to incorporate multimorbidity in annuities, long–term care and life care annuity products. Therefore, our research questions are as follows:

1. What is the impact of multimorbidity on transition rates in a three state model of health status and functional disability that controls for age and gender? Functional disability is defined as the inability to perform 2 or more activities of daily living.

2. To what extent does a five state model of multimorbidity and functional disability capture differences in mortality and functional disability risks?

3. What are the pricing and life expectancy implications of using the two different methods of capturing the effect of multimorbidity on transition rates in multiple state health models?

The remainder of the chapter is structured as follows: Section 5.2 briefly describes the dataset. Section 5.3 describes the methods used to estimate the multistate Markov models. Section 5.4 shows the results. We discuss the results in Section 5.5 and conclude in Section 5.6.

5.2 Data

The Health and Retirement Study (HRS) is a biennial longitudinal survey of Americans aged 50 and above. As is standard practice, the HRS omits individuals in institutions such as prisons or aged care facilities. However, if a member transitions from a regular household to an institution, they continue to be interviewed despite their new residency status. This means transitions to institutions and subsequent transitions are adequately captured in the data. We use the final release of the RAND HRS Longitudinal File 2016 (V2) which contains 13 waves. There are seven cohorts in the dataset and the oldest cohort was born during 1931 and 1941. The most recent cohort was born during 1960 and 1965 and represents the Late Baby Boomers. We define multimorbidity as the presence of more than one chronic conditions from the following doctor diagnosed conditions: high blood pressure, diabetes, cancer, lung disease, heart problems, stroke, psychiatric problems, and arthritis. The respondent self declares that the doctor has diagnosed the illness. However, this might be incorrect due to recall bias. While this list is not exhaustive, it is suitable for our analysis. We fit multistate models using data from waves 4 to 13 since earlier waves have inconsistencies in defining activities of daily living. Functional disability is defined as the failure to perform 2 or more activities of daily living (walking across the room, dressing, eating, getting in and out of bed, using the toilet and bathing or showering). To clean the data, we remove participants who provide either inappropriate interview status responses or do not respond in any wave. We also exclude individuals who do not appear at least twice and those who have not turned 45 by the final interview date of 31 December 2018. The final sample used for model estimation contains $31,969$ individuals and $537,622$ observations. All data cleaning procedures were performed using Stata:Release 17. The actual variables from the HRS that we use in this analysis are summarised in Table 5.1 while Figure 5.18 summarises our sample selection process.

Table 5.1: Variable description

Variable	Description
Socio-demographic	
HHIDPN	Person specific identifier in each household
RAGENDER	Sex
HACOHORT	Cohort
RABMONTH	Birth month
RABYEAR	Birth year
RABDATE	Date of birth
RADMONTH	Month of death
RADYEAR	Year of death
RADDATE	Date of death
Interview	
RxIWEND	End date of interview
RxIWENDY	End year of interview
RxIWSTAT	Interview status
Activities of daily living	
RxWALKRA	Some difficulty walking across room
RxDRESSA	Some difficulty dressing
RxEATA	Some difficulty eating
RxBEDA	Some difficulty getting in and out of bed
RxTOILTA	Some difficulty using the toilet
RxBATHA	Some difficulty bathing or showering
Doctor diagnosed health conditions	
RxHIBPE	Ever had high blood pressure
RxDIABE	Ever had diabetes
RxCANCRE	Ever had cancer
RxLUNGE	Ever had lung disease
RxHEARTE	Ever had heart problems
RxSTROKE	Ever had stroke
RxPSYCHE	Ever had psychiatric problems
RxARTHRE	Ever had arthritis

* x indicates the wave of the HRS dataset

5.3 Methodology

We use a multistate latent factor intensity model with a proportional hazard specification defined as follows: the transition intensity of type $s = 1, \ldots, S$ for an individual k for $k = 1, \ldots, K$ at time t years is modelled as

$$\lambda_{k,s}(t) = \exp(\beta_s + \gamma_s' w_k(t) + \alpha_s \psi(t)) H_{k,s}(t),$$

where β_s is the time invariant baseline log–intensity for transition type s, $w_k(t)$ is a vector of the observed predictors for each individual k, $\psi(t)$ is frailty which is a stochastic latent process, γ_s is a vector measuring the sensitivity of $\lambda_{k,s}(t)$ with respect to $w_k(t)$, α_s is a scalar measuring the sensitivity of $\lambda_{k,s}(t)$ with respect to $\psi(t)$ and the generalised baseline hazard function for duration dependence $H_{k,s}(t) = 1$ due to having a Markovian property (Koopman et al., 2008b; Z. Li et al., 2017; Sherris & Wei, 2021). We estimate the transition rates and the relevant parameters using code developed by Fu et al. (2021) which is available at Functional Disability China US. Adjustments are made to allow the static, trend and frailty models to incorporate the multimorbidity covariates. The likelihood function for the static and trend models is

$$L(\theta|F_J) = \prod_{k=1}^{K} \prod_{s=1}^{S} \prod_{j=1}^{J} \exp\Bigg\{ Y_{k,s,j} \ln \lambda_{k,s}(\hat{t}_j) - R_{k,s}(t_j) \int_{t_j}^{\hat{t}_j} \lambda_{k,s}(u) du - R_{k,s}(\hat{t}_j) \int_{t_j}^{t_{j+1}} \lambda_{k,s}(u) du \Bigg\}, \quad (5.1)$$

where θ is set of parameters to be estimated, F_J denotes all the information up to time t_J, $Y_{k,s,j}$ is an indicator function and $Y_{k,s,j} = 1$ if transition type s is observed for the k^{th} individual between the j^{th} and $(j+1)^{th}$ interviews, $R_{k,s}(t)$ is an indicator function and $R_{k,s}(t) = 1$ if the individual is exposed to the risk of transition type s at time t, $\hat{t}_j$ is the time of transition if a transition occurs.

For the frailty model, the likelihood function conditional on the complete path of the frailty is given by

$$L(\theta|F_J,\Psi)=\prod_{k=1}^{K}\prod_{s=1}^{S}\prod_{j=1}^{J}\exp\Big\{Y_{k,s,j}\ln\lambda_{k,s}\hat{(}t_j)-R_{k,s}(t_j)\int_{t_j}^{\hat{t}_j}\lambda_{k,s}(u)du- R_{k,s}(\hat{t}_j)\int_{t_j}^{t_{j+1}}\lambda_{k,s}(u)du\Big\}, \quad (5.2)$$

where F_J denotes all the information up to time t_J, $\Psi=(\Psi(t_j : j=0,1,\ldots,J))$. This means that the likelihood function of the frailty model is

$$L(\theta|F_J)=\int L(\theta|F_J,\Psi)dP(\Psi). \quad (5.3)$$

We use a Monte Carlo simulation technique to determine the maximum likelihood and define this as

$$\hat{L}(\theta|F_J)=\frac{1}{M}\sum_{m=1}^{M}L(\theta|F_J,\Psi^{(m)}), \quad (5.4)$$

where M is the number of simulated paths $\Psi^{(1)},\ldots,\Psi^{(M)}$. Further details for the recovery of the frailty process are provided by Sherris & Wei (2021) and Fu et al. (2021).

5.3.1 Extending health and functional disability state models

To determine the impact of multimorbidity on log transition rates whilst controlling for age and gender we fit a Cox regression model. Figure 5.1 shows the three state functional disability model used to fit the empirical data and the 4 types of transitions that can occur. An individual is considered healthy (H) if they have no functional disability. Functional disability (F) is defined as the inability to perform 2 or more activities of daily living. The absorbing state is the Dead (D) state. Disability rates are described by transition type $s=1$ while recovery rates are shown by transition type $s=2$. Mortality rates for the healthy and functionally disabled are shown by transition types $s=3$ and $s=4$, respectively. Note that for this model, healthy individuals are exposed to risk of transition types 1 and 3 only while functionally disabled individuals are exposed to risk of transition types 2 and 4 only.

For the static model, we assume that transition rates depend on age, gender and multimorbidity such that

$$\ln \lambda_{k,s}(t) = \beta_s + \gamma_s^{age} x_k(t) + \gamma_s^{female} F_k + \gamma_s^{multimorbidity} M_k, \tag{5.5}$$

where $\ln \lambda_{k,s}(t)$ is the log transition rate, β_s is the reference level of each transition type s, $x_k(t)$ is the age for the k^{th} individual at time t, $F_k = 1$ if the k^{th} individual is female and 0 otherwise while $M_k = 1$ if the k^{th} individual is multimorbid and 0 otherwise.

In the trend model, we add a time variable to capture sensitivity of log transition rates $\ln \lambda_{k,s}(t)$ to variations in time so that

$$\ln \lambda_{k,s}(t) = \beta_s + \gamma_s^{age} x_k(t) + \gamma_s^{female} F_k + \gamma_s^{multimorbidity} M_k + \phi_s^{time} t, \tag{5.6}$$

where t is the time trend.

The frailty model captures the effect of systematic uncertainty on the transition rates:

$$\ln \lambda_{k,s}(t) = \beta_s + \gamma_s^{age} x_k(t) + \gamma_s^{female} F_k + \gamma_s^{multimorbidity} M_k + \phi_s^{time} t + \alpha_s \psi_t, \tag{5.7}$$

where ψ_t is the frailty factor that is modelled as a random walk process with a drift term, ϵ_j and

$$\psi_t = \psi_j = \rho^{t_j - t_{j-1}} \psi_{j-1} + \epsilon_j, \quad \epsilon_j \sim N(0, \sigma_j^2), \quad \psi_0 = 0, \quad t \in (t_{j-1}, t_j), \tag{5.8}$$

where t_j is measured in years and denotes the time of the j^{th} interview, ρ is the autoregressive parameter. Similar to Fu et al. (2021), we scale the age covariate $x_k(t)$ to $\frac{x_k^{last}(t)-65}{10}$ where $x_k^{last}(t)$ is the age at last birthday and $t = t/10$. Table 5.2 shows the conversion of the calendar year to time t. We use estimated parameters from the health and functional disability models to determine the transition probabilities. Consequently, we can then derive future lifetime statistics of people with and without morbidity.

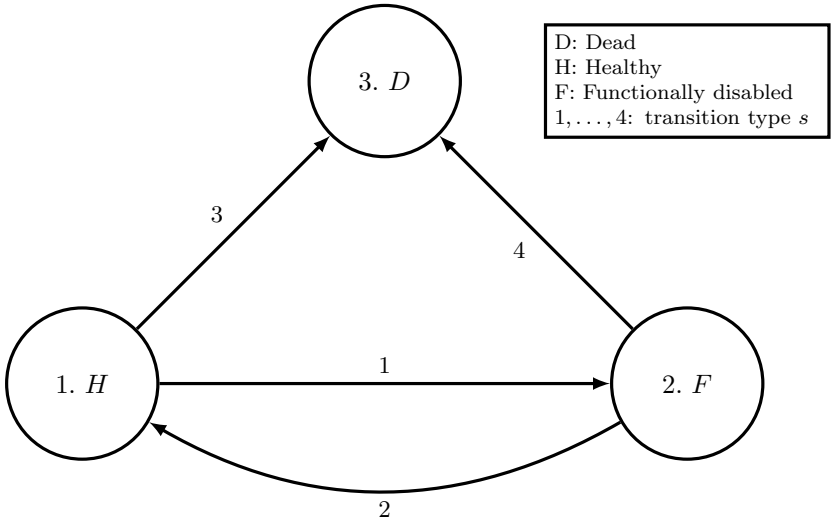


Figure 5.1: Three state health and functional disability model

Table 5.2: Conversion of calendar year to time

Year	**Time**
1998–1999	1
2000–2001	3
2002–2003	5
2004–2005	7
2006–2007	9
2008–2009	11
2010–2011	13
2012–2013	15
2014–2015	17
2016–2017	19

In a three state functional disability model, for a healthy individual aged x

- ${}^{11}_{t}p_x$ = the probability of remaining healthy at age $x + t$,
- ${}^{11}_{t-1}p_x$ = the probability of remaining healthy at age $x + t - 1$,
- ${}^{12}_{t}p_x$ = the probability of becoming functionally disabled at age $x + t$,
- ${}^{12}_{t-1}p_x$ = the probability of becoming functionally disabled at age $x + t + 1$ and
- ${}^{13}_{t}p_x$ = the probability of dying at age $x + t$.

For a functionally disabled individual aged x

- ${}^{21}_{t}p_x$ = the probability of becoming healthy at age $x+t$,
- ${}^{21}_{t-1}p_x$ = the probability of becoming healthy at age $x+t-1$,
- ${}^{22}_{t}p_x$ = the occupancy probability at age $x+t$,
- ${}^{22}_{t-1}p_x$ = the occupancy probability at age $x+t-1$, and
- ${}^{23}_{t}p_x$ = the probability of dying at age $x+t$.

Multi–year transition probabilities are defined by the following Chapman–Kolmogorov equations for one year and are valid for $t \geq 1$:

$$ {}^{11}_{t}p_x = {}^{11}_{t-1}p_x \; {}^{11}p_{x+t-1} + {}^{12}_{t-1}p_x \; {}^{21}p_{x+t-1}, \tag{5.9} $$

$$ {}^{12}_{t}p_x = {}^{11}_{t-1}p_x \; {}^{12}p_{x+t-1} + {}^{12}p_x \; {}^{22}_{t-1}p_{x+t-1}, \tag{5.10} $$

$$ {}^{21}_{t}p_x = {}^{22}_{t-1}p_x \; {}^{21}p_{x+t-1} + {}^{21}_{t-1}p_x \; {}^{11}p_{x+t-1}, \tag{5.11} $$

$$ {}^{22}_{t}p_x = {}^{22}_{t-1}p_x \; {}^{22}p_{x+t-1} + {}^{21}_{t-1}p_x \; {}^{12}p_{x+t-1}, \tag{5.12} $$

where the limiting age ω is 110, ${}^{11}_{0}p_x = 1$, ${}^{12}_{0}p_x = 0$, ${}^{22}_{0}p_x = 1$ and ${}^{21}_{0}p_x = 0$.

Table 5.3 shows the conditional probabilities related to the pricing of long term care products using three state functional disability models.

Table 5.3: Conditional probabilities related to the pricing of annuities, long term care insurance and life care annuities for three state model

State at age x	**State at x + 1**		
	H	**F**	**D**
H	${}^{11}p_x$	${}^{12}p_x$	${}^{13}p_x$
F	${}^{21}p_x$	${}^{22}p_x$	${}^{23}p_x$
D	0	0	1

We consider three different products: a life annuity, a long term care insurance and a life care annuity that combines long term care insurance and a life annuity. The life annuity pays a benefit $B = \$12,000$ annually if the individual is healthy (State 1). The life annuity continues to pay the benefit $B = \$12,000$ when the individual transitions to the

functionally disabled state from the healthy state. The long term care insurance pays a benefit $B' = \$36,000$ annually once an individual is functionally disabled (State 2). The annual effective interest rate is 3% and premiums are paid as lump sums. Benefits are unlimited and there is no waiting period.

For a life annuity, the expected present value W for a healthy individual (H) is

$$\mathbb{E}[W] = \sum_{t=1}^{n} B\, {}^{11}_{t}p_x\, v^t + \sum_{t=1}^{n} B\, {}^{12}_{t}p_x\, v^t - P_x, \tag{5.13}$$

while the expected present value Y for a long term insurance product purchased by a healthy individual (H) is

$$\mathbb{E}[Y] = \sum_{t=1}^{n} B'\, {}^{12}_{t}p_x\, v^t - P_x, \tag{5.14}$$

where $n = \omega - x$, $v = \frac{1}{1+i}$, $i = 3\%$ and P_x is the once–off premium.

For a life care annuity, the expected present value Z for a healthy individual (H) is

$$\mathbb{E}[Z] = \sum_{t=1}^{n} B\, {}^{11}_{t}p_x\, v^t + \sum_{t=1}^{n} B\, {}^{12}_{t}p_x\, v^t + \sum_{t=1}^{n} (B' - B)\, {}^{12}_{t}p_x\, v^t - P_x, \tag{5.15}$$

$$\mathbb{E}[Z] = \sum_{t=1}^{n} B\, {}^{11}_{t}p_x\, v^t + \sum_{t=1}^{n} B'\, {}^{12}_{t}p_x\, v^t - P_x,$$

where $n = \omega - x$, $v = \frac{1}{1+i}$, $i = 3\%$ and P_x is the once–off premium at age x. Note that changes in time for an individual aged x are captured in the multiyear transition probabilities.

5.3.2 Proposed five state model of multimorbidity and functional disability

Static, trend and frailty models for the proposed five state model of multimorbidity and functional disability are defined by the following equations using previously defined notation:

$$\ln \lambda_{k,s}(t) = \beta_s + \gamma_s^{age} x_k(t) + \gamma_s^{female} F_k, \tag{5.16}$$

$$\ln \lambda_{k,s}(t) = \beta_s + \gamma_s^{age} x_k(t) + \gamma_s^{female} F_k + \phi_s^{time} t, \tag{5.17}$$

$$\ln \lambda_{k,s}(t) = \beta_s + \gamma_s^{age} x_k(t) + \gamma_s^{female} F_k + \phi_s^{time} t + \alpha_s \psi_t. \tag{5.18}$$

Figure 5.2 shows the five state multimorbidity and functional disability model used to fit the empirical data along with the 12 possible transitions amongst the states. Functional disability is triggered when an individual has any of two activities of living specified in Table 5.1. Similar to the three state model we define multimorbidity as the presence of any one of the 8 doctor diagnosed conditions in the HRS data.

Note that for the five state model, the state space is $s = \{1, \ldots, 12\}$ which means that

1. Individuals who are healthy (H), that is, not multimorbid and not functionally disabled are exposed to risk of transition types $1, 2, 3$ and 4;

2. Individuals who are not multimorbid and functionally disabled (F) are exposed to risk of transition types $7, 8, 9$ and 10;

3. Individuals who are multimorbid and not functionally disabled (M) are exposed to risk of transition types 5 and 6; and

4. Individuals who are multimorbid and functionally disabled (MF) are exposed to risk of transition types 11 and 12.

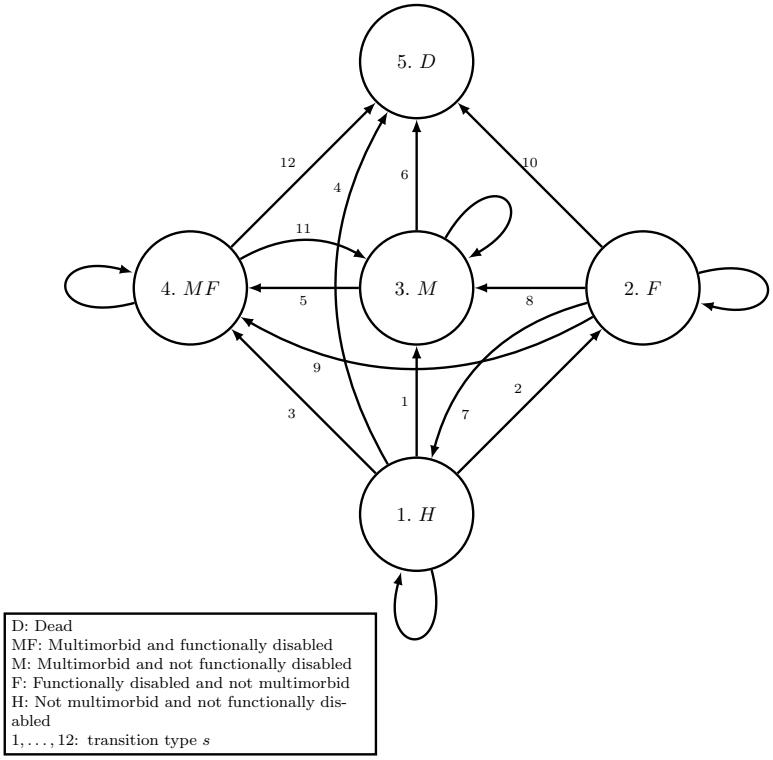


Figure 5.2: Five state multimorbidity and functional disability model

For a healthy individual aged x

- ${}^{11}_{t}p_x$ = the probability of remaining healthy (neither multimorbid nor functionally disabled) at age $x + t$,
- ${}^{11}_{t-1}p_x$ = the probability of remaining healthy at age $x + t - 1$,
- ${}^{12}_{t}p_x$ = the probability of becoming functionally disabled and not multimorbid at age $x + t$,
- ${}^{12}_{t-1}p_x$ = the probability of becoming functionally disabled and not multimorbid at age $x + t - 1$,
- ${}^{13}_{t}p_x$ = the probability of becoming multimorbid and not functionally disabled at age $x + t$,
- ${}^{13}_{t-1}p_x$ = the probability of becoming multimorbid and not functionally disabled at age $x + t - 1$,
- ${}^{14}_{t}p_x$ = the probability of becoming multimorbid and functionally disabled at age $x + t$, and
- ${}^{14}_{t-1}p_x$ = the probability of becoming multimorbid and functionally disabled at age $x + t - 1$,.

For a functionally disabled and not multimorbid individual aged x

- ${}^{21}_{t}p_x$ = the probability of becoming healthy at age $x+t$,
- ${}^{21}_{t-1}p_x$ = the probability of becoming healthy at age $x+t-1$,
- ${}^{22}_{t}p_x$ = the occupancy probability at age $x+t$,
- ${}^{22}_{t-1}p_x$ = the occupancy probability at age $x+t-1$,
- ${}^{23}_{t}p_x$ = the probability of becoming multimorbid and not functionally disabled at age $x+t$,
- ${}^{23}_{t-1}p_x$ = the probability of becoming multimorbid and not functionally disabled at age $x+t-1$,
- ${}^{24}_{t}p_x$ = the probability of becoming multimorbid and functionally disabled at age $x+t$, and
- ${}^{24}_{t-1}p_x$ = the probability of becoming multimorbid and functionally disabled at age $x+t-1$.

For a multimorbid and not functionally disabled individual aged x

- ${}^{33}_{t}p_x$ = the occupancy probability at age $x+t$,
- ${}^{33}_{t-1}p_x$ = the occupancy probability at age $x+t-1$,

- ${}^{34}_{t}p_x$ = the probability of becoming multimorbid and functionally disabled at age $x+t$ and
- ${}^{34}_{t-1}p_x$ = the probability of becoming multimorbid and functionally disabled at age $x+t-1$

For a multimorbid and functionally disabled individual aged x

- ${}^{43}_{t}p_x$ = the probability of becoming multimorbid and not functionally disabled at age $x+t$,
- ${}^{43}_{t-1}p_x$ = the probability of becoming multimorbid and not functionally disabled at age $x+t-1$,
- ${}^{44}_{t}p_x$ = the occupancy probability at age $x+t$ and
- ${}^{44}_{t-1}p_x$ = the occupancy probability at age $x+t-1$.

Multi–year transition probabilities are defined by the following Chapman–Kolmogorov equations for one year and are valid for $t \geq 1$:

$$
\begin{aligned}
{}^{11}_{t}p_x &= {}^{11}_{t-1}p_x \; {}^{11}p_{x+t-1} + {}^{12}_{t-1}p_x \; {}^{21}p_{x+t-1}, && (5.19)\\
{}^{12}_{t}p_x &= {}^{11}_{t-1}p_x \; {}^{12}p_{x+t-1} + {}^{12}p_x \; {}^{22}_{t-1}p_{x+t-1}, && (5.20)\\
{}^{13}_{t}p_x &= {}^{11}_{t-1}p_x \; {}^{13}p_{x+t-1} + {}^{13}_{t-1}p_x \; {}^{33}p_{x+t-1} + {}^{12}_{t-1}p_x \; {}^{23}p_{x+t-1} + {}^{14}_{t-1}p_x \; {}^{43}p_{x+t-1}, && (5.21)\\
{}^{14}_{t}p_x &= {}^{11}_{t-1}p_x \; {}^{14}p_{x+t-1} + {}^{14}_{t-1}p_x \; {}^{44}p_{x+t-1} + {}^{12}_{t-1}p_x \; {}^{24}p_{x+t-1} + {}^{13}_{t-1}p_x \; {}^{34}p_{x+t-1}, && (5.22)\\
{}^{21}_{t}p_x &= {}^{22}_{t-1}p_x \; {}^{21}p_{x+t-1} + {}^{21}_{t-1}p_x \; {}^{11}p_{x+t-1}, && (5.23)\\
{}^{22}_{t}p_x &= {}^{22}_{t-1}p_x \; {}^{22}p_{x+t-1} + {}^{21}_{t-1}p_x \; {}^{12}p_{x+t-1}, && (5.24)\\
{}^{23}_{t}p_x &= {}^{22}_{t-1}p_x \; {}^{23}p_{x+t-1} + {}^{23}_{t-1}p_x \; {}^{33}p_{x+t-1} + {}^{24}_{t-1}p_x \; {}^{43}p_{x+t-1} + {}^{21}_{t-1}p_x \; {}^{13}p_{x+t-1}, && (5.25)\\
{}^{24}_{t}p_x &= {}^{21}_{t-1}p_x \; {}^{14}p_{x+t-1} + {}^{22}_{t-1}p_x \; {}^{24}p_{x+t-1} + {}^{24}_{t-1}p_x \; {}^{44}p_{x+t-1} + {}^{23}_{t-1}p_x \; {}^{34}p_{x+t-1}, && (5.26)\\
{}^{33}_{t}p_x &= {}^{33}_{t-1}p_x \; {}^{33}p_{x+t-1} + {}^{34}_{t-1}p_x \; {}^{43}p_{x+t-1}, && (5.27)\\
{}^{34}_{t}p_x &= {}^{33}_{t-1}p_x \; {}^{34}p_{x+t-1} + {}^{34}_{t-1}p_x \; {}^{44}p_{x+t-1} && (5.28)\\
{}^{43}_{t}p_x &= {}^{44}_{t-1}p_x \; {}^{43}p_{x+t-1} + {}^{43}_{t-1}p_x \; {}^{33}p_{x+t-1}, && (5.29)\\
{}^{44}_{t}p_x &= {}^{44}_{t-1}p_x \; {}^{44}p_{x+t-1} + {}^{43}_{t-1}p_x \; {}^{34}p_{x+t-1}, && (5.30)
\end{aligned}
$$

where the limiting age ω is 110, ${}^{11}_{0}p_x = 1$, ${}^{12}_{0}p_x = 0$, ${}^{13}_{0}p_x = 0$, ${}^{14}_{0}p_x = 0$, ${}^{21}_{0}p_x = 0$, ${}^{22}_{0}p_x = 1$, ${}^{23}_{0}p_x = 0$, ${}^{24}_{0}p_x = 0$, ${}^{33}_{0}p_x = 1$, ${}^{34}_{0}p_x = 0$, ${}^{43}_{0}p_x = 0$ and ${}^{44}_{0}p_x = 1$.

Table 5.4 shows the conditional probabilities related to the pricing of long term care products using the five state multimorbidity and functional disability models.

Table 5.4: Conditional probabilities related to the pricing of annuities, long term care insurance and life care annuities for five state model

State at age x	**State at age x + 1**				
	H	**F**	**M**	**MF**	**D**
H	${}^{11}p_x$	${}^{12}p_x$	${}^{13}p_x$	${}^{14}p_x$	${}^{15}p_x$
F	${}^{21}p_x$	${}^{22}p_x$	${}^{23}p_x$	${}^{24}p_x$	${}^{25}p_x$
M	0	0	${}^{33}p_x$	${}^{34}p_x$	${}^{35}p_x$
MF	0	0	${}^{43}p_x$	${}^{44}p_x$	${}^{45}p_x$
D	0	0	0	0	1

As in the previous case, we consider three different products: a life annuity, a long term care insurance and a life care annuity that combines long term care insurance and a life annuity. The long term care insurance pays $B' = \$36,000$ per annum once an individual is functionally disabled (State F) while the life annuity pays $B = \$12,000$ per annum if the individual is alive. An individual who is both multimorbid and functionally disabled receives a benefit $B'' = \$48,000$. The annual effective interest rate is 3% and premiums are paid as lump sums. Benefits are unlimited and there is no waiting period.

For a life annuity, the expected present value W for a healthy individual in State 1 at age x is

$$ {}^{1}W_x = \sum_{t=1}^{n} B\, {}^{11}_{t}p_x\, v^t + \sum_{t=1}^{n} B\, {}^{12}_{t}p_x\, v^t + \sum_{t=1}^{n} B\, {}^{13}_{t}p_x\, v^t + \sum_{t=1}^{n} B\, {}^{14}_{t}p_x\, v^t - {}^{1}P_x, \tag{5.31} $$

and the expected present value W for a multimorbid and not functionally disabled individual in State 3 at age x is

$$ {}^{3}W_x = \sum_{t=1}^{n} B\, {}^{33}_{t}p_x\, v^t + \sum_{t=1}^{n} B\, {}^{34}_{t}p_x\, v^t - {}^{3}P_x, \tag{5.32} $$

where $n = \omega - x$, $v^t = \frac{1}{1+i}^{\,t}$ is the discount factor at time t, $i = 3\%$ is the annual effective interest rate, ${}^{1}P_x$ is the once–off premium in State 1 at age x and ${}^{3}P_x$ is the once–off premium in State 3 at age x.

For a stand alone long term care insurance product, the expected present value Y for a healthy individual in State 1 at age x is

$$ {}^{1}Y_x = \sum_{t=1}^{n} B'\, {}^{12}_{t}p_x\, v^t + \sum_{t=1}^{n} B''\, {}^{14}_{t}p_x\, v^t - {}^{1}P_x \tag{5.33} $$

and the expected present value Y for a multimorbid and not functionally disabled individual in State 3 at age x is

$$ {}^{3}Y_x = \sum_{t=1}^{n} B''\, {}^{34}_{t}p_x\, v^t - {}^{3}P_x, \tag{5.34} $$

where $n = \omega - x$, $v^t = \frac{1}{1+i}^{\,t}$ is the discount factor at time t, $i = 3\%$ is the annual effective interest rate, ${}^{1}P_x$ is the once–off premium in State 1 at age x and ${}^{3}P_x$ is the once–off

premium in State 3 at age x.

For a life care annuity, the expected present value Z for a healthy individual in State 1 at age x is

$$
\begin{aligned}
{}^{1}Z_x = \sum_{t=1}^{n} B\, {}^{11}_{t}p_x\, v^t + \sum_{t=1}^{n} B\, {}^{12}_{t}p_x\, v^t + \sum_{t=1}^{n}(B' - B){}^{12}_{t}p_x\, v^t + \sum_{t=1}^{n} B\, {}^{13}_{t}p_x\, v^t + \sum_{t=1}^{n} B\, {}^{14}_{t}p_x\, v^t \\
+ \sum_{t=1}^{n}(B'' - B){}^{14}_{t}p_x\, v^t - {}^{1}P_x, \quad (5.35)
\end{aligned}
$$

$$
{}^{1}Z_x = \sum_{t=1}^{n} B\, {}^{11}_{t}p_x\, v^t + \sum_{t=1}^{n} B'\, {}^{12}_{t}p_x\, v^t + \sum_{t=1}^{n} B\, {}^{13}_{t}p_x\, v^t + \sum_{t=1}^{n} B''\, {}^{14}_{t}p_x\, v^t - {}^{1}P_x,
$$

the expected present value Z for a multimorbid and not functionally disabled individual in State 3 at age x is

$$
{}^{3}Z_x = \sum_{t=1}^{n} B\, {}^{33}_{t}p_x\, v^t + \sum_{t=1}^{n} B\, {}^{34}_{t}p_x\, v^t + \sum_{t=1}^{n}(B'' - B){}^{34}_{t}p_x\, v^t - {}^{3}P_x, \quad (5.36)
$$

$$
{}^{3}Z_x = \sum_{t=1}^{n} B\, {}^{33}_{t}p_x\, v^t + \sum_{t=1}^{n} B''\, {}^{34}_{t}p_x\, v^t - {}^{3}P_x,
$$

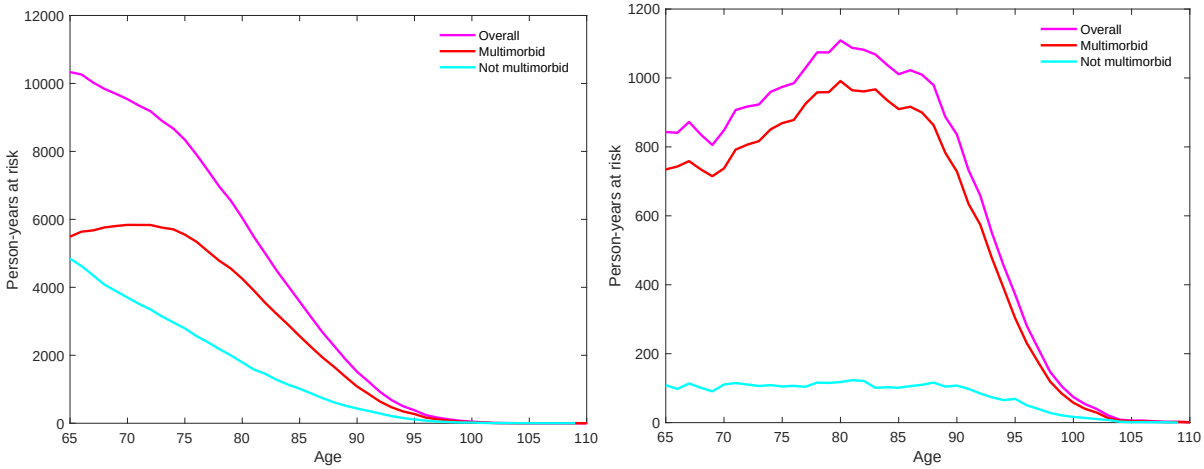
where $n = \omega - x$, $v^t = \frac{1}{1+i}^t$ is the discount factor at time t, $i = 3\%$ is the annual effective interest rate, ${}^{1}P_x$ is the once–off premium in State 1 at age x and ${}^{3}P_x$ is the once–off premium in State 3 at age x.

5.4 Results

5.4.1 Exploratory data analysis for three state health and functional disability model

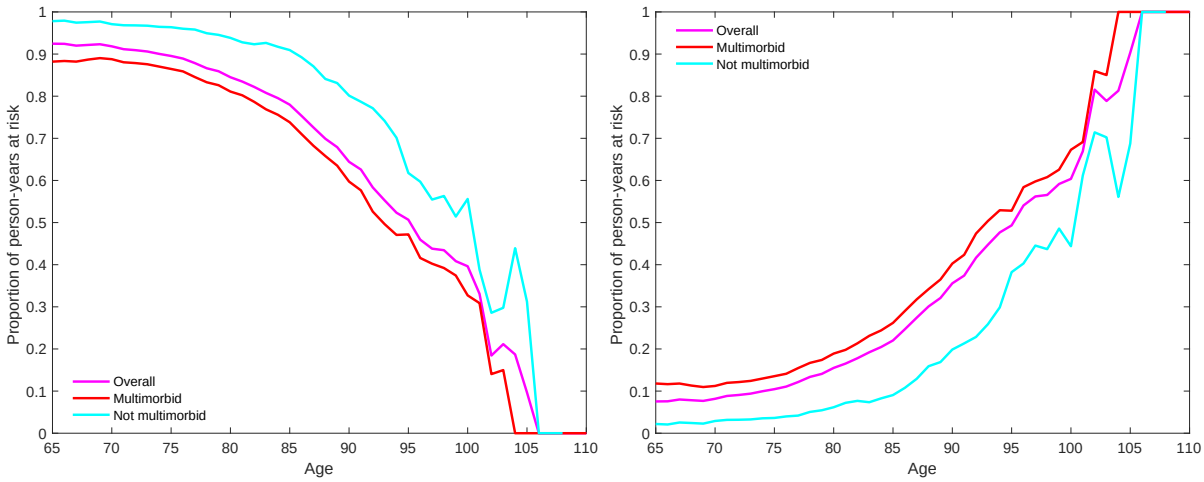
Figure 5.3 shows the exposed to risk at different ages for the overall dataset, multimorbid and non–multimorbid groups. All curves for person years at risk are downward sloping in the healthy state. However, the total exposed to risk for individuals who are not multimorbid is much lower than that of the multimorbid groups particularly in the disabled state. Figure 5.4 shows the proportions of person years at risk for different ages using

the three state model. Individuals who are not multimorbid have a higher proportion of years spent in the healthy state and a lower proportion spent in the disabled state.



(a) Healthy state **(b)** Functionally disabled state

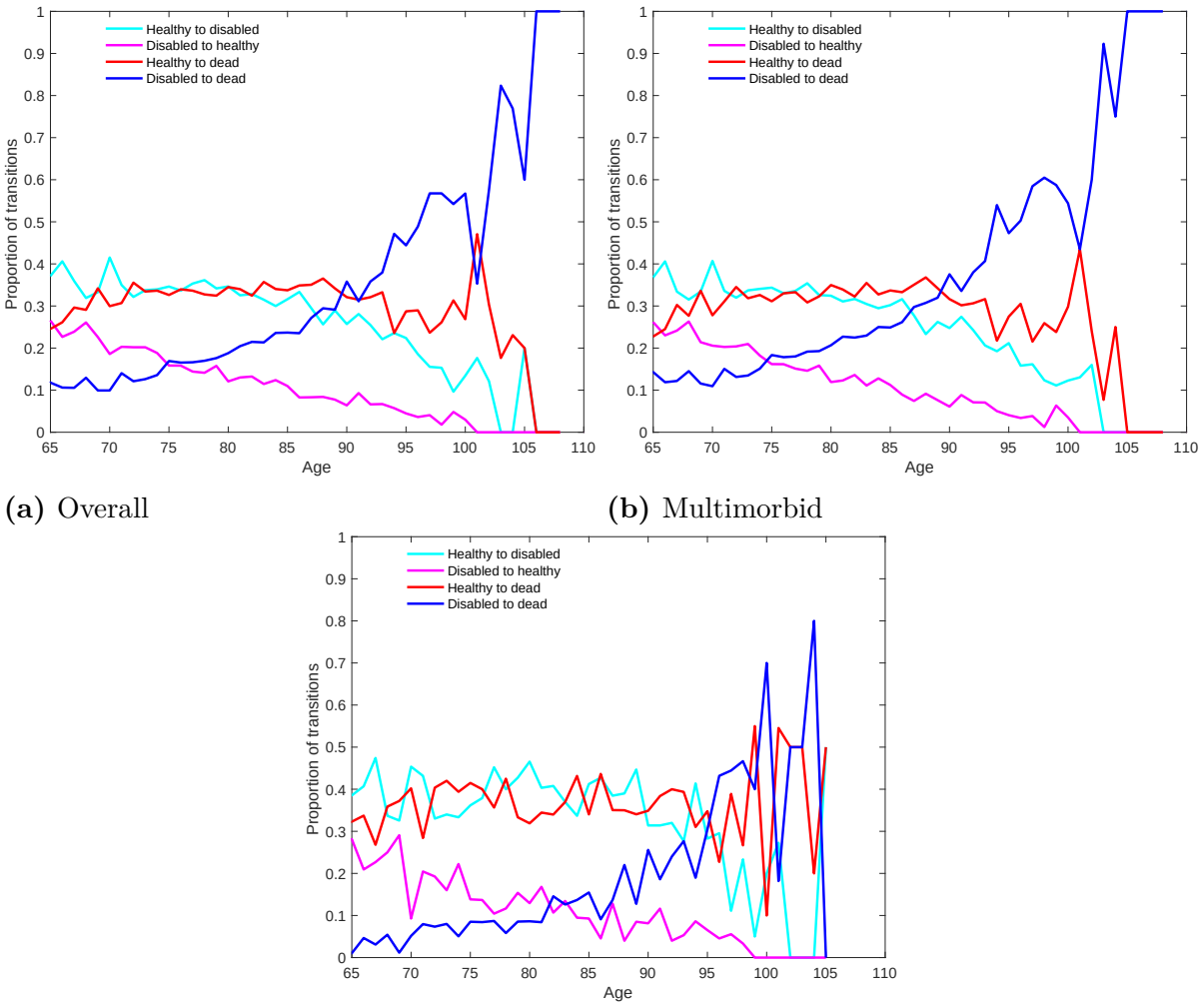
Figure 5.3: Total exposed to risk at different ages for three state model



(a) Healthy state **(b)** Functionally disabled state

Figure 5.4: Proportions of person years at risk for different ages using three state model

Figure 5.5 shows the proportions of transitions between states for the multimorbid sample, the not multimorbid sample and the overall sample. Transitions between states for the overall dataset and multimorbimorbid dataset are almost mirror images of each other. However, whilst there is a jump in the proportions of transition counts in older ages from either the healthy state or functionally diabled state in the not multimorbid group, we observe a more dramatic shift in the overall and multimorbid groups.



(a) Overall **(b)** Multimorbid

(c) Not multimorbid

Figure 5.5: Proportions of transitions at different ages using three state model

Figure 5.6 and Figure 5.7 show the crude health transition rates for males and females for the three state model. The curves for the multimorbid and not multimorbid groups appear to cross over at older ages which suggests that we might need an interaction term between age and multimorbidity in the 3 state model. This effect is magnified in females more than males. Once we fit the model as specified by Equation (5.5) we notice in Figure 5.8 that while the model is reasonable for most groups it tends to overestimate mortality and disability rates for females who are not multimorbid. This implies that we have to remain aware of this discrepancy when interpreting life expectancy statistic and pricing for females. By graphical inspection, we note that the proportional hazard assumption is also valid when we plot crude transitions against multimorbidity as shown in Figure 5.9.

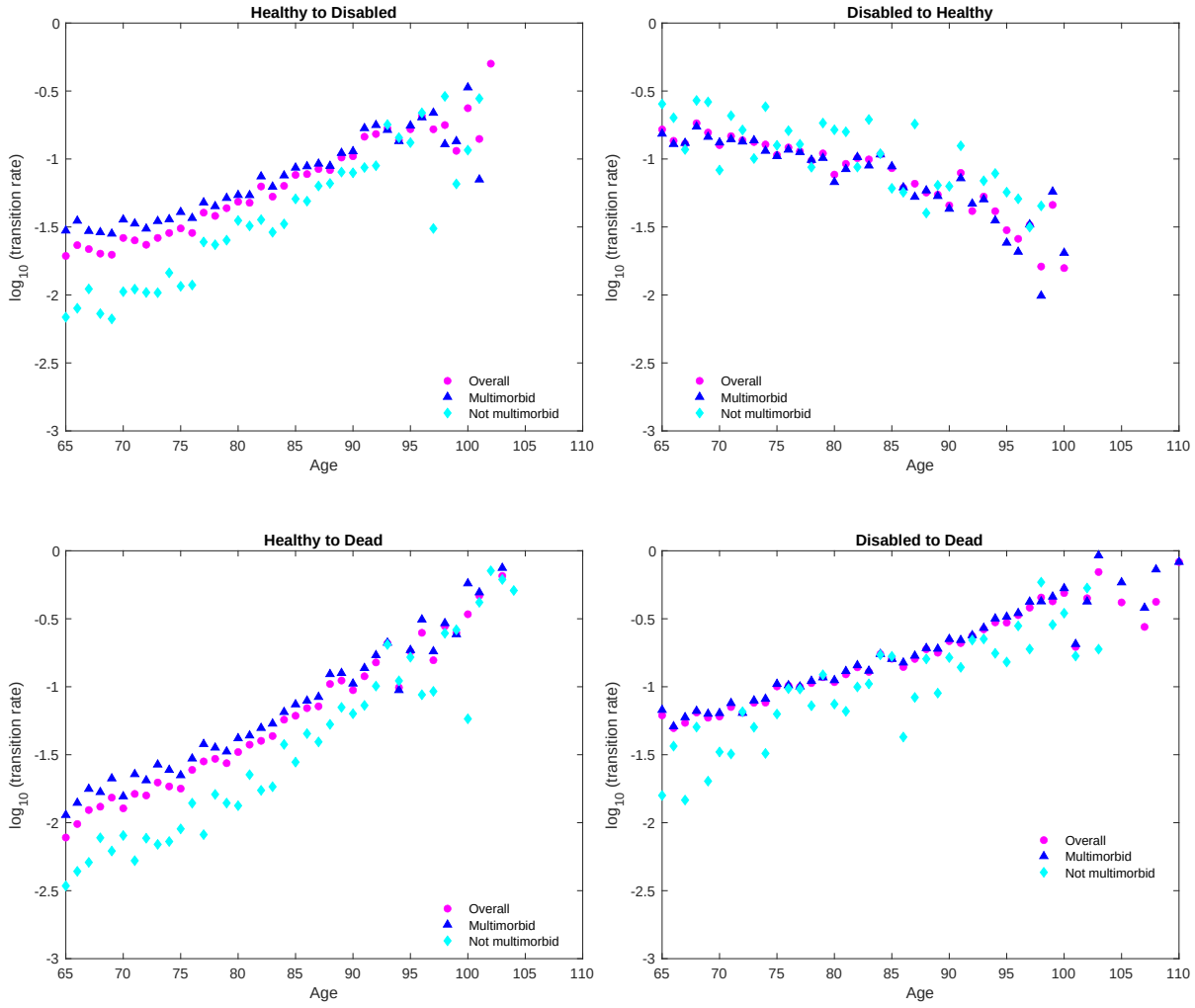


Figure 5.6: Crude transition rates at different ages for females using three state model

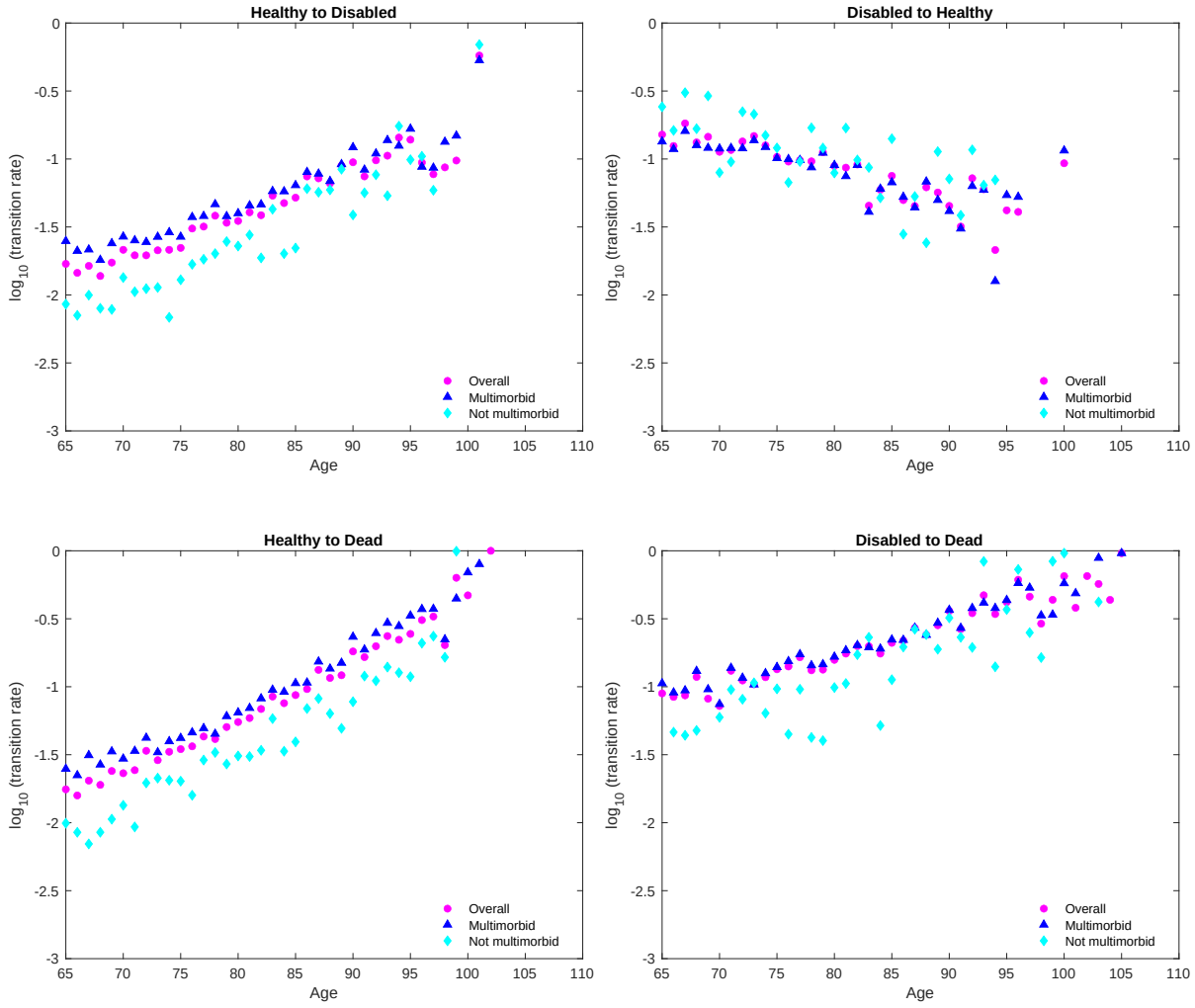


Figure 5.7: Crude transition rates at different ages for males using three state model

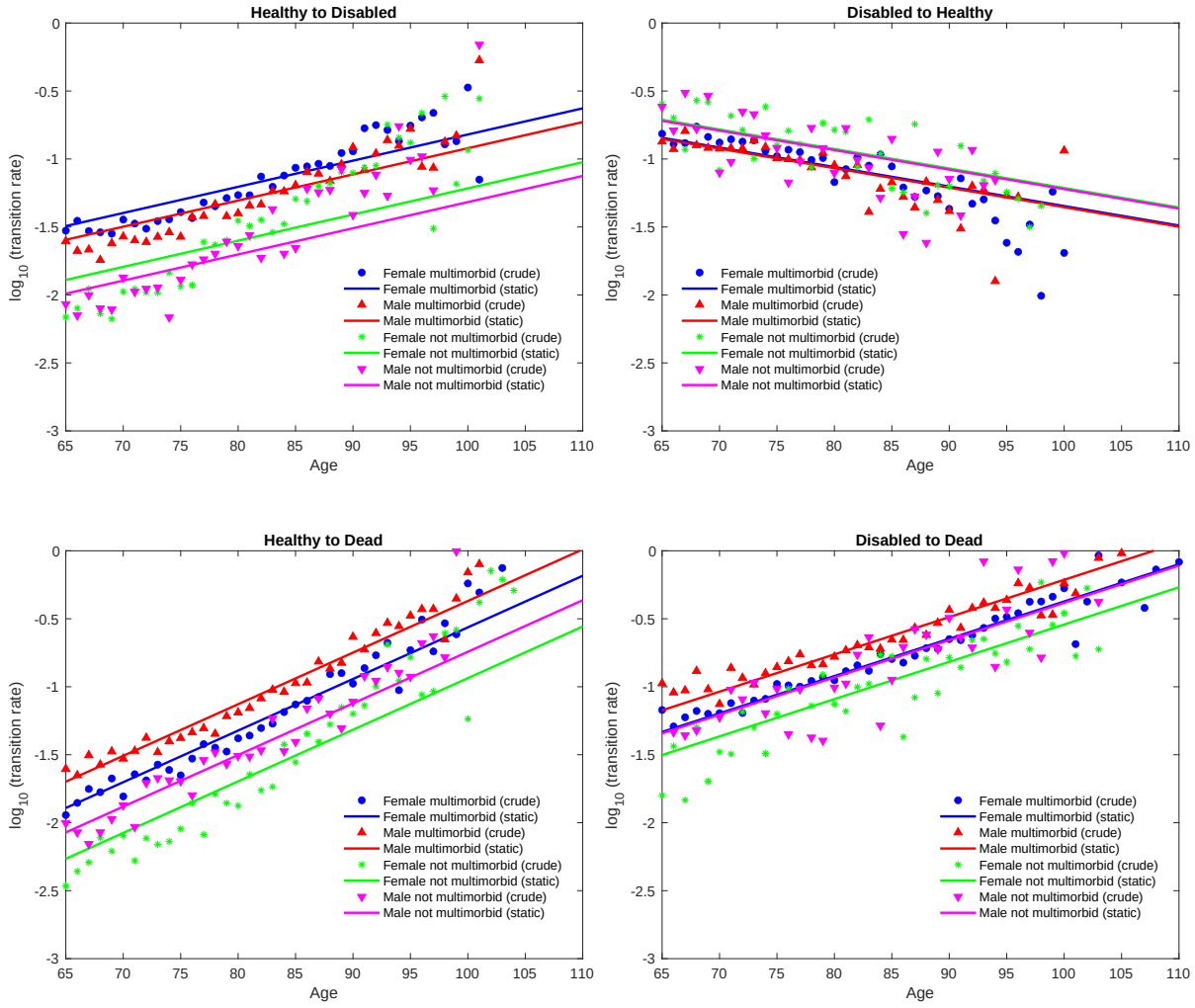


Figure 5.8: Comparison of crude transition rates and estimated static model for three state with multimorbidity predictor

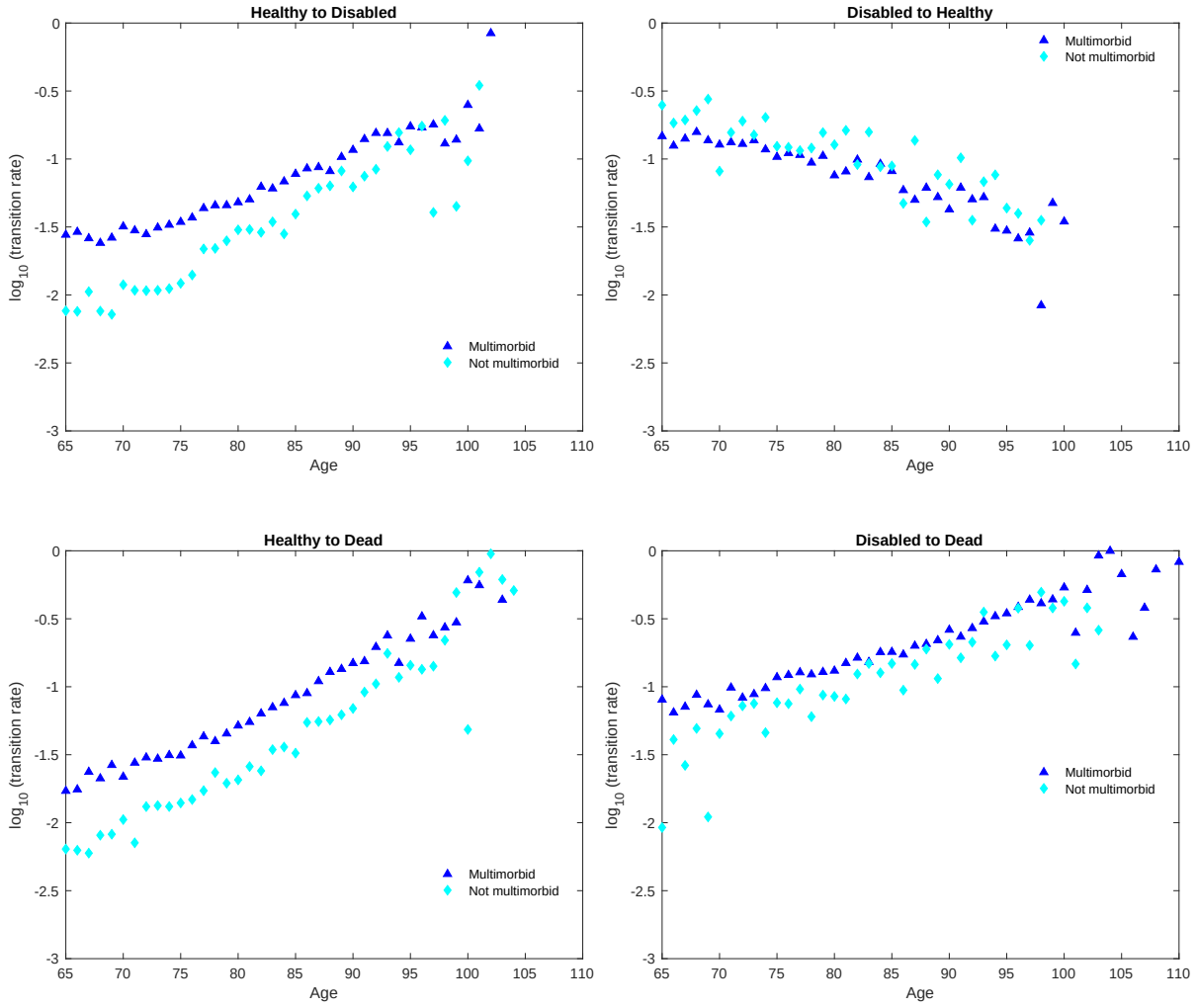
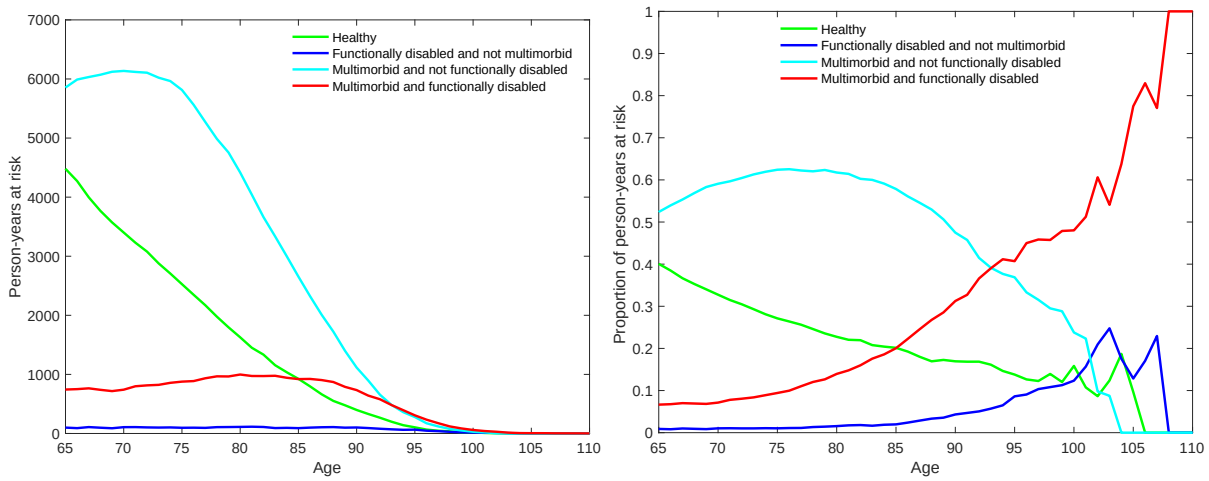


Figure 5.9: Crude transition rates versus multimorbidity using three state model

5.4.2 Exploratory data analysis for five state multimorbidity and functional disability model

Figure 5.10a shows the total exposed to risk for all states in the 5 state model except for the dead state. The functionally disabled and not multimorbid state has the lowest exposure for all ages. The curves for the healthy state and the functionally disabled and not multimorbid state are downward sloping. The person years at risk for the multimorbid and functionally disabled state are stable for ages 65 to 90 and then decrease steadily in later years. Figure 5.10b shows the proportion of person years at risk for states 1 to 4 in the 5 state model. The curves for states 2 and 4 are upward sloping while the opposite is true for states 1 and 3.



(a) Total exposed to risk **(b)** Proportion of person years at risk

Figure 5.10: Total exposed to risk and proportion of person years at risk at different ages for 5 state model

Figure 5.11 and Figure 5.12 show the crude health transition rates for males and females for the five state model. There are fewer transitions at older ages and this effect is more pronounced for the individuals who are functionally disabled and not multimorbid (State 2). By graphical inspection, we note that log transition rates have a linear trend with age for both males and females for most transition types. This allows us to estimate the models as previously specified. Recovery from state 2 to either the healthy state or multimorbidity state becomes less likely as age increases. Mortality rates increase with age for all 4 states for males and females. Healthy females have a higher chance of becoming functionally disabled than healthy males. Healthy males have a slighly higher chance of becoming multimorbid and not functionally disabled than healthy females. Figure 5.13 and Figure 5.14 show the crude health transition rates overlaid by the static model which verifies that our model is reasonable and valid.

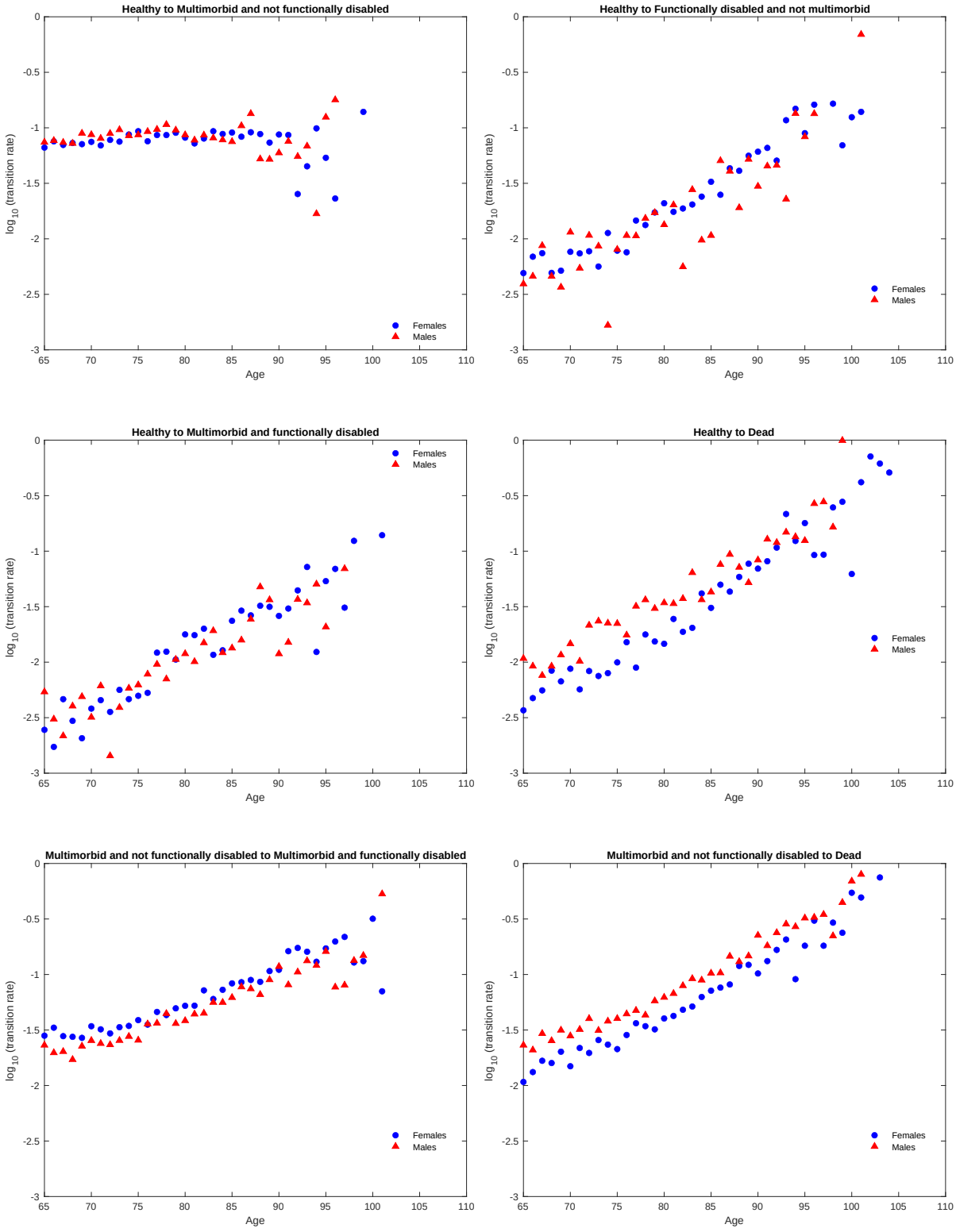


Figure 5.11: Crude transition rates at different ages for males and females using 5 state model (s=1 to s=6)

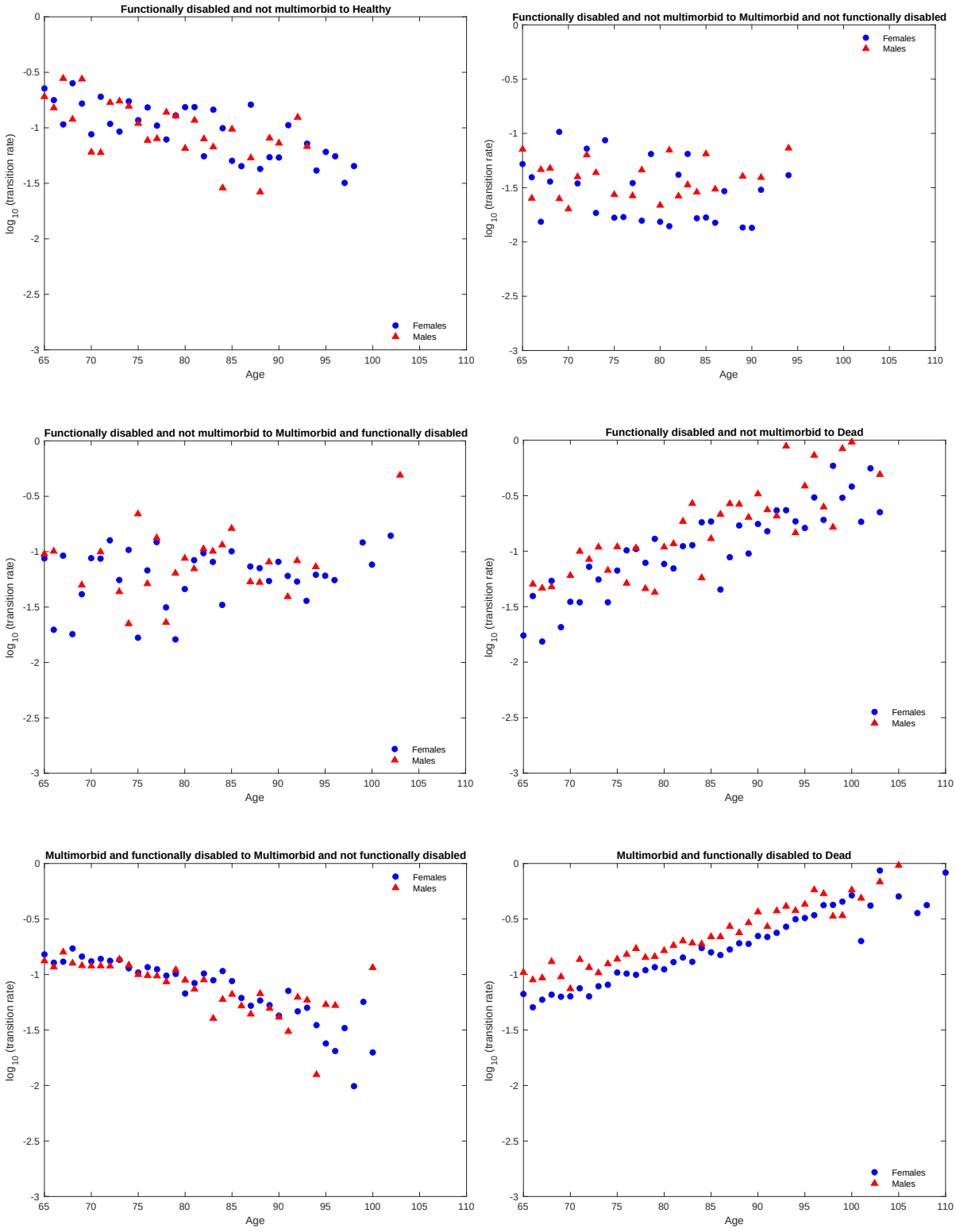


Figure 5.12: Crude transition rates at different ages for males and females using 5 state model (s=7 to s=12)

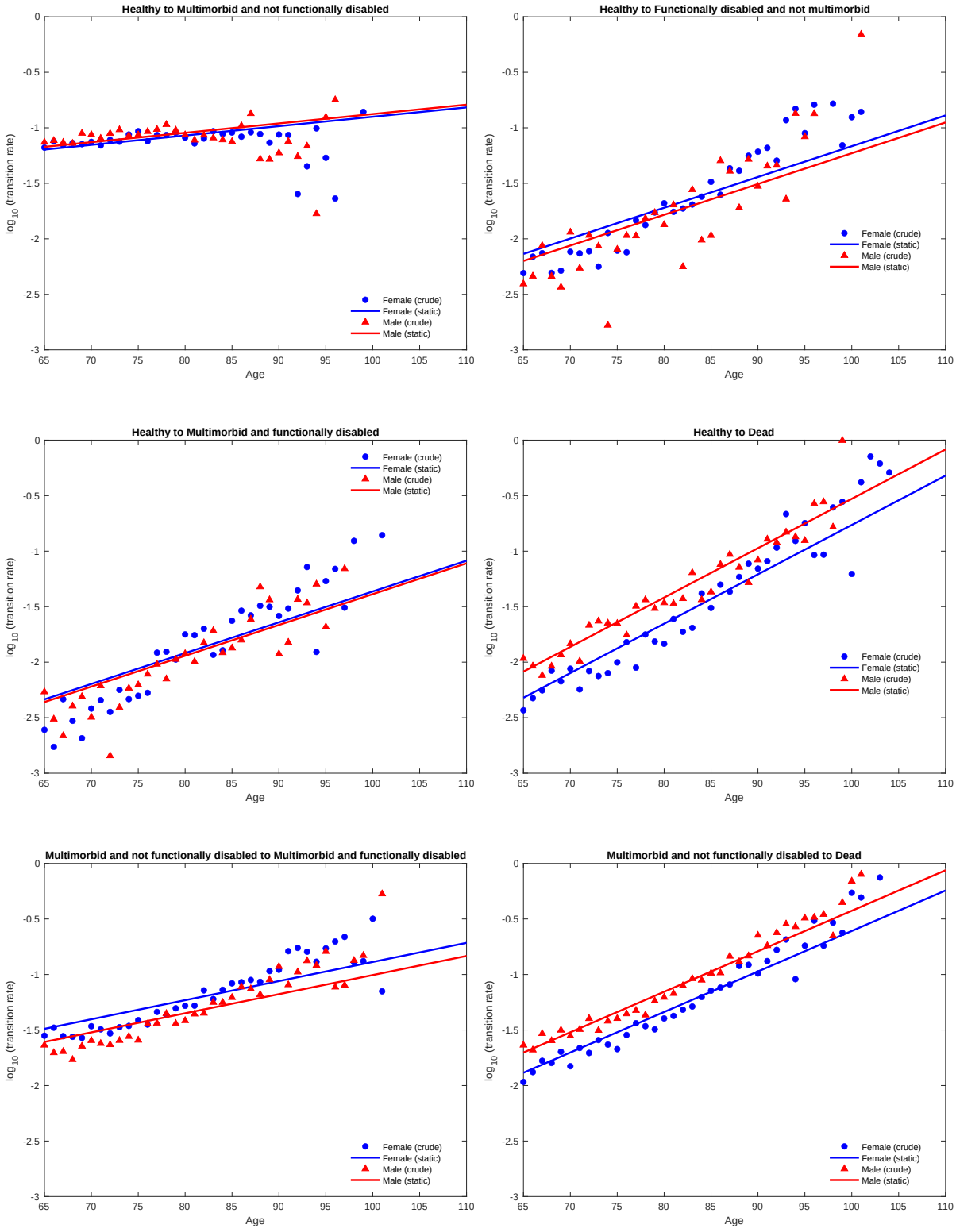


Figure 5.13: Crude transition rates at different ages for males and females using 5 state model (s=1 to s=6) overlaid by static

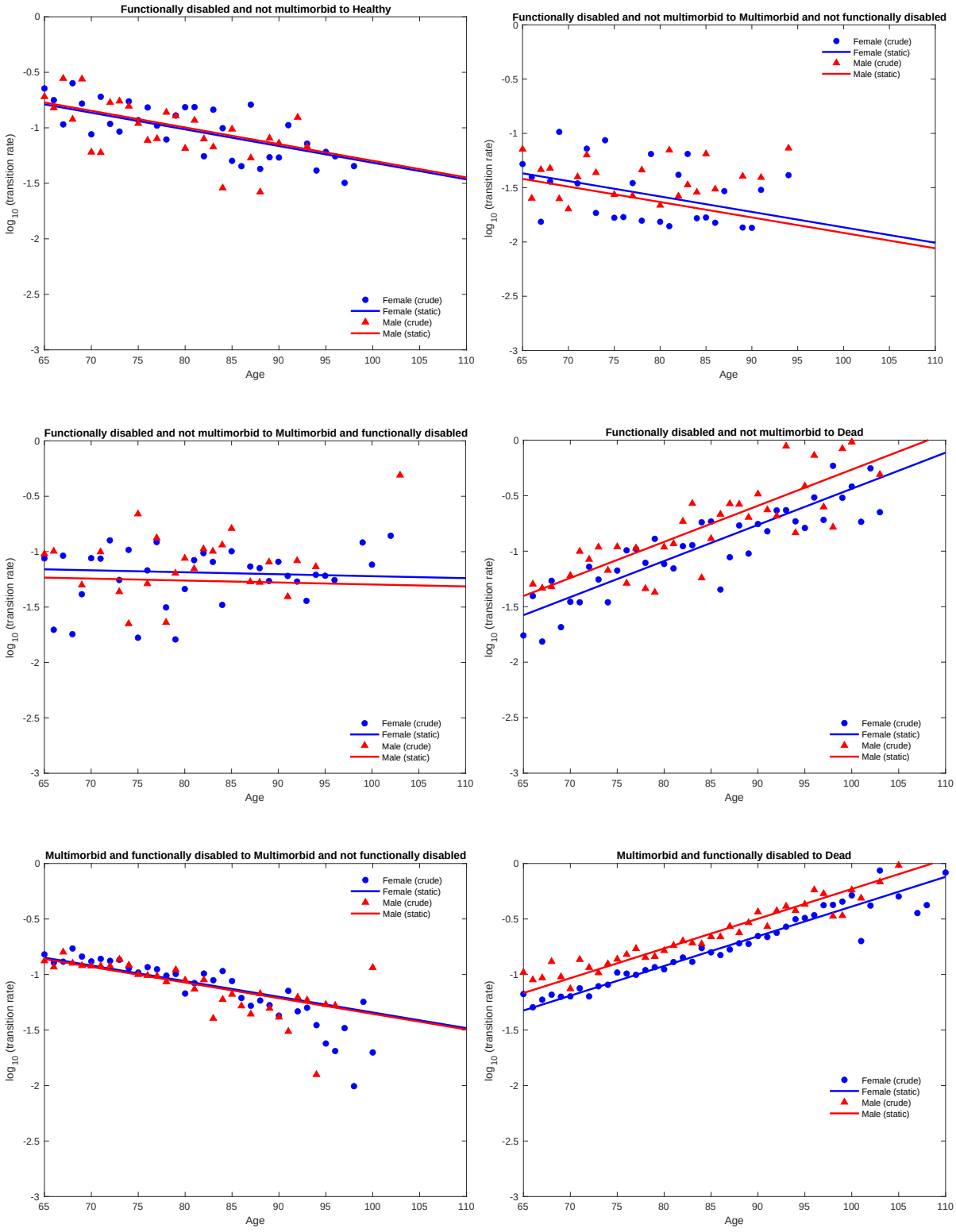


Figure 5.14: Crude transition rates at different ages for males and females using 5 state model (s=7 to s=12) overlaid by static

5.4.3 Model parameters

Table 5.5 and Table 5.6 show the estimated parameters for the three state static and trend models with a multimorbidity predictor, respectively. Coefficients for the baseline log intensity using static and trend models reveal that functionally disabled individuals have higher mortality rates than healthy individuals. Disability rates increase with age but recovery rates decrease with age. Females are more likely to become disabled than males. The trend model shows that disability and mortality rates decrease with time and recovery is less likely as time increases. These results corroborate previous findings in the literature using three state models of functional disability (Fong et al., 2015; Fu et al., 2021; Z. Li et al., 2017; Sherris & Wei, 2021). However, through our addition of a multimorbidity predictor we show that multimorbidity increases disability and mortality rates while it also decreases the chance of recovery. Both static and trend models with multimorbidity have lower values of Akaike information Criterion (AIC) and Bayesian Information Criterion (BIC) than models without implying better fits to the data (Akaike, 1974; Schwarz, 1978).

Table 5.5: Three state static model with multimorbidity: estimated parameters with standard errors in parentheses

Transition	H→F		F→H		H→D		F→D	
s	1		2		3		4	
$\hat{\beta}_s$	-4.5836	***	-1.6566	***	-4.7699	***	-3.0919	***
	(0.0276)		(0.0394)		(0.0307)		(0.0564)	
$\hat{\gamma}_s^{\text{age}}$	0.4431	***	-0.3302	***	0.8740	***	0.6315	***
	(0.0105)		(0.0121)		(0.0119)		(0.0139)	
$\hat{\gamma}_s^{\text{female}}$	0.2326	***	0.0192		-0.4462	***	-0.3672	***
	(0.0224)		(0.0310)		(0.0233)		(0.0302)	
$\hat{\gamma}_s^{\text{multimorbidity}}$	0.9124	***	-0.3064	***	0.8601	***	0.3900	***
	(0.0270)		(0.0381)		(0.0299)		(0.0506)	
Log likelihood	-95674							

† *Note:* $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$;

†† Age covariate is calculated using age last birthday.

Table 5.6: Three state trend model with multimorbidity: estimated parameters with standard errors in parentheses

Transition	H→F		F→H		H→D		F→D	
$\hat{\beta}_s$	-4.4453	***	-1.4445	***	-4.6476	***	-2.9798	***
	(0.0321)		(0.0449)		(0.0353)		(0.0605)	
$\hat{\gamma}_s^{\text{age}}$	0.4448	***	-0.3345	***	0.8770	***	0.6320	***
	(0.0105)		(0.0121)		(0.0120)		(0.0140)	
$\hat{\gamma}_s^{\text{female}}$	0.2331	***	0.0103		-0.4469	***	-0.3729	***
	(0.0224)		(0.0311)		(0.0233)		(0.0302)	
$\hat{\gamma}_s^{\text{multimorbidity}}$	0.9357	***	-0.2609	***	0.8827	***	0.4148	***
	(0.0271)		(0.0384)		(0.0300)		(0.0509)	
$\hat{\gamma}_s^{\text{time}}$	-0.1710	***	-0.2596	***	-0.1536	***	-0.1385	***
	(0.0207)		(0.0275)		(0.0224)		(0.0276)	
Log likelihood	-95559							

† *Note:* $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$;

†† Age covariate is calculated using age last birthday.

Table 5.7 and Table 5.8 report the estimated parameters for the five state static and trend models of multimorbidity and functional disability. Mortality rates increase with age (H→D, M→D, F→D and MF→D). Disability rates (H→F, H→MF and M→MF) also increase with age. Multimorbidity rates from the healthy state (H→M, H→MF) increase with age while the multimorbidity rate from the functionally disabled state (F→MF) decreases with age. Recovery rates from functional disabilty (F→H, F→M and MF→M) decrease with age. While the conclusions are similar to Sherris & Wei (2021) due to the consistency in the signs of the coefficients, there are differences in the size of the effects for all parameters. This could be due to the use of an updated dataset and difference in definition of illness. Moreover, here we consider multimorbidity while Sherris & Wei (2021) consider illness as the presence of only one disease from a pool of four diseases, namely: diabetes, heart problems, lung disease and stroke. Consequently this affects the size and number of all transitions amongst the states.

On sex, females have lower mortality rates than males. Females are more likely to become disabled than males. Females are less likely to become multimorbid than males (H→M). However, females are more likely than males to become multimorbid and functionally disabled (H→MF and M→MF). The effect of gender on recovery rates is not significant.

The estimated trend model shows that mortality rates have decreased with time. Generally disability rates have decreased with time but the same effect is not observed for the transition from the healthy state to the multimorbid and functionally disabled state (H→MF). The effect of time on multimorbidity is not significant and ambiguous (H→M, H→MF, and F→MF). Generally recovery rates improve with time except for the transition from F→M.

The baseline intensity parameter shows that individuals who are multimorbid and functionally disabled are the most likely to die. Individuals who are exclusively functionally disabled are more likely to recover into the healthy state than the multimorbid state. Disability rates are highest among those who are multimorbid.

Table 5.7: Five state model for s=1 to s=6: estimated parameters with standard errors in parentheses

Transition	H→M		H→F		H→MF		H→D		M→MF		M→D	
s	1		2		3		4		5		6	
Static												
$\hat{\beta}_s$	-2.6980	***	-5.0637	***	-5.4317	***	-4.8005	***	-3.7013	***	-3.9213	***
	(0.0161)		(0.0508)		(0.0613)		(0.0442)		(0.0231)		(0.0250)	
$\hat{\gamma}_s^{\text{age}}$	0.1943	***	0.6382	***	0.6394	***	1.0238	***	0.3963	***	0.8401	***
	(0.0100)		(0.0266)		(0.0328)		(0.0240)		(0.0121)		(0.0136)	
$\hat{\gamma}_s^{\text{female}}$	-0.0558	***	0.1456	**	0.0550		-0.5423	***	0.2716	***	-0.4190	***
	(0.0215)		(0.0616)		(0.0754)		(0.0535)		(0.0255)		(0.0258)	
Log likelihood	-131,332											
Trend												
$\hat{\beta}_s$	-2.6875	***	-4.9033	***	-5.4352	***	-4.7387	***	-3.5367	***	-3.7694	***
	(0.0234)		(0.0685)		(0.0856)		(0.0607)		(0.0314)		(0.0337)	
$\hat{\gamma}_s^{\text{age}}$	0.1941	***	0.6362	***	0.6394	***	1.0238	***	0.3990	***	0.8441	***
	(0.0100)		(0.0266)		(0.0328)		(0.0240)		(0.0121)		(0.0137)	
$\hat{\gamma}_s^{\text{female}}$	-0.0559	***	0.1451	**	0.0550		-0.5427	***	0.2725	***	-0.4196	***
	(0.0215)		(0.0616)		(0.0754)		(0.0535)		(0.0255)		(0.0258)	
$\hat{\gamma}_s^{\text{time}}$	-0.0124		-0.1972	***	0.0042		-0.0753		-0.1756	***	-0.1627	***
	(0.0202)		(0.0584)		(0.0710)		(0.0513)		(0.0234)		(0.0249)	
Log likelihood	-131,217											
Frailty												

† *Note:* $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$;

†† Age covariate is calculated using age last birthday.

Table 5.8: Five state model for s=7 to s=12: estimated parameters with standard errors in parentheses

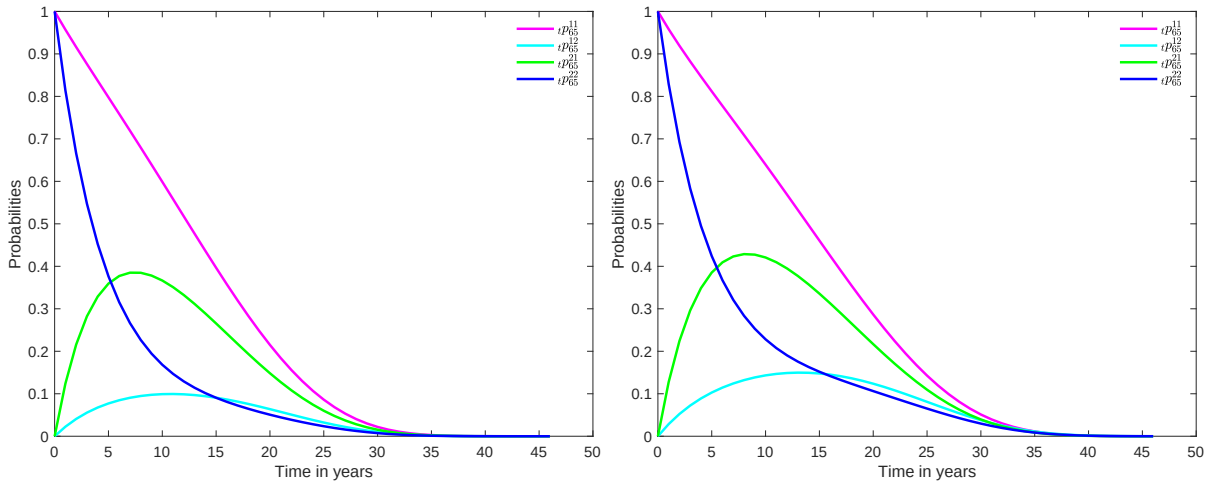
Transition	F→H		F→M		F→MF		F→D		MF→M		MF→D	
s	7		8		9		10		11		12	
Static												
$\hat{\beta}_s$	-1.7771	***	-3.2671	***	-2.8417	***	-3.2331	***	-1.9824	***	-2.6866	***
	(0.0610)		(0.1283)		(0.1012)		(0.1070)		(0.0290)		(0.0340)	
$\hat{\gamma}_s^{\text{age}}$	-0.3457	***	-0.3273	***	-0.0408		0.7491	***	-0.3248	***	0.6162	***
	(0.0297)		(0.0587)		(0.0414)		(0.0433)		(0.0136)		(0.0147)	
$\hat{\gamma}_s^{\text{female}}$	-0.0396		0.1178		0.1723		-0.3970	***	0.0291		-0.3639	***
	(0.0783)		(0.1599)		(0.1214)		(0.0995)		(0.0346)		(0.0317)	
Log likelihood	-131332											
Trend												
$\hat{\beta}_s$	-1.5973	***	-3.2612	***	-2.9376	***	-3.2078	***	-1.7117	***	-2.5433	***
	(0.0852)		(0.1801)		(0.1404)		(0.1325)		(0.0408)		(0.0440)	
$\hat{\gamma}_s^{\text{age}}$	-0.3486	***	-0.3274	***	-0.0393		0.7488	***	-0.3297	***	0.6169	***
	(0.0297)		(0.0587)		(0.0415)		(0.0433)		(0.0136)		(0.0148)	
$\hat{\gamma}_s^{\text{female}}$	-0.0554		0.1173		0.1791		-0.3981	***	0.0220		-0.3698	***
	(0.0785)		(0.1603)		(0.1217)		(0.0995)		(0.0346)		(0.0317)	
$\hat{\gamma}_s^{\text{time}}$	-0.2129	***	-0.0067		0.1088		-0.0302		-0.2748	***	-0.1454	***
	(0.0733)		(0.1447)		(0.1083)		(0.0936)		(0.0303)		(0.0289)	
Log likelihood	-131,217											
Frailty												

† *Note:* $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$;

†† Age covariate is calculated using age last birthday.

5.4.4 Multiyear transition probabilities

Figure 5.15 shows the three state multiyear transition probabilities for males and females aged 65 fitted using the static model. The curve for recovery probabilities is concave downward and recovery rates (${}^{21}_{t}p_x$) are higher than disability rates (${}^{12}_{t}p_x$). The curves showing occupancy probabilities are concave upwards in both healthy and functionally disabled states. Occupancy probabilities (${}^{11}_{t}p_x$) in the healthy state are higher than both recovery and disability probabilities. Females have higher disability and recovery probabilities than males.



(a) Males **(b)** Females

Figure 5.15: Multiyear transition probabilities for males and females aged 65 using three state static model

Figure 5.16 and Figure 5.17 show the five state multiyear transition probabilities for males and females aged 65 fitted using the static model. The probability of becoming multimorbid and not functionally disabled (${}^{13}_{t}p_x$) is consistently higher than that of either becoming functionally disabled (${}^{12}_{t}p_x$) or multimorbid and functionally disabled (${}^{14}_{t}p_x$). Recovery rates (${}^{21}_{t}p_x$) are higher than disability rates. The probability of death is highest for individuals who are multimorbid and functionally disabled and lowest for individuals who are healthy.

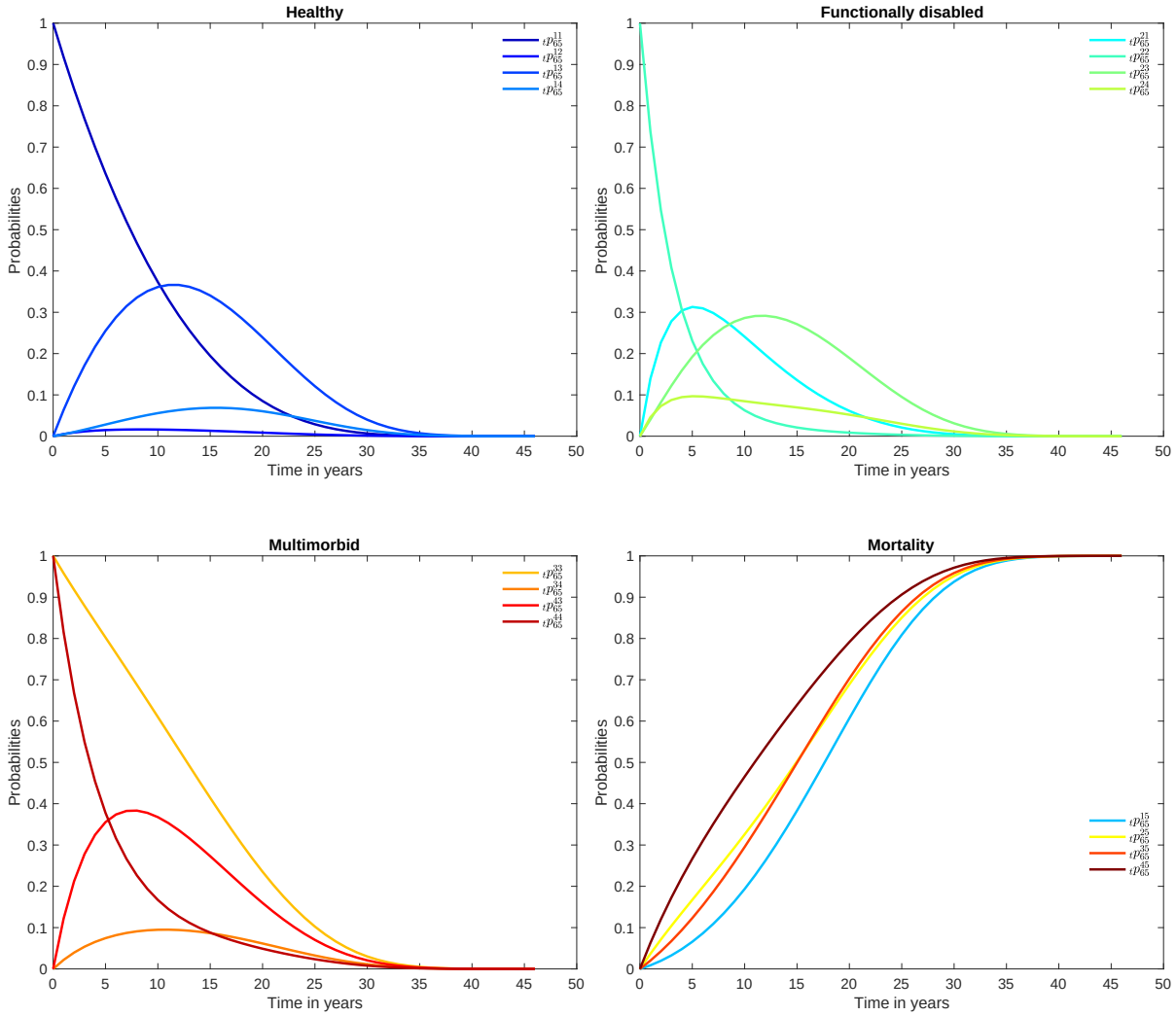


Figure 5.16: Multiyear transition probabilities for males aged 65 using five state static model

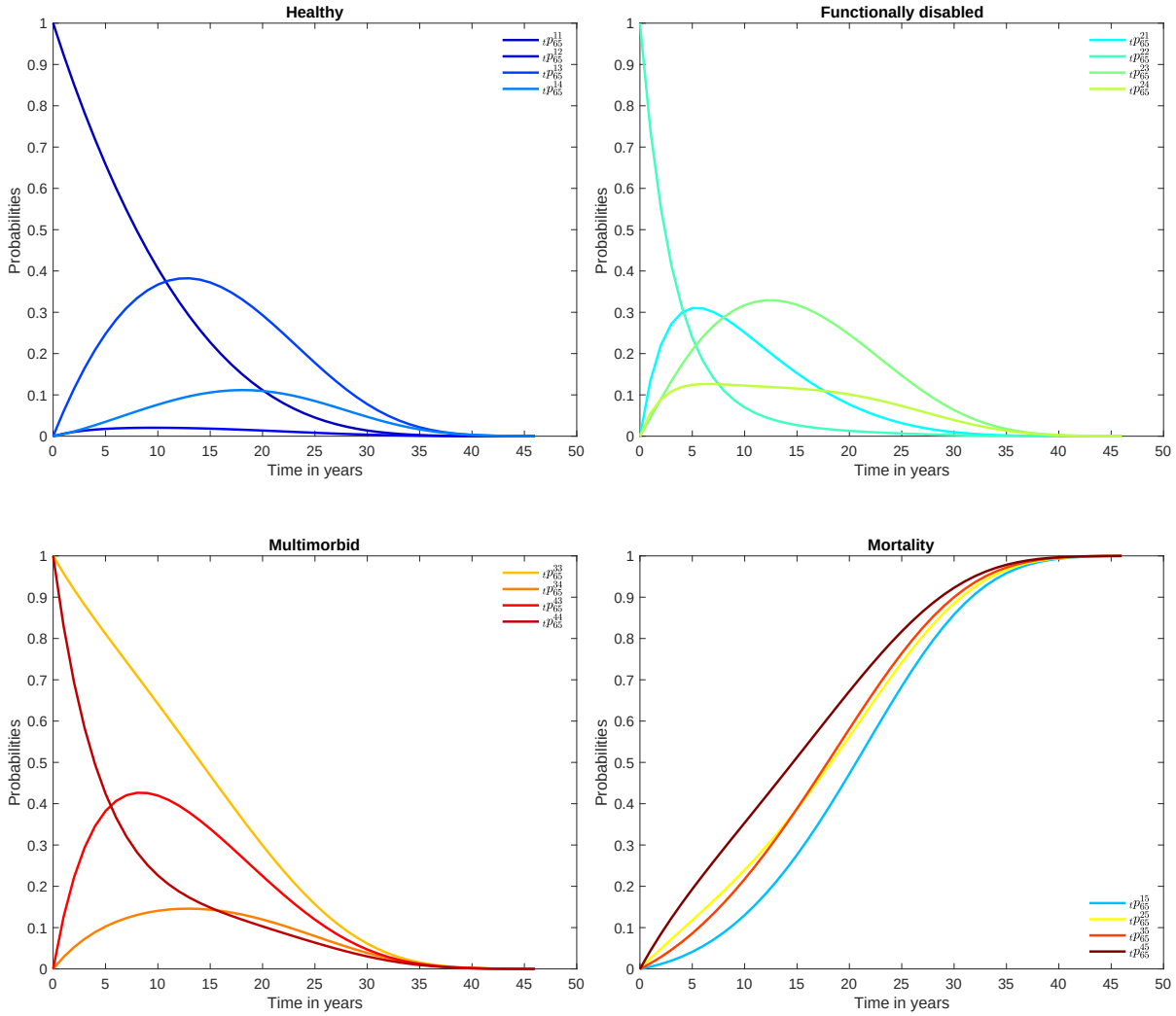


Figure 5.17: Multiyear transition probabilities for females aged 65 using five state static model

5.4.5 Lifetime statistics

Table 5.9 shows the future lifetime statistics for healthy 65–year olds fitted using the three state functional disability static models. Consistent with previous studies using HRS data, generally females have longer life expectancy and spend more time disabled than males (Fong et al., 2015; Fu et al., 2021; Z. Li et al., 2017). We extend these studies by quantifying differences in future lifetime statistics by multimorbidity status. We find that the gap in life expectancy at age 65 between non–multimorbid males and multimorbid males is approximately 7.2 years while that for females is approximately 7.7 years. Additionally, the gap in healthy life expectancy at age 65 between non–multimorbid males and multi-

morbid males is approximately 7.760 years while that for females is approximately 8.574 years. We observe that individuals who are not multimorbid live longer and healthier lives than individuals who are multimorbid. The proportion of Healthy Life Expectancy over Total Life Expectancy (HLE/TLE) of individuals who are not multimorbid is higher than both the multimorbid and the overall group. Individuals who are multimorbid become functionally disabled much earlier than everyone else.

Table 5.9: Future lifetime statistics for 65–year old healthy males and females using static three state models

	Overall		**Multimorbid**		**Not multimorbid**	
	Female	**Male**	**Female**	**Male**	**Female**	**Male**
Life expectancy						
Mean	19.437	16.736	17.762	14.957	25.453	22.166
Standard error	0.090	0.083	0.088	0.080	0.107	0.102
Standard deviation	9.013	8.309	8.817	8.011	10.689	10.155
Healthy life expectancy						
Mean	16.416	14.975	14.436	12.970	23.010	20.730
Standard error	0.085	0.080	0.081	0.076	0.105	0.100
Standard deviation	8.473	8.021	8.089	7.580	10.496	10.043
Disabled life expectancy						
Mean	3.021	1.760	3.327	1.987	2.443	1.436
Standard error	0.044	0.033	0.046	0.035	0.040	0.030
Standard deviation	4.356	3.262	4.587	3.519	3.987	3.032
Healthy life expectancy over life expectancy						
Mean	0.853	0.901	0.827	0.879	0.904	0.934
Standard error	0.002	0.002	0.002	0.002	0.002	0.001
Standard deviation	0.205	0.181	0.226	0.204	0.160	0.141
Age at onset of disability conditional on becoming disabled						
Mean	79.023	78.068	77.129	76.228	83.736	82.035
Standard error	0.112	0.126	0.101	0.110	0.152	0.171
Standard deviation	8.206	7.706	7.588	6.949	10.332	9.632

* Overall represents the model without multimorbidity predictor;

† Maximal age is 110 years.

Table 5.10 and Table 5.11 show the life expectancy results from fitting the three state trend models using the trend in 1998 and 2016, respectively. We find that the gap in life

expectancy at age 65 between non–multimorbid males and multimorbid males is approximately 8.276 years while that for females is approximately 8.855 years using the trend in 1998. Additionally, the gap in healthy life expectancy at age 65 between non–multimorbid males and multimorbid males is approximately 8.793 years while that for females is approximately 9.712 years. We note that life expectancy using the recent time trend from 2016 captures the improvements in mortality as shown by the higher life expectancy and healthy life expectancy for both males and females.

Table 5.10: Future lifetime statistics for healthy 65–year old males and females in 1998 using three state trend models

	Overall		**Multimorbid**		**Not multimorbid**	
	Female	**Male**	**Female**	**Male**	**Female**	**Male**
Life expectancy						
Mean	19.821	16.953	18.336	15.206	27.191	23.482
Standard error	0.095	0.086	0.097	0.087	0.119	0.114
Standard deviation	9.464	8.649	9.737	8.710	11.930	11.358
Healthy life expectancy						
Mean	16.516	15.066	14.691	13.096	24.403	21.889
Standard error	0.087	0.083	0.088	0.082	0.118	0.112
Standard deviation	8.732	8.260	8.813	8.148	11.784	11.226
Disabled life expectancy						
Mean	3.306	1.887	3.646	2.110	2.789	1.593
Standard error	0.048	0.035	0.052	0.038	0.047	0.034
Standard deviation	4.835	3.499	5.176	3.772	4.725	3.411
Healthy life expectancy over life expectancy						
Mean	0.846	0.897	0.819	0.875	0.898	0.931
Standard error	0.002	0.002	0.002	0.002	0.002	0.002
Standard deviation	0.211	0.184	0.234	0.208	0.171	0.147
Age at onset of disability conditional on becoming disabled						
Mean	79.120	78.087	77.048	76.225	83.921	82.433
Standard error	0.116	0.128	0.108	0.117	0.167	0.188
Standard deviation	8.484	7.878	8.058	7.405	11.118	10.467

* Overall represents the model without multimorbidity predictor;

† Maximal age is 110 years.

Even though lifetimes are increasing for both males and females, the number of years spent

disabled is also increasing. The ratio of HLE/TLE is decreasing for the overall model and relatively stable for multimorbid and non–multimorbid groups which suggests morbidity expansion between 1998 and 2016. These results disagree with those in Z. Li et al. (2017) who conclude that there is no strong evidence in the HRS to suggest either morbidity compression or expansion. However, their models cover a shorter duration (1998–2012). Fu et al. (2021) conclude that there is morbidity compression but their study also covers a shorter period (1998-2014).

Table 5.11: Future lifetime statistics for healthy 65–year old males and females in 2016 using three state trend models

	Overall		**Multimorbid**		**Not multimorbid**	
	Female	**Male**	**Female**	**Male**	**Female**	**Male**
Life expectancy						
Mean	20.966	17.956	20.946	17.594	30.006	26.303
Standard error	0.099	0.090	0.106	0.096	0.121	0.119
Standard deviation	9.902	9.007	10.577	9.570	12.110	11.906
Healthy life expectancy						
Mean	17.005	15.710	16.825	15.226	26.870	24.471
Standard error	0.092	0.087	0.100	0.092	0.125	0.120
Standard deviation	9.202	8.675	9.972	9.229	12.460	12.007
Disabled life expectancy						
Mean	3.961	2.246	4.120	2.368	3.136	1.831
Standard error	0.058	0.042	0.060	0.044	0.057	0.041
Standard deviation	5.836	4.205	6.025	4.369	5.687	4.093
Healthy life expectancy over life expectancy						
Mean	0.828	0.885	0.818	0.875	0.894	0.929
Standard error	0.002	0.002	0.002	0.002	0.002	0.002
Standard deviation	0.235	0.205	0.246	0.216	0.188	0.159
Age at onset of disability conditional on becoming disabled						
Mean	79.901	78.779	79.123	77.934	85.311	84.616
Standard error	0.123	0.136	0.126	0.136	0.184	0.210
Standard deviation	8.899	8.271	9.152	8.389	11.616	11.260

* Overall represents the model without multimorbidity predictor;

† Maximal age is 110 years.

Table 5.12 shows the future life statistics of healthy 65–year old individuals using five

state static and trend models. Consistent with other findings, females have higher life expectancy than males (Sherris & Wei, 2021). We find that the gap in healthy life expectancy at age 65 between the healthy state males and multimorbid and not functionally disabled males is approximately 1.993 years while that for females is approximately 1.586 years. Individuals spend more time multimorbid (10.378 years) than they do healthy (9.594 years). Using the most recent trend in 2016 results in a much higher life expectancy for both males and females than that based on the trend in 1998. However, the proportion of years spent multimorbid over total life expectancy increases from 1998 to 2016 which suggests morbidity expansion. This means that gains in life expectancy are counteracted by increases in multimorbidity. Similarly, for individuals in State 1, while there are gains in healthy life expectancy, the proportion of healthy life expectancy over life expectancy decreases from 1998 to 2016 despite the gains in total life expectancy during the same period. This confirms morbidity expansion during the period 1998 to 2016. The mean time spent disabled increases during the same period for both males and females. As observed in Fu et al. (2021), the life expectancy of the trend model in 1998 is higher than that estimated by the static model.

Table 5.12: Future lifetime statistics for healthy 65–year old individuals using five state static and trend models

	Static		**Trend in 1998**		**Trend in 2016**	
	Female	**Male**	**Female**	**Male**	**Female**	**Male**
Life expectancy						
Mean	20.372	17.500	21.247	18.090	23.533	20.166
Standard error	0.087	0.080	0.096	0.088	0.104	0.096
Standard deviation	8.663	8.030	9.591	8.787	10.421	9.630
Healthy life expectancy						
Mean	9.594	8.885	9.607	8.903	9.966	9.243
Standard error	0.074	0.068	0.075	0.069	0.078	0.072
Standard deviation	7.389	6.830	7.525	6.943	7.833	7.221
Functionally disabled and not multimorbid life expectancy						
Mean	0.400	0.302	0.412	0.300	0.320	0.238
Standard error	0.016	0.014	0.017	0.014	0.015	0.013
Standard deviation	1.616	1.348	1.708	1.378	1.496	1.316
Multimorbid and not functionally disabled life expectancy						

Table 5.12: Future lifetime statistics for healthy 65–year old individuals using five state static and trend models *(continued)*

	Female	**Male**	**Female**	**Male**	**Female**	**Male**
Mean	8.008	6.982	8.585	7.432	10.054	8.937
Standard error	0.078	0.073	0.085	0.078	0.097	0.091
Standard deviation	7.795	7.262	8.515	7.834	9.742	9.103
Multimorbid and functionally disabled life expectancy						
Mean	2.369	1.331	2.643	1.454	3.193	1.747
Standard error	0.038	0.028	0.043	0.031	0.053	0.037
Standard deviation	3.853	2.783	4.344	3.082	5.325	3.728
Healthy life expectancy over total life expectancy						
Mean	0.505	0.549	0.494	0.540	0.471	0.515
Standard error	0.003	0.003	0.003	0.003	0.003	0.003
Standard deviation	0.323	0.333	0.323	0.334	0.323	0.336
Mean time spent disabled						
Mean	2.769	1.634	3.055	1.754	3.513	1.986
Standard error	0.055	0.041	0.060	0.045	0.068	0.050
Standard deviation	5.469	4.131	6.052	4.460	6.821	5.044
Mean time spent multimorbid						
Mean	10.378	8.313	11.228	8.886	13.248	10.684
Standard error	0.116	0.100	0.129	0.109	0.151	0.128
Standard deviation	11.648	10.045	12.859	10.917	15.067	12.831
Multimorbidity over life expectancy						
Mean	0.474	0.434	0.485	0.442	0.514	0.471
Standard error	0.003	0.003	0.003	0.003	0.003	0.003
Standard deviation	0.333	0.340	0.333	0.340	0.330	0.342
Age at onset of multimorbidity conditional on becoming multimorbid						
Mean	74.793	73.970	74.756	74.053	75.116	74.365
Standard error	0.079	0.074	0.079	0.076	0.083	0.078
Standard deviation	6.856	6.225	6.850	6.354	7.220	6.606
Age at onset of disability conditional on becoming disabled						
Mean	76.796	75.649	76.088	74.900	76.386	75.136
Standard error	0.240	0.245	0.241	0.244	0.285	0.290
Standard deviation	7.496	6.827	7.453	6.644	7.686	6.815

* Overall represents the model without multimorbidity predictor;

† Maximal age is 110 years.

5.4.6 Pricing implications

Table 5.13 and Table 5.14 show the premiums of various health and longevity risk products using the three state and five state models, respectively. Life annuities premiums are lower for males than females due to a higher mortality experience. Long term care is more expensive for females than males due to females spending more time disabled than males. Premiums using trend models are always higher due to improvements in mortality rates over time. However, these differences are not as magnified as those reported in Sherris & Wei (2021). As observed in other studies, purchasing a life care annuity is cheaper than buying stand alone life annuities or long term care insurance (J. Brown & Warshawsky, 2013; Spillman et al., 2003). Table 5.14 also shows that life annuities are cheaper for individuals who are multimorbid and functionally disabled than for individuals in the healthy state. Long term care is cheaper for the healthy state and more expensive for individuals who are multimorbid.

Table 5.13: Premiums for insurance products using three state model of functional disability for males and females at age 65

	Males		Females	
State	**Static**	**Trend**	**Static**	**Trend**
Life annuity				
Healthy	$131,586	$132,392	$151,488	$153,926
Difference from static		0.61%		1.61%
Long term care				
Healthy	$47,488	$49,524	$77,344	$82,179
Difference from static		4.29%		6.25%
Life care annuity				
Healthy	$163,245	$165,408	$203,051	$208,712
Difference from static		1.33%		2.79%

Table 5.14: Premiums for insurance products using five state model of multimorbidity and functional disability for males and females at age 65

	Males		Females	
State	**Static**	**Trend**	**Static**	**Trend**
Life annuity				
Healthy	151,555	154,540	170,619	175,091
Multimorbid and not functionally disabled	133,878	134,977	152,603	155,336
Long term care				
Healthy	48,576	51,278	78,750	84,870
Multimorbid and not functionally disabled	61,165	64,010	101,471	108,212
Life care annuity				
Healthy	187,344	192,352	228,806	237,872
Multimorbid and not functionally disabled	179,751	182,984	228,706	236,494

Table 5.15 shows the comparison of premiums for insurance products using the three state and five state model. We find that premiums for life annuities, long term care insurance and life care annuities from the five state model are always higher than the three state model premiums for both static and trend models irrespective of gender. For life annuities, what is deemed healthy in the three state model is much closer to the multimorbid and not functionally disabled state than it is to the healthy state. For long term care insurance, healthy individuals have a similar risk of functional disability to those in the three state model. However, those who are multimorbid and not functionally disabled have exorbitantly higher premiums (28.80%). For life care annuities, healthy individuals pay more than those who are multimorbid and not functionally disabled. This suggests that the costs of the life annuity portion of the life care annuity outweighs the costs of the long term care in a combined product.

Table 5.15: Comparison of premiums for insurance products using three state model and five state model for males and females at age 65

	Males		Females	
State	**Static**	**Trend**	**Static**	**Trend**
Difference from healthy state for life annuity				
Healthy	15.18%	16.73%	12.63%	13.75%
Multimorbid and not functionally disabled	1.74%	1.95%	0.74%	0.92%
Difference from healthy state for long term care				
Healthy	2.29%	3.54%	1.82%	3.27%
Multimorbid and not functionally disabled	28.80%	29.25%	31.19%	31.68%
Difference from healthy state for life care annuity				
Healthy	14.76%	16.29%	12.68%	13.97%
Multimorbid and not functionally disabled	10.11%	10.63%	12.63%	13.31%

[a] All premiums compared to healthy state in three state model

5.5 Discussion

5.5.1 Contributions

In this chapter we compare two methods of integrating multimorbidity in multiple state health modelling. We observe that conclusions reached on life expectancy and pricing are very sensitive to the selected multimorbidity model. In particular, we find that there is some weak evidence of morbidity expansion with the three state functional disability model. However, the use of a more nuanced approach such as the proposed five state model of multimorbidity and functional disability provides strong evidence that gains in life expectancy are being lost to increasing multimorbidity which supports the theory of morbidity expansion. These results concur with similar studies on the expansion of multimorbidity albeit with data from different countries (Kingston et al., 2018; Tetzlaff et al., 2017). The five state model of multimorbidity and functional disability is better able to capture the dynamics of how multimorbidity and functional disability evolve over time. This allows us to estimate more realistic values of life expectancy and prices of insurance products. Consequently, beyond the health policy implications, there are pricing

implications that actuaries need to consider when pricing or reserving longevity and health linked insurance products. While the three state functional disability model correctly shows that there are distinct life expectancy and healthy life expectancy patterns between morbid and non–morbid groups; the five state model of multimorbidity and functional disability improves this model by showing that providers can expect to makes significant losses on life annuities, long term care and lifecare annuities due to mispricing of mortality and functional disability risks. In particular, healthy individuals who are free of functional disability or multimorbidity, can cause losses of up to 15.18% for annnuities, 2.29% for long term care and 14.76% for life care annuities when priced using the static model as shown in Table 5.15. Likewise, multimorbid and not functionally disabled individuals can cause losses of up to 1.74% for annnuities, 28.80% for long term care and 10.11% for life care annuities. These losses are even higher with the trend model. This highlights the serious modelling pitfalls and challenges that come with incorporating multimorbidity in long term care pricing.

5.5.2 Limitations

Prior studies in the epidemiology literature have relied on much larger datasets with more than $100,000$ individuals particularly in the UK. We did not have access to such a large dataset but are able to make robust conclusions given that the HRS is a nationally representative study of older adults in the United States. However, for insurance processes, this dataset is quite rich in comparison to the datasets most insurance companies have access to. Secondly, the estimation of mortality at old ages is not robust due to fewer observations at older ages which affects the pricing and life expectancy implications. Models which are more sensitive to mortality at older ages can be considered and adapted. Female mortality might be underestimated by the three state health and functional disability model. Future work will consider adding an interaction term between sex and age to improve the fit. It would also be interesting to see what factors are driving the differences between multimorbid and not–multimorbid groups.

5.6 Conclusion

It is imperative the actuaries account for multimorbidity in long term care due to the increasing prevalence of multimorbidity across the world and its impact on life expectancy. Multimorbidity increases health care costs and reduces quality of life during older age. For the more appropriate design and pricing of actuarially fair products, we need to know how best to incorporate multimorbidity in multiple health state modelling. This has the potential to increase demand for annuities and other health linked longevity products. We compare two methods of multistate methods with recovery from functional disability. We develop a five state model of multimorbidity and functional disability and extend the three state model with functional disability by adding a multimorbidity predictor. We find that the five state multimorbidity and functional disability model is able to capture the dynamics of health over time more accurately than the three state health and functional disability model. The results from the later model strongly suggest morbidity expansion. Our five state model of multimorbidity and functional disability improves the three state models by showing that life annuities, stand alone long term care products and life care annuities are grossly mispriced which can lead to significant losses for annuity and insurance providers.

5.7 Disclosure statement

No potential conflict of interest is reported by the authors.

5.8 Funding

This research is supported by the Australian Research Council Centre of Excellence in Population Ageing Research (project number CE170100005) and the UNSW Business School 2022-23 Sustainable Development Goals Research Grant.

5.9 Acknowledgements

This research includes computations using the computational cluster Katana supported by Research Technology Services at UNSW Sydney. The authors also acknowledge the valuable comments on an earlier version of this chapter from Dr Francesco Ungolo.

Appendix

5.A Data cleaning

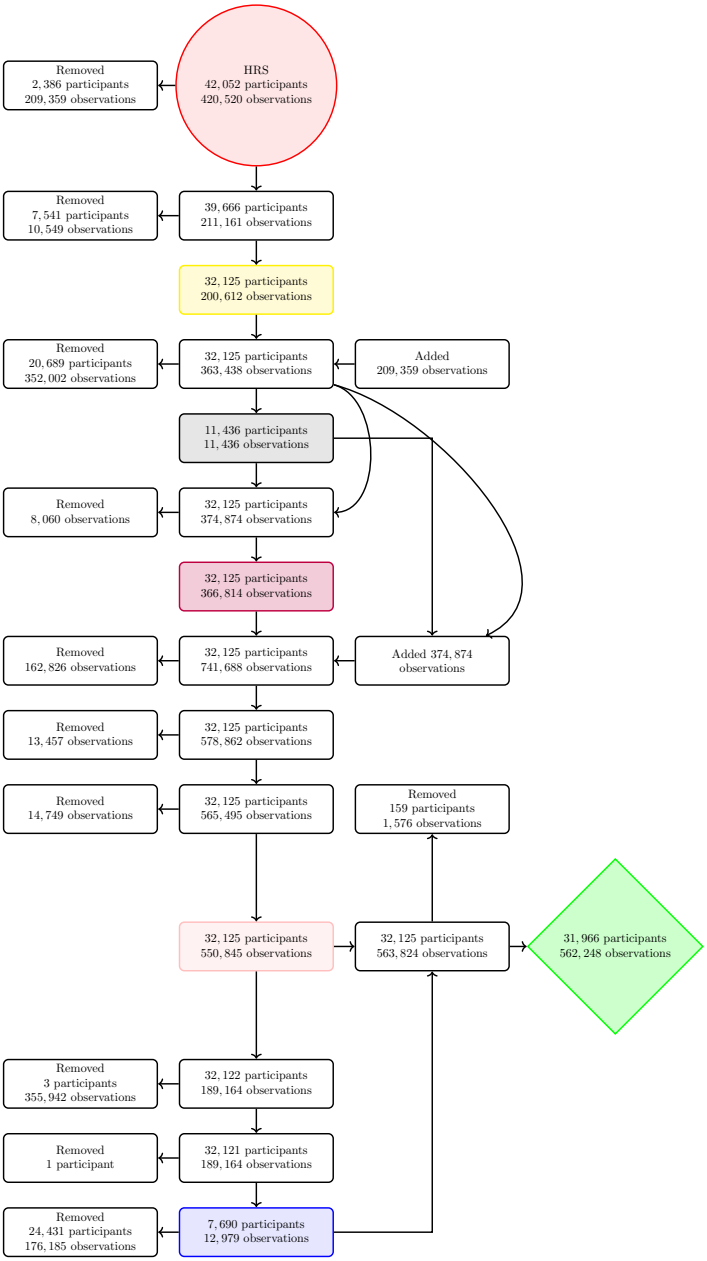


Figure 5.18: Sample selection for HRS dataset using procedure outlined in Appendix 5.B

5.B Sample selection procedure

1. Step 1
    a) Exclude observations from individuals who provide inappropriate responses and do respond (except for those that are non–respondent and died this wave) in any wave.
    b) Exclude participants without consecutive interview dates
    c) Fill in missing calendar years between interview dates
    d) Use date of death as interview date for people who are non-respondent and died this wave
    e) Save file with interviews and filled in calendar years as "~/mydata/interviews.dta"
    f) Save file with death events only as "~/mydata/deathdays.dta"
2. Step 2
    a) Combine interviews with deathdays
    b) Remove duplicates in calendar years
    c) Create Event: "Birthday" in each calendar year including year of death and year of first interview
    d) Save file as "~/mydata/birthdays.dta"
3. Step 3
    a) Combine birthdays with deaths and interviews
    b) Remove individuals with missing dates
    c) Remove birthdays occurring before first interview
    d) Remove events after death
    e) Save as "~/mydata/allexceptmidpoints.dta"
4. Step 4
    a) Keep interview information
    b) Keep individuals who are alive and either healthy or functionally disabled
    c) Create midpoint date for transitions between healthy and disabled states
    d) Extract individuals who have transitioned and save as "~/mydata/midpoints.dta"
5. Step 5

a) Combine file with midpoints with the file with all interviews, birthdays and death days
b) Remove individuals who haven't reached age 45 (using RxAGE) by the end of wave 13, that is, 31/12/2017

5.C Future lifetime statistics for three state health and functional disability model

Table 5.16: Future lifetime statistics for 75–year old healthy individuals using three state static models

	Overall		Multimorbid		Not multimorbid	
	Female	**Male**	**Female**	**Male**	**Female**	**Male**
Life expectancy						
Mean	12.609	10.380	11.520	9.273	17.893	15.023
Standard error	0.070	0.062	0.067	0.058	0.088	0.081
Standard deviation	6.994	6.178	6.728	5.809	8.780	8.062
Healthy life expectancy						
Mean	10.256	9.099	9.042	7.922	15.868	13.902
Standard error	0.064	0.059	0.060	0.054	0.086	0.080
Standard deviation	6.400	5.871	6.001	5.410	8.580	7.946
Disabled life expectancy						
Mean	2.353	1.281	2.479	1.351	2.025	1.121
Standard error	0.037	0.026	0.038	0.026	0.035	0.025
Standard deviation	3.657	2.576	3.749	2.641	3.506	2.500
Healthy life expectancy over life expectancy						
Mean	0.832	0.889	0.811	0.874	0.889	0.926
Standard error	0.002	0.002	0.003	0.002	0.002	0.002
Standard deviation	0.241	0.211	0.259	0.228	0.190	0.164
Age at onset of disability conditional on becoming disabled						
Mean	84.362	83.599	83.368	82.605	88.379	86.990
Standard error	0.086	0.098	0.079	0.085	0.124	0.137
Standard deviation	5.934	5.552	5.511	4.919	8.020	7.255

* Overall represents the model without multimorbidity predictor;

† Maximal age is 110 years.

Table 5.17: Future lifetime statistics for healthy 75–year old individuals in 1998 using three state trend models

	Overall		**Multimorbid**		**Not multimorbid**	
	Female	**Male**	**Female**	**Male**	**Female**	**Male**
Life expectancy						
Mean	12.646	10.342	11.422	9.040	18.527	15.428
Standard error	0.072	0.063	0.071	0.060	0.095	0.088
Standard deviation	7.224	6.314	7.104	6.037	9.501	8.763
Healthy life expectancy						
Mean	10.232	9.048	8.918	7.699	16.326	14.199
Standard error	0.065	0.060	0.063	0.056	0.093	0.086
Standard deviation	6.534	5.960	6.286	5.579	9.311	8.597
Disabled life expectancy						
Mean	2.414	1.294	2.503	1.341	2.201	1.229
Standard error	0.038	0.026	0.038	0.027	0.039	0.028
Standard deviation	3.777	2.606	3.841	2.655	3.866	2.777
Healthy life expectancy over life expectancy						
Mean	0.830	0.887	0.808	0.873	0.884	0.923
Standard error	0.002	0.002	0.003	0.002	0.002	0.002
Standard deviation	0.242	0.211	0.262	0.230	0.197	0.168
Age at onset of disability conditional on becoming disabled						
Mean	84.331	83.523	83.129	82.406	88.253	87.042
Standard error	0.087	0.098	0.080	0.087	0.129	0.144
Standard deviation	6.048	5.613	5.598	4.995	8.236	7.631

* Overall represents the model without multimorbidity predictor;

† Maximal age is 110 years.

Table 5.18: Future lifetime statistics for healthy 75–year old individuals in 2016 using three state trend models

	Overall		**Multimorbid**		**Not multimorbid**	
	Female	**Male**	**Female**	**Male**	**Female**	**Male**
Life expectancy						
Mean	13.668	11.161	13.517	10.825	20.899	17.792
Standard error	0.076	0.067	0.080	0.069	0.100	0.094
Standard deviation	7.613	6.688	7.993	6.891	9.958	9.442
Healthy life expectancy						
Mean	10.749	9.610	10.552	9.268	18.472	16.410
Standard error	0.068	0.063	0.072	0.065	0.100	0.094
Standard deviation	6.851	6.284	7.240	6.504	10.034	9.432
Disabled life expectancy						
Mean	2.919	1.551	2.965	1.557	2.427	1.382
Standard error	0.045	0.032	0.046	0.032	0.045	0.033
Standard deviation	4.525	3.147	4.607	3.178	4.516	3.263
Healthy life expectancy over life expectancy						
Mean	0.814	0.880	0.807	0.874	0.885	0.923
Standard error	0.003	0.002	0.003	0.002	0.002	0.002
Standard deviation	0.260	0.224	0.268	0.234	0.206	0.177
Age at onset of disability conditional on becoming disabled						
Mean	84.964	84.046	84.606	83.599	89.927	88.590
Standard error	0.092	0.103	0.096	0.104	0.145	0.165
Standard deviation	6.389	5.862	6.618	5.907	8.921	8.419

* Overall represents the model without multimorbidity predictor;

† Maximal age is 110 years.

5.D Future lifetime statistics for five state multimorbidity and functional disability model

Table 5.19: Future lifetime statistics for healthy 75–year old individuals using five state static and trend models

	Static		Trend in 1998		Trend in 2016	
	Female	**Male**	**Female**	**Male**	**Female**	**Male**
Life expectancy						
Mean	13.871	11.464	14.092	11.501	15.850	12.979
Standard error	0.069	0.062	0.074	0.065	0.081	0.073
Standard deviation	6.922	6.206	7.386	6.539	8.126	7.272
Healthy life expectancy						
Mean	7.106	6.432	7.096	6.398	7.406	6.739
Standard error	0.055	0.050	0.056	0.051	0.059	0.053
Standard deviation	5.533	5.010	5.613	5.060	5.881	5.316
Functionally disabled and not multimorbid life expectancy						
Mean	0.466	0.306	0.460	0.332	0.390	0.257
Standard error	0.017	0.013	0.017	0.014	0.016	0.012
Standard deviation	1.667	1.274	1.658	1.349	1.548	1.236
Multimorbid and not functionally disabled life expectancy						
Mean	4.608	3.832	4.709	3.837	5.821	4.836
Standard error	0.055	0.049	0.057	0.051	0.068	0.061
Standard deviation	5.463	4.932	5.681	5.092	6.836	6.104
Multimorbid and functionally disabled life expectancy						
Mean	1.692	0.893	1.826	0.934	2.232	1.147
Standard error	0.031	0.021	0.034	0.022	0.041	0.027
Standard deviation	3.110	2.112	3.384	2.254	4.068	2.725
Healthy life expectancy over total life expectancy						
Mean	0.558	0.615	0.555	0.615	0.528	0.590
Standard error	0.003	0.003	0.003	0.003	0.003	0.003
Standard deviation	0.331	0.338	0.332	0.338	0.336	0.344
Mean time spent disabled						
Mean	2.158	1.200	2.287	1.266	2.623	1.404
Standard error	0.048	0.034	0.050	0.036	0.056	0.040
Standard deviation	4.777	3.386	5.043	3.603	5.616	3.961

Table 5.19: Future lifetime statistics for healthy 75–year old individuals using five state static and trend models *(continued)*

	Female	**Male**	**Female**	**Male**	**Female**	**Male**
Mean time spent multimorbid						
Mean	6.300	4.726	6.536	4.772	8.053	5.983
Standard error	0.086	0.070	0.091	0.074	0.109	0.088
Standard deviation	8.572	7.044	9.065	7.346	10.904	8.829
Multimorbidity over life expectancy						
Mean	0.408	0.358	0.410	0.355	0.445	0.388
Standard error	0.003	0.003	0.003	0.003	0.003	0.004
Standard deviation	0.341	0.343	0.341	0.343	0.344	0.349
Age at onset of multimorbidity conditional on becoming multimorbid						
Mean	82.516	81.745	82.558	81.760	82.783	82.032
Standard error	0.064	0.061	0.065	0.062	0.068	0.063
Standard deviation	5.117	4.593	5.185	4.667	5.432	4.813
Age at onset of disability conditional on becoming disabled						
Mean	83.314	82.510	82.739	82.283	83.261	82.735
Standard error	0.160	0.166	0.154	0.162	0.180	0.206
Standard deviation	5.502	4.945	5.288	4.953	5.512	5.374

[*] Overall represents the model without multimorbidity predictor;

[†] Maximal age is 110 years.

Chapter 6

Conclusion

We set out to demonstrate how actuaries can merge revolutionary ideas from the disparate fields of machine learning, epidemiology and statistics to improve product design and pricing. Our main motivation is to encourage the wider adoption and coverage of annuities and other longevity risk products by taking advantage of the existing mortality heterogeneity in the population. The use of health trajectories from longitudinal individual level data is relatively unexplored in the actuarial literature. This thesis introduces this concept to an actuarial audience and elucidates how we can identify risk profiles using determinants of mortality and morbidity such as body mass index, multimorbidity and self reported health.

In Chapter 3 we applied a k–means clustering algorithm adapted for longitudinal data to determine risk profiles that demonstrate mortality heterogeneity. After significant trial and error, we settled on trajectories of self reported health and body mass index. To our knowledge, this is the first study in the actuarial literature that uses health trajectories to determine mortality risk profiles. Earlier work used subjective determination of risk groups or used static single variables such as education, income and deprivation indices to determine different risk profiles. With this method, we are able to identify three clusters that are unique across different socio–economic characteristics: a normal, stable BMI and declining very good health (A), a normal, stable BMI and declining fair health (B) and a high, increasing BMI and declining good health (C). The most interesting result from this analysis is that relying on trajectories of BMI alone masks the effects of having fair health and a normal BMI. Without considering their health status, these individuals would

be treated as healthy and overcharged for annuities despite having the worst longevity prospects. Estimated probabilities of death are even worse than those with increasing BMI and good health.

Moreover, in Chapter 4 we build on our findings in Chapter 3 by using a more robust technique that enables us to estimate meaningful distances between trajectories of categorical variables. We use hidden Markov models with covariates which are a powerful statistical tool that allows us to map each individual's health trajectory onto a hidden Markov model from which we can determine the likelihood of the hidden Markov model. After establishing a distance matrix between trajectories, we are able to perform k–medoids clustering to determine risk profiles. Using multistate models of functional disability, we are able to demonstrate that models with clustering have a better fit to empirical data than those without. The frailty model with clustering has the most superior performance. The static model has no time trends and is unable to capture the changes in the composition of the data over time. Multivariate health trajectories are able to capture mortality heterogeneity amongst different clusters. Using these clusters warrants us to offer better insurance products that cover all risk levels and increase the demand of longevity risk products.

A very interesting result that emanates from work in Chapters 3 and 4 is that clusters determined from just two variables are able to show differences in other socio–economic characteristics such as education, wealth and income across clusters. This demonstrates the effectiveness of clustering in that you might not need all these other variables in actuarial modelling if you have selected the right clustering variables. This is particularly important given the usually limiting restrictions on which variables are used to price products in the insurance industry. However, the decision of which clustering variables to use requires significant judgement and exploratory analysis.

Finally, in Chapter 5 we compare two strategies on the incorporation of multimorbidity in the actuarial pricing models for life annuities, long term care insurance and lifecare annuities. Multimorbidity is common in old age and requires complex treatment which in turn makes health care costs expensive. While there is a wealth of literature showing morbidity expansion, increased mortality risks and socio–economic inequalities associated with multimorbidity; it has been omitted in the actuarial pricing of longevity risk products despite its known impact. Using our proposed five state model of multimorbidity and

functional disability, we find that the omission of multimorbidity in pricing can cause significant losses for insurance and annuity providers. These losses mainly stem from a failure to distinguish between the risk profile of a healthy person (free from both functional disability and multimorbidity) with that of a multimorbid and not functionally disabled individual. In our extension of the three state model of functional disability by adding a multimorbidity predictor, we find that these two risk profiles are merged into one profile which causes providers to assume a lower risk than what is observed in reality. While our results show morbidity expansion, this should not be taken negatively as multimorbidity is the rule and not the exception. Increased longevity should be celebrated. Insurers should take up this challenge and be more innovative in designing longevity and health linked products that meet the needs of most retirees.

Overall, this thesis makes a timely contribution given the vast amounts of data, increase in computational power and the need for fit for purpose insurance products. We acknowledge that while some of the techniques introduced in this thesis are not necessarily new per se, they have not been applied as demonstrated in this thesis. Therefore, this thesis provides a solid foundation on which future work in actuarial precision analytics can be advanced. As such, while there are multiple avenues for future work, we identify three main areas of extension that also continue the trend of working across disciplines.

Both Chapters 3 and 4 utilise trajectories of body mass index for the generation of risk profiles. However, recent research show the limitations of using body mass index as a measure of obesity. Future work could consider other obesity measures that are more robust than BMI which include waist to hip ratio and waist circumference. Since these measures are more sensitive than BMI, we expect to see significant pricing and longevity implications.

The estimation of hidden Markov models with covariates requires large computational resources which are not easily available even when using high performance computing. Therefore, more robust techniques should be employed particularly when the number of individuals is high.

Based on the work in Chapter 5, one natural extension would be to identify the driving factors behind the differences in life expectancy between morbid and non–morbid groups in old age and link this with the pricing of longevity linked products. Some work has

already demonstrated differences in life expectancy due to socio–economic inequalities, smoking status and education and longitudinal data (Chan et al., 2019). However, pricing implications of such analyses are rarely observed in the literature.

Actuaries have long worked in the area of mortality modelling and life insurance; making novel and significant contributions in this field. In this thesis we continue this spirit and encourage researchers and practitioners to work at the intersection of actuarial science, epidemiology and machine learning to gain more insights into mortality heterogeneity. Advances in these areas when combined can help in the achievement of the Sustainable Development Goals, particularly Sustainable Development Goal 1 on eliminating poverty and Sustainable Development Goal 10 on reducing inequalities.

Bibliography

Abbafati, C., Machado, D. B., Cislaghi, B., Salman, O. M., Karanikolos, M., McKee, M., Abbas, K. M., Brady, O. J., Larson, H. J., Trias-Llimós, S., Cummins, S., Langan, S. M., Sartorius, B., Hafiz, A., Jenabi, E., Mohammad Gholi Mezerji, N., Borzouei, S., Azarian, G., Khazaei, S., ... Zhu, C. (2020). Global age-sex-specific fertility, mortality, healthy life expectancy (HALE), and population estimates in 204 countries and territories, 1950–2019: a comprehensive demographic analysis for the Global Burden of Disease Study 2019. *The Lancet*, *396*(10258), 1160–1203. https://doi.org/10.1016/S0140-6736(20)30977-6

Akaike, H. (1974). A New Look at the Statistical Model Identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Akiko, T., Yoshiyuko, O., & JACC Study Group. (2004). Self-reported sleep duration as a predictor of all-cause mortality: results from the JACC study, Japan. *Sleep*, *27*, 51–54.

Alai, D. H., Arnold (-Gaille), S., Bajekal, M., & Villegas, A. M. (2018). Mind the Gap: A Study of Cause-Specific Mortality by Socioeconomic Circumstances. *North American Actuarial Journal*, *22*(2), 161–181. https://doi.org/10.1080/10920277.2017.1377621

Alai, D. H., Arnold (-Gaille), S., & Sherris, M. (2015). Modelling cause-of-death mortality and the impact of cause-elimination. *Annals of Actuarial Science*, *9*(1), 167–186. https://doi.org/10.1017/s174849951400027x

Arcadu, F., Benmansour, F., Maunz, A., Willis, J., Haskova, Z., & Prunotto, M. (2019). Deep learning algorithm predicts diabetic retinopathy progression in individual patients. *Npj Digital Medicine*, *2*(1), 1–9. https://doi.org/10.1038/s41746-019-0172-3

Arnold (-Gaille), S., & Sherris, M. (2013). Forecasting Mortality Trends Allowing for Cause-of-Death Mortality Dependence. *North American Actuarial Journal*, *17*(4),

273–282. https://doi.org/10.1080/10920277.2013.838141

Arnold (-Gaille), S., & Sherris, M. (2015). Causes-of-Death Mortality: What Do We Know on Their Dependence? *North American Actuarial Journal*, *19*(2), 116–128. https://doi.org/10.1080/10920277.2015.1011279

Aro, H., Djehiche, B., & Löfdahl, B. (2014). Stochastic modelling of disability insurance in a multi-period framework. *Scandinavian Actuarial Journal*, *2015*(1), 88–106. https://doi.org/10.1080/03461238.2013.779594

Barro, R. J. (1991). Economic Growth in a Cross Section of Countries. *The Quarterly Journal of Economics*, *106*(2), 407. https://doi.org/10.2307/2937943

Basellini, U., Camarda, C. G., & Booth, H. (2022). Thirty years on: A review of the Lee–Carter method for forecasting mortality. *International Journal of Forecasting*. https://doi.org/10.1016/J.IJFORECAST.2022.11.002

Baum, L. E., & Eagon, J. A. (1967). An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology. *Bulletin of the American Mathematical Society*, *73*(3), 360–363. https://doi.org/bams/1183528841

Bélisle-Pipon, J. C., Vayena, E., Green, R. C., & Cohen, I. G. (2019). Genetic testing, insurance discrimination and medical research: what the United States can learn from peer countries. *Nature Medicine 2019 25:8*, *25*(8), 1198–1204. https://doi.org/10.1038/s41591-019-0534-z

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(8), 1798–1828. https://doi.org/10.1109/TPAMI.2013.50

Benjamini, Y. (2010). Discovering the false discovery rate. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, *72*(4), 405–416. https://doi.org/10.1111/J.1467-9868.2010.00746.X

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/J.2517-6161.1995.TB02031.X

Bergeron-Boucher, M. P., Canudas-Romo, V., Oeppen, J., & Vaupel, J. W. (2017). Coherent forecasts of mortality with compositional data analysis. *Demographic Research*,

37(1), 527–566. https://doi.org/10.4054/DEMRES.2017.37.17

Bergeron-Boucher, M.-P., Kjærgaard, S., Oeppen, J., & Vaupel, J. W. (2019). The impact of the choice of life table statistics when forecasting mortality. *Demographic Research*, *41*, 1235–1268. https://doi.org/10.4054/demres.2019.41.43

Berkman, L. F. (1984). Assessing the Physical Health Effects of Social Networks and Social Support. *Annual Review of Public Health*, *5*(1), 413–432. https://doi.org/10.1146/annurev.pu.05.050184.002213

Bishop, C. M. (1994). *Mixture Density Networks*. Aston University. http://www.ncrg.aston.ac.uk/

Blackburn, C., & Sherris, M. (2013). Consistent dynamic affine mortality models for longevity risk applications. *Insurance: Mathematics and Economics*, *53*(1), 64–73. https://doi.org/10.1016/j.insmatheco.2013.04.007

Booth, H., Maindonald, J., & Smith, L. (2002). Applying Lee-Carter under conditions of variable mortality decline. *Population Studies*, *56*(3), 325–336. https://doi.org/10.1080/00324720215935

Booth, H., & Tickle, L. (2008). Mortality Modelling and Forecasting: a Review of Methods. *Annals of Actuarial Science*, *3*(1-2), 3–43. https://doi.org/10.1017/s1748499500000440

Brouhns, N., Denuit, M., & Vermunt, J. K. (2002). A Poisson log-bilinear regression approach to the construction of projected lifetables. *Insurance: Mathematics and Economics*, *31*(3), 373–393. https://doi.org/10.1016/S0167-6687(02)00185-3

Brown, J. R. (2009). Financial Education and Annuities. *OECD Journal: General Papers*, *2008*(3), 173–215. https://doi.org/10.1787/gen_papers-v2008-art20-en

Brown, J., & Warshawsky, M. (2013). The Life Care Annuity: A New Empirical Examination of an Insurance Innovation That Addresses Problems in the Markets for Life Annuities and Long-Term Care Insurance. *Journal of Risk and Insurance*, *80*(3), 677–704. https://doi.org/10.1111/J.1539-6975.2013.12013.X

Brown, R. L., & McDaid, J. (2003). Factors Affecting Retirement Mortality. *North American Actuarial Journal*, *7*(2), 24–43. https://doi.org/10.1080/10920277.2003.10596083

Brugiavini, A. (1993). Uncertainty resolution and the timing of annuity purchases. *Journal of Public Economics*, *50*(1), 31–62. https://doi.org/10.1016/0047-2727(93)90059-3

Cairns, A. J. G., Blake, D., & Dowd, K. (2006). A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration. *The Journal of Risk and Insurance*, *73*(4), 687–718.

Cairns, A. J. G., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., & Khalaf-Allah, M. (2011). Mortality density forecasts: An analysis of six stochastic mortality models. *Insurance: Mathematics and Economics*, *48*(3), 355–367. https://doi.org/10.1016/j.insmatheco.2010.12.005

Cairns, A. J. G., Blake, D., Dowd, K., Coughlan, G. D., Epstein, D., Ong, A., & Balevich, I. (2009). A Quantitative Comparison of Stochastic Mortality Models Using Data From England and Wales and the United States. *North American Actuarial Journal*, *13*(1), 1–35. https://doi.org/10.1080/10920277.2009.10597538

Cairns, A. J. G., Kallestrup-Lamb, M., Rosenskjold, C., Blake, D., & Dowd, K. (2019). *Modelling socio-economic differences in mortality using a new affluence index* (No. 3; Vol. 49, pp. 555–590). Cambridge University Press. https://doi.org/10.1017/asb.2019.14

Caliñski, T., & Harabasz, J. (1974). A Dendrite Method For Cluster Analysis. *Communications in Statistics*, *3*(1), 1–27. https://doi.org/10.1080/03610927408827101

Cannon, E. S., & Tonks, Ian. (2008). *Annuity markets* (p. 295). Oxford University Press.

Cappelletti, G., Guazzarotti, G., & Tommasino, P. (2013). What determines annuity demand at retirement? *Geneva Papers on Risk and Insurance: Issues and Practice*, *38*(4), 777–802. https://doi.org/10.1057/gpp.2012.43

Cappuccio, F. P., D'Elia, L., Strazzullo, P., & Miller, M. A. (2010). Sleep Duration and All-Cause Mortality: A Systematic Review and Meta-Analysis of Prospective Studies. *Sleep*, *33*(5), 585–592. https://doi.org/10.1093/sleep/33.5.585

Carroll, K. J. (2003). On the use and utility of the Weibull model in the analysis of survival data. *Controlled Clinical Trials*, *24*(6), 682–701. https://doi.org/10.1016/S0197-2456(03)00072-2

Cerhan, J. R., Steven, C., Eric, J., & Cari, M. (2014). *A Pooled Analysis of Waist Circumference and Mortality in 650 , 000 Adults*.

Cesare, M. D., & Murphy, M. (2009). Forecasting Mortality, Different Approaches for Different Cause of Deaths? The Cases of Lung Cancer; Influenza, Pneumonia, and Bronchitis; and Motor Vehicle Accidents. *British Actuarial Journal*, *15*(S1), 185–211.

https://doi.org/10.1017/s1357321700005560

Chan, M. S., Van Den Hout, A., Pujades-Rodriguez, M., Jones, M. M., Matthews, F. E., Jagger, C., Raine, R., & Bajekal, M. (2019). Socio-economic inequalities in life expectancy of older adults with and without multimorbidity: a record linkage study of 1.1 million people in England. *International Journal of Epidemiology*, *48*(4), 1340–1351. https://doi.org/10.1093/IJE/DYZ052

Charlson, M. E., Pompei, P., Ales, K. L., & MacKenzie, C. R. (1987). A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *Journal of Chronic Diseases*, *40*(5), 373–383. https://doi.org/10.1016/0021-9681(87)90171-8

Chassan, M., & Concordet, D. (2023). Computational Statistics and Data Analysis How to test the missing data mechanism in a hidden Markov model. *Computational Statistics and Data Analysis*, *182*, 107723. https://doi.org/10.1016/j.csda.2023.107723

Chau, J. Y., Grunseit, A. C., Chey, T., Stamatakis, E., Brown, W. J., Matthews, C. E., Bauman, A. E., & Van Der Ploeg, H. P. (2013). Daily sitting time and all-cause mortality: A meta-analysis. *PLoS ONE*, *8*(11). https://doi.org/10.1371/journal.pone.0080000

Chen, M. L., Doddi, A., Royer, J., Freschi, L., Schito, M., Ezewudo, M., Kohane, I. S., Beam, A., & Farhat, M. (2019). Beyond multidrug resistance: Leveraging rare variants with machine and statistical learning models in Mycobacterium tuberculosis resistance prediction. *EBioMedicine*, *43*, 356–369. https://doi.org/10.1016/j.ebiom.2019.04.016

Cheng, Y.-J., Chen, Z.-G., Wu, S.-H., Mei, W.-Y., Yao, F.-J., Zhang, M., & Luo, D.-L. (2021). Body mass index trajectories during mid to late life and risks of mortality and cardiovascular outcomes: Results from four prospective cohorts. *EClinicalMedicine*, *33*, 100790. https://doi.org/10.1016/J.ECLINM.2021.100790

Chetty, R., Stepner, M., Abraham, S., Lin, S., Scuderi, B., Turner, N., Bergeron, A., & Cutler, D. (2016). The association between income and life expectancy in the United States, 2001-2014. *JAMA - Journal of the American Medical Association*, *315*(16), 1750–1766. https://doi.org/10.1001/jama.2016.4226

Coale, A. J., & Kisker, E. E. (1990). Defects in Data on Old-Age Mortality in the United States: New Procedures for Calculating Mortality Schedules and Life Tables at the Highest Ages. *Asian and Pacific Population Forum*, *4*(1), 1–32. http://scholarspace.

manoa.hawaii.edu/handle/10125/3598

Coutinho, T., Goel, K., Corrêa De Sá, D., Carter, R. E., Hodge, D. O., Kragelund, C., Kanaya, A. M., Zeller, M., Park, J. S., Kober, L., Torp-Pedersen, C., Cottin, Y., Lorgis, L., Lee, S. H., Kim, Y. J., Thomas, R., Roger, V. L., Somers, V. K., & Lopez-Jimenez, F. (2013). Combining body mass index with measures of central obesity in the assessment of mortality in subjects with coronary disease: Role of "normal weight central obesity". *Journal of the American College of Cardiology*, *61*(5), 553–560. https://doi.org/10.1016/j.jacc.2012.10.035

D'Albis, H., Esso, L. J., & Pifarré I Arolas, H. (2012). *Mortality Convergence Across High-Income Countries : An Econometric Approach*. https://halshs.archives-ouvertes.fr/halshs-00755682

Daal, E. A., & Madan, D. B. (2005). An empirical examination of the variance-gamma model for foreign currency options. *Journal of Business*, *78*(6), 2121–2152. https://doi.org/10.1086/497039

Davies, D. L., & Bouldin, D. W. (1979). A Cluster Separation Measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *PAMI-1*(2), 224–227. https://doi.org/10.1109/TPAMI.1979.4766909

De Groot, V., Beckerman, H., Lankhorst, G. J., & Bouter, L. M. (2003). How to measure comorbidity: a critical review of available methods. *Journal of Clinical Epidemiology*, *56*(3), 221–229. https://doi.org/10.1016/S0895-4356(02)00585-1

De Jong, P., & Tickle, L. (2006). Extending Lee–Carter Mortality Forecasting. *Mathematical Population Studies*, *13*(1), 1–18. https://doi.org/10.1080/08898480500452109

Deaton, A., & Paxson, C. (1999). *Mortality, Education, Income, and Inequality among American Cohorts*. National Bureau of Economic Research. https://doi.org/10.3386/w7140

Deprez, P., Shevchenko, P. V., & Wüthrich, M. V. (2017). Machine learning techniques for mortality modeling. *European Actuarial Journal*, *7*(2), 337–352. https://doi.org/10.1007/s13385-017-0152-4

Diggle, Peter. (2002). *The analysis of longitudinal data*. Oxford University Press.

Dowd, K., Cairns, A. J. G., Blake, D., Coughlan, G. D., Epstein, D., & Khalaf-Allah, M. (2010a). Backtesting stochastic mortality models: An ex post evaluation of multiperiod-ahead density forecasts. *North American Actuarial Journal*, *14*(3),

281–298. https://doi.org/10.1080/10920277.2010.10597592

Dowd, K., Cairns, A. J. G., Blake, D., Coughlan, G. D., Epstein, D., & Khalaf-Allah, M. (2010b). Evaluating the goodness of fit of stochastic mortality models. *Insurance: Mathematics and Economics*, *47*(3), 255–265. https://doi.org/10.1016/j.insmatheco.2010.06.006

Dugravot, A., Fayosse, A., Dumurgier, J., Bouillon, K., Rayana, T. B., Schnitzler, A., Kivimaki, M., Sabia, S., & Singh-Manoux, A. (2020). Social inequalities in multimorbidity, frailty, disability, and transitions to mortality: a 24-year follow-up of the Whitehall II cohort study. *The Lancet Public Health*, *5*(1), e42–e50. https://doi.org/10.1016/S2468-2667(19)30226-9/ATTACHMENT/2E47C273-E211-4F78-B50C-CE760835005F/MMC1.PDF

Dunn, J. C. (1973). A Fuzzy Relative of the ISODATA Process and Its Use in Detecting Compact Well-Separated Clusters. *Journal of Cybernetics*, *3*(3), 32–57. https://doi.org/10.1080/01969727308546046

Dunn, J. C. (1974). Well-Separated Clusters and Optimal Fuzzy Partitions. *Journal of Cybernetics*, *4*(1), 95–104. https://doi.org/10.1080/01969727408546059

Duru, O. K., Harawa, N. T., Kermah, D., & Norris, K. C. (2012). Allostatic load burden and racial disparities in mortality. *Journal of the National Medical Association*, *104*(1-2), 89–95. https://doi.org/10.1016/S0027-9684(15)30120-6

Edwards, R. D., & Tuljapurkar, S. (2005). Inequality in Life Spans and a New Perspective on Mortality Convergence Across Industrialized Countries. *Population and Development Review*, *31*(4), 645–674. https://doi.org/10.1111/j.1728-4457.2005.00092.x

Faraggi, D., & Simon, R. (1995). A neural network model for survival data. *Statistics in Medicine*, *14*(1), 73–82. https://doi.org/10.1002/SIM.4780140108

Fisher, R. A. (1919). XV.—The Correlation between Relatives on the Supposition of Mendelian Inheritance. *Earth and Environmental Science Transactions of The Royal Society of Edinburgh*, *52*(2), 399–433. https://doi.org/10.1017/S0080456800012163

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2011). *Applied longitudinal analysis* (p. 701). Wiley.

Fong, J. H., Shao, A. W., & Sherris, M. (2015). Multistate Actuarial Models of Functional Disability. *North American Actuarial Journal*, *19*(1), 41–59. https://doi.org/10.1080/10920277.2014.978025

Forfar, D. O., McCutcheon, J. J., & Wilkie, A. D. (1988). On graduation by mathematical formula. *Journal of the Institute of Actuaries*, *115*(1), 1–149. https://doi.org/10.1017/s0020268100042633

Fortin, M., Stewart, M., Poitras, M. E., Almirall, J., & Maddocks, H. (2012). A Systematic Review of Prevalence Studies on Multimorbidity: Toward a More Uniform Methodology. *The Annals of Family Medicine*, *10*(2), 142–151. https://doi.org/10.1370/AFM.1337

Fu, Y., Sherris, M., & Xu, M. (2021). Functional disability with systematic trends and uncertainty: a comparison between China and the US. *Annals of Actuarial Science*, 1–30. https://doi.org/10.1017/S1748499521000233

Gallicchio, L., & Kalesan, B. (2009). Sleep duration and mortality: a systematic review and meta-analysis. *Journal of Sleep Research*, *18*(2), 148–158. https://doi.org/10.1111/j.1365-2869.2008.00732.x

García-García, D., Parrado Hernández, E., & Díaz-de María, F. (2009). A new distance measure for model-based sequence clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *31*(7), 1325–1331. https://doi.org/10.1109/TPAMI.2008.268

Gavrilov, L. A., & Gavrilova, N. S. (2019). New Trend in Old-Age Mortality: Gompertzialization of Mortality Trajectory. *Gerontology*, *65*(5), 451–457. https://doi.org/10.1159/000500141

Genolini, C., Alacoque, X., Sentenac, M., & Arnaud, C. (2015). Kml and kml3d: R packages to cluster longitudinal data. *Journal of Statistical Software*, *65*(4), 1–34. https://doi.org/10.18637/jss.v065.i04

Genolini, C., Pingault, J. B., Driss, T., Côté, S., Tremblay, R. E., Vitaro, F., Arnaud, C., & Falissard, B. (2013). KmL3D: A non-parametric algorithm for clustering joint trajectories. *Computer Methods and Programs in Biomedicine*, *109*(1), 104–111. https://doi.org/10.1016/j.cmpb.2012.08.016

Geronimus, A. T., Hicken, M., Keene, D., & Bound, J. (2006). "Weathering" and age patterns of allostatic load scores among blacks and whites in the United States. *American Journal of Public Health*, *96*(5), 826–833. https://doi.org/10.2105/AJPH.2004.060749

Ghassempour, S., Girosi, F., & Maeder, A. (2014). Clustering multivariate time series using Hidden Markov Models. *International Journal of Environmental Research and*

Public Health, *11*(3), 2741–2763. https://doi.org/10.3390/ijerph110302741

Girosi, F., & King, G. (2008). *Demographic Forecasting*. Princeton University Press; Princeton University Press.

Glei, D. A., & Horiuchi, S. (2007). The narrowing sex differential in life expectancy in high-income populations: Effects of differences in the age pattern of mortality. *Population Studies*, *61*(2), 141–159. https://doi.org/10.1080/00324720701331433

Gompertz, B. (1825). On the Nature of the Function Expressive of the Law of Human Mortality, and on a New Mode of Determining the Value of Life Contingencies. *Philosophical Transactions of the Royal Society of London*, *115*, 513–583. http://www.jstor.org/stable/107756

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. The MIT Press.

Greenfield, S., Apolone, G., McNeil, B. J., & Cleary, P. D. (1993). The importance of co-existent disease in the occurrence of postoperative complications and one-year recovery in patients undergoing total hip replacement. Comorbidity and outcomes after hip replacement. *Medical Care*, *31*(2), 141–154. https://doi.org/10.1097/00005650-199302000-00005

Gruenberg, E. M. (1977). The failures of success. *The Milbank Memorial Fund Quarterly. Health and Society*, *55*(1), 3–24. https://doi.org/10.2307/3349592

Gutterman, S., & Vanderhoof, I. T. (1998). Forecasting changes in mortality: A search for a law of causes and effects. *North American Actuarial Journal*, *2*(4), 135–138. https://doi.org/10.1080/10920277.1998.10595759

Haberman, S., & Renshaw, A. (2009). On age-period-cohort parametric mortality rate projections. *Insurance: Mathematics and Economics*, *45*(2), 255–270. https://doi.org/10.1016/j.insmatheco.2009.07.006

Haberman, S., & Renshaw, A. (2011). A comparative study of parametric mortality projection models. *Insurance: Mathematics and Economics*, *48*(1), 35–55. https://doi.org/10.1016/j.insmatheco.2010.09.003

Haberman, S., & Sibbett, T. A. (1995). *History of actuarial science*. W. Pickering & Chatto.

Hainaut, D. (2018). A neural-network analyzer for mortality forecast. *ASTIN Bulletin*, *48*(2), 481–508. https://doi.org/10.1017/asb.2017.45

Hall, M., Hodgson, T., Urwin, R., Deane-Williams, P., Gao, J., Caloisi, A., Martin, I.,

& Johnson, A. (2023). *Global Pension Assets Study 2020* (p. 43). Thinking Ahead Institute Willis Towers Watson. https://www.thinkingaheadinstitute.org/en/Library/Public/Research-and-Ideas/2020/01/Global-Pension-Asset-Study-2020

Hamilton, J. D. (1994). *Time series analysis*. Princeton University Press.

Hanewald, K., Shao, A. W., & Li, H. (2019). Modelling multi-state health transitions in China: A generalised linear model with time trends. *Annals of Actuarial Science*, *13*(1), 145–165. https://doi.org/10.1017/S1748499518000167

Harrell, F. E., Califf, R. M., Pryor, D. B., Lee, K. L., & Rosati, R. A. (1982). Evaluating the Yield of Medical Tests. *JAMA*, *247*(18), 2543–2546. https://doi.org/10.1001/JAMA.1982.03320430047030

Hastie, T., & Stuetzle, W. (1989). Principal curves. *Journal of the American Statistical Association*, *84*(406), 502–516. https://doi.org/10.1080/01621459.1989.10478797

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning Data Mining, Inference, and Prediction* (Second, p. 282). Springer. https://doi.org/10.1007/978-0-387-84858-7

Health and Retirement Study (RAND HRS Longitudinal 2016 (V2)) public use dataset. (2016). *Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740). Ann Arbor, MI, (2016).*

Heligman, L., & Pollard, J. H. (1980). The age pattern of mortality. *Journal of the Institute of Actuaries*, *107*(1), 49–80. https://doi.org/10.1017/s0020268100040257

Hill, K., Barker, B., & Vos, T. (2007). Excess Indigenous mortality: are Indigenous Australians more severely disadvantaged than other Indigenous populations? *International Journal of Epidemiology*, *36*(3), 580–589. https://doi.org/10.1093/ije/dym011

Holm, S. (1979). A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70. http://www.jstor.org/stable/4615733

Horiuchi, S., & Wilmoth, J. R. (1998). Deceleration in the age pattern of mortality at older ages. *Demography*, *35*(4), 391–412. https://doi.org/10.2307/3004009

Huisman, M., Kunst, A. E., Bopp, M., Borgan, J.-K., Borrell, C., Costa, G., Deboosere, P., Gadeyne, S., Glickman, M., Marinacci, C., Minder, C., Regidor, E., Valkonen, T., & Mackenbach, J. P. (2005). Educational inequalities in cause-specific mortality in middle-aged and older men and women in eight western European populations. *The*

Lancet, *365*(9458), 493–500. https://doi.org/10.1016/s0140-6736(05)17867-2

Huque, M. H., Carlin, J. B., Simpson, J. A., & Lee, K. J. (2018). A comparison of multiple imputation methods for missing data in longitudinal studies 01 Mathematical Sciences. *BMC Medical Research Methodology*, *18*(1), 1–16. https://doi.org/10.1186/s12874-018-0615-6

Hyndman, R. J., Booth, H., & Yasmeen, F. (2013). Coherent Mortality Forecasting: The Product-Ratio Method With Functional Time Series Models. *Demography*, *50*(1), 261–283. https://doi.org/10.1007/s13524-012-0145-5

Hyndman, R. J., & Ullah, S. Md. (2007). Robust forecasting of mortality and fertility rates: A functional data approach. *Computational Statistics and Data Analysis*, *51*(10), 4942–4956. https://doi.org/10.1016/j.csda.2006.07.028

Imamura, K., McKinnon, M., Middleton, R., & Black, N. (1997). Reliability of a comorbidity measure: The index of co-existent disease (ICED). *Journal of Clinical Epidemiology*, *50*(9), 1011–1016. https://doi.org/10.1016/S0895-4356(97)00128-5

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning with Applications in R* (G. Casella, S. Fienberg, & I. Olkin, Eds.; Vol. 1, p. 426). Springer. https://doi.org/10.1007/978-1-4614-7138-7

Kaufman, L., & Rousseeuw, P. J. (1987). Clustering by means of medoids. In Y. Dodge (Ed.), *Statistical data analysis based on the L1 norm* (pp. 405–416). North-Holland.

Kingston, A., Robinson, L., Booth, H., Knapp, M., Jagger, C., Adelaja, B., Avendano, M., Bamford, S. M., Banerjee, S., Berwald, S., Bowling, A., Burgon, C., Bustard, E., Comas-Herrera, A., Dangoor, M., Dixon, J., Farina, N., Greengross, S., Grundy, E., … Wittenberg, R. (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

Kleinow, T. (2015). A common age effect model for the mortality of multiple populations. *Insurance: Mathematics and Economics*, *63*, 147–152. https://doi.org/10.1016/j.insmatheco.2015.03.023

Knox, D., & Tomlin, A. (1997). *An Analysis of Pensioner Mortality by Pre-retirement Income*. University of Melbourne. https://fbe.unimelb.edu.au/%7B/_%7D%7B/_%7Ddata/assets/pdf%7B/_%7Dfile/0009/2592135/44.pdf

Kogure, A., Fushimi, T., & Kamiya, S. (2019). Mortality Forecasts for Long-Term Care

Subpopulations with Longevity Risk: A Bayesian Approach. *North American Actuarial Journal*, 1–11. https://doi.org/10.1080/10920277.2019.1653202

Koller, D., Schön, G., Schäfer, I., Glaeske, G., Van Den Bussche, H., & Hansen, H. (2014). Multimorbidity and long-term care dependency - A five-year follow-up. *BMC Geriatrics*, *14*(1), 1–9. https://doi.org/10.1186/1471-2318-14-70/TABLES/3

Koopman, S. J., Lucas, A., & Monteiro, A. (2008a). The multi-state latent factor intensity model for credit rating transitions. *Journal of Econometrics*, *142*(1), 399–424. https://doi.org/10.1016/j.jeconom.2007.07.001

Koopman, S. J., Lucas, A., & Monteiro, A. (2008b). The multi-state latent factor intensity model for credit rating transitions. *Journal of Econometrics*, *142*(1), 399–424. https://doi.org/10.1016/J.JECONOM.2007.07.001

Kramer, M. A. (1991). Nonlinear principal component analysis using autoassociative neural networks. *AIChE Journal*, *37*(2), 233–243. https://doi.org/10.1002/aic.690370209

Kryszczuk, K., & Hurley, P. (2010). *Estimation of the Number of Clusters Using Multiple Clustering Validity Indices* (pp. 114–123). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-12127-2_12

Kulhánová, I., Hoffmann, R., Eikemo, T. A., Menvielle, G., & Mackenbach, J. P. (2014). Educational inequalities in mortality by cause of death: first national data for the Netherlands. *International Journal of Public Health*, *59*(5), 687–696. https://doi.org/10.1007/s00038-014-0576-4

Kulinskaya, E., Gitsels, L. A., Bakbergenuly, I., & Wright, N. R. (2020). Calculation of changes in life expectancy based on proportional hazards model of an intervention. *Insurance: Mathematics and Economics*, *93*, 27–35. https://doi.org/10.1016/j.insmatheco.2020.04.006

Kulinskaya, E., Gitsels, L. A., Bakbergenuly, I., & Wright, N. R. (2021). Dynamic hazards modelling for predictive longevity risk assessment. *Insurance: Mathematics and Economics*, *96*, 222–231. https://doi.org/10.1016/j.insmatheco.2020.11.001

Kullback, S., & Leibler, R. A. (1951). On Information and Sufficiency. *The Annals of Mathematical Statistics*, *22*(1), 79–86. https://doi.org/10.1214/AOMS/1177729694

Le Bras, H. (1976). Lois de mortalité et age limite. *Population*, *31*(3), 655–692. https://doi.org/10.2307/1530761

Lee, R. D., & Carter, L. R. (1992). Modeling and Forecasting U. S. Mortality. *Journal of*

the American Statistical Association, *87*(419), 659. https://doi.org/10.2307/2290201

Lee, R. D., & Miller, T. (2001). Evaluating the Performance of the Lee-Carter Method for Forecasting Mortality. *Demography*, *38*(4), 537–549. https://doi.org/10.1353/dem.2001.0036

Levantesi, S., & Pizzorusso, V. (2019). Application of Machine Learning to Mortality Modeling and Forecasting. *Risks*, *7*(1), 26. https://doi.org/10.3390/risks7010026

Li, H., & Lu, Y. (2019). Modeling cause-of-death mortality using hierarchical Archimedean copula. *Scandinavian Actuarial Journal*, *2019*(3), 247–272. https://doi.org/10.1080/03461238.2018.1546224

Li, N., & Lee, R. (2005). Coherent mortality forecasts for a group of populations: An extension of the lee-carter method. *Demography*, *42*(3), 575–594. https://doi.org/10.1353/dem.2005.0021

Li, Z., Shao, A. W., & Sherris, M. (2017). The Impact of Systematic Trend and Uncertainty on Mortality and Disability in a Multistate Latent Factor Model for Transition Rates. *North American Actuarial Journal*, *21*(4), 594–610. https://doi.org/10.1080/10920277.2017.1330157

Lin, X. S., & Liu, X. (2007). Markov Aging Process and Phase-Type Law of Mortality. *North American Actuarial Journal*, *11*(4), 92–109. https://doi.org/10.1080/10920277.2007.10597486

Liu, X., & Lin, X. S. (2012). A subordinated Markov model for stochastic mortality. *European Actuarial Journal*, *2*(1), 105–127. https://doi.org/10.1007/s13385-012-0047-3

Ludkovski, M., Risk, J., & Zail, H. (2018). Gaussian process models for mortality rates and improvement factors. *ASTIN Bulletin*, *48*(3), 1307–1347. https://doi.org/10.1017/asb.2018.24

Lütkepohl, H. (2005). *New introduction to multiple time series analysis* (pp. 1–764). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-27752-1

Lyu, P., De Waegenaere, A., & Melenberg, B. (2020). A Multi-population Approach to Forecasting All-Cause Mortality Using Cause-of-Death Mortality Data. *North American Actuarial Journal*. https://doi.org/10.1080/10920277.2019.1662316

Macdonald, A. S. (2014). Genetics and Insurance. In *Wiley StatsRef: Statistics reference online*. American Cancer Society. https://doi.org/10.1002/9781118445112.stat04342

MacDonald, B. J., Jones, B., Morrison, R. J., Brown, R. L., & Hardy, M. (2013). Re-

search and Reality: A Literature Review on Drawing Down Retirement Financial Savings. *North American Actuarial Journal*, *17*(3), 181–215. https://doi.org/10.1080/10920277.2013.821938

Mackenbach, J. P., Stirbu, I., Roskam, A.-J. R., Schaap, M. M., Menvielle, G., Leinsalu, M., & Kunst, A. E. (2008). Socioeconomic Inequalities in Health in 22 European Countries. *New England Journal of Medicine*, *358*(23), 2468–2481. https://doi.org/10.1056/NEJMsa0707519

Madan, D. B., Carr, P. P., & Chang, E. C. (1998). The Variance Gamma Process and Option Pricing. *Review of Finance*, *2*(1), 79–105. https://doi.org/10.1023/a:1009703431535

Madan, D. B., & Milne, F. (1991). Option Pricing With V. G. Martingale Components. *Mathematical Finance*, *1*(4), 39–55. https://doi.org/10.1111/j.1467-9965.1991.tb00018.x

Madrigal, A. M., Matthews, F. E., Patel, D. D., Gaches, A. T., & Baxter, S. D. (2011). What Longevity Predictors Should Be Allowed for When Valuing Pension Scheme Liabilities? *British Actuarial Journal*, *16*(1), 1–38. https://doi.org/10.1017/s1357321711000018

Makeham, W. M. (1867). On the Law of Mortality. *Journal of the Institute of Actuaries (1866)*, *13*(06), 325–358. https://doi.org/10.1017/s2046166600003238

Marengoni, A., Angleman, S., Melis, R., Mangialasche, F., Karp, A., Garmen, A., Meinow, B., & Fratiglioni, L. (2011). Aging with multimorbidity: A systematic review of the literature. *Ageing Research Reviews*, *10*(4), 430–439. https://doi.org/10.1016/J.ARR.2011.03.003

Meyricke, R., & Sherris, M. (2013). The determinants of mortality heterogeneity and implications for pricing annuities. *Insurance: Mathematics and Economics*, *53*(2), 379–387. https://doi.org/10.1016/j.insmatheco.2013.06.002

Moreno-Betancur, M., Sadaoui, H., Piffaretti, C., & Rey, G. (2017). Survival Analysis with Multiple Causes of Death. *Epidemiology*, *28*(1), 12–19. https://doi.org/10.1097/EDE.0000000000000531

Murray, C. J. L., & Lopez, A. D. (1997a). Alternative projections of mortality and disability by cause 1990-2020: Global Burden of Disease Study. *Lancet*, *349*(9064), 1498–1504. https://doi.org/10.1016/S0140-6736(96)07492-2

Murray, C. J. L., & Lopez, A. D. (1997b). Global mortality, disability, and the contribution of risk factors: Global burden of disease study. *Lancet*, *349*(9063), 1436–1442. https://doi.org/10.1016/S0140-6736(96)07495-8

Murray, C. J. L., & Lopez, A. D. (1997c). Mortality by cause for eight regions of the world: Global Burden of Disease Study. *Lancet*, *349*(9061), 1269–1276. https://doi.org/10.1016/S0140-6736(96)07493-4

Nagel, G., Peter, R., Braig, S., Hermann, S., Rohrmann, S., & Linseisen, J. (2008). The impact of education on risk factors and the occurrence of multimorbidity in the EPIC-Heidelberg cohort. *BMC Public Health*, *8*(1), 1–10. https://doi.org/10.1186/1471-2458-8-384/TABLES/5

Neira, M., & Prüss-Ustün, A. (2016). *Preventing disease through healthy environments: A global assessment of the environmental burden of disease* (Vol. 259, p. S1). World Health Organization. https://doi.org/10.1016/j.toxlet.2016.07.028

Nelder, J. A., & Wedderburn, R. W. M. (1972). Generalized Linear Models. *Journal of the Royal Statistical Society. Series A (General)*, *135*(3), 370. https://doi.org/10.2307/2344614

Nigri, A., Levantesi, S., Marino, M., Scognamiglio, S., & Perla, F. (2019). A deep learning integrated lee–carter model. *Risks*, *7*(1), 33. https://doi.org/10.3390/risks7010033

O'Meara, T., Sharma, A., & Bruhn, A. (2015). Australia's Piece of the Puzzle – Why Don't Australians Buy Annuities? *Australian Journal of Actuarial Practice*, *3*, 47–57.

Oeppen, J., & Vaupel, J. W. (2002). Demography: Broken limits to life expectancy. *Science*, *296*(5570), 1029–1031. https://doi.org/10.1126/science.1069675

Oksuzyan, A., Juel, K., Vaupel, J. W., & Christensen, K. (2008). *Men: Good health and high mortality. Sex differences in health and aging* (No. 2; Vol. 20, pp. 91–102). Springer International Publishing. https://doi.org/10.1007/BF03324754

Olivieri, A., & Pitacco, E. (2016). Frailty and Risk Classification for Life Annuity Portfolios. *Risks*, *4*(4), 39. https://doi.org/10.3390/risks4040039

Olshansky, S. J., Antonucci, T., Berkman, L., Binstock, R. H., Boersch-Supan, A., Cacioppo, J. T., Carnes, B. A., Carstensen, L. L., Fried, L. P., Goldman, D. P., Jackson, J., Kohli, M., Rother, J., Zheng, Y., & Rowe, J. (2012). Differences In Life Expectancy Due To Race And Educational Differences Are Widening, And Many May Not Catch Up. *Health Affairs*, *31*(8), 1803–1813. https://doi.org/10.1377/hlthaff.2011.0746

Pan, W. (2001). Akaike's Information Criterion in Generalized Estimating Equations. *Biometrics*, *57*(1), 120–125. https://doi.org/10.1111/J.0006-341X.2001.00120.X

Pandolfi, S., Bartolucci, F., & Pennoni, F. (2023). A hidden Markov model for continuous longitudinal data with missing responses and dropout. *Biometrical Journal*, *65*(5), 2200016. https://doi.org/10.1002/BIMJ.202200016

Pearson, K. (1900). X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *50*(302), 157–175. https://doi.org/10.1080/14786440009463897

Perks, W. (1932). On Some Experiments in the Graduation of Mortality Statistics. *Journal of the Institute of Actuaries*, *63*(1), 12–57. https://doi.org/10.1017/s0020268100046680

Pitacco, E. (2004). Survival models in a dynamic context: A survey. *Insurance: Mathematics and Economics*, *35*(2 SPEC. ISS.), 279–298. https://doi.org/10.1016/j.insmatheco.2004.04.001

Pitacco, E. (2019). Heterogeneity in mortality: A survey with an actuarial focus. *European Actuarial Journal*, *9*(1), 3–30. https://doi.org/10.1007/s13385-019-00207-z

Plat, R. (2009). On stochastic mortality modeling. *Insurance: Mathematics and Economics*, *45*(3), 393–404. https://doi.org/10.1016/j.insmatheco.2009.08.006

Preston, S. H., & Elo, I. T. (1995). Are Educational Differentials in Adult Mortality Increasing in the United States? *Journal of Aging and Health*, *7*(4), 476–496. https://doi.org/10.1177/089826439500700402

Pritchard, D. J. (2006). Modeling Disability in Long-Term Care Insurance. *North American Actuarial Journal*, *10*(4), 48–75. https://doi.org/10.1080/10920277.2006.10597413

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rabiner, L. R., & Juang, B. H. (1986). An Introduction to Hidden Markov Models. *IEEE ASSP Magazine*, *3*(1), 4–16. https://doi.org/10.1109/MASSP.1986.1165342

Raftery, A. E., Chunn, J. L., Gerland, P., & Ševčíková, H. (2013). Bayesian Probabilistic Projections of Life Expectancy for All Countries. *Demography*, *50*(3), 777–801. https:

//doi.org/10.1007/s13524-012-0193-x

Rasmussen, C. E. (2004). Gaussian Processes in Machine Learning. In O. Bousquet, U. von Luxburg, & G. Rätsch (Eds.), *Advanced lectures on machine learning: ML summer schools 2003, canberra, australia, february 2 - 14, 2003, tübingen, germany, august 4 - 16, 2003, revised lectures* (pp. 63–71). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-28650-9_4

Ray, S., & Turi, R. H. (1999). Determination of number of clusters in K-means clustering and application in colour segmentation. *The 4th International Conference on Advances in Pattern Recognition and Digital Techniques*, 143. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.587.3517

Renshaw, A. E., & Haberman, S. (1995). On the graduations associated with a multiple state model for permanent health insurance. *Insurance: Mathematics and Economics*, *17*(1), 1–17. https://doi.org/10.1016/0167-6687(95)00011-G

Renshaw, A. E., & Haberman, S. (2003a). Lee-Carter mortality forecasting with age-specific enhancement. *Insurance: Mathematics and Economics*, *33*(2), 255–272. https://doi.org/10.1016/S0167-6687(03)00138-0

Renshaw, A. E., & Haberman, S. (2003b). Lee-carter mortality forecasting: A parallel generalized linear modelling approach for England and Wales mortality projections. *Journal of the Royal Statistical Society. Series C: Applied Statistics*, *52*(1), 119–137. https://doi.org/10.1111/1467-9876.00393

Renshaw, A. E., & Haberman, S. (2006). A cohort-based extension to the Lee-Carter model for mortality reduction factors. *Insurance: Mathematics and Economics*, *38*(3), 556–570. https://doi.org/10.1016/j.insmatheco.2005.12.001

Retirement Plans Experience Committee. (2015). *Mortality Improvement Scale MP-2015* (pp. 1–7). Society of Actuaries. https://www.soa.org/globalassets/assets/files/research/exp-study/research-2015-mp-report.pdf

Richman, R., & Wüthrich, M. V. (2019). A neural network extension of the Lee-Cartermodel to multiple populations. *Annals of Actuarial Science*, 1–21. https://doi.org/10.1017/S1748499519000071

Rickayzen, B. D., & Walsh, D. E. P. (2002). A Multi-State Model of Disability for the United Kingdom: Implications for Future Need for Long-Term Care for the Elderly. *British Actuarial Journal*, *8*(2), 341–393. https://doi.org/10.1017/s1357321700003755

Rieker, P. P., & Bird, C. E. (2005). Rethinking Gender Differences in Health: Why We Need to Integrate Social and Biological Perspectives. *The Journals of Gerontology: Series B*, *60*(Special_Issue_2), S40–S47. https://doi.org/10.1093/geronb/60.special_issue_2.s40

Rimella, L. (2021). *High-dimensional hidden Markov models: methodology, computational issues, solutions and applications* [PhD thesis]. The University of Bristol.

Rocca, W. A., Boyd, C. M., Grossardt, B. R., Bobo, W. V., Finney Rutten, L. J., Roger, V. L., Ebbert, J. O., Therneau, T. M., Yawn, B. P., & St. Sauver, J. L. (2014). Prevalence of Multimorbidity in a Geographically Defined American Population: Patterns by Age, Sex, and Race/Ethnicity. *Mayo Clinic Proceedings*, *89*(10), 1336–1349. https://doi.org/10.1016/J.MAYOCP.2014.07.010

Rogers, R. G. (1995). Marriage, Sex, and Mortality. *Journal of Marriage and the Family*, *57*(2), 515. https://doi.org/10.2307/353703

Rogers, R. G. (1996). The Effects of Family Composition, Health, and Social Support Linkages on Mortality. *Journal of Health and Social Behavior*, *37*(4), 326–338. https://doi.org/10.2307/2137260

Rousseeuw, P. J. (1987). Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, *20*(C), 53–65. https://doi.org/10.1016/0377-0427(87)90125-7

Sahakyan, K. R., Somers, V. K., Rodriguez-Escudero, J. P., Hodge, D. O., Carter, R. E., Sochor, O., Coutinho, T., Jensen, M. D., Roger, V. L., Singh, P., & Lopez-Jimenez, F. (2015). Normal-weight central obesity: Implications for total and cardiovascular mortality. *Annals of Internal Medicine*, *163*(11), 827–835. https://doi.org/10.7326/M14-2525

Schäfer, I., Hansen, H., Schön, G., Höfels, S., Altiner, A., Dahlhaus, A., Gensichen, J., Riedel-Heller, S., Weyerer, S., Blank, W. A., König, H. H., Von Dem Knesebeck, O., Wegscheider, K., Scherer, M., Van Den Bussche, H., & Wiese, B. (2012). The influence of age, gender and socio-economic status on multimorbidity patterns in primary care. first results from the multicare cohort study. *BMC Health Services Research*, *12*(1), 1–15. https://doi.org/10.1186/1472-6963-12-89/TABLES/4

Schiøtz, M. L., Stockmarr, A., Høst, D., Glümer, C., & Frølich, A. (2017). Social disparities in the prevalence of multimorbidity - A register-based population study. *BMC*

Public Health, *17*(1), 1–11. https://doi.org/10.1186/S12889-017-4314-8/TABLES/4

Schünemann, J., Strulik, H., & Trimborn, T. (2017). The gender gap in mortality: How much is explained by behavior? *Journal of Health Economics*, *54*, 79–90. https://doi.org/10.1016/j.jhealeco.2017.04.002

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), 461–464. https://doi.org/10.1214/AOS/1176344136

Shao, A. W., Sherris, M., & Fong, J. H. (2017). Product pricing and solvency capital requirements for long-term care insurance. *Scandinavian Actuarial Journal*, *2017*(2), 175–208. https://doi.org/10.1080/03461238.2015.1095793

Sherris, M., & Wei, P. (2021). A Multi-state Model of Functional Disability and Health Status in the Presence of Systematic Trend and Uncertainty. *North American Actuarial Journal*, *25*(1), 1–23. https://doi.org/10.1080/10920277.2019.1708755

Sherris, M., & Zhou, Q. (2014). Model Risk, Mortality Heterogeneity, and Implications for Solvency and Tail Risk. In *Recreating sustainable retirement*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198719243.003.0007

Siah, K. W., Wong, C. H., Gupta, J., & Lo, A. W. (2022). Multimorbidity and mortality: A data science perspective. *Journal of Multimorbidity and Comorbidity*, *12*, 1–20. https://doi.org/10.1177/26335565221105431

Siler, W. (2007). Parameters of mortality in human populations with widely varying life spans. *Statistics in Medicine*, *2*(3), 373–380. https://doi.org/10.1002/sim.4780020309

Simon, N., Friedman, J., Hastie, T., & Tibshirani, R. (2011). Regularization Paths for Cox's Proportional Hazards Model via Coordinate Descent. *Journal of Statistical Software*, *39*(5), 1–13. https://doi.org/10.18637/JSS.V039.I05

Sithole, T. Z., Haberman, S., & Verrall, R. J. (2000). An investigation into parametric models for mortality projections, with applications to immediate annuitants' and life office pensioners' data. *Insurance: Mathematics and Economics*, *27*(3), 285–312. https://doi.org/10.1016/S0167-6687(00)00054-8

Smith, T. W. (1991). Adult sexual behavior in 1989: number of partners, frequency of intercourse and risk of AIDS. *Family Planning Perspectives*, *23*(3), 102–107. http://www.ncbi.nlm.nih.gov/pubmed/1860475

Spillman, B. C., Murtaugh, C. M., & Warshawsky, M. J. (2003). Policy Implications of an Annuity Approach to Integrating Long-Term Care Financ-

ing and Retirement Income. *Journal of Aging and Health*, *15*(1), 45–73. https://doi.org/10.1177/0898264302239014

Städler, N., & Mukherjee, S. (2013). Penalized estimation in high-dimensional hidden Markov models with state-specific graphical models. *The Annals of Applied Statistics*, *7*(4), 2157–2179. https://doi.org/10.1214/13-AOAS662

Steinorth, P. (2012). The demand for enhanced annuities. *Journal of Public Economics*, *96*(11-12), 973–980. https://doi.org/10.1016/j.jpubeco.2012.07.007

Sterne, J. A. C., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., Wood, A. M., & Carpenter, J. R. (2009). *Multiple imputation for missing data in epidemiological and clinical research: Potential and pitfalls* (No. 7713; Vol. 339, pp. 157–160). British Medical Journal Publishing Group. https://doi.org/10.1136/bmj.b2393

Stoeldraijer, L., Duin, C. van, Wissen, L. J. G. van, & Janssen, F. (2013). Impact of different mortality forecasting methods and explicit assumptions on projected future life expectancy: The case of the Netherlands. *Demographic Research*, *29*, 323–354. https://doi.org/10.4054/demres.2013.29.13

Struyf, A., Hubert, M., & Rousseeuw, P. J. (1997). Clustering in an Object-Oriented Environment. *Journal of Statistical Software*, *1*, 1–30. https://doi.org/10.18637/JSS.V001.I04

Su, S., & Sherris, M. (2012). Heterogeneity of Australian population mortality and implications for a viable life annuity market. *Insurance: Mathematics and Economics*, *51*(2), 322–332. https://doi.org/10.1016/j.insmatheco.2012.05.006

Tabeau, E. (2001). *A Review of Demographic Forecasting Models for Mortality* (pp. 1–32). https://doi.org/10.1007/0-306-47562-6_1

Tabeau, E., Willekens, F., & Poppel, F. van. (2002). *Parameterisation as a tool in analysing age, period and cohort effects on mortality: A case study of the Netherlands* (pp. 141–169). Springer, Dordrecht. https://doi.org/10.1007/978-94-017-3381-6_7

Tetzlaff, J., Muschik, D., Epping, J., Eberhard, S., & Geyer, S. (2017). Expansion or compression of multimorbidity? 10-year development of life years spent in multimorbidity based on health insurance claims data of Lower Saxony, Germany. *International Journal of Public Health*, *62*(6), 679–686. https://doi.org/10.1007/S00038-017-0962-9/METRICS

Thaler, R. H. (1999). Mental accounting matters. *Journal of Behavioral Decision Making*, *12*(3), 183–206. https://doi.org/10.1002/(SICI)1099-0771(199909)12:3%3C183::AID-BDM318%3E3.0.CO;2-F

Thatcher, A. R., Kannisto, Vaino., & Vaupel, J. W. (1998). *The force of mortality at ages 80 to 120*. Odense University Press.

Tiller, J., Morris, S., Rice, T., Barter, K., Riaz, M., Keogh, L., Delatycki, M. B., Otlowski, M., & Lacaze, P. (2019). Genetic discrimination by Australian insurance companies: a survey of consumer experiences. *European Journal of Human Genetics 2019 28:1*, *28*(1), 108–113. https://doi.org/10.1038/s41431-019-0426-1

Trovato, F., & Lalu, N. M. (1998). Contribution of Cause-specific Mortality to Changing Sex Differences in Life Expectancy: Seven Nations Case Study. *Social Biology*, *45*(1-2), 1–20. https://doi.org/10.1080/19485565.1998.9988961

Tukey, J. W. (1949). Comparing Individual Means in the Analysis of Variance. *Biometrics*, *5*(2), 99. https://doi.org/10.2307/3001913

Tuljapurkar, S., & Boe, C. (1998). Mortality change and forecasting: How much and how little do we know? *North American Actuarial Journal*, *2*(4), 13–47. https://doi.org/10.1080/10920277.1998.10595752

Tuljapurkar, S., Li, N., & Boe, C. (2000). A universal pattern of mortality decline in the G7 countries. *Nature*, *405*(6788), 789–792. https://doi.org/10.1038/35015561

Vaillant, G. E., & Mukamal, K. (2001). Successful aging. *American Journal of Psychiatry*, *158*(6), 839–847. https://doi.org/10.1176/appi.ajp.158.6.839

Valderas, J. M., Starfield, B., Sibbald, B., Salisbury, C., & Roland, M. (2009). Defining Comorbidity: Implications for Understanding Health and Health Services. *The Annals of Family Medicine*, *7*(4), 357–363. https://doi.org/10.1370/AFM.983

Van Den Akker, M., Buntinx, F., Metsemakers, J. F. M., & Knottnerus, J. A. (2000). Marginal impact of psychosocial factors on multimorbidity: results of an explorative nested case-control study. *Social Science & Medicine*, *50*(11), 1679–1693. https://doi.org/10.1016/S0277-9536(99)00408-6

Van den Akker, M., Buntix, F., Metsemakers, J. F. M., Roos, S., & Knottnerus, J. A. (1998). Multimorbidity in General Practice: Prevalence, Incidence, and Determinants of Co-Occurring Chronic and Recurrent Diseases. *Journal of Clinical Epidemiology*, *51*(5), 367–375. https://doi.org/10.1016/S0895-4356(97)00306-5

Vaupel, J. W., Manton, K. G., & Stallard, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography*, *16*(3), 439–454. https://doi.org/10.2307/2061224

Villegas, A. M., & Haberman, S. (2014). On the Modeling and Forecasting of Socioeconomic Mortality Differentials: An Application to Deprivation and Mortality in England. *North American Actuarial Journal*, *18*(1), 168–193. https://doi.org/10.1080/10920277.2013.866034

Waldron, I. (1991). Patterns and causes of gender differences in smoking. *Social Science and Medicine*, *32*(9), 989–1005. https://doi.org/10.1016/0277-9536(91)90157-8

Waldron, I. (2012). Contributions of Changing Gender Differences in Behavior and Social Roles to Changing Gender Differences in Mortality. In *Men's health and illness: Gender, power, and the body* (pp. 22–45). SAGE Publications, Inc. https://doi.org/10.4135/9781452243757.n2

Wang, H. C., Yue, J. C., & Tsai, Y. H. (2016). Marital status as a risk factor in life insurance: An empirical study in Taiwan. *ASTIN Bulletin*, *46*(2), 487–505. https://doi.org/10.1017/asb.2016.3

Wang, Q., Hanewald, K., & Wang, X. (2022). Multistate health transition modeling using neural networks. *Journal of Risk and Insurance*, *89*(2), 475–504. https://doi.org/10.1111/JORI.12364

Ward, J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of the American Statistical Association*, *58*(301), 236–244. https://doi.org/10.1080/01621459.1963.10500845

Wei, L. J. (1992). The accelerated failure time model: A useful alternative to the cox regression model in survival analysis. *Statistics in Medicine*, *11*(14-15), 1871–1879. https://doi.org/10.1002/SIM.4780111409

Wen, J., Cairns, A. J. G., & Kleinow, T. (2021). Fitting multi-population mortality models to socio-economic groups. *Annals of Actuarial Science*, *15*(1), 144–172. https://doi.org/10.1017/S1748499520000184

Weng, S. F., Vaz, L., Qureshi, N., & Kai, J. (2019). Prediction of premature all-cause mortality: A prospective general population cohort study comparing machine-learning and standard epidemiological approaches. *PLOS ONE*, *14*(3), e0214365. https://doi.org/10.1371/journal.pone.0214365

White, I. R., Royston, P., & Wood, A. M. (2011). Multiple imputation using chained equations: Issues and guidance for practice. *Statistics in Medicine*, *30*(4), 377–399. https://doi.org/10.1002/sim.4067

Wilmoth, J. R. (1993). *Computational methods for fitting and extrapolating the Lee-Carter model of mortality change* [PhD thesis]. University of California, Berkeley.

Wilmoth, J. R. (1995). Are Mortality Projections always More Pessimistic when Disaggregated by Cause of Death? *Mathematical Population Studies*, *5*(4), 293–319. https://doi.org/10.1080/08898489509525409

Wilson, L. A., Lawson, I. R., & Brass, W. (1962). Multiple disorders in the elderly: A Clinical and Statistical Study. *The Lancet*, *280*(7261), 841–843. https://doi.org/10.1016/S0140-6736(62)90627-X

Wiśniowski, A., Smith, P. W. F., Bijak, J., Raymer, J., & Forster, J. J. (2015). Bayesian Population Forecasting: Extending the Lee-Carter Method. *Demography*, *52*(3), 1035–1059. https://doi.org/10.1007/s13524-015-0389-y

Wong-Fupuy, C., & Haberman, S. (2004). Projecting Mortality Trends: Recent Developments in the United Kingdom and the United States. *North American Actuarial Journal*, *8*(2), 56–83. https://doi.org/10.1080/10920277.2004.10596137

Wu, R. (2016). *Gaussian Process and Functional Data Methods for Mortality Modelling* [PhD thesis]. University of Leicester.

Xu, M., Sherris, M., & Meyricke, R. (2019). Systematic Mortality Improvement Trends and Mortality Heterogeneity: Insights from Individual-Level HRS Data. *North American Actuarial Journal*, *23*(2), 197–219. https://doi.org/10.1080/10920277.2018.1513369

Xu, Y., Sherris, M., & Ziveyi, J. (2019). Continuous-time multi-cohort mortality modelling with affine processes. *Scandinavian Actuarial Journal*, 1–27. https://doi.org/10.1080/03461238.2019.1696223

Yashin, A. I., Vaupel, J. W., & Iachine, I. A. (1994). A duality in aging: the equivalence of mortality models based on radically different concepts. *Mechanisms of Ageing and Development*, *74*(1-2), 1–14. https://doi.org/10.1016/0047-6374(94)90094-9

Zajacova, A., Huzurbazar, S., Greenwood, M., & Nguyen, H. (2015). Long-Term BMI Trajectories and Health in Older Adults: Hierarchical Clustering of Functional Curves. *Journal of Aging and Health*, *27*(8), 1443–1461.

https://doi.org/10.1177/0898264315584329

Zeger, S. L., & Liang, K.-Y. (1986). Longitudinal Data Analysis for Discrete and Continuous Outcomes. *Biometrics*, *42*(1), 121. https://doi.org/10.2307/2531248

Zhao, J., Feng, Q. P., Wu, P., Lupu, R. A., Wilke, R. A., Wells, Q. S., Denny, J. C., & Wei, W. Q. (2019). Learning from Longitudinal Data in Electronic Health Record and Genetic Data to Improve Cardiovascular Event Prediction. *Scientific Reports*, *9*(1), 1–10. https://doi.org/10.1038/s41598-018-36745-x

Zhao, J., Salter-Townshend, M., & O'Hagan, A. (2023). A simulation study for multifactorial genetic disorders to quantify the impact of polygenic risk scores on critical illness insurance. *European Actuarial Journal*, 1–39. https://doi.org/10.1007/S13385-023-00345-5/FIGURES/11

Zheng, H., Tumin, D., & Qian, Z. (2013). Obesity and mortality risk: New findings from body mass index trajectories. *American Journal of Epidemiology*, *178*(11), 1591–1599. https://doi.org/10.1093/aje/kwt179

Zheng, M., & Klein, J. P. (1995). Estimates of Marginal Survival for Dependent Competing Risks Based on an Assumed Copula. *Biometrika*, *82*(1), 127. https://doi.org/10.2307/2337633