Modelling the Impacts of Demographic Ageing on the Demand for Health Care Services

Stephen Darren Clark

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

The University of Leeds

School of Geography

July 2015

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

Acknowledgements

I would like to thank those here at the School of Geography who have been helpful in the provision of guidance and advice during my time undertaking this PhD. I would like to thank Professor Mark Birkin, Dr Alison Heppenstall, Dr Kirk Harland and Professor Phil Rees who, in various combinations, have formed part of my supervision team. Professor Rees has been particularly helpful in the later stages of the PhD to motivate me towards completion. Thanks are also due to the members of my Research and Support team, Dr Paul Norman, Dr Andrew Evans, Dr Myles Gould and Professor John Stillwell and my examiners. I would also like to thank Rachel Homer who was always on hand for sensible and reassuring advice on aspects of the statistics within this PhD. The staff at the Office for National Statistics and the National Centre for Social Research have provided much needed assistance and clarification with the data used in this thesis.

Thanks are also due to my fellow students for welcoming me into the School and providing help through their inspiration. My roommates, Crystal, Ying and Holly and neighbour Emma were always available for a chat when things were not at their best. I have also found my lunch time conversations with fellow student Chris to be reassuring and uplifting. Also Bill, who, like me, came to a PhD late and is someone whom I still miss in so many ways.

Finally the most heartfelt thanks are due to my parents, Maurice and Sheila, who have shown some forbearance in the choices and decisions of their youngest child.

Abstract

This thesis presents a methodology that predicts the number of individuals aged 50 and older who have one or more of three morbidities within each English local authority district to 2031. The three morbidities are cardiovascular disease, diabetes or high blood sugar and respiratory illnesses.

The methodology uses spatial microsimulation to create a representative 2011 base population in each district. This population is then dynamically simulated through time using a process that: ages the population, changes its morbidity status, restructures its composition along demographic lines and replenishes the population at younger ages. An accounting system is used to examine how the demographic changes within each district influence its health status.

In terms of prevalence counts and rates the prediction is for significant reductions in both these measures for CVD. For respiratory illness, the prevalence count remains fairly constant but due the increases in the size of the population at risk, the prevalence rate decreases. With diabetes or high blood sugar, both the prevalence counts and rates increase. Examination of the demographic changes affecting these prevalences shows that for the more ethnically diverse districts the changing ethnic structure has a large impact whilst for the more prospering districts the changing age structure has the largest impact.

These results suggest that public health messages on circulatory and heart conditions and the reduction in smoking will have beneficial health effects in the future, which will help to mitigate the strains placed on the health care system in England. The prospects for diabetes or high blood sugar are however not so good and some considerations on how best to utilise scarce resources to prevent or treat this morbidity are urgently required.

Table of Contents

List of Tables

List of Figures

1 RESEARCH AIMS

Questions and Objectives

This thesis is an applied piece of methodological research in the area of health planning. In particular it is concerned with gaining an understanding of the impacts of population ageing on local health care demand.

This research has the potential to make an important contribution to the understanding of a crucial area of health policy. This importance is linked to recognition that an ageing society is an inevitable process in the context of the United Kingdom population. If anticipated well, there is the potential to mitigate against the negative impacts from such ageing, both in terms of the treatments available to those affected and the efficient use of financial, technological and labour resources.

1.1 MEASURING HEALTH CARE

Health care is a significant item of public and private expenditure in most countries. Recent World Bank (2012) statistics for 2012 estimate that the countries belonging to the Organisation for Economic Co-operation and Development (OECD) spent 12.6% of their Gross Domestic Product (GDP) on health care related costs, with the United States spending the highest at 17.9%. The United Kingdom (UK) is estimated to have spent around 9.4% of its GDP on health, which is below both the average OECD value and that for the European Union of 10.2%.

Whilst the above figures are precise in terms of measuring the costs at a national level, the practical definition of the health of a nation is vaguer. The World Health Organization (WHO) has a definition of health that was formalised over half a century ago; "*a complete state of physical, mental and social well-being, and not merely the absence of disease or infirmity.*" (WHO, 1946). Under this umbrella definition, personal perceptions of health and ill health can however be wide ranging and can be perceived in either an objective or subjective manner. Objectively, it may be measured as the presence of one or more morbidities, usually diagnosed by a health practitioner. These may be present as physical morbidities, that can be either chronic or acute, or as a psychological illness. A subjective perception may be obtained by asking individuals to rate their own health or declare the presence of an illness that might limit their day to day activities. This latter form of health definition is potentially subjective since two individuals with ostensibly the same set of illnesses may describe their category of general health or the limiting nature of these illnesses differently.

Irrespective of how the presence of an illness is defined, once the number of individuals who consider their health to be impaired is known then organisations, both public and private, can begin to plan their services to meet this need. This need maybe met through the primary care sector, such as a local family doctor, nurse or other medical practitioner or may require a referral to secondary provider such as a local clinic or hospital. Accompanying this medical care there may also be the need for social care to provide the patient with convalesce or on-going care in their home or residential institution.

In this study health is defined in an objective manner as the medical practitioner diagnosed presence of one or more of three morbidities that commonly affect the aged population, namely cardio vascular disease (CVD), diabetes or high blood sugar (DHBS) and respiratory illness. This measure of health can be quantified as an actual prevalence count of people with a morbidity or as a prevalence rate, which is this count expressed relative to the population at risk.

1.2 FACTORS AFFECTING POPULATION HEALTH

There are many factors that have the potential to influence a person's health outcome – particularly when looking at specific morbidities. Some factors can be categorised as socio-demographic in nature, such as a person's gender, age or ethnicity. Others can be categorised as more socio-economic, for example the level of education, occupation, wealth or income. A third category covers factors related to the lifestyle of the person, either currently or historically. These factors include their smoking status, the presence of obesity and the degree of alcohol consumption. The remaining factors that impact on health include such determinants as their access to health care (including medication), living or working environment (exposure to harmful or hazardous materials) or injury (work-related or traffic).

1.3 POPULATION OF INTEREST

This study is concerned with those individuals who are aged 50 or older and live in England. This age threshold is chosen to capture the ages at which the morbidities commonly associated with ageing begin to become apparent. A reflection of this is that both data sources used in this study, the 2011 population Census and the English Longitudinal Study of Ageing (ELSA), use this threshold for tabulations (in the case of the Census) and for participation (in the case of ELSA). In terms of geographic area, spending and decision making in the UK National Health Service (NHS) is devolved to the four home countries. Of these home countries, England is chosen as the area for analysis since it is a large, diverse and autonomous home country within the UK with regards to health service provision. Within England, there are a range of health

geographies employed to cover the commissioning, oversight and delivery of health care services. In particular, responsibility for public health planning is devolved to the 326 local authority districts (LADs) in England. In recognition of this role and level of responsibility, the analysis is applied here to each of these LADs. Estimates of prevalence counts and rates for morbidities in LADs are not common in the literature, and even less common are forecasts. This thesis aims to address this deficiency by providing such estimates of current prevalences and forecasting future prevalences.

1.4 RESEARCH QUESTIONS

This thesis provides this contribution by addressing a number of policy related research questions:

1. What is the likely number of people aged 50 and older, who will have one or more of these specific morbidities: CVD; DHBS; or respiratory illness in each English LAD, up to 2031?
2. How do these morbidities vary by the gender, age, ethnicity and other socio-demographic characteristics of the population in the area?
3. How are these predictions influenced by both changes in the demographic composition of an area and by changing trends in health outcomes?
4. What are the public policy implications of the work? How can the scarce health care resources, both financial and personal, be best used to meet this demand?

1.5 OBJECTIVES

To answer these research questions, a number of objectives have been set:

a) The first objective is to gain an understanding of the anticipated challenge associated with an ageing population, and paint a broad, aggregate, picture of the likely size and health status of this population. Allied to this understanding is a review into the literature that describes the trends, determinants and needs associated with health care in an ageing population. Here also there is a requirement to look at the methodologies that have been used in the literature. It is important that the review involves a consideration of both the domestic UK and international literature in this area.
b) The second objective is to identify the data that are available to model the health status of individuals. Data that is both aggregate and dis-aggregate in nature is included in the assessment with a discussion of the strengths and weaknesses of each data set. In concert with the following objective, that explores methodologies, a range of data sources useful to the study are identified.
c) The third objective builds on what has been learnt to identify and outline a methodology to estimate the future prevalence for a range of morbidities.

d) The fourth objective is to apply the methodology and illustrate the outcomes in both a tabular and geographic manner that allows trends and features to be explored.
e) The final objective is to provide an overall summary of the research findings, highlighting the public policy implications of the work, and illustrates how the work may be enhanced and taken forward.

1.6 THESIS STRUCTURE

Following on from this introductory chapter, the second chapter sets the context for the research. This chapter draws mainly on the available official statistics for England collected and published by such organisations as the UK Office for National Statistics (ONS) and the Health and Social Care Information Centre (HSCIC), which is part of the NHS. The third chapter provides a review of published studies relevant to this area of study. This review includes evidence for the significant health care determinants and the types of data and models used in such studies. It also includes a review of the important field of population projection, since the future composition of the population is likely to be an important driver of health care use. These two chapters relate to the first objective.

The fourth chapter discusses in detail the two sources of data used in this study, the 2011 Census and ELSA, which meets the second objective. Objective three is dealt with in Chapter 5, which outlines the modelling framework adopted in this study, breaking each component down and showing how they interact.

There then follow five results chapters that achieve the penultimate objective. The sixth chapter is concerned with the spatial microsimulation that establishes the base population in 2011 for each LAD. The seventh chapter is concerned with the adjustments applied to an ethnic group population projection to account for the performance of the projection methodology between the 2001 and 2011 Censuses. The eighth chapter covers the estimation and calculation of the probabilities to be used in the Monte Carlo simulations. The next results chapter, Chapter 9, presents the headline outcomes of the research - the prevalence for a range of morbidities in each LAD - biennially from 2013 to 2031. The final results chapter, ten, decomposes these headline estimates to examine the relative contribution to changes in health prevalences due to changes in the demographic nature of the LAD.

Following these results chapters, the eleventh chapter provides a discussion on the research and satisfies the final objective. This chapter summarises the thesis and includes a discussion of the policy implications of the work and how it might be enhanced. After the final chapter there is a glossary, reference maps and the thesis references.

2 RESEARCH CONTEXT

Population Ageing, Population Projections and Health Forecast

2.1 INTRODUCTION

The first objective of this thesis is to gain an understanding of the likely challenges associated with an ageing population and the likely trends in health care need associated with this ageing. This chapter provides the demographic context of this study which will help to provide this understanding.

The chapter primarily involves an examination of the size and spatial distribution of the older population within England and how this population is expected to change, both in absolute and relative terms, over the medium term. This change will be explored using official Principal population projections and a number of variants to this Principal projection. There is recognition, however, that the size of the older population will not itself fully determine future health care demand and it is also necessary to consider trends in the health status of the population. This status is explored using the concept of healthy or disability free life expectancy. Use is also made of administrative health data to show where both the demand for care of the older population and the provision of services is concentrated.

The section 2.2 of this chapter provides the international context to the research and illustrates how the demographic shift to an ageing population in the UK compares with other EU nations. Sections 2.3 to 2.7 use a range of official ONS population projections to show how the population structure of the English population is anticipated to change over the period 2011 to 2031, and also provide a geographic perspective on how these changes may vary across England. In section 2.8 a historic and geographic perspective is provided on variations in life and disability free life expectancies, again using official ONS estimates. Section 2.9 presents administrative data obtained from the HSCIC to map the demand for health care from the older population and where this care is provided. The final section, 2.10, provides a summary of the findings identified in this chapter. The interested reader may also like to view the slides on the health and health care of older people produced by the HSCIC (2014a).

2.2 INTERNATIONAL PERSPECTIVE

Many western societies are predicting an important shift in the structure of their populations. The clear and anticipated trend is for the elderly population to increase both in terms of the number of such people and as a proportion of the total population

(Rechel et al, 2009). Published 2015 projected trends in this demographic, calculated using a common set of assumptions and methodologies, for the 28 EU nations (plus Iceland, Norway and Switzerland) are illustrated in Table 2-1 (European Commission, 2014). This figure illustrates the universality of this trend amongst these countries, at the extremes the percentage point growths vary from a low of a 2.0% increase for Sweden to a 9.2% increase for Lithuania. For the EU as a whole, the proportion of the population in the age 65 and older age band is anticipated to increase from 19% of the total population in 2015 to 24% in 2031. Using these projections and highlighting the position of the UK, the proportion of those aged 65 and older is seen to increase modestly from 17.7% to 21.6% (+3.9%), moving the UK from the country with the 18th highest percentage to the 23rd highest in 2031. One possible driver for this relative change in the position of the UK is that it has historically had higher levels of inward net migration at younger age bands than other countries (Wadsworth and Vaitilingam, 2014) helping inflate the denominator in this calculation.

Turning to alternative population projections that are concerned with just England, and broadening the age range to those aged 50 and older, the latest Principal 2012 population projections from the ONS project that by 2031 the number of such older people in England will be 24,027,000 (ONS, 2013a). This is an increase of nearly a third on 2011 estimates and in 2031 those over 50 will make up nearly 40% of the population.

The question arises as to what this ageing phenomenon will mean for society at large (Rutherford, 2011 and House of Lords, 2013 and, in an English regional context, Rees et al., 2013a). Whilst longer life expectancies are to be celebrated, the ageing of the population is, at best, seen as a challenge (Christensen et al., 2009) and at worst, a threat (Laurence, 2002). Much of the challenge lies in how this longer life expectancy is utilised. If the individual has the necessary skills, work is available and they are healthy, these three factors mean that some of these extra years of life expectancy may be spent working, assuaging concerns over productivity (Department of Business, Innovation and Science, 2010a and 2010b) and pension commitments (Harper et al., 2011).

In terms of provision of services for an ageing population, health is an area where the impact of an ageing population may be most keenly felt (Wanless, 2004 and Craig and Mindell, 2005). As hypothesised above, if many of the extra years of life expectancy are spent in a healthy or disability-free state, then concerns over the health, activity and care levels of society may be addressed. If, however, an ageing population is associated with a less healthy population, it may also impact on the ability of the remaining population to physically and financially support and care for the aged population.

Table 2-1 : Trends in the projected proportion of population aged 65 or older, EU28 plus Norway, Switzerland and Iceland, 2015 to 2031

Country	2015	2019	2023	2027	2031	Change 2015 to 2031
Sweden	**20%**	**20%**	**21%**	**21%**	**22%**	**2.0%**
Luxembourg	14%	15%	15%	16%	17%	2.8%
Norway	16%	17%	18%	18%	19%	3.0%
Belgium	18%	19%	19%	20%	21%	3.5%
United Kingdom	**18%**	**18%**	**19%**	**20%**	**22%**	**3.8%**
Romania	17%	18%	20%	21%	21%	3.9%
Denmark	19%	20%	21%	21%	23%	4.0%
Italy	21%	22%	23%	24%	26%	4.1%
Hungary	18%	19%	21%	22%	22%	4.1%
Bulgaria	20%	21%	22%	24%	24%	4.5%
Czech Republic	18%	20%	21%	22%	22%	4.6%
Finland	20%	22%	23%	24%	25%	4.6%
Switzerland	18%	19%	20%	21%	23%	4.7%
France	18%	20%	21%	22%	23%	5.0%
Austria	19%	19%	20%	22%	24%	5.2%
Greece	21%	22%	23%	24%	26%	5.3%
EU28	**19%**	**20%**	**21%**	**23%**	**24%**	**5.4%**
Estonia	19%	20%	22%	23%	24%	5.7%
Croatia	19%	20%	22%	23%	24%	5.7%
Malta	19%	21%	22%	24%	24%	5.9%
Iceland	14%	15%	17%	18%	20%	6.4%
Latvia	19%	20%	22%	24%	26%	6.4%
Cyprus	14%	16%	17%	19%	21%	6.6%
Ireland	13%	14%	16%	18%	20%	6.7%
Netherlands	18%	19%	21%	23%	25%	6.7%
Portugal	20%	22%	23%	25%	27%	6.8%
Germany	21%	23%	24%	26%	28%	6.9%
Spain	18%	20%	21%	23%	26%	7.1%
Slovenia	18%	20%	22%	24%	25%	7.4%
Poland	15%	17%	20%	22%	23%	7.6%
Slovakia	14%	16%	18%	20%	22%	7.9%
Lithuania	**19%**	**20%**	**22%**	**25%**	**28%**	**9.2%**

(source : Health in Europe: Information and Data Interface, 2013)

2.3 POPULATION AGE AND GENDER STRUCTURE

How the structure of the English population is projected to change is illustrated in the population structure diagram in Figure 2-1. This shows the English male and female population structure in 2011 and by how this is anticipated to change by 2031. For the ages 50 and older, the net change in the size of the population from 2011 to 2031 for each five year age band is shown within the bars of this diagram.

The structure of the population does not alter radically between 2011 and 2031, with most age bands showing some growth over this period. There is, however, a broadening of the structure at the older age bands. These projections suggest that the 50 year and older population age band will increase by 5.7 million by 2031. Whilst the younger age population of 20 to 50 year olds will also grow, this growth is expected to be a more modest 323,000. Calculations based on these data shown in Figure 2-1, estimates that the mean male age was 38.8 years in 2011, and the prediction is that this will increase to 41.6 in 2031. For females, the mean age will increase from 40.9 years in 2011 to 43.6 in 2031. The difference in the mean age between the genders will therefore decrease slightly from 2.1 years to 2.0 by 2031.

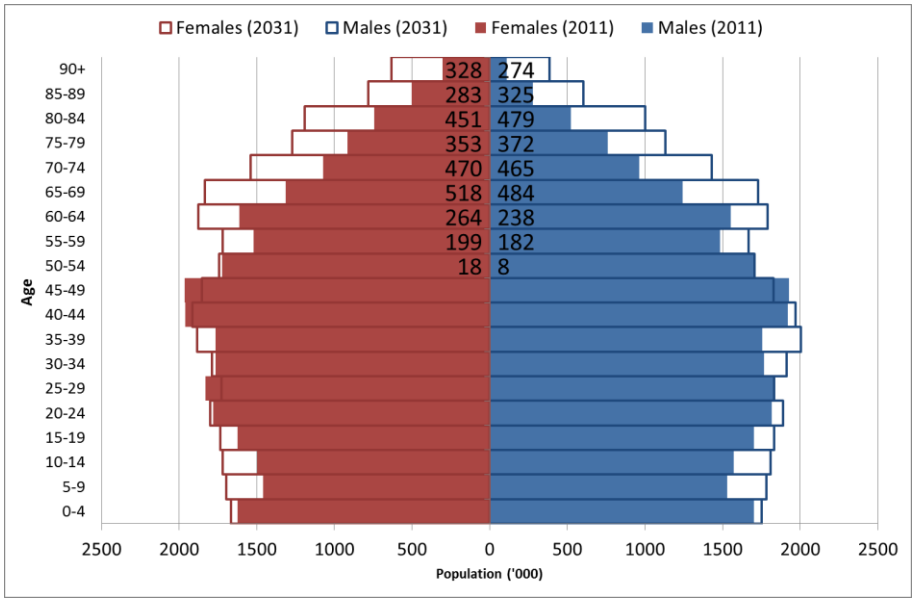


Figure 2-1 : Projected Principal population structure, England, 2011 and 2031

(source : 2011 counts are 2011 mid-year estimates, ONS, 2012a, and 2031 are Principal 2012-based projections, ONS, 2013a)

Thus whilst most age bands in the population are predicted to increase and therefore require additional health care resources, the increase in the size of the older population is particularly large. The evidence from studies suggests that the use of health care resources by the older population is disproportionately larger than that of the population as a whole (for hospital and social care utilisation at the end of life see Dixon et al., 2004, and Bardsley et al., 2010, whilst for general practice consultations see Hippisley-

Cox and Vinogradova, 2008). The expectation is therefore that, all other things being equal, there will be a larger increase in health costs in comparison to population size, occurring at a time of expected tighter government fiscal circumstances (Appleby, 2013 and Wittenberg et al., 2013).

2.4 GEOGRAPHIC DISTRIBUTION

This growth in the older population will vary by location. The ONS provide Principal sub national population projections for English LADs (ONS, 2014a). Figure 2-2 below shows the number of older people aged 50 or older projected to be living in each of these 324 local authorities in 2011 and 2031 (in this thesis the City of London is merged with Westminster and the Isles of Scilly are merged with Cornwall).

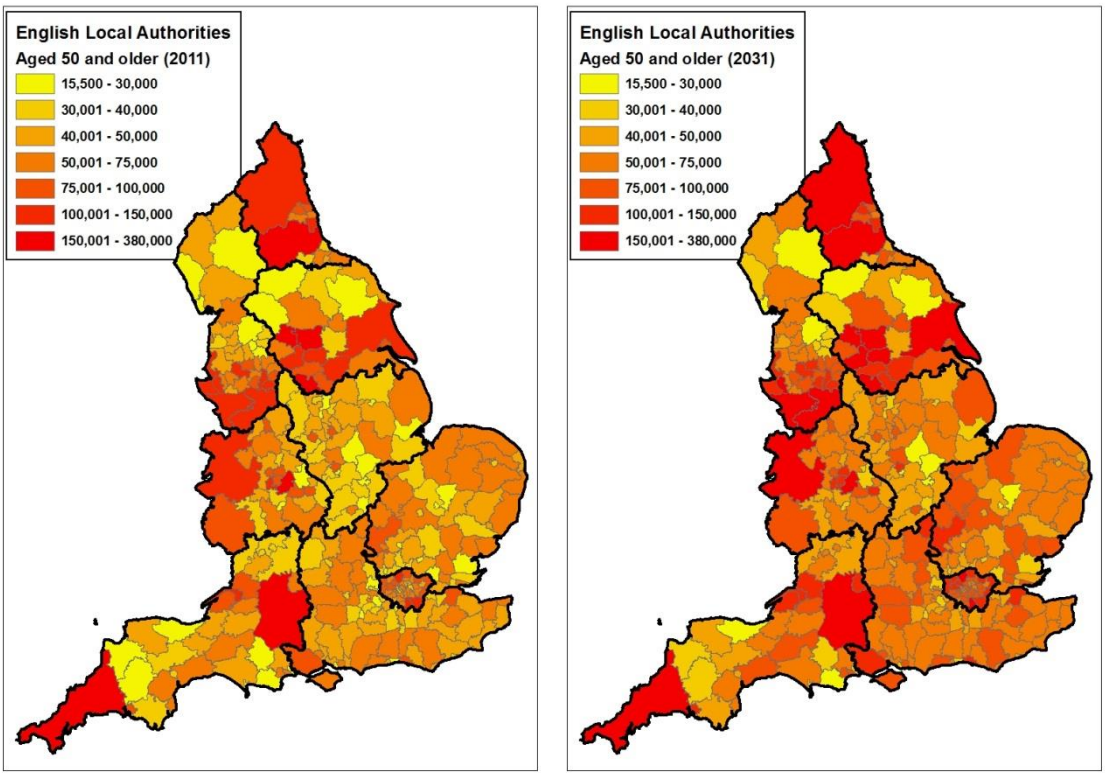


Figure 2-2 : Number of people aged 50 and older, LADs, 2011 (lhs) and 2031 (rhs)

(source : 2011 counts are 2011 interim sub national projections, ONS, 2012b, and 2031 are Principal 2012-based sub national projections, ONS, 2014a)

In 2011 the areas with the largest population of older people are either the larger metropolitan urban areas (e.g. Greater Manchester; West Midlands and South and West Yorkshire) or the geographically larger local authorities (e.g. Cornwall, Wiltshire; East Yorkshire; Durham and Northumberland). This is to be expected since these areas have high population by virtue of them being very densely populated or covering a large geographic area. By 2031 these areas have been joined by other of authorities. The authorities around core cities (e.g. Nottingham; Leicester; Derby and Bristol) as well as

many more authorities in the South and East of England now have a large older population. The fringes of Greater London also show a much larger older population. The growth in the older population between 2011 and 2031 for each local authority is shown in Figure 2-3 (lhs) as an absolute change in numbers and in Figure 2-3 (rhs) as a proportionate change, relative to the 2011 elderly population.

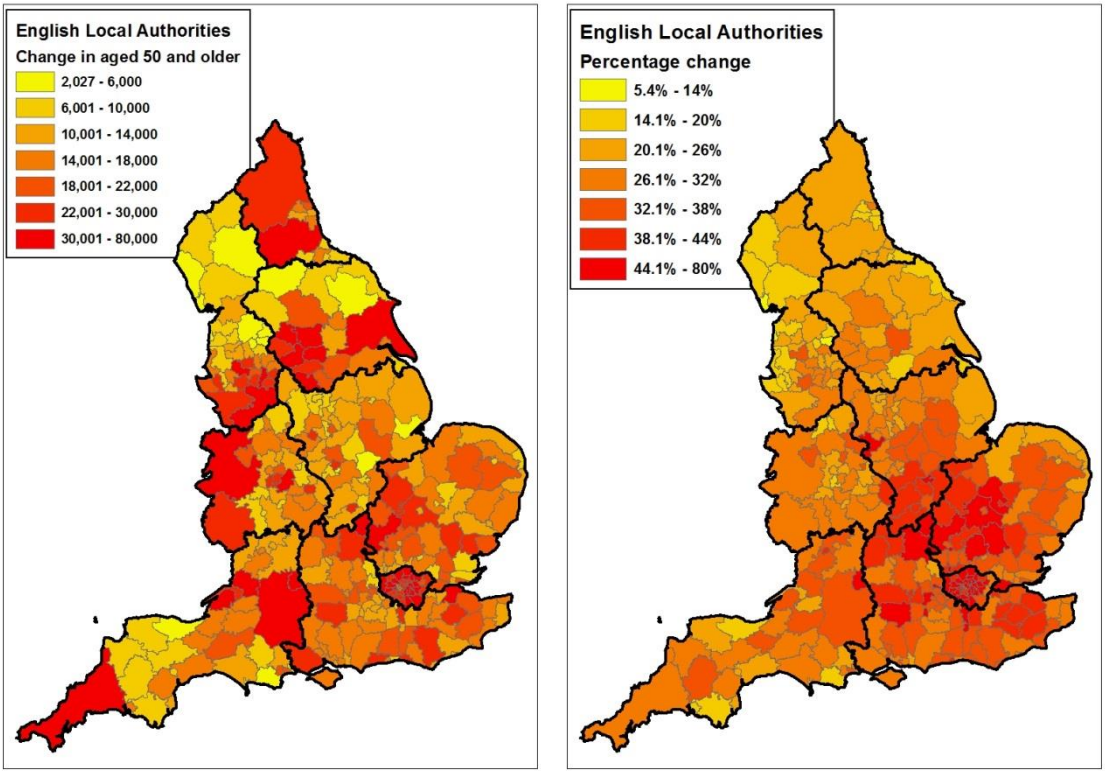


Figure 2-3 : Change (lhs) and proportionate growth (rhs) in 50 year and older population, LADs, 2011 to 2031

(source : 2011 counts are 2011 interim sub national projections, ONS, 2012b, and 2031 are Principal 2012-based sub national projections, ONS, 2014a)

The map showing the size of the older population in 2011 (Figure 2-2, lhs) is not dissimilar from the map here that shows the growth in the size of the population (Figure 2-3, lhs), in that larger areas, either in terms of population or geographic area show the largest growths in the older population. This feature is confirmed in Figure 2-4 (lhs) which shows a scatter plot of the population in 2011 against the absolute change in the size of the 50 and older population, and has a positive correlation.

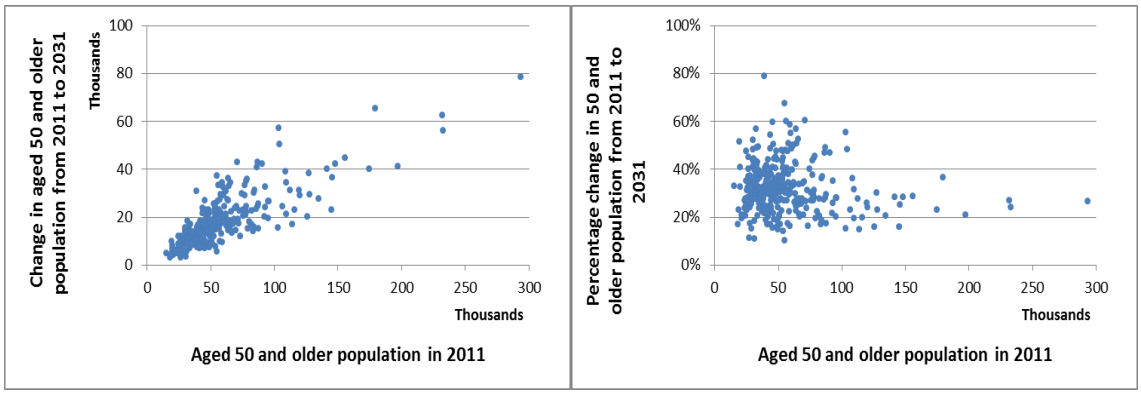


Figure 2-4 : 2011 population verses change in population to 2031 (lhs) and percentage change to 2021 (rhs)

(source : 2011 counts are 2011 interim sub national projections, ONS, 2012b, and 2031 are Principal 2012-based sub national projections, ONS, 2014a)

Looking at the map of the proportionate growth in the older population (Figure 2-3, rhs) and the scatter plot of the 2011 population verses percentage growth (Figure 2-4, lhs) however shows a different pattern. Here there is a significant and disproportionate growth in a band of local authorities just to the north of London and the LADs with small older populations in 2011. These types of areas are starting from a low population in 2011 and may also be seen as attractive locations for those aged over 50, particularly the recently retired, to relocate to. Many local authorities in this area will see their older population increase by a third over this 20 year period. A consideration of Figure 2-2, Figure 2-3 and Figure 2-4 would indicate that those areas that already have a high older population are going to see the biggest increases in this population, but some LADS with smaller older population and those located in the northern Home Counties are going to see a larger proportionate growth.

The changes seen here are the result of two processes, the ageing of the in-situ population who become older and the in- and out- flows of elderly people. Using a log-linear model of both 2001 Census and NHS registration data, Raymer et al. (2007) showed trends in the tendency for migration by those aged 65 or older between 12 ONS area classifications (ONS, 2003). They concentrated on two flows of particular interest: the younger, healthier, migrating from centres of industry; and the older elderly, with health issues, migrating from coastal and countryside areas. Over their short time horizon between 2000/01 and 2003/4 the first flow was seen to increase by 8% whilst the second flow decreased slightly by 2%. Work by Baylis and Slay (2010) in their Figure 1.10 show large migrations of residents aged 50 years and older from London to the East and South West Regions. This migration flow may be triggered by a variety of life course events, e.g. re-location of family, location of amenities, health status or need for long term care. Evandrou et al. (2010) used panel survey data to measure internal migration of the elderly participants and how this activity relates to life events and

socio-demographic status and Stockdale (2011) highlights issues around the geographic variation in ageing – particularly in connection with migration; labour markets and health and care status.

2.5 AGEING

Using the 2011-interim and 2012 sub national population projections published by ONS it is possible to estimate the mean age of those who are aged 50 and older in each LAD from 2011 to 2031. Since the ages are banded in 5 year age bands up to 90, the mid age in each band is used as the estimated age for those in the band. For the open age band of 90 years old the estimated age is 95 in 2011, and to reflect the ageing in this open age band, is increased annually by 0.1 of a year to age 97 by 2031. Figure 2-5 shows the geographical distribution of this mean age estimate for LADs in 2011 and 2031.

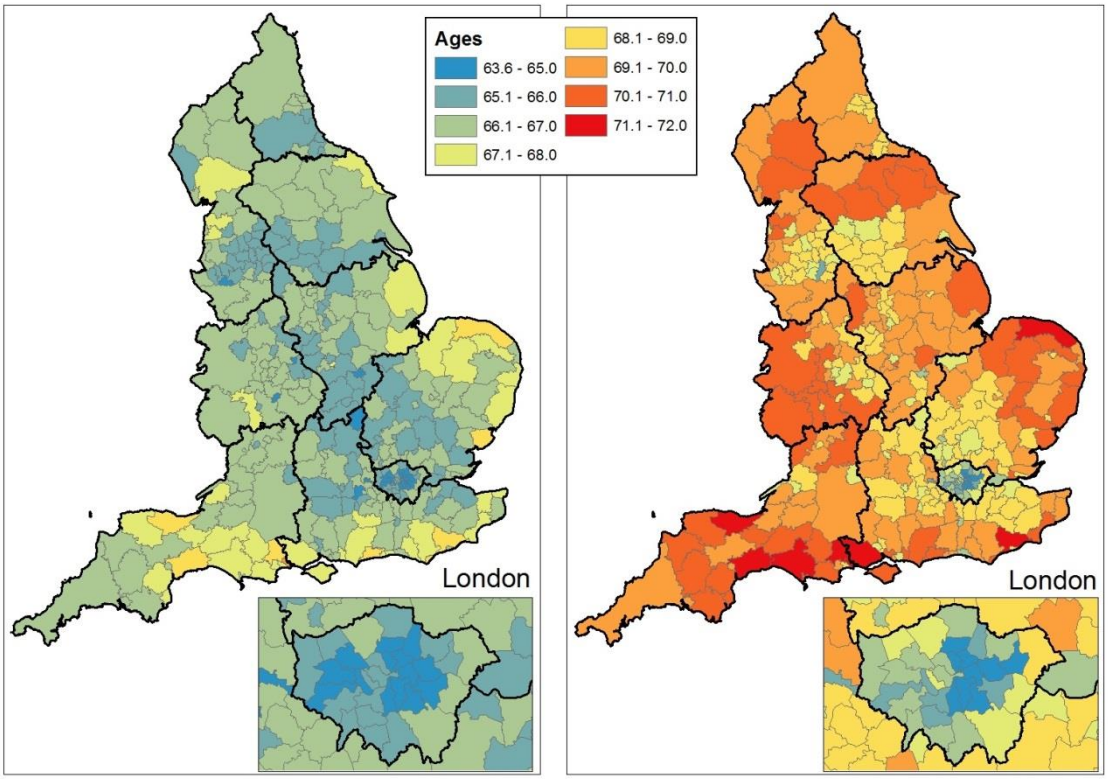


Figure 2-5 : Mean ages, LADs, 2011 (lhs) and 2031 (rhs)

(source : 2011 ages based on 2011-interim mid-year estimates, ONS, 2012b, and 2031 based on Principal 2012-based projections, ONS, 2014a)

In 2011 the LAD with the youngest population is Newham in central London (mean age of 63.7 years old) whilst the oldest is Christchurch in the South west (69.4 years). By 2031 another London Borough, Tower Hamlets has the youngest population (63.2 years) (and actually shows a decrease in the mean age over this 20 year period by 1.1 years) whilst North Norfolk has the oldest population (71.3 years). Eden in the North West shows the greatest increase in the mean age of an additional 4.0 years. More

generally, central London and the more urban LADs are seen to retain a youthful population whilst the large rural authorities outside the South East of England not only have an older population but also show the greatest ageing.

2.6 POPULATION RATIO

Dependency ratios are commonly used to measure the relative sizes of the young (aged 0 to 16) and old population (aged 65 and older) against that of a broader working middle aged population in an area (OECD, 2007). The rationale is that support is provided to these two groups by the middle aged population, both from their family members and also more generally through taxation and labour resources.

In the context of this study where a threshold of aged 50 is adopted for a transition into an older age group the term dependency ratio is not appropriate. Many people are still economically active and also proving support to younger or older family members during their 50's and 60's. Instead a more neutral term of 'population ratio' is adopted. This balance of growth between the generations is therefore expressed as the ratio of the number of people of an older population, aged 50 years or older, to a middle age population (taken to be 20 to 49 years old here) (see equation 2-1). Based on ONS Principal population projections, the anticipated future trends in this population ratio for England are shown in Figure 2-6. In 2011 the national ratio is 0.830 and rises rapidly, until it reaches a value close to 1.0 in 2021. By 2031 the ratio is 1.073.

$$\text{Population ratio} = \frac{\text{Population aged 50 and older}}{\text{Population aged 20 to 49}} \qquad 2\text{-}1$$

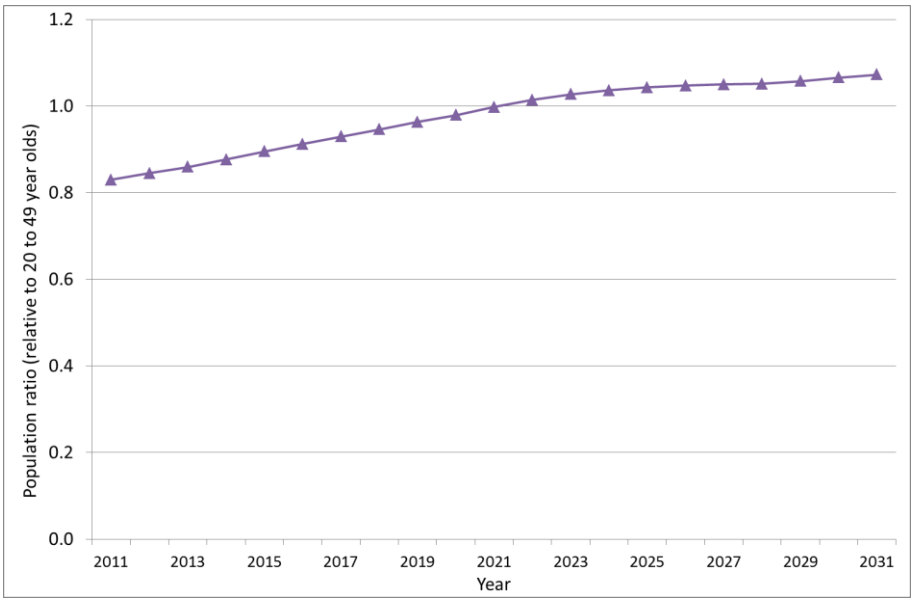


Figure 2-6 : Trend in population ratio, England, 2011 to 2031

(source : 2011 counts are 2011-interim mid-year estimates, ONS, 2012b, and 2012 to 2031 are Principal 2012-based projections, ONS, 2014a)

Just as the number of older people will vary by location as shown in Figure 2-2 so will the population ratio. The provision of funding for elderly health care in the United Kingdom is provided centrally (to avoid perpetuating disadvantage and implementing a 'postcode lottery' – see Barr, et al. (2014), for a discussion of the rationale behind NHS resource allocation) and as such the local population ratio is not relevant to the resourcing of most health care provision. The ratio may, however, be relevant in regard to how the care services needed by an ageing population can be met from within its own catchment population, particularly in terms of labour resources. Figure 2-7 shows how this ratio varies by each local authority in 2011 and the projected case for 2031.

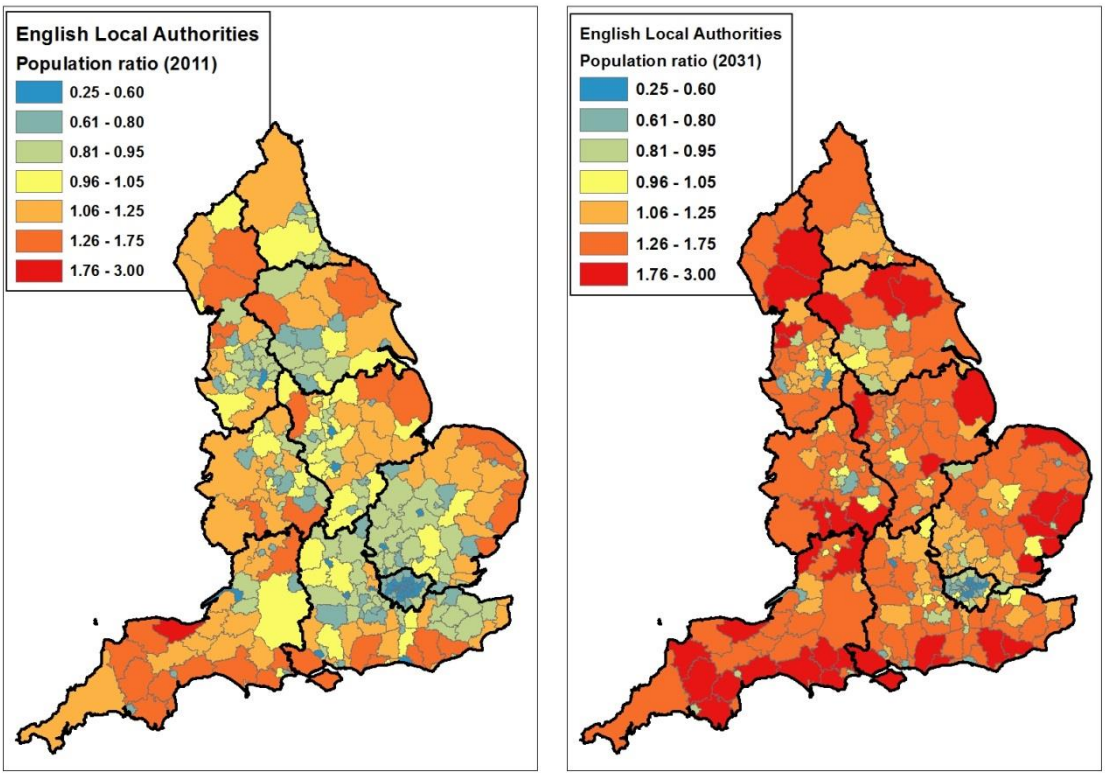


Figure 2-7 : Older (50 or older) to middle age (20-49) population ratio, LADs, 2011 (lhs) and 2031 (rhs)

(source : 2011 counts are 2011 interim sub national projections, ONS, 2012b, and 2031 are Principal 2012-based sub national projections, ONS, 2014a)

The impact of the increasing older population is particularly stark. Whilst only 10 LADs, mainly in coastal locations in the south and east of England, have a population ratio above 1.5 in 2011 (3%), there are 91 such authorities in 2031 (28%). The areas where the ratio remains low are the larger urban areas and inner London. In fact, the urban areas that were seen to have both high older populations (Figure 2-2) and a high absolute growth in the older population (Figure 2-3 lhs) (e.g. West and South Yorkshire and the Greater Manchester area) are seen to have low population ratios in both 2011 and 2031. This suggests that the overall growth in the older population in urban areas is

better balanced with increases in the middle age population. This is further illustrated in Figure 2-6 which plots the population of the LAD against its population ratio in 2011 and 2031. The smaller authorities tend to have the higher population ratios.

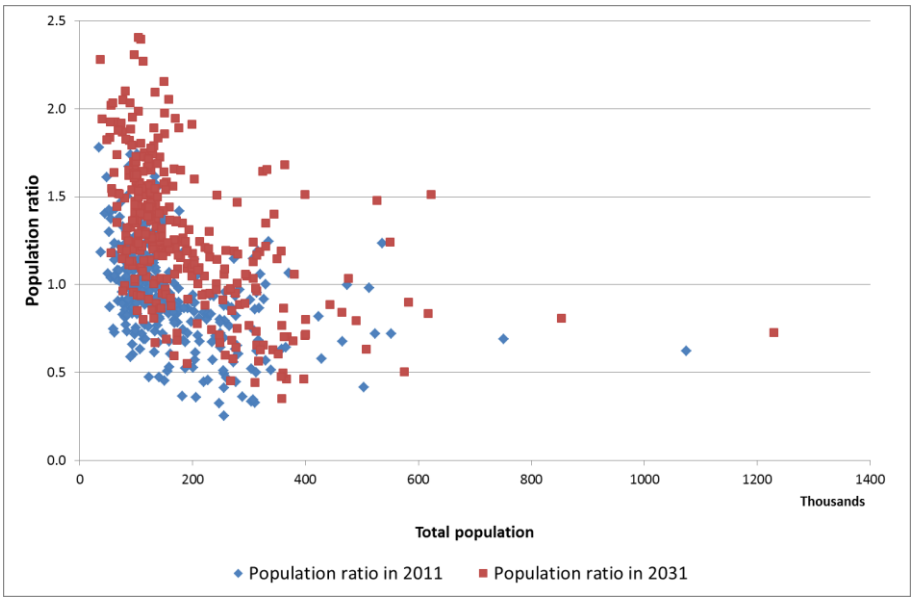


Figure 2-8 : Plot of total population aged 50 and older against the population ratio, LADs, 2011 and 2031

(source : 2011 counts are 2011 interim sub national projections, ONS, 2012b, and 2031 are Principal 2012-based sub national projections, ONS, 2014a)

On this evidence, it is likely that these larger LADs with low population ratios will be more able to provide the health care professionals and care support needed for an ageing population from their own local catchment. The spatial arrangement of this feature is most visible when the ratios are plotted on a cartogram map of England where the size of each area is proportional to its 2001 Census population rather than its geographic area (Social and Spatial Inequalities Research Group, 2004). Figure 2-9 shows the 2011 and 2031 population ratios in Figure 2-7 in this cartogram format. The lower population ratios are clearly more prominently visible in London and the other urban areas of England in this representation. The large visual impact of local authorities with large population ratios is diminished in this cartographic representation.

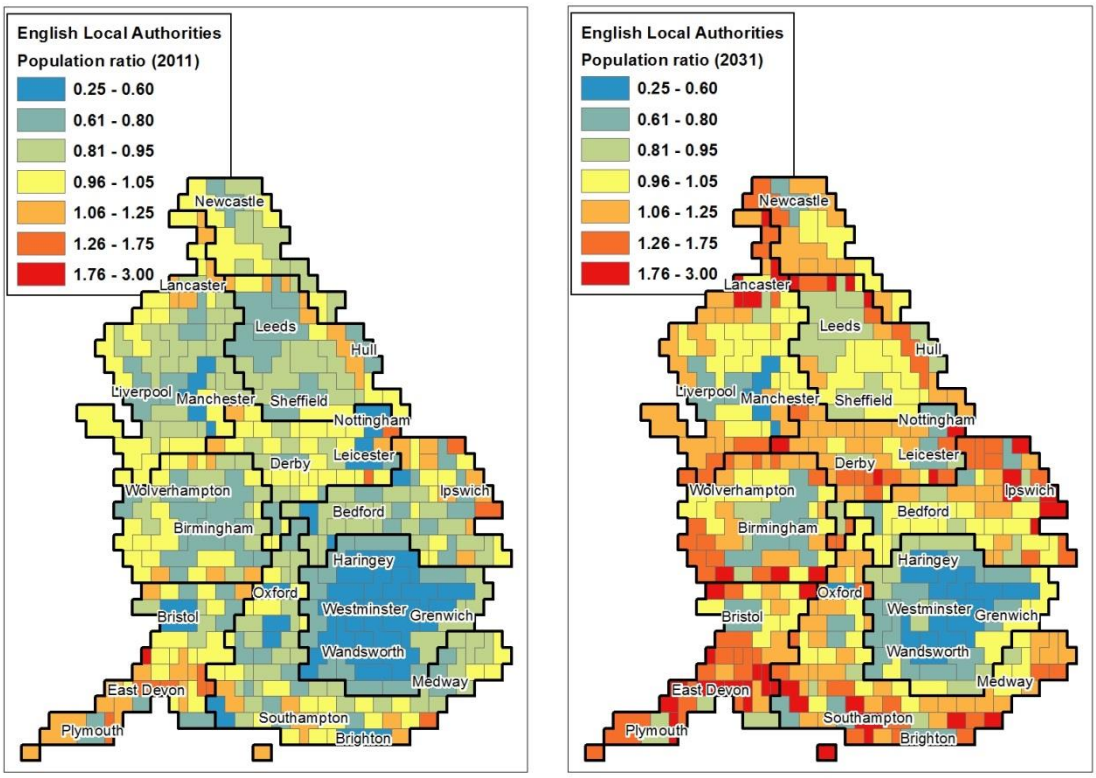


Figure 2-9 : Cartogram of older (50 or older) to middle age (20-49) population ratio, LADs, 2011 (lhs) and 2031 (rhs)

(source : 2011 counts are 2011 interim sub national projections, ONS, 2012b, and 2031 are Principal 2012-based sub national projections, ONS, 2014a)

2.7 VARIANT POPULATION PROJECTIONS

The population projections used so far are just one of many variant population projections produced by the ONS (ONS, 2013b). Each population variant is based on a set of assumptions regarding the three driving components of population change: fertility, life expectancy and migration. Each of these components will have a Principal assumption and an alternative high and low assumption. In addition there are other assumptions: that fertility occurs at a replacement or at a constant (un-changing) level; that there is no improvement in life expectancies and that net migration is zero.

These components are then used to produce a number of variant projections. There is a Principal projection (which is the version used up to now in this chapter); six single variant projections (where only one of the three components at a time are varied to a low or high level); a high and low variant (all components varied at the same time to either the low or high level) and a variant that assumes zero net migration. A further set of special variants assume the special cases with regard to fertility, life expectancies and migration. Table 2-2 specifies which assumptions make up each variant and the resulting total population in 2031 and the old and elderly population in 2031.The range of population projections from these variants is wide. The total population in 2031

could be between 57.7 million and 63.1 million if this collection of variants is sufficient to cover most eventualities. The range of the older population is also wide, being 2.0 million for those aged 50 and older and 1.8 million for those 65 and older. This wide range in the population projections to 2031 will be significant in predicting future health care needs since the size of the population is one of the most important drivers of total health care need. It is not, however, known in a probabilistic sense where these variant outcomes lie in a probability distribution of possible outcomes.

2.8 LIFE AND DISABILITY FREE LIFE EXPECTANCIES

Whilst the size of the older population will undoubtedly have an impact on the need for health care, the health condition of the older population will also have an impact. A large but generally healthy population will have a need of fewer health care resources than one of a similar size but which is poorer in health (Mayhew, 2009). Regular estimates are made by the ONS to establish the life and disability free life expectancy of the population (ONS, 2014b). To calculate the disability free life expectancy, information is obtained from range of household surveys that include an assessment of any long-standing illnesses, disability or infirmity. Recent trends in these ONS expectancies for males and for females are shown in Figure 2-10 – here the age is the age that someone who is aged 65 can expected to live to or be free of disability.

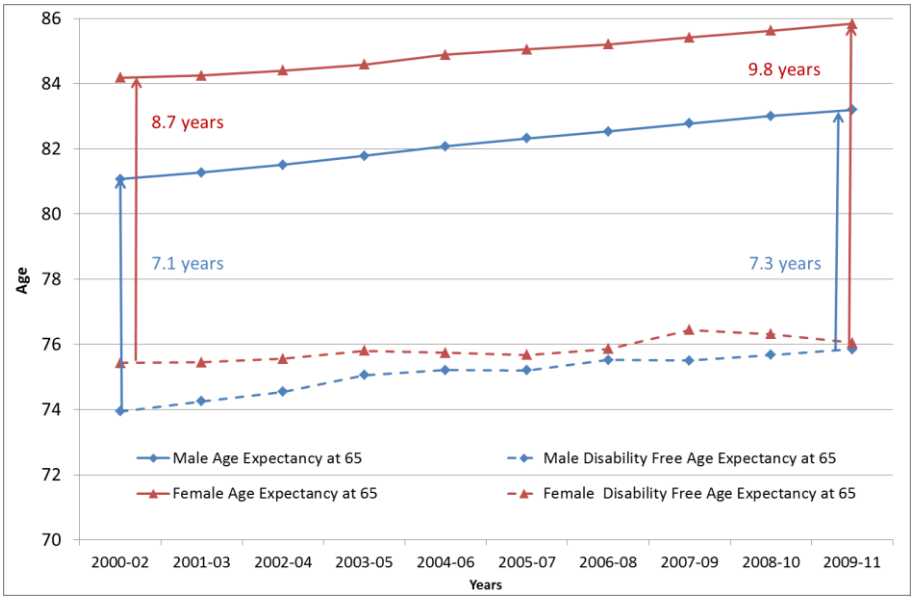


Figure 2-10 : Trends in male and female age and disability free age expectancies at age 65, England, 2000-02 to 2009-11

(source : ONS, 2014b)

Table 2-2 : Summary Statistics from ONS 2012 variant population projections, England, 2031

Variant	Fertility	Life Expectancy	Migration	2031 population projection			2031 population projection (% of Principal)		
				All ages	Aged 50 or older	Aged 65 or older	All ages	Aged 50 or older	Aged 65 or older
No mortality improvement	P	N	P	58,821	22,500	12,130	97%	94%	90%
No change	C	N	P	58,789	22,500	12,130	97%	94%	90%
Low population	−	−	−	57,747	23,528	13,167	96%	98%	97%
Low life expectancy	P	−	P	60,052	23,678	13,212	99%	99%	98%
Young population	+	−	+	62,409	23,832	13,258	103%	99%	98%
Low migration	P	P	−	59,261	23,875	13,485	98%	99%	100%
Zero immigration	P	P	Z	56,721	24,013	13,507	94%	100%	100%
Low fertility	−	P	P	59,248	24,027	13,530	98%	100%	100%
Principal	P	P	P	60,419	24,027	13,530	100%	100%	100%
Constant fertility	C	P	P	60,386	24,027	13,530	100%	100%	100%
Replacement fertility	R	P	P	61,599	24,027	13,530	102%	100%	100%
High fertility	+	P	P	61,584	24,027	13,530	102%	100%	100%
High migration	P	P	+	61,589	24,181	13,576	102%	101%	100%
Older population	−	+	−	58,467	24,215	13,795	97%	101%	102%
High life expectancy	P	+	P	60,775	24,367	13,842	101%	101%	102%
High population	+	+	+	63,135	24,521	13,888	104%	102%	103%
Range				6,414	2,022	1,759	11%	8%	13%

Keys :

Principal projection
Single variant
Combined variant
Special variant

Fertility : P principal; − low; + high; R replacement; C constant
Life expectancies : P principal; − low; + high; N no improvement
Migration : P principal; − low; + high; Z zero (net)

(source : ONS, 2013b)

This figure shows that females have a consistently longer age expectancy than males, although this difference has narrowed over time, from 65 year old females living 3.1 years longer than males in 2000-02, to living longer by just 2.6 years in 2009-11. Trends used in ONS population projections predict that by 2030 males will still have a age expectancy three years fewer than females although Mayhew and Smith (2012) estimate that by 2030, this gender gap in such expectancies at age 30 will almost have vanished. Females also have longer disability free age expectancies than men, but more recently this difference in disability age expectancy between the genders has become less pronounced.

The number of years spent in disability is the difference between these two trends for each gender. This interpretation means that females spend much longer in a disabled condition than males, on average, 9.8 years for females versus 7.3 for men ('*men die, women suffer*', Bruggink, 2011). There is also a trend for this disability gap to increase over time, starting at 7.1 years for men in 2000-02 and growing to 7.3 in 2008-10; and for females, growing from 8.7 years in 2000-02 to 9.8 in 2008-10.

The ONS has also produced a set of experimental statistics that provide both life and disability free life expectancies during 2010-12 (ONS, 2014c) for health care commissioning consortia termed Clinical Commissioning Groups (CCG) (Boyle, 2011). Figure 2-11 shows these expectancies, again as ages, for someone who has reached the age of 50 across the 211 CCGs in England, sorted by age expectancy (hence the smoothness of this series in the figure). The differences in the two expectancies are variable across the CCGs groups. Fifty year old Males living in South Reading and in Knowsley, Merseyside, can both expect to live to 79½ years, however in South Reading a 50 year old man can expect to live to age 68½ years without a disability but this figure is over four years lower in Knowsley, at only 64 years. For females there are similar disparities, the East Riding of Yorkshire and Newham, both have a similar female age expectancy of just greater than 84 years but different disability free Age expectancies, 70 years in the East Riding but only 64½ years in Newham. These age expectancies can also be mapped to examine if there are any geographical patterns. Figure 2-12 displays these two age expectancies for males and for females.

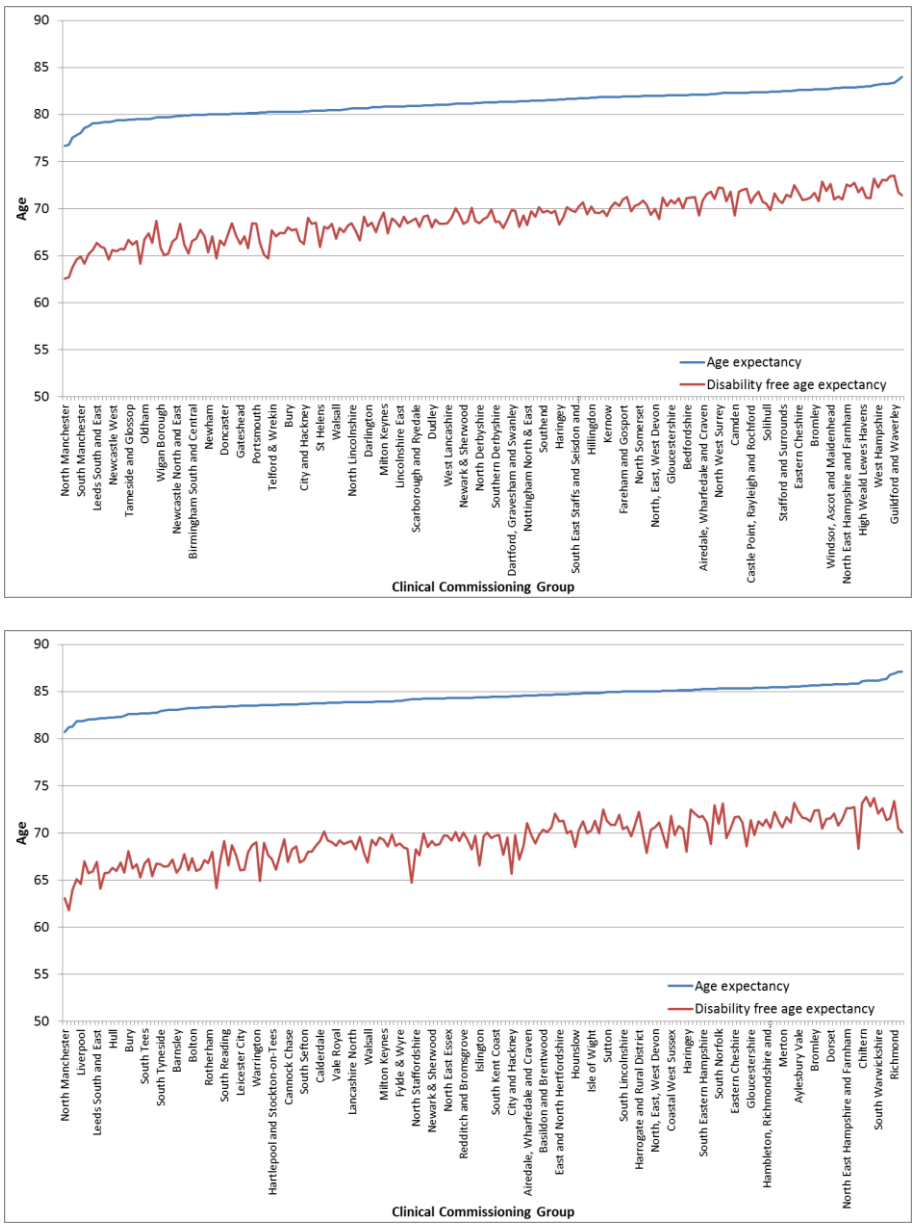


Figure 2-11 : Life and disability free age expectancies at aged 50 years for male (top) and female (bottom), Clinical Commissioning Groups, 2009-11

(source : ONS, 2014c)

As shown by Figure 2-12, top row, in all CCGs the age expectancies of females aged 50 exceed that of males. Both males and females have lower age expectancies in the areas around Durham and Newcastle, Lancashire and parts of Manchester, South and West Yorkshire and a band through the East Midlands and into the West Midlands. These areas were all once heavily industrialised areas with a large proportion of the workforce engaged in manual trades such as textiles, coal mining and steel making. Working in these occupations, particularly the work undertaken by males, would contribute to a range of disabling medical conditions that would tend to lead to lower life and disability free life expectancies (Johnson, 2011).

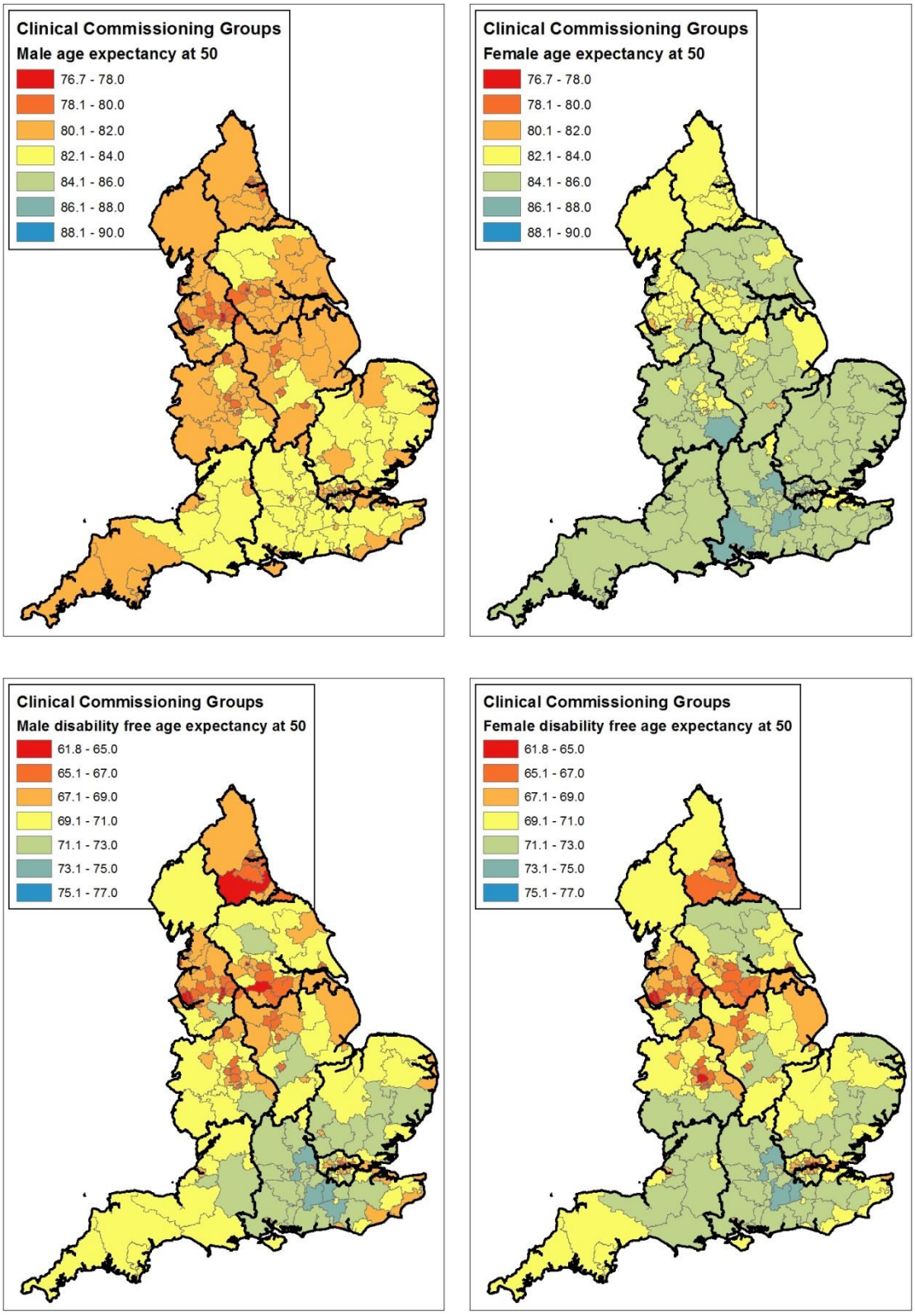


Figure 2-12 : Age expectancies at aged 50, CCGs, 2010-12, clockwise from top right : male life expectancy; female life expectancy; female disability free life expectancy and male disability free life expectancy,

(source : ONS, 2014c)

Looking at the geographic spread of disability free age expectancies the differences between the genders is once again, as Figure 2-10 shows, less pronounced. What is most striking about the disability free age expectancies shown in Figure 2-12 is the greater variation between locations and this variation is much more pronounced than for the related age expectancies. Once again the previously industrialised areas have much lower disability free age expectancies and the more rural, Home Counties and South West authorities which have disability free life expectancies not that much lower than their already elevated life expectancies.

By subtracting the number of years of disability free living from the life expectancy it is possible to measure the number of years in a disabled condition. The significance of this measure is that it is likely that individuals with a disabling condition will require greater use of social and medical services. How this measure varies by CCG is shown in Figure 2-13. In the left hand side map, the pattern of increased male years spent living with a disability is apparent in the more industrialised areas of northern England. In other areas, males die at older ages, but they do not become disabled until these later years. The contrast that is perhaps most insightful here is that between the genders, where once again this shows that females spend many more years with a disabling condition than males do, and also that there is a weaker of pattern in the geographic variation for this measure for females.

2.9 HEALTH CARE

As mentioned at the beginning of this chapter, one of the most important impacts of an ageing population will be felt on health care demand and delivery. The United Kingdom possess a well-resourced and integrated national health service, whose activities range from community based general practitioner teams through local clinics and hospitals and onto world renowned treatment centres. This high quality and integrated approach to health care provision puts the NHS in a good position to deliver and co-ordinate health care in an efficient manner (Karlsson et al 2007).

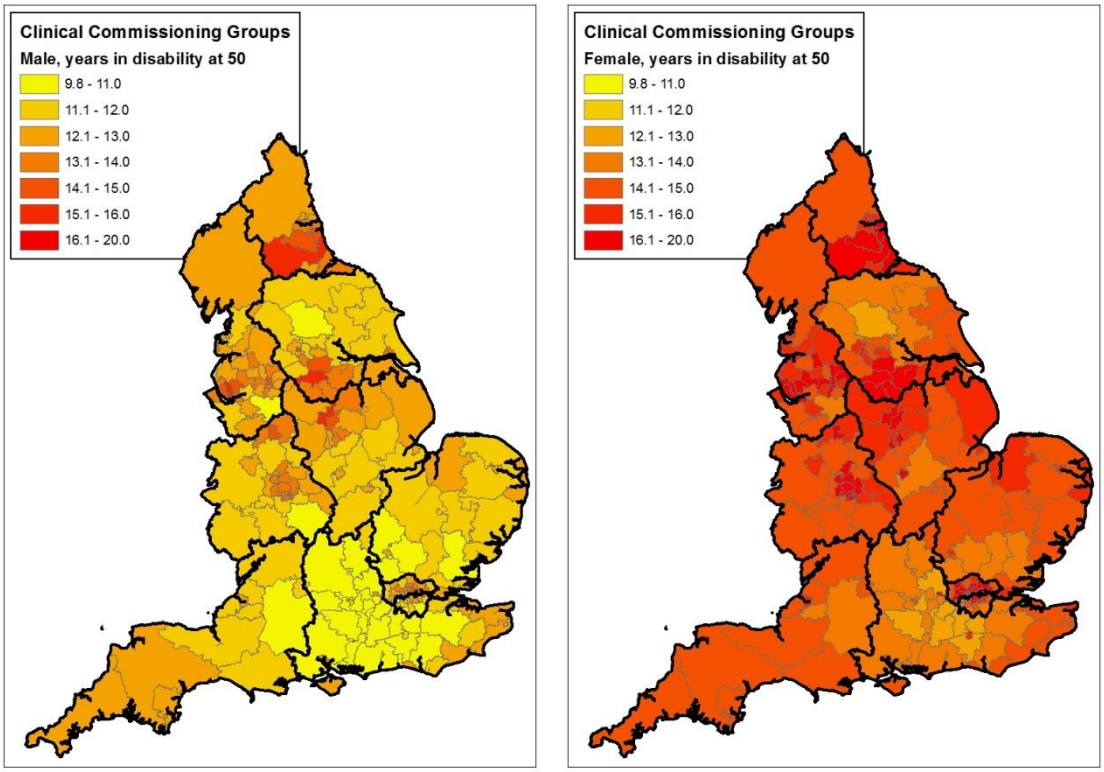


Figure 2-13 : Years spent in a disabled condition for male (lhs) and female (rhs), by CCG, 2010-12

(source : ONS, 2014c)

2.9.1 Trends in geriatric provision

Figure 2-14 shows the recent trend in the number of geriatric in-patient beds available within the NHS in England.

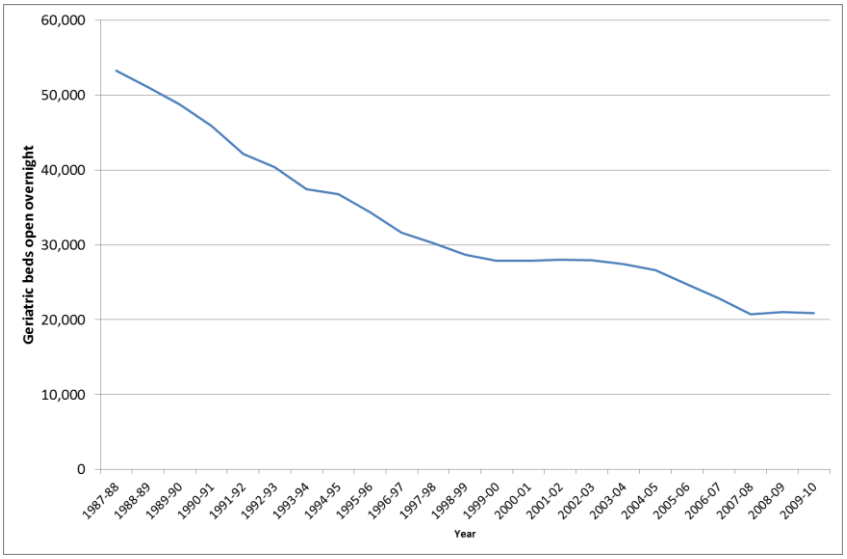


Figure 2-14 : The number of geriatric beds available overnight, England, 1987/88 to 2009/10

(source : Harker, 2009)

In the decade from 1987/88, the number of geriatric beds has more than halved in English NHS hospitals, but the trend in more recent times has been for this number to stabilise at around 20,000 beds. This reduction cannot, however, be assumed to be a reflection of a reduced need for elderly hospital care. A recent trend in many medical treatments has been a move towards day surgery cases, where an overnight stay in hospital is not necessary. These patients are usually accommodated in day beds and the use of such day beds by the elderly is not reflected in these figures (over this time period the number of NHS day beds has increased from 2,000 per day to nearly 11,000).

2.9.2 Geographic provision

Geographically referenced information on the demand for health care is made available by the HSCIC, through commissioning statistics and, separately, the delivery of health care, through provider statistics (HSCIC, 2015).

The primary measure of inpatient health care use is the inpatient Finished Consultant Episodes (FCEs) which are derived from the Hospital Episode Statistics (HES) (Cowper, 2004). An FCE consolidates together a series of related treatments that an individual may receive over a short to medium period of time, all whilst under the direction of the same consultant. Thus each FCE may be regarded as a series of medical interventions over a short term, but all related to the same illness.

2.9.3 Commissioner inpatient data

Looking at the commissioning of health care, Figure 2-15 shows the number of commissioned FCEs by CCG during 2012-13 for the 50 to 65 (lhs) and the 65 year and older age bands (rhs), both counts divided by the ONS's estimate of the CCGs 2012 mid-year age band population. This provides a measure of the geographic demand for health care, relative to its population - the higher the measure the greater the demand for per capita health care.

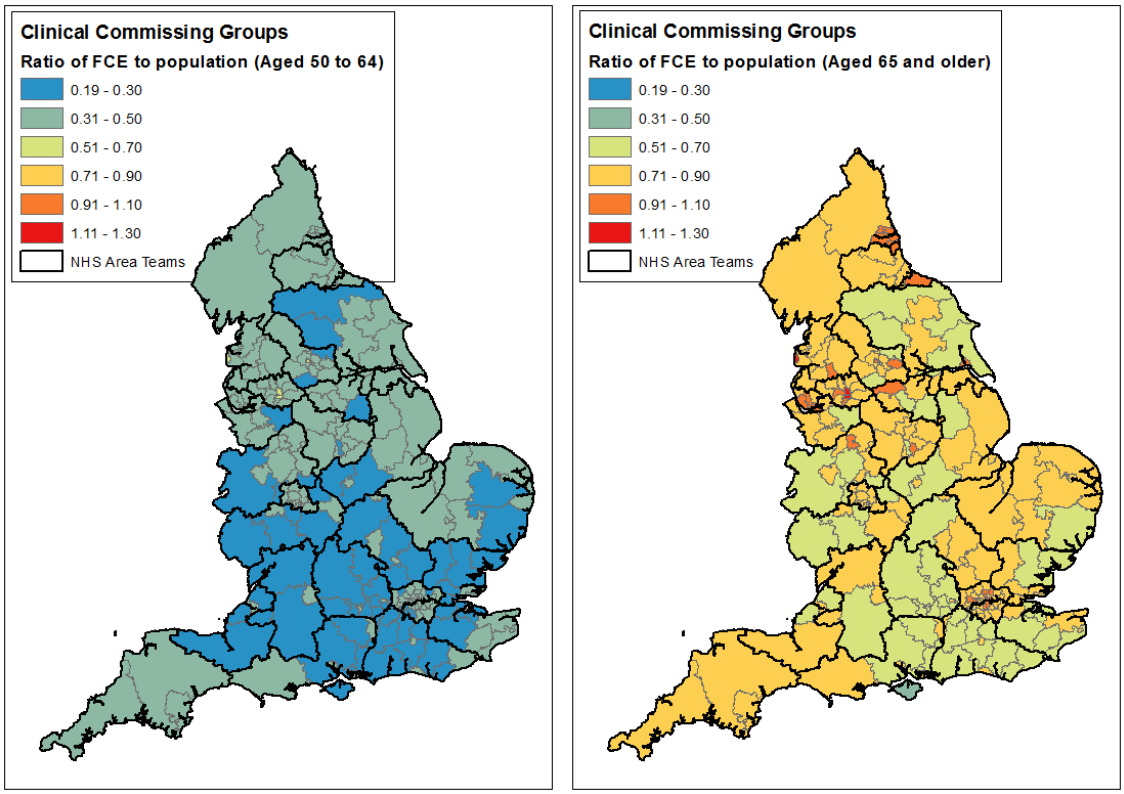


Figure 2-15 : Ratio of commissioned inpatient FCEs per capita for aged 50 to 64 (lhs) and aged 65 or older (rhs), Clinical Commissioning Groups, 2012

(source : HSCIC, 2015, and ONS, 2012b)

For the 50 to 64 age band in only four CCGs is the rate above 0.5 FCEs per head of population (Manchester North and South, Bradford City and Blackpool). More generally, CCGs where the number of inpatient FCEs per person are low tend to be in the Home Counties surrounding Greater London, along the Welsh border and in North Yorkshire. This picture is not dis-similar to that shown in Figure 2-13 which shows a measure of duration of disability in each CCG by gender. For the oldest age band, the number of FCEs per capita is higher, which is to be expected because of their greater need for medical treatment and the smaller population size of this age band. In this age band, the urban centres of northern England tend to have the highest ratios, whilst the Isle of White has a particularly low rate, at 0.4913 FCEs per person. Areas in southern England tend to have lower FCEs demand.

There may not be an exact match between the local need (as quantified in Figure 2-13) and the number of treatments required (as quantified in Figure 2-15). This most often manifests as unmet demand where the capacity to deliver health care services in an area is less than the demand.

2.9.4 Provider inpatient data

As well as providing information on the commissioning of health care provision, the HSCIC also provides information on where the FCE's are delivered, this is usually in a hospital or clinic context (HSCIC, 2015). Figure 2-16 shows the number of inpatient FCEs provided at hospitals and clinics in England for the period 2012-2013 for the 60 to 74 year age band (lhs) and the 75 and older age band (rhs). These figures show where the FCE's are delivered for patients. As would be expected they tend to cluster strongly in cities and towns where the facilities are located (insert maps provide greater geographic detail for London).

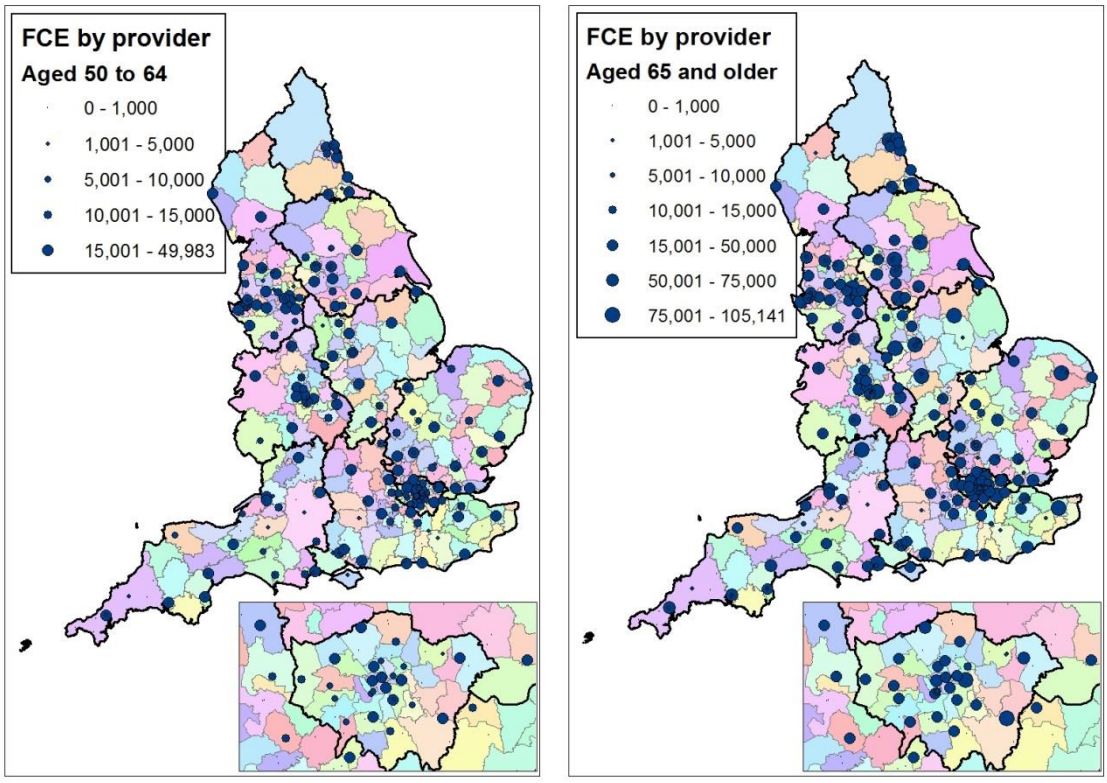


Figure 2-16 : Location and extent of inpatient FCEs provision for aged 50 to 64 (lhs) and aged 65 or older (rhs), England, 2012/13

(source : HSCIC, 2015)

Note : The shading of LADs is for discrimination and identification purposes only and has no meaning attached.

First inspection shows that there is actually very little difference in the relative volume and the spatial distribution of FCEs from these two age groups. There are large clusters of providers in the urban centres, particularly in Greater Manchester and London. It is, however, not easy to scale this measure by population since the catchment of each provider is not well defined.

2.9.5 Provider outpatient data

In a similar manner to the inpatient FCEs, the HSCIC also provides information on the location and volume of outpatient appointments. The location of these appointments for the 50 years and older age group in 2012-2013 is shown in Figure 2-17.

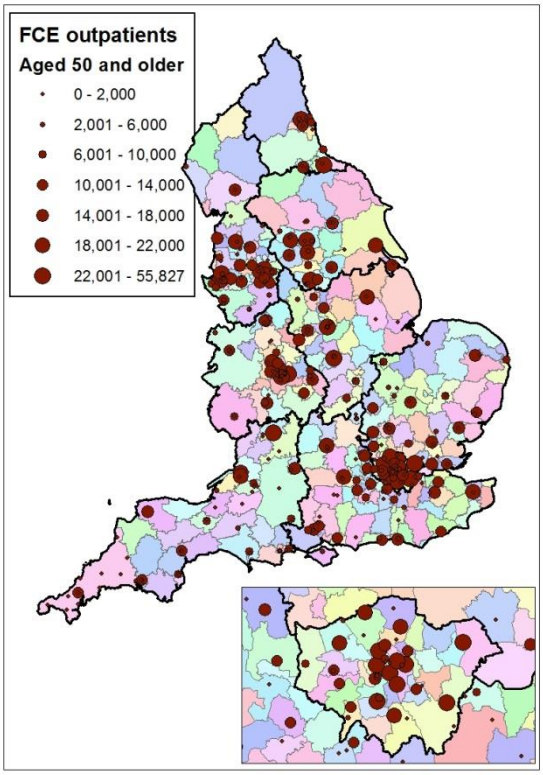


Figure 2-17 : Location of outpatient appointment provision for aged 50 and older, England, 2012/13

(source : HSCIC, 2015)

Note : The shading of LADs is for discrimination and identification purposes only and has no meaning attached.

Once again, many of the providers of outpatient care are concentrated in the urban centres, as was the case for the inpatient care. This is not a surprising finding since many of the providers will provide both inpatient and outpatient services.

2.10 CONCLUSION

Looking at the range of official statistics presented in the chapter a number of themes have emerged.

Whilst there is some uncertainty about the exact numbers, there is a widely held consensus that population ageing within England is inevitable. Modest increases in the mean age of the population as a whole mask a disproportionate large growth in the older population aged 50 or older. Not only is this growth concentrated in a particular age band but also the growth is not uniform across England. Larger local authorities,

determined by high population density or large geographic size, will have the largest growth in the number of older people, although for the dense urban areas the population ratio (defined in equation 2-1) will remain fairly constant, as these large centres attract young adults and experience an out-migration of retirees.

Recent trends in the life and disability free life expectancies suggest two features. Firstly, that the difference in both of these life expectancies between males and females is narrowing and secondly, that the gap between the two expectancies is growing, for both genders. This latter feature indicates that individuals will spend more years with a disabling condition. People who live in the traditionally industrialised cities of the north and midlands tend to have lower life and disability free life expectancies than people who live in the south and east of England. When considering time spent in disability these regional differences are still present, and females tend to spend more years in a disabled condition than males.

Both the demand and delivery of health care for the old tend to be concentrated in the urban centres of England, even when (in the case of demand) accounting for the size of the aged population of the area. This concentration is for historic reasons, since many of the providers are long established in these locations, and given their situation, they have a large catchment population.

Considering how these features will impact on future health care provision highlights a number of likely trends:

1. By 2031, there will be a larger older population spending a greater length of time with disabling conditions;
2. Areas that already have a large older population in 2011 will see the largest increases in the older population by 2031, however, relative to their size, these increases will be more modest than in smaller authorities;
3. The population ratio will remain stable in the larger urban areas, which will facilitate the availability of formal and informal care support in such areas;
4. The differences in health care use by gender will narrow over time;
5. Provision of health care is currently concentrated in urban centres and this is likely to be the case for the foreseeable future;

This chapter has begun to meet the first objective set out in section 1.5, to understand the challenges associated with an ageing population. This is further enhanced in the next chapter which provides a review of the published literature in the area of ageing and its impact on health. The review also includes material on the modelling methodologies that have been adopted to look at the issues of an ageing population and their health care needs, which also goes towards meeting the first objective.

3 LITERATURE REVIEW

What is known about Elderly Health and Health Care

3.1 INTRODUCTION

Chapter 2 of this thesis presents a picture of the demographic characteristics for a future ageing population and a flavour of its current health care delivery. As discussed in section 2.2, this ageing will have profound consequences in several areas. For example, the will be financial pressures associated with pension provision and the ability of the workforce to support an ageing population. One sector in particular that these pressures will be most keenly felt is in the health care sector. To aid in an understanding of this issue, this chapter provides a consideration of the existing literature on the determinants of the health needs of an ageing population and on the delivery of health care to an ageing population.

Health care is a multi-faceted phenomenon. The definitions used can be inconsistent and conflicting – arising from the diversity of the backgrounds of the researchers, their geographic context or the period during which the work was conducted. Within the scope of this work, constant definitions are used: *health care* is used to refer to all aspects of the delivery of health care, be it inpatient; outpatient, nursing, residential, pharmaceutical or other services. The term *primary care* is used to denote care that takes place within a general practitioner setting, which, in the UK context, is an individual's usual first point of contact with the health care system. Similarly, *secondary care* is used to refer to those services traditionally provided either in a hospital or a clinic setting, usually after referral by a general practitioner and can range from consultations with specialists through to participating in leading edge treatments. *Residential care* refers to care delivered in a residential setting (e.g. the persons own home, a nursing home or a residential care home). Residential care may be further sub divided into *institutional* (e.g. in a nursing or residential home) or *home based* (if provided in the person's own home). This home based care can be either *formal* (provided by an outside agency, usually at a cost) or *informal* (provided by a family member or friend). Occasionally there may be references to *other costs*, such as prescription drugs and optical or dental costs, and where necessary these will be explicitly mentioned.

Increasingly in the literature there is recognition that the health care needs of the elderly population are not homogeneous with respect to the important determinant of age. Many studies differentiate between the *young-old* and the *old-old* population (in a similar

manner, Baltes and Smith 2003, highlight the contrasts between the *third* and *fourth* ages of the life course). The exact distinction between these two populations is difficult to define and is dynamic over time, but unless otherwise stated, the assumption here is that the young-old constitute those aged between 65 and 74 years, whilst the old-old are aged 85 or more. Those aged 75 to 84 constitute the *middle-old*.

The first substantive section of this chapter, section 3.2, reviews the concepts of life expectancies and healthy life expectancies, concepts which help to determine both the size of the elderly population and its health status. Section 3.3 considers studies that have attempted to model the demand for and cost of health care in an elderly population. The factors reported to be significant in this relationship are considered in turn. These two sections inform the choice of constraint variables for the spatial microsimulation (Chapter 6) and the probability modelling (Chapter 8). This section is followed by section 3.4 which examines the characteristics of the three case study morbidities used in this thesis, CVD, DHBS and respiratory illnesses. Section 3.5 contains a discussion of the range of health data that have been employed in the literature (in Chapter 4 two data sets arising from this consideration, and of particular value to this thesis, are described in more detail). The review then moves on to a consideration of the types of models that have been used to understand and predict future health care needs (section 3.6). This will inform the choice of methodology adopted in the thesis which is described in detail in Chapter 5. Since the future size and composition of the elderly population will be an important factor in determining the amount of health care required, the following section, 3.7, examines the range of English population projections. One of these projections and a mechanism to revise its projected population is covered in Chapter 7. The penultimate section, 3.8, looks at studies that have considered how to meet the demand for health care once a need has been established. Finally section 3.9 provides a summary of these review findings.

3.2 LIFE EXPECTANCIES

Many studies use the well understood concept of life expectancy to explore the impacts of an ageing population on society (Mayhew, 2005; Mayhew, 2009, and Christensen et al., 2009). The general trend in recent years has been for consistent increases in life expectancy, however, there is some debate as to whether this trend will continue into the far future. This continuation is proposed by Oeppen and Vaupel (2002) and Wilmoth et al. (2000) whilst others propose that at some stage a maximum life expectancy will be reached, either through a biological mechanism or societal changes (see Olshansky et al., 2005, where the hypothesis is that increases in obesity at younger ages may lead to shorter life expectancies).

When life expectancy is coupled with the concept of healthy life expectancy, it is possible to begin to examine future health care needs. If healthy life expectancy increases by the same amount as life expectancy then individuals will roughly spend a constant amount of time in an unhealthy condition (termed *morbidity postponement*, where ill health is moved to later ages, Payne et al., 2007). If, however, the life expectancy increases at a faster rate than healthy life expectancy then individuals can expect to spend more years in an unhealthy condition (termed *morbidity expansion*, Gruenberg, 1977 and Kramer, 1980). Conversely if healthy life expectancy increases at a faster rate than life expectancy then the more years will be spent in a healthy condition (*morbidity compression,* Fries, 1980). A fourth possibility is beginning to emerge in the literature, *morbidity equilibrium* (Manton, 1982), where the expectation is that more years will be spent in a disabled condition but that the severity of this disability will be more varied and dynamic.

A complication arises if these changes are measured not as an absolute number of years in morbidity but are measured relative to either total life expectancy or remaining years of life (Hoffmann and Nachtmann, 2010). Thus there is the concept of *relative expansion* or *relative compression* of morbidity.

3.2.1 Life expectancy

The Office for National Statistics (ONS) regularly produces estimates of life expectancies in the UK and Shaw (2007) examined the accuracy of the official national life expectancy projections over the last 50 years. Figure 5 from Shaw (2007) shows life expectancy estimates from successive reports from the ONS between 1971 and 2004 for males and females (here updated to cover projections in the period 1994 to 2012 and shown as Figure 3-1). In both cases, over time there is a consistent pattern of increases in life expectancy from one estimate to the next. This work highlights the fact that improvements in mortality, particularly at older ages, have been consistently underestimated in the UK and would point to adopting more optimistic assumptions with regard to future to life expectancies. This suggestion is reinforced in work by Mayhew and Smith (2012) and Bennett et al. (2015) who provide modelling evidence that there will be a convergence in the life expectancies between the genders over the next 20 years to an extent greater than that envisaged by ONS.

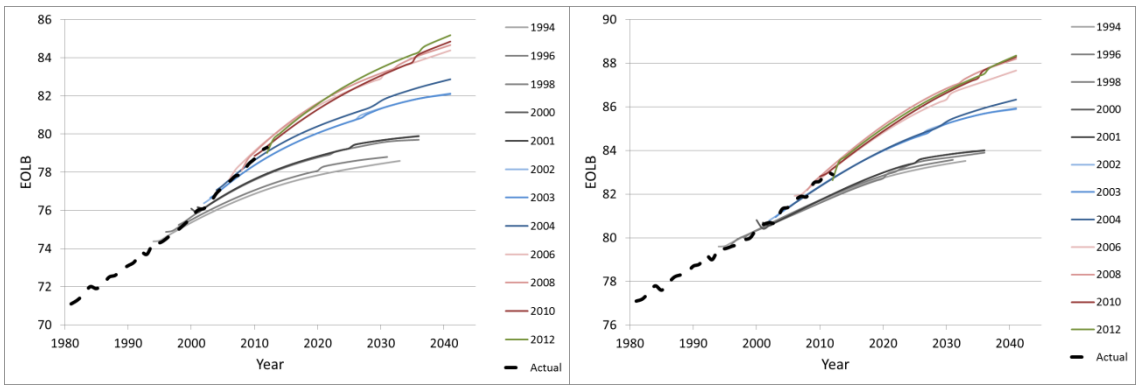


Figure 3-1 : Historical revisions to expected years of life at birth (EOLB), males (lhs) and females (rhs), England, 1994 to 2012 projections.

(after : Shaw, 2007)

Specific conditions may, however, limit the increases in life expectancy (Hoffmann and Nachtmann, 2010, cite obesity and an 'impending' global influenza pandemic as possibilities). Lubitz et al., 2003, predicted how life expectancy at 70 years of age was affected by the ability of the individual to perform certain tasks; whether they were cared for in an institutional care setting and what their self-reported health status was. They found that the life expectancy of an individual with good functional status was greater than for an individual with poor functional status, but the total cost of health care beyond age 70 was similar across all function groups, with the exception that the cost for those in institutional care was seen to be the largest, although they had the lowest life expectancy.

3.2.2 Healthy life expectancy

The calculations of healthy life expectancy are subject to some sampling variability since they use the results of survey data which sometimes makes absolute conclusions with regards to past trends difficult to draw (Parliamentary Office for Science and Technology, 2006). Caley and Sidhu (2010) claim that the trend in the UK over the 25 years to 2007 is for an absolute morbidity expansion, with life expectancy at birth increasing by 5.05 years but disabled free life expectancy increasing by only 3.64 years. ONS (2008), in consideration of similar data covering the period 2000/2 to 2007/9 implies that there has been a relative compression of morbidity for both males and females at age 65. Clark et al. (2004) asserts that in the US the trend is for morbidity compression, "*This trend now suggests an improvement in functioning and mobility of older Americans*." (p275 and comments on p323). Murphy and Martikainen (2010), report that in their opinion "*more experts expect a reduction in the proportion (although not the absolute numbers) of people with poor health*". This relative compression is consistent with the scenario that the number of years in disability may be increasing (Caley and Sidhu, 2010) but the proportion of life in disability is

decreasing (ONS, 2008). Hoffmann and Nachtmann (2010) review and present their own evidence from Germany and whilst acknowledging some contradictory results, suggest that the morbidity trend is "*positive*".

The study by Lubitz et al. (2003) estimates the proportion of remaining life spent in poor/fair; good and very good/excellent health based on an individual's self-reported health status at the age of 70. This research shows that better health leads not only to greater life expectancy but also higher quality of life. Individuals who report their health as being excellent spent just 2.7 of their remaining 13.8 years in either a fair or poor state of health. For those whose reported health was poor however, over half of their remaining 10 years of life were spent in poor or fair health.

A consideration of these studies would suggest that the two most likely outcomes in the future are for a morbidity expansion with either a relative expansion or a relative compression. In their report to the Wanless Social Care Review on the future cost of residential elderly health care, Jagger et al. (2006), state that "*It thus seems unlikely that a compression of disability will occur unless the severity of disability associated with diseases diminishes*." (page vi). For this thesis it is likely that the trends measured as absolute rather than relative morbidity will be the more relevant since the amount of actual time in a disabled condition will more accurately determine the health care needs. This interpretation is in line with that used in the study of future pension provision by Harper, Howse and Baxter (2011) "... *historical data for the UK suggests that for both men and women the increases in 'healthy life expectancy' (HLE), and 'disability free life expectancy' (DFLE) in particular, have not kept pace with total gains in life expectancy. This is important as both of these measures provide an indication of the length of time an individual remains 'healthy'* ..." (pages 8-9).

3.2.3 Summary

The expectation for the medium term over the next 20 to 30 years is that life expectancies in the UK will continue to increase at least at the same rate as seen in recent years. Whilst there have also been increases in healthy and disability free life expectancies, these increases have been at a slower rate than life expectancies; this means that individuals are likely to spend more time in an unhealthy or disabled condition.

3.3 ELDERLY HEALTH CARE DETERMINANTS

In this section, a wide range of factors that influence an individual's use of health care resources will be considered. These are related to the characteristics of the individual themselves, such as age and ethnicity, but may also relate to their employment history, lifestyle or environment in which they live.

Many of the studies reported below measure the current and future need for health care for the elderly population in terms of costs rather than direct measures such as the number of procedures carried out, hospital admissions or bed days. Care is however needed when trying to equate increasing costs with poorer health status (Westerhout, 2014). Certainly as the health of a population deteriorates, the total costs associated with treatment of these illnesses would be expected to increase. However there may be other factors that drive up health care costs. Firstly, there is the rising cost of treatments, as medicines become more sophisticated and as life spans increase, longer term use of such treatments will drive up costs. Secondly, the 'over treatment' by doctors practicing 'defensive medicine' who, either yield to a desire from the patients to receive some form of treatment or, particular in a United States context, wish to minimise potential mal practice actions (Garcia-Retamero and Galesic, 2014). Thirdly, the medical profession and its institutions are a quasi-monopoly, able to inflate their remuneration, even in the context of stable workloads. Finally, there may be incentives in the system that reward treating illness rather than maintaining good health, e.g. the Quality Outcomes Framework in the UK which financially rewards aspects such as diagnosis and treatments (HSCIC, 2014b) which may shift the judgement boundary towards the diagnosis of an illness.

3.3.1 Age since birth

One of the most often cited factors in determining an individual's health care need is age. A study by Alemayehu and Warner (2004) apportions whole life health care costs by age and gender. Their results show that just over half the total lifetime health care expenditure for an individual took place by the age of 65. Between 65 and 84 over a third of the remaining expenditure occurred, with 12% occurring after the age of 84. This may appear at odds with a study by Seshamani and Gray (2004a) which indicates that increased longevity after the age of 65 does not cause significantly additional hospital care costs over time. However, the Alemayehu and Warner (2004) study found that for the elderly population, a significant proportion of the high costs was attributable to Institutional care rather than hospital care, and even more so for the old-old. This phenomenon is also found in a study by Spillman and Lubitz (2000) who considered the cost of health care provision based on the age of the individual at death. Unsurprisingly the longer someone lives, the higher their cumulative health costs, although for someone who dies at 90 the majority of their costs are from sources such as residential care, prescription drugs and other services rather than hospital care. Further corroboration is found in Murphy and Martikainen (2010) who analyse the number of bed days in hospital and residential care in Finland and show that generally, up until the age of 85, more bed days are spent in hospital care whilst at subsequent ages the greater number of bed days are spent in residential care.

After some methodological modifications by Seshamani and Gray (2004a) to work by Zweifel et al. (1999) and using a two-part estimation model of the likelihood to incur costs (i.e. be admitted into hospital) and to estimate the costs once in hospital, they found that a non-linear function of age was a significant factor (as was time to death). The cost shape was concave down, peaking in the 85-89 age band, indicating an escalation of costs up to this age band, followed by reductions at latter ages.

These studies suggest that age is a significant factor in determining health care costs (although see below for counter arguments) and that as individuals age these health care costs increase. However, the mix of costs varies, with institutional care taking a larger proportion of costs at old-old ages, even to the extent that hospital care would decrease amongst the old-old.

3.3.2 Time to death

A second factor that has been often quoted as a significant determinant of health care costs is how close the individual is to dying. Using linked data from NHS care and local authority-funded social care in three primary care trusts in England, Bardsley et al. (2010) were able to identify patterns in hospital and residential care use by over 16,000 individuals in the period up to their death. The pattern found was for increasing costs during this period - for hospital care this increase began in the final two months of life whilst for residential care costs, it began in the final 12 months. They also noted changes in the mix of the two types of provision by age at death and that there was some variation in patterns of use between the three trusts, with one trust having a distinctly higher pattern of hospital care usage than the other two.

Whilst the modelling studies reported above have identified age as a significant factor in health costs, other work has suggested that age is acting merely as a proxy for some other feature. Zweifel et al. (1999) found that for the old population, once time to death (measured in quarter years to death) was used in the modelling framework, age specific regression terms were no longer significant in explaining hospital costs. This is also generally true when the components of health care are broken down (Werblow et al., 2007). Spillman and Lubitz (2000) found that health care costs in the final two years of life did not vary by age at death; however beyond the age of 90 at death, it was more likely that these costs would be from institutional care rather than hospital care. Seshamani and Gray (2004b) estimated that the 5% of patients aged 65 and over who were in their last year of life accounted for approximately half of the hospital care expenditure for all those in this age band. A similar result was found by Clark et al. (2004) studying the US, who quote that 28% of total acute hospital and other care costs were incurred by the 5% of state Medicare members who were in the last year of their life (p 276). In their study Alemayehu and Warner (2004) estimated remaining health

care costs for the surviving population and the total cohort population (which includes both survivors and those who will die within the year). The remaining costs for survivors are seen to be lower than the total population, and by significant amounts for the old (75%) and the old-old (45%). Conversely in the age vs. time to death debate, Atella and Conti (2013) examined the relative impacts of age and time to death for outpatient costs in Italy and found that age had a bigger impact on costs (a 500% increase from age 40 to 80) than time to death (a 30% increase in the last four years of life).

The importance of any potential time to death effect, where death and hence health care use may shift to later ages, can be illustrated by using data presented by McCulloch (2012a) on population ageing and mortality. This article examines the trends in age at death in England and Wales between 1968 and 2010 and presents two age at death population diagrams (reproduced below as Figure 3-2) which show how the age at death has shifted significantly to the older age bands, especially so for males. Whilst historically the majority of males would experience their final year of life sometime between the ages of 65 and 80, in recent times this final year of death is more likely to occur after the age of 75.

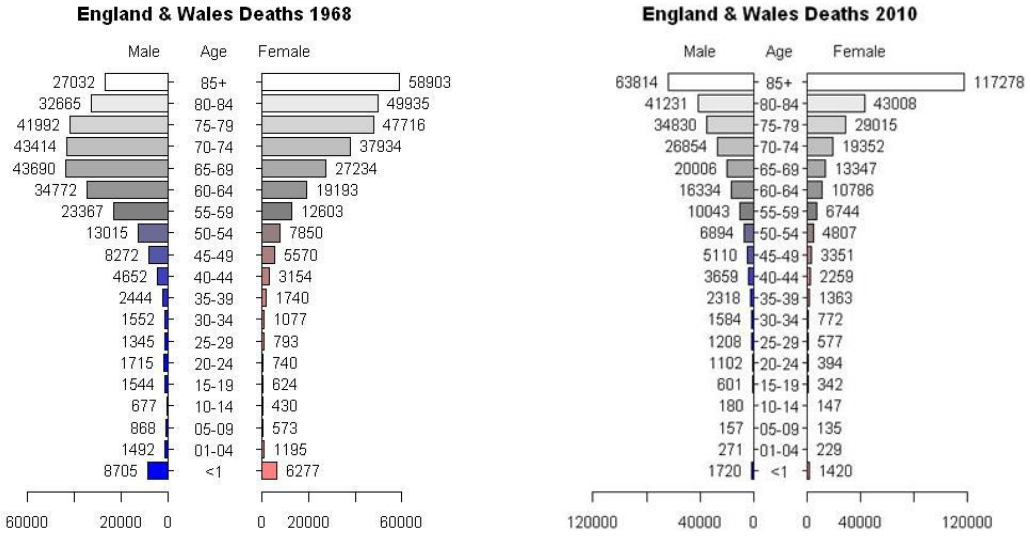


Figure 3-2 : Population pyramid showing age distribution of mortality, England and Wales, 1968 (lhs) and 2010 (rhs)

(source : McCulloch, 2012a)

Whilst the studies reported above have highlighted the significance of time to death rather than just age in determining health care costs, de Meijer et al. (2011) in a Dutch study of the cost of long term residential care (and separately by institutional and formal home based care) found that for formal home based care the time to death for an

individual became less significant in determining costs, after controlling for age, co-habitation status, cause of death and most interestingly, disability.

There appears to be an acceptance in the literature that there is some overlapping role to be played by age, time to death and disability in determining an individual's health status. Incorporating time to death in a model is seen to impact on the age specific effect, and the time to death effect is itself weakened when a measure of disability status is included. In the absence of time to death information, Dormont et al. (2006) advocate the use of a Death Risk and chronic illness indicator terms.

3.3.3 Gender

It is well recognised that females tend to have longer life expectancies than their male counterparts. What is also beginning to emerge is that the healthy life expectancy for females is, however, not that much different than that of males. This indicates that females spend a greater number of years of their later life in poor health.

The study of functional status and self-reported health at age 70 by Lubitz et al. (2003) examined differences between the genders. Their work shows that females had a longer life expectancy at age 70 than males (14.3 years vs. 11.8) but spend fewer of these years in an active state (6.5 years vs. 7.4). This is supported by the earlier finding by Spillman and Lubitz (2000) that cumulative health care expenditure at death, controlling for greater longevity amongst females, is between 15 and 27 percent higher for females than males. In a study of use of general and special physician services in Switzerland, Schellhorn et al. (2000) estimated that females made 15% more visits to medical practitioners. A consideration of the use of bed days in Murphy and Martikainen (2010) showed that, on average, females used 20% more of this resource than males of the same age. The whole life health care costs for females was estimated at 34% higher than that for males in the study by Alemayehu and Warner (2004), with the majority of this additional cost attributable to both greater female life expectancy and much greater expenditure on residential care costs in later life. Hoffmann and Nachtmann (2010) implies that females tend to need greater care than males for a number of reasons: firstly, they tend to live longer; secondly, they suffer more often than males from non-fatal chronic illnesses or comorbidities; and thirdly, since they tend to outlive their male partners, are less able to depend on such partners for help and support.

This evidence suggests that females have historically had greater health care needs than males, due mainly to their increased longevity and some specific morbidities linked to their gender. However, the recent trend is for this differential between males and females to narrow (particularly in life expectancies) (Karlsson et al., 2009) – and given

that males are now living longer, previously late onset, male only morbidities, such as prostate cancer, may become more significant in the larger old-old male population.

3.3.4 Disability

In modelling of the influence of greater health costs later in life and changes in morbidity, Caley and Sidhu (2010) note that "*The potential for different conditions to compress or expand morbidity varies greatly and the future consequences of modern healthcare are still unpredictable.*"

In a model of the health care costs associated with specific morbidities, the cause of death was incorporated into models of residential care costs by de Meijer et al. (2011). Diabetes, mental illness, cardiovascular disease, a respiratory or digestive disease were seen to have significantly higher total lifetime care costs whilst cancer deaths have lower costs. Looking at just hospital costs however, Seshamani and Gray (2004a) show that cancer, stroke and other diagnoses are associated with higher costs relative to respiratory and heart conditions. Denton et al. (2002) present a consideration of the average individual's physician costs for 19 categories of physician services in Ontario by gender and age band, derived from administrative data. The cost of many of these services is seen to peak in the elderly population (and decline for the old-old), the only exceptions being paediatrics, psychiatry and obstetrics/gynaecology which show peaks at earlier ages.

Looking at comorbidities, work by Rasulo et al. (2009) estimated disease free life expectancy (which is potentially less limiting in the activities that an individual can undertake than disability free life expectancy) broken down by both single disease morbidities (cardio vascular; infections; cancer; respiratory; other acute) and by comorbidity groupings of these morbidities. This work suggests that years of life in morbidity due to a single morbidity has declined in the UK since the 1990's, but this decline has been matched by an increase in years of life with a variety of comorbidities.

Care is needed when projecting past trends in the prevalence of morbidities. Based on the extrapolation of recent trends in standardised death-ratios, McCulloch (2012b) presents a calculation of such ratios for cancer and coronary heart disease (CHD). These standardised death ratios are country specific death rates applied to an international standard population profile of 100,000 individuals – so as to eliminate the effects of changing population structures both between countries and over time that a raw death-ratio would provide. These calculations show that by simply extrapolating past trends, deaths due to morbidities such as coronary heart disease (CHD) may be almost eliminated in the not too distant future! (His figure is reproduced here as Figure 3-3.)

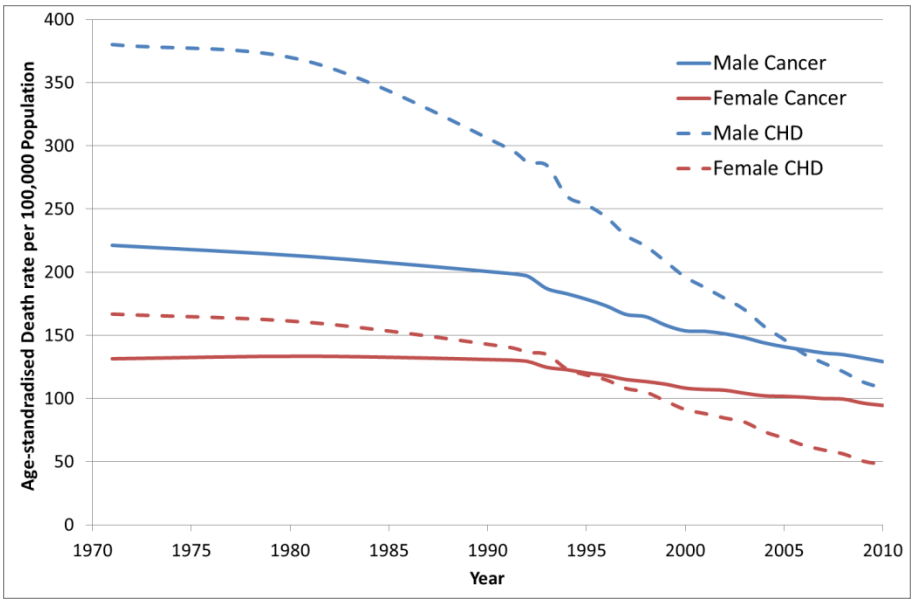


Figure 3-3 : Trends in standardised death ratios for Cancer and CHD, England and Wales, 2010

(source : McCulloch, 2012b)

The reductions in deaths with CHD as the main cause of death seen in Figure 3-3 are dramatic, especially for males. These reductions may result from two mechanisms: firstly, individuals may be less likely to develop CHD in the first place, driven by changes to occupations, lifestyles or pharmacological interventions (e.g. the introduction of cholesterol lowering stains, see Law et al, 2003) or secondly, that the morbidity is managed better, so that individuals with CHD live to die from other morbidities.

The differing impact of various morbidities on health care is well understood. Some morbidities such as cancer can cause an intensive use of resources and thereafter the use is minimal, either because the individual has recovered or deceased. Other morbidities, such as respiratory illnesses, are much less intensive in their use of health care but the period of use is more protracted. However, the combined impact of comorbidities is less well understood – is the impact simply the worst amongst the morbidities or to what extent does the presence of other morbidities increase the impact?

3.3.5 Socio-demographic status

As well as age and gender, there will potentially be other demographic factors that affect an individual's use of health care. Grundy and Sloggett (2003) showed that lack of education qualifications and lack of social support systems were linked to increased likelihood of reported bad health for the old population. Additionally the influence of marital status (married; single; widowed and divorced or separated) for females in particular, was found to be weak or counter intuitive (with married females enjoying no

significant better health than those in other categories). A study by Murphy and Martikainen (2010) however showed that being married or widowed greatly reduced the number of bed days (both in hospital and in a residential setting) compared to their divorced or single counterparts. Seshamani and Gray (2004b) were only able to demonstrate reduced hospital costs for married individuals over those who were single, widowed or divorced (however widowed showed the smallest difference). Simeonova (2013) explored a possible link between a bereavement event and subsequent deteriorations in health outcomes.

A consideration of health statistics indicates that health status varies by ethnicity (HSCIC, 2005a and 2005b, Parliamentary Office for Science and Technology, 2007). These studies found that generally poorer health and increased likelihood of an acute sickness is reported amongst the elderly (55 or older) south Asian population of both genders. This population also experienced greater incidences of cardio-vascular morbidities, stroke and diabetes (this latter morbidity was also common amongst the black Caribbean population). The Chinese population were seen to have generally better health and lower incidence of these morbidities. There is little consensus so far on how much these differences are attributable to lifestyle, socio-economic factors or genetic pre-dispositions. Modelling may help to dis-entangle these impacts. With regard to ethnicity, probit modelling of survival, cohabitation, working and health status by Karlsson et al. (2009) shows that controlling for many factors, white ethnicity (and higher education) were associated with greater healthy life expectancy. Lubitz et al. (2003) estimated that members of the black community had lower life expectancy at age 70 than the white community (11.5 years vs. 13.5) and fewer of these years were active (5.4 years vs. 7.2).

3.3.6 Socio-economic status

The economic status of an individual is likely to impact on their need for health care, not just in their elder years but throughout life. McCulloch (2011) illustrates this by plotting the proportion of residents on benefits against the number of recorded FCEs for English LADs. His chart (together with a similar chart for the proportion of the population aged 75 and older versus FCEs) is reproduced below as Figure 3-4. In this chart, there is clear evidence of a positive, but weak, correlation between levels of receipt of income benefits and the number of FCEs. A consideration of the socio-economic impacts on self-reported and objective measures of health by Grundy and Sloggett (2003) also suggests that health outcomes are poorer for those in receipt of income support payments and those housed with social sector housing tenancies. This finding was relatively consistent across all the measures of health outcomes and for both males and females.

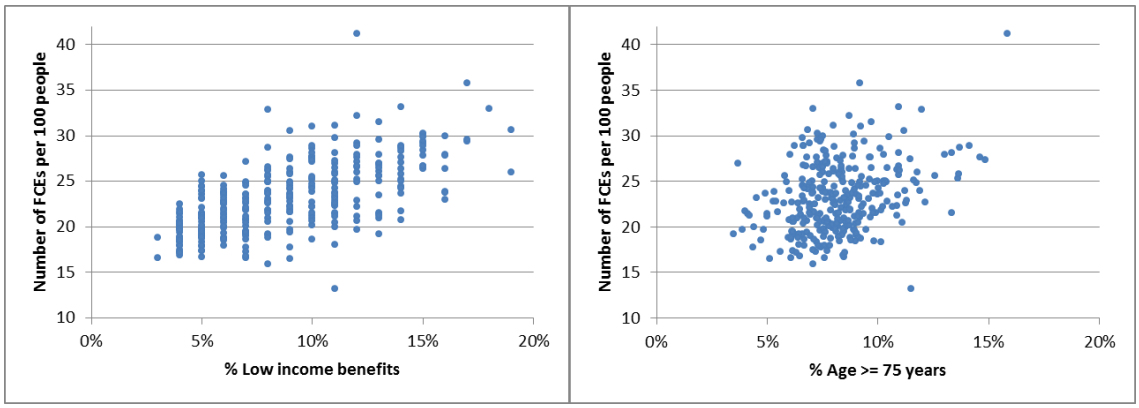


Figure 3-4 : Hospital admissions per 100 people, by income (lhs) and by age (rhs), LADs, 2008

(source : McCulloch, 2011)

Melzer et al. (2000) using two point survey data and NHS death records showed that males aged 65-69 from the first quartile of social status (professional or managerial) not only enjoyed greater life expectancies than those with other social status (15.0 years vs. 13.1 years) but they also enjoyed fewer years in disability (1.0 year vs. 1.6). Whilst females from the first quartile also enjoyed longer lives (18.7 years vs. 17.2) the numbers of years in disability were similar amongst the two groups (3.4 years vs. 3.2). These general findings were seen to be invariant to the degree and definition of disability. An ONS study on disabled free life expectancies also showed that these expectancies were lower in the more deprived the areas of England (Smith et al., 2011a).

Possessing a higher education qualification and having higher incomes was shown to reduce the number of visits by individuals to medical practitioners by Schellhorn et al. (2000), with a further indication that such people were more likely to access specialist practitioners than the population at large. A simulation study by Karlsson et al. (2009) showed that a 50 year old male who has a university degree and is working, cohabiting and healthy can expect to live for 15 years more than their male opposite number (lowest education, non-working, non-cohabiting and ill), and remarkably, spend 25 more years in a healthy condition.

A comparison between the number of bed days by manual and non-manual workers in Murphy and Martikainen (2010) shows that whilst manual workers had greater use of bed days, this difference was only slight. Seshamani and Gray (2004b) also estimated that non-manual workers had significantly lower quarterly hospital costs.

Clearly higher incomes and wealth, however they are measured (living environment; tenure of accommodation; educational attainment; employment) are positively related to a healthier lifestyle and a consequent reduction in health care needs.

3.3.7 Geography

What emerges from many of the studies is the lack of geographic differentiation in the reported findings. When considering the needs and the delivery of health care provision the variability between locations is critically important. Many reported studies simply model data from a specific region (using a bespoke survey) or a country (using a national data set). Where a study is conducted for a specific region, using locally collected data, there is the potential for arguments over whether these results would be still valid at a national scale, but even in such a case however, they would still not allow for any geographic differentiation (e.g. Seshamani and Gray, 2004a, who claim that their study which is based on Oxfordshire data, is representative of England and Wales as a whole).

One study that has both national coverage and a detailed geography is that by Smith et al. (2011a) who used a geographically specific measure of deprivation to highlight health inequalities, showing greater ill health in more deprived neighbourhoods. A similar piece of work by Webber (2004) also examines how total hospital admissions with certain diagnosis vary by type of neighbourhood, as characterised by the MOSAIC classification (Experian, 2003). This work also provides a methodology to identify which neighbourhoods need targeted resources so as to effectively tackle a range of 19 diagnoses. The morbidities of Chronic Obstructive Pulmonary Disease (COPD), strokes, emergency admissions and injuries and poisonings are seen to have particular high admission rates in elderly neighbourhood types. Staying with classifications, work by CACI (2007) classifies postcodes within the UK according to their health status and diet, so that again, health improvement resources may be targeted at local areas. The headline classification categories used are: Existing problems (15% of postcodes); Future problems (12%); Possible future concerns (35%) and Healthy (37%, not 33% as given in the report). Selected sub categories contain elderly traits, e.g. "Elderly with associated health issues" (3%); "Affluent healthy pensioners, dining out" (7%). The authors comment that "*Looking at existing illness alone is unlikely to focus resource on those areas that will present major health issues over the next 20 years*."

3.3.8 Summary

Whilst many of the early studies into the health care costs of the elderly population cite age as an important factor, subsequent work suggests that once time to death is considered, age itself becomes less of a contributory factor. For some studies this finding presents a data challenge, since knowledge that an individual is within say, one year of death, is only available retrospectively. Subsequent work that revisited this issue, restored age to some significance in explaining health care costs, but time to death was still a more important factor. The most recent work has however suggests that

time to death is itself not the real driver for health care costs; instead it is the disability of the individual that is important. This simplifies the task of assembling a data set for modelling purposes since disability can be observed contemporaneously. The conclusion here is that some accurate assessment of an individual's state of disability is an important factor in not only assessing their 'on-going' health care need but also is critically important in assessing their magnified health care need in the period leading to their death.

Females currently have longer life expectancies than males but they also tend to have longer periods in disability and this means that they have greater lifetime health care costs than males. The trend, however, is for this difference to narrow.

The type of morbidity affects the profile of need for health care. Some morbidities may require an intense period of treatment followed by a return to almost full health or death (e.g. cancer); some may also be linked to a specific health event but the prognosis is for continuing treatments (e.g. a stroke) whilst others will be more subtle but still debilitating over a longer period of time (e.g. dementia, heart and respiratory diseases). This latter category probably constitutes the larger group in terms of health care use.

Considering socio-economic and demographic factors together, individuals who are white, have higher levels of qualification, higher incomes or wealth and do or did do non-manual work will tend to have longer life expectancies and spend many more years in a healthy state than their opposite counterparts. To an extent all these five factors will be measuring similar effects.

Many of the quoted studies here do not tend to incorporate any geographic differentiation in their findings. This is primarily a data issue with the required geographic detail not being present in these data or these data relate to just one specific location. This deficiency is important since it is likely that the makeup of an area is determined by its history, culture and population, and therefore health outcomes will vary by area.

3.4 CASE STUDY MORBIDITIES

In this thesis three case study morbidities that particularly affect the elderly population are studied. Examining statistics on the causes of death in 2012 for those aged 50 and older, the three most common causes of death are: neoplasms (also known as cancers) (ICD-10 codes C00-D48) which account for 29% of such deaths, diseases of the circulatory system (ICD-10 codes I00-I99) 28% of deaths, and diseases of the respiratory system (ICD-10 codes J00-J99) 15% (ONS, 2014d and Table 5 of the associated data table). Examination of usage of secondary care, measured through FCEs

shows that in 2013-14, 15% of such episodes are associated with neoplasms, 12% with CVD and 8% with respiratory illnesses (HSCIC, 2015).

In terms of specific definitions, for circulatory diseases the term CVD is used to cover both ischaemic heart (IHD) and cerebrovascular diseases, the former including the occurrence of a myocardial infarction (a heart attack) or angina whilst the latter is commonly results from a stroke. For respiratory diseases, these are typically signified by the presence of a lung disease or asthma. Cancers can occur in many locations within the body and can have diverse causes, some of which are lifestyle related whilst others are part of the ageing process. It is this variety and range of causal factors that make it challenging to attribute the onset of the generality of cancer to specific and significant factors. Of these three major causes of death, only the morbidities of the circulatory and respiratory systems are studied here.

Deaths attributable to endocrine, nutritional and metabolic diseases (ICD-10 codes E00-E99) such as diabetes are less common at 2% of all deaths and 2% of FCEs for those aged 50 and older. However diabetes is still a significant morbidity and was estimated to be a direct cost to the NHS of £13.9 billion in 2010 (Kanavos et al., 2012). These costs do not include the additional indirect costs associated with common complications such as glaucoma, foot conditions and strokes and the economic costs of absenteeism, early retirement and increased social benefits. Notwithstanding the low proportion of deaths and hospital admissions, but recognising the cost of diabetes and, the comorbidity links between diabetes and CVD, DHBS is the third case study adopted in this thesis.

In the remainder of this section, the determinants found in the literature that are particularly associated with the onset of these three case study morbidities are considered.

3.4.1 Cardio vascular disease

The now abolished Yorkshire and Humber Public Health Observatory published a series of fact sheets for CVD in 2013 (Public Health England, 2013a) that report on findings in the literature around CVD. Each sheet covers a significant influence that determines an individual's propensity to acquire CVD. Some are themed on lifestyle or behavioural influences (sheets 1 to 5), whilst others to non-behavioural influences (sheet 6) and some to information on linked morbidities (sheets 7 to 16).

3.4.1.1 Behaviour

There are five main lifestyle or behaviour influences that contribute to an individual developing CVD. The first influence is smoking – studies have estimated that current smokers increase their risk by a factor of up to three times when compared with those

who have never smoked. Even ex-smokers have a near doubling in risk. These increased risks are however moderated by age, as individuals get older the relative risk associated with smoking tends to decrease. Exanimation of long terms trends in smoking status suggests that fewer individuals will have smoked in future cohorts which means that the prevalence of CVD should be lower in future (Murphy and Di Cesare, 2012, illustrate in figures 1 to 5 the likely reductions in future smoking behaviour and go onto to estimate that much of the improvement in the mortality of those born in the 1930s is due to variations in smoking).

The remaining behavioural influences covered in fact sheets 2 to 5 are linked together and concern obesity, physical activity, nutrition and alcohol consumption. Obesity primarily impacts on CVD through type-2 diabetes, with females who are obese having a 12.7 increased risk of type-2 diabetes and obese males a 5.2 increased risk. Unlike smoking, the trends are for increasing obesity in the English population. Regular physical exercise combats the effects of obesity and thereby helps to reduce CVD. It can also help to lower blood pressure and reduce stress further helping to guard against CVD. Like lack of physical exercise, poor nutrition contributes to obesity and hence CVD, however diet can also have a direct impact on CVD. The adoption of a 'Mediterranean' diet based on consumption of vegetables, nuts, seeds, fish and certain types of unsaturated fats is seen to be beneficial in protecting against CVD whilst salt, which increases blood pressure, increases the risk of CVD, as do saturated fats. Low to moderate alcohol consumption is reported to offer some protection against CVD but this benefit quickly turns to harm with heavy consumption. It is estimated that 5.8% (54,700) of all hospital admissions with a primary diagnosis of CVD in 2011/12 could be attributed to alcohol, and to provide some context, 137,400 could be attributable to smoking.

3.4.1.2 Non-behavioural

Non-behaviour factors include gender, age, ethnicity and deprivation. There appears to be little difference in the prevalence rate of CVD between males and females, although at older ages the morbidity is more prevalent for males. By ethnic group the higher rates of CVD overall are seen in the male Pakistani and Indian groups whilst the lowest rates are reported in the Black African and Chinese groups. For stroke, however, older Black Caribbean males have the highest prevalence, followed by Bangladeshi and Indian females. Evidence points to a social gradient with CVD, people living in areas that are more deprived and less affluent tend to have higher prevalences for CVD.

The separation of behavioural and non-behavioural factors that contribute to CVD can be problematic. For example there may be a (non-behavioural) genetic component that links to a propensity to become obese (behavioural). Also (non-behavioural) deprivation

can impact on an individual's ability to eat healthier and take regular physical exercise (behavioural).

3.4.2 Diabetes or high blood sugar

The campaigning charity Diabetes UK produces a periodic report on the statistics related to diabetes (Diabetes UK, 2012). The report covers the extent of diabetes in the population and various sub populations, distinguished by age and region. The report also highlights the risk factors associated with the morbidity and how it impacts on individual's lives, both directly and through medical complications. The report also contains a comprehensive list of references to the major studies into the morbidity.

Given the cited link between CVD and diabetes it is not surprising that there are commonalities in the risk factors that influence the prevalence of DHBS (see Morrissey et al., 2014, who explicitly modelled this comorbidity). The behavioural factors of obesity and nutrition are commonly cited; increasing levels of obesity linked to poor diet increases the chances of acquiring diabetes or having high blood sugar. The link with smoking, physical activity and alcohol is less certain, other than how these factors affect obesity. With regards to non-behavioural factors, once again those of a south Asian ethnic background have a higher risk of DHBS by a factor of up to 6 times (see Khunti et al., 2009, for recommendations on how to explore this finding). Those of a Black African or Caribbean ethnicity also have an increased risk factor of 3 times in comparison to the general population (HSCIC, 2005a). Also genetics plays a greater role with DHBS - having a parent or a sibling with the morbidity increases the risk.

3.4.3 Respiratory Illness

Respiratory diseases manifest in a number of forms. Beside cancers the commonest forms are COPD, pneumonia, tuberculosis and asthma. COPD manifests commonly in two forms, either bronchitis or emphysema. A statistical consideration of the diseases is provided in the Burden of Lung Disease report published by the British Thoracic Society (2006).

Unsurprisingly smoking is the most important contributory factor to developing a respiratory illness. Half of smokers develop some form of COPD whilst 10%-20% develop clinical symptoms (Devereux, 2006). Other causes of COPD include exposure to air pollution and from employment (e.g. inhaling noxious gases and dust particles) although the exposure to these causes in the UK has diminished over time. Both pneumonia and tuberculosis are infectious diseases, spread through contact with the bacterial agent. Whilst not caused directly by smoking there is evidence that smoking not only compromises the ability of the lungs to function whilst suffering an infection,

but smoking makes the acquisition of the infection in the first place more likely (Nuorti et al., 2000, Kolappan and Gopi, 2002, and Maurya et al., 2002).

Whilst it is unclear what causes asthma, the commonly cited cause is a result of an atopic reaction to living in a too hygienic environment – the 'hygiene hypothesis' (Strachan, 2000). Some studies have suggested that the causes of asthma are otherwise primarily linked to either a family history of the morbidity or childhood events. These childhood events include: having bronchiolitis as a child; childhood exposure to tobacco smoke; if your mother smoked during pregnancy; being born prematurely; or having a low birth weight. The extent of these childhood events varies over time – people aged 50 in 2010 for example will have been children during the 1960's when these events, particularly in regards to smoking, were more common. After a steady growth in the incidence of diagnosis of asthma in the population until the early part of this century, the trend is now for reduced incident diagnosis (Anderson et al., 2007 and Simpson and Sheikh, 2010).

More generally, there is little difference between the genders in the prevalence of respiratory illness but as people age the prevalences increase. The female south Asian and Chinese populations have the lowest rates of such diseases, linked to their low rates of smoking (reported in HSCIC, 2005a, and shown later in Table 8-15). Socio-economic status will influence the chances of acquiring a respiratory illness, those employed in manufacturing industries having a higher prevalence than those who work in the retail, service or professional sectors.

3.4.4 Summary

Two of the three morbidities selected for study in this thesis together account for nearly a half of all deaths in the 50 year and older population of England. CVD is just behind cancer in terms of the number of deaths whilst respiratory illness is the third largest cause. The morbidity of DHBS is often linked with CVD and so has also been included as the third case study morbidity. There is little evidence to suggest general gender or age differences in the incidence of these morbidities, although, if the morbidity does not significantly increase mortality, the prevalence of the morbidity will naturally increase with age. Lifestyle or behavioural factors are seen to have a large influence on the morbidities. The higher levels of obesity and its associated factors of less physical activity and poor nutrition are seen to increase the incidence of the linked morbidities of CVD and DHBS. Smoking has an impact on all three morbidities – more so for those who are current smokers. The most consistent and important non-behavioural factor that emerges for consideration is ethnicity. This finding may have an underlying genetic or lifestyle cause. The south Asian population have greater risk than the general population for CVD and DHBS but less risk for respiratory illness. Black ethnic groups have lower

risk for CVD (but within CVD, higher risk of stroke) but higher risk of DHBS. The Chinese ethnic group appears to have a lower risk for all three morbidities.

3.5 HEALTH DATA

This section provides a review of the types of data that have been used to describe and model elderly health care use. To a degree there will be an overlap here with section 3.3 on elderly health care determinants and the following section on modelling approaches. However, this section will also present an opportunity to step back and consider the specific features of a fuller range of health care data. In particular a consideration will be given of the strengths and weaknesses of each type of data and how any weaknesses may be addressed.

There are four potential sources for data on elderly health care which are most commonly identified in the literature. The first source is sample survey data that are collected for specific health related purpose within a defined region or locality such as the Oxfordshire (Seshamani and Gray, 2004a) and EIGER (Schellhorn et al., 2000) data sets.

The second source is also sample survey data but is collected on a national level. These surveys may be health specific (e.g. the Health Survey for England, HSfE) (Grundy and Sloggett, 2003) or may be general but still containing a significant health element (e.g. ELSA) (Institute for Fiscal Studies, 2015a and 2015b). One important consideration in connection with an elderly population is the nature of the population which is being sampled, often the surveys may not include individuals who are resident in an institutional care setting.

An important distinction for both regional and national survey data sets concerns whether they are cross sectional or longitudinal in nature (ONS, 2010a, section B.3). Cross sectional data may be collected as either as a one-off exercise or periodically, but the important feature is that each individual participates in the survey just once. With longitudinal surveys there is a desire to re-survey the same panel of individuals at a number of different points in time. The advantages of cross sectional surveys are that they are easier to implement and they produce larger samples than longitudinal surveys. A longitudinal survey however, allows a life course picture to be built up for individuals and transitions within this life course can be observed contemporaneously (rather than through recall). Also, estimates derived from longitudinal surveys tend to be more precise and less variable since they do not incorporate the person to person variability that is present in cross sectional surveys. The disadvantages for longitudinal surveys are related to the administrative work involved in locating the same participants at each survey ‘wave’, how to accommodate dynamic shifts in people’s arrangements (e.g.

marriages and divorces) and the need to refresh the panel as attrition takes place over time.

The third source of data arises from administrative data sets collected as part of the health administration process. Examples of these are the Hospital Episode Statistics in the UK (Cowper, 2004), Medicare related databases in the USA (Spillman and Lubitz, 2000) and various European systems (Zweifel et al., 1999, de Meijer et al., 2011, Ziegler and Doblhammer, 2010, Hoffmann and Nachtmann, 2010 and Murphy and Martikainen, 2010).

A fourth source of health data are those collected as part of a periodic census of the population. For example, the English and Welsh 2011 Census contained questions on self-reported general health status, long-term illness and provision of care (ONS, 2010b).

3.5.1 Survey data

Most on-going general household surveys will capture some information on the state of health or the disability status of the members of the household. The largest such data set within the UK is the Integrated Household Survey which is composed of a set of 100 core responses extracted and combined from a number of ONS surveys. This combination of responses from a number of surveys allows for a higher precision in any estimates derived from these data and the potential for greater geographic detail (to local authority level) than any single survey (ONS, 2012c). The range of health questions in this survey is limited; concerned primarily with a self-assessment of the degree of ill health, its duration and the limitations it places on daily activities.

The United Kingdom Household Longitudinal Study (UKHLS) (branded as 'Understanding Society' and previously known as the British Household Panel Survey) (Brice et al., 2010) is an on-going longitudinal survey designed to capture the changing behaviour of individuals and households in the UK. The survey collects a similar range of health related data to the now discontinued General Household Survey (ONS, 2011). These data have been used to study income-inequalities in the use of health care amongst the elderly (Allin et al., 2006 and Hanratty et al., 2008); provision of care for the elderly (Hirst, 2004, Glaser et al., 2008) and how social capital influences health care needs (Sessions et al., 2011, Gray, 2009 and Pevalin and Rose, 2004).

Moving away from general household surveys, the HSfE is perhaps the most detailed source of health related survey information for England (equivalent surveys are conducted in Scotland and Wales) (Joint Health Surveys Unit, 2012). The HSfE is a cross sectional survey begun in 1991 and was initially made up of a sample of 16,000 individuals but this was reduced to 8,600 individuals in 2009. The survey not only

contains information on self-assessed health conditions that are commonly found in other such survey it is also supplemented with a nurse practitioner collected medication survey and a samples for biomarkers. These biomarkers include: body shape; physiology; blood anylates: physical function; sensory tests and samples of saliva, cheek cells, cortisol and DNA. Occasional booster samples are used to gain insight into a particular topic, for example in 2005 there was a sample boost for elderly respondents (Craig and Mindell, 2005). In 2011 the HSfE was reformed as the Health and Social Care Survey (HSCS) resulting in a revision to the set of core questions within the survey and the introduction of additional questions around social care

ELSA is a survey dedicated to capturing the behaviour and lifestyle of an ageing population. The survey was started in 1998 and is currently in wave 7 (2014-15). The initial sample members were recruited from respondents to the HSfE in 1998, 1999 and 2001 and ELSA is periodically refreshed using respondents from the HSfE. Like the HSfE, the survey also periodically collects medication and biomarker data from individuals at waves 0, 2, 4 and 6. Recent studies using ELSA have included work by Tanaka et al. (2011) that used ELSA data to estimate how the socio-demographic status of an individual affected their susceptibility to develop diabetes. Shankar et al. (2010) used ELSA to report on the likelihood of individuals engaging in risky health behaviours (smoking; excessive alcohol consumption and lack of exercise) as a result of their socio-economic and demographic background. Steel et al. (2006) used self-reported data by individuals within ELSA on the quality of the health care that they received in regards to various health conditions.

3.5.2 Administrative data

As a result of how the American health care system is funded, many American based studies use data held by medical insurance companies or provided through state (Medicaid) and federal (Medicare) programmes, with the Medicare scheme being primarily concerned with the 65 year and older population. Many studies using these data sources have been concerned with the modelling of cost for health care (Lubitz et al., 2003). Where available, some studies have advanced the work by considering the cost of individual physician services as recorded in such databases e.g. in Ontario, Canada using a Health Insurance Plan database constructed to administer fee-for-services system (Denton et al., 2002).

A more ambitious undertaking is to attempt to link together a number of administrative data sources. For a population model de Meijer et al. (2011) used three types of administrative data linked at the individual level. These data are able to provide the cost of care (an Administrative Office Exceptional Medical Expenses data set from 2004);

the time to death and cause of death (the Death Causes Registration 2004–2007 data set) and household and individual characteristics (the Municipality Register 1998–2006).

With regard to hospital admissions in the UK, the English HES data provides a comprehensive source of in and out patient data, measured in terms of FCEs. The primary intended purpose of these data are to provide a mechanism to enable health commissioning bodies to reimburse clinical providers for the cost of care and to provide data for various performance indicators. These data do however have the potential to be a rich source of data for research purposes (Garratt et al., 2010). The HES contains admitted patient care data from 1989 onwards, with more than 14 million new records added each year, and outpatient attendance data from 2003 onwards, with more than 40 million new records added to this database each year. These data are made available either in aggregated tabular format or, under certain conditions, at individual FCE level. The data quality for inpatient diagnosis is good (since payment is linked to the quality of this field that provider's record) but the data quality for outpatient diagnosis is not so good.

Looking at studies of morbidities affecting an elderly population that have used the HES data, many studies use HES data to assess clinical outcomes or discharge destination. For example Jeyarajah et al. (2009) and Faiz et al. (2010) predict as outcomes the length of stay/re-admission/mortality for digestive conditions that dis-proportionately affects the elderly population. Gilbert et al. (2009) predict the probability of a discharge to the patient's home or to an institutional setting after a fall using HES data from a single hospital in the South east of England.

However, the studies which are perhaps most relevant to this thesis are those that consider factors that lead to admissions or re-admissions to hospital. An early study by Birrell et al. (1999) used population projections and age/sex specific rates of hip replacements, derived from HES, to forecast the need for such operations to 2006, 2016 and 2026, assuming rates remained constant. As part of a wider consideration of 'unintentional' falls amongst the elderly, Scuffham et al. (2003) used HES data to estimate the incidence and cost of hospital admissions due to falls, by four elderly aged bands. Hospital admission rates were seen to increase by an order of magnitude between the 60-64 age band and the 75 plus age band. Roland et al. (2005) investigated the re-admission pattern of a high-risk elderly group (those having two or more hospital admissions in one year) against a general elderly population. Their study, using HES data from 1997 to 2003, found that eventually the admission pattern of the high-risk group (after 3 to 5 years) is not dissimilar to that of the general elderly population. Billings et al. (2006) used a 10% sample of HES data to also predict future likelihood of hospital re-admission across a range of morbidities. From a sequence of logistic

regressions they found that of all the age bands, only belonging to the two oldest age bands, aged 65-74, and 75 or older, were significant in influencing the likelihood of re-admission.

3.5.3 Summary

Each type of data set has its own strengths and weaknesses. The general survey data sets are generally large and able to provide accurate estimates for general health at some intermediate level of geography. Unfortunately, in these types of survey the range of health information is usually limited to self-assessments of the individual's general health condition and whether they have a number of functional limitations. The specific health and elderly survey data sets are much richer in their coverage of health, using both self-reported measures of health status and more objective biomarkers. These data sets also record use of health care resources. However, such surveys are usually small in size and are less able to produce reliable estimates for certain variables, particularly when segregated by some characteristic or geography. Whilst the administrative data sets are primarily designed for payment and monitoring purposes, they can be useful for research purposes and are able to provide an almost universal coverage of health care use by a population, especially when health care is free at the point of use. This universality enables the information to be provided at detailed geographic level (and at a greater regularity than periodic censuses). The detail of the characteristics of the patient may however be lacking in this data set, as can, surprisingly, the diagnosis accuracy, see Mathur et al. (2014), for an example that considers the quality of ethnicity information in HES and Khan et al. (2010) and Herrett et al. (2009), for an assessment of the quality of diagnosis information in the General Practice Research Database. A later study by Herrett et al. (2015) concludes that the quality of the data in the (rebranded) Clinical Practice Research Datalink is improving over time as its research value becomes apparent.

From this consideration it is clear that no one data set is likely to be able to provide the entire picture on elderly health care and how this use is influenced by various factors. Ways of combining these data sets are required so as to bring their varied strengths to bear on these considerations, whilst minimising their weakness.

3.6 MODELLING METHODOLOGIES

Due to its economic and policy importance, health care modelling is a fertile field of academic study. Brailsford et al. (2009) report a study that both defines a strategy for conducting a review of health care modelling and reports the outcome from such a review. The need for a systematic approach is highlighted by their estimate that a simple '*((healthcare or health care) and (modelling or modeling or simulation))*', search term produces an extra 30 articles per day. After a distillation of nearly 10,000 articles they

conducted full reviews of over 300 to identify date, country of origin, primary and secondary analysis methodologies, funding sources and level of implementation. They find that the vast majority (nearly 200) of articles use either statistical modelling or statistical analysis techniques, with the third most common methodology being simulation. They describe as "*depressing*" the low level of implementation of the results of the studies (5.3%), which is ascribed to the lack of grey literature in their initial search.

As Brailsford et al. (2009), identified the predominant techniques used in studies of elderly health care have been either statistical or 'mathematical'. Mathematical techniques are those that involve some form of simple calculations. This may be the use of population estimates or projections which are then multiplied by some estimated age and gender specific rates of disease prevalence or health care use. This is the approach developed by the Projecting Older People Population Information organisation (Poppi User Guide, 2011) to predict the prevalence for a range of morbidities to 2030, and used by the Mental Health Observatory (Glover, 2008) to produce forecasts of dementia suffers to 2025. These rates are commonly static over time (which can lead to over estimates, see Mason et al., 2015) but may be evolved along some identified or extrapolated trend.

Some of the studies identified by Brailsford et al. (2009) involve the use of published life tables. These tables show, at each age, the probability that a person in a given population will die before their next birthday. They can be used to calculate both life expectancy and (with certain assumptions concerning disability rates) healthy life expectancy for people of different ages (Jagger et al., 2007). An example of such an application is Ziegler and Doblhammer (2010) who used life tables/population projections and dementia prevalence assumptions (that evolve over time) to estimate future levels of dementia within the German population to 2047.

Another approach is to model how a population evolves over time using transition probabilities in a Markov context (Sato and Zouain, 2010). Such studies involve the enumeration of a number of mutually exclusive states that individuals can occupy and the calculation or estimation of the probabilities that individuals transition between these states over time. For their study, Lubitz et al. (2003) developed a series of age, gender and race specific state transition probabilities, with states ranging from no health limitation to institutionalised. A Monte-Carlo/Markov Chain simulation exercise was then conducted to see how the health of a population of 100,000 over 70 year olds evolved, with individuals moving between states of health using a comparison between sampled pseudo random numbers and these transition probabilities. Another simulation study, conditioned by probit parameter estimates, was also used by Karlsson et al.

(2009) to examine how life and healthy life expectancies varied by socio-economic characteristics.

3.6.1 Statistical modelling

Many of the studies reported in the earlier sections of this review used statistical estimation techniques. This is to be expected since such techniques are well understood and have a strong theoretical grounding. They allow for the modelling of both individual and aggregate data and for an objective consideration of the importance of influences through significance levels.

For modelling aggregate data such as health care costs or length of stay, traditional linear or generalised linear models are often used. For individual level data a common approach adopted in the modelling of elderly health care use involves a two stage process. The first stage establishes the likelihood of accessing health care, usually by the use of a choice model. A second stage is then to estimate the extent of the health care use (usually though a cost measure) by the subset of users identified in the first stage, using a regression model, either ordinary least squares or a more generalised linear model (Seshamani et al., 2004a and 2004b, Dormont et al., 2006, and deMeijer, 2011).

3.6.2 Static microsimulation

Static microsimulation systems are complex mathematical systems that are able to measure the impact of policy changes on individual's behaviour. Such models are widely used for a variety of purposes including pension planning; changes to tax-benefit systems; higher education funding and inter-general issues (Anderson and Hicks, 2011). Many models attempt to synthesise or predict a population's demographic structure, since such a structure is expected to have a large impact on welfare, pension and health care need.

These models are typically based on a regional or national survey data set that contains detailed information about individuals or their households. These data sets are normally issued with a set of weights attached that measure how this individual or household represents the population as a whole. Techniques are then applied to these weights to evolve the surveyed population, typically through time, to represent alternative or future populations. Where these data are financial, such as incomes or costs these can also be updated based on gross domestic product or inflation growth assumptions. Once this evolved population has been estimated, it can be examined to explore issues such as its age and gender make-up, its labour force status, health status and income distribution. Also the cumulative wealth of the population can be established by accumulating information as the evolution occurs. The impact of policy changes (e.g. tax and benefit

rates) can be modelled by changing the rules that are applied to the populations through time.

Many examples of the application of static microsimulations in the area of health care are provided in the book Gupta and Harding (2007). The book includes papers that consider:

- the measurement of health status (through analysis of birth cohorts in Britain and the care needs and delivery of care for the elderly in Scotland);
- the costs of pharmaceutical treatments (predominantly within the Australian health care system);
- the financing of health care, through a short horizon (2002-2006) simulation of medical insurance costs in China;
- various studies into the availability of care for an ageing population, covering both formal care (the number of nurses and physicians) and informal care (the size of the 'middle' generation);
- data issues, around non-response and how the desire for confidentiality impacts on data quality;

The book also contains a section that provides a review of ten static microsimulation models, providing summaries of their history, scope, data source, technical detail and uses.

Elsewhere, Pearson et al. (2011) describe a static microsimulation model of the 2002 New Zealand population built to measure patient visits to general practitioners (GPs); practitioners activities and the range of morbidities in the population. The verified and validated model was then used to explore various family support and GP retention scenarios to the year 2021. Shi et al. (2011) used a microsimulation of the Californian population to estimate the future prevalence of type-2 diabetes under various Body Mass Index trends to the year 2021. A microsimulation model of the future elderly population (FEM) is developed in Goldman et al. (2004) to help predict the likely impact of a wide range of medical breakthroughs on health care status and costs. The same model is used in Michaud et al. (2009) to compare health status and life expectancies in Europe and the USA.

3.6.3 Dynamic microsimulation

The next step beyond static microsimulation is dynamic microsimulation (O'Donoghue, 2001, and Zucchelli et al., 2010). The major difference in a dynamic microsimulation is that individuals move between 'major life events', e.g. marriage, parenthood, divorce, labour force participation, morbidity and mortality. This is primarily accomplished

using information on transition probabilities between each state. These probabilities may also be conditional on other characteristics such as age and gender.

Given the complexity of such models they are less commonly found in the literature than static models. Zaidi and Rake (2001), Spielauer (2007) and Rutter et al. (2010) contain reviews of the use of such models in the context of an ageing population and health status. As with static microsimulations, a further review of the details and technicalities of seven dynamic microsimulation models is provided in Gupta and Harding (2007).

Like static models, dynamic models are able to provide basic statistics on the projected age and gender breakdown of a population with some able to additionally supply information on levels of elderly residential status within the population and disability (for example the Canadian LifePaths model, Légaré, 2011). The details on health status are more limited. It is possible with most of these models to augment them so as to produce more detailed health projections. A paper by Walker (2007) describes an illustrative study of the use of a health 'add in' to a dynamic microsimulation model to measure the ability of older Australian citizens to stay in the work force to the year 2018. Fukawa (2011) used a dynamic microsimulation model to simulate the health care expenditure costs on the elderly in Japan to 2050, claiming that, under various scenarios, these costs would increase from 8% of GDP in 2006 to near 11% in 2050, which appears to be an optimistic outcome.

3.6.4 Small area population estimation

As mentioned previously many of the studies in elderly health care have estimated either national models or region specific models, with no consideration of geographic variation. Commonly this is either because there is insufficient detail of data at smaller geographies to provide reliable estimates or the data set used did not actually contain any detailed geographic information. In planning where future demand for health care is likely to occur, a consideration of sub-national geographies is essential. This leads to a requirement that any such modelling should have a strong spatial element (Rahman et al., 2010).

Whilst the England and Wales decennial population Census and some administrative data sets are able to provide information at a very detailed level of geography, certainly sufficient for health care planning, their richness in the field of health information is limited. Self-reported counts are available to highlight individual's general health (on a five point scale) or the presence of a long term limiting illness (on a three point scale). Other data sources, such as HSfE and ELSA are able to provide the required richness of information, but unfortunately the detail of the geographic coverage is poor. To

overcome these limitations a range of synthesis techniques are available to augment or simulate a richer, small area geography data set.

A simple approach to combining information from both these two types of data is to use the survey data to establish a rate of occurrence of a measure of interest by a set of characteristics common to both data sets (e.g. how incidence of Alzheimer's relates to age and gender). These rates can then be applied to a cross tabulation of these characteristics from the population data set, which is available at a detailed geography, to arrive at a local estimate of the measure of interest.

A slightly more sophisticated approach involves modelling the relationship between the measure of interest and the common characteristics. The process begins by estimating a relationship between the measure of interest and any common characteristics using just the survey data set. This model can then be used to estimate the measure of interest for any population that possesses these same common characteristics, at any level of geography. This technique was used by Scholes et al. (2007) to estimate four health indicators (smoking, drinking, obesity and healthy eating) for Middle layer Super Output Areas in England and by Schofield et al. (2012) to estimate the ease of access to general practitioner services in rural Australia. A similar approach for disaggregate data involves the estimation of choice models using the survey data and using these models to predict whether individuals from a population data set will have the attribute of interest. This approach was used by Charlton (1998) to estimate the prevalence of serious illness for English local authorities. The disadvantage of this approach is that models have to be separately estimated for each measure of interest in the survey data.

A further technique that is able to produce estimates of a measure of interest at a detailed geographic level is termed small area microsimulation (Birkin and Clarke, 1989 and Hermes and Poulsen, 2012a). This technique attempts to allocate individuals from a survey data set (which has a coarse geography) to a geographic area based upon the commonality of various attributes. The aim is that individuals will be wholly or proportionally allocated to those areas whose inhabitants they have the most in common with. What this requires is a set of characteristics common to the survey data and the area in order to establish this similarity. Rather than providing one or more modelled measures of interest, here all the attributes of the individuals newly acquired by an area are then available to describe the area.

The use of small area population synthesis methods have been used widely in the field of health (Brown, 2011). These techniques are able to produce synthesised local populations which may then be used in further modelling exercises. Such uses include producing evaluations of policy interventions or projections of future populations. Ballas et al. (2006) used SimBritain (a national small area microsimulation model) and

Ballas et al. (2005) examine how health status was related to income for electoral wards within the City of York in northern England. The comparison was made using British Household Panel Survey data simulated at the ward level. Other similar health studies have included investigating the prevalence of obesity (Procter et al., 2008; Edwards and Clarke, 2009, Edwards et al., 2011); diabetes (Brown et al., 2009); people who smoke (Tomintz et al., 2008; Hermes and Poulsen, 2012b; and Smith et al., 2011b); care needs (Lymer et al., 2008, and Lymer et al., 2009) and general practitioner utilisation rates (Morrissey et al., 2008).

3.6.5 Spatial microsimulations

Many of the microsimulation models so far considered have used a national or regional data set to model policy interventions and demographic trends. There are a further range of models that attempt to model these impacts at a variety of locations, allowing for some geographic differentiation. Since data with the geographic detail required for such models is not commonly available, some small area population microsimulation may be required to produce the information rich base population for an area. Reviews of some of these models and their specific health relevance can be found in Gupta and Harding (2007).

Applications of spatial microsimulation models are relatively rare. A study by Harding et al. (2011) used a spatial microsimulation technique to project to 2027 a population age structure for each of the 1,422 Statistical Local Areas in Australia. The purpose of these projections was to assess both the health care needs of an ageing population and the childcare needs of young families. Future projections of the population age structure, ethnicity and disability status to 2031 are presented as example outcomes from an (event-driven) spatial microsimulation model of electoral wards in Leeds described in a paper by Townend et al. (2009).

3.6.6 Simulation

Brailsford (2007) provides a review of simulation case studies in health care, concentrating on Discrete Event Simulation (DES) and Systems Dynamics (SD). A DES models entities (e.g. patients) as flows through a system of queues prior to obtaining some kind of service (e.g. treatment). Each entity has its own characteristics that determine its path through the system and monitoring data can be collected on the entity and the service units. Such DES models usually have a random or stochastic element that simulates the choices available to the entity. A different approach is adopted by SD models. Here there are no individual entities, instead groups of entities move through a system spending time occupied in a variety of states. Speed of movement is usually controlled by restrictions on this flow of individuals. Such models do not require a stochastic element since there is no concept of an entity taking

independent decisions in these models. Brailsford (2007) reviewed how these two models have been applied in a health care context. They considered models at increasing levels of scale from disease models (patients) at the lowest scale, through capacity planning models (medical units) and onto strategic models (whole systems). They found that lower level models were more common and that these tended to be DES models whilst models at the third level were rarer and tended to be SD models.

Homer and Hirsch (2006) outline various systems dynamic approaches to modelling health care. They outline a range of suitable application areas, such as: disease and substance abuse epidemiology; patient flows and capacity studies. The article gives an example of a small hypothetical system to explore how best to deploy resources to meet an epidemic over a 50 year time span. The choice is between using resources to mitigate the complications from the disease or to fund prevention programmes. The model found that resources are more effectively spent on prevention rather than mitigation. The article refers to more comprehensive systems that are able to model wider health care systems, one such being Health Care Microworld (Hirsch and Immediato, 1999), which has been used as a training/what-if tool for American health care providers.

3.6.7 Agent based models

Agent based modelling is a population modelling technique that has seen application in a wide variety of fields (Crooks and Heppenstall, 2012 and Macal and North 2010). These models are implemented in software and are composed of a number of agents each of which represents an 'actor' in the system. In a health context these agents could be patients (with their own coded characteristics, e.g. gender, age, ethnicity); medical practitioners (e.g. with a speciality and cost) or health clinics or hospitals (with treatment specialities and capacities). Rules are then defined on now the various agents evolve over time and interact with other agents (of the same type or different). The advantages of this technique are that it: is flexible; it allows for the emergence of complex behaviours from a seemingly simple set of rules; it can model phenomena at different and multiple scales or time frames; and it may contain a notion of location or space that influences behaviour (Crooks et al., 2008).

Applications in the field of health care are limited (see Auchincloss and Diez Roux, 2008). A paper by Bilge and Saka (2006) refers to two Agent Based models, one of COPD prevalence and a second on the social network in a hospital department. For the COPD model, their simulator consisted of three elements: patients (with socio-demographic characteristics and morbidities); health care units (providing treatments) and events (such as weather, flu epidemics). They were able to forecast, over a 15-20 year time span, regional incidents of COPD on a weekly basis and claim that *"[The*

simulator] is producing plausible estimates of number of COPD patients for the future."

The study by Charfeddine and Montreuil (2010) also looks at COPD prevalence and contains two 'agent-orientated' elements. A population model firstly defines a total population (with various socio-demographic, clinical and life-style characteristics) that evolves into vulnerable, affected and registered populations (the latter being a population that has accessed the health care system). The second model is a health care delivery model consisting of agents that interact to provide care, e.g. patients; administrative and clinical staff. Rules define how an individual in the affected population accesses the agents within the health care delivery model. The model utilises regional data related to prevalence of COPD and the regional availability of treatment centres. Using the AnyLogic simulation platform the model is able to show the daily evolution of COPD vulnerable, affected or registered individuals and the utilisation of treatment centres.

Meng et al. (2010) used agent based models to examine the daily colonisation of drug resistant infections in a hospital ward environment. The model incorporates a patient location and movement model and a patient colonisation, detection and treatment model and is able to model the rate of colonisation within the ward. Kanagarajah et al. (2006) present an agent based simulation model of the hourly use of an emergency reception unit in a hospital. They consider the interaction of various resources, e.g. doctors; nurses; treatment rooms and administrative staff and also patient arrival rates and spikes. Within their model, agents are able to adapt to the environment, e.g. as queues grow doctors spend less time on initial consultations and additional on-call doctors can be utilised. Outcomes are measured in terms of waiting times and resource utilisation.

What is noticeable in a consideration of these reported studies is how few have gone through a formal "*verification, calibration and validation*" process (Crooks and Heppenstall, 2012) and perhaps explains why so few have been implemented (Brailsford et al., 2009)

3.6.8 Summary

This section has highlighted the range and diversity of modelling approaches. Firstly there are mathematical or mechanistic models that apply some formula to data to arrive at an outcome. Many of these models are based on the use of life tables. Secondly, statistical modelling is often performed in order to summarise the relationships within these data, and highlight the important factors that contribute to elderly health care.

The third most often used technique is simulation. This technique is implemented in a variety of ways and under different titles but essentially involves representing the

population that is of interest and using rules to modify this population, implemented using the Monte-Carlo technique. The parameters of this simulation can be informed by expert judgement, access to relevant data or by outcomes from statistical models.

In many cases the existing data may not be sufficiently rich to perform this modelling, estimation or simulation task and some initial manipulation of the data sets may be required. One such task is the ability to synthesise a population using a range of data sets that possess the required richness of both information and geography. Once this geographic and information rich data are available then techniques of simulation, (static/dynamic) microsimulation and Agent Based modelling are able to evolve the population and examine its interaction with health care provision.

3.7 POPULATION PROJECTIONS

The earlier sections of this review identified those characteristics of the population that have an important influence on morbidity. Whilst such a consideration is, in itself, essential to establishing future health care need, there is an additional requirement to establish the size of the population which will, potentially, experience these morbidities. These projections can be used to establish the size of at risk populations or to act as constraints on other population projection methodologies (e.g. dynamic microsimulation or agent based models).

3.7.1 Official projections and variants

The ONS is the body responsible for the production of official English population projections. These cohort component projections are normally produced biennially, with the latest estimates calculated on a 2012-base, and are broken down by age and gender. In the production of recent projections the ONS has been able to produce future projections that incorporate variations in future trends in fertility, migration and mortality. This means that in addition to the Principal projection, variants are also available (as highlighted in section 2.7). These variants are categorised as standard ‘single component’ variants; standard ‘combination’ variants and special case variants (ONS, 2013a and 2013b). By way of illustration, Figure 3-5(lhs) shows the projected population to 2031 for a range of these variant whilst Figure 3-5 (rhs) shows the variant population projections expressed as the percentage of the population aged 65 and over. This shows a similar picture with regard to the range of the projections, with the estimate of the proportion of old people in the population by 2031 varying from 21.5% to 23.8%. Clearly, any forecast of health care need will be highly sensitive to which projection variant is used

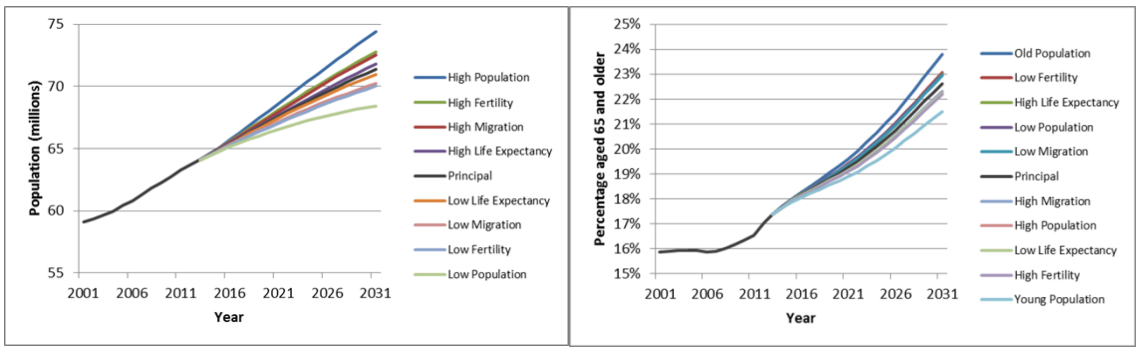


Figure 3-5 : Variant total population projections, counts (lhs) and percentage (rhs), UK, 2001 to 2031

(source : ONS,2013a and ONS, 2013b)

3.7.2 Projections of disability

A consideration of studies in the area of elderly health care need suggests that factors beyond age and gender are important: time to death, disability, socio-economic or socio-demographic status being some of the more prominent. Unfortunately, there are few readily available population projections for the population disaggregated by these characteristics.

Taking time to death and disability, standard actuarial techniques are available to predict what proportion of an age band is likely to die within the year, denoted by the death-rate at age x, m_x. Changes in these can be projected into the future to estimate what proportion of a future population is likely to die within the year (ONS, 2014e). For measuring disability itself, Marshall (2009) produced a set of projections to 2021 of disability rates for eight categories of disability, by age and gender, for each LAD.

The POPGROUP EXCEL software (Simpson, 2004) is able to produce a range of population projections, of both people and households. These projections can be disaggregated by household composition (and hence co-habitation status) and by economic activity. They are also available by age bands for UK LADs. Whilst the software is pre-configured to produce these sets of projections as standard, there appears to be the functionality to produce projections by other characteristics after an investment of expertise and time to set up a ‘derived projection’.

3.7.3 Projections by ethnicity

Studies have shown that some ethnic groups are more predisposed towards certain morbidities than others. Given these differences it would appear that some consideration of the future ethnic composition of any population projection would be essential when establishing the prevalence of certain morbidities. Lievesley (2010) produced projections of the population of England and Wales by 16 ethnic groups to 2051 under three migration variants. Whilst not constrained to the ONS projections, the total

populations were see to be close to those produced by the ONS. The report notes that currently the ethnic population has a generally younger population profile than that of the population as a whole, and that this feature will impact on future predictions for older members of these populations beyond 2030. Other population projections for the ethnic population have attempted to produce projections of the ethnic composition for small, sub national, geographic areas. Norman et al., 2010 produce five variant population projections to 2051 by 16 ethnic population groupings for English LADs. With the variant that was designed to most closely match the methodology used by the ONS, TREND-EF, the geographically aggregated difference between the two projections was small (0.6 million) and attributable to known issues. However, for the other four variants, the two of the four that were benchmark projections were seen to produce very much smaller aggregate population projections than the ONS whilst the two trend projections were seen to produce slightly larger projections. Comparisons in this study were also made with other ethnic group projections, such as those from the Greater London Authority (Klodawski, 2009) (where the agreement was good) and those by Coleman (2010) (where the agreement was not so good).

3.7.4 Summary

What begins to emerge from a consideration of this topic is that population prediction is a challenging task and the long term outcomes are sensitive to a range of assumptions. The value in the variant projections is that they provide a range of alternative scenarios that may be used in the conduct of model sensitivity analysis.

The challenge of prediction is made all the more difficult when there is a desire to disaggregate the projections by some characteristic. At best it would appear that in addition to age and gender only one other characteristic can be projected, with projections by more than three variables, especially at a local level, being unavailable. There may be scope to additionally incorporate some consideration of likelihood of death within the year, through national or regional death rate estimates and projections (which are already part of the methodology to produce population projections). Thus information on: gender; age; ethnicity and survivor status may be available.

3.8 HEALTH CARE DELIVERY

How to meet the health care needs of a future ageing population efficiently is as substantial a task as predicting what the actual demand for health care might be. Mayhew (2009) in section 3c, notes that *"... estimating health care costs is notoriously difficult, because consumption of health care tends to reflect the resources available and not underlying need. So to constrain actual demand to these totals would require health improvements in combination with supply side adjustments running in*

parallel.". Linking both the demand and the delivery of health care is therefore important in establishing either. Howse (2012) highlights a need to also consider health care in a wider context, recognising the value of earlier interventions in life and the treatment of people rather diseases, in particular that "*[health care services] have to get better at proactively seeking need rather than responding to demand, at involving patients in their own care, and at using alternative care settings to hospitals*.". Concerns and suggestions on how to "*fundamentally shift*" the health care delivery in the UK for an ageing population is provided in Oliver et al. (2014) and Mayhew and Rickayzen (2012).

3.8.1 Health care organisation in England

A summary of the previous and current structure of the demand (commissioner) and supply (provider) model of health care in England is provided by Roland et al. (2012). The paper comments on identifiable strengths and weaknesses in the previous arrangements and anticipates the challenges associated with a change to a model of general practitioner commissioning and greater competition within the health service. Additionally, Pollock et al. (2012) argue specifically that the move to area based CCGs of GPs will pose problems in the equitable allocation and monitoring of health care resources. Erler et al. (2011) present some issues and challenges on how primary care may develop, not just in England, but in the Netherlands and the USA. Challenges that are highlighted include an ageing population, the greater number of patients with complex chronic morbidities, patients with comorbidities and increases in patient expectations. The health systems in England and the Netherlands are characterised as "*strong*" with a well-developed and regarded primary care infrastructure. However, none of these systems are seen to be immune to the financial and resourcing challenges of an ageing population.

3.8.2 Modelling demand and supply

At a local level it is important to match together both the demand for health care and its supply. The study by Judge et al. (2010) used ELSA survey data to predict the need for hip and knee operations and HES data to estimate the provision for such operations. Using these predictions they were able to calculate for each local authority the equity of provision, expressed as a provision to need ratio. The highest values for this ratio were seen to cluster in local authorities within the South East of England, suggesting disproportionate over-provision of such treatments in these locations.

Tomintz et al. (2008) use small area estimation techniques to model the residential location of people who smoke, thereby establishing the potential demand for smoking cessation services. They then use a location-allocation model to establish how well the existing provision of smoking cessation services meets this need and suggest more

optimal locations (in terms of minimum travel distances). A recent study by Morrissey et al. (2015) uses a similar small area estimation approach coupled with location of acute psychiatric facilities to establish how the access to such facilities impacted on individual's own mental health.

Many agent based models in health care are able to model simultaneously both the patient demand for services and the utilisation of the resources that provide the treatments (i.e. the hospital or clinic) (Charfeddine and Montreuil, 2010, and Bilge and Saka, 2006). Using this approach it is also possible to model different delivery frameworks and monitor their impact, which can be measured in terms of treatment volume and efficiency for both the users and providers.

3.8.3 Health care innovations

Predicting how pharmaceutical and treatment innovations may impact on the need for health care is problematic. One approach is to establish some historic trends in treatment effectiveness and project these trends into the near future. Another more challenging aspect is to try and imagine what new and innovative technologies may become available. The report by Goldman et al. (2004) provides a table listing the outcome of a consideration of review material by an expert medical panel on likely medical breakthroughs, e.g. stem cell and genome derived treatments. The morbidities they considered included cardiovascular, ageing, cancer and neurological conditions. These breakthroughs were quantified in terms of the likelihood of them occurring within the next 20 years and the impact on the prevalence and severity of the morbidity.

One other areas of potential expansion in health care are provided by: the increased use of technology in both the monitoring and delivery of health care (Lindeman et al., 2011a and 2011b, Darkins et al., 2008 and Sood et al., 2007); education in drug and disease management (Bodenheimer et al., 2002 and Lam et al., 2011); greater assisted living (Audit Commission, 2004 and Department of Health, 2011) and greater involvement of local pharmacists in diagnosis and treatments (Barr et al., 2012). Less invasive medical technologies (endoscopes, MRI/CT and key hole surgery) may also make some medical procedures for a vulnerable elderly population more feasible (Clark et al., 2004), thereby increasing the number of such interventions but delivering each at a lower unit cost.

3.8.4 Summary

Whilst most of the previous sections have dealt with gaining an understanding of the drivers of health care demand this section has highlighted the issues around how this health care is delivered to meet these demands. This is a set of multi-faceted and complex issues. An important issue is the institutional organisation of the health care

system. The nature of the split between the organisations that commission health care (demand) and the organisations that provide the health care (supply) will define a high level architecture for the coordination of health care provision. For the providers there are a range of possible models, from a large diversified centrally located facility to more disparate clinics located in the community. This latter point also stresses the importance of establishing the local demand and supply for health care so that they can be effectively matched.

A further dimension of uncertainty is around the future developments in medical and pharmacological technology and how this may transform the current prevalence, severity and treatment pathways associated with certain morbidities.

3.9 CONCLUSION

Past trends in the United Kingdom point to the likelihood that there will be further increases in the life expectancy of individuals, particularly for males, and to the extent that such gender differences may be eliminated. Of these extra years, past trends also indicate that more of this time will be spent in ill health. These trends imply an expansion of absolute morbidity.

If age is the significant factor in determining health care costs then the prospect of an ageing population would appear to indicate that these costs and the associated need for medical interventions will increase substantially. If however time to death is more important than age then these increases in costs are likely to be more modest, with some of the costs 'pushed back' to the later years of life. Years to death, however, may not be the true measure of need; it may be simply acting as a proxy for an increasing state of disability towards the end of life.

Gender has been found to play a role in health care need, either in differences between the genders or in the ability to form support mechanisms through mutual care and co-habitation. In future, however, gender differences may no longer be a significant factor in explaining health care need. Ethnicity may however begin to play a more significant role as the size of the elderly ethnic population increases with a consequential re-balancing of the relative occurrence of certain medical morbidities. There is also a distinction between the later life health outcomes for manual and non-manual occupations, and as the nature of the workforce changes in this regard, another re-balancing will occur.

Trends in a person's lifestyle choices may increase the prevalence of some morbidities (e.g. increased obesity, linked to diabetes) whilst decreasing others (e.g. less smoking, linked to COPD). Different morbidities may also become more or less amenable to

enhanced or new technological and pharmaceutical interventions, which, whilst not offering cures, can help alleviate the symptoms of the morbidity.

There appears to be a wealth of data available to measure health care, through access to administrative data sets or household surveys. Whilst many of the surveys collect information on disability, only a few surveys are devoted to an in-depth consideration of wider health care needs and use. Fortunately, some well-developed and understood techniques are available to allow the strengths of the two data forms to be combined.

Much of the literature that explores the determinants of elderly health care use (measured primarily through cost) involves the building of mathematical or statistical models. Such models are well suited to identifying these determinants. Other techniques are, however, more suited to the exploration of future 'what-if' type scenarios. Such techniques include microsimulation (with an additional spatial consideration) and agent based modelling. The advantage of this latter approach is that the link between health care need and delivery can be explicitly modelled.

In terms of first research question set out in Chapter 1, which is to estimate some measure of health care demand, this thesis will model the prevalence counts and rates for a range of morbidities (see section 3.4). This is in contrast to the alternative measures of demand such as cost of health care or general health/disability which are used in some studies. For planning purposes the concept of general health provides little guidance in distinguishing the type of care required: primary, secondary, long term or acute. Health care costs are also a measure which provides little specificity in terms of resources requirements beyond the financial, and may be an imperfect measure of health demand. In terms of methodology, a simulation approach is adopted in preference to a statistical modelling approach so that it will be possible to segment the simulated spatial population of the local area by its characteristics, e.g. gender, age or ethnicity, as required by the second research question. The availability of the spatial population also allows the health impacts of demographic population change to be investigated, as required by the third research question.

Looking at the specific objectives, this chapter, together with the previous chapter on the research context has gone towards meeting the first objective of this thesis which was to paint a broad picture of the likely size and health condition of the elderly population of England and identify those factors that directly influence its health status. Parts of the second objective, to identify potentially useful data sources, and the third, to gain an understanding of the range of modelling methodologies used in this area of study, have also been addressed here.

The following chapter contains a detailed description and discussion about three sources of data that will be used extensively in this thesis: the 2011 population Census, ELSA and the ETHPOP population projections. As part of this discussion there is a consideration of the commonality of information between these data sets and how this knowledge helps to determine the extent of morbidities in the elderly population.

4 DATA

The 2011 Population Census and the English Longitudinal Study of Ageing

4.1 INTRODUCTION

The three data sets that are highlighted in section 3.5 and section 3.7.3 as being of particular relevance for studies into the impacts of ageing in the English population are described in detail here. Firstly, there is the annual Censuses of Population, which provides a detailed and quality assured picture of the English population at a range of geographic scales (see sections 4.2 and 4.3). Secondly, there is ELSA which has an almost unique depth of information on the nature of the older population of England (section 4.4). Through its longitudinal nature it is able to provide information on participant's changes in circumstances and how these may impact on their health status. The compatibility of these two data sets is explored in section 4.5. A third data set that was referred to in the review section on population forecasts is also relevant to this study. This is the ETHPOP population projections for English LADs that provide projections for the size of the population for a range of ethnic groups in each district. This data set is introduced in section 4.6. This chapter finishes with a discussion of its findings, in section 4.7.

In line with the second objective in section 1.5, this chapter provides the reader with an introduction to these three data sets and covers the nature of each data set, their composition and the advantages and disadvantages that they offer. As the following Methodology chapter will illustrate, ELSA data are used to provide a sample population for use in a spatial microsimulation and in the estimation of hazard models. The Census data are used to construct equivalent constraint tables for the spatial microsimulation and to guide the revision of the original ETHPOP projections. These revised ETHPOP projections are used to ensure that the future population structure in each LAD evolves to represent changes in its demographic composition.

4.2 2011 POPULATION CENSUS

The twenty second UK national Census took place on Sunday the 27th March, 2011. The primary purpose of each Census is to collect detailed information on the characteristics of individuals and households that are resident in the UK on the night of the Census (White, 2009). Each household within the United Kingdom was required to return a Census form, either on paper or electronically, that provided information on the household composition, household attributes (e.g. type of accommodation, car

ownership) and the characteristics of all individuals within the household (e.g. employment status, educational qualifications). A companion data collection exercise also took place for those individuals living in institutional establishments such as prisons and care homes.

4.2.1 Data outputs

The outputs from the Census come in many forms. The predominant dissemination format is through published tabular counts of individuals and households. These predefined tables are either univariate or multivariate in nature. For the 2011 Census, the Key Statistics (KS) and Quick Statistics (QS) versions of the tables are univariate and they are usually the tables that provide the most detail and are available for a wide range of population scales. The Detailed and Local Characteristics (DC and LC) tables are multivariate, presenting counts on a number of related census topics in a cross tabulation. Of the two, the DC tables, as implied by the name, contain categorical detail within each topic almost as great as that for the KS and QS tables. The DC tables are, however, more restricted in their population scale, typically only providing counts for area types with larger populations. By contrast, the LC tables provide multivariate counts for a full range of population scales but the compromise here is that the categorisation within each topic is less detailed.

A number of different population bases are provided by the Census. The most common basis is the residential population of an area, i.e. those people who usually live in a property located within its boundaries. A further distinction is made between the population who are resident in private households and those resident in institutional establishments such as nursing homes, student halls of residence, military establishments or prisons. Residents in such establishments may be further classified as to whether they are residents as a result of their employment, e.g. they own the establishment (which may also be their home), or they are clients of the establishment. Also there are tabulations that are based on the characteristics of a household reference person, with the choice of which individual in the household who satisfies this definition being left to each household.

After the residential basis the next most commonly used population basis is the workplace population in an area. This is simply the people who are in employment and work in the area. A third population basis is the day time population which is composed of the workplace population plus those who would normally be expected to be at their usual place of residence during the day, e.g. the retired and unemployed. The 2011 census innovated on previous Censuses by collecting information on second residences. This would cover children who live at separate addresses during the week, those who spend time at a holiday home, or those that have a part time, work related, residence.

In addition to the tabulated counts described above, a sub-sample of individual records is provided as an output from the Census. This information is available in either a longitudinal or cross-sectional format. Starting with the 1971 Census, the ONS Longitudinal Study (ONS, 2014f) has been available for individuals in England and Wales. This data set links together records for individuals across a number of Censuses and also links in data from other 'vital statistics' sources, such as births, deaths and a limited range of health events (e.g. information from the Cancer Registry). Cross-sectional samples of individual's records are also available (ONS, 2014g).

4.2.2 Geographic outputs

In the discussion above on data outputs reference is made to the geographic areas on which data are reported. Within the 2011 Census there are a range of geographies that define these areas. Some of these areas are hierarchical, in that one level of geography nests within a higher geography, whilst others are completely separate and be-spoke geographies. The ONS provides a range of products to help navigate these geographies, ranging from simple look-ups between geographies to complete GIS format files (ONS, 2015).

4.2.2.1 OA based statistical geographies

The basic building block of the 2011 Census reporting geographies is the Output Area (OA). The OA is a purely statistical construct devised for the release of Census data and was first used in the 2001 Census - it is not an attempt to define neighbourhoods - but their boundaries do respect certain administrative divisions, e.g. local authority boundaries. In the original design of the OAs the ideal was to have close to 125 households within each OA, with a strict requirement that there should not be fewer than 40 households and, additionally, more than 100 residents (Martin et al., 2001). At the time of the 2011 Census ONS recognised that it was not feasible to maintain all the OAs created at the time of the 2001 Census. As the household population within areas grew and shrank, some OAs became either too large to be of sufficient detail or too small so as to become potentially disclosive. The strategy adopted by ONS was to use 2011 Census returns to split those OAs identified as being too large into two or more new OAs and merge smaller OAs with neighbouring OAs (Cockings et al., 2011). The final set of 2011 Census OAs are therefore largely consistent with the 2001 versions but their populations were also within the desired ranges.

In 2001 the ONS designed a further statistical geography of larger areas on which to report statistics termed Lower Super OAs and Middle Super OAs (LSOA and MSOA). The LSOA geography was constructed by the aggregation of a small number of neighbouring OAs, typically between 4 and 6. Similarly, the MSOA geography was

constructed by the aggregation of between 4 and 6 neighbouring LSOAs. There are 2011 Census equivalents for LSOA and MSOAs.

4.2.2.2 Administrative based geographies

Whilst the OA geographies described above are a statistical creation, designed to remain stable over time, counts and measures based on alternative geographies are also available that reflect the organisation of political and governmental bodies. These areas are more likely to readily identifiable by people but the boundaries of these areas could easily vary over time. These areas are thus less stable than the OAs and this makes statistical comparisons at a local level across Censuses difficult using these administrative geographies.

The most significant administrative geography within England is the local authority district (the LADs). There are single tier unitary authorities who have control over all the functions of a local authority within their area. The alternative is a two-tier structure consisting of an upper tier of county authorities and within each county there are a number of lower tier districts. In this latter case some functions are carried out by the district authority (e.g. housing and planning) whilst others by the county (e.g. education and highways). For organisation and efficiency purposes this authority structure can change through local government reorganisations. The tendency recently has been towards the creation of unitary authorities.

Beyond the local authority there are other larger geographic entities, e.g. the former metropolitan counties or former government office regions, but these have no administrative function and counts on these geographies are largely maintained for continuity and convenience purposes only.

The two remaining important administrative geographies on which 2011 Census data are released are the Parliamentary constituencies (both for Westminster and the devolved Parliaments/Assemblies) and health geographies (see below).

4.2.3 Health geographies

Both the 2001 and 2011 Censuses released counts for administrative areas based on health geographies. In the 2001 Census these were the 28 Strategic Health Authorities (SHA) and 304 Primary Care Organisations (PCO) in England. By the time of the release of 2011 Census data the health geography within England had undergone a major re-organisation (Shapiro, 2013). The PCOs (which latterly became Primary Care Trusts) were completely replaced by a set of 211 Clinical Commissioning Groups (CCGs); and the SHAs replaced by 25 NHS Commissioning Area Teams (NHS England, 2012).

4.2.3.1 NHS commissioning groups

The area covered by each CCG is constructed from an aggregation of the LSOAs that are judged to provide a reasonable coverage of the constituent GP practice's patient lists. Some CCGs span a number of LADs, particularly in rural areas, whilst more urban LADs contain a number of CCGs within their boundaries. There are also a limited number of cases where a CCG has a boundary that does not align with a LAD boundary. The population of these CCGs, based on the ONS 2010 population projections for 2013, ranges from NHS Corby with 67,800 residents through to NHS North, East and West Devon with 901,200 residents (NHS England, 2012). The average population is 260,644.

In regards to 2011 Census data, there is an almost complete release of the Local Characteristics tables from the 2011 Census tables available on the CCG geography. Tabular counts for the 25 NHS Area Commissioning teams are not provided as a standard release by ONS but they may be built by aggregating together the equivalent counts for constituent CCGs.

4.2.3.2 Public health

The health sector in England also has a parallel structure of public health bodies. These bodies have a responsibility to promoted behaviours that encourage healthy outcomes. There are four Public Health Regions, covering northern England, the Midlands and the East of England, southern England and London. The regions have a strategic view of the wider public health system, and work with partner organisations such as NHS England, LADs and health education providers to coordinate activity. Within each region are a number of Centres (15 in total throughout England) that mainly work with LADs by providing guidance and expertise. They also work with other bodies such as education authorities, training organisations and academic institutions to further public health initiatives.

In terms of the delivery of the bulk of public health services, this rests with the LADs (Department of Health, 2012). The role of the authorities is to provide services in the three main areas that form part of the public health remit: health improvement, health protection and public health. They also have responsibility for some treatment services such as sexual health services and alcohol and drug abuse/addition services. Whilst such authorities had previously had some involvement in promoting public health, the newly introduced responsibilities represented a step change in responsibility for the authorities. The Local Government Association (2014) has recently published a report on how successfully these roles have been tackled in the first year and what lessons could be learnt.

4.2.4 Statistical disclosure control

A legislative requirement of the Census is that it should not be possible to discover something previously unknown about an individual simply by examining the Census outputs. This desire to enforce the non-disclosive nature of the Census requires the implementation of a number of techniques to minimise this possibility (Shlomo, 2007). Commonly, these techniques can applied either pre or post-tabular. With the pre-tabular approach the necessary changes to these data to avoid disclosure are carried out at the individual level of data. The quality of these data are purposely degraded to avoid knowledge on one aspect of information within a sub-group of the population (e.g. small ethnic communities or large families) revealing other knowledge. This is commonly achieved by identifying such groups and: altering their data; randomly swapping information amongst this group; or moving households or individuals between neighbouring areas.

Post-tabular techniques make changes to a table after its production but before its publication. One common approach is the small cell adjustment methodology (SCAM) (Stillwell and Duke-Williams, 2007). This methodology requires that any count in a table below a certain threshold should be rounded down to 0 or up to the threshold. In which direction the adjustment is applied is based on a Monte-Carlo probabilistic outcome.

For the 2001 Census a combination of both pre- and post-tabular disclosure control was applied by ONS, however for the 2011 Census only pre-tabular disclosure control was applied.

4.3 2011 CENSUS TABULATIONS

The technique of spatial microsimulation used in this study requires the availability of a constraint population data set that covers the entire population of interest and is available at a useful population scale. The 2011 Census is able to provide such constraint tables, either as univariate (KS or QS) or multivariate (DC or LC) formats. In regards to which format is most suitable, Hermes and Poulsen (2012a) provides a thorough overview of the processes and considerations necessary when carrying out a spatial microsimulation. One of the recommendations is to use multivariate cross classification tables that capture not only the main effects of the significant determinants but also the interactions between these determinants. In the current context for example, gender and age are thought to influence the morbidity of the older population but there may also be an interaction effect, where the impact of a particular morbidity as an individual ages varies by gender. This impact can be captured if some knowledge of the joint gender and age structure of a population is known.

Ideally these tabulations should be provided separately for both the household resident population and the institutionally resident population. This is not always the case, with tabulations of the whole population (all people) or the former group (all people in households) being more common.

4.3.1 Household resident population

The review of literature suggests that the important determinants of health care need amongst the older population are potentially: age, gender, general health, illness, ethnicity, lifestyle, wealth and income. The 2011 Census has good coverage of the first five of these determinants but asks no direct questions on lifestyles (e.g. diet, exercise or smoking); wealth (e.g. assets or savings) or income (e.g. wages or pensions). Proxies are available that may be correlated with these determinants. For lifestyle, counts by marital status, the social grade or highest level of educational qualification may act as a discriminatory measure capturing some aspects of an individual's lifestyle. Tenure of accommodation is a strong indicator of wealth, with an implied positive wealth gradient starting with low wealth for those living in social/private rented accommodation, through mortgaged and onto owned houses. For income, social grade (again) or level of vehicle ownership may be useful proxies but they are not perfect in this regard. For example vehicle ownership also varies with the residential density of an area and the availability of public transport.

4.3.2 Institutional resident population

The nature of the institutionally resident population is complex and diverse. The first consideration is the type of institutional establishment. The 2011 Census provides counts according to a breakdown of types of establishment. Those in medical and care establishments comprise about forty percent of the institutional population, the rest are accommodated in judicial, educational or leisure settings. Within the medical category the significant locations for those aged 50 and older are NHS run hospitals, nursing homes, residential care homes (these two latter separated between local authority and private run homes) and housing association properties. As well as distinguishing by the type of establishment and the age structure of their residents, the Census also provides separate headcounts of staff and families resident in these establishments. The availability of information beyond gender and age counts in the 2011 Census for institutional residents is limited. Some information on general health, long term limiting illness, marital status, and education qualification are available. There is, obviously, no information on household attributes, such as tenure or vehicle ownership.

4.4 ENGLISH LONGITUDINAL STUDY OF AGEING

ELSA was established as a study of the English population aged 50 and older who were resident in a private household during the period 2002 to 2003. The survey is half funded by a consortium of UK Government departments and half by the National Institute on Aging in the US. The purpose of the survey is to examine the lifestyles of the ageing population of England in order to better understand the impact of this changing demographic. This interest is not just limited to physical and psychological health outcomes but involves a range of objective and subjective measures concerning: household and participant demographics, work and pensions, income and assets, housing and social participation.

It is designed as a longitudinal survey in that it attempts to re-contact the same participants through time, measured as survey waves. The aim of the survey is to collect information on complete households where one of the members of the household is aged 50 or older. The core participants, who are followed up at subsequent waves, are those who are aged 50 or older within the household and have provided a baseline survey at their initial contact wave. Information is also collected on the other members of these households but, unless they qualify later as new core members, they are not necessarily followed up at subsequent waves if they move out of the core participant's household.

The first survey wave of ELSA surveys took place from March 2002 to March 2003 and recruited eligible participants from the 1998, 1999 and 2001 HSfE. The interview period for each subsequent wave is given in Table 4-1. Wave 1 data was made available in 2004. Subsequent waves were reported every two years, i.e. 2006, 2008, 2010, 2012 and 2014, currently providing six waves of data.

Table 4-1 : ELSA survey periods

	Survey Period	
	Start date	End date
Wave 1	March 2002	March 2003
Wave 2	June 2004	July 2005
Wave 3	May 2006	August 2007
Wave 4	June 2008	July 2009
Wave 5	July 2010	June 2011
Wave 6	May 2012	June 2013

Source : ELSA Main Reports (Institute for Fiscal Studies, 2015a)

Ideally, on visiting the household the interviewer attempts an interview with each participant in person. The interview is primarily conducted face-to-face and is computer

assisted. There is also a self-completion questionnaire that the participant completes in private, on the same occasion as the face-to-face interview. Where it is not possible to conduct the interview in person (the participant may be in poor health, absent from home or died) then a proxy interview is conducted with a nominated individual.

Periodically the main interview survey is supplemented with other surveys. The most important of these are a nurse visit to collect biomarker data from the participant, e.g. physical body dimensions; health related physical performance; blood and saliva samples. These visits are conducted at even numbered waves for those participants who have consented to such visits, so currently there have been such visits at waves 2, 4 and 6. Aside from these nurse visits, during wave 3 core participants were asked to complete a life history questionnaire that captured significant events in their lives before their enrolment in ELSA (Wood, 2011). This survey asked the participant to provide information on their marriage history, places of residence, working patterns and child rearing. Also information was collected on overall health behaviours, e.g. if and when they started smoking.

4.4.1 Survey size (cross-sectional)

Like any longitudinal survey, ELSA may also be regarded as a series of cross-sectional surveys. Whilst the use of such surveys as cross-sectional data sets is not ideal - they suffer from selection bias (through attrition) and are typically of a smaller size due to the expense of such surveys (Scholes et al, 2009) - they can be of use to analysts. For each ELSA wave, Table 4-2 list the (unweighted) achieved sample size for the wave. Typically each wave consists of around 10,000 interviews.

Table 4-2 : Cross sectional size of ELSA survey at each wave

	In person			By proxy		
	Full interview	Partial interview	Institutional interview	Full interview	Partial interview	Institutional interview
Wave 1	11,709	215	(Note 1)	171	4	(Note 1)
Wave 2	9,265	42	(Note 2)	123	2	(Note 2)
Wave 3	9,431	93	15	199	0	33
Wave 4	10,532	58	14	389	3	54
Wave 5	9,623	102	12	472	5	60
Wave 6	9,923	52	12	550	1	63

Notes:
1 - All wave 1 participants are resident in a household;
2 - At wave 2 there is no distinction between interviews in person or by proxy for institutional residents, but the total number of institutional interviews is 54.

Source : Tabulated from individual wave core data files using each wave's individual outcome variable.

To facilitate the cross-sectional analysis of ELSA, the National Centre for Social Research provides a set of cross-sectional weights to correct for non-survey response at each wave (Cheshire et al, 2012). Since it is the longitudinal nature of ELSA that is of interest here, no more will be said about the cross-sectional nature of the survey.

4.4.2 ELSA composition

To fully understand the size and complexity of longitudinal nature of ELSA requires an understanding of various concepts and contexts. What is presented here is a distillation of these concepts, a fuller picture is available in the various wave Technical Reports (Institute for Fiscal Studies, 2015b).

Within ELSA there are four cohorts. The first cohort was recruited from the 1998, 1999 and 2001 HSfE and participated in ELSA wave 1. The next cohort was a refreshment cohort and they were recruited at wave 3 by virtue of their 2001 to 2004 HSfE participation if they were about to become 50 to 54 years old at wave 3 (this age range ensures that ELSA still represented all ages of 50 and above). A further cohort was recruited at wave 4 from 2006 HSfE participants aged 50 to 51 to refresh the sample and also to top-up the sample using HSfE participants aged 52 to 74. The final cohort was constructed in a similar manner to the second, its purpose to ensure that at wave 6 the age range for ELSA still included a suitable number of those aged 50 to 54.

Each cohort is composed of a range of different types of participant. The most important are the core members. These are participants who are aged 50 years or older, volunteered to be part of ELSA, gave an ELSA baseline interview at their initial wave and were not living in an institution. These people are variously described as cohort 1 core members (C1CM); cohort 3 core members (C3CM); cohort 4 core members (C4CM) or cohort 6 core members (C6CM). Strenuous efforts are made to ensure that Core members are contacted and interviewed at every possible subsequent wave.

Associated with the core members are other ELSA participants who are usually living in the same household as the core member, but are not eligible to be core members. They are termed as core partners, younger partners or core sample members. Core partners are participants who are living with a core member and would be eligible to participate in ELSA but they did not provide an ELSA interview at their first wave, so they are unable to provide a baseline. These participants are termed cohort 1 core partners (C1CP), cohort 3 core partners (C3CP), cohort 4 core partners (C4CP) or cohort 6 core partners (C6CP). Interviews are attempted with them whilst they are still a partner of a core member but if they are no longer a partner of the core member then an attempt at an ‘exit’ interview is made, in person or in proxy, and they are not followed at waves subsequent to this exit. Younger partners are similar to core partners but by virtue of their age (i.e. they are younger than 50 years of age) they do not qualify as a core

member or partner. Like core partners, they are surveyed until they exit from the sample. They are termed C1YP, C3YP, C4YP and C6YP. Core sample members are also similar to core partners but they refuse to participate in any wave of ELSA and only limited information is known about them, e.g. gender, age and year of death. They are termed C1SM, C3SM, C4SM or C6SM.

As well as participants leaving an eligible household, individuals may also join a household that contains a core member. They are considered new partners of these core members. They can be introduced into the ELSA sample at any of the waves and this is designated in their type coding. New partners of cohort 1 core members at wave 1 (i.e. they joined the core member after the core member's HSfE interview but before ELSA wave 1 interview) are termed cohort 1 new partners at wave 1 (C1NP1). Similarly partners who joined a cohort 1 core member after wave 1 but before wave 2 are termed cohort 1 new partner at wave 2 (C1NP2). By extension there are also new partners termed C1NP3; C1NP4; C1NP5 and C1NP6 for subsequent partners of cohort 1 core members. The natural extension to this are the new partners of cohorts 3; 4 and 6 core members, these are termed C3NP3; C3NP4; C3NP5; C3NP6 and C4NP4; C4NP5; C4NP6 and C6NP6.

A type of participant introduced for cohorts 3; 4 and 6 are the older partners of cohort 3; cohort 4 and cohort 6 members, who are designated C3OP; C4OP and C6OP. The final category of individual are those who are non-eligible to be part of ELSA (i.e. they are not a sample member or a partner of a core member) and are not surveyed but like, sample members, there is some limited information available about them. These individuals tend to be of a different, younger, generation to the core member.

It is rare, but possible that participants in ELSA may change their type, e.g. at wave 3 some 104 participants who were classified as C1YP were reclassified as C3CM at subsequent waves since they were, at that stage, 50 years of age or older, in an ELSA household and had a productive baseline interview at wave 3.

The number of participants within each of these categories at wave 6 is given in Table 4-3. The largest group in the survey are the cohort 1 core members. The next biggest group are the sample members at wave 1, those who are resident in a core member's household but refused to take part in the survey.

This rather complex designation structure in ELSA is there to provide flexibility on who should and should not be included in any analysis. In most analysis it is the core members that are of most value, but information on partners may also be of some value, even if this is only provided by virtue of their association with the core member. Also with partners, the distinction between (ordinary) partners, younger or older partners may

also be important. Such complex participation designations are not uncommon in longitudinal surveys. The British Household Panel Survey and its successor, Understanding Society, also have rules concerning different types of panel members and how they are treated for follow up at subsequent waves, e.g. Original Sample Members (OSM) vs Temporary Sample Members (TSM) (Buck and McFall, 2012).

Table 4-3 : Final response status after wave 6

Core Members	
C1CM	11,391
C3CM	1,276
C4CM	2,290
C6CM	826
Core member Partners	
C1CP	366
C3CP	31
C4CP	64
C6CP	28
Core member Younger Partners	
C1YP	1,147
C3YP	569
C4YP	236
C6YP	146
Core member Sample Members	
C1SM	8,148
C3SM	1,228
C4SM	1,380
C6SM	(Note 2)
Not eligible (Note 3)	8,836

Core member Older Partner	
C1OP	(Note 1)
C3OP	325
C4OP	298
C6OP	144
Core member New Partners	
C1NP1	106
C1NP2	58
C1NP3	41
C1NP4	26
C1NP5	14
C1NP6	10
C3NP3	34
C3NP4	18
C3NP5	16
C3NP6	3
C4NP4	24
C4NP5	16
C4NP6	7
C6NP6	10
TOTAL	39,112

Notes:
1 - C1OP are interviewed as Core or Sample Members at wave 1;
2 - Data to derive this count is not available;
3 - Not a sample member or partner.

Source : Tabulated from the wave 0 to 5 index file and the wave 6 core data file using the final status of respondent variable.

4.4.3 ELSA attrition

Table 4-3 provides a picture of the make-up of the ELSA sample in terms of their recruitment into the survey and the type of participation that they provide. Separate to this is the pattern of their participation at each wave.

Reasons for non-participation or response at a wave can be a result of known events such as a death or a move out of England. In such cases an 'exit' interview is attempted, either with the individual concerned or, in the event of death, with a nominated survivor. Other reasons for non-response for a survey member can be a refusal to

participate, protracted illness during the interview period or they become un-contactable. Figure 4-1 below shows the contact pattern for cohort 1 core members (C1CM). In this coding, where there is a numeral in place at position i then this indicates participation at wave i. Where there is an X at position i, then this indicates a non-response at the wave. Thus the pattern 1 2 3 signifies that at wave 3 there are full responses at all three waves, whilst 1 2 3 4 5 6 at wave 6 indicates full responses at waves 1 to 6. The pattern 1 2 3 X at wave 4 signifies a response at waves 1, 2 and 3 but no response at wave 4 and the pattern 1 X 3 X 5 at wave 5 indicates responses at waves 1, 3, and 5 but no responses at waves 2 and 4. In these figures only the largest patterns of response have been labelled and only where space permits.

Figure 4-1 shows that the number of cohort 1 core members who completed an interview at all waves (coded without an X) has decreased gradually from wave to wave so that at wave 6 just over 40% of the original cohort 1 core members have fully participated in all waves. The next most common pattern is of participants contributing up to a particular wave and then not participating subsequently (e.g. there are 1,988 participants who only complete wave 1 and are never interviewed again). There are also a small number of participants who do not provide a responses at a particular wave but do return at a subsequent wave, e.g. 229 people respond at waves 1, 2, 3, and 5 but not at wave 4 (pattern 1 2 3 X 5).

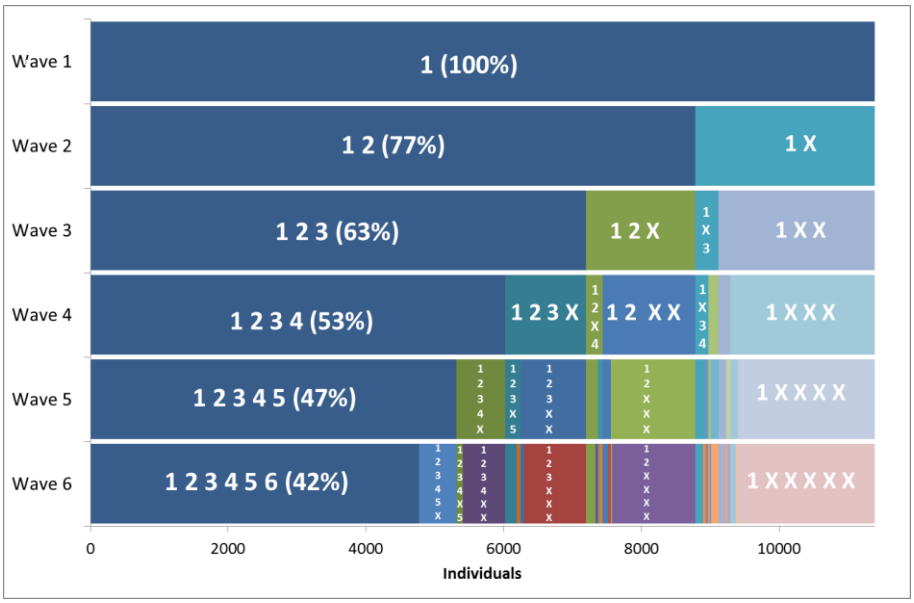


Figure 4-1 : Pattern of participation across six waves by cohort 1 core members

Source : Tabulated from the wave 0 to 5 index file and the wave 6 core data file using the individual outcome variable.

Similar tables and figures are provided for cohort 3 and 4 core members. Figure 4-2 shows this pattern for cohort 3 core members and Figure 4-3 shows this for cohort 4 core members. In Figure 4-2 the retention rate for cohort 3 core members for two

subsequent waves is similar to that in Figure 4-1, 76% provide responses for the first two waves and 67% provide responses for the first three waves, whilst a slightly improved percentage here of 61% provide responses at all four subsequent waves. The retention rate in Figure 4-3 for cohort 4 core members is greater for the two subsequent waves, at 84% and 75%, than the two preceding cohorts.

The National Centre for Social Research provides a set of longitudinal weights at each wave (Cheshire et al, 2012). At wave 6, the purpose of these weights is to ensure that the 42% of cohort 1 core members who responded from waves 1 to 6, when weighted, are representative of all the cohort 1 core members. These weights are therefore able to correct to some degree the impact of attrition across the waves. Such attrition weights are only provided for cohort 1 core members who responded at all waves and as such this weighted sample probably provides the most representative longitudinal sample within ELSA. In the statistical analysis that is reported here, it is these weighted cohort 1 core members who are of most interest.

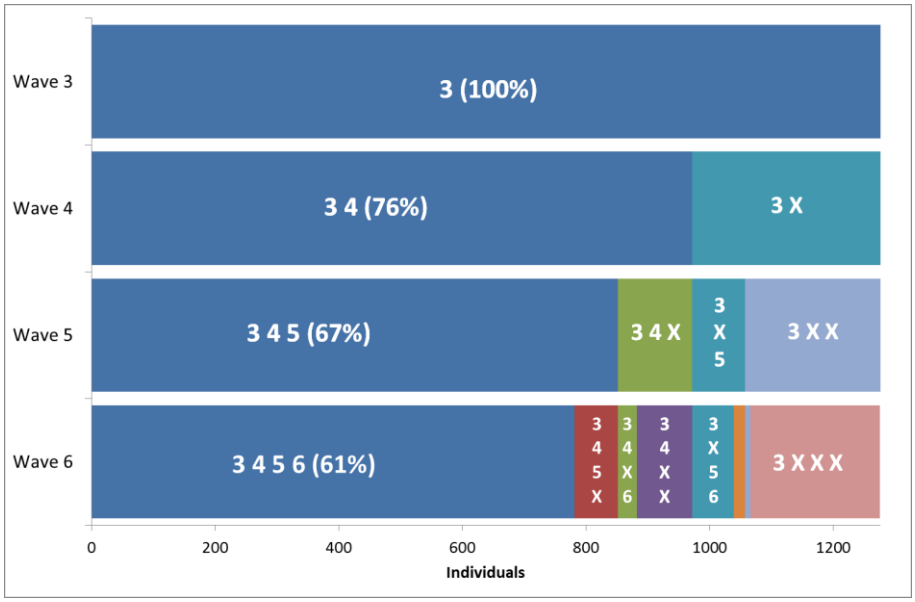


Figure 4-2 : Pattern of participation across four waves by cohort 3 core members

Source : Tabulated from the wave 0 to 5 index file and the wave 6 core data file using the individual outcome variable.

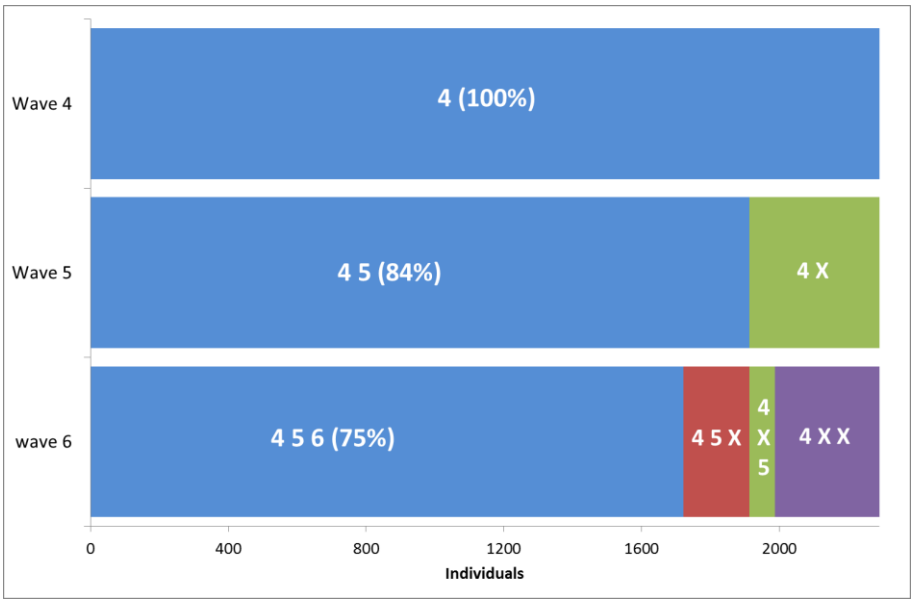


Figure 4-3 : Pattern of participation across three waves by cohort 4 core members

Source : Tabulated from the wave 0 to 5 index file and the wave 6 core data file using the individual outcome variable.

4.4.4 Institutional residents

One condition of recruitment into ELSA is that the participant is resident in a household and not in an institution (e.g. a nursing home). Thus, at wave 1, no cohort 1 core members are living in an institution, and at wave 3, no cohort 3 core members are living in an institution, etcetera for wave 4 and 6 core members. At subsequent waves however core members who have moved to become a resident in an institution are interviewed. In these cases the interview is still conducted in person by the interviewer visiting the institution or, if this is not possible, by proxy with a nominated individual. These interviews are attempted at all subsequent waves even if the respondent is still living in an institution. Non-core members who move into an institution are surveyed one last time as an 'exit' interview.

Whilst ELSA is to be commended for actively collecting data from institutional residents (Moore and Hanratty, 2013), the number of productive interviews from institutional residents is low. Table 4-4 shows the number of responses from institutional residents. The coding here is similar to that above except X now includes a response outside an institutional setting, as well as refusals and ineligibility. A numeral at position i indicates a productive interview with someone in an institutional setting at wave i. All the patterns start with an X at position 1 since people cannot be recruited into ELSA at wave 1 if they are resident in an institution.

Table 4-4 : Pattern of productive institutional interviews

Pattern	Count
X 2 3 4 5 X	1
X 2 3 4 X X	2
X 2 3 X X X	9
X 2 X 4 X X	1
X 2 X X X X	41
X X 3 4 5 6	1
X X 3 4 5 X	1
X X 3 4 X 6	1
X X 3 4 X X	10
X X 3 X 5 X	1
X X 3 X X X	21
X X X 4 5 6	7
X X X 4 5 X	14
X X X 4 X X	30
X X X X 5 6	19
X X X X 5 X	28
X X X X X 6	47
X X X X X X	37,704
Total	37,938

Source : Tabulated from individual wave core data files using each wave's individual outcome variable

There is no one who was interviewed in a household at the first wave and then spent all the subsequent waves in an institution (would be coded as X 2 3 4 5 6). Institutional interviews become more common at later waves.

4.4.5 Health in ELSA

ELSA is a large data set containing over 5,700 variables of original content; content fed forward from the previous waves; and some derived information. About 700 of these questions directly relate to self-reported health status and outcomes. In the health outcome sections of the wave reports that accompany the release of each wave's data, the commonly reported prevalences for a subset of health morbidities that affect the 50 years and older population are provided (Institute for Fiscal Studies, 2012). The morbidities are: coronary heart disease; diabetes; cancer; respiratory illness; arthritis and depression. The definition used for prevalence is, "*Respondents were asked whether a doctor had **ever** told them that they suffered from any of the following conditions: coronary heart disease (angina or myocardial infarction), diabetes, cancer, respiratory illness (asthma or pulmonary disease), arthritis and depression*." (my bold emphasis). The definition of diabetes also includes the presence of high blood sugar and the definition for arthritis includes osteoarthritis, rheumatoid arthritis and other kinds of arthritis.

At wave 1 there is a simple coding of a range of variables that enumerate whether a participant has one or more of these morbidities. For waves 2 to 5 the standard logic used to establish if the respondent has a morbidity uses a combination of the respondent's reply for this morbidity at the previous wave (if such a reply exists) and whether, when prompted, they mention that the morbidity has been newly diagnosed since the previous wave. This is termed 'dependent interviewing' and allows the presence of a morbidity to be established without having to make explicit reference to the participant's history of diagnosis at all previous waves.

There are, however, three exceptions to this rule in the published prevalences. Firstly, for depression no information is fed forward, instead reliance is placed on a single question at each wave as to whether they currently have depression. Secondly, the published figure for diabetes in wave 2, the diagnosis depends on the answer to a single question as to whether the participant has ever been told by a doctor that they have diabetes, with no feed forward information used and no explicit reference to high blood sugars. Finally, for reported wave 6 prevalence rates a participant was recorded as having a morbidity if this information had been feed forward or newly diagnosed at **any** previous wave, not just the previous wave. This requires an explicit reference back to the participant's diagnosis history at all previous waves. The prevalences reported in the ELSA Reports are crude rates and not standardised to any standard population profile (Bains, 1990).

4.4.6 Re-producing ELSA prevalences

Using the ELSA data sets and the logic outline above, it is possible to reproduce the majority of the published prevalences reported in the ELSA wave reports. There are three exceptions:

- For cancer in wave 2, there was an error in the syntax code provided by the ELSA team. This error meant newly diagnosed cancer cases were not counted. Here the error is corrected.
- For diabetes or high blood sugar in wave 2, the ELSA team used the response to the single question asking "*Whether ever been told has diabetes by doctor*". Here instead the wave 2 to 5 logic of feed forward plus newly diagnosed cases is used, and high blood sugar is included in the diagnosis.
- For arthritis in waves 4 and 5, the wave 2 to 5 logic is used but the published rates are not re-producible. There is no available explanation or validation of this finding from the ELSA team.

The differences between the published prevalences and those calculated here are shown in Table 4-5. The Published rows are the prevalences in the appropriate wave report

whilst the Revised rows show the re-calculation using corrected syntax (cancer at wave 2) or the feed-forward plus new cases logic (the rest).

Table 4-5 : Reported ELSA and revised prevalence percentage rates for those waves and morbidities where they are different

Wave and morbidity	Gender	Source	50/52 to 54	55 to 59	60 to 64	65 to 69	70 to 74	75 to 79	80 plus	Total
Wave 2 Diabetes	Males	Published	5.0	6.8	7.3	11.6	13.9	14.4	10.3	9.5
		Revised	5.4	6.5	6.6	10.5	12.2	12.9	8.6	8.7
	Females	Published	2.9	3.9	6.8	7.0	9.8	10.2	9.0	7.0
		Revised	3.2	3.3	6.2	5.9	8.2	7.6	7.4	5.9
Wave 2 Cancer	Males	Published	1.7	2.0	3.7	5.7	6.7	8.3	6.8	4.6
		Revised	2.0	3.1	5.6	8.4	10.3	11.3	9.7	6.8
	Females	Published	3.4	5.7	7.9	6.8	6.1	8.0	7.6	6.6
		Revised	3.9	7.2	9.6	8.4	8.7	10.6	9.0	8.4
Wave 4 Arthritis	Males	Published	15.3	18.1	25.8	27.7	30.8	20.6	38.8	24.9
		Revised	17.5	19.0	28.1	29.9	33.7	33.0	42.3	27.2
	Females	Published	23.3	31.8	38.9	42.2	47.7	52.4	53.1	39.9
		Revised	24.7	32.9	41.1	44.2	50.2	56.2	56.9	42.4
Wave 5 Arthritis	Males	Published	13.5	18.1	27.2	29.0	33.5	35.5	38.1	27.6
		Revised	15.8	20.3	28.8	30.9	35.7	38.0	41.9	30.9
	Females	Published	25.1	30.7	39.0	45.6	49.8	54.2	57.5	43.5
		Revised	26.3	32.3	41.1	47.4	51.4	57.7	60.3	45.8

Source : Published rates are taken from ELSA Main reports (Institute for Fiscal Studies, 2015a) whilst the Revised rates are derived using feed forward and newly reported diagnosis information in each wave's core data file

The newly calculated cancer prevalences are always higher than those published since they now account for new cases since wave 1. There is also an expectation that, since the calculated rates for diabetes at wave 2 now include a diagnosis of high blood sugars, the calculated prevalences would also be higher. This is not the case and may be explained by a fault in the feeding forward of wave 1 information to wave 2 (see next section). The revised arthritis rates are consistently higher than those published.

4.4.7 Feed forward mechanism in ELSA

If participants are consistent in their recollections and the interviewing technology is correctly configured to record these recollections, then the dependent interviewing techniques is ideal. It reduces the recall burden on the interviewee and allows the current health diagnosis status of the participant to be known at any wave by examining just two pieces of information that accompany the wave. Unfortunately circumstances are not always ideal.

There is evidence in the diagnosis pattern for certain morbidities of a systematic failing that has the potential to affect the prevalence rate. Looking at the logic to establish when a participant acquires a morbidity, it is clear that it is not possible for a participant to be

cured of a morbidity (and also consequently, they should not relapse back into the morbidity). However, there are a number of instances where this does actually occur. In particular there is a pattern of participants acquiring a morbidity at wave 1, appearing not to have the morbidity at wave 2, and then re-acquiring the morbidity at subsequent waves. Contact with researchers at the National Centre for Social Research revealed that there were issues with certain morbidities not being fed forward for some participants between waves 1 and 2.

To explore this further, the individual diagnosis pattern for CHD, using the same logic as that used to reproduce the published prevalence rates, is constructed for each participant from waves 1 to 5. Wave 6 is not included in this examination of the feed forward logic since the logic to define the diagnosis of a morbidity changed at wave 6.The pattern adopted to represent the history of diagnosis is similar to that for participation and institutional residence. If a participant has the diagnosis of CHD at wave i, the numeral at position i is set to be the wave number. If they do not have CHD at wave i then the character at position i is N. If they did not participate at wave i then the character at position i is X. Thus a pattern of N N 3 4 5 represents someone who participated in all waves and was first diagnosed with CHD at wave 3. The pattern N X N 4 5 is someone who participated in all waves except wave 2 and was diagnosed with CHD at wave 4. Clearly if the logic of dependant interviewing is correct and the participant is consistent in their responses across waves, a pattern where an N follows a numeral is not possible - once someone has the morbidity then that information should be feed forward to all subsequent waves. Such a pattern where an N follows a numeral reveals a *false negative.* For the CHD diagnosis there are a number of these false negative patterns. The top ten such patterns by their frequency are shown in Table 4-6 along with the total number of participants with such a false negative pattern.

Table 4-6 : Frequency of occurrence for the ten most common false negative patterns

Pattern	Frequency
1 N 3 4 5	144
1 N X X X	72
1 N 3 X X	56
1 N 3 4 X	36
1 N 3 N N	19
1 2 N N N	18
1 N 3 N 5	14
N 2 3 N N	12
1 2 3 N N	10
1 2 N N 5	9
Total	507

Source : Derived using feed forward and newly reported diagnosis information from each wave's core data file.

This table identifies 144 participants who had CHD at wave 1, 'recovered' at wave 2 and then had CHD for the remaining three waves. Looking at these patterns, there is clearly something systematic that is causing a diagnosis at wave 1 of CHD to be forgotten at wave 2 before being recorded at wave 3 and thereafter. There also appears to be other issues at subsequent waves.

The task is therefore to try and reconcile the irreconcilable or gain a true as possible picture of the participant's pattern of morbidity diagnosis. The method adopted here is a modification of that adopted for the wave 6 report. In that method a participant has a diagnosis at wave 6 if at any wave from 1 to 6, the morbidity was fed forward or recorded as a new diagnosis. Of the two pieces of information, the feed forward would appear to be the least reliable, leading to a number of false negatives in the diagnosis. Therefore here the feed forward information is not used. Thus a participant has the morbidity at wave i if they have reported a new diagnosis of the morbidity at any wave up to and including wave i. Using this new definition, all the top ten patterns in Table 4-1, (except the pattern N 2 3 N N) are re-defined as having a CHD diagnosis at all waves (1 2 3 4 5), since the wave 1 diagnosis feeds forward to all subsequent waves. The 12 participants with the pattern N 2 3 N N have a diagnosis of CHD starting at wave 2 and become N 2 3 4 5. This is exactly how the feed forward mechanism should operate.

A calculation of the prevalence rates at wave 6 using the wave 6 report methodology (which takes account of feed forward information at all waves) and the methodology adopted here (which ignores all feed forward information) shows that only for CHD is there a difference in the rates at wave 6. These differences are reported in Table 4-7.

Table 4-7 : Comparison of ELSA wave 6 prevalences and those calculated using the revised methodology

Wave and morbidity	Gender	Source	50-54	55-59	60-64	65-69	70-74	75-79	80+	Total
Wave 6 CHD	Men	ELSA	2.2	8.2	11.9	18.1	24.3	25.9	38.5	15.7
		Revised	2.2	5.7	9.3	14.3	19.1	19.7	31.2	12.4
	Females	ELSA	1.3	3.4	5.8	9.1	13.7	21.1	29.6	10.8
		Revised	1.3	2.4	4.1	6.5	10.2	15.3	25.0	8.4

Source : ELSA rates are taken from ELSA wave 6 main report (The Institute for Fiscal Studies, 2015) whilst the Revised rates are derived using only newly reported diagnosis information in each wave's core data file

Incorporating the fed forward information for CHD shows that there are also some *false positives* at previous waves. An examination of instances where there is a record of CHD being fed forward from the previous wave but there is no evidence of CHD at any previous wave (either as new diagnosis or feed forward) reveals issues at wave 5. Table

4-8 shows that there are nearly 240 participants who have CHD fed forward at the two later waves (5 and 6) but have no previous diagnosis of CHD.

Table 4-8 : Instances where diagnosis information is falsely fed forward

Circumstance	Count
Fed forward at wave 5 but no new diagnosis at waves 1 to 4	195
Fed forward at waves 5 and 6 but no new diagnosis at waves 1 to 5	40
Fed forward at wave 6 but no new diagnosis at waves 1 to 5	4

Source : Derived from the waves diagnosis fed forward and newly diagnosed variables in the core data files.

Given the ability to reproduce the published wave 6 prevalences for the other four morbidities implies that there are no issues of false positives with these other morbidities.

Correspondence with those responsible for producing the wave 6 prevalences on this issue has once again raised concerns over the accuracy of the feed forward information, both in generating false negatives and false positives. This re-enforces the rationale behind ignoring the feed forward information in establishing the diagnosis of a morbidity. More generally it emphasises the importance of understanding these data and the generating mechanisms within these data and being able to validate the operation (or otherwise) of the mechanism.

4.4.8 Geography in ELSA

The standard version of ELSA data only identifies the (now abolished) Government Office Region of residence for the participant. The reason for this is to avoid inadvertently disclosing the identity of the participant through a combination of geographic and demographic information. It is however possible to make a special request for a subset of geographic information that can be linked to the participant without necessarily disclosing their identity. The level of disclosure risk is classified from Low, to High and onto Very High. The example provided of Very Highly disclosive geographic information is the OA in which the participant lives. Examples of Highly disclosive information are a post code sector; the 2003 statistical ward of residence and the index of multiple deprivation for the LSOA in which the participant lives. Such data are not commonly made available to researchers outside a secure setting. Information that is of low disclosivity can, however, be requested. This information includes the local authority of residence; quintiles of postcode population density; quintile of index of deprivation and type of area (urban or rural).

4.5 COMPATIBILITY OF 2011 CENSUS AND ELSA WAVE 5

As part of this study it is necessary to link together information from the 2011 Census and ELSA to carry out a spatial microsimulation. This is achieved by defining a population base and identifying information that is in common to both surveys. There is some merit in trying to replicate a population at a household level, especially when the composition of such households can have an important impact on matters such as wealth, income and health. As mentioned earlier there are some tabulations from the census that describe each LAD in terms of the characteristics of its households, for example table DC4201 provides a count of "Tenure by ethnic group by age - Household Reference Persons" however this count is just of the ethic group of the household reference person, no other family members are included in this table. Census tables on "Families" also have issues in that they do not allow the segmentation of the 50 year and older population and they are primarily concerned with the characterisation of dependent children within the family. Even if it were possible to usefully construct a picture of a LAD's households on all the potentially useful characteristics there may be issues with constructing and equivalent study population from within ELSA. Recall that whilst ELSA is in essence a household survey it is possible that not everyone in the household will be surveyed to the required degree. In particular there are sample members (SM) within households, for whom very little information is known, and it would be difficult to use such members to complete a household with some desired characteristics. These issues on both the Census and ELSA side lead to a move toward a methodology that matched individuals rather than households between the two data sets.

For this association of individual information between both the Census and ELSA to be successful it is important that the degree of compatibility in how these attributes are recorded in each survey is understood. Usefully, age and gender are defined in the same way in both data sets.

The Census contains two questions directly related to health: an assessment of general health and the presence of a limiting long term illness. Both the Census and ELSA use a 5-point scale for rating general health, but the descriptive text associated with the scales are different. The Census allows the individual to choose one of: very good; good; fair; bad; or very bad, in rating their general health whilst the ELSA participant has a choice of: excellent; very good; good; fair; or poor. The obvious approach to harmonising these two scales is by merging together adjacent categories but is not always straight forward (Smith and White, 2009).

There is also a discrepancy in how information on the presence of a limiting long term illness is recorded in each data set. The Census has three possible responses: limited a lot; limited a little and not limited, whilst ELSA asks a two part question - if they have a

long term illness and, if so, is this illness limiting? Again there are issues on how to harmonise these two questions. Clearly those with no limiting long term illness or those whose illness limits them a lot are easy to reconcile between the two surveys. However, should those who describe their illness as limiting them a little in the Census be classed as someone with or without a long term limiting illness in the ELSA sense? Here, there is an assumption that even though the reported limitation in the Census is not severe it is still a limitation and as such is grouped with those who are limited a lot.

The Census has a broad list of sixteen different ethnic groups. The range of ethnic groups in ELSA is, however, more limited. Each wave contains an indication as to whether the participant's ethnicity is white or non-white (i.e. Black or Minority Ethnic, BME). Wave 5 however, contains feed forward information on ethnicity with a broader range of 7 classes: White, Mixed, Black, Black British, Asian, Asian British and any other group. The wave 5 questionnaire document (NATCEN, 2012) clarifies further that Asian covers Indian; Pakistani; Bangladeshi and African-Indian whilst 'Any other' includes Chinese, Japanese, Philippino and Vietnamese. It is therefore possible to reconcile the ethnicity coding of the two data sets to a consistent five class ethnicity classification of: White, Mixed, Black, Asian or Other (including Chinese).

Education is categorised using slightly different philosophies in the two data sets. The Census defines a set of four levels of highest education plus having no qualifications and an 'other' category. Actual qualifications are split between levels, e.g. having 1 to 4 O levels/GCSEs is Level 1 whilst having five or more is Level 2. In ELSA, however, the highest education is solely determined by the type of qualification, e.g. having any number of O level/GCSE qualifications yields the same highest qualification categorisation. Thus categories of qualification as defined in ELSA span categories as defined in the Census. However, it is possible to reconcile these categories into one of four categories: no qualifications, qualification below degree level, degree level or higher qualification and foreign/other qualification.

The National Statistics socio-economic classification (NS-SEC) is a well-defined and utilised categorisation of an individual based on their current or past employment. The Census publishes counts according to the 8-class scheme with class 1 (Higher managerial, administrative and professional occupations) further divided into two sub classes and the 8th class (never worked and full time unemployed) split into three sub classes. ELSA has NS-SEC coded as the long form 'Operational categories and sub-categories classes' and from this the 8-; 5-; and 3-class 'Analytical classes' are derived. Therefore there is a good agreement for this attribute in both data sets.

In terms of the provision of unpaid care to other people, the Census categorises the amount of such care by its duration, using categories of 1 to 19 hours; 20 to 49 hours

and more than 50 hours, plus a count of those who do not provide any care. In ELSA the amount of unpaid care is specified as an explicit amount in hours and thus can easily be categorised in accordance with the Census definitions.

Within the Census marital status or living arrangements are defined using a combination of information collected about the relationships of individuals within the household and their individually declared marital status. Individuals are counted as either being in a couple or not. If they are in a couple, then their marital is status is used to define whether they are married or cohabiting. If they are not in a couple then the marital status is again used to decide if they are single (never married), married, separated, divorced or widowed (each with civil partner equivalents). ELSA also has information on the living arrangements of participants, whether they are living as a married couple, cohabiting or neither of these, and also information on their marital status. It is therefore possible to define an ELSA participant's living arrangements in a manner consistent with the Census definition.

Finally, ownership or having access to a vehicle from within the household is a categorised variable in the Census. Depending on the tabulation, this can be five categories that range from none to four or more cars or van; to a more limited range of: no cars or vans to a highest category of two or more cars or vans. In ELSA each vehicle owned by the household is separately identified and such vehicles may include motorcycles (but not motorised scooters) in addition to cars and vans. An important distinction in ELSA is that a vehicle owned by someone outside the interviewee's household is only counted if the interviewee's household is the sole or main user of that vehicle. This is clearly a more restricted definition of use than the one employed by the Census. None the less, in ELSA it is possible to count the number of cars or vans owned by the household and categorise them according to whatever tabulation is needed.

4.6 POPULATION PROJECTIONS BY GENDER, AGE AND ETHNICITY

In the review chapter a number of studies and reports were highlighted that indicated a difference in prevalence rates for some morbidities by ethnicity. If this is the case, it is highly likely that the prevalence rate in an area for these morbidities will be heavily influenced by the size of the ethnic population within the area.

The ONS only produces sub national population projections for LADs by gender and age. Alternative projections are, however, available that do provide a breakdown of projected local authority populations by ethnicity. One such projection is from the ESRC project funded under the Understanding Population Trends and Processes programme, UPTAP. These projections were subsequently made available as ETHPOP projections (Rees et al, 2011).

The methodology employed uses a bi-region cohort component model to produce, for each English LAD, gender specific, single year of age population projections, up to age 100 (plus a 100 years and older band), for 16 ethnic groups. Four scenario projections are provided, one that assumes that trends in the components of population change established during the period 2002-07 continue to 2051. A second scenario aligns the trends as much as possible to those assumed in the ONS 2008 national population projection methodology. The two remaining scenarios are those that are most distinctive to the work. Both use specifically derived ethnic estimates for the various components, but the difference between the two scenarios is that one variant treats emigration as a flow whilst the alternative treats it as a proportion of the 'population at risk'. There are thus four possible projections arising out of the work and some consideration is required into which projection or combination of projections is most suitable.

4.7 CONCLUSION

This chapter has provided the background and some depth of knowledge on three data sets that are of particular relevance to this study in order to satisfy the second objective of this thesis. The 2011 Census is seen to be a representative and comprehensive enumeration of the English population with information being provided for a wide range of geographies. It is representative in that it attempts to capture details on the entire population of an area. It is informative in the range of tabulated counts that are provided, both in detailed univariate tables and multivariate cross-tabulations. However, it is unable to provide all the detail that is required for this study. It is primarily published in the form of aggregate counts of the population making it difficult, beyond a few dimensions, link the behaviour by patterns of characteristics in the population. Also, of particular relevance to this study is that it does not gather detailed information on health outcomes in the population; in particular no information is provided for actual morbidity status.

ELSA is, however, able to provide a wide range of information on the population of interest at the disaggregate level of the individual. Since it is an individual survey it is possible to tie together information on the same individual and because of its longitudinal nature it is possible to study the evolution of these ties over time. There are, however, issues with ELSA that also mean it is unable to provide all the information requited. Firstly, there is coverage. ELSA is a national survey consisting of around 16,000 core members, and, whilst this appears to be a large survey it is not in itself of sufficient size to provide reliable prevalence rate estimates at LAD level. Secondly, the composition of the survey is complex, possessing a range of different categories of participant, from qualifying core members to ancillary sample members. In any analysis careful though is required into which categories of membership to include and what

trade-offs are consequential on this decision. The third issue is inherent in all longitudinal surveys and concerns the commitment of the participants. Do they provide responses at all waves or do they dip in and out of the survey? The final issue is around the consistency of the information provided when it is correlated across waves. These inconsistencies may be a result of the interviewee being inaccurate or mistaken in their knowledge or recollection. In such circumstances it is difficult to identify the true situation. The related issue is when inconsistency is introduced as part of the survey process. Just looking at one aspect of the survey, self-reported diagnosis of six morbidities, reveals evidence of system generated false negatives and false positives. This concern over the quality of diagnosis in health related databases is common (see Khan et al. (2010) and Herrett et al. (2009) for similar concerns in a general practitioner administrative database). Fortunately it is possible to identify such cases using the longitudinal nature of ELSA and use a revised logic that corrects for the inconsistency.

In gaining an understanding of future trends in the prevalence of a morbidity in an area, after age and gender, it is perhaps the ethnic makeup of an area that will have a sizeable impact. It is therefore useful to know what the ethnic composition of an area is likely to be in order to accurately account for this demographic effect.

The following Methodology chapter illustrates how these three data sets interact within a modelling framework to produce a set of prevalence forecasts in the 50 years and older population in each English LAD from 2013 to 2031.

5 METHODOLOGY

A Combined Microsimulation and Macrosimulation Approach

5.1 INTRODUCTION

This chapter introduces the overall modelling methodology that is used to produce the forecasts of the number of older people in each English LAD with a range of morbidities to the year 2031. The third objective in section 1.5 requires that this methodology be identified and outlined.

The methodology is composed of a number of components and each component is described in turn. The first component is to obtain a spatially relevant individual microsimulated population of each LAD. This population is defined by a LAD specific, re-weighting of a sample of the wave 5 ELSA participants. The next component is to evolve each of these LAD populations through time by ageing the population and adapting the weights to accord with an externally derived aggregate population projection. As the population is evolved, the morbidity status of individuals is updated in accordance with their individual characteristics. In this model the demographic dynamics are captured through the initial spatial microsimulation and the adaptation of individual weights over time, whilst the health dynamics are captured through the updating of individual's health status over time.

This model is able to forecast any discrete health outcomes and for illustrative purposes the following three morbidities of: CVD, DHBS and respiratory illness are chosen. CVD is indicated by a doctor diagnosis of a heart attack, angina or a stroke whilst respiratory illness results from a diagnosis of lung disease or asthma. Data on the causes of death (ONS, 2014d) shows that nearly 30% of all deaths of those aged 50 and over have an underlying cause related to diseases of the circulatory system (including CVD) and 15% are attributable to diseases of the respiratory system. A much lower 2% die from diseases related to endocrine, nutritional and metabolic causes, which includes DHBS.

The resulting health status of the population can be gauged using a number of measures. For health planning purposes it is the actual prevalence count within the population that have a particular morbidity, possibly categorised by some characteristic of the population. However, in comparing the relative health of the population between each LAD, the prevalence rate is more insightful. Using the approach proposed in this study, it is possible to provide both these measures.

The overall structure of the modelling methodology is presented in Figure 5-1. The rectangular sharp angle cornered boxes represent processes, e.g. a statistical model estimation (the purple boxes) or a computation (the blue boxes), whilst the rounded corner rectangles represent a data source. The outputs from the models are shown in the angular boxes, some of which are intermediate. The layered angular boxes represent different 'versions' of the intermediate output. Surrounding a number of the elements in this figure is a thick dashed outline, this compartmentalises the components of the model and identifies which of the chapters in this thesis report the results from this component.

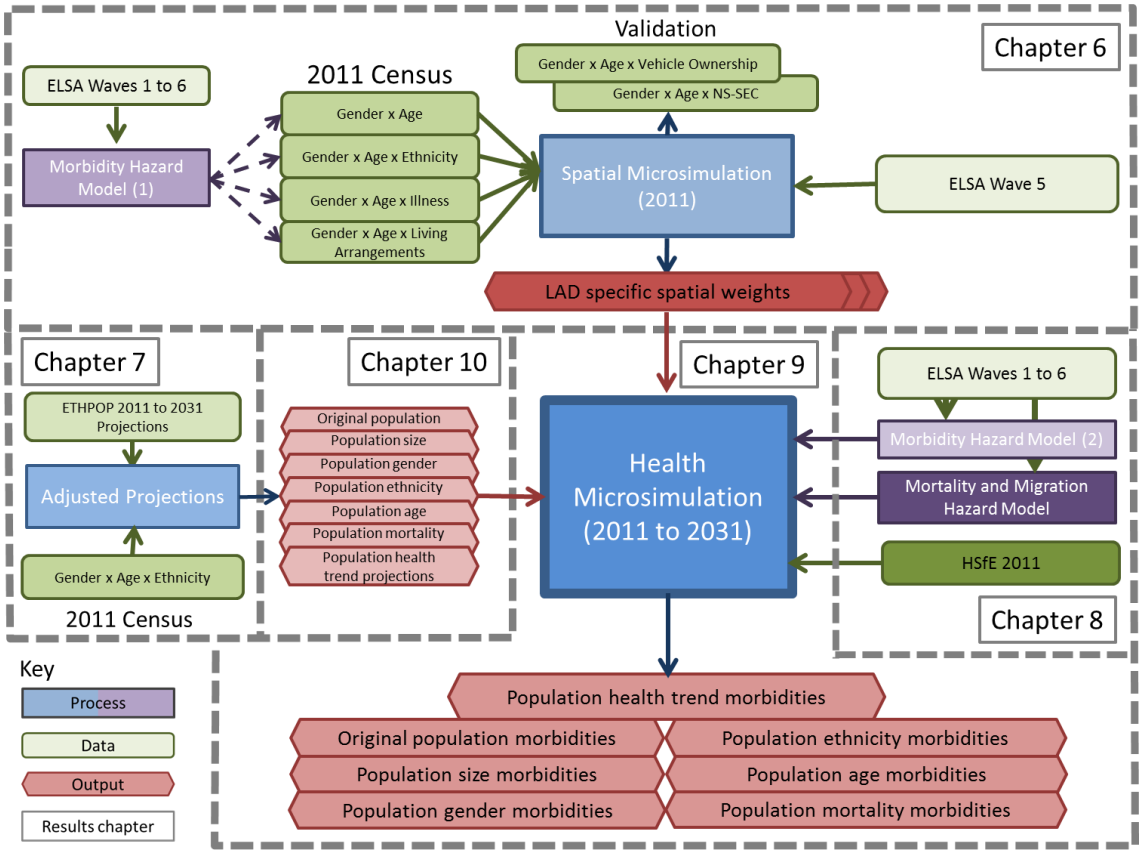


Figure 5-1 : Modelling methodology summary

The structure of this chapter mirrors the components of this figure. Following on from this introduction, the definition of a hazard model is provided in section 5.2 and an explanation is given on how these models are used in the modelling process (the results are to be found in Chapter 8). Section 5.3 describes the spatial microsimulation that re-weights the ELSA data for each LAD to produce the spatially relevant populations in each LAD in 2011 (the results of this are reported in Chapter 6). In addition to having a spatially relevant population for each LAD there is also the need for a population projection in each LAD to 2031. This is provided by applying revisions to an external population projection and this revision process is described in section 5.4 (Chapter 7 describes the results of this process). Section 5.5 explains the mechanisms that bring

together the hazard models, the spatially relevant LAD populations and the population projections to produce the morbidity forecasts (see Chapter 9 for the results). Section 5.6 presents a way to decompose the overall trends in morbidity outcomes by various demographic and health trends (see Chapter 10 for the results). The chapter finishes with section 5.7 on the strategy for presenting the results and section 5.8 which provides a summary of the methodology and its rationale

5.2 HAZARD MODELS

In Figure 5-1 hazard models are used for a number of purposes: to identify significant constraint variables for the spatial microsimulation, to predict the acquisition of a morbidity and to predict how likely an individual is to die or migrate. In this section a general overview of the utility and fitting of such models is provided (also see Singer and Willett, 2003).

5.2.1 Functional form of hazard models

Discrete-time hazard models are estimated on individual level data where the outcome of interest is typically a binary event. Such models are often used in the context of longitudinal data analysis where: individuals are present in a data set for a number of time periods or waves; and the event is 'absorptive', i.e. once an event has occurred it cannot be undone (e.g. death or, as discussed in section 4.4.5 in the context of ELSA, the acquisition of a morbidity).

Using such models it is possible to estimate the impact and significance of other variables on the outcome and estimate a probability that the outcome occurs. This is achieved using the concept of 'logits'. A logit is not dis-similar to traditional ordinary least squares regression in that the logit for an individual is expressed as a linear combination of a number of explanatory variables.

$$\text{logit}_i = \log_e\left(\frac{\text{probability}_i}{1 - \text{probability}_i}\right) = \alpha + \beta_1 x_{1,i} + \beta_2 x_{2,i} + \ldots + \beta_k x_{k,i} + \varepsilon_i \qquad 5\text{-}1$$

where logit_i is the estimated logit for individual i;
probability_i is the probability that the event occurs for individual i;
$\alpha, \beta_1, \beta_2, \ldots, \beta_k$, are parameters to be estimated;
$x_{1,i}, x_{2,i}, \ldots, x_{3,i}$ are explanatory variables associated with individual i;
ε_i is an error term for individual i;
$\log_e$ is the natural logarithm.

The explanatory variables may be a continuous measure (e.g. income) or categorical (e.g. male/female). Depending on the assumptions concerning the distribution of the error term, ε_i, these logits can then be transformed into other, more meaningful, measures such as odds or probabilities. The odds measure the ratio between the

probabilities of the event occurring verses the probability that it does. More straightforwardly, the probability represents the likelihood that the event occurs. If the error term is assumed to follow a logistic distribution then the logit or probability may be transformed into the odds as follows:

$$\text{odds}_i = e^{\text{logit}_i} = \frac{\text{probability}_i}{1\text{-probability}_i} \qquad 5\text{-}2$$

and the probability can be expressed using either odds or logits.:

$$\text{probability}_i = \frac{\text{odds}_i}{1+\text{odds}_i} = \frac{e^{\text{logit}_i}}{1+e^{\text{logit}_i}} = \frac{1}{1+e^{-\text{logit}_i}} \qquad 5\text{-}3$$

The values of logit_i can vary from minus infinity to plus infinity, the odds_i are strictly positive whilst the probability is constrained to lie between zero and plus one. These relationships can be seen in Figure 5-2 which shows the form of these transformations based on the logit function with a single explanatory variable, $x_{1,i}$.

$$\text{logit}_i = -1.0 + 1.0\, x_{1,i} \qquad 5\text{-}4$$

It is clear that as $x_{1,i}$ increases in value so do all the measures, although in a non-linear fashion. Most notable, low values of logits are associated with low probabilities whilst high values are associated with high probabilities.

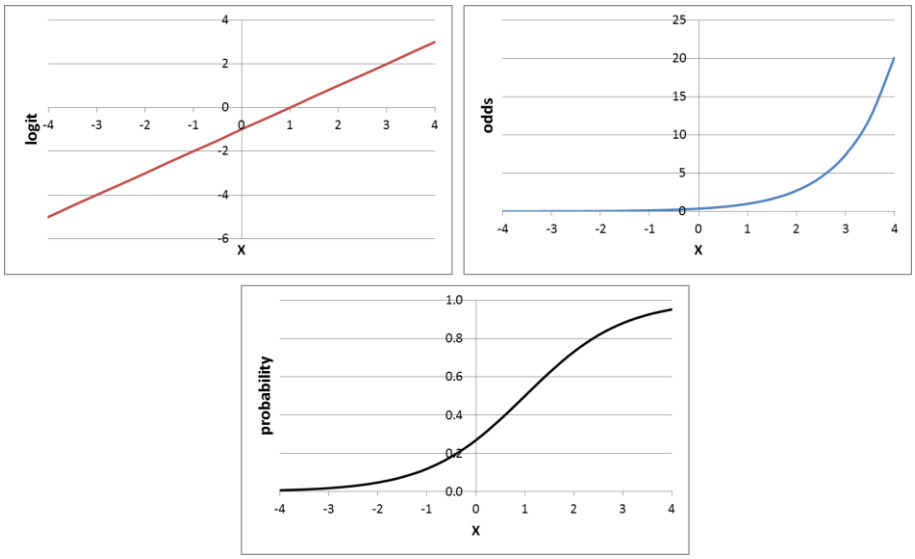


Figure 5-2 : Graphical relationship between logits, odds and probabilities

5.2.2 Estimation of hazard models

The estimation of discrete time logit hazard models is relatively straight forward in most statistical packages once these data are formatted correctly. Typically, these data should be arranged in a person-period format (see Box 5-1) rather than the alternative person-level format.

Once these data have been arranged in this format the models can be estimated using the standard logistic regression algorithms found in most statistical software.

Although the formulation of the hazard model in logits is linear, the standard ordinary least squares regression equations cannot be used to estimate the parameters of the model. Instead a maximum likelihood estimation technique is often used.

This technique formulates the model as a likelihood function that measures how likely the observed outcomes are, given a set of parameter estimates. The estimation algorithm optimises the choice of these parameters so as to achieve the greatest likelihood of generating the observed outcomes. The value of this likelihood is termed the model's maximum likelihood value and is useful in comparing the fit of alternative models.

Box 5-1 Person-Period Format

A person-period dataset has these data arranged so that each row represents the information on each person for one time period. Information on the same person is usually linked together using an identification field. An example of such a dataset is given below. In this example, individual 23 contributed at waves 1 and 2 and was alive at the time of the wave 2 interview. Individual 24 only contributed at wave 1. Finally, individual 28 contributed to waves 1, 2, 3 and 5, but at wave 5 it was found they died aged 69.

id	wave	gender	age	ethnicity	Dead?
23	1	Male	54	White	No
23	2	Male	56	White	No
24	1	Female	62	Black	No
28	1	Female	61	Other	No
28	2	Female	63	Other	No
28	4	Female	68	Other	No
28	5	Female	69	Other	Yes

Since hazard models are commonly used to model the absorptive transition from one state to another (i.e. there is no possibly of returning to the earlier state e.g. alive to dead) these data are normally censored after the occurrence of the event, i.e. there is no information to be gained from knowing that someone persisted in the state. In a health context this is akin to using the hazard model to predict the incidence of a morbidity, i.e. the first time occurrence of the morbidity rather than. Of course the two are related, accumulated incidence measures prevalence (which is exploited here as described later in section 5.5.3).

5.2.3 Goodness of fit for hazard models

In comparing the fit of ordinary least squares regression models use is commonly made of t-ratios, R^2_{adj} values and ANOVA tables. With hazard models such comparisons are

not so straightforward. Whilst the traditional interpretation of t-ratios from hazard models is possible, there is no direct equivalent of an R^2_{adj} or the ANOVA.

A method for comparing the goodness of fit of two alternative models is however possible using a function of the maximum likelihood value obtained from the fitted model. For this comparison to be valid both models must be estimated using the same data set (i.e. they must be use the same observations). The models must also be nested, in the sense that one model can be obtained from the other by setting a sub-set of the parameters to zero.

Box 5-2 : Test for improvement of fit

Suppose model A includes gender and age as explanatory variables and its maximum log likelihood value is -245.0. Model B is then estimated that includes ethnicity as a two category variable and its maximum log likelihood is -242.0. The respective deviances are 490.0 and 484.0, with a difference of 6.0. The extra number of parameters in model B is one, and the appropriate critical values are $\chi^2_1(10\%)=2.706$; $\chi^2_1(5\%)=3.841$ and $\chi^2_1(1\%)=6.635$. Therefore at the 10% and 5% levels model B is preferable to model A whilst at the 1% level model A is preferable.

The first task in conducting this test is to convert the likelihood value into a deviance value by taking its logarithm and multiplying this by -2.0. The difference in this deviance value between the two models is then calculated. This value follows a χ^2 distribution with the degrees of freedom equal to the difference in the number of parameters in each model. If the observed value of the deviance statistic is greater than the critical value of this distribution then the model with the extra parameters is deemed to be a better model in that it reduces the deviance in proportion to the change in the number of parameters. An example of this test is provided in Box 5-2.

A further assessment of the accuracy of prediction of an individual model is the Homser-Lemshow test (Homser and Lemshow, 1980). This test works by firstly using the estimated model to predict the probably of the event occurring. The data set is then sorted in increasing order by these probabilities. Once sorted the data set is then divided into an equal number of groups, usually ten groups. The first group contains the 10% of observations that are least likely to have occurred, whilst the top group contains the 10% most likely. The estimated probabilities in each group are then summed and this can be viewed as the expected number of occurrences within the group. A count is also made of the actual number of occurrences observed in each group to obtain the observed number of occurrences. A χ^2 test is then conducted to see if there is a significant difference between the expected and observed counts across the ten groups. In a model that fits well, there should be no significant difference in these counts.

5.2.4 Choice of constraint variables

Anderson, 2007 outlined four criteria to be used in the selection of appropriate constraint variables for use in his spatial microsimulation of household incomes. The four criteria are:

1. Compatibility of definition in the constraint and sample populations;
2. Be available at the unit of analysis (in his case, households);
3. Known to be reasonable predictors at the small area level;
4. Good predictors at the level of the household.

A judgement on the first point is largely mechanistic and can be made after studying the definitions of the candidate variables in both populations. The second point can be addressed by excluding those constraints that are not consistent with the proposed unit of analysis, here individuals, rather than households. Point three is informed by a review of the literature.

For the fourth point Anderson used logistic regression techniques to identify which combination of potential constraint variables best predicted the outcome of interest in his sample population. A stepwise fitting procedure was adopted which orders these candidate constraints by increasing usefulness

In this study, a similar approach is adopted. Issues around points 1 and 2 are largely addressed in Section 4-5. The literature review chapter also provides some insight into what characteristics are strong determinants. What remains is the adoption of his logistic regression modelling approach to identity the important set of constraint variables. In this study this is done by fitting hazard models to waves 1 to 6 of ELSA data, using the absence and presence of a morbidity as the outcome of interest. Here the criteria used for assessing the contribution of each constraint to the goodness of fit is measured by the decrease in the deviance and also the Hosmer-Lemshow test.

5.2.5 Morbidity hazard

During the health microsimulation stage of the model there is a need to update an individual's morbidity status to simulate the acquisition of the morbidity. This updating takes place using hazard models. To an extent such models are not dis-similar to those that identified which constraints to use in the spatial microsimulation but are different in two respects. Firstly, the range of constraints available at this stage of the health microsimulation are more restricted than those at the spatial microsimulation stage. For example the NS-SEC of an individual is known in 2011 in both the Census and wave 5 of ELSA, however in future such information about the individual may have changed or be unknown. Thus NS-SEC may be used as a constraint to estimate the 2011 spatially relevant population but may not be used to predict future morbidity diagnoses.

Secondly, the purpose of these models is predictive whilst the models to identify significant constraints are explanatory. There is therefore greater emphasis on the predictive power of such models here since they are used to estimate a probability of morbidity diagnosis for an individual.

5.2.6 Mortality and migration hazard

As described below, during the health microsimulation it is necessary to select and de-select individuals so as to achieve a desired population demographic. Given a number of potential candidates for selection or de-selection the question arises as to who to select. The most straight forward approach is to perform the selection or de-selection with equal weight. An alternative is to vary this weight according to some characteristics of the individual. To do this it is necessary to estimate a probability and this is done using hazard models. Hazard models are required to predict the probability of death (a de-selection outcome) and to predict the probability of migration (both a selection and de-selection outcome).

5.3 SPATIAL MICROSIMULATION

The top half of Figure 5-1 shows the spatial microsimulation modelling. The purpose of the spatial microsimulation is to estimate LAD specific weights for individuals within the wave 5 ELSA sample population so that the aggregate counts for the constraint variables in the weighted sample accurately reproduce the equivalent LAD specific 2011 Census counts in a range of constraint tables. Wave 5 of ELSA is chosen in preference to the other waves since it is the most contemporaneous with the 2011 Census. There are a number of techniques available to estimate the weights in this 9,987 by 324 table of weights (here there are 9,987 individuals within the sample population and 324 LADs), the two most common are iterative proportional fitting (IPF) and combinatorial optimisation (CO).

IPF is a long established methodology for adjusting a set of weights to meet a series of constraints (Deming and Stephan, 1940 and Zaloznik, 2011) and has been adapted for application in the field of spatial microsimulation (Ballas et al., 2007). Its name is fairly descriptive in that it is an iterative approach that requires a number of passes through these data to achieve some level of convergence. Each pass operates on a number of tables (e.g. gender, ethnicity) and categories within these tables (e.g. male/female; white/BME). The convergence is improved at each step by proportionally adjusting selected weights in the sample population so that aggregated weighted counts agree with constraint counts (equation 5-5).

$$\left[\text{weight}_{i+1}(n)=\text{weight}_i(n) \times \frac{\text{table(t,c)}}{\sum_{n\in t,c} \text{weight}_i(n)}\right] \forall\ t\in T, c\in C(t) \qquad 5\text{-}5$$

where $\text{weight}_i(n)$ is the weight at iteration i for individual n who has attribute c in constraint table t;
t is the constraint table;
c is the category within the constraint table;
table(t,c) is the count of constraint c in table t;
T is the set of tables;
C(t) is the set of constraints in table t.

These adjustments are conducted in sequence for all the constraints C(t) in all the tables T. The iterations stop when the adjustments that are applied to successive weights are small or, equivalently, that some threshold measure of overall fit between aggregate weights and tabulated counts is achieved.

The alternative combinatorial optimisation approach re-formats the task as a optimisation problem (Voas and Williamson, 2000). An objective function is defined that measures how well the aggregate counts in a weighted sample population agree with constraint counts. Commonly the Total Absolute Error (TAE) formulation is chosen as the form of this objective function and for each constraint table, the absolute value of the difference between the aggregate counts and the tabulated counts is used as a measure of goodness of fit (equation 5-6).

$$TAE_i = \sum_{t\in T, c\in C(t)} \left| \left\{ \sum_{n\in t,c} weight_i(n) \right\} - table(t,c) \right| \qquad 5\text{-}6$$

where TAE_i is the absolute error at iteration i;
$\text{weight}_i(n)$ is the weight at iteration i for individual n who has attribute c in constraint table t;
t is the constraint table;
c is the category within the constraint table;
T is the set of tables;
C(t) is the set of constraints in table t.

The outer sum ranges over all tables (T) and constraints within a given table, t. The inner sum aggregates the weights for those individuals who have the same attribute as the constraint c in table t. The objective is to minimise the value of TAE to zero, or as close to zero as possible.

In this study, the combinatorial optimisation approach has been adopted. The reasons behind this choice are given below.

Integer sample weights required. This is perhaps the most important consideration. The health microsimulation operates by selecting and de-selecting individuals to achieve a desired demographic profile for an area. This is best achieved if individuals are represented as complete whole units in the weighted sample. The operation of IPF produces non-integer weights whilst CO produces integer weights. Of course, various schemes exist to integerise these non-integer ipf weights (Ballas et al., 2006) but this post processing has the potential to introduce further error into the fitting procedure.

Zero weights are perpetuated through the sample. Using the IPF approach, if in equation 5-5 the count of category c within table t is zero then the estimated weight for individuals with this category is zero. This in itself is not a problem since there is no requirement for individuals in this category. However the person possesses other characteristics that are categories in other tables and since weights are updated by multiplication, such an individual's weight can never be other than zero (see Box 5-3). With CO, at each iteration, the weight of an individual can be adjusted to any value to achieve an improvement in fit and is not constrained by the value at any previous iteration.

Box 5-3 : Impact of zero weights

Suppose in a particular LAD there are no individuals with an ethnicity of Black. When the weights are adjusted using the Black category within the ethnicity table, the weights are set to zero for all individuals in the sampling population who are Black. However some of these Black individuals may be male and when it comes to re-weighting to achieve a given number of males, Black males cannot be given a non-zero weight. Thus such individuals are excluded from contributing to any constraint count and therefore the diversity of the sampling population is reduced.

Division by zero. If there are no individuals with the desired characteristics in the sample population then the sum of weights in the denominator of equation 5-5 is zero and the proportionate adjustment is undefined. The usual solution to this is to either re-design the sample population by merging categories of individuals to ensure that individuals with the desired characteristics are present in the sample population or abandon this iteration step. Using the CO approach there is no need to implement a special mechanism to account for such a situation. The inner sum of weights in equation 5-6 will simply be zero and the inability to represent individuals with the desired characteristics in the sample population is captured by an increased TAE.

Importance of table order. The outcome of IPF is sensitive to the order in which the tables are considered, with a perfect fit achieved for the final table used in the iterative

procedure. In the absence of any guidance on what this order should be this decision becomes a subjective decision of the analyst. (Although the stepwise procedure adopted by Anderson, 2007, does provide such guidance to the analyst.) In CO all tables are considered jointly and equally in assessing the quality of fit at each iteration.

Execution time. Traditionally the drawback for CO has been the long execution time of algorithms to minimise the goodness of fit function. This is understandable given that the problem represents an optimisation in N dimensions where N is the size of the sample population. However the adoption of efficient optimisation algorithms and efficient implementation of such algorithms on multi-core computers has tended to overcome this deficiency (Harland, 2013).

5.3.1 Construction of constraint tables

In section 4.5 a range of candidate constraint variables are identified that have the potential to be related to the morbidity diagnosis of an individuals and are compatible with equivalent ELSA variables. Table 5-1 lists those 2011 Census tables that the literature identifies as potentially informative constraint variables and have their equivalents in the wave 5 ELSA data. All these tables are multi-dimensional cross tabulations which follows from recommendations by Hermes and Poulsen (2012a).

As well as compatibility in the definition of constraints between the two populations (Anderson, 2007, point 1), another desirable objective in constructing the constraint tables is that there is consistency amongst the tables. This consistency should include having the same population size irrespective of the constraint table. This is was not possible with 2001 Census tabular output since each output table was individually adjusted before publication to mask small counts but it is possible with the 2011 Census, since only pre-tabular disclosure control has been used (see Section 4.2.6).

Most of the constraint tables are based on a population of usual residents living in an area, there being 18,229,893 such residents aged 50 and over in England in 2011. There are however five tables where the population is based on residents in households only (signified with a '*' in Table 5-1) and one which is based on residents in institutional establishments only (there are 17,857,816 people resident in households and 372,077 people resident in institutional establishments).

Table 5-1 : Summary of 2011 Census constraint tables

Table	Age bands	Categories
DC1117 (Age structure)	Single year of age	Not applicable
LC3101 (Disability)*	50, 55, 60, 65, 70, 75, 80, 85 and older	No LLTI, Some LLTI
DC2101 (Ethnicity)		English/Welsh/Scottish/Northern Irish/British, Irish, Gypsy or Irish Traveller, Other White, White and Black Caribbean, White and Black African, White and Asian, Other Mixed, Indian, Pakistani, Bangladeshi, Chinese, Other Asian, African Caribbean, Other Black, Arab, Any other ethnic group
DC1108 (Living Arrangements)*		Married couple, Cohabiting, Single, Married (not a couple), Separated, Divorced, Widowed
DC4109 (Vehicle ownership or use)*		No vehicle, 1 vehicle, 2 or more vehicles
DC6114 (NS-Sec)	50, 65 and older	1.1 Large employers and higher managerial and administrative occupations, 1.2 Higher professional occupations, 2. Lower managerial, administrative and professional occupations, 3. Intermediate occupations, 4. Small employers and own account workers, 5. Lower supervisory and technical occupations, 6. Semi-routine occupations, 7. Routine occupations, 8. Never worked and long-term unemployed, L14.1 Never worked, L14.2 Long-term unemployed, L15 Full-time students
LC3301 (Unpaid care)*		Provides no unpaid care, Provides 1 to 19 hours unpaid care a week, Provides 20 to 49 hours unpaid care a week, Provides 50 or more hours unpaid care a week
DC3408 (Tenure)*		Owned outright; Owned with a mortgage or loan or shared ownership; Social rented; Private rented or living rent free
DC5107 (Education)	50, 55, 60, 65 and older	No qualifications, Level 1 qualifications, Level 2 qualifications, Apprenticeship, Level 3 qualifications, Level 4 qualifications and above, Other qualifications
DC3402 (Disability) +	50, 65, 75, 85 and older	No LLTI, Some LLTI (a little), Some LLTI (a lot), staff member, family member

Note :
* population base is residents in households;
+ population base is residents in institutional establishments

After some consideration it is clear why this distinction is made. With the living arrangements table, one component used to categorise an individual is the relationships between household members recorded within the household grid. Clearly, if someone is not resident in the household then they are not present in this grid and their living arrangements cannot be defined. Tenure and the ownership or use of a vehicle are attributes that are associated with households and thus only household residents can be included in these counts. The provision of unpaid care is something that occurs outside an institutional context, so is only reported for those who live in households. Finally the presence of a disability is separately tabulated for both household residents and

residents in an institutional establishment (signified with a '+' in Table 5-1). In the context of disability this distinction between household and institutional establishment residents is important – it is likely that the health profile of these two populations are different, with residents in institutional establishment having greater disability (Gordon et al, 2014).

The issue then arises on how to reconcile these two population basis. One answer is to have within each table based on residents in households, a 'residual' category of people who are resident in the area but not living in a household. However, some individuals who live in institutional establishments do so by virtue of their employment or through ownership of the establishment, in fact there are 35,774 people age 50 or older in England who live in an institutional establishment as owners, staff or their family. The question is whether they resemble more closely, on attributes such as car ownership and disability, the household or the institutionally resident 'client' population. If the assumption is the former then there needs to be a mechanism to identify how many people there are in this category (institutional residents who are not clients) and a way to infer the household characteristics for such individuals. This mechanism is now discussed.

The task of identifying how many and the gender and age distribution of such individuals can be achieved by the differencing of two published Census tables. The Census table on Residence type (DC1104) provides a gender by age by residence type (Lives in a household; Lives in an institutional establishment) count for each LAD. Unfortunately for residents in an institutional establishment, there is no differentiation between those who reside there as clients and those who reside by virtue of their employment or ownership. A further census table, DC3402 however separately counts those institutional residents who are clients and those who are not clients into four post 50 age bands. By differencing the counts from these two tables, it is possible to construct the true age distribution of those who are living in an institutional establishment but are not clients in the establishment. In the case of one LAD, Plymouth, the calculations in Box 5-4 shows that in the 50 years and older population there are 85,250 individuals who live in households plus 1,773 individuals who are clients in an institutional establishment and 154 who are resident in a institutional establishment but not a client. The distributions of the gender and age counts derived from these calculations are not estimates but true counts since the counts are have been derived by the differencing of published counts from two Census tables.

Box 5-4 : Residence type counts for Plymouth

Male	DC1104		DC3402	Difference	Female	DC1104		DC3402	Difference
	In household	In Institutional Establishment	Clients	Not Clients		In household	In Institutional Establishment	Clients	Not Clients
50 to 54	8,259	51	139	29	50 to 54	8,010	47	99	32
55 to 59	6,730	52			55 to 59	7,131	42		
60 to 64	7,377	65			60 to 64	7,656	42		
65 to 69	5,944	56	84	5	65 to 69	6,104	52	122	7
70 to 74	4,565	33			70 to 74	5,168	77		
75 to 79	3,565	63	163	7	75 to 79	4,388	104	286	22
80 to 84	2,384	107			80 to 84	3,283	204		
85 and older	1,589	168	162	6	85 and older	3,097	764	718	46
Total	40,413	595	548	47	Total	44,837	1,332	1,225	107

The total household population aged 50 and over in Plymouth is 85,250. Additionally there are 1,927 residents living in institutional establishments. Of those 1,773 are clients of the establishments whilst 154 are employed by, own the establishment or are family members. This latter figure its gender and age distribution (derived here) are **not** an estimate but would be the actual number of individuals aged 50 and older who live in institutional establishments, but are not clients within the establishment, if such a count was provided.

The question then arises of how to account for these non-client residents in institutional establishments in those tables where they are not included in the population base. Using disability status as an illustration, Census table LC3301 already categorises the 85,250 Plymouth household residents who live in households according to their disability status. There is also Census table DC3402 that categorises the 1,773 individuals who are clients into disability categories according to their disability status. However, there is no information in the Census (either explicitly or derivable) on the disability status by age of the 154 individuals who are resident in an institutional establishment but are not clients within the establishment. For these individuals the solution adopted here is to allocate each gender and age band counts in proportion to that recorded for household residents shown in table LC3301. Thus in the case of presence of a disability, those individuals who are male and aged 50 to 64 who are resident in a institutional establishment, but not clients, are allocated to a category of having a limiting long term illness or not in the same ratio as that for male household residents aged 50 to 64 (see Box 5-5 for an example of how this applies to males in Plymouth). These final counts are not true counts but estimates based on this allocation assumption.

In summary the Census tables used as constraints have been handled in the following ways:

1) Census tables based on the usual resident population are as published (see Table 13-11 for an example);

2) Census tables based on the household resident population (tables denoted with a '*' in Table 5-1 and see Table 13-12 for an extract) are constructed from:
    a) Published counts of household residents;
    b) Residents of institutional establishments who are not clients allocated according to the gender and age specific splits in the household resident population;
    c) A univariate count of residents of institutional establishments who are clients.
3) Census tables based on the institutional establishment populations (just table DC3402, see Table 13-13) are constructed from:
    a) Published counts of clients resident in institutional establishments;
    b) A univariate count of household residents and non-client residents in institutional establishments.

Box 5-5 : Allocation of non-client institutional residents for Plymouth

Males	No LLTI	LLTI	Total
50 to 64	16,326	6,040	22,366
65 to 69	5,926	4,583	10,509
75 to 74	2,195	3,754	5,949
85 and older	296	1,293	1,589

The distribution of disability by age for males living in households.

Males	No LLTI	LLTI	Total
50 to 64	73.0%	27.0%	100%
65 to 69	56.4%	43.6%	100%
75 to 74	36.9%	63.1%	100%
85 and older	18.6%	81.4%	100%

These counts expressed as a percentage.

Males	No LLTI	LLTI	Total
50 to 64	21.17	7.83	29
65 to 69	2.82	2.18	5
75 to 74	2.58	4.42	7
85 and older	1.12	4.88	6

The counts of males who are resident in a institutional establishment, but are not clients (see Box 5-4) distributed according to these percentages.

Males	No LLTI	LLTI	Total
50 to 64	21	8	29
65 to 69	3	2	5
75 to 74	3	4	7
85 and older	1	5	6

The counts above converted integers.

Males	No LLTI	LLTI	Total
50 to 64	16,347	6,048	22,395
65 to 69	5,928	4,586	10,514
75 to 74	2,198	3,758	5,956
85 and older	297	1,298	1,595

The counts above added into the household residents' counts, giving the final estimate of the distribution of disability for non-clients of institutional establishments.

The final table here is not a true count from the Census but an estimate based on the assumption that those who are not clients, but are residing in an institutional establishment, have similar disability status to those who live in households.

5.3.2 Definition of sample population from ELSA

The second data requirement in a spatial microsimulation is a sample population. These are the individuals whose weights are adjusted to reproduce the constraint counts provided by the 2011 Census tables. The sample population here is taken from wave 5 of ELSA. It consists of participants of any type (core members and new partners) who are aged 50 or older and have a valid response to any of the constraint variables. This latter requirement excludes all sample members, about whom very little is known. In addition, information is required in the health microsimulation on whether the participant has ever smoked. The correspondences between the structure of the constraint tables and the coding of individuals within the sample population are shown in Table 13-1 to Table 13-13 located in the Appendix.

Due to the sparse nature of the ELSA population in regards to certain types of individual, it is necessary to merge some age band categories. This is particularly so for the ethnic categorisation. The gender, age and ethnic composition of the sample population is provided in Table 5-2 and it is apparent that some combinations of these characteristics are not represented in this sample population (shown as **0**). This represents a problem for the spatial microsimulation since it is unable to find any individuals with the required combination of characteristics. To ensure that there are individuals available for all combinations of characteristics, categories within each characteristic may be merged. For example the ethnic classification uses just two age bands (Table 13-2) rather than the eight shown in Table 5-2.

An extract of the sampling population is provided in Table 13-14 which categorises each individual according to the structures in Table 13-1 to Table 13-13. Note that this extract contains two individuals (103730 and 103955) who are resident in an institution as clients and therefore have generic information (INSTITIUTION) for those constraint tables that are based on the household resident population (tables of type 2 in Section 5.3.1). Similarly residents in households or non-clients resident in institutional establishments have generic information (HOUSEHOLD) for table DC3402. This structure helps to ensure that the process of choosing sample weights recognises the existence of the two sub-populations, those who are clients in an institutional establishment and those who are not.

Table 5-2 : Composition of the sample population by gender, age and ethnic group

	Age band	White	Mixed	Asian	Black	Other	Total
Male	Aged 50 to 54	153	1	6	1	1	162
	Aged 55 to 59	843	2	31	7	6	889
	Aged 60 to 64	963	6	14	4	5	992
	Aged 65 to 69	735	2	10	10	3	760
	Aged 70 to 74	665	1	12	5	2	685
	Aged 75 to 79	483	**0**	7	2	2	494
	Aged 80 to 84	272	**0**	2	2	1	277
	Aged 85 and older	206	**0**	1	2	**0**	209
Female	Aged 50 to 54	350	1	18	4	5	378
	Aged 55 to 59	1,045	2	34	13	8	1,102
	Aged 60 to 64	1,107	4	10	11	4	1,136
	Aged 65 to 69	808	1	12	4	3	828
	Aged 70 to 74	745	1	5	5	**0**	756
	Aged 75 to 79	561	1	2	9	3	576
	Aged 80 to 84	390	1	1	2	**0**	394
	Aged 85 and older	348	**0**	1	**0**	**0**	349
Total		9,674	23	166	81	43	9,987

5.3.3 Validation of spatial population

To validate the robustness of the spatially microsimulated population it is useful to examine how well the simulated population within each LAD reproduces 2011 Census counts for a table not used as a constraint (a procedure outlined in Edwards et al., 2011). The degree of agreement between the 2011 Census counts and the synthetically derived equivalent counts can be assessed in a number of ways, but perhaps the simplest is to scatter plot one count against the other. Ideally these points should lie close to a 45º line that passes through the origin.

5.4 ETHPOP PROJECTION REVISIONS

The ETHPOP sub national population projections produced in 2010 provide projections of the sub national population in each English LAD by gender, age and ethnic group for each year to 2051 (Rees et al., 2011). The important role for these projections in the health microsimulation is to ensure that the gender, age and ethnic structure of future populations in each LAD evolves to represent the dynamic nature of these demographic trends, see Figure 5-1.

Whilst the projections are available for single years of age, by 16 ethnic groups and 352 English LADs, here these single years are aggregated to five year age bands, the number of ethnic groups is reduced to five broad groups and the information is aggregated to the 324 LADs in place for 2013 rather than the actual current 326. This reduction results

from a merger of the City of London with Westminster and the Scilly Isles with Cornwall. The time frame of the microsimulation here is also limited to 2031 as the final year of estimation.

The published ETHPOP projections used a mix of information: 2001 Census tabulations, both published and commissioned; vital statistics on fertility and mortality; and administrative data sources e.g. national insurance registrations and the National Health Service Central Register. Unfortunately the data from the 2011 Census had yet to be collected when the projections were carried out so it was not possible to use these counts. Now, however, the performance of the ETHPOP projections for 2011 may be judged against 2011 Census counts (this comparison is explored in detail in Section 7.2). In light of this performance, LAD specific revisions are proposed to the trends in these original ETHPOP projections, producing a set of revised ETHPOP projections. This revision process consists of a number of elements (Rees and Clark, 2014):

Step 1. Re-bases projections to 2011 Census counts;
Step 2. Utilises subsequent ETHPOP cohort projections;
Step 3. Incorporates a projection adjustment factor based on 2001 to 2011 performance; and
Step 4. Constrains the population by gender and age to ONS 2012-sub national mid-year population projections (ONS, 2014a).

5.4.1 Yearly adjustment factor

For each LAD and ethnic group, the growth observed in the 40 years and older population between the 2001 and 2011 Censuses and the 2001 and 2011 ETHPOP projections is compared using equation 5-7.

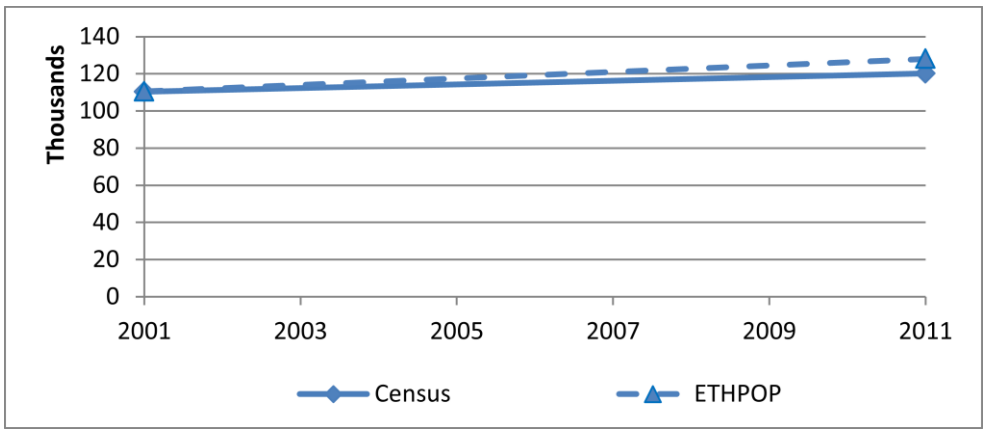
$$\text{yearly adjustment} = \left[\frac{\frac{2011\ \text{Census}}{2001\ \text{Census}}}{\frac{2011\ \text{ETHPOP}}{2001\ \text{ETHPOP}}}\right]^{\frac{1}{10}} \qquad 5\text{-}7$$

where 2001 Census is the ethnic specific count of the 40 and older population in the LAD from the 2001 Census (Table S101);
2011 Census is the ethnic specific count of the 40 and older population in the LAD from the 2011 Census (Table DC2101);
2001 ETHPOP is the mean of the TREND-EF and UPTAP-ER 2001 projection of the 40 and older population in the LAD from ETHPOP;
2011 ETHPOP is the mean of the TREND-EF and UPTAP-ER 2011 projection of the 40 and older population in the LAD from ETHPOP.

This yearly adjustment factor reflects how the ETHPOP projection needs to be adjusted to reproduce the 2011 outcome (see Box 5-6 for an example calculation):

Box 5-6 : Yearly adjustment for the White population of Plymouth

White	2001	2011	growth	yearly adjustment
Census	110,314	120,221	$\frac{120{,}221}{110{,}314} = 1.090$	$\left[\frac{1.090}{1.158}\right]^{\frac{1}{10}} = 0.9940$
ETHPOP	110,506	128,000	$\frac{128{,}000}{110{,}506} = 1.158$	

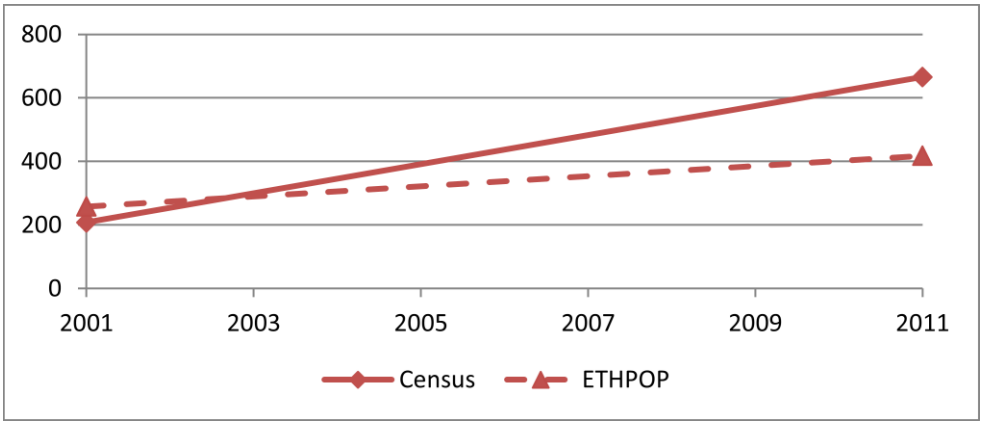


To reproduce the observed growth between the 2001 and 2011 Census, the ETHPOP projections need to be reduced by 0.0061 (1-0.9939) or 0.61% per annum.

In practice some of these adjustment factors can be very large or small, due to the calculation involving small counts in the denominators. The repercussions are that there is the potential for some large growths (or declines) in some sub-populations. To limit this impact information about the area type of the LAD is used in order to apply a constraint around these adjustment factors (ONS, 2003). The hypothesis is that the bulk of LADs of the same type experience the same shifts in population growth and hence should have similar correction factors. To identify those LADs that are untypical of similar LADs, the raw adjustment factors of each authority and ethnicity are calculated (as shown in Box 5-6) and within each LAD area type a 10% and 90% empirical confidence interval of these raw factors is calculated for each ethnicity. If the authority's raw adjustment factor is outside this range then it is re-set to the appropriate lower (10%) or upper (90%) bound (see Box 5-7).

Box 5-7 : Yearly adjustment for the Asian population of Plymouth

White	2001	2011	growth	yearly adjustment
Census	208	666	$\frac{666}{208} = 3.202$	$\left[\frac{3.202}{1.620}\right]^{\frac{1}{10}} = 1.070$
ETHPOP	258	418	$\frac{418}{258} = 1.620$	



To reproduce the observed growth between the 2001 and 2011 Census, the ETHPOP projections need to be adjusted by +7.0% pa. The [10%,90%] range for Asian in Cities and Services type LADs is [-0.3%,+4.5%], therefore actual yearly correction factor for Plymouth Asian population is adjusted from +7.0% down to +4.5%.

5.4.2 Re-base and adjust ETHPOP projection

Once these yearly correction factors are known for each ethnicity within each LAD, the appropriate ethnic population in a LAD for year y (y=2011; 2013, …, 2031) may then calculated using equation 5-8.

$$EP_e^{LAD,y} = C_e^{LAD\ 2011} * \left(\frac{E_e^{LAD,y}}{E_e^{LAD,2011}}\right) * (AF_e^{LAD})^{(y-2011)} \qquad 5\text{-}8$$

Where $EP_e^{LAD,y}$ is the adjusted ETHPOP sub-population aged 40 and older of ethnicity e, in LAD for year y;

$C_e^{LAD,y}$ is the 2011 Census sub-population aged 40 and older of ethnicity e, in LAD (step 1);

$E_e^{LAD,y}$ is the original ETHPOP sub-population aged 40 and older of ethnicity e, in LAD for year y (step 2);

$AF_{g,e}^{LAD}$ is the adjustment factor for adjust for 2001 to 2011 ETHPOP projection performance for ethnicity e, in LAD (step 3).

The quantity EP is the size of the total sub-population aged 40 and older of a given ethnicity. This is then apportioned to each 5-year age band using the ETHPOP LAD gender, age and ethnic specific counts (equation 5-9).

$$EP_{g,a,e}^{LAD,y} = EP_{e}^{LAD,y} * \frac{E_{g,a,e}^{LAD,y}}{E_{+,+,e}^{LAD,y}} \quad \text{5-9}$$

An illustration of this process is provided in Box 5-8 .

Box 5-8 : Adjustment to ETHPOP projection

2011 Census	ETHPOP GROWTH	Adjustment factor	ETHPOP population share	Revised estimate
equation 5-8			equation 5-9	
White people aged 40 and older in Plymouth (step 1)	White people aged 40 and older in Plymouth from 2011 to 2015 (step 2)	(step 3)	White males aged 50 to 54 relative to all white people aged 40 and older	
120,221 *	$\frac{133,697}{128,000}$	* 0.9939^4 *	$\frac{9,112}{133,697}$	= 8,352.1

The original ETHPOP projection of white males aged 50 to 54 in Plymouth in 2015 is revised from 9,112 to 8,352.1.

5.4.3 Constrain to ONS sub national population projections

The final step is to constrain the gender and five year age bands to the ONS 2012 sub national population projections (ONS, 2014a):

$$P_{g,a,e}^{LAD,y} = \frac{EP_{g,a,e}^{LAD,y}}{EP_{g,a,+}^{LAD,y}} * SNPP_{g,a}^{LAD,y} \quad \text{5-10}$$

where $P_{g,a,e}^{LAD,y}$ is the revised population of gender g, age a, ethnicity e, in LAD for year y;

$EP_{g,a,e}^{LAD,y}$ is the ETHPOP population of gender g, age a, ethnicity e, in LAD for year y;

$SNPP_{g,a}^{LAD,y}$ is the ONS 2012-based sub national population projection of gender g, age a, in LAD for year y.

This process is illustrated in Box 5-9.

Box 5-9 : Constrained to ONS 2012-SNPP

Males aged 50 to 54	$EP^{Plymouth,2015}_{Male,50\ to\ 54,e}$	$\frac{EP^{Plymouth,2015}_{Male,50\ to\ 54,e}}{EP^{Plymouth,2015}_{Male,50\ to\ 54,+}}$	$SNPP^{Plymouth,2015}_{Male,50\ to\ 54}$	$P^{Plymouth,2015}_{Male,50\ to\ 54,e}$
White	8,352.1	0.9694	8,400	8,142.8 (9,112)
Mixed	67.5	0.0078	8,400	65.8 (40)
Asian	91.3	0.0106	8,400	89.0 (48)
Black	50.3	0.0058	8,400	49.1 (31)
Other	54.7	0.0063	8,400	53.3 (67)
Total	8,615.9	1.0000		8,400 (9,304)

The total male population aged 50 to 54 for Plymouth in 2015, of 8,615.9 is, pro-rata, reduced to equal the 2012-SNPP of 8,400. To give an indication of the scale of the adjustments, the original ETHPOP projections are given in ()'s in the final column.

Since the health microsimulation is a microsimulation of the behaviour of individuals, there is a requirement that the population projections are composed of integer counts. This achieved by rounding the counts $P^{LAD,y}_{g,a,e}$ to the nearest integer.

5.5 HEALTH MICROSIMULATION

The outcome of the spatial microsimulation is a series of LAD specific weights for the ELSA wave 5 derived sample population. These weights synthesise the makeup of each LADs 50 years and older population in March 2011. The task of the health microsimulation is to evolve this population through time to represent future LAD population structures. This is achieved by firstly ageing the population as time progresses, replenishing the younger ages in the population, and adjusting the weights to reproduce future population structures. Since this is a health microsimulation and not just a demographic projection, the health status of individuals is also updated. In Figure 5-1 the health microsimulation is shown at the heart of the figure and is seen to take information from the spatial microsimulation, the adjusted ETHPOP projections, information from the 2011 HSfE and uses hazard models for morbidity, mortality and migration propensities.

5.5.1 Ageing

In this study, each LAD's population is evolved through time to produce biennial mid-year forecasts of the health status of its population for a 20 year time horizon. Thus the 2011 March Census population is evolved to 2013, 2015, … , 2031 mid-year populations. This adoption of a two year step in forecasts is in keeping with the two-year wave cycle of ELSA. In terms of ageing, there is a slight mismatch for the first microsimulation step from March 2011 to mid-year 2013. Between these two dates an individual will of aged 2¼ years rather than 2. To infer someone's age mid-year 2013 age from their age in March 2011 requires information on their month of birth. Unfortunately ELSA does not provide information on the date or month of birth so it is not possible to perform this calculation. In any case the time spans in this study are already to a degree inconsistent - ELSA data are collected over a one year period that starts 9 months before the 2011 Census and finishes 3 months after.

5.5.2 Morbidity status

As the synthetic population is aged, the morbidity status of the individuals who do not have a specific morbidity is updated. This is achieved using a Monte-Carlo approach. In this approach, hazard models are estimated that use gender, age, time (wave) trend, ethnicity, smoking status, the presence of a comorbidity, and the type of LAD to estimate the probability that an individual acquired a morbidity in the past two years. A pseudo random number taking a value between 0.000 and 1.000 is then generated (Matsumoto and Kurita, 1994) and if the value is less than the hazard probability, the individual is deemed to have acquired the morbidity, otherwise they continue without the morbidity.

The two morbidities of CVD and DHBS are linked - individuals with DHBS are more likely to develop CVD. To build on this linkage, the presence of DHBS at the previous wave is used as a potential predictor of CVD. For the remaining morbidity of having a respiratory illness, there is no similar comorbidity link.

5.5.3 Replenishers

A consequence of the ageing of participants in the LAD population is that the population no longer covers the full age range. At each microsimulation step the age of the youngest individuals increases by 2 years. Thus by 2031 there would be no one younger than 70 in the population. There therefore needs to be some mechanism to replenish the synthetic population with younger participants and the problem then becomes how to give sufficient detail to these Replenishers so that they can be fully functioning individuals within the population.

The gender, age and ethnicity makeup of the 50 and 51 years old Replenishers can be determined easily enough by the revised ETHPOP population projections. Within this

study these populations are only known in five year age bands. To estimate the single year of age populations for 50 and 51 years olds the methodology proposed by Sprague (1880) is used (Calot and Sardon, 2003). In application, the method uses the five year age band containing the single year of age required (here the 50 to 54 year age band) plus the two five year age bands below this (aged 40 to 49 and aged 45 to 49) and the two five year age bands above (aged 55 to 59 and aged 60 to 64). This requirement is the reason that the revised ETHPOP projections covered the 40 years and older group and not just the more obvious 50 years and older group.

Reliable future predictions on other aspects of the population are more problematic. Considering the determinants that are discussed in Chapter 3, it is difficult to infer them for future populations. The marital status of individuals changes through time, so it is a challenge to accurately ascertain their living arrangements at age 50 and 51 and how these arrangements change subsequently. There are similar issues with all the other main socio-economic factors : highest education; NS-SEC classification; tenure; care giving and vehicle ownership or use.

One determinant that is health related and has some form of stability is whether the individual has ever smoked. To reliably use this information in any model the hypothesis that people rarely begin smoking after the age of 30 needs to be true. The age 30 is significant because it is the smoking status of those aged 30 in 2011 who will be 50 in 2031, the final Replenishers. If it is known how many people aged 30 to 50 in 2011 who had ever smoked (say from the HSfE) then, as long as no one in this age group starts smoking after 2011, their smoking status when they replenish the synthetic population will also be known. Evidence on how true this hypothesis is can be found in the life course data collected during the ELSA wave 3 interview. In that interview a question was asked on when the individual first started to smoke cigarettes. Box **5-10** tabulates the weighted counts of wave 3 life course participant's age by the age at which they started to smoke cigarettes. This provides evidence that amongst this cohort, a very small percentage of participants (2.5%) started smoking after they were aged 30. Statistics based on the UK General Household Survey quoted in Chatterjee et al. (2008) also shows that few people start smoking after the age of 30, there, 9% of males and 15% of female smokers started smoking after the age of 25. The assumption here is that the Replenishers, aged 30 to 50 in 2011, who do not already smoke, are unlikely to start.

Box 5-10 : Age at which ELSA wave 3 participants started smoking

Age at wave 3	Don't know	Never smoked	Aged started smoking cigarettes							Total
			Younger than 10	Aged between 10 and 19	Aged between 20 and 29	Aged between 30 and 39	Aged between 40 and 49	Aged between 50 and 59	Aged 60 or older	
Aged 50 to 59	3%	40%	1%	47%	8%	1.0%	0.5%	0.0%	0.0%	2,288
Aged 60 to 69	3%	37%	1%	47%	11%	1.5%	0.4%	0.2%	0.0%	2,252
Aged 70 to 79	3%	39%	1%	41%	12%	2.8%	1.3%	0.3%	0.1%	1,580
Aged 80 to 89	4%	36%	0%	40%	16%	2.4%	0.5%	0.4%	0.0%	805
Aged 90 and older	40%	38%	0%	15%	5%	2.4%	0.0%	0.0%	0.0%	127
All	4%	38%	1%	44%	11%	1.7%	0.6%	0.2%	0.0%	7,053

Of those aged 50 to 90, the greatest percentage (nearly a half) started smoking cigarettes when they were aged between 10 and 19. About 2.5% started smoking aged 30 or older. A consistent percentage, around 40%, never smoked cigarettes. The recall information for those aged 90 and older is poor, however the never smoked percentage for this very old age group looks to be in line with the percentage reported for younger cohorts.

To gain some understanding of the smoking prevalence within this younger cohort, use is made of the 2011 HSfE. The percentage of those aged 30 to 50 in 2011 who have ever smoked is likely to provide accurate estimates of the percentage who smoke when they become the Replenishers during the period 2013 to 2031, since very few change their status from never having smoked to smoked during that period (see Box **5-10**). Using the HSfE it is possible to estimate separate rates by gender; age cohort (aged 30 to 39 and aged 40 to 49) and ethnicity. When introducing Replenishers into the LAD population these HSfE derived smoking rates can be used to ensure they are representative. This smoking status is assigned using a Monte-Carlo approach in combination with these rates. Clearly this information on whether someone has ever smoked is not the most powerful determinant of health status. Whether they are a current smoker, the number of years they have smoked or the amount smoked would be better determinants, however all these measures have the capability of changing in unknown ways over a life course.

A further requirement is the need to designate a proportion of these Replenishers as having morbidities when they are introduced into the synthetic population. It would be unreasonable to assume that all the 50 and 51 year olds are completely healthy. To estimate this probability use is made of the morbidity hazard models. Starting at the age of 30, with an assumption of not having the morbidity, and repeatedly applying the hazard model to predict the probability of acquiring a morbidity through the ages 30 to 49 and across the waves, the accumulated probability over 20 years of acquiring a morbidity can be estimated. When the Replenishers are introduced into the synthetic population, the Monte-Carlo approach using these accumulated probabilities is used to decide whether the individual has the morbidity at age 50 or 51.

5.5.4 Constraining weights

Ageing and replenishing the synthetic population are not the only demographic processes that need to occur. There is the need to ensure that the population is of the correct size and structure. This is achieved by selecting individuals and adjusting their weights to represent growth in the population (cloning) or decline (removal), to ensure that the population is structured

Box 5-11 : Impact of using unequal selection probabilities

Supposing the population was 1,000 individuals, split 900 without and 100 with a morbidity, and the size of the population needed to be reduced by 100. With equal probabilities this would produce, on average, a split of 810 vs 90 (each is reduced by 10%), with an estimated prevalence of 90 / 900 = 10%. However supposing that much MORE of the reduction is in the sub population with the morbidity (by ¼ vs ¾) then the split would be 875 vs 25, a much lower prevalence of 25 / 900 = 2.8%.

correctly. Removal from the population represents two demographic processes, those of death and out-migration from the LAD, whilst cloning represents in-migration. The question then arises of how to select the individual to clone or remove. The simplest approach would be to select with equal probability. This would however not be a true reflection of the dynamics within the population. Consider the need to remove an individual of a particular type, potentially representing a death. It is more likely that death occurs to someone who is older or who has a morbidity. This suggests that there is the need to estimate the probability of mortality as a function of the characteristics of the individual. Once again this is achieved using a hazard model to predict the probability of death. Selection then takes place not with equal probability but with a probability equal to that estimated by the mortality hazard model. How this change in approach has the potential to influence estimated prevalences is illustrated in Box 5-11. There is similarly a need to estimate the probability of migration. When selecting an individual for cloning, only the migration probability is used, whilst selection for removal, both the mortality and migration probabilities are used. It is important to recognise that these mortality and migration probabilities are not used to determine the size of the population in future years - that is determined by the adjusted ETHPOP projections. They are merely used to determine who should be cloned or removed to achieve a population structure that resembles the adjusted ETHPOP projections.

A cloned individual inherits all the characteristics of the original individual (obviously gender, age and ethnicity) including their smoking status and morbidity status. After the cloning however, they become independent, thus if at a later stage the source individual acquires a morbidity or is removed from the population, the clone does not necessarily do the same (and vice-versa).

A final issue arises when there are no suitable individuals within the sample population to create a clone from. It is entirely possible that in the spatially microsimulated population that represents the 2011 population within the LAD there is no requirement for individuals with a combination of attributes (e.g. older males of mixed ethnicity). However, in future years the ETHPOP projections may indicate that there will be people living in the LAD with this combination, but there may be none available for cloning. The approach here is to offer, in these exceptional circumstances, the original spatial microsimulation sample population as sources for clones. These ‘shadow’ individuals are only cloned if there are no suitable individuals already in the sample population available for cloning (either in the original synthetic population or from Replenishers). Before their potential use, the shadow individuals are subject to the same processes (e.g. ageing and changes to their morbidity status) as the synthetic population, so that when and if they are cloned they are truly representative. Also, since ageing

takes place there is a need for ‘shadow Replenishers’. Once a shadow individual has been cloned they are treated as an original individual within the synthetic population.

5.5.5 Merged categories

A matter not unrelated to the need for shadow individuals is apparent from looking at Table 5-2. Outside the White and younger Asian populations the representation of the population is fairly sparse and for some combinations there is no representation in the sample population. Older, mixed ethnicity, males are particularly poorly represented and the ageing of the younger individuals in this demographic will only start to fill out this age range at 2023 (the oldest mixed ethnicity male is aged 74 in 2011 and will be 86 in 2023). It is therefore impossible to re-shape the future population of this demographic because there are no individuals (shadow or otherwise) in the sample. More generally even where there are some individuals available for cloning if this pool of individuals is small then the sample becomes too homogeneous – one or two individuals are repeatedly cloned. To overcome these concerns rather than attempting to reproduce all gender, age and ethnic combination populations using the adjusted ETHPOP projections, certain age bands are merged to ensure a diversity of individuals within the synthetic population from which to create clones. These merged categories and the number of sample population individuals within the merged categories is shown in Table 5-3.

Table 5-3 : Composition of sample population with merged categories

<table>
<tr><th></th><th></th><th>White</th><th>Mixed</th><th>Asian</th><th>Black</th><th>Other</th></tr>
<tr><td rowspan="12">Male</td><td>Aged 50 to 54</td><td>153</td><td rowspan="3">9</td><td>6</td><td rowspan="3">12</td><td rowspan="3">12</td></tr>
<tr><td>Aged 55 to 59</td><td>843</td><td>31</td></tr>
<tr><td>Aged 60 to 64</td><td>963</td><td>14</td></tr>
<tr><td>Aged 65 to 69</td><td>735</td><td rowspan="5">3</td><td>10</td><td rowspan="2">15</td><td rowspan="5">8</td></tr>
<tr><td>Aged 70 to 74</td><td>665</td><td>12</td></tr>
<tr><td>Aged 75 to 79</td><td>483</td><td rowspan="3">10</td><td rowspan="3">6</td></tr>
<tr><td>Aged 80 to 84</td><td>272</td></tr>
<tr><td>Aged 85 and Older</td><td>206</td></tr>
<tr><td rowspan="12">Female</td><td>Aged 50 to 54</td><td>350</td><td rowspan="3">7</td><td>18</td><td rowspan="3">28</td><td rowspan="3">17</td></tr>
<tr><td>Aged 55 to 59</td><td>1,045</td><td>34</td></tr>
<tr><td>Aged 60 to 64</td><td>1,107</td><td>10</td></tr>
<tr><td>Aged 65 to 69</td><td>808</td><td rowspan="5">4</td><td>12</td><td rowspan="2">9</td><td rowspan="5">6</td></tr>
<tr><td>Aged 70 to 74</td><td>745</td><td>5</td></tr>
<tr><td>Aged 75 to 79</td><td>561</td><td rowspan="3">4</td><td rowspan="3">11</td></tr>
<tr><td>Aged 80 to 84</td><td>390</td></tr>
<tr><td>Aged 85 and Older</td><td>348</td></tr>
</table>

For the White ethnicity the selection and de-selection of individuals reproduces the full gender, age and ethnicity projections. For the other more sparse ethnicities the selection and de-selection reproduces the merged projections. For the mixed ethnicity this means that the ETHPOP projections for those younger than 65 are merged and those over 65 are merged so that, in the worst case, initially 3 individuals are available as potential clones to represent the older mixed male ethnicity (although over time younger mixed ethnicity males will age into this age group and become eligible for cloning).

5.6 HEALTH IMPACT DECOMPOSITION

The health microsimulation described above has two main drivers that affect the prevalence of morbidities. One is a demographic dimension that is driven by the change in the gender, age and ethnic structure of a LAD and the second is the trends in the morbidity hazard driven by the estimated hazard equation. Some insight into the relative contribution of these two drivers can be obtained using an accounting system.

The demographic drivers have a number of elements. Firstly there is the size of the population. Naturally as the population grows the number of individuals with morbidities also grows. If the composition within this population (in terms of gender, age and ethnic splits) does not change (an unlikely occurrence) then the prevalence rates would remain unaffected by this growth. However, the structure of the population will change over time and these changes can be decomposed into a number of scenarios enabling the health contribution of each component of demographic change to be assessed (Bongaarts and Bulatao, 1999).

The final scenario revolves around assumption on health trends within the populations. The morbidity hazard change is driven by a time trend in the hazard equations and how these equations differentiate by gender, age and ethnicity in the estimation of hazard probabilities. If this hazard function is not used to update morbidity status; the population is not aged; and selection probabilities are not used, then this is equivalent to assuming constant morbidity prevalence within each sub population, fixed at 2011 estimates.

5.6.1 Demographic drivers

The approach taken here to measure the impact of various demographic factors on the prevalence rates is to start with an assumption that the demographic structure and size of the population of a LAD resembles that in the 2011 Census. Incrementally these assumptions are relaxed, the order of which, can to a degree, be arbitrary. The first change is trivial and assumes that the total size of the 50 and older population increases in line with the revised ETHPOP projection with no mortality improvement, but the 2011 Census gender, age and ethnic composition persists. Next, the first critical

demographic characteristic to be relaxed is the gender structure of the population. Thus the population follows the gender structure predicted by the adjusted ETHPOP projections for the forecast year but the age and ethnic structures are those from the 2011 Census. The next stage is to adjust both the gender and ethnic structures to those in the ETHPOP projections whilst retaining the 2011 Census age structure. The final stage is to relax all three demographic factors, gender, ethnicity and age. However, the age structure is relaxed in two stages. The ageing of the population occurs through two processes, the first is by cohort replacement and migration, where the natural ageing of a population and migration change its structure; the second is through mortality improvements where over time, individuals of a given age are less likely to die. The latter process of mortality improvement can be negated by the use of factors derived from the no mortality improvement variant of ONS's 2012-based national population projections (ONS, 2013b). These factors are gender and age specific but not specific for a particular LAD or ethnic group.

It should be emphasised that these alternative versions of the future population structure presented here are not in any way thought to be plausible structures (except for $P_{g,a,e}^{LAD,y}$), they are just artificial constructs to help to quantify the contribution of various components to trends in health outcomes. In a similar manner to how Rees et al. (2013b) decomposed the components of population change, the health outcomes from using each of these versions in the health microsimulation enables the estimation of the incremental or proportionate impact of relaxing each assumption.

Constructing populations that meet these population constraint structures is a challenging task and is accomplished in two stages. The first stage is to calculate broad adjustment factors that can be applied to a previous population to bring it more into line with a new structure. The second stage is to use this adjusted population as the starting point in an optimisation procedure to exactly produce the new structure. To help with the description of the first stage, some notation is introduced here (which follows on from equation 5-10) (also to aid in the process of understanding the mechanisms and how they relate to this notation, a hypothetical scenario is provided in Table 5-4 to Table 5-7).

$P_{g,a,e}^{LAD,y}$ is the revised ETHPOP population for local authority LAD in year y of gender g, age band a and ethnicity e (see Table 5-5);

$P_{g,a,e}^{LAD,2011}$ the population for local authority LAD in the 2011 Census of gender g, age band a and ethnicity e (see Table 5-4);

$NMI_{g,a,\circ}^{\circ,y}$ the no mortality improvement factor in year y for gender g, age band a (see Table 5-6).

The dot ($^{\circ}$) denotes a characteristic that does not vary by the relevant attribute. This means that in the notation above, the no mortality improvement factors are the same for all LAD's and for each ethnic group, but vary by year, gender and age. In what follows, a plus (+) signifies a summation across all categories for the characteristic. Thus $P_{g,+,+}^{LAD,2011}$ is the sum of males (g=male) or females (g=female) for local authority LAD in the 2011 Census (over all age bands and ethnic groups).

Table 5-4 : Hypothetical 2011 Census Counts

2011	Male		Female		age by gender		age by ethnicity		
Census	White	BME	White	BME	Male	Female	White	BME	age
Aged 50 to 64	200	20	250	20	220	270	450	40	490
Aged 65 to 75	150	15	200	20	165	220	350	35	385
Aged 75 and older	100	5	150	10	105	160	250	15	265
gender by ethnicity	450	40	600	50					
gender	490		650						
ethnicity	White	1050	BME	90					
							1140		

Table 5-5 : Hypothetical 2021 ETHPOP projections

2021	Male		Female		age by gender		age by ethnicity		
Principal /ETHPOP	White	BME	White	BME	Male	Female	White	BME	age
Aged 50 to 64	150	25	200	35	175	235	350	60	410
Aged 65 to 75	200	20	250	25	220	275	450	45	495
Aged 75 and older	150	15	200	20	165	220	350	35	385
gender by ethnicity	500	60	650	80					
gender	560		730						
ethnicity	White	1150	BME	140					
							1290		

Table 5-6 : Hypothetical factors to remove mortality improvement (NMI)

2021	Principal		NMI variant			
Principal to NMI	Male	Female	Male	Female	Male	Female
Aged 50 to 64	175	235	157	211	0.9	0.9
Aged 65 to 75	220	275	176	248	0.8	0.9
Aged 75 and older	165	220	116	132	0.7	0.6

Table 5-7 : 2021 ETHPOP projection without mortality improvement

2021	Male		Female		age by gender		age by ethnicity		
NMI	White	BME	White	BME	Male	Female	White	BME	age
Aged 50 to 64	135	22.5	180	31.5	157.5	211.5	315	54	369
Aged 65 to 75	160	16	225	22.5	176	247.5	385	38.5	423.5
Aged 75 and older	105	10.5	120	12	115.5	132	225	22.5	247.5
gender by ethnicity	400	49	525	66					
gender	449		591						
ethnicity	White	925	BME	115					
							1040		

These quantities are then aggregated to produce some useful totals. The first is the size of the population under the assumption of no mortality improvement, i.e. life expectancies remain unchanged (1,040 in Table 5-7).

$$PN_{g,a,e}^{LAD,y} = P_{g,a,e}^{LAD,y} \times NMI_{g,a,\bullet}^{\bullet,y} \quad 5\text{-}11$$

and

$$PN_{+,+,+}^{LAD,y} = \sum_{g,a,e} \left(PN_{g,a,e}^{LAD,y} \right) \quad 5\text{-}12$$

5.6.1.1 Original population

The starting point is the assumption that the size and structure of future populations is exactly the same as that in the 2011 Census (Table 5-4):

$$PO_{g,a,e}^{LAD,y} = P_{g,a,e}^{LAD,2011} \quad 5\text{-}13$$

5.6.1.2 Population size

The next step is to grow the population in future years so that it reflects the change in the size of the population without any mortality improvement but the gender, age and ethnic distribution of the population is the same as the 2011 Census (Table 5-8).

$$PS_{g,a,e}^{LAD,y} = \frac{P_{g,a,e}^{LAD,2011}}{P_{+,+,+}^{LAD,2011}} \times PN_{+,+,+}^{LAD,y} \quad 5\text{-}14$$

5.6.1.3 Population gender

The third step is to create a population that reflects the change in the population size and its gender structure, but retains the ethnic and age structure from the 2011 Census (Table 5-9).

$$PG_{g,a,e}^{LAD,y} = \frac{P_{g,a,e}^{LAD,2011}}{P_{g,+,+}^{LAD,2011}} \times PN_{g,+,+}^{LAD,y} \quad 5\text{-}15$$

5.6.1.4 Population ethnicity

The fourth step is to create a population that reflects the change in the population size and its gender and ethnic structure, but retains the age structure from the 2011 Census (Table 5-10).

$$PE_{g,a,e}^{LAD,y} = \frac{P_{g,a,e}^{LAD,y}}{P_{g,+,e}^{LAD,y}} \times PN_{g,+,e}^{LAD,y} \quad 5\text{-}16$$

5.6.1.5 Population age

The penultimate step is a population that reflects the change in the population size and the three demographic characteristics of gender, age and ethnicity (Table 5-7) but no mortality improvement. This is simply $PN_{g,a,e}^{LAD,y}$.

Table 5-8 : The 2011 Census gender, age and ethnic structure re-sized to produce the 2021 population with no mortality improvement

2021	Male		Female		age by gender		age by ethnicity		
SIZE	White	BME	White	BME	Male	Female	White	BME	age
Aged 50 to 64	182.5	18.2	228.1	18.2	200.7	246.3	410.5	36.5	447.0
Aged 65 to 75	136.8	13.7	182.5	18.2	150.5	200.7	319.3	31.9	351.2
Aged 75 and older	91.2	4.6	136.8	9.1	95.8	146.0	228.1	13.7	241.8
gender by ethnicity	410.5	36.5	547.4	45.6	**1040**				
Gender	447.0		593.0						
ethnicity	White	957.9	BME	82.1					

Table 5-9 : The 2011 Census age and ethnic structure BUT the gender split and 2021 population size is the same as the no mortality improvement

2021	Male		Female		age by gender		age by ethnicity		
GENDER	White	BME	White	BME	Male	Female	White	BME	age
Aged 50 to 64	183.3	18.3	227.3	18.2	201.6	245.5	410.6	36.5	447.1
Aged 65 to 75	137.4	13.7	181.8	18.2	151.2	200.0	319.3	31.9	351.2
Aged 75 and older	91.6	4.6	136.4	9.1	96.2	145.5	228.0	13.7	241.7
gender by ethnicity	412.3	36.7	545.5	45.5	**1040**				
gender	**449.0**		**591.0**						
ethnicity	White	957.9	BME	82.1					

Table 5-10 : The 2011 Census age structure BUT the gender and ethnic split and 2021 population size is the same as the no mortality improvement

2021	Male		Female		age by gender		age by ethnicity		
ETHNIC	White	BME	White	BME	Male	Female	White	BME	age
Aged 50 to 64	177.8	24.5	218.8	26.4	202.3	245.2	396.5	50.9	447.4
Aged 65 to 75	133.3	18.4	175.0	26.4	151.7	201.4	308.3	44.8	353.1
Aged 75 and older	88.9	6.1	131.3	13.2	95.0	144.5	220.1	19.3	239.5
gender by ethnicity	**400.0**	**49.0**	**525.0**	**66.0**	**1040**				
gender	**449.0**		**591.0**						
ethnicity	White	**925.0**	BME	**115.0**					

5.6.1.6 Population mortality

The final structure is a population that reflects changes in the three demographic characteristics and also incorporates mortality improvement (Table 5-5). This is simply $P_{g,a,e}^{LAD,y}$.

5.6.1.7 Optimisation

If these adjustments work perfectly then there should be exact agreement between the relevant marginal totals in Table 5-4 to Table 5-10. The gender totals in Table 5-9 should match with the gender totals in Table 5-7 (they do) and the age, gender and age by ethnicity totals in Table 5-9 should match with those in Table 5-8 (they don't quite). Similarly the gender, ethnic and gender by ethnicity totals in Table 5-10 should match those in Table 5-7 (they do) whilst the age totals in Table 5-10 should match those in Table 5-8 (they don't quite).

Whilst the factoring described above has, in this hypothetical example, produced population structures that nearly meet the required constraints of re-producing a mix of population structures, this cannot be guaranteed and whilst they may be a good starting point for a suitable structure, further work is required.

Here this extra level of agreement is achieved by optimising the counts within the table so that they meet the required margin totals dictated by the required structure. This is achieved by optimising the cells in the table to minimise the differences between the optimised and required totals. The starting values in this optimisation are those derived above using the factoring approach. For the Population Gender counts this involves:

Choose $PG^{*\,LAD,y}_{g,a,e}$ such that they minimise:

$$TAE_G = \sum_{g=1}^{2} \left| PG^{*\,LAD,y}_{g,+,+} - PN^{LAD,y}_{g,+,+} \right| + \sum_{a=1}^{8} \sum_{e=1}^{5} \left| PG^{*\,LAD,y}_{+,a,e} - PS^{LAD,y}_{+,a,e} \right| \qquad 5\text{-}17$$

The first part of equation 5-17 attempts to reproduce the gender totals of Table 5-7 whilst the second part attempts to reproduce the age by ethnicity marginal totals of Table 5-8. For Population ethnicity counts the expression is:

Choose $PE^{*\,LAD,y}_{g,a,e}$ such that they minimise:

$$TAE_{GE} = \sum_{g=1}^{2} \sum_{e=1}^{5} \left| PE^{*\,LAD,y}_{g,+,e} - PN^{LAD,y}_{g,+,e} \right| + \sum_{a=1}^{8} \left| PE^{*\,LAD,y}_{+,a,+} - PS^{LAD,y}_{+,a,+} \right| \qquad 5\text{-}18$$

The first part of equation 5-18 attempts to reproduce the gender by ethnicity marginal totals of Table 5-7 whilst the second part attempts to reproduce the age totals of Table 5-8.

The success of the optimisation can be judged by the values of equations 5-17 or 5-18 at the minimum and the impact of the changes can be judged by the difference between the initial $PG^{LAD,y}_{g,a,e}$ and the optimised $PG^{*\,LAD,y}_{g,a,e}$ (and similarly between $PE^{LAD,y}_{g,a,e}$ and $PE^{*\,LAD,y}_{g,a,e}$). It is hoped that these changes are small.

For the hypothetical example using Table 5-10 as the starting point and minimising the function in equation 5-18 using the EXCEL solver add-in (FrontlineSolvers, 2015) produces the population structure $PE^{*\,LAD,2021}_{g,a,e}$ in Table 5-11.

Table 5-11 : The optimised 2011 Census age structure BUT the gender and ethnic split and 2021 population size is the same as the no mortality improvement

2021	Male		Female		age by gender		age by ethnicity		
ETHNIC (OPT)	White	BME	White	BME	Male	Female	White	BME	age
Aged 50 to 64	177.6	24.4	218.5	26.4	202.1	244.9	396.2	50.8	447.0
Aged 65 to 75	132.6	18.4	173.9	26.4	151.0	200.2	306.4	44.8	351.2
Aged 75 and older	89.8	6.2	132.6	13.2	95.9	145.8	222.4	19.4	241.8
gender by ethnicity	400.0	49.0	525.0	66.0	1040				
gender	449.0		591.0						
ethnicity	White	925.0	BME	115.0					

The value of the objective function value of equation 5-18 with Table 5-10 is 4.58 and after the optimisation the value with Table 5-11 it is 0.00058. The sum of the absolute changes made between Table 5-10 and Table 5-11 is 4.66; which is just 0.4% of the population size.

5.6.2 Health change drivers

The impact of no changes in health status is measured by not ageing individuals and not applying the mortality, morbidity or migration probabilities in the health microsimulation and consequently not introducing Replenishers. This approach provides the same cohort of individuals to represent each gender, age and ethnic population in the LAD for all forecast years. Coupled with this is the need to have equal probability of selection for mortality and migration, else otherwise the prevalence rate within the sub population would change over times by virtue of disproportionately selecting individuals by their morbidity.

5.6.3 Health decomposition implementation

The health microsimulation can be used to estimate the number of individuals with a morbidity for all these health decomposition scenarios. The starting point is using the original population (Table 5-4) as input to the health microsimulation instead of the adjusted ETHPOP projections and not implementing any of the ageing, replenishment, ageing or selection mechanisms (in fact these mechanisms are not implemented until the final scenario). The second scenario is to repeat the exercise but use the 2011 Census population that is re-sized to the size of a population with no mortality improvement (Table 5-8). The third scenario substitutes the population with one that has undergone gender adjustment (Table 5-9); the fourth substitutes with a gender and ethnicity adjustment (Table 5-10); the fifth with a no mortality improvement structure adjustment (Table 5-7). The penultimate scenario uses the revised ETHPOP projections (Table 5-5)

without the health change drivers (e.g. ageing or updating of morbidities). The final scenario is the full model that uses the adjusted ETHPOP projections and the health change drivers e.g. ageing, replenishing the population, unequal selection probabilities and updating morbidities.

The impact of each demographic and health driver can be quantified in terms of an incremental (I) and a relative (R) change in the size of the population at risk and those affected with the morbidity. The equations for the absolute changes are:

$$IS_{g,a,e}^{LAD,y} = MS_{g,a,e}^{LAD,y} - MO_{g,a,e}^{LAD,y} \quad 5\text{-}19a$$

$$IG_{g,a,e}^{LAD,y} = MG_{g,a,e}^{LAD,y} - MS_{g,a,e}^{LAD,y} \quad 5\text{-}19b$$

$$IE_{g,a,e}^{LAD,y} = ME_{g,a,e}^{LAD,y} - MG_{g,a,e}^{LAD,y} \quad 5\text{-}19c$$

$$IN_{g,a,e}^{LAD,y} = MN_{g,a,e}^{LAD,y} - ME_{g,a,e}^{LAD,y} \quad 5\text{-}19d$$

$$IM_{g,a,e}^{LAD,y} = MM_{g,a,e}^{LAD,y} - MN_{g,a,e}^{LAD,y} \quad 5\text{-}19e$$

$$IH_{g,a,e}^{LAD,y} = MH_{g,a,e}^{LAD,y} - MM_{g,a,e}^{LAD,y} \quad 5\text{-}19f$$

where $MO_{g,a,e}^{LAD,y}$ is the number of individuals with the morbidity based on the 2011 Census population;

$MS_{g,a,e}^{LAD,y}$ is the number of individuals with the morbidity based on a re-sized population without mortality improvement;

$MG_{g,a,e}^{LAD,y}$ is the number of individuals with the morbidity based on a population without mortality improvement adjusted by gender;

$ME_{g,a,e}^{LAD,y}$ is the number of individuals with the morbidity based on a population without mortality improvement adjusted by gender and ethnicity;

$MN_{g,a,e}^{LAD,y}$ is the number of individuals with the morbidity based on a population without mortality improvement adjusted by gender, ethnicity and age;

$MM_{g,a,e}^{LAD,y}$ is the number of individuals with the morbidity based on a population with mortality improvement adjusted by gender, ethnicity and age but no health improvement;

$MH_{g,a,e}^{LAD,y}$ is the number of individuals with the morbidity based on a population with mortality improvement adjusted by gender, ethnicity and age with health improvement.

The final estimate of the number of individuals with a morbidity can be reconstructed from the original estimate $MO_{g,a,e}^{LAD,y}$ and the incremental estimates in equations 5-19a to 5-19f using equation 5-20.

$$MH_{g,a,e}^{LAD,y} = MO_{g,e,q}^{LAD,y} + IS_{g,e,q}^{LAD,y} + IG_{g,e,q}^{LAD,y} + IE_{g,e,q}^{LAD,y} + IN_{g,e,q}^{LAD,y} + IM_{g,e,q}^{LAD,y} + IH_{g,e,q}^{LAD,y} \quad 5\text{-}20$$

Alternatively, the impacts can be derived as a ratios:

$$RS_{g,a,e}^{LAD,y} = MS_{g,a,e}^{LAD,y} \setminus MO_{g,a,e}^{LAD,y} \quad 5\text{-}21a$$

$$RG_{g,a,e}^{LAD,y} = MG_{g,a,e}^{LAD,y} \setminus MS_{g,a,e}^{LAD,y} \quad 5\text{-}21b$$

$$RE_{g,a,e}^{LAD,y} = ME_{g,a,e}^{LAD,y} \setminus MG_{g,a,e}^{LAD,y} \quad 5\text{-}21c$$

$$RN_{g,a,e}^{LAD,y} = MN_{g,a,e}^{LAD,y} \setminus ME_{g,a,e}^{LAD,y} \quad 5\text{-}21d$$

$$RM_{g,a,e}^{LAD,y} = MM_{g,a,e}^{LAD,y} \setminus MN_{g,a,e}^{LAD,y} \quad 5\text{-}21e$$

$$RH_{g,a,e}^{LAD,y} = MH_{g,a,e}^{LAD,y} \setminus MM_{g,a,e}^{LAD,y} \quad 5\text{-}21f$$

In which case the final estimate of the number of individuals with a morbidity can be re-constructed using equation 5-22.

$$MH_{g,a,e}^{LAD,y} = MO_{g,e,q}^{LAD,y} \times RS_{g,e,q}^{LAD,y} \times RG_{g,e,q}^{LAD,y} \times RE_{g,e,q}^{LAD,y} \times RN_{g,e,q}^{LAD,y} \times RM_{g,e,q}^{LAD,y} \times RH_{g,e,q}^{LAD,y} \quad 5\text{-}22$$

5.6.4 Health decomposition scenarios

It is possible to examine these ratios to gauge to what extent health outcomes are inevitable and which are modifiable. The multipliers $RS_{g,e,q}^{LAD,y} \times RG_{g,e,q}^{LAD,y} \times RE_{g,e,q}^{LAD,y} \times RN_{g,e,q}^{LAD,y}$ are largely outside the control of the health services and other agencies, representing the natural change in population structure. The multiplier, $RM_{g,e,q}^{LAD,y}$, is a consequence of the postponement of death through: lifestyle and occupation changes, environmental improvements and good health care.

If the value of the final multiplier, $RH_{g,e,q}^{LAD,y}$, is less than 1.00 then this measures the potential impact of intervening to reduce morbidity prevalence, or if it is greater than 1.00, how much needs to be done to mitigate the future prevalence for the morbidity.

5.7 PRESENTATION OF RESULTS

The health microsimulation produces a large volume of output. The basic outputs from the health microsimulation are counts of individuals who do and do not have each of the three morbidities in the 324 LADs for the 10 forecast years. Since the microsimulation works at the level of individuals, these morbidity statistics can also be presented separately by gender, age, ethnicity and smoking status.

To compare the outputs between areas, the prevalence counts can be converted into rates and expressed as a percentage using equation 5-23.

$$\text{prevalence rate}\%_{\text{morbidity}}^{LAD,y} = \frac{\text{prevalence count}}{\text{population at risk}} \times 100 \quad 5\text{-}23$$

To provide insight into the outcomes of the modelling process, a number of differing formats are used. These outputs are tabular, graphical and maps but use is also made of

case study LADs. To this end, seven LADs are selected, one each from the ONS area type classification at the level of Super Groups (ONS, 2003). In each Super Group, the LAD 'closest' to the centre of the cluster is chosen as the case study LAD. These LADs and their distance from the centre are given in Table 5-12. Since there are three Super Groups orientated towards London, there is a large percentage of these case studies form Greater London.

Table 5-12 : LADs cluster and distance to cluster centres

Super Group	LAD (distance)
Cities and Services	Leeds (1.22)
London Suburbs	Redbridge (3.07)
London Centre	Camden (1.49)
London Cosmopolitan	Haringey (3.87)
Prospering UK	Wiltshire (0.76)
Coastal and Countryside	Teignbridge (1.41)
Mining and Manufacturing	Rotherham (1.12)

5.8 CONCLUSION

In this chapter each component of the structure of the health microsimulation model has been introduced in order to satisfy the third objective. Diagrammatically, this structure is illustrated in Figure 5-1. The structure involves many components: a spatial microsimulation to define a starting population for each LAD; a collection of ethnic population projections to define the structure of future LAD populations; a mechanism to evolve the population through time to 2031; and a series of hazard models to both identify significant characteristics in the population and to modify the status of the individuals in the population over time.

A range of constraint tables are available for use in the spatial microsimulation and special care is taken in constructing these tables to ensure that the population size in consistent in each table. Individuals who are resident in institutional establishments as clients are included in the spatial microsimulation so here the spatial microsimulation covers the entire population, not just those resident in households. Too often these individuals are omitted from health models and, although small in number, their health care needs are distinct and critical.

Use is made of an external aggregate population projection to ensure that the future population structure within each LAD evolves over time. In light of the subsequent release of 2011 Census tables these projections are re-based, adjusted and constrained to ONS sub-national population projections.

The health microsimulation takes as input these LAD starting populations and the adjusted population projections to evolve the population from 2011 to 2031 in biennial steps. This is achieved by: ageing the population; changing the morbidity status of individuals; replenishing the population at the younger ages; and constraining the population to the projected structure.

In an attempt to understand how each of these dynamics affect the future morbidity status of the population, the changes in the structure of the population are decomposed into scenarios that reflect changes due to the size of the population and its gender, ethnic and age structure. An additional scenario is developed which does not implement any trends in improved morbidity outcomes.

This thesis continues with a series of five results chapters which meet the fourth objective set out in section 1.4. The first three results chapters deal with one aspect of the methodology described in this chapter. They are the spatial microsimulation that defines the LADs base population in 2011 (Chapter 6); the revisions to the ETHPOP projections (Chapter 7); and the probability aspects of the methodology, e.g. estimation of hazard models and representing the chances of having smoked in the Replenishers (Chapter 8). Chapter 9 presents the outcome of the health microsimulation and the final results chapter, Chapter 10, describes the decomposition of the morbidity outcomes using an accounting system.

6 SPATIAL MICROSIMULATION

Creating the Initial Population

6.1 INTRODUCTION

This chapter presents the result of the spatial microsimulation that establishes a 2011 base population of those aged 50 and older for each LAD in England. This will be achieved by using a set of constraint tables derived from the 2011 Census and a sample population taken from wave 5 of ELSA. The aim of the spatial microsimulation is to estimate a series of LAD specific weights for the sample population so that the composition of the weighted sample population re-produces the LADs counts in the 2011 Census tables.

The chapter begins in section 6.2 with a re-cap of the candidate 2011 Census tables that have the potential to be used in the spatial microsimulation. Attention then turns in section 6.3 to the ELSA wave 5 data that are used to identify those constraints that contribute significantly to explaining the three outcome morbidities of CVD, DHBS and respiratory illness. In the spirit of Anderson (2007), the literature review and the outcome of hazard model fits are used to identify a set of constraint tables for use in both the spatial microsimulation and for validation purposes. Section 6.4 reports on the calibration fit of the spatial microsimulation providing a visual assessment of the path to convergence for each LAD. Section 6.5 reports the outcome of the spatial microsimulation summarised using the estimated prevalence of the three morbidities for each LAD in 2011. The important issue of how robust the weight estimation process is to the choice of random numbers is covered in section 6.6. The validation against a set of constraint tables that are not used in the spatial microsimulation is reported in section 6.8. The final section, 6.9, provides a summary of the finding of this chapter and point to how they will be used in a subsequent chapter on health microsimulation.

6.2 CONSTRAINT TABLES

In section 4.5 an overview is provided of the variables that are both consistent between the 2011 Census and ELSA and that have been found to influence individual's health outcomes in the literature. More detail on how this compatibility is achieved is provided in sections 5.3.1 and 5.3.2 which describe the reformatting of the 2011 Census data and the wave 5 ELSA data.

All the constraint tables are cross tabulations for two reasons. Firstly, it follows the advice of Hermes and Poulsen (2012a) who highlight that such cross tabulation may

account for an unequal distribution among the categories that is not apparent in a series of univariate tables. Secondly, this requirement implicitly accepts that such tabulations represent not just potential but actual significant interactions between their constituent variables which motivates the ONS to continue to produce such cross tabulations.

Commonly two of the dimensions used in tables published by the ONS are gender and age. This age banding is of crucial importance here since there is a requirement to isolate the count information to just the 50 years and older population. The third constraint variable is then one of the potential socio-demographic (ethnicity, disability or living status) or socio-economic variables (highest level of qualification, vehicle ownership, tenure of accommodation or amount of unpaid care provided).

6.3 CHOICE OF CONSTRAINT TABLES

Looking at the health of the population using a general measure such as disability free life expectancy shows that there is a small difference in health status by gender. In section 2.8 the historic difference, with females having a slightly longer disability free life expectancy than males, is seen to persist but this may be diminishing over time. The literature review also suggests that there are potentially differing health outcomes for males and females for specific morbidities.

Age is also a possible factor that may influence health status – as people get older their health tends to deteriorate. There are also some morbidities, diabetes being a case in point, that has different prevalence by ethnicity. Other factors include various socio-demographic factors that relate to affluence or deprivation through accumulated wealth or income, e.g. educational qualification; social classification or tenure.

6.3.1 Initial constraint selection

Using morbidity outcomes in waves 1 to 6 of ELSA, hazard models are fitted to these data that attempts to measure the influence of various factors in predicting the incidence of these morbidities. Thinking forward to the health microsimulation it is important that the population of each LAD, in its composition, accurately reflects nature of the area so that future incidence and hence prevalence can be estimated. An initial set of variables that are strong candidates for inclusion arising out of a consideration of the literature are shown in Table 6-1 for each of the three morbidities. The table shows the deviance values, the degrees of freedom (dof) for the model, the change in deviance over the preceding model and its the significance level.

Table 6-1 : Model deviance values for initial constraints

CVD	deviance	dof	change		significance
constant	4000.9	1			
wave	3999.8	2	1.10		29.4%
wave + age	3915.7	3	84.08	***	0.0%
wave + age + gender	3903.0	4	12.69	***	0.0%
wave + age + gender + disability	3788.9	5	114.11	***	0.0%
wave + age + gender + disability + ethnicity[1]	3786.1	8	2.85		41.5%
Diabetes or high blood sugar					
constant	4081.8	1			
wave	4079.3	2	2.50		11.4%
wave + age	4078.9	3	0.41		52.1%
wave + age + gender	4070.4	4	8.50	***	0.4%
wave + age + gender + disability	4011.6	5	58.82	***	0.0%
wave + age + gender + disability + ethnicity[1]	3986.6	8	24.94	***	0.0%
Respiratory illness					
constant	3139.5	1			
wave	3129.5	2	9.95	***	0.2%
wave + age	3128.3	3	1.20		27.4%
wave + age + gender	3128.1	4	0.26		60.8%
wave + age + gender + disability	3056.2	5	71.81	***	0.0%
wave + age + gender + disability + ethnicity[1]	3053.7	8	2.54		46.8%

Note :
1 The three ethnic groups are: White; Asian; and Black, Mixed or Other. There are too few occurrences of the morbidities to reliably separate out the Mixed and the Other grouping.

* significant at the 10% level, ** significant at the 5% level, *** significant at the 1% level

The pattern of impact in Table 6-1 for each morbidity is inconsistent. The time trend, captured by the wave counter, is only significant for respiratory illnesses. This finding indicates that over time the probability of acquiring a respiratory illness has changed, all other things being equal. Age in only significant for the acquisition of CVD meaning that the probability of acquiring CVD varies by age. It is important to understand that this latter result does not mean the prevalence of these other morbidities does not vary by age. Recall that the purpose of the hazard model is to identify those factors that influence the incidence of the morbidity. Where age is not a significant factor then this says that, all other things being equal, the probability of acquiring the morbidity does not change with age. However, looking at the prevalence in the population as a whole, and given that the acquisition of a morbidity is an absorbive event, people will carry any morbidities with them through the age ranges until they leave the population, either through death or migration. Incidences at younger ages therefore accumulate to higher prevalences at older ages. Gender is important for two morbidities, CVD and DHBS – for these morbidities, males have a different probability of acquiring these two morbidities than females. Unsurprisingly the presence of a disability is significant in predicting the onset of a morbidity – since the presence of one of these morbidities

would incline people to describe their day to day activities as being limited. The final variable in this initial set is ethnicity which is shown to only be influential in predicting the onset DHBS. All the constraints in this initial set are seen to be influential for at least one morbidity – none can be rejected as being totally irrelevant. CVD justifies the inclusion of all the potential constraints except ethnicity and DHBS justifies the inclusion of ethnicity.

6.3.2 Socio-economic constraint selection

As described in section 4.5 and Table 5-1, there remain a number of candidate socio-economic constraint variables to help constrain the spatial population. These are: highest educational qualification, vehicle ownership and use, living arrangements, provision of unpaid care, social economic classification and tenure. From these candidate constraints it is useful to select both an additional constraint variable and one or two variables for validation purposes as outlined in section 5.3.3. As is done above in section 6.3.1 hazards models that build on the five variable models in that section are used to identify these constraints. However, here each candidate will be considered in turn and in isolation since there is a desire not to over-fit these data which can dilute the impact of other more informative variables, induce correlations between variables and make the model calibration more challenging than it needs to be. There is also the desire to retain some constraints, outwith the spatial microsimulation fit, for validation purposes. In Table 6-2 these hazard modelling results are shown in Table 6-1.

As seen for Table 6-1, once again there is little unanimity in the outcomes. Few variables are significant for CVD whilst many are significant for DHBS. Vehicle ownership is significant at least at the 10% level for all morbidities. Living arrangements and NS-SEC are significant at least at the 5% level for DHBS and respiratory illness. Highest qualification and tenure are only significant for one morbidity. The provision of unpaid care is never significant.

The results in Table 6-2 shows that vehicle ownership is the most import socio-economic variable, followed by either living arrangements or NS-SEC. The obvious choice would be to include vehicle ownership as a constraint variable in the spatial microsimulation, but here it is used instead as a measure to validate the model outputs. It is important that such a validation measure is itself significant and influential in determining the outcome of interest; otherwise its validation performance would be irrelevant. However there is still a need for a socio-economic constraint to include as a constraint variable. So, the next best constraint is used as the socio-economic constraint in the spatial microsimulation, leaving the third most important constraint available as a second validation. This logic means that living arrangements is chosen as the socio economic constraint variable to use in the spatial microsimulation since it outperforms

NS-SEC on two of the morbidities, leaving NS-SEC as the second validation constraint, in addition to vehicle ownership.

Table 6-2 : Model deviance values for candidate socio-economic variables

CVD	deviance	dof	Change		significance
Base (from Table 6-1)	3786.06	8			
Highest qualification	3783.10	11	2.96		39.9%
Vehicle ownership	3770.48	10	15.58	***	0.04%
Living arrangements	3778.98	14	7.07		31.4%
Unpaid care	3782.30	11	3.76		28.9%
NS-SEC	3781.85	12	4.20		37.9%
Tenure	3785.67	11	0.39		94.3%
Diabetes or high blood sugar					
Base (from Table 6-1)	3986.61	8			
Highest qualification	3974.70	11	11.91	***	0.77%
Vehicle ownership	3973.68	10	12.93	***	0.16%
Living arrangements	3964.09	14	22.52	***	0.10%
Unpaid care	3985.37	11	1.24		74.3%
NS-SEC	3966.22	12	20.39	***	0.04%
Tenure	3935.52	11	51.09	***	0.00%
Respiratory illness					
Base (from Table 6-1)	3053.70	8			
Highest qualification	3052.44	11	1.26		73.8%
Vehicle ownership	3048.40	10	5.31	***	7.04%
Living arrangements	3036.93	14	16.77	**	1.02%
Unpaid care	3052.58	11	1.12		77.2%
NS-SEC	3041.43	12	12.28	**	1.54%
Tenure	3049.22	11	4.49		21.3%

* significant at the 10% level, ** significant at the 5% level, *** significant at the 1% level

6.4 SPATIAL MICROSIMULATION ESTIMATION

Following on from the analysis in the previous section the constraint tables used in the spatial microsimulation are: gender by five year age band (LC1117) (Table 13-1); gender by age by ethnicity (DC2101) (Table 13-2); gender by age by disability (household residents) (LC3101) (Table 13-3); gender by age by disability (institutional residents who are clients) (DC3402) (Table 13-13) and gender by age by living arrangements (DC1108) (Table 13-10).

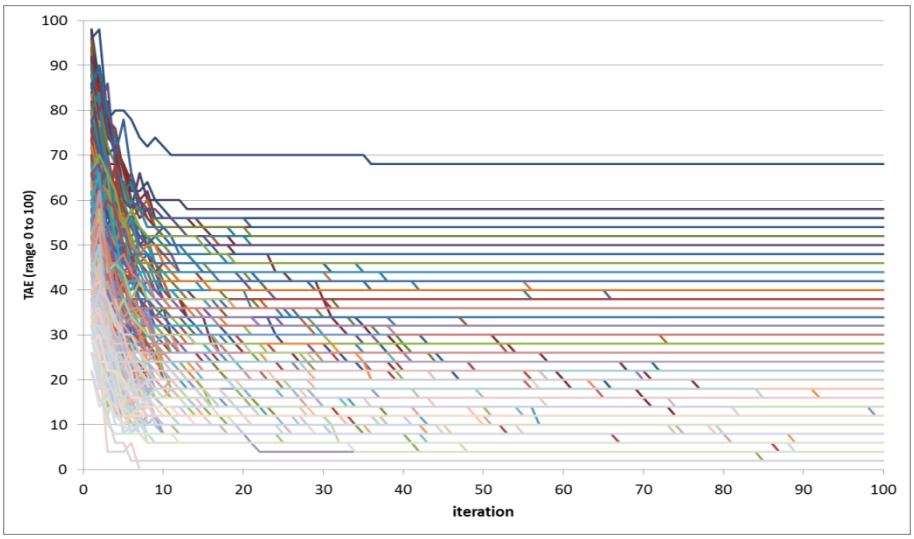
The Flexible Modelling Framework (FMF) software package is used to derive the LAD specific spatial weights, using the default optimisation parameters for the software (Harland, 2013). This implementation of the CO approach to spatial microsimulation has been used in health planning (for local prevalence estimates, Clark et al., 2014 and Shulman et al., 2015), transport (to estimate the capacity for walking and cycling in Leeds, Phillips, 2015 and estimating carbon emissions in Beijing, Ma et al., 2014) and to examine the impacts of urban renewal (Jordan et al., 2010).

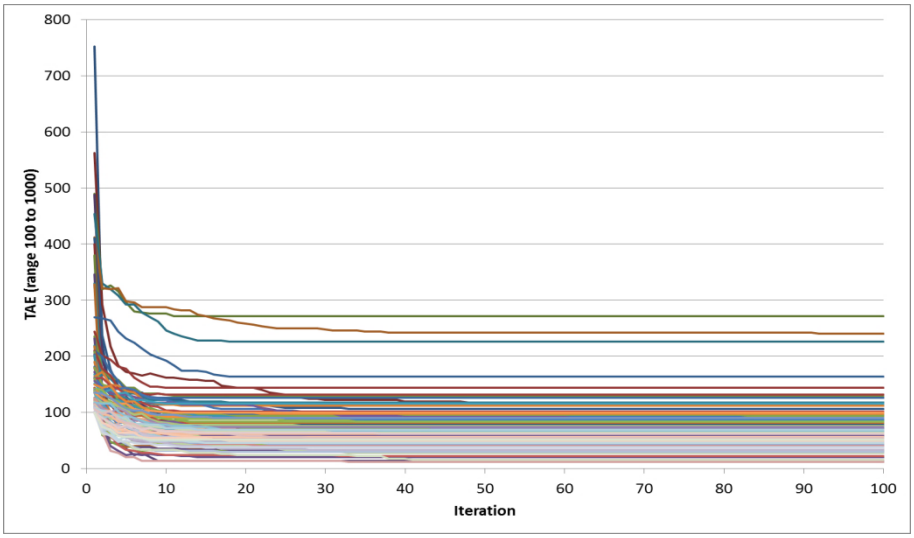
To derive the spatial weights for a population of just over 18 million people, the FMF takes 12 hours on a quad core i5-2300 PC with 4GB of RAM running Windows 7. The outputs of the FMF are contained in two files: a sample weights file that contains pairs of zone identifiers (i.e. LADs) and sample member identifiers (i.e. ELSA participants), the number of rows of which equals the size of the constraint population (in this case 18,299,893); and a fit statistic file that reports the Total Absolute Error (TAE) at the end of each annealing stage for each zone (the TAE here is analogous to that of 5-5). Given the requirement that the size of the 50 and older population in each LAD is maintained, there is some degree of double counting in the TAE, i.e. if an individuals is miss-allocated then the fit is out by |+1| for one constraint in a table (an over count), then for another constraint in the same table it must also be out by |-1| (an under count), making a total contribution to the TAE of +2.

Figure 6-1 shows the rate of convergence of the TAE over the 100 annealing stages, categorised by the TAE at the end of the first iteration. There are 226 LADs whose initial TAE is below 100 and further iterations make modest reductions in this already good TAE. There are just 17 LADs with a TAE above 1,000, and even with these challenging LADs the FMF is able to quickly reduce the TAE to a range of between 6 and 202. Recall that the TAE measures the differences between the implied counts (derived using the set of weights) and the actual constraint counts, summed over 16 + 20 + 17 + 11 + 36 = 100 constraint counts. Whilst these TAEs are not zero, they are very small when compared to the size of the population within each LAD, which averages 56,000 and ranges from 15,443 to 292,565. In the context of the number of constraints to be satisfied and the size of the population, the fit can be declared good.

6.5 ESTIMATING DISEASE PREVALENCE

Using information in ELSA for waves 1 to 5 it is possible to establish at the time of the wave 5 interview whether an ELSA participant had a practitioner diagnosed morbidity of CVD, DHBS or respiratory illness. The weighted count of those individuals with a morbidity within the spatially microsimulated population for each LAD enables an estimate of the prevalence count for a morbidity to be made and this can be converted into a prevalence rate by dividing by the size of the 50 and older population. These prevalence counts and rates are shown in Figure 6-2. These maps show the prevalence for each morbidity as a colour scale and also the number of individuals with the morbidity as a circle, where the area of the circle is proportional to the number of individuals with the morbidity. Four maps are provided, one for each single morbidity and one for the CVD and DHBS comorbidity. Inset maps are provided for London and Greater Manchester and Merseyside.





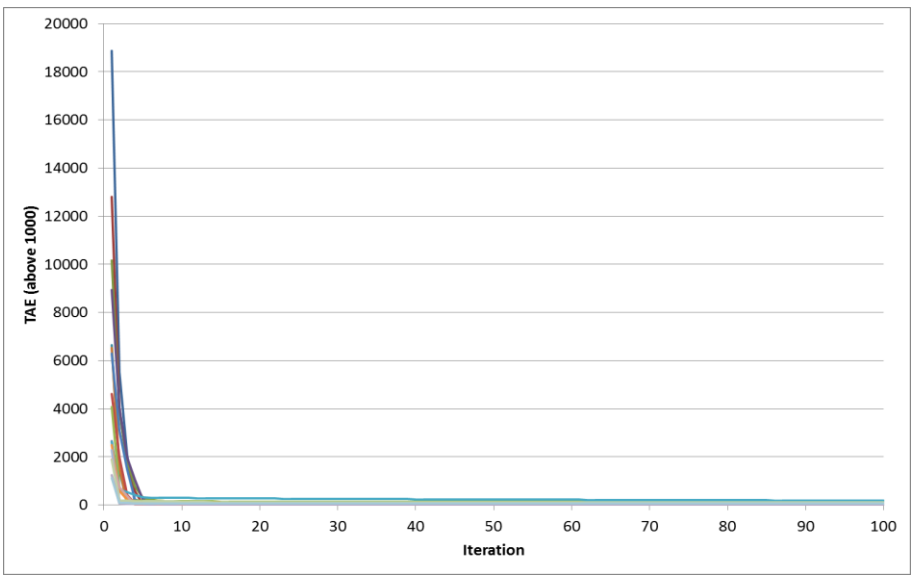


Figure 6-1 : Change in TAE from the spatial microsimulation, categorised by initial TAE

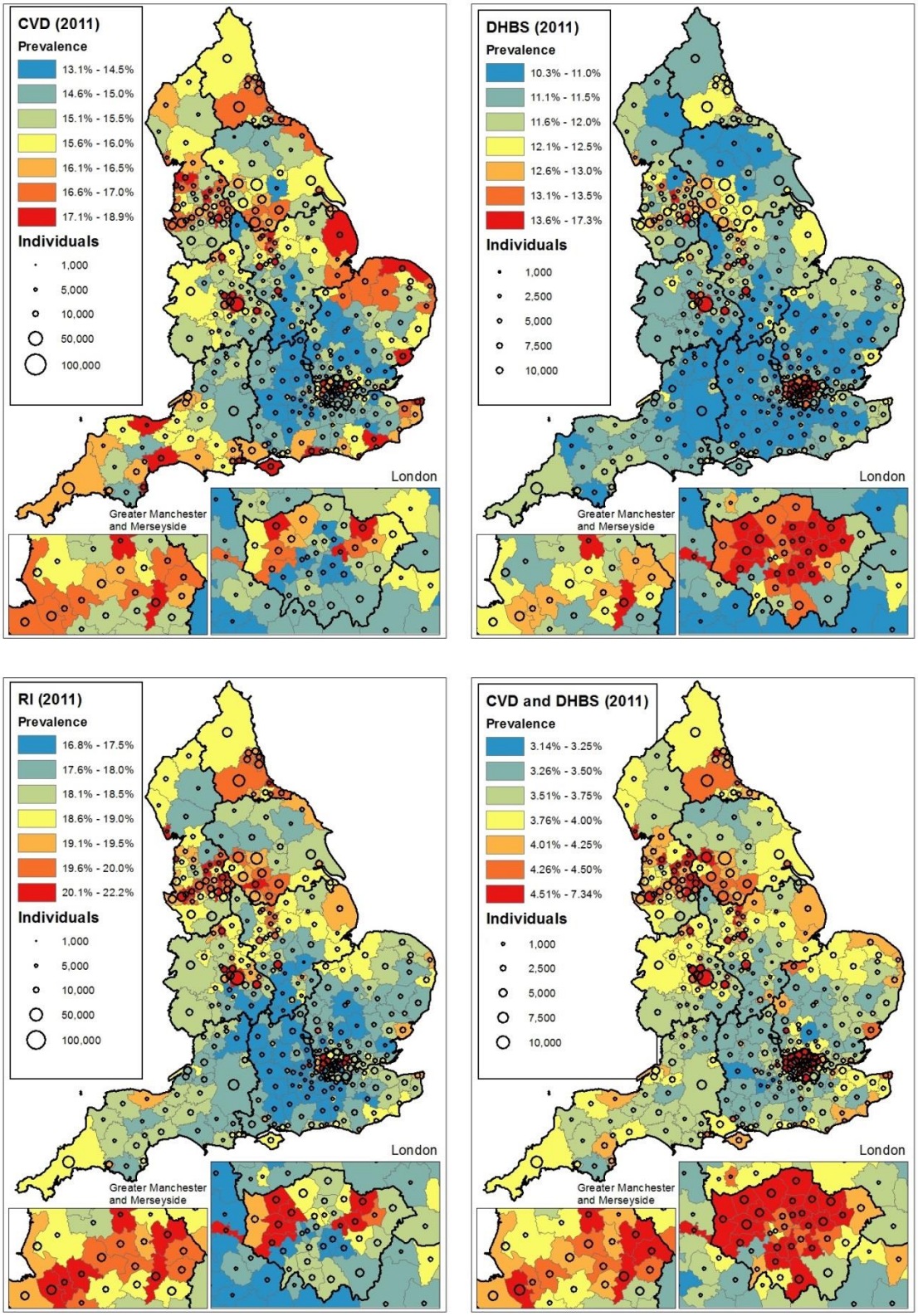


Figure 6-2 : Prevalences and number of individuals with morbidities, LADs, 2011, from top left, CVD, top right, diabetes or high blood sugar, bottom left, respiratory illness and botton right CVD and diabetes or high blood sugar

The 10 LADs with the highest and lowest estimated prevalences rates are shown in Table 6-3. For those LADs with high prevalences of CVD there is a range of types of

LAD, including urban cities, London boroughs and semi-rural LADs in southern England. There could be a number of explanations for this mix of LADs. Firstly, some LADs will have a population which is generally in poorer health, the causes of which could be employment history or lifestyle factors (e.g. Tendering). The second explanation could be due to the ethnic makeup of the area, with some ethnic groups tending to have a higher prevalence of the conditions that contribute to CVD (e.g. the south Asian community in Tower Hamlets which makes up a quarter of its 50 and older population). The final explanation could be related to a generally aged population living in the area, with the incidences of CVD accumulating in an older age structure (e.g. Eastbourne).

For DHBS perhaps the sole defining characteristic is the ethnic composition of the LAD, with all the top ten LADs having a high proportion of their population from the BME community, in particular the south Asian community, which tends to have higher rates of DHBS than the general population (see table 2.7 of HSCIC, 2005b). With respiratory illness there appears to be both an ethnic and deprivation dimension to those LADs in the top ten, indicating that this morbidity affects individuals differently depending on their ethnicity or deprivation. The LADs with the highest prevalence for the comorbidity of CVD and DHBS appears to be those most influenced by patterns of DHBS prevalence, with very similar LADs featuring in the top ten for both DHBS and CVD and DHBS.

The pattern for 10 LADs with the lowest prevalences appears clearer than for the 10 highest. The vast majority of LADs are prosperous authorities located in Southern England, but there are also some London boroughs reporting low prevalences. Clearly the very diverse and cosmopolitan nature of this category of area type can lead to some divergence in the expected outcomes with respect to some morbidities. A case in point is the London borough of Southwark which has the lowest estimated prevalence rate for respiratory illness and a low prevalence for CVD. Southwark is an authority with a large black African and Caribbean population (in the 2011 Census this ethnic group makes up 17% of the 50 and older population in Southwark). Again the report by the HSCIC (2005b) shows that illnesses due to the heart and circulatory system and to the respiratory system are lower for this ethnic group than the general population. Further, a study of standardised mortality ratios by Chaturved (2003) shows that people from this ethnic group, particularly men, have a mortality risk of dying from CVD of half that of the general population. In other Cosmopolitan London boroughs with a different ethnic mix (say a larger south Asian population) a different prevalence profile is relevant and this is borne out by an inspection of the top 10 prevalences in the same table.

Table 6-3 : Highest and lowest prevalences for CVD, diabetes or high blood sugar, respiratory illness and CVD and diabetes or high blood sugar, LADs, 2011

Rank	CVD		Diabetes or high blood sugar		Respiratory illness		CVD and DHBS	
1	Leicester (CI)	18.9%	Newham (LC2)	17.3%	Leicester (CI)	22.2%	Newham (LC)	7.3%
2	Tendring (CC)	18.1%	Brent (LC2)	17.1%	Tower Hamlets (LC)	22.0%	Leicester (CI)	7.2%
3	Sandwell (CI)	18.0%	Leicester (CI)	16.4%	Newham (LC2)	21.3%	Tower Hamlets (LC)	7.1%
4	Tower Hamlets (LC)	18.0%	Tower Hamlets (LC)	16.4%	Sandwell (CI)	20.9%	Brent (LC)	6.6%
5	Eastbourne (RC)	18.0%	Harrow (LS)	15.6%	Wolverhampton (CI)	20.9%	Harrow (LS)	6.2%
6	Wolverhampton (CI)	17.9%	Slough (LS)	15.5%	Blackburn with Darwen (CI)	20.8%	Hounslow (LS)	6.1%
7	Rother (CC)	17.9%	Hackney (LC2)	15.5%	Knowsley (IH)	20.7%	Ealing (LS)	6.1%
8	Christchurch (CC)	17.9%	Ealing (LS)	15.4%	Hounslow (LS)	20.7%	Slough (LS)	6.1%
9	Thanet (CC)	17.8%	Hounslow (LS)	15.3%	Birmingham (CI)	20.6%	Redbridge (LS)	5.9%
10	Wyre (CC)	17.6%	Redbridge (LS)	15.2%	Slough (LS)	20.6%	Manchester (CI)	5.6%
315	South Northamptonshire (PST)	13.6%	Elmbridge (PSE)	10.5%	Surrey Heath (PSE)	17.1%	South Hams (CC)	3.3%
316	Daventry (PST)	13.6%	Stroud (PST)	10.5%	South Northamptonshire (PST)	17.1%	South Cambridgeshire (PSE)	3.3%
317	Basingstoke and Deane (PSE)	13.5%	Bracknell Forest (PSE)	10.5%	Waverley (PSE)	17.1%	Bracknell Forest (PSE)	3.3%
318	Milton Keynes (NGT)	13.4%	South Northamptonshire (PST)	10.5%	Mid Sussex (PSE)	17.0%	Guildford (PSE)	3.3%
319	Lambeth (LC2)	13.4%	Wokingham (PSE)	10.5%	Guildford (PSE)	17.0%	South Northamptonshire (PST)	3.3%
320	Southwark (LC2)	13.3%	Uttlesford (PSE)	10.4%	Kensington and Chelsea (LC)	16.9%	Basingstoke and Deane (PSE)	3.3%
321	Hart (PSE)	13.3%	Chiltern (PSE)	10.4%	Hart (PSE)	16.9%	East Hertfordshire (PSE)	3.2%
322	West Berkshire (PSE)	13.2%	Mid Sussex (PSE)	10.4%	Rutland (PST)	16.8%	West Berkshire (PSE)	3.2%
323	Wokingham (PSE)	13.2%	East Hertfordshire (PSE)	10.4%	Tandridge (PSE)	16.8%	Surrey Heath (PSE)	3.2%
324	Bracknell Forest (PSE)	13.1%	West Berkshire (PSE)	10.3%	Southwark (LC2)	16.8%	Hart (PSE)	3.1%

Note : CI Centres with Industry; CC Coastal and Countryside; LC London Centre; RC Regional Centres; PST Prospering Smaller Towns; PSE Prospering Southern England; NGT New and Growing Towns; LC2 London Cosmopolitan; LS London Suburbs; IH Industrial Hinterlands

6.6 ROBUSTNESS OF SPATIAL POPULATION

The approach adopted by the FMF involves a stochastic mechanism that operates both through the selecting and de-selecting individuals and the 'annealing' decision on whether to accept a deterioration in model fit. These decisions are made through the generation of a sequence of pseudo random numbers and the values in this sequence are determined by a seed value. By varying this seed value a different sequence of random numbers will be used in the FMF. To explore the sensitivity of the outcomes to this seed value two additional runs of the FMF are conducted with differing seed values and, to assess the sensitivity of the final fit to this value, two comparisons are made. Firstly the final TAE value for each LAD in each run is compared and secondly the sensitivity of an outcome of interest, the prevalence of CVD, DHBS and respiratory illness is compared across runs.

6.6.1 Stochastic TAE fit

By comparing the TAE at the end of the spatial microsimulation an aggregate picture is obtained of the consistency of the optimisation process when using different seeds. Ideally that each optimisation would converge to a similar TAE value, and a scatter plot of these values will lie on a 45° line that passes through the origin. Figure 6-3 shows these scatter plots between pairs of spatial microsimulations runs with different seed values and the correlation coefficient (r^2) between the two measures.

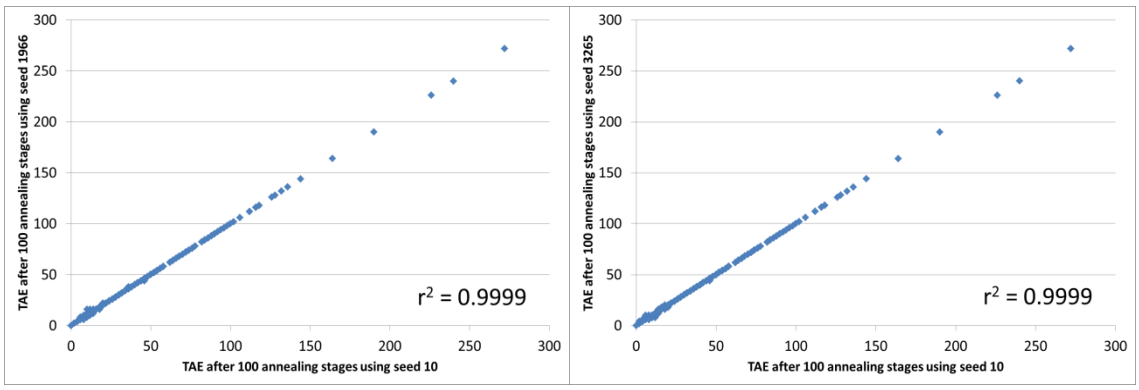


Figure 6-3 : Comparison of final TAEs from three spatial microsimulations

These plots clearly show that the quality of the final model fit is not sensitive to the random number seed used, with the correlations being near perfect.

6.6.2 Stochastic prevalence estimates

Whilst Figure 6-3 is re-assuring that the model fits in aggregate are consistent, there is value in performing a more detailed comparison of how an outcome of interest is influenced by the random number seed. In section 6.5 a method of estimating a morbidity outcome is defined, and measured as a prevalence rate for each LAD. Here

the spatially microsimulated populations generated for the TAE comparison in section 6.6.1 is used to estimate two additional set of LAD prevalences. Once again a comparison of the consistency of these estimates is made using a scatter plot and again the points should line on a 45° straight line that passes through the origin. These plots are shown in Figure 6-4 along with the linear correlation coefficient r^2. Whilst there is some scatter, the correlation between the two sources of estimates is reassuring and high. Once again this provides the reassurance that the composition of the spatially microsimulated population is not sensitive to the random number seed chosen.

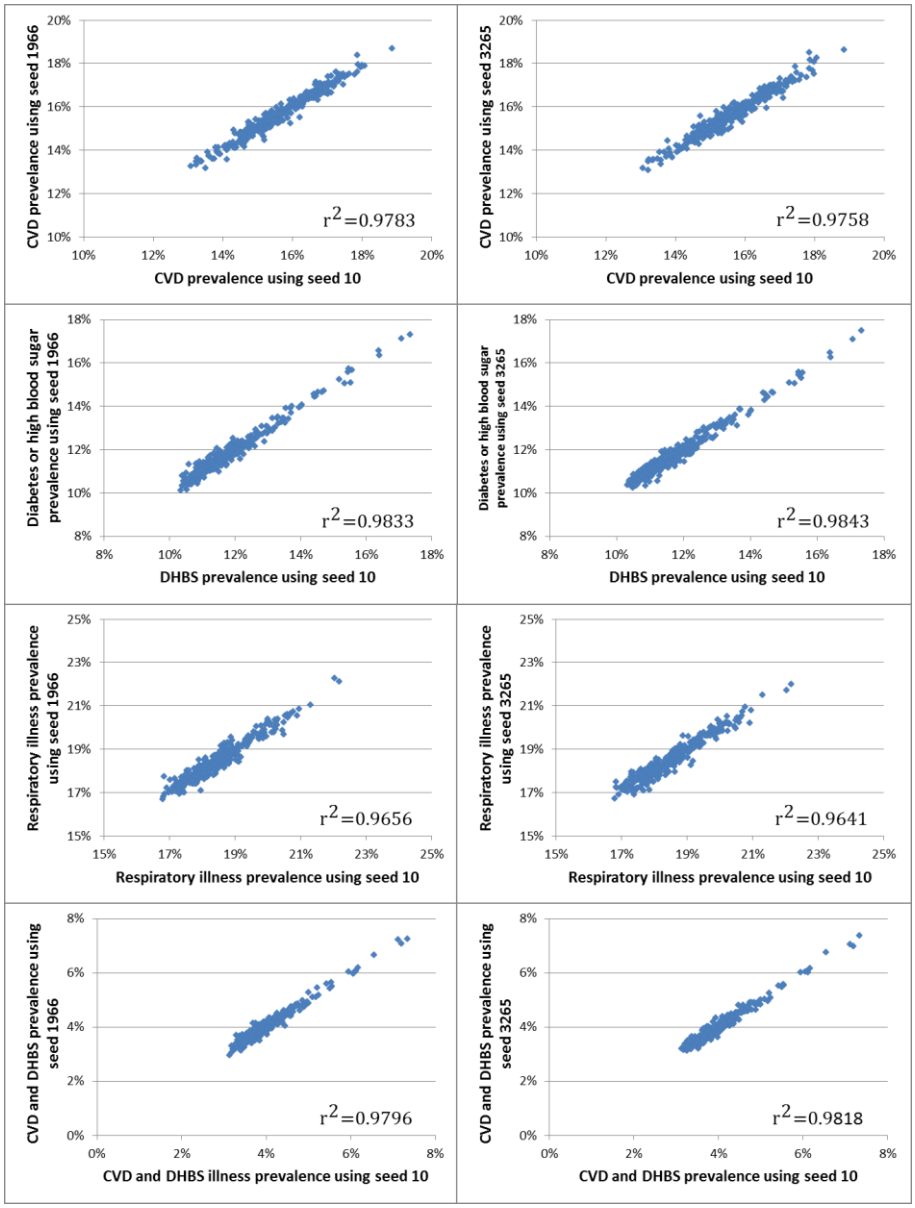


Figure 6-4 : Comparison of LAD prevalences from three spatial microsimulations, from top : CVD, diabetes or high blood sugar, respiratory illness and CVD and diabetes or high blood sugar

6.7 MODEL PARAMETERS

In addition to specifying the random number seed the FMF also allows for the variation of various parameters used to govern the optimisation process. The first three parameters (maximum number of steps, maximum number of improvement steps and number of improvement to the optimisation before continuing) control the looping within the optimisation (Harland, 2013, page 30) and the larger the values the greater the search space that is explored. To check that the default parameters are sufficient to explore the search space, an additional run was conducted with these values increased by 30% (from 100 to 130 for the first two parameters and from 10 to 13 for the third, and the execution time increases from 12 to 20 hours as a consequence). The output from this run is compared with the run using the default parameters and a random seed of 10.

The fourth parameter is the annealing reduction factor (set at the recommended value of 0.9). The larger this factor the more the search space is explored. This factor is varied to a larger value of 0.95 and two smaller values of 0.85 and 0.80. With a factor of 0.95 the execution time increases to 18 hours, whilst with factors of 0.80 and 0.85 the execution time reduces to 10 hours. Again these three runs are compared with the run using the default parameters and a random seed of 10.

Table 6-4 : Comparison of final TAE from alternative parameter runs

Alternative model	Correlation with default parameters and seed of 10
Looping parameters plus 30%	0.9999
Annealing reduction factor of 0.95	0.9999
Annealing reduction factor of 0.85	0.9993
Annealing reduction factor of 0.80	0.9992

Table 6-4 provides the linear correlation coefficient in the final TAE values between the run with the default parameters and a seed of 10 and the four alternative runs described here. The top two rows show that the increase in the search space does not change the final TAE's by much. The bottom two rows show marginal changes in the final TAE, the more so with the lowest factor. A similar comparison of the morbidity outcomes from these four alternative runs with the default parameter run is provided in Table 6-5. Again there is a reassuring agreement between the estimates from the default parameter run and these alternatives.

Table 6-5 : Comparison of prevalence counts from alternative parameter runs

Alternative model	CVD	DHBS	Respiratory Illness
Looping parameters plus 30%	0.9731	0.9865	0.9629
Annealing reduction factor of 0.95	0.9781	0.9829	0.9659
Annealing reduction factor of 0.85	0.9775	0.9852	0.9673
Annealing reduction factor of 0.80	0.9773	0.9868	0.9642

From the results in Table 6-4 and Table 6-5 there is reassurance that the optimisation process is robust to the choice of the parameters. There is some evidence to suggest that the default parameters maybe too ambitious, causing the FMF to do more work than is strictly necessary, with lower parameter values possibly being acceptable. However the extra work is not too arduous (12 hour runs are eminently doable) and if the parameter values are reduced too far then the differentiation between LADs contained in the constraints would not be re-produced.

6.8 VALIDATION

In section 6.3 vehicle ownership or use and NS-SEC are chosen as validation variables to test how well the spatially microsimulated population re-produces a measure of interest that is not used as a constraint. To assess the level of agreement between the actual 2011 Census counts and the derived counts from the spatially microsimulated population a series of scatter plots are used. If the level of agreement between the two counts is good then most of these points should lie along a straight 45° line that passes through the origin. The degree of scatter away from this 45° line can be measured by the correlation coefficient.

6.8.1 Vehicle ownership or use

Figure 6-5 shows four scatter plots for the levels of vehicle ownership within the four sub-populations that are males aged between 50 and 64 (M50), males aged 65 and older (M65), females aged 50 to 64 (F50) and females aged 65 and older (F65). The same counts are also shown in Figure 6-6, however here each of the three scatter plot contains gender and age plots for each of the three levels of ownership. Looking at these charts it is clear that the ownership of one vehicle is the most common category, irrespective of gender and age. Also owning two or more vehicles is more common amongst the younger age groups, irrespective of gender.

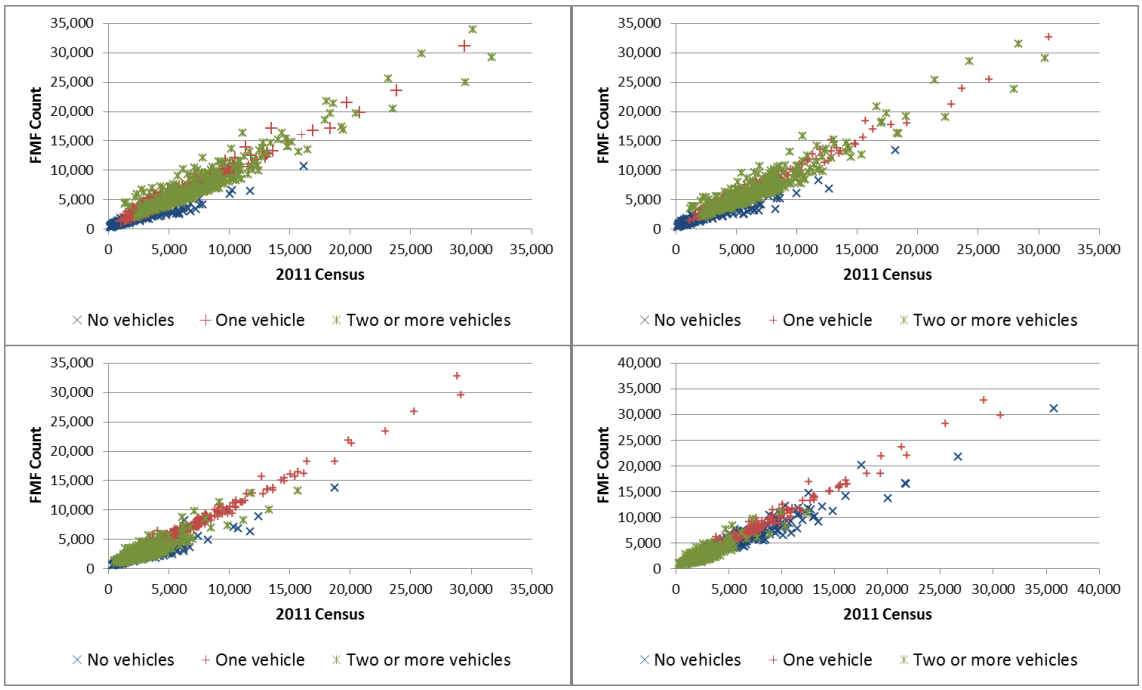


Figure 6-5 : Scatter plot of vehicle ownership or use by gender and age categories, LADs, 2011, clockwise from top left: males aged 50 to 64, females aged 50 to 64, females aged 65 and older and males aged 65 and older

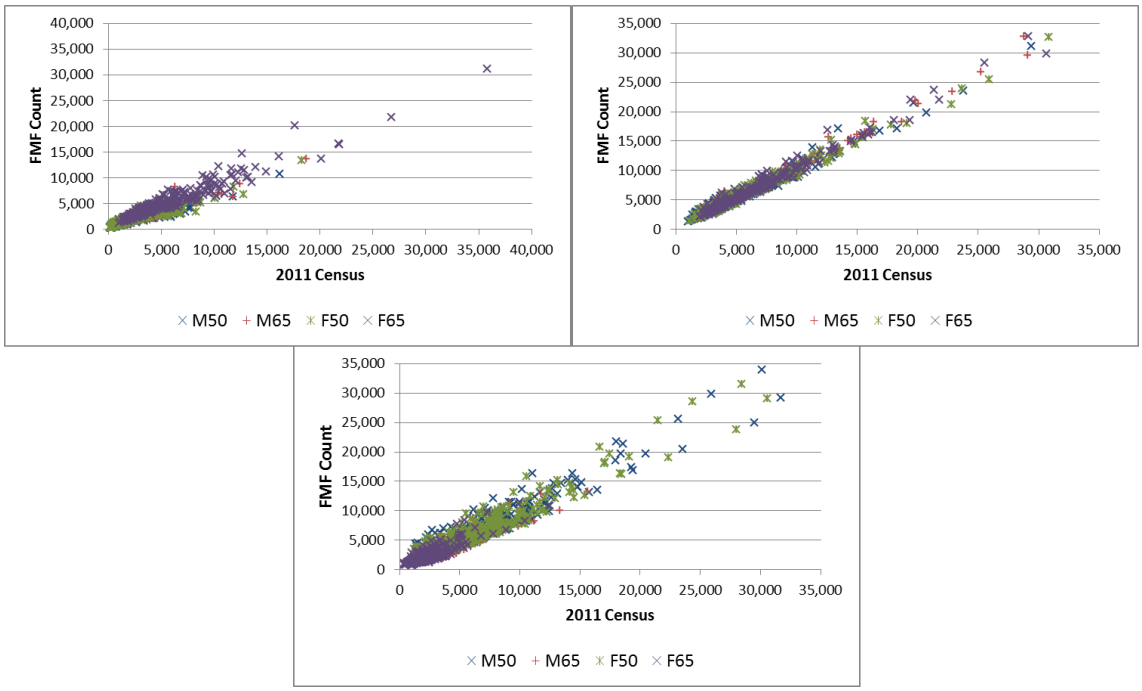


Figure 6-6 : Scatter plot of vehicle ownership or use by level of ownership or use categories, LADs, 2011, top row, left to right: no vehicles, one vehicle, bottom row, two or more vehicles

The level of agreement between the two counts is generally good, with a slight tendency for the spatially microsimulated population to under-predict the number of individuals living in households without a vehicle and a consequent slight over prediction of the

number of multiple vehicle owning households. The agreement for individuals living in households with one vehicle is particularly good. The correlation coefficients in Table 6-6 also shows that the degree of spread is least for individuals living in one vehicle households.

Table 6-6 : Correlation coefficient between the census counts and the FMF derived counts for vehicle ownership

Gender	Age	Vehicle Ownership	r^2
Male	Aged 50 to 64	No vehicles	0.9465
Male	Aged 50 to 64	One vehicle	0.9830
Male	Aged 50 to 64	Two or more vehicles	0.9440
Male	Aged 65 and older	No vehicles	0.9232
Male	Aged 65 and older	One vehicle	0.9913
Male	Aged 65 and older	Two or more vehicles	0.8777
Female	Aged 50 to 64	No vehicles	0.9383
Female	Aged 50 to 64	One vehicle	0.9871
Female	Aged 50 to 64	Two or more vehicles	0.9441
Female	Aged 65 and older	No vehicles	0.9610
Female	Aged 65 and older	One vehicle	0.9891
Female	Aged 65 and older	Two or more vehicles	0.8631

6.8.2 NS-SEC

Figure 6-7 shows four gender by age scatter plots for NS-SEC with the four sub-populations (the equivalent of Figure 6-5) and Figure 6-8 shows six NS-SEC scatter plots within which are plotted for the four gender by age categories (the equivalent of Figure 6-6). The level of agreement between the two counts here is less good than for vehicle ownership or use. There appears to be good agreement for the intermediate (class 2) and semi-routine and routine (class 5) classes. The other class of never worked or long-term unemployed (class O) is particularly poorly re-produced. This under prediction is coupled with an over prediction of higher managerial, administrative and professional occupations (class 1) and lower supervisory and technical occupations (class 4). The remaining class, small employers and own account workers (class 3) appears to be slightly under estimated in the spatially microsimulated populations.

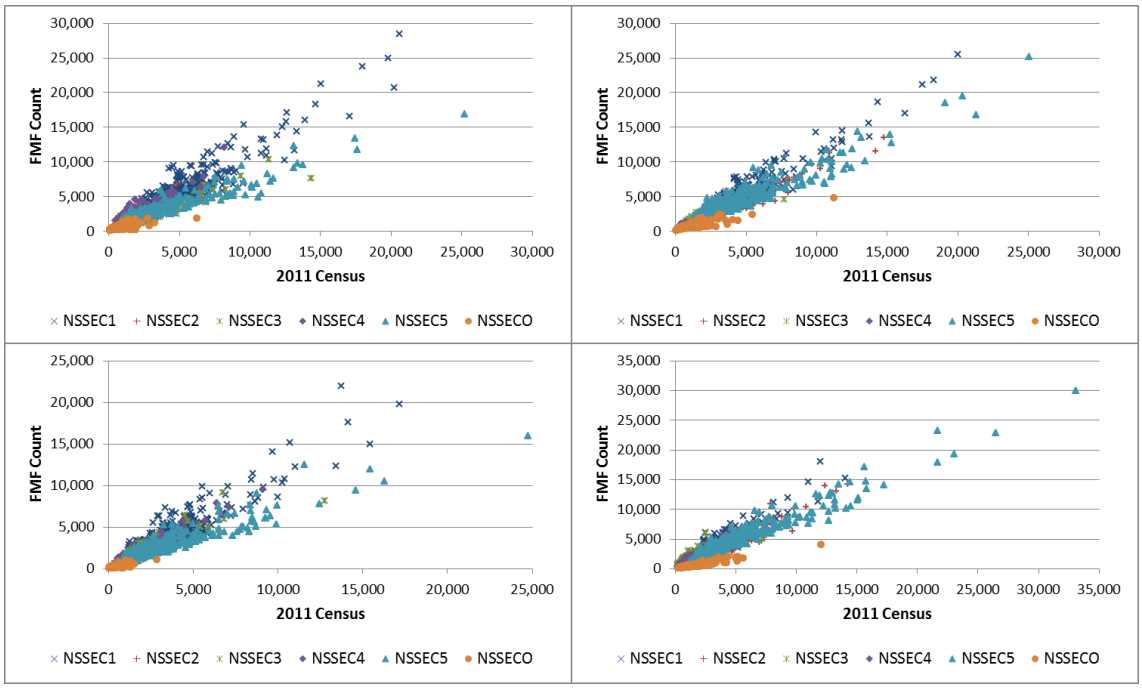


Figure 6-7 : Scatter plot of NS-SEC by gender and age categories, LADs, 2011, clockwise from top left: males aged 50 to 64, females aged 50 to 64, females aged 65 and older and males aged 65 and older

Clearly the spatially microsimulated population has not re-produced well the nature of some LADs by their NS-SEC population. The performance is particularly poor for the group that defines those who have never worked or are long term unemployed. It is reasonable to assume that some of the reasons why individuals find themselves in this group are health related.

Looking at the degree of scatter measured by the correlation coefficients in Table 6-7 the scatter is greatest for class 3 and class O categories, which depending on the gender and age range.

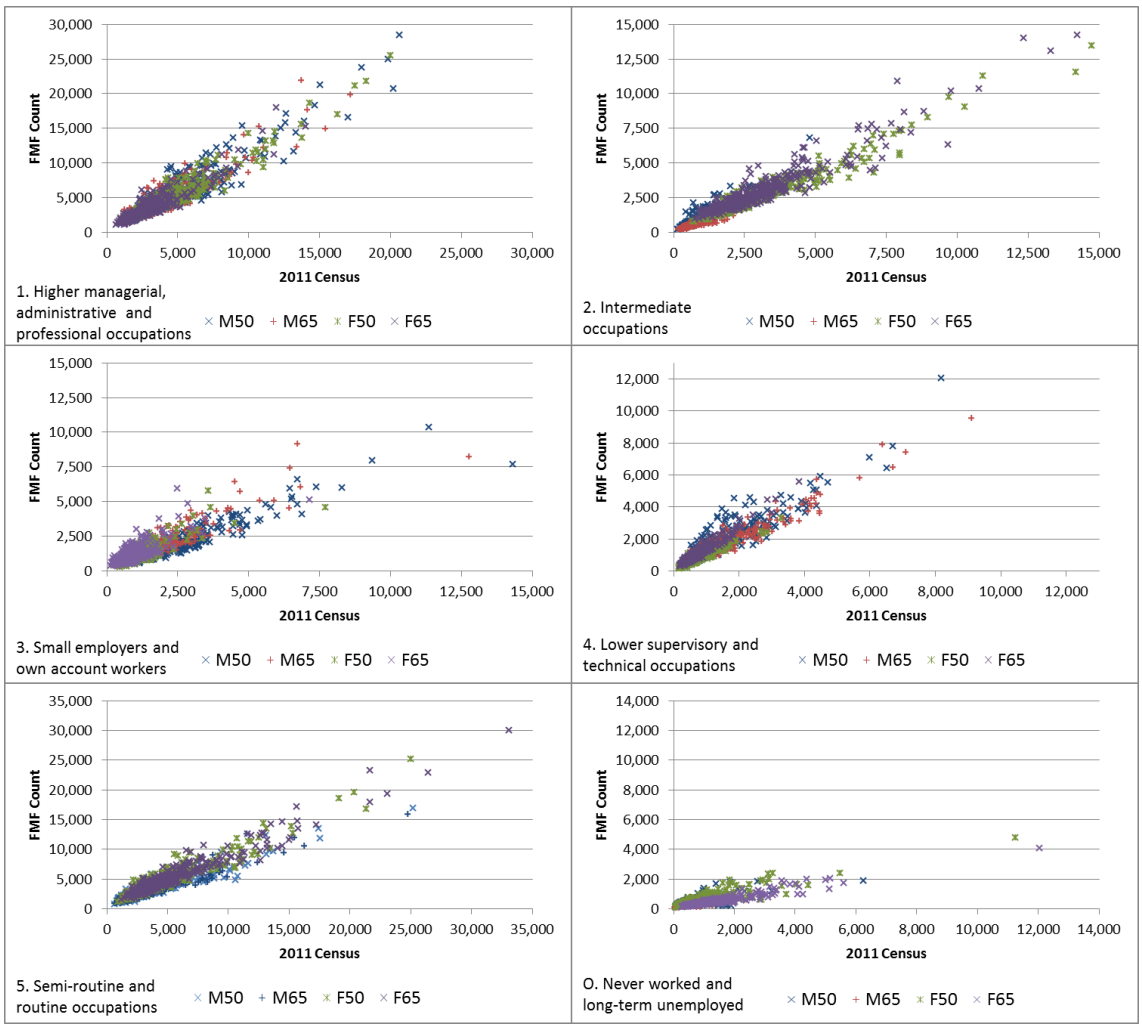


Figure 6-8 : Scatter plot of NS-SEC by NS-SEC category, LADs, 2011 (as labelled).

6.9 CONCLUSION

In this chapter a description is provided of the process used to define and select the set of constraints to use for the calibration of the LAD specific weights and those to use for validation. This is in essence a two stage process. The literature informs the choice of constraint variables in the first stage, covering gender, age, ethnicity and disability. Here these constraints are largely givens, with the hazard model providing some context to the relative strength of each of these variables in determining the incidence for the morbidities. The literature also suggests that there are also socio-economic factors that contribute to health outcomes in an area. A range of such factors are considered in turn and using the outcome of hazard modelling, three are selected, either as constraint or validation variables. From an inconsistent set of results, the variable living arrangements is chosen as the final constraint variable, leaving vehicle ownership or use and NS-SEC classification as validation variables.

Table 6-7 : Correlation coefficient between the census counts and the FMF derived counts for NS-SEC

Gender	Age	NS-SEC	r^2
Male	Aged 50 to 64	class 1	0.9203
Male	Aged 50 to 64	class 2	0.9237
Male	Aged 50 to 64	class 3	0.9355
Male	Aged 50 to 64	class 4	0.9089
Male	Aged 50 to 64	class 5	0.9401
Male	Aged 50 to 64	class O	0.6954
Male	Aged 65 and older	class 1	0.9049
Male	Aged 65 and older	class 2	0.9240
Male	Aged 65 and older	class 3	0.8830
Male	Aged 65 and older	class 4	0.9515
Male	Aged 65 and older	class 5	0.9246
Male	Aged 65 and older	class O	0.7149
Female	Aged 50 to 64	class 1	0.9499
Female	Aged 50 to 64	class 2	0.9679
Female	Aged 50 to 64	class 3	0.8189
Female	Aged 50 to 64	class 4	0.9517
Female	Aged 50 to 64	class 5	0.9468
Female	Aged 50 to 64	class O	0.9063
Female	Aged 65 and older	class 1	0.9211
Female	Aged 65 and older	class 2	0.9413
Female	Aged 65 and older	class 3	0.7742
Female	Aged 65 and older	class 4	0.9644
Female	Aged 65 and older	class 5	0.9620
Female	Aged 65 and older	class O	0.9452

The weight estimation process is seen to produce fits that, whilst not perfect, are very good, with small TAE's when measured relative to the size of the population being synthesised. Given the availability of a spatially microsimulated population for each LAD and information on the morbidity status of the individuals in this population, it is possible to estimate prevalence counts and rates for the morbidities in each LAD. The general picture is for these prevalence rates to be lowest in the more affluent LADs in southern England. The picture for those authorities with the highest prevalence is more mixed and specific to the morbidity. CVD is seen to be potentially influenced by a variety of factors, including deprivation; ethnicity and an aged population. With DHBS the picture is simpler; LADs with a high percentage south Asian population have high prevalences. For respiratory illness the main influence appears to be living in a densely populated urban area, which itself is probably an occupation or lifestyle impact.

Recognising that the spatial microsimulation process is driven to a large degree by a stochastic mechanism, a number of duplicate model runs are conducted with differing random number seeds. The consistency of these runs is measured in aggregate through the final TAE for model fits and in more detail through the morbidity estimates mentioned above. In both cases the agreement between the outcomes from each model run is seen to be satisfactory, allaying any concerns over the sensitivity of the outputs to this random number seed value.

The final piece of work reported on here are the validation counts derived from the spatially microsimulated population against 2011 Census counts of vehicle ownership or use and the NS-SEC. The validation for vehicle ownership is seen to be good, with a slight under estimate in the spatially microsimulated population of individuals in households without a vehicle and a commensurate increase in those with one vehicle. For NS-SEC the validation is less good, in particular the number of people who had never worked or are long-term unemployed (class O) are underrepresented in the spatially microsimulated population.

Notwithstanding this issue with NS-SEC class O, the outcome from this exercise is a spatially population that reflects the composition of each LADs 50 years and older population at the time of the 2011 Census. This is the base population that the health microsimulation evolves through time. This process is described in section 5.5 and the results shown in Chapter 9. Before then, however, the impact of the revision process on the ETHPOP population projection is described in Chapter 7 and the justification for the probabilities used in the health microsimulation is described in Chapter 8.

7 ETHPOP REVISIONS

Development of Projected Ethnic Populations

7.1 INTRODUCTION

As discussed in the methodology chapter, the purpose of the ETHPOP population projections is to re-structure the future year's population of those aged 50 and older in each LAD so that it represents some projected ethnic structure and size. The motivation behind this desire is that, firstly, the size of the population will have an impact on the number of individuals with a specific morbidity. Secondly, the literature has shown that the diversity of the population, by gender, age and ethnicity also impacts on the prevalence rates for some morbidities (HSCIC, 2005b).

This chapter begins in section 7.2 with a discussion of the performance of the original ETHPOP population projections in re-producing the ethnicity counts in the 2011 Census. In this section the performance is assessed using aggregate measures and a ranking of best and worst prediction performances and, in section 7.3, the predictive performance for the seven case study LADs are discussed in detail. The chapter then provides a discussion of the adjustment factors that form a pivotal role in revising the ETHPOP projections from 2013 to 2031 in section 7.4. Section 7.5 gives an overview of the impact of the revisions by comparing the original and revised ETHPOP projections for 2031. This impact is further explored in the context of the seven case study LADs in section 7.6. The chapter finishes with a summary of these findings and the potential issues raised. In these analyses, the ETHPOP data was obtained as a bulk download of the ethnic population projections available from the ETHPOP web site (University of Leeds, 2011).

7.2 COMPARISON OF 2011 ETHPOP PROJECTIONS AND 2011 CENSUS COUNTS

At the time of the production of the original ETHPOP population projections, the 2011 Census outputs were not available as a potential source of information on medium term trends in the structure of the ethnic population within LADs. The subsequent publication of 2011 Census data does, however, allow a validation check on the performance of the ETHPOP projections to be carried out.

The task of producing population projections is a complex one that is subject to a number of assumptions that, whilst having historic veracity, many not hold well even for the short or medium term. Of the three main drivers of the demographic process, it is the mortality and migration drivers that are likely to influence most the size and

structure of the 50 year and older population. Figure 3-1 illustrates these national mortality assumptions used in ONS projections. The general trend is for these assumptions to become more optimistic with each new set of projections – life expectancies increase. Migration is an even more complex phenomenon. Not only is it difficult to forecast what the future trends in migration might be, it also difficult to measure historic or current trends (see ONS, 2014h, of a review of the performance of migration methodologies). These difficulties in forecasting mortality and particularly migration are further compounded by the desire to measure and apply these at a sub-national geography, e.g. English LADs. In particular migration now needs to capture the hitherto irrelevant internal migrations between these sub-national geographies.

Added to the complexity of a detailed geography, there is further complexity if the definition of the population is expanded beyond gender and age, as is the requirement here by ethnicity. There is thus a need to measure and project these two significant demographic processes by ethnicity. Ethnic identification can itself be fluid, with the same individual identifying themselves as different ethnic groups over time, depending on the range of groups offered or their own assessment (Simpson et al., 2014).

From an initial comparison at a national level between the 2011 ETHPOP projections and the 2011 Census output (Rees et al, 2013c), the recommendation is that the best outcome at prediction is reached if the mean of the TREND and UPTAP-ER scenarios are taken. This advice is followed in this study. The same report hints at some of the reasons behind the discrepancies identified by the comparison. The most relevant here is a tendency for the trended life expectancies of the projected population to be too high, particularly for the White ethnic group.

Here the correspondence between the averaged ETHPOP TREND and UPTAP-ER projections and the 2011 Census outputs is judged using Census table DC2101, which provides cross tabulations of gender by age by ethnic group for LADs in England. The ages are banded into five year bands, providing eight such bands for the over 50's. The ethnic groups are detailed by sixteen classes but for comparative purposes these groups are aggregated here into five broad groups to form: White, Mixed, Asian, Black and Other groups with the Chinese class re-allocated from Asian to Other (to maintain consistency with ELSA categorisations).

7.2.1 Difference by structure

The first comparison is at a national scale between the ETHPOP projections and the 2011 Census counts and is provided in the form of population structure diagrams. Looking at this figure, the ETHPOP projections are good for the younger 50 to 64 White ethnic group, but thereafter there is a tendency for the ETHPOP projections to be greater than the Census counts. This confirms that ETHPOP has under estimated the mortality for these older age bands. For the Mixed group, ETHPOP under predicts at all ages above 50, especially for the youngest 50 to 54 year old age band. The prediction performance for the Asian group is the opposite of that for the White group. Here the prediction performance is worst for the younger age bands (ETHPOP under predicts) but is generally good for the older age bands. The prediction performance for the Black ethnic group is good, especially for females. The prediction performance for Other females is also good but for Other males the performance is poor, with ETHPOP consistently under predicting the numbers in this ethnic group.

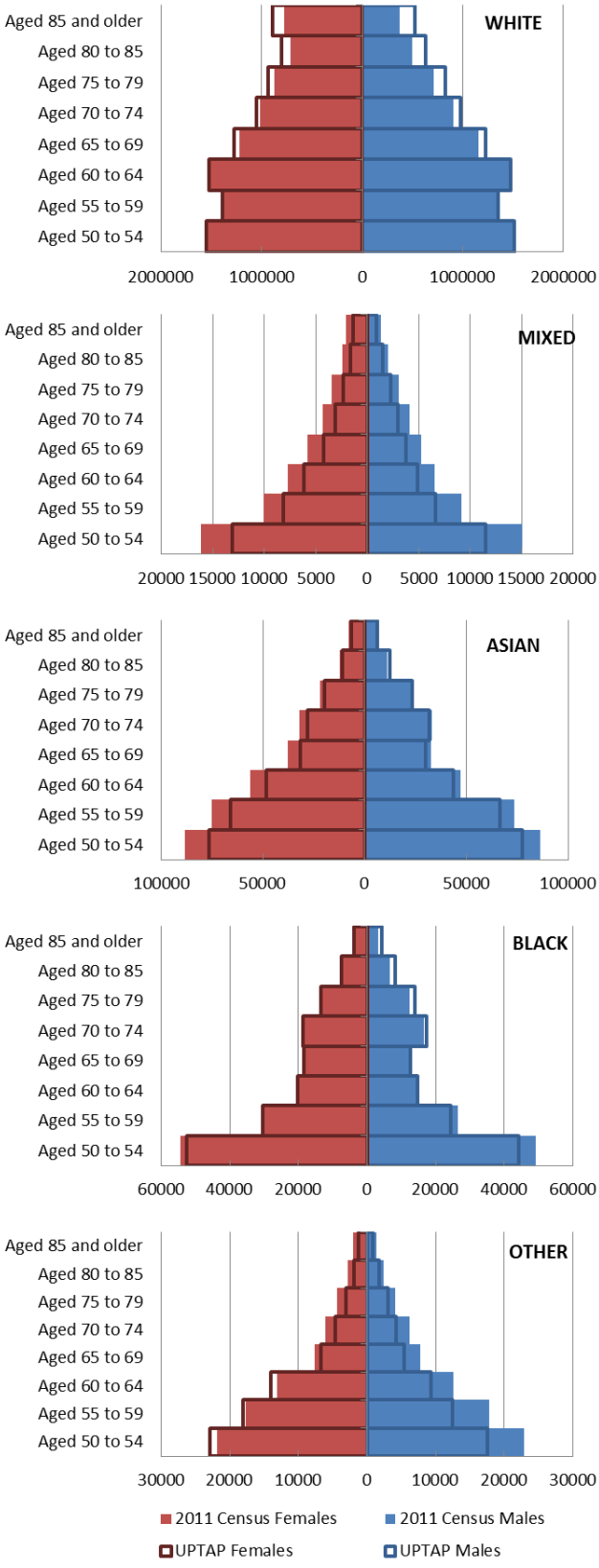


Figure 7-1 : Comparison of English population structures, 2011, from top: White, Mixed, Asian, Black and Other

7.2.2 Direction of differences

Subject to the qualifications mentioned above, the projection performance on a national scale appears reasonable; however, there may be local variations in this performance. In the remainder of this section these local measures of difference between the ETHPOP projections and the 2011 Census counts are explored. To quantify the agreement between the two counts a Total Error (TE) statistic is calculated for each LAD using equation 7-1.

$$TE\%_e^{LAD} = \frac{\sum_{g=1}^{2}\sum_{a=1}^{8}(E_{g,a,e}^{LAD,2011} - C_{g,a,e}^{LAD,2011})}{\sum_{g=1}^{2}\sum_{a=1}^{8}(C_{g,a,e}^{LAD,2011})} \quad 7\text{-}1$$

where: $TE\%_e^{LAD}$ is the total error percentage between the ETHPOP and Census population counts for ethnicity e in local authority LAD;

$E_{g,a,e}^{LAD,2011}$ is the ETHPOP sub-population of gender g, aged a, of ethnicity e, in LAD for 2011;

$C_{g,a,e}^{LAD,2011}$ is the 2011 Census sub-population of gender g, aged a, of ethnicity e, in LAD for 2011.

Note : $C_{g,a,e}^{LAD,2011}$ and $E_{g,a,e}^{LAD,2011}$ are not contemporaneous (27th March vs 20th June respectively). Whilst this is a cause of genuine difference, it is likely to be of little practical importance when compared with other sources of variation.

This measure provides an indication of the size of the aggregate discrepancy relative to the size of the population and the direction of this discrepancy is preserved, either as an under or an over prediction. If the percentage is positive then the ETHPOP projections are, in aggregate, over estimates (i.e. greater than those from the 2011 Census) and if it is negative then the ETHPOP projections are under estimates. To prevent any errors calculated on small counts inflating the outcome, if the Census population (i.e. the denominator of equation 7-1) is less than 160 then the calculation is not performed (this value is the equivalent of an average of ten people in each gender and age band combination).

A frequency tabulation of TE% for each of the five broad ethnic groups is shown in Table 7-1. The smallest range of errors is for the White ethnic group. This is also the most symmetric group but still with a bias towards over prediction by ETHPOP. In contrast, all the remaining ethnic groups have been under predicted by ETHPOP. The scale of under prediction is particularly poor for the Mixed group, where 75 LADs have an under prediction of more than 30%. There are also a large number of LADs where the size of the aged 50 and older ethnic populations, for other than the White and Asian groups, is fewer than 160.

Table 7-1 : Distribution of LADs by total error percentages for each ethnic group

		White	Mixed	Asian	Black	Other	ALL
	Population fewer than 160	0	139	74	181	134	0
ETHPOP under predicts the	Less than -30%	0	75	39	15	53	0
	Between -25% and -30%	0	18	27	13	21	0
	Between -20% and -25%	0	23	22	13	18	0
	Between -15% and -20%	2	28	27	14	16	2
	Between -10% and -15%	5	20	39	21	11	4
	Between -5% and -10%	37	9	27	21	9	39
	Between 5% and -5%	151	10	39	25	20	162
ETHPOP over predicts the	Between 5% and 10%	54	1	7	7	11	59
	Between 10% and 15%	33	1	7	7	6	32
	Between 15% and 20%	14	0	4	4	5	9
	Between 20% and 25%	9	0	1	1	5	9
	Between 25% and 30%	7	0	4	0	5	6
	More than 30%	12	0	7	2	10	2

Looking at the worst performing LAD's both in terms of under and over prediction in Table 7-2, some commonalities are present. The ten LADs that are most under predicted for the White ethnic group are located in the south of England (except Ryedale in North Yorkshire) and all are prosperous in nature. The most over predicted are from Greater London or close by (Slough) and each is characterised by a cosmopolitan mix of ethnic groups. For the remaining four ethnicities there appears to be less of a pattern, with LADs that are geographically dis-located, affluent and deprived, and a range of ethnic diversities appearing in both the top ten under and over predictions. The performance of Cornwall and the Isles of Scilly is particularly bad for the black ethnicity, but this is due to the small population for the denominator in equation 7-1 (the values for the calculation are 393/179 = 220%).

7.2.3 Extent of difference

The calculation of the total error percentage in equation 7-1 allows cancellation of under and over predictions amongst the gender and age band combination to take place. If however the absolute value of these differences is taken, then this cancellation effect is removed:

$$TAE\%_e^{LAD} = \frac{\sum_{g=1}^{2}\sum_{a=1}^{8}\left|E_{g,a,e}^{LAD,2011} - C_{g,a,e}^{LAD,2011}\right|}{\sum_{g=1}^{2}\sum_{a=1}^{8}(C_{g,a,e}^{LAD,2011})} \qquad 7\text{-}2$$

where $TAE\%_e^{LAD}$ is the total absolute error percentage between the ETHPOP and Census population counts for ethnicity e in local authority LAD.

Table 7-2 : Top and bottom ten LADs by total error percentage performance

	White		Mixed		Asian		Black		Other	
1	South Hams	-17%	East Lindsey	-58%	Rushmoor	-74%	Eastbourne	-64%	Wolverhampton	-71%
2	Rother	-16%	Thanet	-55%	Chichester	-63%	Basildon	-47%	Sandwell	-68%
3	South Somerset	-12%	Blackpool	-54%	Vale of White Horse	-47%	Southend-on-Sea	-46%	Oadby & Wigston	-68%
4	Ryedale	-11%	Barnsley	-54%	Tunbridge Wells	-45%	Woking	-46%	South Tyneside	-66%
5	Mid Suffolk	-11%	Wirral	-53%	Swindon	-45%	Wakefield	-43%	Poole	-61%
6	West Dorset	-11%	Test Valley	-53%	Ashford	-44%	Broxbourne	-41%	Blackburn with Darwen	-61%
7	Purbeck	-10%	Lewes	-52%	Surrey Heath	-44%	Dartford	-34%	Bury	-56%
8	East Devon	-10%	Eastleigh	-51%	Mole Valley	-43%	Hertsmere	-34%	Gravesham	-56%
9	Shepway	-9%	South Kesteven	-51%	Harborough	-43%	Thurrock	-33%	Tonbridge & Malling	-55%
10	Adur	-9%	Halton	-51%	Worthing	-42%	Stoke on Trent	-33%	Enfield	-55%
315	Lewisham	35%	Cherwell	-3%	South Tyneside	27%	Luton	14%	Shropshire	31%
316	Newham	36%	Brighton & Hove	-2%	Sheffield	27%	Kingston upon Hull	14%	Canterbury	34%
317	Tower Hamlets	38%	Norwich	-2%	Portsmouth	27%	Slough	15%	Merton	40%
318	Islington	39%	Luton	-1%	Harrogate	31%	Newham	17%	Swindon	41%
319	Slough	39%	Guildford	-1%	Exeter	39%	Bolton	18%	Oxford	41%
320	Southwark	42%	Woking	-1%	Kingston upon Hull	46%	Redbridge	18%	Runnymede	41%
321	Hammersmith & Fulham	42%	Aylesbury Vale	0%	County Durham	62%	Cheshire East	20%	Reading	46%
322	Wandsworth	50%	Sheffield	0%	Northumberland	64%	Wigan	22%	Newham	50%
323	Hackney	58%	Kingston upon Thames	6%	Shropshire	79%	Wiltshire	31%	County Durham	60%
324	Lambeth	65%	Ealing	12%	Cornwall & Isles of Scilly	105%	Cornwall & Isles of Scilly	220%	Norwich	71%

The distribution of TAE% is shown in Table 7-3. Here there is no direction of prediction, either under or over. Instead this calculation assesses the size of the error, irrespective of direction. Very few LADs have a TAE% of less than 5% and even for the best performing White ethnic group, the minority have TAE%s less than 10%. The remaining ethnicities have substantial numbers of LADs with TAE%s greater than 30%.

Table 7-3 : Distribution of LADs by total absolute error percentages for each ethnic group

	White	Mixed	Asian	Black	Other	ALL
Population fewer than 160	0	139	74	181	134	0
Less than 5%	8	0	0	1	0	5
Between 5% and 10%	138	0	23	12	0	140
Between 10% and 15%	120	5	32	20	2	127
Between 15% and 20%	26	19	43	25	11	30
Between 20% and 25%	13	17	42	26	36	11
Between 25% and 30%	7	31	38	16	32	6
More than 30%	12	113	72	43	109	5

The ten best LAD and ten worst LAD performers on the TAE% calculation are shown in Table 7-4. The pattern seen for the White ethnic group in Table 7-2 holds true in this table too. Looking at the remaining ethnicities, the large urban authorities tend to perform best on this calculation, particularly for the Asian and Black ethnicities. The poor performance for Cornwall and the Isles of Scilly with the Asian and Black ethnicities persists into this table.

The ETHPOP authors are aware of the errors in their projections (Rees et al., 2015) and are working on a complete revision of their projections but these new results are not expected until September 2015. Therefore in the meantime there is a need to implement an approximate correction that revises the ETHPOP projections so that they align with published 2001 and 2011 Census outputs, this correction was described in Section 5.4 and the outcome of which is described here in Section 7.5.

7.3 COMPARISON FOR CASE STUDY AUTHORITIES

Whilst an overall assessment of projection performance might be gained from the appearance of Figure 7-1, the evidence in Table 7-2 and Table 7-4 suggests that there are local differences in this projection performance. In section 5.7 a number of case study LADs are identified using their representativeness of other LADs with the same ONS Super Group area classification. Here the population structure diagrams that compare the averaged 2011 ETHPOP projections and 2011 Census counts for these LADs are provided, with an accompanying discussion.

Table 7-4 : Best ten and worst ten LADs by total absolute error percentage performance

	White		Mixed		Asian		Black		Other	
1	South Gloucestershire	3%	Brighton & Hove	11%	Bradford	6%	Waltham Forest	5%	Lambeth	12%
2	Huntingdonshire	3%	Hillingdon	12%	Birmingham	6%	Croydon	6%	Southwark	12%
3	Erewash	4%	Sheffield	12%	Wolverhampton	6%	Southwark	6%	Epsom & Ewell	16%
4	Swindon	4%	Manchester	14%	Coventry	6%	Haringey	7%	Cambridge	16%
5	Waverley	4%	Ealing	15%	Sandwell	7%	Birmingham	8%	Lewisham	16%
6	North Somerset	4%	Haringey	15%	Luton	7%	Lambeth	8%	Elmbridge	17%
7	Woking	5%	Birmingham	16%	Bristol	7%	Barnet	8%	Greenwich	18%
8	Winchester	5%	Leicester	16%	Warwick	7%	Islington	9%	Harrow	18%
9	Basildon	5%	Wandsworth	16%	Tower Hamlets	7%	Leicester	9%	Barnet	18%
10	Rushcliffe	5%	Kingston upon Thames	16%	Leeds	7%	Hackney	9%	Liverpool	19%
315	Lewisham	35%	Lewes	53%	Vale of White Horse	47%	Nuneaton & Bedworth	42%	Bury	57%
316	Tower Hamlets	38%	Halton	54%	Exeter	48%	Broxbourne	44%	Tonbridge & Malling	58%
317	Newham	38%	South Kesteven	54%	Kingston upon Hull	48%	Maidstone	46%	Blackburn with Darwen	61%
318	Islington	39%	Rother	54%	Isle of Wight	49%	Plymouth	46%	Poole	62%
319	Slough	39%	Wirral	55%	County Durham	62%	Basildon	47%	South Tyneside	66%
320	Southwark	42%	Thanet	55%	Chichester	63%	Woking	49%	County Durham	66%
321	Hammersmith & Fulham	42%	Blackpool	57%	Northumberland	64%	Wakefield	50%	Oadby & Wigston	68%
322	Wandsworth	50%	Test Valley	58%	Rushmoor	74%	Southend-on-Sea	54%	Sandwell	68%
323	Hackney	58%	East Lindsey	58%	Shropshire	79%	Eastbourne	67%	Wolverhampton	71%
324	Lambeth	65%	Barnsley	59%	Cornwall & Isles of Scilly	105%	Cornwall & Isles of Scilly	220%	Norwich	71%

7.3.1 Leeds (Cities and Services)

The relative predictive performance for the White ethnic group in Leeds is reasonable for the younger age bands, but the performance deteriorates at the oldest age bands, particularly for those aged 80 and older. For males aged 85 and older the ETHPOP projection is over twice the 2011 Census count. Conversely, the Mixed ethnic group performs well at older ages but poorly at younger ages – although the difference when quantified as numbers is small. Of the five ethnic groups the best prediction performance is probably for the Asian group. The correspondence is good for all ages in this ethnic group that, after the White group, is the most numerous in Leeds for the 50 and older population. The Black ethnic group is the third largest in Leeds and, like the Mixed group, the ETHPOP performance is good at older ages but under predicts for the two younger age bands, especially for males. This pattern of under prediction of males is also evident within the Other group.

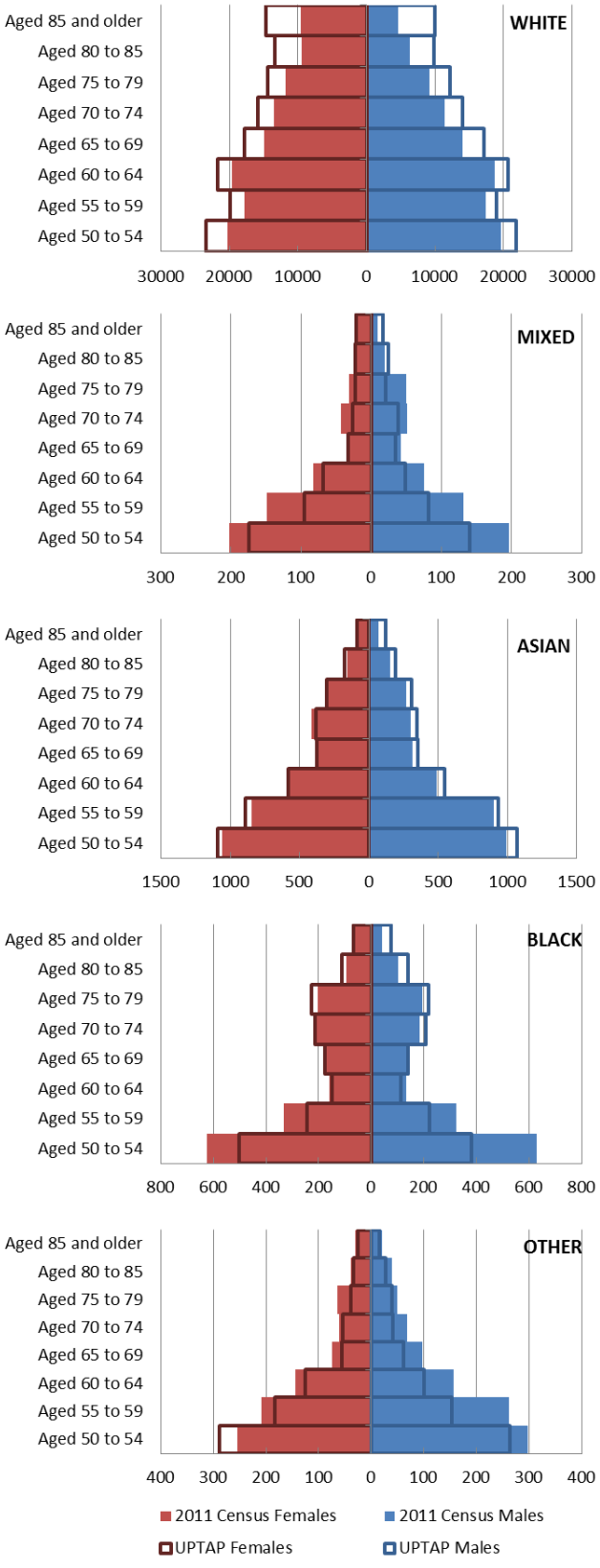


Figure 7-2 : Comparison of Leeds population structures, 2011, from top: White; Mixed; Asian, Black and Other ethnicities

7.3.2 Redbridge (London Suburb)

In Redbridge, the ETHPOP projections for the White ethnic group over predicts the population at most age bands, with the performance at predicting 65 to 69 years olds being poor in both absolute and relative terms. However, in contrast to Leeds, the performance for the very old is very good indeed. The performance for the Mixed group is reasonable except for the very youngest age band where ETHPOP under predicts by a significant percentage, although the numbers involved are again small. As for Leeds, the performance with the Asian group is remarkably good for Redbridge whose Asian group is beginning to be as large as the White group. The Black group is also another sizeable demographic and the predictive performance of ETHPOP is better for the middle and older age bands. Relatively speaking the performance at predicting the Other group is the poorest of the five, ETHPOP under predicts by around a half the number of people in this group.

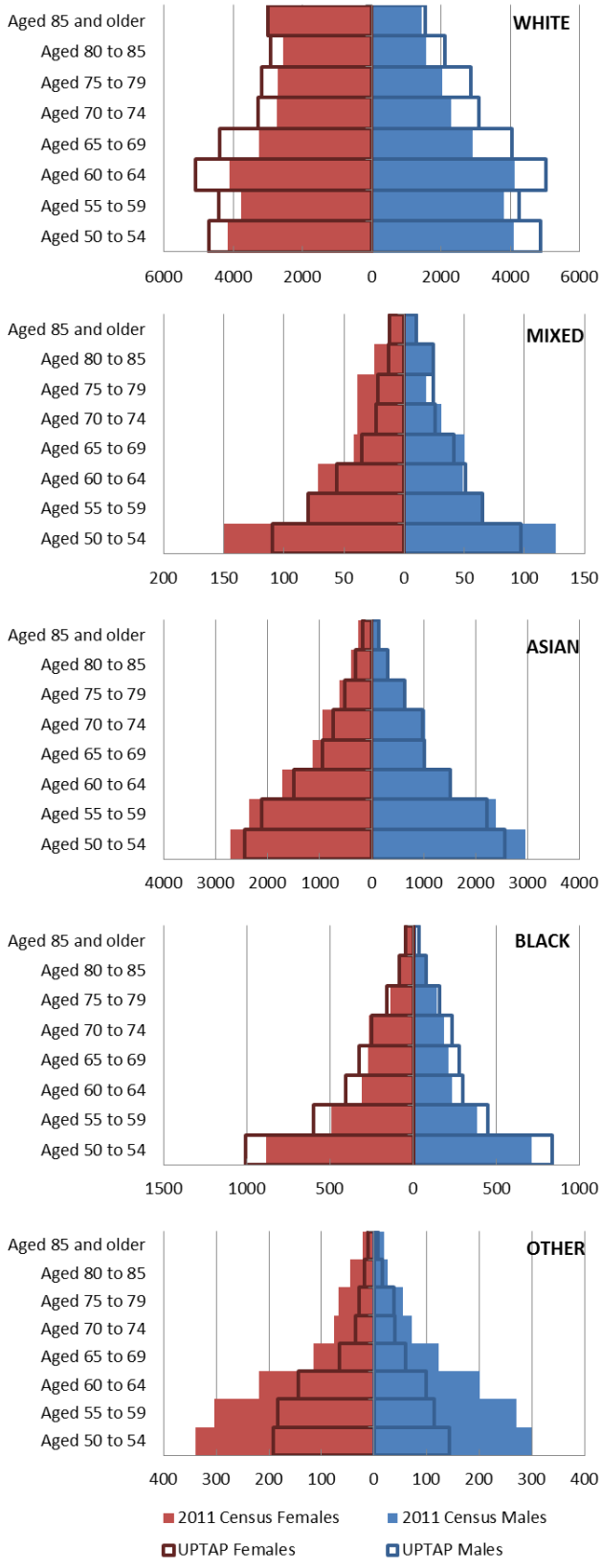


Figure 7-3 : Comparison of Redbridge population structures, 2011, from top: White, Mixed, Asian, Black and Other ethnicities

7.3.3 Camden (London Centre)

The White ethnic group is by far the largest in Camden, with none of the remaining ethnic groups being larger than 600 for any gender and age combination. Again for this group, the tendency is for the ETHPOP projections to exceed the Census counts, and this is perhaps most pronounced for older males. With the Mixed group the performance is poor for the youngest age band (an under prediction by ETHPOP) but does improve for older age bands. Looking at the 2011 Census counts for the Asian group, there appears to be a disproportionately large number of females in the younger age bands. This feature has been replicated by the ETHPOP predictions, but not to the same extent. For the Black group at younger ages, the performance of ETHPOP reverses that seen for the Mixed group, with an over prediction at younger age bands. With the remaining Other group, the predictive performance for males is good but less so for females.

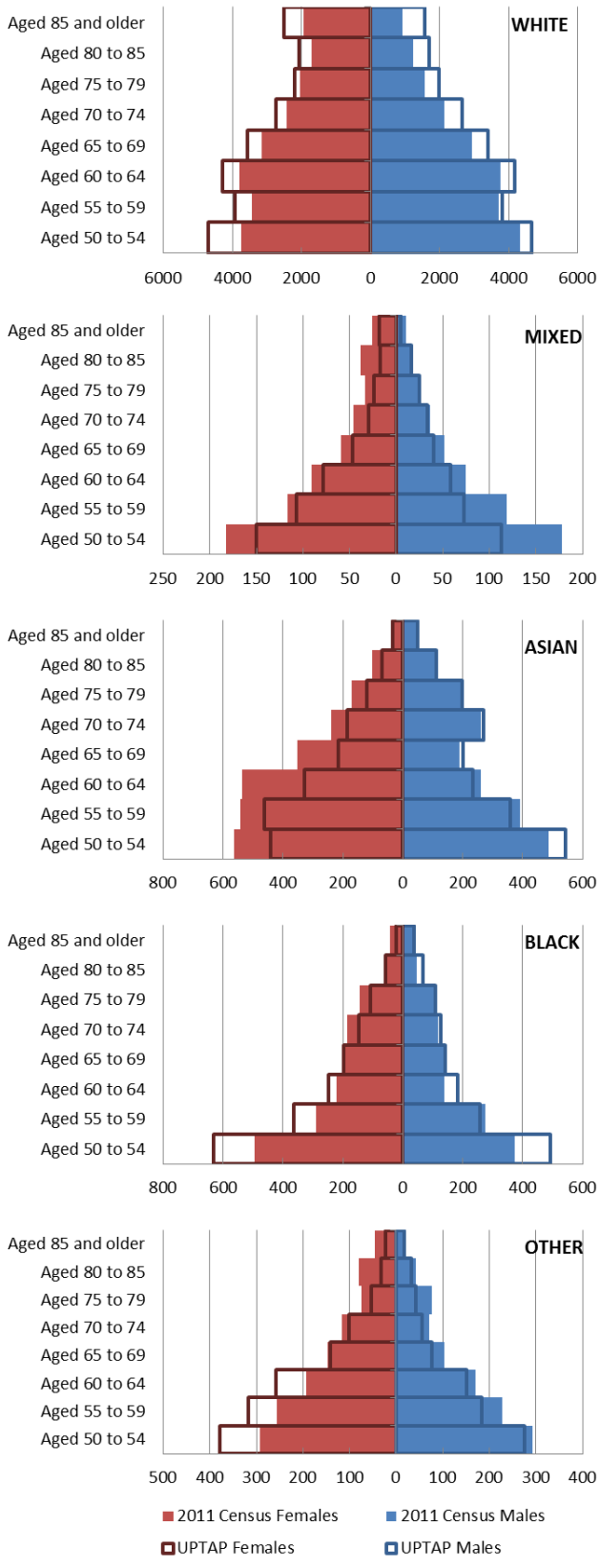


Figure 7-4 : Comparison of Camden population structures, 2011, from top: White, Mixed, Asian, Black and Other ethnicities

7.3.4 Haringey (London Cosmopolitan)

Here, the ETHPOP projections consistently over predict the number of individuals in the White ethnic group. Whilst these differences are similar in absolute number for each gender and age band, the proportionate impact is greater for the older age band. The prediction performance is much better for the Mixed group, particularly females. As with Camden, Haringey shows a disproportion number of females in the younger age Asian groups when compared to males. Other than this feature, the performance is reasonable. Looking at the structure diagram for the Black ethnicity, this appears to be the group that shows evidence of a transition in its importance for Haringey. The 50 to 54 year old age band is almost twice that of the next oldest band and this dramatic transition has been well predicted by ETHPOP. ETHPOP consistently under predicts the size of the Other ethnic group.

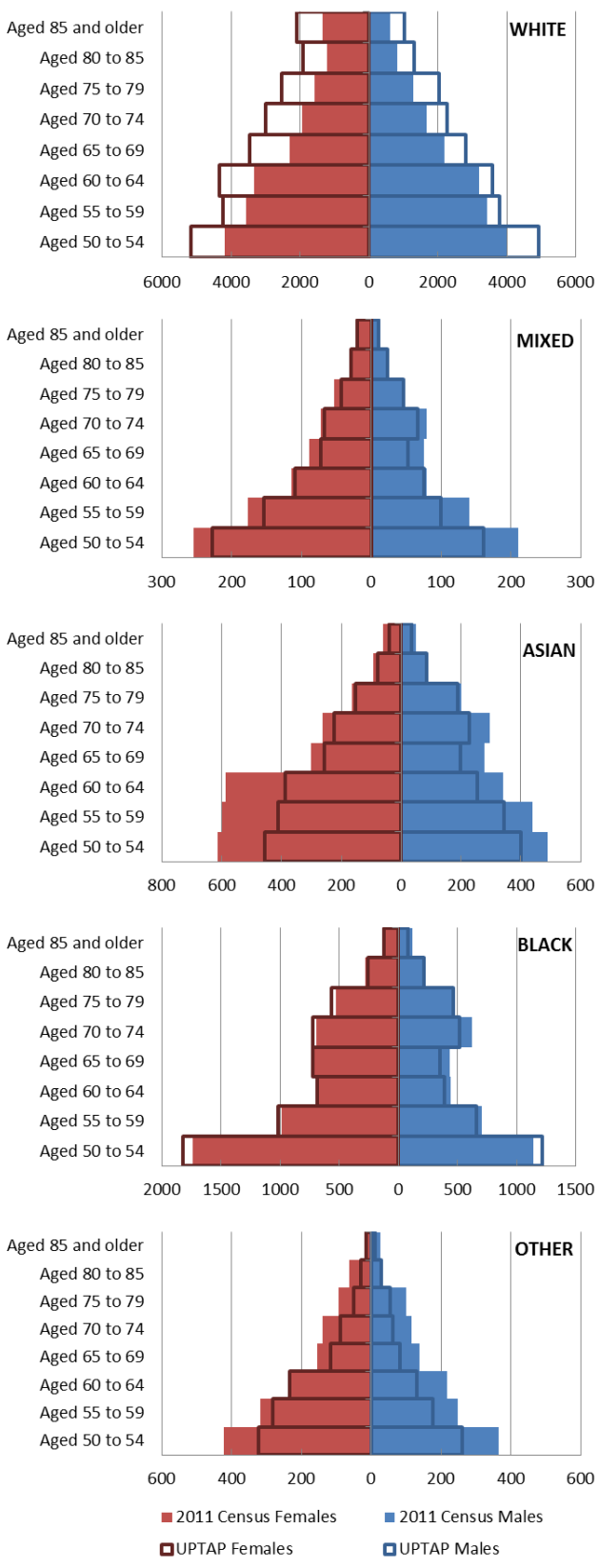


Figure 7-5 : Comparison of Haringey population structures, 2011, from top: White; Mixed; Asian, Black and Other ethnicities

7.3.5 Wiltshire (Prospering UK)

Wiltshire is a large authority (both terms of geography and population) with a predominately White ethnic population. The resultant scale of the structure diagrams tends to mask the size of some of the differences between the ETHPOP projections and the 2011 Census. For example the Female, aged 60 to 64, White ethnic group projections are 1,200 fewer than the 2011 Census counts. The remaining four ethnicities are much smaller in size and there is a pattern of under prediction for the Mixed and Asian ethnic groups and an over prediction for the Black and Other groups. The transition in the size of the Black ethnic group between ages 50 to 54 and 55 to 59 as seen in other case study authorities, is also present here in rural Wiltshire, although to a less pronounced degree due to the smaller population size.

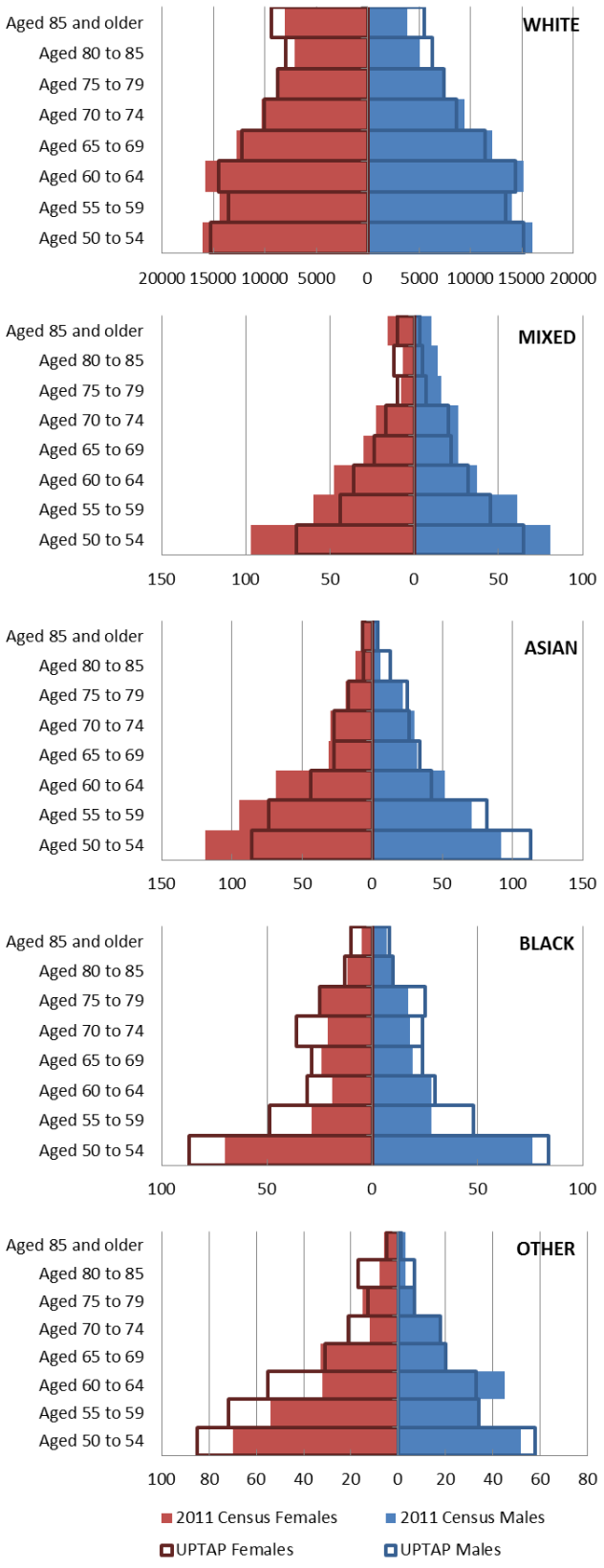


Figure 7-6 : Comparison of Wiltshire population structures, 2011, from top: White, Mixed, Asian, Black and Other ethnicities

7.3.6 Teignbridge (Coastal and Countryside)

The coastal and countryside area classification is probably the most heterogeneous of the ONS defined areas, ranging from the economically comfortable towns of Teignmouth in Devon to the more socially and economically challenged resort of Blackpool in the North West. Teignbridge is, like Wiltshire, a LAD with a predominantly White ethnic population, although it is smaller in size than Wiltshire. The ETHPOP projections are generally good for the White ethnic group, until the two older age bands, with the oldest age band, 85 and older, being greatly over predicted by ETHPOP. The populations of the remaining four ethnicities are very small; indeed they are fewer than 160, the cut off used in Table 7-1 and Table 7-3. This makes it difficult to identify any reliable features in these population structures.

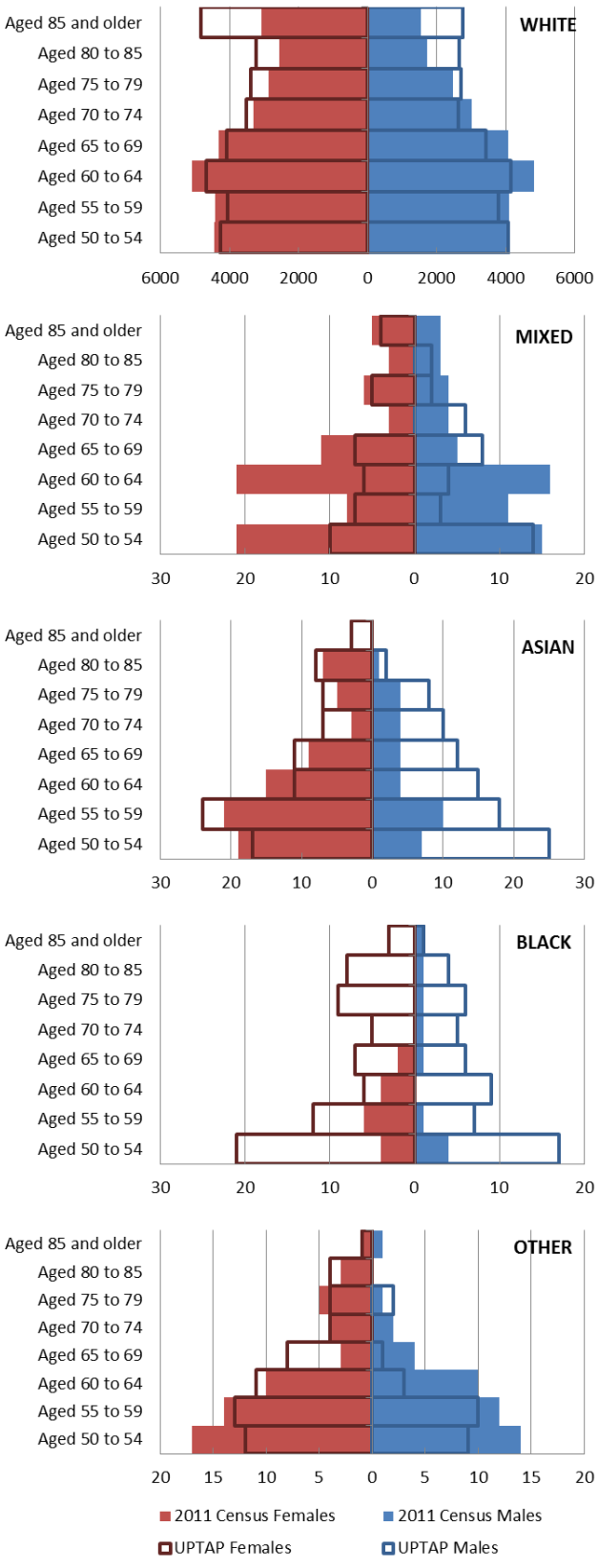


Figure 7-7 : Comparison of Teignbridge population structures, 2011, from top: White; Mixed; Asian, Black and Other ethnicities

7.3.7 Rotherham (Mining and Manufacturing)

Rotherham is once again a LAD where the White ethnic group dominates the 50 and over population. The predictive performance of ETHPOP is good for all White ethnic group ages and genders except for a sizeable under prediction of males aged 60 to 64 and females aged 85 and older. After the age of 54 the Mixed ethnic group has a small population size that is not well predicted by ETHPOP. The Asian ethnic group is more substantial in number and the pattern is for ETHPOP to under predict this population. As for the Mixed ethnic group, the Black ethnic group is also small after age 54, but does again show this transition pattern between ages 50 54 and 55 to 59 mentioned elsewhere. For the remaining Other group, the populations are once again small (but greater than 160 in aggregate) with Rotherham showing a disproportionate large number of males aged 50 to 59 relative to females. This feature has not been predicted by the ETHPOP projections.

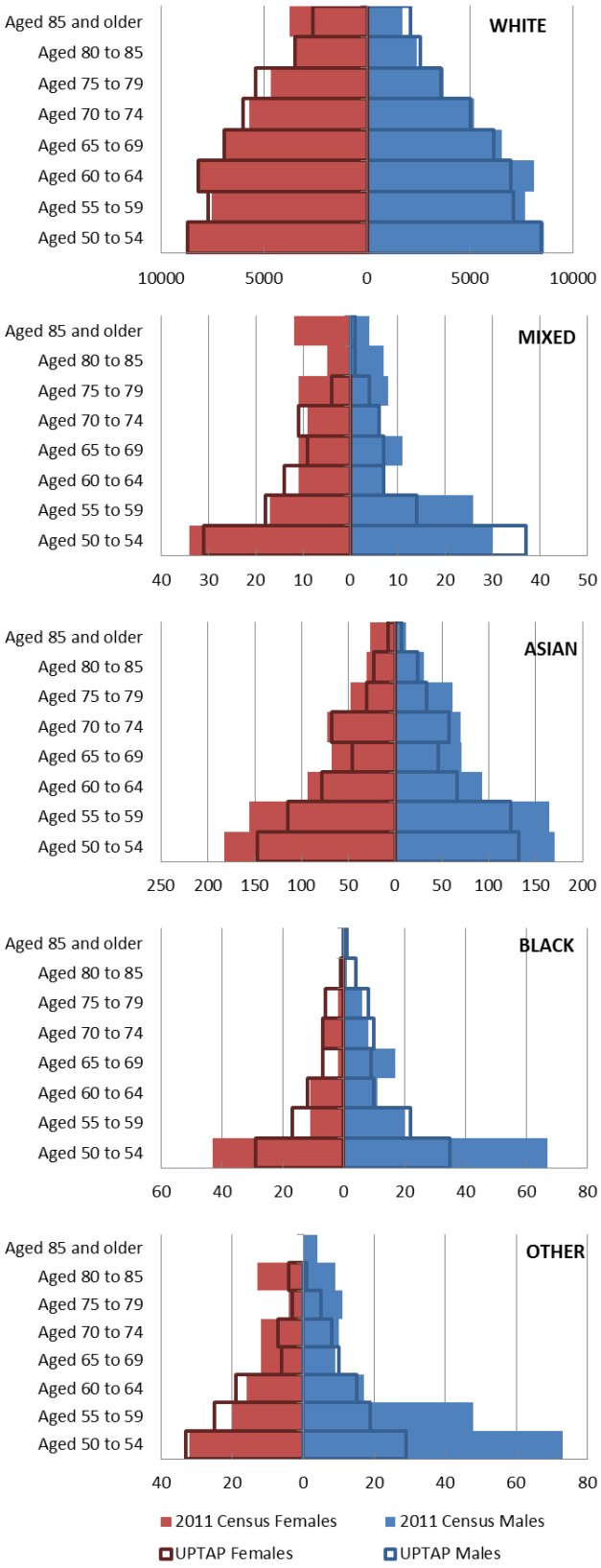


Figure 7-8 : Comparison of Rotherham population structures, 2011, from top: White, Mixed, Asian, Black and Other ethnicities

Looking at these representative LADs from within each of the ONS area classifications has revealed some common issues. Firstly the ETHPOP projections have tended to overestimate the size of the White ethnic group in each LAD, particularly for the oldest age bands, which is most probably influenced by the mortality assumptions used in the projections. The pattern for the remaining ethnic groups is less clear cut, but to some extent does reinforce the picture in Table 7-1 and Table 7-3, that the ETHPOP projections under predict these four ethnic groups. However whilst the percentages in these tables appear large, in many cases the number of people involved in these under predictions are small. Ethnic switching (Simpson et al., 2014) will also influence the relative projection performance between groups if individuals chose to define their ethnicity differently over time, particularly those in the Mixed group.

7.4 PROJECTION ADJUSTMENT FACTORS

In 5.4 the concept of an adjustment factor is defined to correct for the difference between the ETHPOP and 2011 Census population growths from 2001 to 2011 is introduced. This adjustment factor is itself a measure of the correspondence between the ETHPOP 2011 projections and the 2011 Census counts, the greater the departure from 1.00 for this adjustment factor, the poorer the correspondence.

However before these raw adjustments are used to adjust the trends in population change beyond 2011, the factors for each LAD are constrained to lie within a [10%,90%] area type specific confidence interval empirically derived for each ethnicity. In Figure 7-9 to Figure 7-13 a series of box plots type charts are provided that capture these distributions. In these charts, the centre line marks the median adjustment factor. The box marks the lower and upper quartiles, and the height of this box is the inter-quartile range (IQR). The lower whisker marks the 10% ile whilst the upper whisker marks the 90% ile. The asterisks below the 10% ile are the 10% of factors that are less than the 10% ile whilst those above are the 10% of factors that are greater than the 90% ile. As outlined in section 5.4.1, it is these outlying values that are truncated to the 10 and 90% ile. The difference between the highest and lowest factor is the range.

If there is no relationship between these adjustment factors and a grouping of these factors by some criteria (here the ONS area type classification) each of the box type plots for each grouping would be similar, both in terms of their median level and the spread – in essence the grouping would just be a random sample of factors (the Other ethnic group shown in Figure 7-13 perhaps comes closest to this idea of there being no discriminatory information in the area type). Differences in the median level and/or spread indicate either a clustering of factors around a particular level and/or an inherent degree of variation in their estimation.

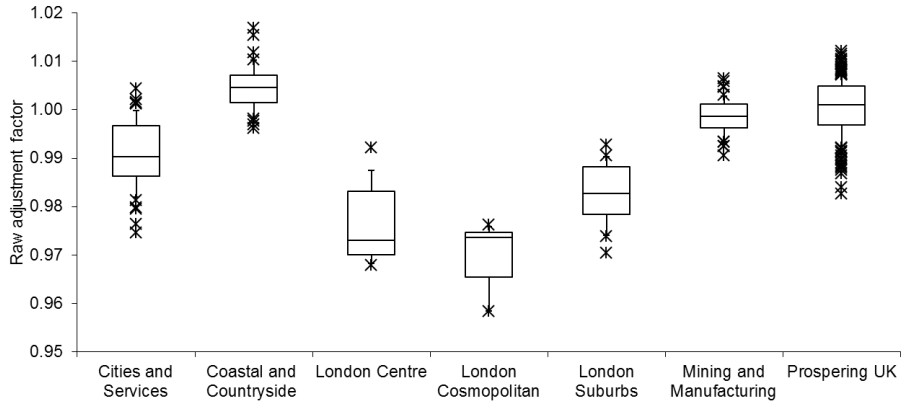


Figure 7-9 : Box type plots of the distribution of the adjustment factors in each area type for the White ethnic group

The White ethnic group shows a wide variation in the median level of these adjustment factors. If this factor is greater than 1.0 then the ETHPOP projections are under estimates of the 2011 Census White population (the Census counts are larger than the ETHPOP projections in equation 5-6). The median adjustment factor is greater than 1.0 only for the Prospering UK (1.0047) and Coastal and Countryside (1.0011) area types – ETHPOP has under estimated these populations in 2011. The IQRs also vary by area type, being the smallest for Mining and Manufacturing (0.0049) and Coastal and Countryside (0.0056). The IQR is greatest for London Centre (0.0132). Compared to some of the remaining ethnic groups, the range of factors by area type for the White group are fairly consistent.

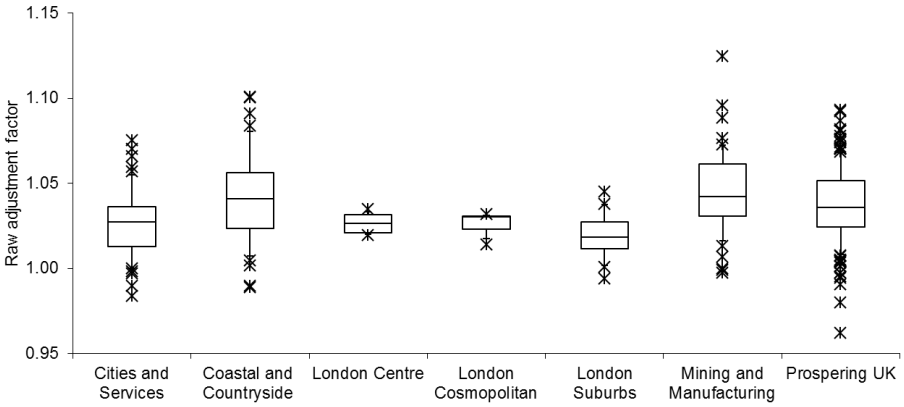


Figure 7-10 : Box type plots of the distribution of the adjustment factors in each area type for the Mixed ethnic group

Like the remaining four ethnicities the contrast in median levels for the Mixed ethnic group is not so pronounced as for the White group. However, the spread in some types of area are quite different, with the London centric areas having a much narrower range of factors than the remaining areas.

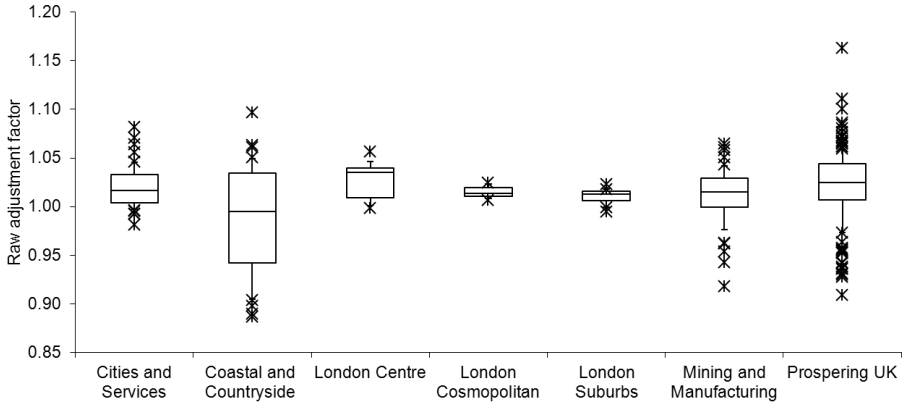


Figure 7-11 : Box type plots of the distribution of the adjustment factors in each area type for the Asian ethnic group

For the Asian ethnic group, the comments made for the Mixed group apply equally here except for the Coastal and Countryside area type where the IQR is much wider than that for any remaining area type. This suggests that the ETHPOP projection performance for LADs of this type in 2011 was particularly inconsistent.

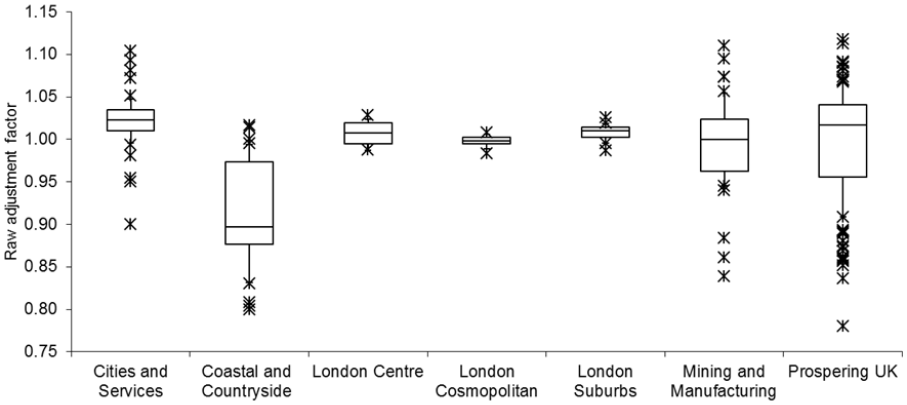


Figure 7-12 : Box type plots of the distribution of the adjustment factors in each area type for the Black ethnic group

Once again, the Coastal and Countryside area type is distinct from the non-London centric area types, however, on this occasion the median level of the adjustment factors appears to be much lower here than the rest, signifying that ETHPOP has almost consistently over predicted this ethnic group for this area type.

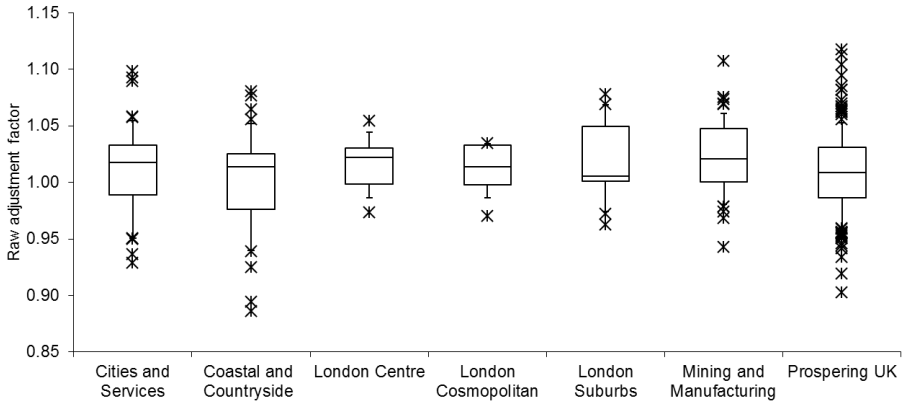


Figure 7-13 : Box type plots of the distribution of the adjustment factors in each area type for the Other ethnic group

Perhaps of all five ethnicities this Other group is the one where the distinction by area type is least informative. The median levels are very similar and the contrasts in the spread are not so stark as for the Mixed, Asian and Black ethnic groups.

A more formal assessment of the agreement between the distributions of the adjustment factors within each area classification can be made by use of the Kruskal Wallis test for the equivalence of two distributions (Kruskal and Wallis, 1952). This test is used in preference to a one way ANOVA since there is evidence in the box type plots to indicate that some of these distributions are not normal. The null hypothesis is that the distribution of these adjustment factors is the same within each area type whilst the alternative hypothesis is that is at least one area type whose distribution is different to at least one other. The results of this test are shown Table 7-5.

The results here corroborate the visual assessment of Figure 7-9 to Figure 7-13, with only the Other ethnic group having similar distributions within each area type.

Table 7-5 : Results of the Kruskal Wallis test for the difference in the distribution of adjustment factors by area type

Ethnic group	Test statistic (K)	Degrees of freedom	Significance level
White	138.231	6	0.000
Mixed	29.704	6	0.000
Asian	20.292	6	0.020
Black	55.115	6	0.000
Other	8.647	6	0.194

Both the box type plots and the results in Table 7-5 do justify the constraining of the raw adjustment factors based on the characteristics that are captured in the area type.

Other groupings may exist that produce similar behaviour; however the advantage of using these area types is their established provenance from ONS and longevity of use.

7.5 IMPACT OF ETHPOP REVISIONS

Section 5.4.2 outlines the methodology that uses these adjustment factors, along with 2011 Census counts, ETHPOP cohort projections and ONS sub national population projections to revise the ETHPOP projections from 2013 to 2031.

The extent and impact of these revisions to the original ETHPOP projections for the final forecast year of 2031 can be calculated as a percentage change between the original 2031 ETHPOP 50 and over population and the revised projection (see equation 7-3).

$$TR\%_e^{LAD} = \frac{\sum_{g=1}^{2}\sum_{a=1}^{8}(P_{g,a,e}^{LAD,2031} - E_{g,a,e}^{LAD,2031})}{\sum_{g=1}^{2}\sum_{a=1}^{8}(E_{g,a,e}^{LAD,2031})} \quad 7\text{-}3$$

Where: $TR\%_e^{LAD}$ is the total revision percentage between the original ETHPOP and revised ETHPOP population counts for ethnicity e in local authority LAD;

$P_{g,a,e}^{LAD,2031}$ is the revised ETHPOP sub-population of gender g, aged a, of ethnicity e, in LAD for 2031;

$E_{g,a,e}^{LAD,2031}$ is the original ETHPOP sub-population of gender g, aged a, of ethnicity e, in LAD for 2031.

This equation is similar to equation 7-1 and generally it is reasonable to assume that a positive TR% indicates an upward revision to the original ETHPOP projections whilst a negative value indicates a downward revision. For the White ethnic group, the distribution of this TR% value within each area type is shown as traditional box plots in Figure 7-14 (the whiskers now represent the additional range of 1.5 times the IQR and there are no outliers present).

Unsurprisingly the trends in this figure match those in Figure 7-9 since the adjustment factors shown there are used in the revision process. The greatest reductions in the size of the White ethnic group are those area types associated with London. Reductions are also seen in the Cities and Services area type, and to a lesser degree, the Mining and Manufacturing area type. Upward revisions are commonly seen for both the Prospering UK area type and, especially, the Coastal and Countryside area type.

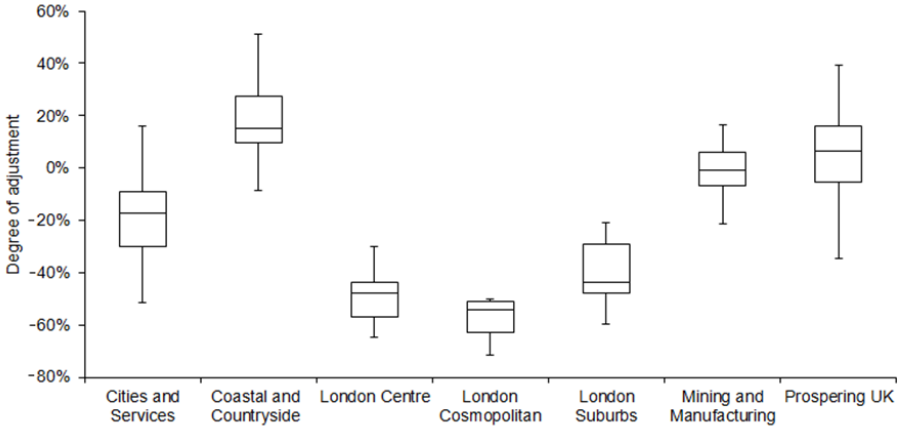


Figure 7-14 : Distribution of the impact of the adjustment process on the White ethnic group

The ten LADs with the greatest upward and downward revisions measured by TR% are identified for each ethnicity in Table 7-6. The changes seen for the White ethnic group are consistent with those highlighted in Figure 7-14 in that the London centre type of authority show the greatest downward revision in the size of this population whilst the Coastal and Countryside and Prospering UK type of authority show upward revisions. Revisions to the Mixed group tend to be consistently revised upward, with only three LADs showing a downward revision.

Rushmoor shows a noteworthy ten fold revision in the size of its south Asian 50 and over population. The ETHPOP projections for 2001 and 2011 for the over 40 population in Rushmoor are for an increase from 494 to 718 (recall that the adjustment factors are calculated to adjust the size of the 40 and older population, not just the 50 and older). The equivalent Census counts are an increase from 468 to 3,081. This translates to a yearly adjustment factor on +16.3% which is constrained to a value of 5.9% (Rushmoor is the outlier for the Prospering UK area type in Figure 7-11). Coupled with this is the re-basing of the projections from 2011 to a population of 3,081 rather than 718 which also causes subsequent upward revisions to the projection. Clearly the observed growth in the Asian population between 2001 and 2011 in Rushmoor is not sustainable and the constraining stage has helped to moderate this substantial growth. Also, in a number of LADs the already small Black and Other ethnic populations have been eliminated as a consequence of these revisions.

7.6 IMPACT OF REVISION ON CASE STUDY AUTHORITIES

To further illustrate the impact of these revisions, the population structure diagrams showing the original ETHPOP and revised ETHPOP populations for the seven representative LADs are given below. These diagrams show the size of the population for just two projection years, 2021 and 2031.

Table 7-6 : Top ten LADs by upward and downward total revision percentage

	White		Mixed		Asian		Black		Other	
1	Rother	51.0%	Thanet	994.7%	Rushmoor	1058.9%	Wakefield	690.1%	Sandwell	660.9%
2	South Hams	49.8%	Wirral	885.1%	Tandridge	639.1%	Basildon	678.3%	Bury	594.5%
3	Wealden	39.4%	Knowsley	694.8%	Ashford	617.1%	Dartford	647.0%	Dudley	574.1%
4	Ryedale	38.6%	Arun	572.3%	Tunbridge Wells	606.5%	Broxbourne	641.0%	Wolverhampton	572.5%
5	Mid Suffolk	37.0%	East Riding of Yorkshire	571.8%	Colchester	576.7%	Thurrock	612.1%	Slough	551.2%
6	South Somerset	36.3%	Winchester	528.9%	South Cambridgeshire	575.0%	South Gloucestershire	609.9%	Enfield	538.9%
7	East Dorset	32.3%	Barnsley	528.8%	Harborough	566.5%	Crawley	601.7%	Nuneaton and Bedworth	528.0%
8	Broadland	31.9%	Wigan	521.8%	Chichester	550.0%	West Berkshire	591.2%	Gravesham	511.5%
9	New Forest	31.5%	Three Rivers	485.8%	Chelmsford	521.4%	Stevenage	555.1%	Central Bedfordshire	498.2%
10	Adur	31.0%	Waverley	483.9%	Vale of White Horse	506.5%	Central Bedfordshire	530.1%	North Hertfordshire	495.9%
315	Lewisham	-52.1%	Hillingdon	7.0%	East Devon	-86.2%	South Hams	-98.2%	North Dorset	-78.7%
316	Islington	-52.6%	Ealing	5.3%	Cornwall and Isles of Scilly	-86.5%	Scarborough	-98.2%	Norwich	-80.0%
317	Redbridge	-52.9%	Exeter	5.2%	Weymouth and Portland	-88.8%	Carlisle	-98.4%	Lancaster	-80.6%
318	Southwark	-54.3%	Hounslow	2.7%	West Somerset	-90.3%	East Devon	-99.0%	Preston	-81.5%
319	Brent	-58.5%	Forest Heath	1.5%	South Hams	-92.2%	Craven	-100.0%	East Lindsey	-81.5%
320	Slough	-59.7%	Woking	1.1%	West Devon	-95.1%	Eden	-100.0%	Great Yarmouth	-82.6%
321	Wandsworth	-61.1%	North Dorset	1.1%	North Norfolk	-95.3%	Purbeck	-100.0%	North Norfolk	-83.2%
322	Hammersmith and Fulham	-64.7%	Barking and Dagenham	-0.3%	Ryedale	-95.9%	Ribble Valley	-100.0%	Forest Heath	-84.2%
323	Lambeth	-67.6%	Boston	-10.6%	Purbeck	-96.6%	Torridge	-100.0%	Boston	-85.0%
324	Hackney	-71.6%	Leicester	-11.7%	Torridge	-97.0%	West Devon	-100.0%	Purbeck	-91.6%
Population fewer than 240		0		86		0		22		67

7.6.1 Leeds (Cities and Services)

The impact of the revision process on the White ethic group in Leeds is a reduction in the size of this group for all gender and age combinations. The population share of this group was 91.4% in 2031 using the original ETHPOP projections, but after the revision this falls to 82.0%. The revision prcess has nearly tripled the size of the Mixed ethnic group but even after the revision this is still a modest 2.9% of the population. The impact on the Asian ethnic group appears to be modest too, with a slight increase in this population at the younger ages. The greatest impact of the revision in Leeds is on the Black group. The larger cohort of 50 to 54 year olds seen in Figure 7-2 is seen to age to 60 to 64 in 2021 and to 70 to 74 in 2031, with a following cohort of a substantive Black population. After the revision process the Black ethnic groups share in 2031 increases from 1.5% using the original 2031 ETHPOP projections to 6.2% using these revised projections, more than quadrupling. The size of the Other ethnic group population continues to grow steadily and the revision process makes this group a little larger.

7.6.2 Redbridge (London Suburb)

The most noteworthy impact of the revision process in Redbridge is the reduction in the share of the White ethnic group (the original share was 52.0% and the revised share is 27.6%) and the increase in the Asian group (share from 34.1% to 49.5%) and the Other group (share increases from 2.3% to 13.0%). Examing the dynamics of the White group first, the size of this group in Redbridge fell from 77,999 in 2001 to 65,454 in 2011. This dictates a yearly adjustment factor of -2% (which compunds over 20 year to a 33% reduction). Also since the revised ETHPOP projections are re-based to 2011 Census counts, the initial base for the White group is the much lower value of 65,454 rather than the original ETHPOP projection of 77,949. For the Asian group the year adjustment is +1.4% and the 2011 base is revised up to the 2011 Census count of 34,843 (rather than 30,680). Finally, the yearly adjustment for the Other ethnic group is +6.1% and the base is now 3,702 (not 2,073) (the reason for the large yearly adjustment for the Other group is clear from looking at the bottom diagram in Figure 7-3). The picture is Redbridge is driven primarily by the reduction in the size of the White group (a result of a large reduction in its starting point and a moderate yearly downward adjustment) and significant increases in the Asian and Other populations, this latter dynamic driven by large values of the yearly adjustment. Otherwise, the progression of a smaller 50 to 54 year old Black cohort in 2011 is seen to take place through the age bands.

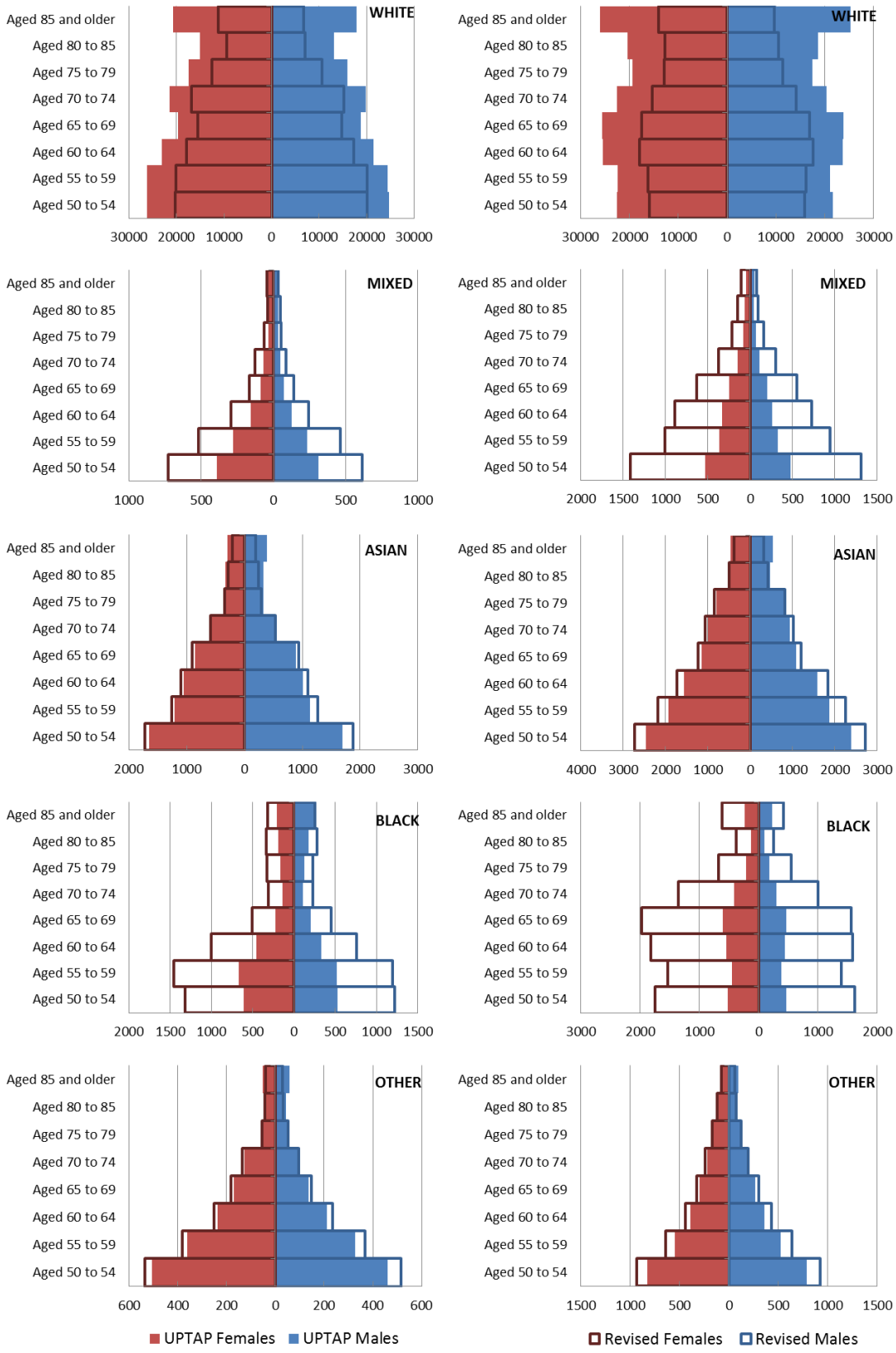


Figure 7-15 : Leeds' 2021 (left) and 2031 (right) original and revised ETHPOP population structure, for White (top), Mixed, Asian, Black and Other (bottom)

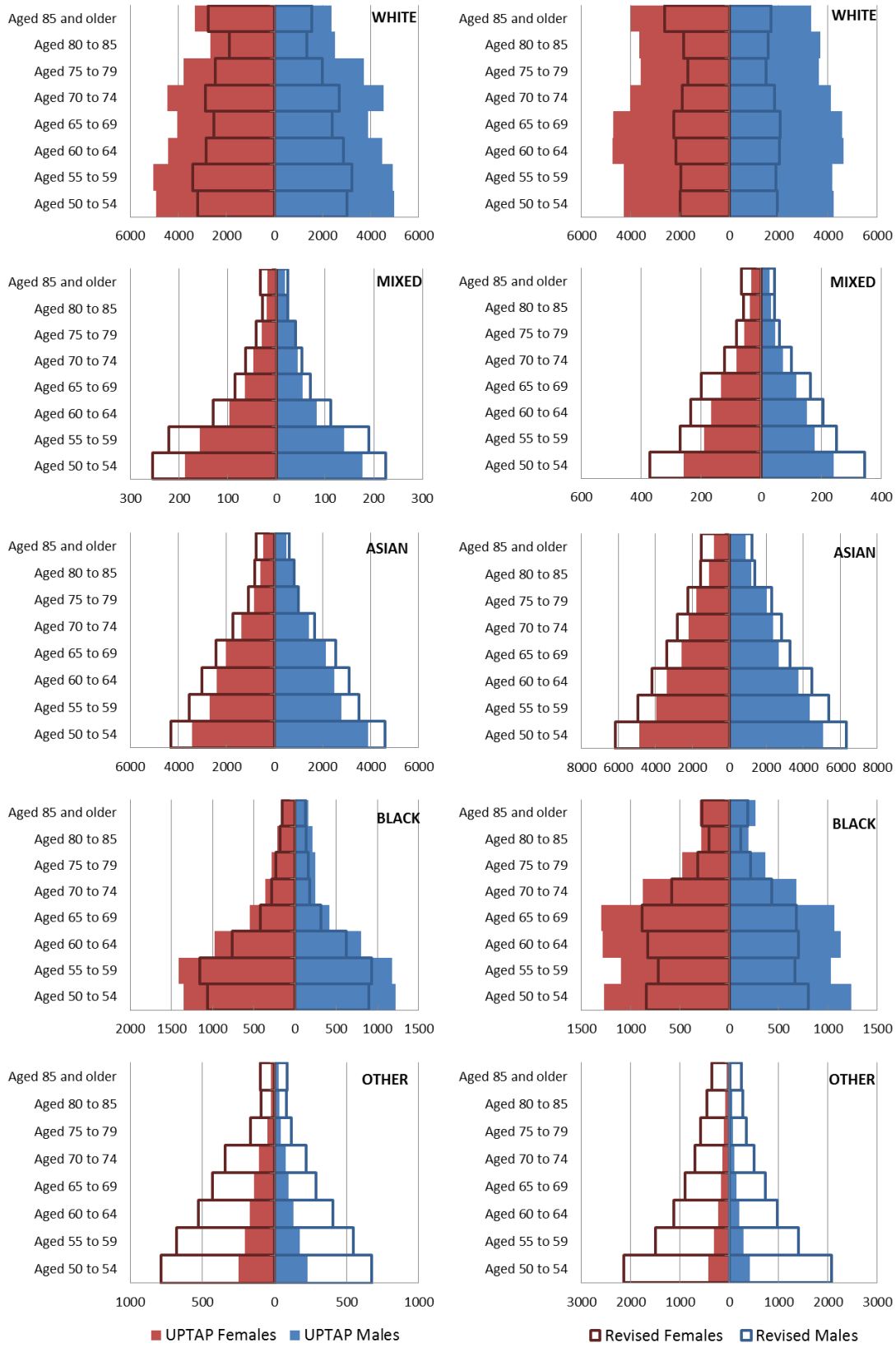


Figure 7-16 : Redbridge's 2021 (left) and 2031 (right) original and revised ETHPOP population structure, for White (top), Mixed, Asian, Black and Other (bottom)

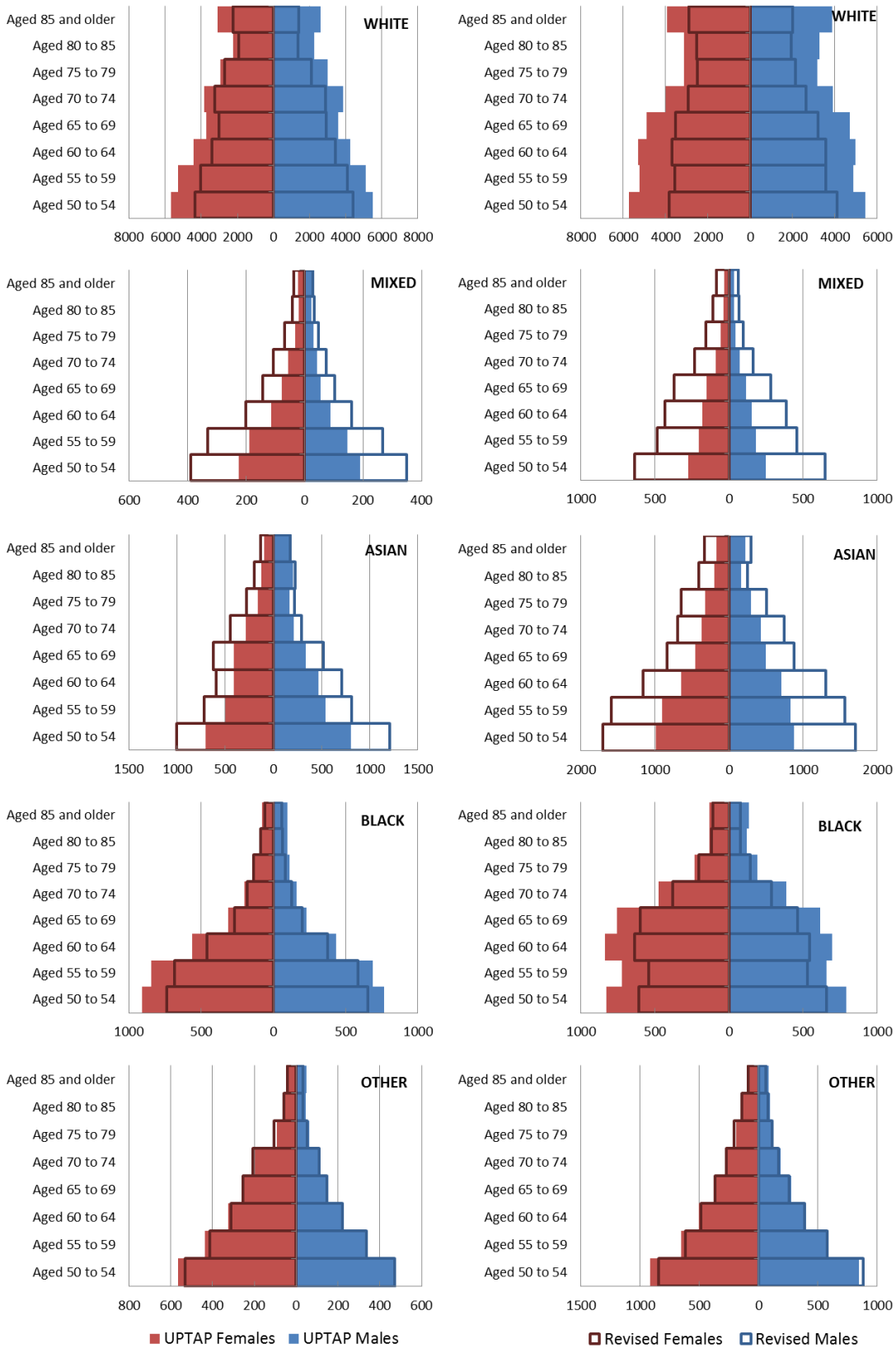


Figure 7-17 : Camden's 2021 (left) and 2031 (right) original and revised ETHPOP population structure, for White (top), Mixed, Asian, Black and Other (bottom)

7.6.3 Camden (London Centre)

For Camden, there is once again a reduction in the size and share of the White ethnic group (the share, using original ETHPOP projections, of 74.8% in 2031 falls to 61.5% after the revisions). Here, increases are apparent for the Mixed and Asian groups, where the share for these two groups grows from the original ETHPOP 2031 projections of a 2.1% and 8.7% share to revised projections of 5.8% and 18.3% respectively. The yearly correction for the Mixed group is +2.7% (which over 20 years compounds to +69.0%). The yearly adjustment factor for the Asian ethnic group is a more modest 1.6% but still compounds to a 37% growth over 20 years. The prediction for the Other ethnic group is good requiring little overall revision, whilst for the Black ethnic group the revised population is smaller for the majority of age bands than the original ETHPOP projection.

7.6.4 Haringey (London Cosmopolitan)

The revision process has once again had a dramatic impact on the size of the White ethnic group in Haringey. As for Redbridge, this is a combination of two factors. The 2011 original ETHPOP population for this group was 76,425 whilst the 2011 Census count (and hence the new 2011 base) is just 59,865. The yearly adjustment factor is also substantial at -2.4%. In 2031 the share of the population which is of a White ethnicity falls from 71.5% to 44.5% as a result of the revision process (by way of context, the 2011 Census has 65.9% of Haringey's aged 50 and over population being White). The consequence of the reduction in the size of the White group and increases in the size of the remaining ethnic groups is an increase in the share of the population of Haringey for these remaining ethnic groups (less so with the Black group). After the White group, the Black group is predicted to be the largest in 2031 at 18.6%, followed by Other, 16.9%, Asian, 13.8% and Mixed at 6.2%.

7.6.5 Wiltshire (Prospering UK)

The revisions to the original ETHPOP projections for the White ethnic group in Wiltshire has increased the size of this group at nearly all ages, although the White group share of the population drops a little from 96.7% to 96.2%. The Mixed and Asian groups have larger populations than the original ETHPOP projections, whilst the Black ethnic group is smaller. The impact of the revisions for the Other group have been minor.

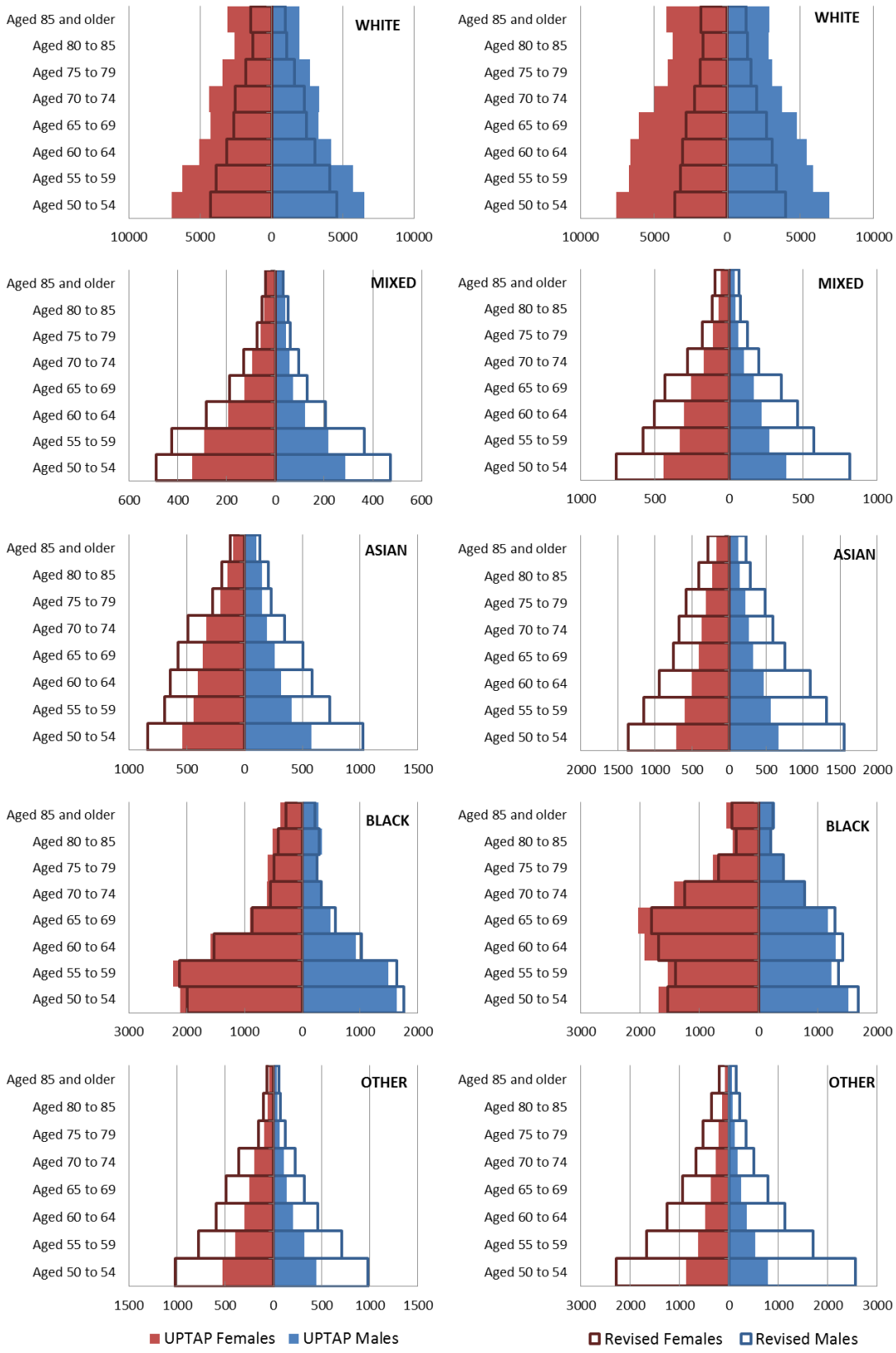


Figure 7-18 : Haringey's 2021 (left) and 2031 (right) original and revised ETHPOP population structure, for White (top), Mixed, Asian, Black and Other (bottom)

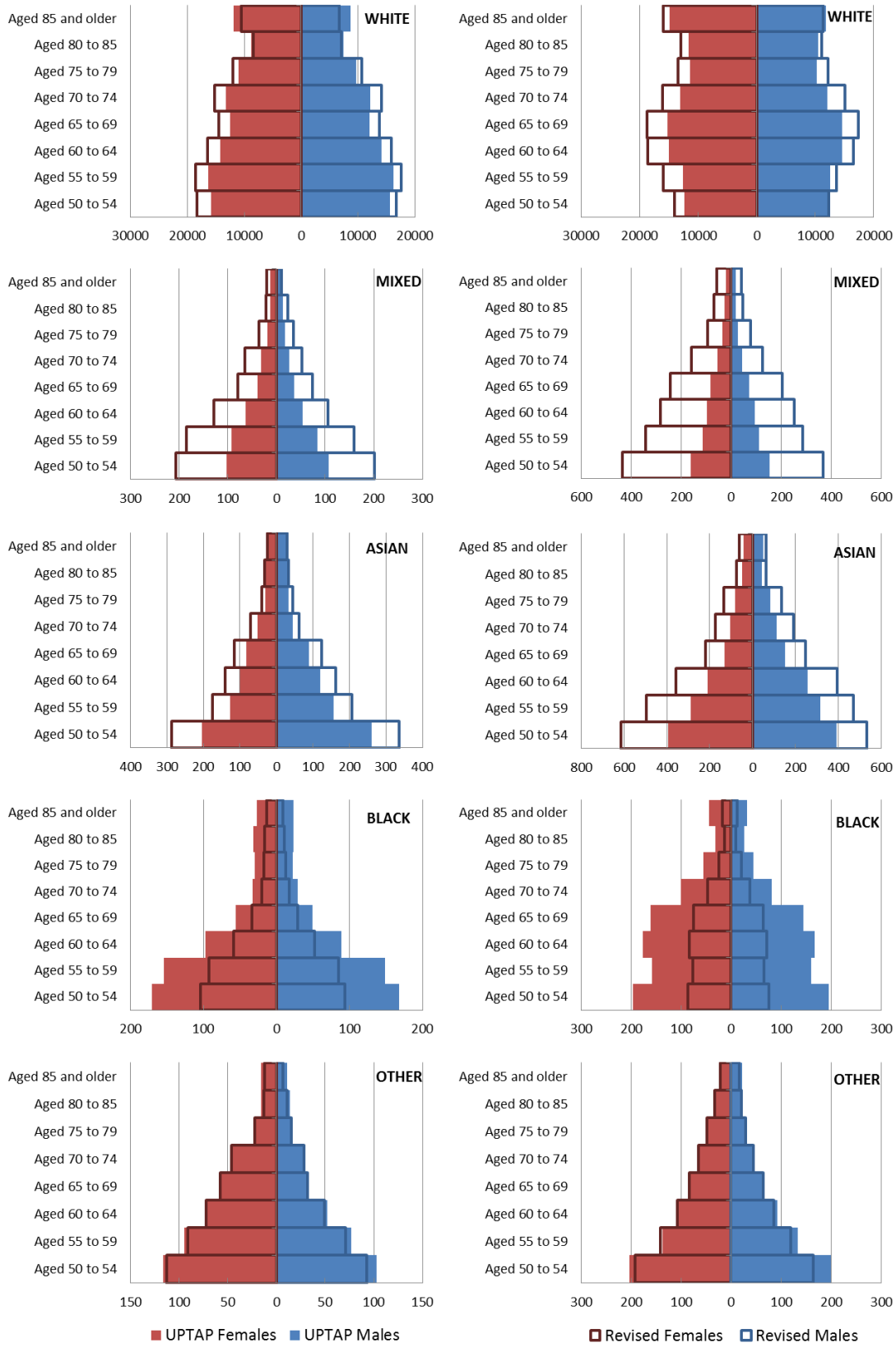


Figure 7-19 : Wiltshire's 2021 (left) and 2031 (right) original and revised ETHPOP population structure, for White (top), Mixed, Asian, Black and Other (bottom)

7.6.6 Teignbridge (Coastal and Countryside)

The greatest impact on the White ethnic group in Teignbridge is a reduction in those aged 85 and older. This is a result of the final step of the revision process where the size of a particular gender and age combination is constrained to the ONS 2012 SNPP. Here the 2012 SNPP for those aged 85 and older is around half the size of the original ETHPOP projections. The impact of the overall revision process on the Mixed and Other ethnic group has been to increase them in size, whilst the remaining two groups, Asian and Black, have decreased. This is most dramatically seen for the Black ethnic group which has been virtually eliminated from the population of Teignbridge as a result of the revision process. Between the 2001 and 2011 Censuses the size of the Black ethnic group in Teignbridge decreased from 75 to 58, rather than the projected ETHPOP change from 83 to 251 in 2011. These values result is a large adjustment factor of -12.8% per annum. These large reductions are not uncommon with Coastal and Countryside type authorities for the Black group, as is evident in Figure 7-12.

7.6.7 Rotherham (Mining and Manufacturing)

The revision process has created larger populations for most ethnic groups in Rotherham. All the adjustment factors are greater than zero, ranging from +0.27% per annum for the White ethnic group to +7.0% for the Other group. Even with this (admittedly low) positive adjustment factor for the White group, the share of this population is revised down from 96.0% in the original ETHPOP population projection to 90.4% revised in the revised projections.

7.6.8 Summary

As was inferred by the comparisons between the 2011 ETHPOP projections and the 2011 Census counts in section 7.2, the revision process has made some substantial changes to the original ETHPOP projections. For many LADs, the size of the White ethnic group has been reduced. Sometimes this reduction is substantial, especially so for the London based LADs (see also Figure 7-14 and Table 7-6). This reduction has come about by a combination of a re-basing of the starting points of the projections to the 2011 Census counts, which can be substantially lower than the 2011 ETHPOP projections. Also since this 2011 ETHPOP White group projection is usually less than the 2011 Census count (see Figure 7-9), the adjustment factor has been large and negative, with small annual percentages compounding to large reductions over a 20 year time span.

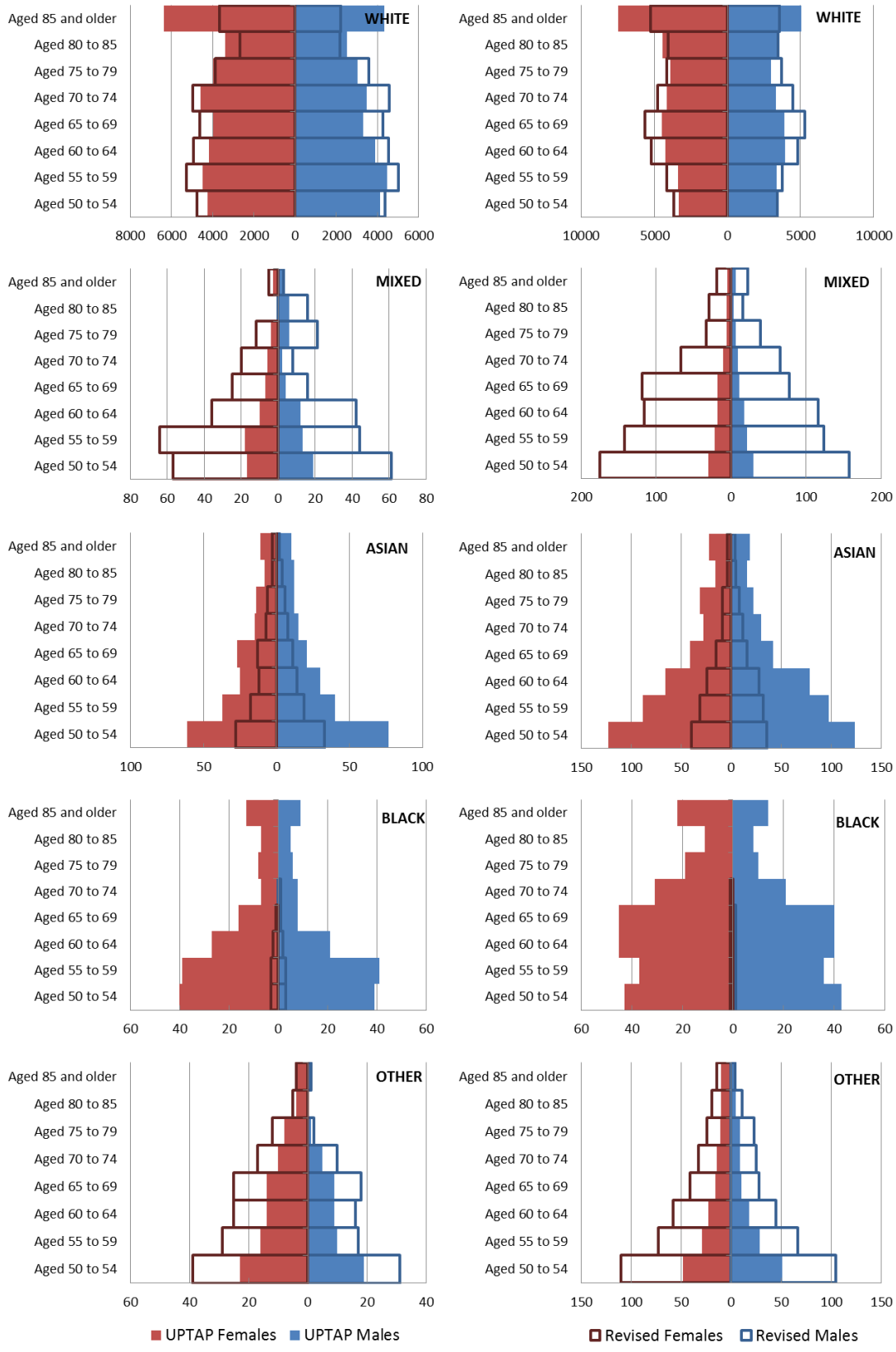


Figure 7-20 : Teignbridge's 2021 (left) and 2031 (right) original and revised ETHPOP population structure, for White (top), Mixed, Asian, Black and Other (bottom)

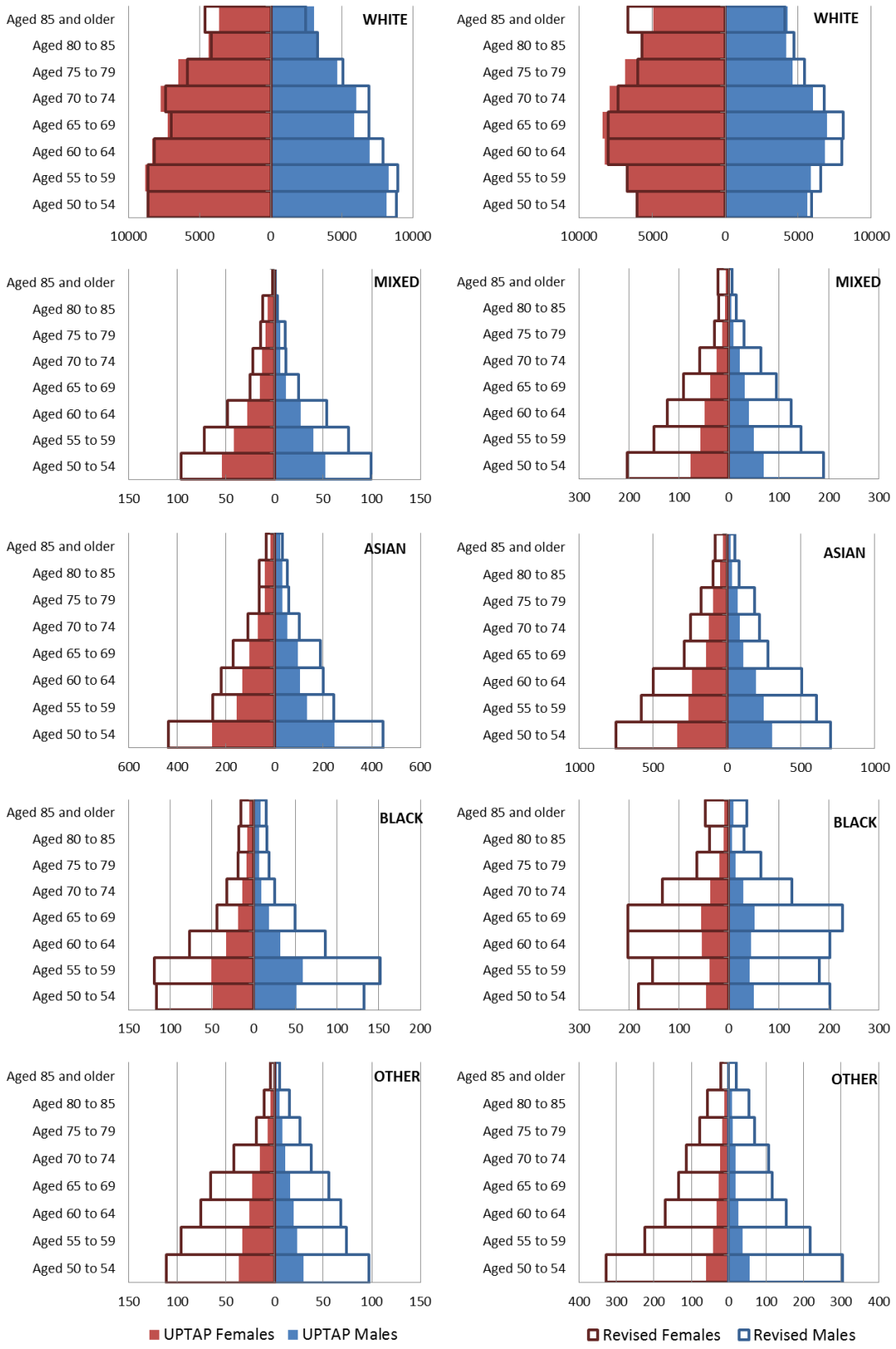


Figure 7-21 : Rotherham's 2021 (left) and 2031 (right) original and revised ETHPOP population structure, for White (top), Mixed, Asian, Black and Other (bottom)

The Mixed ethnic group is consistently under predicted in all the seven case study LADs by the original ETHPOP projections. For the Asian group the picture is less clear. The revised Asian populations are larger for all types of authority, except for the Coastal and Countryside authority of Teignbridge where the population is smaller. With the Black ethnic group, there are variable outcomes. Only in the Cites and Services and Mining and Manufacturing types of authority has the revision increased the size of this group. At the extreme, the Black population of the Coastal and Countryside LAD, Teignbridge, has been virtually eliminated. The Other group is either predicted well or under predicted by the original ETHPOP projections.

7.7 CONCLUSION

In this chapter an assessment has been made of the ETHPOP population projections, which provides an important input into the health microsimulation model, helping to ensure that the future population structure reflects the gender, age and ethnic makeup of each LAD.

The availability of 2011 Census data has allowed a comparison to be made between the ETHPOP projections for 2011 and the appropriate Census counts. The correspondence between these two data sources has been measured in terms of its direction (bias) and size (extent) for each of the five ethnic groups used in this study. The general trend has been for an over prediction of the White ethnic group and an under prediction of the remaining four groups. In percentage terms these differences are least for the White group, although this is undoubtedly due to the relatively large size of this group.

The comparison between the original ETHPOP and 2011 Census data, coupled with similar data from 2001, has enabled the calculation of LAD specific adjustment factors to correct for any discrepancies in the projections over a 10 year time span. How these adjustment factors vary by the ONS area type classification shows that the factors vary in their level and variability by this area type. This, to an extent, justifies the use of [10%,90%] confidence intervals to constrain some of the more extreme adjustment factors. Even with these constraints in place however, some seemingly modest yearly adjustment factors can, when compounded over 20 years from 2011 to 2031, have a dramatic impact on the size of each ethnic group's population in a LAD.

The revision of the original ETHPOP projections involves not just these adjustment factors but a rebasing of projections to 2011 Census counts in 2011 and a final constraint to the ONS 2012 SNPP. Both these actions impact on the extent of the revisions. Overall the results of the revision are the most dramatic for the London centric types of LAD, with big downward revisions in the size of the White ethnic group and commensurate increases in some, but not all, of the remaining ethnic groups.

Perhaps another challenging result is the virtual elimination of already small ethnic populations in some more prosperous LADs by 2031.

Not knowing what the future population in each LAD is makes it difficult to assess the plausibility of the revision process. The process does however, in an intelligent way, use a wide variety of data to achieve its goal. A consideration of how ETHPOP projected the change in population in each LAD in the past informs how these future projections beyond 2011 should be adjusted. Where these adjustments are less than credible they are moderated in the context of similar types of authority. The rebasing to 2011 Census counts makes sense since they are a known and trusted starting point for the revised projections beyond 2011. Finally, the constraint to the 2012-based SNPP provides the reassurance of a population whose gender and age structure matches an external source with good veracity. Whether the revisions (or indeed the original ETHPOP projections) are truly accurate will only become apparent when other more sophisticated/competing projections, or the actual data itself, become available. Indeed the health microsimulation is flexible enough to allow the use of alternative ethnic population projections to these revised estimates, be they either the unrevised ETHPOP projections or ethnic population projections available from other sources (e.g. the Greater London projections, Klodawski, 2009).

So far in this thesis the base 2011 population for each LAD has been synthesised and a series of revised population projections by gender, age and ethnicity have been created. Before the health microsimulation can be conducted, however, further information is required concerning the probabilities for use in the Monte Carlo process that updates the status of individuals in the population. This is covered in the following chapter.

8 PROBABILITY MODELS

The Dynamics of Change through Monte Carlo Simulation

8.1 INTRODUCTION

The health microsimulation model introduced in section 5.1 and illustrated in Figure 5-1 is dynamic in so far as individuals naturally age over time, their morbidity status is changed to reflect their personal circumstances, new individuals are introduced into the population to replenish its membership and individuals leave the population through migration and mortality.

The mechanism adopted to effect these changes in the population is the Monte Carlo approach. This requires an estimate of the probability that an event will take place and a method of sampling a pseudo random number. A comparison between these two probabilities enables the occurrence of an event to be simulated. If the estimated probability of occurrence is say, 0.10, and the sampled random number is less than or equal to 0.10, then the event is deemed to of occurred, otherwise it is deemed not to of occurred. The random sampling is relatively straight forward (akin to rolling a die), using a pseudo random number generation algorithm (Matsumoto and Kurita, 1994). The estimation of the probability of the event is more challenging. In this study, these probabilities are primarily estimated using hazard models fitted to ELSA data, although in one case probabilities based on HSfE survey data are used.

The following section, 8.2, describes the estimation of the hazard models for each of the three morbidities studied in this thesis: CVD, DHBS and respiratory illness. These are the models that update an individual's morbidity status. In section 8.3, hazard models are also described that estimate the probability of an individual dying or migrating. The purpose of these probabilities is to define the selection weights that are used in the process which reshapes the LAD population to some given structure. The fourth section, 8.4 deals with the task of how to characterise the population that replenishes the population as it ages, i.e. introduce those aged 50 and 51. This involves firstly, the task of describing how estimates of morbidity prevalence rates at ages 50 and 51 are calculated for these future Replenishers. This is accomplished using the morbidity models discussed in section 8.2 to accumulate 'historic' morbidity incidence probabilities over time. Secondly, there is the task of estimating the probability that the Replenishers have ever smoked, and here, use is made of the 2011 edition of the HSfE. The final section will recap the work described in this section and lead onto the following chapter which presents the results of the health microsimulation.

8.2 MORBIDITY HAZARD MODELS

As described in section 4.4.7 an adapted version of the ELSA wave 6 mechanism for determining the absence or presence of a morbidity at each wave is used here. The wave 6 definition of the presence of a morbidity differs from that used in previous waves by not relying on the fed forward morbidity information provided at each wave. Instead each previous wave's core data are explicitly examined for a new diagnosis or a feed forward of a diagnosis. The adaptation here is to discount the information on the feed forward status of a morbidity and instead rely only on explicit new diagnosis at the current or previous waves.

The hazard models are estimated using the wave 6 longitudinal weights which correct for attrition of C1CM participants in the ELSA survey (see section 4.4.3). The explanatory variables available are those which are known in the sample or can be inferred for those who replenish the sample (see section 8.4.2). The chronological variables are wave, which acts as a time trend, each wave covering two years, age centred on age 50 (i.e. the value used in the model is the age in years minus 50) and also a centred age squared term. The socio-demographic variables are gender and ethnicity. As seen elsewhere in this thesis, the representation of BME groups is poor in ELSA and in order to ensure that there is sufficient event occurrence (i.e. the acquisition of a morbidity) some merging of the five ethnic groups is required. These three merged groupings are: a White group, an Asian group and Mixed/Black and Other group. Smoking status, recorded as whether someone has ever smoked a cigarette, is one reliable lifestyle indicator available for both existing members of ELSA and those who will replenish the sample at younger ages. In the case of CVD there is also an additional variable which is an indication of whether the individual had DHBS at the previous wave. This is to reflect the finding that those who have DHBS are at a greater risk of developing a CVD (Public Health England, 2013b).

Whilst the above variables provide information about the socio-demographic, economic and lifestyle circumstances of the individual, there is also a potential role for contextual information. This contextual information is provided by information concerning the type of area in which the individual lives. In the more recent releases of ELSA data the (now abolished) Government Office Region of residence of the participant is provided, however, it is unlikely that this information will be sufficient in itself to capture these area type effects. On request, however, the actual LAD of residence at each wave is obtainable from the ELSA team. This extra level of geographical detail has been obtained and allows the inclusion of Group classes from the ONS area type classifications to be included as explanatory terms. It should be noted here that this classification is a general classification and not one aimed at distinguishing LADs by health outcomes (in the classification health is measured using just two variables,

disability and care provision, part of the eight variables for the socio economic domain, which itself is one of six domains used to construct the classification).

In the estimation of statistical models sometimes the models are refined in stages to exclude insignificant terms from the model. The rationale behind this decision is that such terms do not contribute significantly in influencing the outcome but they do 'consume' degrees of freedom thereby potentially undermining the assessment of the quality of the fit of the model. Here, however, the alternative approach is adopted where all terms are kept in the model, even if they are not significant. The justification is that, firstly, even an insignificantly estimated term is a measured effect that has some influence but it cannot be said, with a reasonable degree of confidence, that the effect is large enough to be present. Secondly, ELSA is a large data set, potentially providing 5 observations on each of 5,000 individuals, so the reduction in the degrees of freedom for the regression on retaining some insignificant terms is negligible. The third justification is that by including some insignificant terms the heterogeneity of the outcomes is increased, even slightly, creating a population whose outcomes are more diverse than if the term was not included. However, this retention of insignificant terms is limited only to main effects in the model, interactions between main effects (e.g. wave and age) are only included if they are significant.

When fitting hazard models there are issues around the censoring of data. Censoring happens when the event occurs outside the observation period. Right censorship, where the event occurs subsequent to the period of observation is not as problematic as left censoring, where the event occurs before the period of observation. With right censoring it is known that an event did not occur during the observation period and it is also known the circumstances of the individual during this period. This is valid and compatible information to use in the model. With left censoring it is known that the event occurred prior to the observation period but not necessarily the circumstances of this individual at the point of occurrence – it is outside the observation period and provides no information on occurrence (Singer and Willett, 2003). The usual approach to this problem is to drop those individuals who are left censored. The practical implication of this requirement is that our population starts at wave 2 with those individuals who did not have the morbidity at wave 1. The count at wave 2 will be those individuals who did not have the morbidity at wave 1 but acquired it by the time they were surveyed at wave 2. There are therefore just five waves of observation. The count of the number of individuals who acquired each of the three morbidities and the rate of this occurrence over these five waves is shown in Figure 8-1.

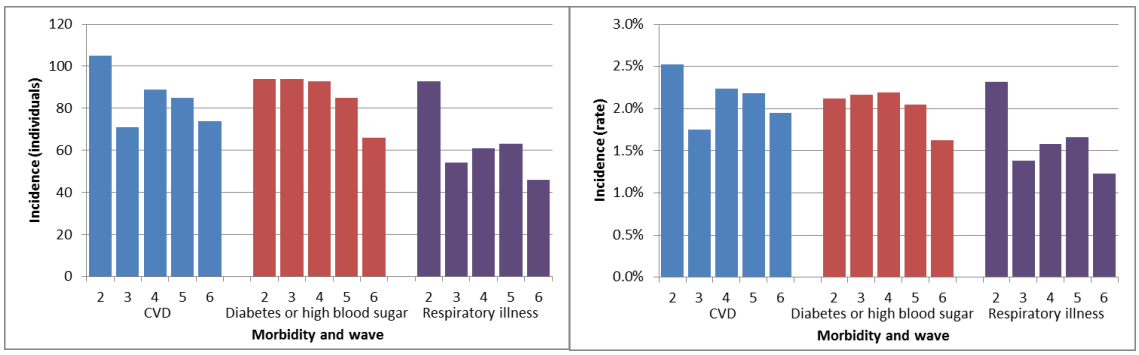


Figure 8-1 : Trends in morbidity incidence, count of individuals (lhs), rate in population (rhs), England.

For CVD there appears to be a spike in incidence at wave 2 followed by a dip at wave 3. Any trend information in these data will be exaggerated by the high incidence at wave 2 and diminished by the low incidence at wave 3 – potentially cancelling each other out. With DHBS, there appears to be a plateau for waves 2 to 4 followed by a reduction for the final two waves. For respiratory illness there is a significant spike in incidence at wave 2, the rate for this wave being nearly 1% higher than the other waves. This high initial peak may have a influence on any time trend term in models for this morbidity.

8.2.1 CVD hazard model

The estimation results for the hazard model to predict the onset of a CVD morbidity are shown in Table 8-1. The age and age squared terms are both significant and are concaved which means that the probability of acquiring a CVD increases to a maximum then decreases. The two year duration wave effect is also highly significant and negative which means that the probability of acquiring a CVD decreases over time. This term is capturing the impact of trends which are not accounted for in the other explanatory terms. These could be related to socio-demographic, socio-economic or health care influences (e.g. here the introduction of statins, Law et al., 2003). People of a south Asian ethnicity have a slightly increased probability of acquiring a CVD when compared to the White ethnic group whilst the remaining ethnic groups have a slightly reduced probability, although these ethnic effects are not statistically significant. The literature does support this enhanced risk for south Asian ethnic groups and reduced risks for the remaining non-White groups (i.e. Mixed, Black and Other (Chinese)). These ethnic parameter estimates are large and so are their standard errors, with the size of the standard errors being a reflection of the small sample size of observations for these ethnic groups.

Having ever smoked or having DHBS at the previous wave both increase the probability of a CVD. Overall the type of LAD is only statistically significant at the 10% level, and individually only the individuals living in Industrial Hinterland type LAD's show a statistically significant increased probability of a CVD over the reference LAD of

Manufacturing Towns. The type of LAD that shows the greatest reduction on the probability, relative to Manufacturing Towns, is Thriving London Periphery. The R^2 value is small, which is usual for such models. The Hosmer-Lemeshow statistic is not significant at the 0.1% or 5% but is significant at the 10% level. This suggests some weak evidence that the ordered expected and observed counts are not similar.

Table 8-1 : Parameter estimates from CVD hazard model

Variable	beta		Se(beta)	Wald	dof	Sig	Odds
Age minus 50	.1001	***	.0221	20.51	1	.000	1.105
(Age minus 50)^2	-.0011	**	.0005	5.16	1	.023	.999
Wave	-.1456	***	.0362	16.15	1	.000	.864
Female	-.2616	**	.1012	6.68	1	.010	.770
Ethnicity				1.23	2	.541	
White (reference)							
South Asian	.4548		.4504	1.02	1	.313	1.576
Mixed/Black/Other	-.1937		.4815	0.16	1	.688	.824
Ever smoked	.2307	**	.1096	4.43	1	.035	1.260
Diabetes or hbs at previous wave	.4326	***	.1527	8.03	1	.005	1.541
LAD Type		*		17.92	11	.083	
Manufacturing Towns (reference)							
Regional Centre	-.3294		.2672	1.52	1	.218	.719
Centre with Industry	-.1899		.2400	0.63	1	.429	.827
Thriving London Periphery	-.6936		.4507	2.37	1	.124	.500
London Suburban	-.4741		.3222	2.16	1	.141	.622
London Centre	.0767		.4829	0.03	1	.874	1.080
London Cosmopolitan	-.5503		.4660	1.39	1	.238	.577
Prospering Smaller Towns	-.0797		.1823	0.19	1	.662	.923
New and Growing Towns	.0040		.2591	0.00	1	.988	1.004
Prospering Southern England	-.0428		.2165	0.04	1	.843	.958
Coastal and Countryside	-.0168		.2218	0.01	1	.939	.983
Industrial Hinterlands	.3717	*	.2094	3.15	1	.076	1.450
Constant	-4.6958		.3116				
Nagelkerke R^2	0.037						
Hosmer-Lemeshow	14.693	*			8	.065	
N	20,018						

* significant at the 10% level, ** significant at the 5% level, *** significant at the 1% level

In the use of this and the remaining morbidity models there is the implicit assumption that the contextual impact on the morbidity estimated through the LAD area classification will remain constant beyond 2011 to 2031. Rees (2011) charts the changes in life expectancies by deprivation quintile and an alternative area classification over the period 1991 to 2007 in his figure 1.3 (reproduced here as Figure 8-2). This chart shows that the relative position by area type and, especially, deprivation quintile does not change much over time. This relative stability in life expectancies lends weight to the

assumption that in the short to medium term the relative influence of these area types remain consistent.

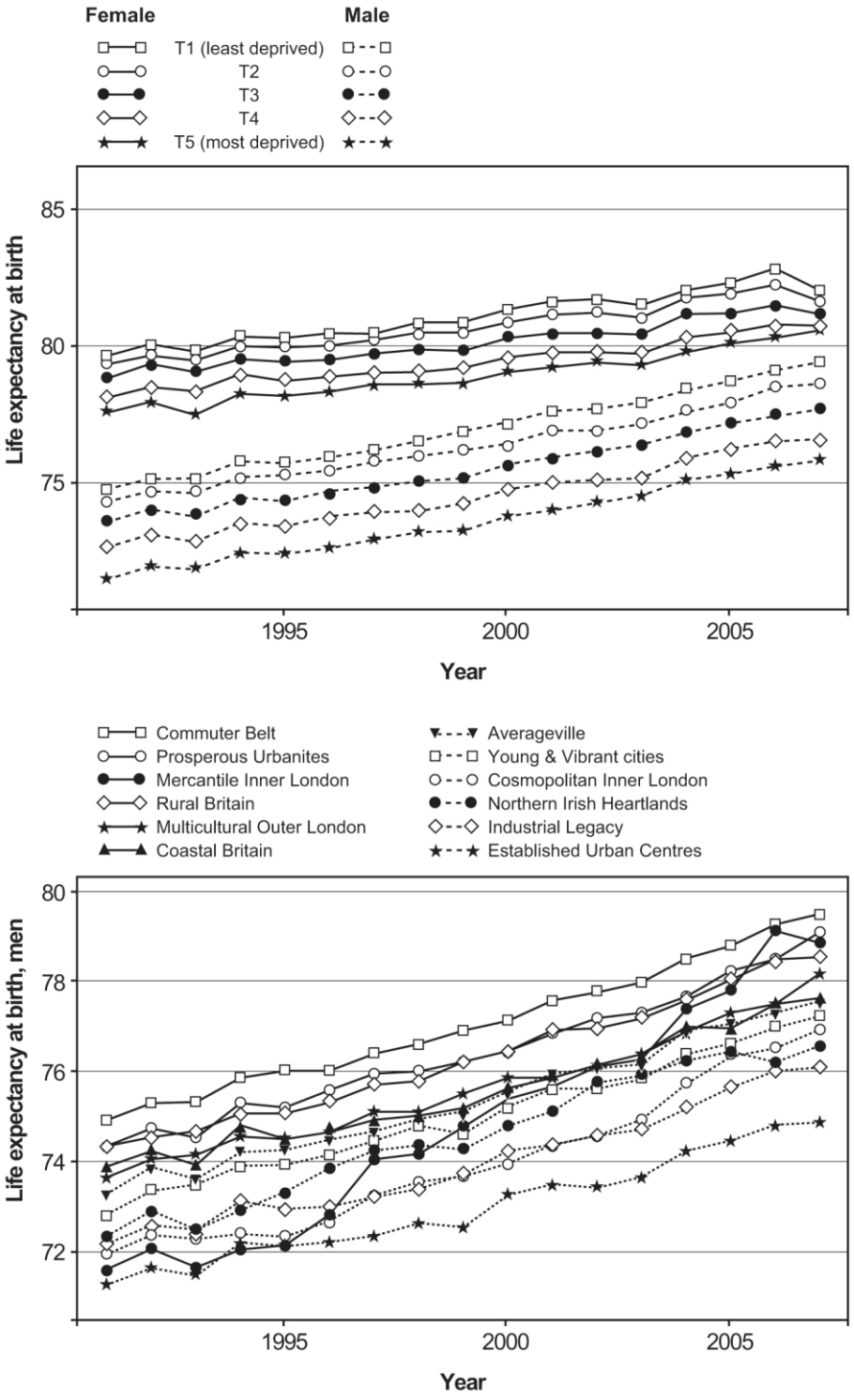


Figure 8-2 : Graphs illustrating the stability of deprivation and area classification rankings over time (Rees, 2011)

Motivated by the trend for the incidence of CVD in Figure 8-1 a revised model was fitted with two intervention terms to account for the potentially anomalous counts at waves 2 and 3. These interventions are coded as 0.0 for all values, except for the wave

they are relevant to, when they are coded as 1.0. These two intervention terms were not significant and the wave time trend remains high at -0.167. In light of this the model of Table 8-1, without the interventions, is preferred.

8.2.2 Diabetes or high blood sugar hazard model

The estimation results for the hazard model to predict the onset of DHBS are shown in Table 8-2.

Table 8-2 : Parameter estimates from diabetes or high blood sugar hazard model

Variable	beta		Se(beta)	Wald	dof	Sig	Odds
Age minus 50	.0525	**	.0223	5.56	1	.018	1.054
(Age minus 50)^2	-.0010	**	.00052	3.99	1	.046	.999
Wave	-.0785	**	.0362	4.70	1	.030	.925
Female	-.2098	**	.0994	4.46	1	.035	.811
Ethnicity		***		24.17	2	.000	
White (reference)							
South Asian	.9377	***	.3102	9.14	1	.003	2.554
Mixed/Black/Other	1.1624	***	.2749	17.88	1	.000	3.198
Ever smoked	.0716		.1053	0.46	1	.497	1.074
LAD Type		*		18.59	11	.069	
Manufacturing Towns (reference)							
Regional Centre	-.3401		.2513	1.83	1	.176	.712
Centre with Industry	.1424		.2005	0.50	1	.478	1.153
Thriving London Periphery	-.7306	*	.4152	3.10	1	.078	.482
London Suburban	-.3527		.2750	1.64	1	.200	.703
London Centre	.3663		.3652	1.01	1	.316	1.442
London Cosmopolitan	-.5144		.3964	1.68	1	.194	.598
Prospering Smaller Towns	-.3361	*	.1741	3.73	1	.054	.715
New and Growing Towns	-.1666		.2490	0.45	1	.503	.847
Prospering Southern England	-.3550	*	.2129	2.78	1	.095	.701
Coastal and Countryside	-.3082		.2225	1.92	1	.166	.735
Industrial Hinterlands	-.1365		.2150	0.40	1	.526	.872
Constant	-3.8541		.2792				
Nagelkerke R^2	0.015						
Hosmer-Lemeshow	11.461				8	.177	
N	21,312						

* significant at the 10% level, ** significant at the 5% level, *** significant at the 1% level

Again, the age effect is concaved and significant and the wave time trend is negative and significant. Females are less likely to acquire DHBS than males. There are strong ethnic trends with this morbidity. Both the south Asian and the remaining group of Mixed/Black or Other have a much increased probability of developing this morbidity relative to the White ethnic group. The impact is strongest for the Mixed/Black or Other group which is contra to some studies that say, whilst the Asian and Black groups both have elevated risk of DHBS, the risk is highest for the south Asian group, but again the

standard errors for these estimates are large making and relative comparison difficult. The lifestyle variable of ever having smoked has only an insignificant small positive impact on the probability of acquiring this morbidity. As with CVD, the LAD area type effects are significant at the 10% level. The LAD type with the greatest increase in probability relative to Manufacturing Towns is London Centre authorities, and the greatest reduction is once more Thriving London Periphery.

8.2.3 Respiratory illness hazard model

The estimation results for the hazard model to predict the onset of respiratory illness are shown in Table 8-3. This model includes a significant intervention term for wave 2. A model without this intervention gives a wave effect of -0.136, a significantly higher trend effect than that shown in Table 8-3, caused by the high incidence at wave 2 which inflates the trend. This high incidence of respiratory illness at wave 2 could be a genuine one-off event or a potential issue in these data. In either case, to mitigate its lasting influence, the model with a wave 2 intervention term, see Table 8-3, is preferred.

Unlike the previous two hazard models, here there is no significant age effect on the probability of developing a respiratory illness. There is also a weak negative time trend effect once the intervention at wave 2 is accounted for. Females are slightly more likely to acquire a respiratory illness although the effect is small and insignificant. Being a member of a BME group increases the probability of a respiratory illness and this is most pronounced for the Mixed/Black or Other group. As expected ever having smoked is a significant and positive contribution to the probability of having a respiratory illness. Overall the LAD area type does have a statistical significant impact of developing a respiratory illness relative to Manufacturing Towns. There are no instances of the onset of a respiratory illness in a London Centre type of authority, so to enable the estimation of a model with a LAD area type term this category is merged with London Cosmopolitan. Surprisingly the Coastal and Countryside area type has the greatest increase in probability and London Suburban the least.

Table 8-3 : Parameter estimates from respiratory illness hazard model

Variable	beta		Se(beta)	Wald	dof	Sig	Odds
Age minus 50	.0308		.0250	1.53	1	.217	1.031
(Age minus 50)^2	-.0006		.0006	1.13	1	.287	.999
Wave	-.0379		.0617	0.38	1	.539	.963
Wave 2	.4186		.1939	4.66	1	.031	1.520
Female	.0570		.1163	0.24	1	.624	1.059
Ethnicity		***			14.60	2	.001
White (reference)							
South Asian	.9878	**	.4290	5.30	1	.021	2.685
Mixed/Black/Other	1.1903	***	.3642	10.68	1	.001	3.288
Ever smoked	.7516		.1394	29.06	1	.000	2.120
LAD Type		**			18.57	10	.046
Manufacturing Towns (reference)							
Regional Centre	-.2473		.2983	0.69	1	.407	.781
Centre with Industry	-.2066		.2684	0.59	1	.441	.813
Thriving London Periphery	-.2011		.4246	0.22	1	.636	.818
London Suburban	-.9331	**	.4050	5.31	1	.021	.393
London Centre and Cosmopolitan	-.6521		.4214	2.39	1	.122	.521
Prospering Smaller Towns	-.2036		.2091	0.95	1	.330	.816
New and Growing Towns	-.4364		.3242	1.81	1	.178	.646
Prospering Southern England	-.0406		.2421	0.03	1	.867	.960
Coastal and Countryside	.3023		.2358	1.64	1	.200	1.353
Industrial Hinterlands	.1228		.2478	0.25	1	.620	1.131
Constant	-4.8359		.4063	141.64			
Nagelkerke R^2	0.024						
Hosmer-Lemeshow	5.300				8	.724	
N	19,379						

* significant at the 10% level, ** significant at the 5% level, *** significant at the 1% level

8.3 SELECTION WEIGHTS

In section 5.5.4 the process of re-shaping the population within each LAD to match a predefined ethnic composition is introduced. This is primarily achieved through the process of cloning existing individuals within the population to mimic both population growth through in migration and population decline through mortality and out migration. Since these selections are stochastic in nature, there is a need to choose individuals in the population to be cloned or removed. This is done here using a Monte Carlo approach with unequal selection probabilities or weights. The justification for the unequal selection is that the characteristics that individuals possess will influence their probability of mortality or migration. The operation of these unequal selection probabilities can have a large impact on the prevalence rates within the population, as exemplified in Box 5-11. In this section the hazard model that predicts the probability of mortality and the one that predicts migration are discussed. It should be emphasised that these models in application are not meant to predict how many people are going to die

or move between LADs, since this is a component implicitly included in the population projections. What these models do is predict which individuals are likely to die or move.

8.3.1 Mortality hazard model

The knowledge of the death of a participant in the ELSA study is recorded well (but see Figure 8-3 below). Not only is the death of a participant potentially apparent when a wave interview is attempted, but the death of all participants, be they core members, partners of core members or sample members, is recorded using information contained in the NHS Central Register. Thus there are fewer issues around attrition for this outcome than for morbidities – indeed there are no suitable weights in ELSA to help to correct for attrition in regards to mortality.

This lack of participation in all waves for those who have died does cause an issue with establishing if they have a morbidity when they die. If someone is recorded as having a morbidity at an earlier wave then due to the absorptive nature of the way the morbidity is defined they will have the morbidity at death. Thus for someone who died between waves 3 and 4 but was only observed at wave 1, at which time they had CVD, then it is correct to assume they had the morbidity when they died. The converse is however not true, if they did not have CVD at wave 1 and were never surveyed again, it is not possible to infer if they had CVD at death. If it is thought that the morbidity status makes a significant contribution to the probability of mortality, then there is a dilemma. A 'quantity' approach keeps individuals in the sample but at the expense that the morbidity information known about the individual up to death is not certain. A 'quality' approach only includes those individuals in the sample whose morbidity status is known during the period up to death. It is this latter approach that is adopted here. This is driven by the potential utility of the morbidity status to influence mortality and hypothesising that the highest quality and relevant information is most likely to lead to an informed outcome. The drawback is that the sample size is diminished by around 50%, but this still leaves a substantial sized sample on which to estimate the hazard model.

The mortality wave variable ('*when died in relation to ELSA wave*') is recorded in the Index File which is only currently available for waves 1 to 5 so the analysis here is limited to the first five waves of ELSA. As with morbidities, left censoring means that the counts for the first wave are not provided. Figure 8-3 plots the number of mortalities at each wave and in this figure there appears to be an unusual spike of mortalities at wave 2. To prevent the presence of this spike having undue influence on the wave trend term in the hazard model, a wave 2 intervention term is used in the model (as was similarly used for the respiratory illness hazard model in section 8.2.3).

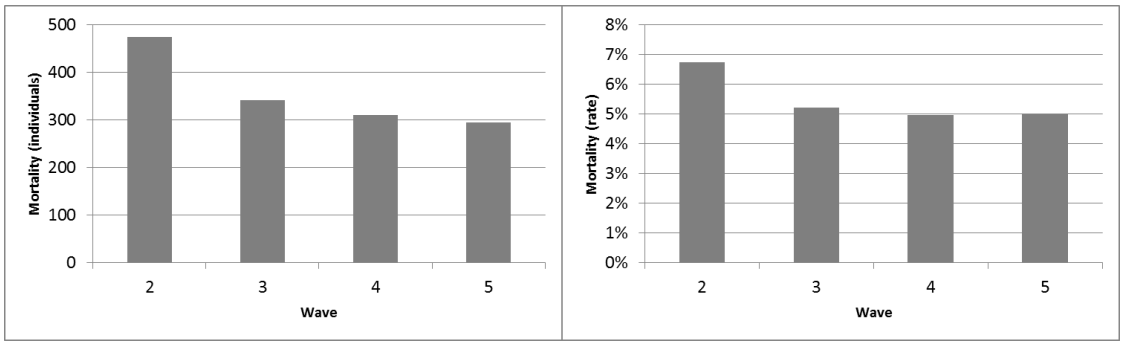


Figure 8-3 : Trends in mortality incidence, count of individuals (lhs), rate in population (rhs), England.

A critical hypothesis is that the probability of death may be influenced by the characteristics of the participant, some of which are not contained in the Index File (e.g. ethnicity or morbidity status). This information is transferred to the participant from the wave core files as outlined in section 4.4.7.

The parameter estimates from the mortality hazard model are shown in Table 8-4. As individuals get older the probability of dying increases, a not unexpected result. Over time however the probability of death is decreasing but there is a significant interaction between these two chronological events. The interaction is positive indicating that the reduction in mortality for older people in future years will be less pronounced. All other things being equal, females are less likely to die than males and there is some evidence of an ethnic influence, with the south Asian population having a smaller probability of death than the remaining ethnic groups. Someone who has ever smoked and has one of the studied morbidities is more likely to die, although DHBS has the least impact (of the three morbidities DHBS is the least likely to be recorded as the main cause of death, ONS, 2014d). The pseudo R^2 statistic is good for models of this kind and the Hosmer-Lemeshow test estimates no significant difference between the expected and observed counts in each ordered decile.

Whilst the estimates provided in Table 8-4 are rational in terms of their significance and direction of impact, there may be issues around the actual scale of the impacts. In particular the time trend reduction looks dramatic, even when the potential outlier at wave 2 is accounted for. However recall the purpose of these probabilities, they provide weights for the probability of selection for removal from the synthesised population and not to predict actual death. The actual mortality and migration impacts on the size of the population are implicitly contained in the population projections that determine how many (but not which) individuals should be selection for removal. The true utility of the results in Table 8-4 come from the smoking and morbidity status variables which do explicitly allow for a differentiation of the weight of selection by these characteristics.

Table 8-4 : Parameter estimates from the mortality hazard model

Variable	beta		Se(beta)	Wald	dof	Sig	Odds
Age minus 50	.0846	***	.0085	99.378	1	.000	1.088
Wave	-.3156	***	.0875	12.998	1	.000	.729
Wave 2	.2362	**	.1047	5.089	1	.024	1.266
Wave by age minus 50	.0062	**	.0025	6.106	1	.013	1.006
Female	-.4281	***	.0599	51.086	1	.000	.652
Ethnicity in 3 classes		*		5.265	2	.072	
White (reference)							
South Asian	-.5658		.4551	1.546	1	.214	.568
Mixed/Black/Other	.5508	*	.2878	3.663	1	.056	1.735
Ever smoked	.4863	***	.0706	47.429	1	.000	1.626
Has CVD	.4132	***	.0615	45.181	1	.000	1.512
Has diabetes or hbs	.1768	**	.0827	4.569	1	.033	1.193
Has respiratory illness	.2559	***	.0682	14.075	1	.000	1.292
Constant	-4.6630	***	.3141	220.355	1	.000	.009
Nagelkerke R^2	0.213						
Hosmer-Lemeshow	7.618				8	0.473	
N	24,383						

* significant at the 10% level, ** significant at the 5% level, *** significant at the 1% level

There may still however be some re-assurance gained by comparing the probability estimates from the hazard model of Table 8-4 with other sources of similar outcomes. One such source is the National Life Tables for England covering the period 2010-2012 (ONS, 2014e). These life tables are produced separately by gender but not by ethnicity. In Figure 8-4 a number of demographic scenarios are presented. In each chart the solid blue and red lines are the male and female estimates of mortality probability by age for the general population and do not vary between the charts in this figure. To calculate this probability, account needs to be taken of the fact that the hazard model estimates the mortality over a two year duration, so the mortality probabilities in the published life tables for two successive ages are compounded (equation 8-1).

$$ {}_2q_x = 1 - \left(1 - {}_1q_x\right) \times \left(1 - {}_1q_{x+1}\right) \qquad 8\text{-}1 $$

where ${}_iq_x$ is the mortality rate between age x and (x +1) over interval i.

The dotted lines are estimates of mortality probabilities for various demographic populations as estimated using the hazard model of Table 8-4 at wave 5. The chart in the top left and corner shows these probabilities for individuals from the south Asian ethnic group who do not smoke and have no morbidities. This is the healthiest combination of characteristics and the morality probability for this group is less than half that of the general population. The next three charts are for the majority ethnic group of White. The top right charts show the mortality probability for the non smokers who have no morbidities, the middle left the probability for smokers without a morbidity and the middle right for smokers with a respiratory illness – showing an

increasing trend for morbidity in each successive chart. It is these trends that influence how more likely a member of the White ethnic group who smokes and has a respiratory illness is to die than someone of the same ethnicity who, say, doesn't smoke or have a respiratory illness. Note that only for the combination of the White ethnic group, who smoke and have a respiratory illness are the probabilities close to those from the national life table. The remaining two charts at the bottom of Figure 8-4 provide two scenarios for the Mixed, Black or Other population. The left hand chart is for someone in this ethnic grouping who smokes but has no morbidities whilst final right hand chart is also for this same combination but instead does have a respiratory illness.

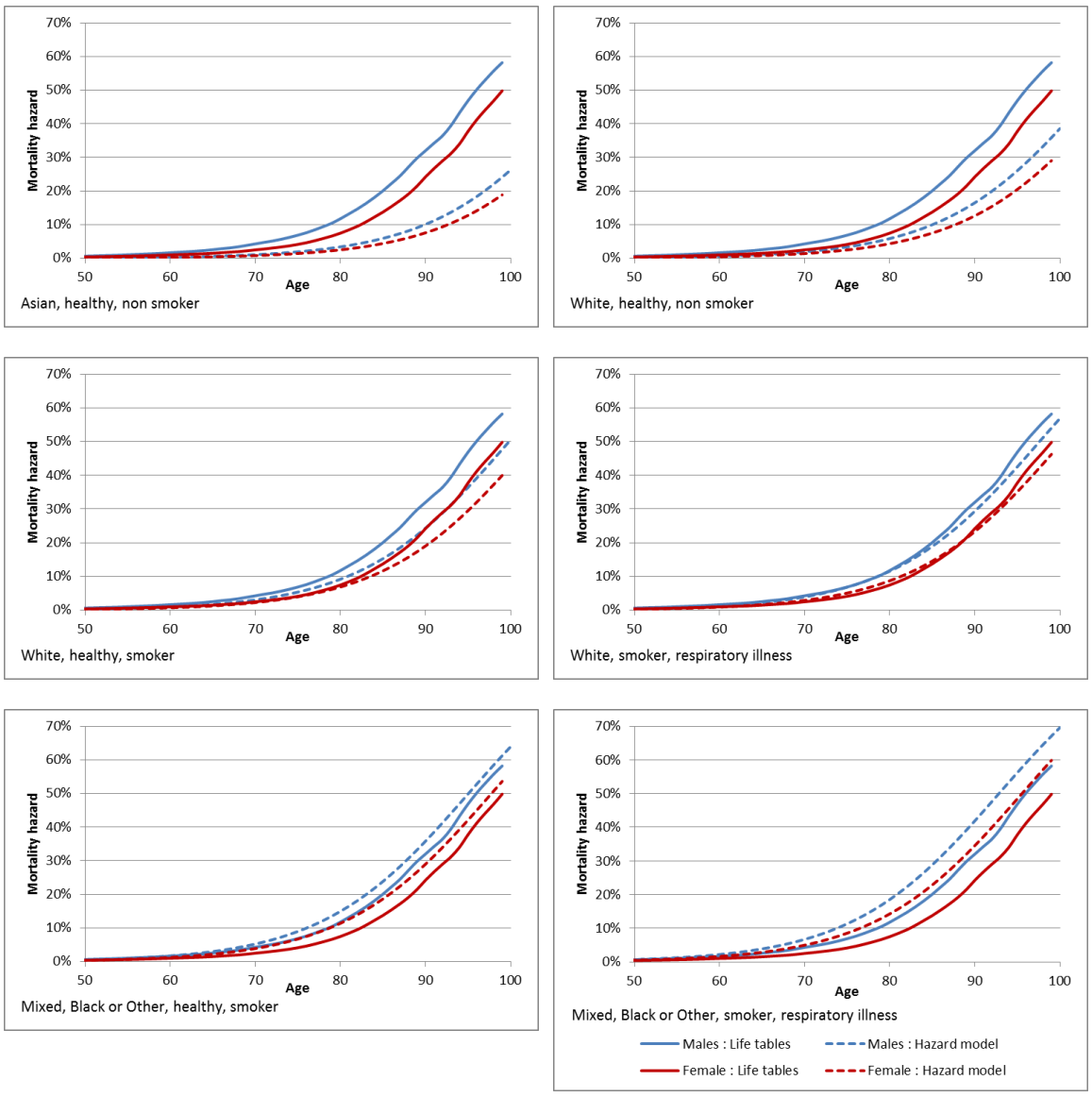


Figure 8-4 : Comparison of ONS life table and hazard model mortality probabilities

Clearly the morbidity probabilities of the general population are a weighted calculation of the various combinations of characteristics, some of which are shown in Figure 8-4.

However, without access to the actual population estimates of the size of these sub populations it is difficult to re-construct a general population estimate for the model in Table 8-4. What the charts do illustrate however is the impact of the smoking and morbidity characteristics on the relative weight for selection during the health microsimulation.

8.3.2 Migration hazard model

In contrast to the case for mortality, where the information on the occurrence of a death is well recorded, and supplemented using an administrative source external to ELSA, the occurrence of a move between LADs is more problematic. Information on the LAD of residence is not normally available with the publicly released ELSA data but on request, the information on LAD of residence may be obtained from the ELSA team. This information has been received for waves 1 to 6 and it is therefore possible to identify those ELSA participants who moved between LADs in the two year period between ELSA interviews. Figure 8-5 shows these migrations as migration counts and rates.

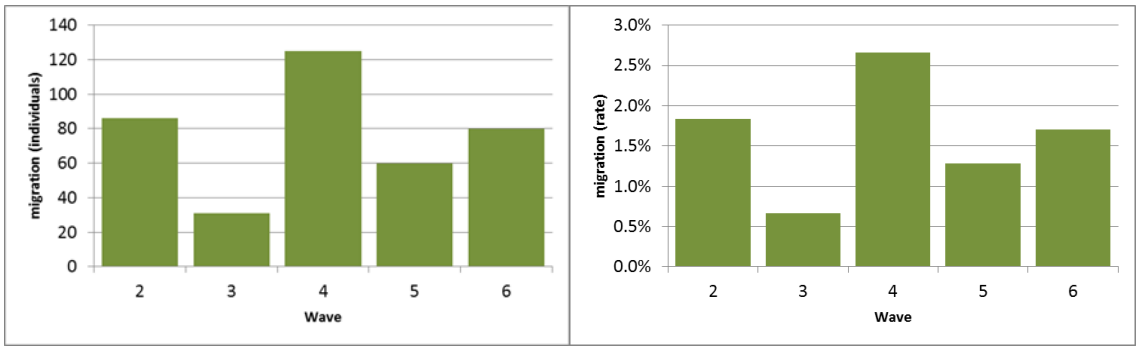


Figure 8-5 : Trends in migration, count of individuals (lhs), rate in population (rhs), England.

These data are quite erratic; with a sharp dip in migration between waves 2 and 3 and a spike between waves 3 and 4. If these two outliers are averaged then the series is more level. As was done with other hazard models, to prevent these features having an undue influence on the wave trend estimate, two intervention terms are included, one for wave 3 and a second for wave 4.

There will also be other participants who left ELSA through having moved house and did not provide a follow up address – an outcome of '*Moved – unable to trace*'. These people are known movers, however use of this information becomes problematic. What is required for the health microsimulation is a probability that a person has moved LAD between waves - moves with the LAD are of no interest. The ELSA outcome cannot distinguish between these inter and intra LAD moves. Also recall that the longitudinal weights that correct for attrition within ELSA are provided as part of the ELSA data set only for those C1CM participants in every wave. Those who left ELSA with a '*Moved* -

unable to trace' outcome will not have a longitudinal weight. If they are to be used in the sample for estimation then it is not possible to use the longitudinal weight and therefore not possible to correct for attrition due to other reasons.

By restricting ourselves to those participants whose migration journey is known, it is likely that these migrations probabilities will be an underestimate of the true migration probabilities – people who do not migrate are more likely to be retained in the sample and those who move and are untraceable will be missed. However, as for the mortality hazard model, the aim here is not to exactly replicate migration probabilities but to differentiate migration probabilities within a population defined by its gender, age band and ethnicity for a particular wave.

The parameter estimates from the migration hazard model are shown in Table 8-5. The occurrence of a move is a rare event and becomes even rarer when disaggregated by various factors, for example, there are only three recorded moves by BME participants who are C1CM and responded at all waves. Thus the level of disaggregation by this ethnic characteristics is restricted.

Table 8-5 : Parameter estimates from the migration hazard model

Variable	beta		Se(beta)	Wald	dof	Sig	Odds
Age minus 50	-.0788	***	.0191	16.94	1	.000	.924
(Age minus 50)^2	.0015	***	.0004	13.83	1	.000	1.002
Wave	.0071		.0422	0.03	1	.867	1.007
Wave 3	-.9267	***	.1977	21.97	1	.000	.396
Wave 4	.5299	***	.1135	21.81	1	.000	1.699
Female	.0978		.1070	0.84	1	.361	1.103
BME	-1.3391	**	.5648	5.62	1	.018	.262
Ever smoked	-.0344		.1101	0.10	1	.755	.966
Has CVD	.2747	**	.1375	3.99	1	.046	1.316
Has diabetes or hbs	-.0004		.1676	0.00	1	.998	1.000
Has respiratory illness	.1153		.1284	0.81	1	.369	1.122
Constant	-3.4444	**	.2356	213.76	1	.000	.032
Nagelkerke R^2	0.025						
Hosmer-Lemeshow	2.332				8	0.969	
N	23,441						

* significant at the 10% level, ** significant at the 5% level, *** significant at the 1% level

The age impact on migration is quadratic and convexed, which indicates that there is an age at which migration is least likely. Once the anomalous observations at wave 3 and 4 are accounted for there appears to be only a weak positive time trend. The gender effect is also insignificant with a slight indication that females have a heightened probability of migration. Perhaps the largest impact is on the migration probability is due to ethnicity. As noted above, moves by such groups are rare and this is reflected in a large estimate in the reduction in the probability of migration for members of the BME group.

The standard error for this estimate is also large, a result of the groups small sample size. The lifestyle characteristic of having smoked and the morbidity statuses are largely insignificant except for having CVD, which, apparently increases the probability of migration. The pseudo R^2 value is not very large but the Hosmer-Lemeshow test is insignificant which shows there is no significant difference between the ordered expected and observed counts.

As was done for the case of the mortality hazard, it is useful to benchmark the migration probabilities against other measures. This is possible using the migration output from the 2011 Census. Table MM01 provides an origin destination matrix of moves between LADs categorised by age. The three age bands used here are 50 to 64, 65 to 74 and 75 and older. The number of individuals moving between LADs is the sum of the off diagonal terms in each of these three matrices. The number of such moves, the population within the appropriate age band, and the resultant estimate of the proportion who move over 1 and 2 years is show in Table 8-6.

Table 8-6 : 2011 Census estimates of migration proportions

Age range	Movers	Population	Migration probabilities (one year) [a]	Migration probabilities (two years) [$1 - (1 - a)^2$]
50 to 64	182,564	9,569,364	1.91%	3.8%
65 to 74	56,440	4,552,283	1.24%	2.5%
75 and older	56,451	4,108,246	1.37%	2.7%

The tendency to migrate is greatest in the youngest age band and least in the middle age band. The initial high rate of migration may be caused by individuals in newly retired households moving home and leaving the LAD. As people get older however the tendency to migrate will naturally decrease, but at the oldest ages there may be a need to migrate once more to access care and social support, particularly from family and friends. This pattern of migration is consistent with the figures in Table 8-6.

The migration probabilities estimated using the hazard model in Table 8-5 for non-smoking individuals without a morbidity of both genders and ethnicities are shown in Figure 8-5 (lhs). The 2011 Census equivalent estimates over a two year wave are shown as horizontal dotted lines. The convexed nature of both the relationships is apparent, although the hazard model probabilities tend to be at their minimum a little after the 2011 Census data would suggest.

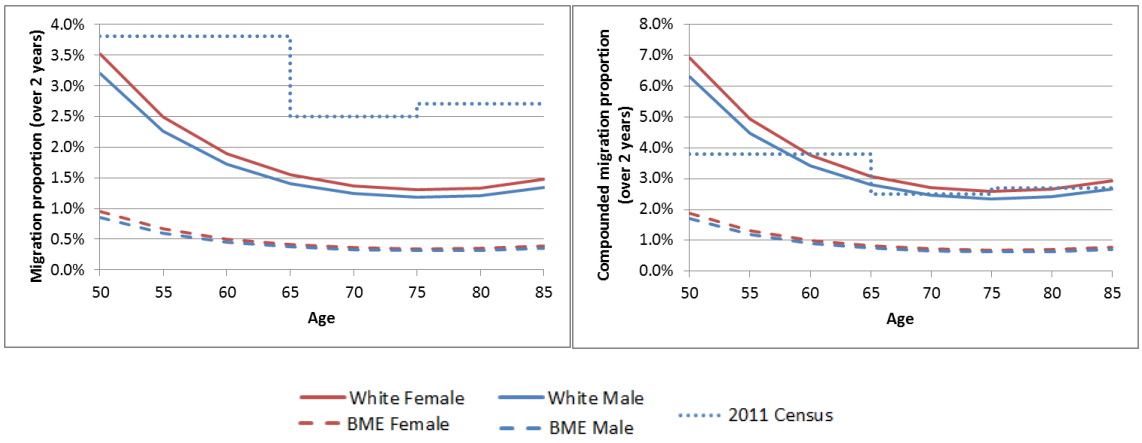


Figure 8-6 : Migration probabilities from the migration hazard model - single wave (lhs) and compounded (rhs), England, 2011 Census

Clearly the hazard model under predicts the probability of migration across the whole age range. This is however to be expected since there are a whole class of movers who leave ELSA and are untraceable. If this under recording is 'corrected' by assuming that the estimated probabilities should be compounded over two waves, then the probabilities are now as shown in Figure 8 5 (rhs). Here the agreement between the hazard probabilities and the 2011 Census derived estimate is good, especially for those aged 65 and older. This compounding of the two year duration wave migrations is not ideal but is driven by these data (using hazard model probabilities only) and corrects for an expected and observed under estimate. The migration probabilities between LADs of the BME ethnic group are substantially less than those of the White group.

Having mentioned earlier that these migration probabilities are not intended to determine the degree of migration but just to differentiate the propensity of individuals to migrate, the question arises as to why is it important that the compounding happens. Instead why not just use the probabilities shown in Figure 8-6 (lhs)? The reasoning is discussed below.

8.3.3 Use of selection weights

As discussed in 5.5.4 the mortality and migration hazard models are used to create weights of selection for individuals for the restructuring of the population to agree with a population projection. Individually these mortality and migration probabilities (after compounding in the latter case) were shown to be 'ball park' when compared with official administrative and 2011 Census data sources. When selecting individuals for removal both the mortality and migration probabilities are combined to produce the weight. This means that across all age ranges these two weights contribute their own influence. At younger ages the migration probability will have the most influence on the individual's weight whilst at older ages the mortality probability will have the most

influence. To ensure that the relative influence of these two demographic effects is correct the relative probabilities of mortality and migration need to be correct. Without compounding the migration impact will be less influential than mortality, especially at younger ages.

8.4 STATUS OF REPLENISHERS

In section 5.5.3 the concept of Replenishers was introduced to ensure that the projected population for future years covered the full age range of 50 years of age and older. The creation of Replenishers required some knowledge of their characteristics. The gender, age and ethnicity are given by the interpolated counts of 50 and 51 year olds from the revised ETHPOP projections estimated using Sprague's method (Sprague, 1880). What is further required is their morbidity status with regards to CVD, DHBS and respiratory illnesses and their smoking status at age 50 or 51. How these requirements are satisfied is explained in the following two sections.

8.4.1 Morbidity status at 50 or 51

It would clearly be mistaken to assume that individuals aged 50 and 51 who replenish the sample at future waves will be in perfect health with no morbidities. Instead it will be necessary to estimate the probability of Replenishers having each morbidity as they enter the sample. This estimate is arrived at using the same hazard models as those used to update the morbidity status of individuals for future waves described in section 8.2. The calculation of prevalence at age 50 or 51 is based on the assumption that the individual did not have the morbidity at age 30 but has the potential to acquire the morbidity thereafter. The morbidity hazard models predict the probability of incidence at a given wave and age (and gender, ethnicity and smoking status) and may be applied historically (since waves are just a manifestation of time) to estimate accumulated incidence between ages 30 and 49 and hence prevalence at age 50 or 51.

For illustrative purposes consider those who will replenish the survey at synthetic wave 6 and will be aged 50 or 51 in the period 2012-13 (see Table 8-7). They would have been aged 30 at wave minus 4 (1992-93) and it is possible to calculate the probability that they acquired the morbidity at that time using the equations in section 8.2. Moving onto when they are aged 32 in wave minus 3 (1994-95), the incidence for those who do not have the morbidity from wave minus 4 can be estimated and combined with those who do have the morbidity at wave minus 4 to give the prevalence at wave minus 3. This is repeated until they arrive at wave 6, with an accumulated morbidity prevalence. The general equation for this is equation 8-2:

$$PR_{50}^{W} = 1 - \prod_{i=W-10}^{W-1} (1 - h_{50-2\times(W-i)}^{i}) \qquad 8\text{-}2$$

where PR_{50}^{W} is the prevalence rate for those aged 50 at synthetic wave W (W>5);

h_a^w is the hazard probability at age a for wave w;

$50 - 2 \times (W-i)$ translates the wave that is i waves previous to W into an age.

There is also an equivalent for those who are aged 51 at wave W (P_{51}^{W}), equation 8-3.

$$PR_{51}^{W} = 1 - \prod_{i=W-10}^{W-1} (1 - h_{51-2\times(W-i)}^{i}) \qquad 8\text{-}3$$

In the same way that someone's morbidity status changes in the full synthetic wave population, a Monte Carlo draw is used to compare against this accumulated prevalence rate and if the random number drawn is less than the rate then the Replenisher has their morbidity status when 'joining' ELSA set to having the morbidity. This same technique is used for all future cohorts of Replenishers.

Table 8-7 : Example of replenishing cohorts

	Wave	Years	First cohort Replenishers' Age		Last cohort Replenishers' Age	
Historic 'waves'	-4	1992-1993	30	31		
	-3	1994-1995	32	33		
	-2	1996-1997	34	35		
	-1	1998-1999	36	37		
	0	2000-2001	38	39		
Actual waves	1	2002-2003	40	41		
	2	2004-2005	42	43		
	3	2006-2007	44	45		
	4	2008-2009	46	47		
	5	2010-2011	48	49	30	31
Synthetic 'waves'	**6**	**2012-2013**	**50**	**51**	**32**	**33**
	7	**2014-2015**			**34**	**35**
	…	**...**			**…**	**…**
	14	**2028-2029**			**48**	**49**
	15	**2030-2031**			**50**	**51**

To see what these future prevalence rates are, a simplified version of the hazard models is used that does not distinguish by LAD. This simplification reduces the combination

of possible Replensher prevalences for illustrative purposes, however in the actual health microsimulation, the LAD terms are used (just as they would be in updating a participant's morbidity status). Staying with the simplified model, Table 8-8 shows that in the simplified model the parameter estimates of the remaining non-LAD terms do not change much when the LAD term is not used. The exceptions to this are the ethnicity parameters and the constant. Since there are few members of the ELSA sample from the BME community who develop a morbidity, the estimates of these parameters are imprecise and a structural change to the model has moved the estimates somewhat. The constant has changed to reflect the loss of what is in effect the LAD specific constants in the models described in section 8.2.

Table 8-8 : Hazard model parameter estimates for models with and without LAD

	CVD		Diabetes or high blood sugar		Respiratory illness	
	With LAD	Without LAD	With LAD	Without LAD	With LAD	Without LAD
Age minus 50	.1001	.0997	.0525	.0515	.0308	.0328
(Age minus 50)^2	-.0011	-.0011	-.0010	-.0010	-.0006	-.0007
Wave	-.1456	-.1454	-.0785	-.0778	-.0379	-.0364
Wave 2					.4186	.4214
Female	.2616	-.2639	-.2098	-.2112	.0570	.0662
South Asian	.4548	.2614	.9377	.9574	.9878	.6407
Mixed/Black/Other	-.1937	-.3549	1.1624	1.2149	1.1903	.8194
Ever smoked	.2307	.2250	.0716	.0697	.7516	.7404
Diabetes at previous wave	.4326	.4532				
Constant	-4.6958	-4.7430	-3.8541	-4.0602	-4.8359	-4.9657

When the simplified hazard models described in Table 8-8 are used to estimate this prevalence rate on Replenishers joining at future synthetic waves of ELSA the resultant prevalence probabilities for 51 year olds in different ethnic groups and with different smoking status are shown in Table 8-9 for CVD, Table 8-10 for DHBS and Table 8-11 for respiratory illness.

The ELSA wave 6 report provides some comparable estimates for 2012-2013 (wave 6) covering the age range 50 to 54, keeping in mind that the rates here are for those joining aged 51. For CHD (which is CVD without taking account of strokes and will thus be an underestimate of CVD) the prevalence rate for males aged 50 to 54 is 2.2% (95% confidence interval, 0.4% to 4.0%) and for females it is 1.3% (0.1% to 2.5%). These ELSA estimates are not differentiated by ethnicity or smoking status so are a weighted

combination of equivalent data on the wave 6 row of Table 8-9. It is likely however that the most weight will be for the White ethnic group and these ELSA wave 6 report estimates compare well with those in Table 8-9. The trend is for significant reductions in these initial prevalences so that by synthetic wave 15, the rates are around a third of those for synthetic wave 6.

Table 8-9 : Estimated prevalence rates for CVD in replenishing population

Male	non-smokers			smokers		
Synthetic wave	White	South Asian	Mixed/ Black/ Other	White	South Asian	Mixed/ Black/ Other
6	2.4%	3.1%	1.7%	2.9%	3.8%	2.1%
7	2.0%	2.6%	1.4%	2.6%	3.3%	1.8%
8	1.8%	2.3%	1.2%	2.2%	2.9%	1.6%
9	1.5%	2.0%	1.1%	1.9%	2.5%	1.3%
10	1.3%	1.7%	0.9%	1.7%	2.1%	1.2%
11	1.1%	1.5%	0.8%	1.4%	1.9%	1.0%
12	1.0%	1.3%	0.7%	1.2%	1.6%	0.9%
13	0.9%	1.1%	0.6%	1.1%	1.4%	0.8%
14	0.7%	1.0%	0.5%	0.9%	1.2%	0.7%
15	0.6%	0.8%	0.5%	0.8%	1.0%	0.6%

Female	non-smokers			smokers		
Synthetic wave	White	South Asian	Mixed/ Black/ Other	White	South Asian	Mixed/ Black/ Other
6	1.8%	2.4%	1.3%	2.3%	2.9%	1.6%
7	1.6%	2.0%	1.1%	2.0%	2.5%	1.4%
8	1.4%	1.8%	1.0%	1.7%	2.2%	1.2%
9	1.2%	1.5%	0.8%	1.5%	1.9%	1.0%
10	1.0%	1.3%	0.7%	1.3%	1.7%	0.9%
11	0.9%	1.1%	0.6%	1.1%	1.4%	0.8%
12	0.8%	1.0%	0.5%	1.0%	1.2%	0.7%
13	0.7%	0.9%	0.5%	0.8%	1.1%	0.6%
14	0.6%	0.7%	0.4%	0.7%	0.9%	0.5%
15	0.5%	0.6%	0.3%	0.6%	0.8%	0.4%

The ELSA wave 6 estimates for DHBS are 4.0% (1.6% to 6.4%) for males and 6.8% (4.3% to 9.4%) for females. The synthetic wave 6 estimates in Table 8-10 for males are greater than this 4.0% estimate, nearly twice as large. For females however the ELSA wave 6 report weights are similar to those in the table. What are also different are the relative prevalence rates by gender for this youngest age band, here they are higher for males than females. However, this higher rate for males is also seen in the ELSA wave 6 report in all the age bands covering 55 to 79, meaning that overall, the prevalence of DHBS is higher for males both here and in the ELSA report. Once again these initial prevalence rates diminish over time, although not to the same degree as seen for CVD.

Table 8-10 : Estimated prevalence rates for diabetes or high blood sugar in replenishing population

Male	non-smokers			smokers		
Synthetic wave	White	South Asian	Mixed/ Black/ Other	White	South Asian	Mixed/ Black/ Other
6	7.8%	19.0%	23.8%	8.4%	20.2%	25.3%
7	7.3%	17.8%	22.3%	7.8%	18.9%	23.7%
8	6.8%	16.6%	20.8%	7.2%	17.6%	22.1%
9	6.3%	15.4%	19.5%	6.7%	16.4%	20.7%
10	5.8%	14.4%	18.2%	6.2%	15.3%	19.3%
11	5.4%	13.4%	16.9%	5.8%	14.3%	18.0%
12	5.0%	12.5%	15.8%	5.3%	13.3%	16.8%
13	4.6%	11.6%	14.7%	5.0%	12.4%	15.7%
14	4.3%	10.8%	13.7%	4.6%	11.5%	14.6%
15	4.0%	10.0%	12.7%	4.3%	10.7%	13.6%

Female	non-smokers			smokers		
Synthetic wave	White	South Asian	Mixed/ Black/ Other	White	South Asian	Mixed/ Black/ Other
6	6.4%	15.7%	19.8%	6.8%	16.8%	21.1%
7	5.9%	14.7%	18.5%	6.4%	15.6%	19.7%
8	5.5%	13.7%	17.3%	5.9%	14.6%	18.4%
9	5.1%	12.7%	16.1%	5.5%	13.6%	17.1%
10	4.7%	11.8%	15.0%	5.1%	12.6%	16.0%
11	4.4%	11.0%	14.0%	4.7%	11.7%	14.9%
12	4.1%	10.2%	13.0%	4.4%	10.9%	13.9%
13	3.8%	9.5%	12.1%	4.0%	10.1%	12.9%
14	3.5%	8.8%	11.2%	3.7%	9.4%	12.0%
15	3.2%	8.2%	10.5%	3.5%	8.8%	11.2%

The comparable ELSA wave 6 report estimates for respiratory illness in males aged 50 to 54 is 10.2% (6.6% to 13.8%) and for females it is 10.7% (7.5% to 13.9%). These estimates are greater than those in Table 8-11 for 51 year olds at synthetic wave 6, especially for non-smokers. This implies that the Replenishers who join at synthetic wave 6 and subsequently may under estimate the true prevalence of this illness for these cohorts. The slight differential in favour of lower rates for males seen in the ELSA report is replicated in Table 8-11.

Table 8-11 : Estimated prevalence rates for respiratory illness in replenishing population

Male	non-smokers			smokers		
Synthetic wave	White	South Asian	Mixed/ Black/ Other	White	South Asian	Mixed/ Black/ Other
6	4.3%	7.9%	9.4%	8.7%	15.8%	18.6%
7	4.1%	7.6%	9.1%	8.4%	15.3%	18.0%
8	4.0%	7.4%	8.7%	8.1%	14.8%	17.4%
9	3.8%	7.1%	8.5%	7.8%	14.3%	16.8%
10	3.7%	6.9%	8.2%	7.6%	13.8%	16.3%
11	3.6%	6.6%	7.9%	7.3%	13.4%	15.8%
12	3.4%	6.4%	7.6%	7.1%	12.9%	15.2%
13	3.3%	6.2%	7.4%	6.8%	12.5%	14.7%
14	3.2%	6.0%	7.1%	6.6%	12.1%	14.3%
15	3.1%	5.8%	6.9%	6.4%	11.7%	13.8%

Female	non-smokers			smokers		
Synthetic wave	White	South Asian	Mixed/ Black/ Other	White	South Asian	Mixed/ Black/ Other
6	4.5%	8.4%	10.0%	9.3%	16.8%	19.7%
7	4.4%	8.1%	9.6%	8.9%	16.2%	19.1%
8	4.2%	7.9%	9.3%	8.6%	15.7%	18.4%
9	4.1%	7.6%	9.0%	8.3%	15.2%	17.9%
10	3.9%	7.3%	8.7%	8.1%	14.7%	17.3%
11	3.8%	7.1%	8.4%	7.8%	14.2%	16.7%
12	3.7%	6.8%	8.1%	7.5%	13.7%	16.2%
13	3.5%	6.6%	7.8%	7.3%	13.3%	15.7%
14	3.4%	6.4%	7.6%	7.0%	12.9%	15.1%
15	3.3%	6.2%	7.3%	6.8%	12.4%	14.7%

8.4.2 Smoking status at 50 or 51

As implied by the evidence presented in Box 5-10, few people begin smoking after the age of 30. Therefore someone's smoking status during the ages 30 to 49 is likely to remain unchanged when they reach 50 or 51. To examine the smoking status of this demographic (which is outside the ELSA age range) use is made of the 2011 HSfE. Here the derived variable of '*Cigarette Smoking Status - Never/Ex-reg/Ex-occ/Current*' is used to indicate smoking status, where 'Never' is taken to be never smoked whilst the remaining three categories are taken to be someone who has smoked. In the tabulations that follow, the counts are weighted using the weight provided for the analysis of the core interview sample (and hence the counts are not always integer). The cross tabulation count of gender, age and ethnicity for the smoking status variable is shown in Table 8-12 and also as percentages in Table 8-13.

Table 8-12 : Cross tabulation counts of smoking status by gender, age and ethnicity

		Ever smoked?									
		White		Mixed		Asian		Black		Other	
		No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Male	Age 30 to 39	247.7	322.3	6.2	5.3	44.5	37.7	10.7	8.4	4.4	5.4
	Age 40 to 49	305.7	379.0	3.1	6.6	38.7	26.2	26.8	15.6	6.7	3.6
Female	Age 30 to 39	274.6	313.0	6.6	5.0	60.7	8.8	26.3	7.6	12.6	2.8
	Age 40 to 49	338.4	363.7	7.8	5.1	53.1	6.3	19.7	8.6	2.2	0.8

Table 8-13 : Cross tabulation percentage of smoking status by gender, age and ethnicity

		Ever smoked?									
		White		Mixed		Asian		Black		Other	
		No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Male	Age 30 to 39	43%	57%	54%	46%	54%	46%	56%	44%	45%	55%
	Age 40 to 49	45%	55%	32%	68%	60%	40%	63%	37%	65%	35%
Female	Age 30 to 39	47%	53%	57%	43%	87%	13%	78%	22%	82%	18%
	Age 40 to 49	48%	52%	61%	39%	89%	11%	70%	30%	74%	26%

Clearly it is possible to use these percentages as thresholds within the Monte Carlo approach to determine if a Replenisher of a particular gender, age cohort and ethnicity combination has ever smoked. However, some of these probability estimates are based on very small samples, particularly the Mixed and Other groups. If these estimates are to be made more reliable then there is a need to merge some categories. A χ^2 test of each of the three dimensions in Table 8-12 shows that only for age is there no significant difference between its levels (Table 8-14).

Table 8-14 : Results of χ^2 test to identify significant factors

Factor	χ^2	dof	Sig
Age cohort	0.059	1	0.808
Gender	12.812^{***}	1	0.000
Ethnicity	87.756^{***}	4	0.000
Custom Groups	116.804^{***}	2	0.000

Recalculating the percentages in Table 8-13 without the age cohort, provides Table 8-15. In this table there is evidence for a potential cluster of three grouping in regards to smoking status. Those with the lowest rates of ever having smoked are Asian, Black and Other females (~80% never smoked). The next lowest is Asian, Black, Other males or

Mixed females (~60% never smoked) whilst the highest are White or Mixed males (~45% never smoked). If this customised grouping is adopted and a χ^2 conducted to test for the significance of this grouping, it is seen to be highly significant in Table 8-14.

Table 8-15 : Cross tabulation percentage of smoking status by gender and ethnicity

	Ever smoked?									
	White		Mixed		Asian		Black		Other	
	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Male	44%	56%	44%	56%	57%	43%	61%	39%	55%	45%
Female	48%	52%	59%	41%	88%	12%	74%	26%	81%	19%

The actual probabilities used are thus estimated based on these groupings and shown in Table 8-16. The advantage of these estimated probabilities over those in Table 8-15 is that they are more reliable, being based on aggregated larger samples that have retained a degree of homogeneity with regard to smoking status.

Table 8-16 : Custom Monte Carlo probabilities to determine the smoking status of Replenishers

Grouping	Never smoked	Smoked	Count
Asian, Black, Other Females	83%	17%	210
Asian, Black, Other Males or Mixed Females	58%	42%	253
White or Mixed Males	46%	54%	2,565

8.5 CONCLUSION

In this chapter the results of further preparatory work for the health microsimulation has been presented. The theme of the material is the estimation of the probabilities that determine how an individual's characteristics change or are defined through the modelling process.

The morbidity hazard models are critical to the operation of the health microsimulation since they model the incidence of morbidities over time which in turn defines the prevalence. Whilst initially the morbidity status of the population within each LAD will largely be defined by the surveyed status of individuals at ELSA wave 5, as the health microsimulation progresses to future synthetic waves the contribution of these survey individuals to the count for each morbidity will diminish. They will be augmented either by those in the same cohort who did not initially have the morbidity but the hazard model predicts they will subsequently acquire the morbidity or by those who replenish the population aged 50 to 51 and become a new ageing cohort. For CVD and DHBS there are significant time impacts, both in the chronological age and the wave trend. There is also evidence that females have lower probabilities of acquiring these two

morbidities. The south Asian ethnic group are more likely to acquire DHBS or a respiratory illness than the White group, but the highest probability for these two morbidities is the Mixed, Black or Other group. With CVD the ethnic grouping is not significant, but the lowest incidence is amongst the Black ethnic group, with a negative estimate, which supports the finding in the literature of the lowest prevalence for this group. Ever having smoked significantly increases the probability of CVD and respiratory illness but not DHBS (where the parameter is still positive, but not significant). The contextual type of LAD has a moderating impact on the probabilities. A general pattern is for those living in the more prosperous types of LAD to have a lower probability of acquiring a morbidity.

The second set of hazard models are used to define the selection weights that determine who is selected for cloning or removal in order to achieve a required population structure. Whilst it is not required that these estimates accord exactly with other similar sources of estimates, some insight is gained by carrying out this comparison. In particular it was found that the expected under counting of migration between LADs was indeed present and substantial, and an adaptation is required to better represent the probabilities of such moves.

The final piece of work reported defined the characteristics of those who replenish the populations at younger ages within each LAD. These characteristics are their morbidity status when they join the population and whether they have ever smoked. The morbidity status was estimated using the existing morbidity hazard models to accumulate their incidences over a 20 year period during which they aged from 30 to 49. In comparison with current prevalences reported in wave 6 of ELSA these prevalences were similar for males and females with CVD and for females with DHBS. For males with DHBS the accumulated hazard estimates were higher than those in the ELSA report and for respiratory illness they were lower for both males and females.

These three preceding results chapters have provided the basis to enable the modelling within the health microsimulation to take place. They have provided a spatially relevant and representative 2011 base population for each LAD; provided projections on how this population will evolve in terms of its gender, age and ethnic composition and here, provided mechanisms to dynamically change the characteristics of this population through time. The following chapter provides the results of implementing this modelling methodology to predict the number of individuals with each of the three morbidities in each LAD to 2031.

9 HEALTH MICROSIMULATION

Estimating Morbidity Outcomes for Local Authorities

9.1 INTRODUCTION

In this chapter the results of the model which forecasts the number of individuals within each English LAD with specific morbidities are presented. This is the fulfilment of the substantive research question of this thesis identified in section 1.4.

The methodology adopted is a hybrid approach where the population is studied at the individual or micro scale, modelling each individual's behaviour, but the composition of the population is adapted over time to an aggregate or macro structure. This is described in detail in Chapter 5 and in order to achieve these results there are a number of pre-requisites. Firstly, there is a need to construct a base population representing the LAD population in 2011. This is provided through the spatially microsimulated population which is described in Chapter 6. Secondly, there is the need to have a projection of the ethnic population in each LAD covering the forecasting period 2011 to 2031. The application of the mechanism to produce these projections is described in Chapter 7. What remains are models that update the morbidity status of individuals through time, creating new incidences of the morbidities and models that allow for the weighted selection of individuals to achieve a desired population structure. The models which provide these estimates are described in Chapter 8. This chapter presents the modelled outcomes, both at a national (section 9.3) and subnational level (sections 9.4 and 9.5), and in some detail for the seven case study LADs (section 9.6).

9.2 HEALTH MICROSIMULATION

The task of running the health microsimulation is done using a piece of bespoke C code. The C code reads in the spatially microsimulated population as output from the FMF for each LAD. It reads in the characteristics of the ELSA sample population to give these individuals their attributes, e.g. gender, age, ethnicity, presence of a morbidity and smoking status (here shadow individuals are also created, see section 5.5.4). The parameter values for the various hazard functions are also read in. In execution, the non-parallelised code takes about 30 minutes to produce the forecast prevalence counts biennially for the years 2011 to 2031. The output consists of gender, five year age band and ethnic group counts of those with and without each morbidity and additionally those with the comorbidity of CVD and DHBS. These prevalence counts may then be interpreted as the demand on local health care services associated with the morbidity or,

when divided by the population at risk, a prevalence rate to allow for a comparison over time and amongst LADs.

9.3 RESULTS FOR ENGLAND

The morbidity prevalence rates for the whole of England, calculated by the aggregation of the counts for each constituent LAD are shown in Table 9-1, Table 9-2, Table 9-4 and Table 9-5 for each biennial forecast year and each five year age band. The values in the cells in the entire table are rendered, with the intensity of the colour increasing as the prevalence increases, from white (low) to blue (CVD), red (DHBS), purple (respiratory illness) or grey (CVD and DHBS).

The general expectation in these tables is that the prevalence will increase with age reading across the tables. Individuals will acquire the morbidity at young ages and, if they survive, carry this morbidity to later ages. The trends reading down forecast years may however differ for each morbidity, once again driven by the ageing of the population. In each table there will exist cohort effects, where the progress of a group with a particularly high or low prevalence may be followed through the table. Such a cohort has been highlighted in the DHBS table. The group aged 50 to 54 during 2015 and 2017 has appreciably higher prevalences than the same aged group in 2013 or 2019. They will all have moved into the 55 to 59 year age band by 2018 to 2022, contributing this higher prevalence for the forecast periods up to 2019, 2021 and 2023. They will similarly be in the 60 to 64 age band for the forecast periods 2025 and 2017. This cohort of higher prevalences are shown as thick boxes in Table 9-2.

Table 9-1 : Forecast CVD prevalences for England

	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85 plus	Total
2011	2.9%	5.7%	8.1%	16.4%	20.6%	30.8%	39.8%	42.1%	18,229,893
2013	3.0%	4.0%	8.6%	12.9%	21.9%	25.9%	39.4%	41.2%	18,964,704
2015	2.4%	4.4%	7.7%	12.0%	20.1%	24.5%	35.5%	41.1%	19,689,593
2017	2.2%	4.5%	6.1%	11.0%	17.7%	23.8%	30.4%	40.8%	20,400,332
2019	1.9%	3.4%	6.0%	10.2%	15.0%	23.2%	27.3%	39.0%	21,083,429
2021	1.6%	3.2%	5.8%	9.9%	13.1%	21.5%	26.1%	35.7%	21,732,680
2023	1.4%	2.6%	5.2%	7.3%	12.6%	17.8%	26.0%	32.5%	22,304,919
2025	1.2%	2.3%	4.3%	7.3%	11.4%	16.3%	23.6%	30.7%	22,755,033
2027	1.1%	2.0%	3.8%	6.8%	9.4%	14.7%	20.7%	28.5%	23,129,460
2029	0.9%	1.7%	3.3%	5.5%	8.8%	13.4%	18.0%	26.7%	23,539,913
2031	0.8%	1.5%	2.9%	4.9%	8.4%	12.7%	16.0%	24.7%	24,029,232

The most dramatic reductions in morbidity to 2031 occur for CVD, with all age bands showing large reductions in the prevalence over the 20 year forecast period (Table 9-1). The youngest age bands show the greatest relative reduction, the rate amongst 50 to 54 years olds is nearly a quarter of that in 2011 whilst for the oldest age band it is nearly a

half of the 2011 prevalence by 2031. This dramatic reduction in the prevalence of CVD is in line with the recent trends in the reduction in standardised death ratios for CHD shown in Figure 3-3.

Table 9-2 : Forecast diabetes or high blood sugar prevalences for England

	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85 plus	Total
2011	7.0%	8.3%	9.7%	13.6%	15.5%	19.7%	14.6%	19.0%	18,229,893
2013	7.2%	9.0%	11.3%	14.0%	16.0%	17.9%	19.9%	19.0%	18,964,704
2015	8.4%	9.6%	11.7%	13.6%	17.2%	18.6%	22.1%	19.3%	19,689,593
2017	8.4%	9.7%	12.7%	13.9%	17.8%	18.3%	22.1%	20.8%	20,400,332
2019	7.3%	10.2%	13.2%	14.8%	17.0%	19.2%	22.2%	22.3%	21,083,429
2021	6.7%	11.4%	12.6%	15.6%	16.9%	20.0%	21.5%	23.0%	21,732,680
2023	6.1%	10.1%	12.8%	15.5%	17.9%	20.0%	21.2%	22.8%	22,304,919
2025	5.4%	9.3%	13.9%	15.4%	18.0%	19.3%	22.2%	23.5%	22,755,033
2027	4.9%	8.5%	13.5%	15.0%	18.9%	19.4%	22.4%	22.9%	23,129,460
2029	4.5%	7.8%	12.3%	15.3%	18.7%	20.1%	21.4%	23.4%	23,539,913
2031	4.0%	7.2%	11.5%	16.5%	17.6%	20.6%	21.2%	23.4%	24,029,232

For DHBS the pattern in less uniform (Table 9-2). For the very youngest age bands, aged 50 to 59, there are reductions in the prevalence rate when comparing the start and end of the forecast period. For the remaining age bands the prevalence increases by between 2% and 4% points from 2011 to 2031. These forecasts suggest that DHBS will become a morbidity much more concentrated in the middle-old and old-old population. An anomaly here is for the 80 to 84 age band where the 2011 prevalence forecast looks untypically low. There is a constant dip in the published DHBS prevalence for those aged 80 or older in the successive ELSA wave reports. Figure 9-1 charts these trends using published figures in table H3a of each ELSA wave report. There is an almost universal trend of a drop in the prevalence for this older age band for both genders and for all waves. (Incidentally, this figure also shows that prevalences for DHBS are increasing over time; wave 5 rates are higher than those in earlier waves). Elsewhere, using the 2008 to 2012 HSfE, Morrissey et al. (2014) also report a drop in prevalence rates for diabetes for males aged 75 or older (although there may be some endogeneity here since ELSA participants are recruited from previous respondents to the HSfE). There are a number of potential explanations for this phenomenon. Firstly, it may be an artefact of the way the information was collected in ELSA. Secondly, this outcome may be correct for the ELSA cohort (and by implication HSfE population) but anomalous for the population as a whole. The third explanation is the existence of a 'golden cohort' whose health outcomes, particularly for diabetes, are better than younger cohorts (Goldring et al, 2011).

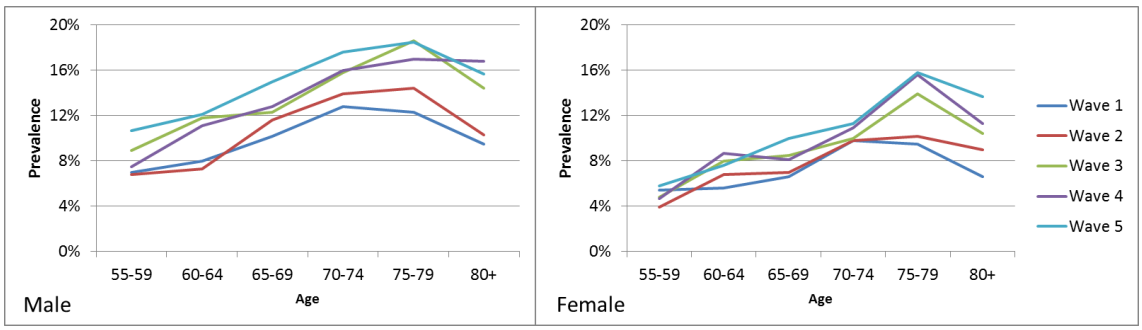


Figure 9-1 : Prevalence rates for diabetes or high blood sugar in each ELSA wave, male (lhs), female (rhs), England

(source : ELSA wave reports, 2003-2014)

To further explore this feature, using the ELSA wave 5 core data files this oldest age band can be split into two, those aged 80 to 84 and those aged 85 and older. These prevalences are provided in Table 9-3 for males. Here it is seen that the published 80 years and older rate of 15.7% is made up of a 15.1% rate for those aged 80 to 84 and a 16.3% rate for those aged 85 and older – the published reduction in prevalence is for the 80 and older population is almost all a result of the low prevalence at ages 80 to 84.

Table 9-3 : Male diabetes or high blood sugar prevalences from ELSA wave 5

Age band	Prevalence rate
Aged 52 to 54	6.9%
Aged 55 to 59	10.7%
Aged 60 to 64	12.0%
Aged 65 to 69	15.0%
Aged 70 to 74	17.6%
Aged 75 to 79	18.5%
Aged 80 to 84	15.1%
Aged 85 and older	16.3%
All	13.8%

The pattern of prevalences for respiratory illnesses is similar to that for DHBS, with decreases in prevalence over the period 2011 to 2031 for the age band 50 to 69; and increases for the age bands after. These reductions in the younger age bands are present here too. As the health microsimulation progresses these younger age bands will be composed almost entirely of Replenishers and it was noted in section 8.4.1 that the prevalence of these Replenishers for this morbidity was less than that found in the equivalent ELSA wave 6 prevalence table.

Table 9-4 : Forecast respiratory illness prevalences for England

	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85 plus	Total
2011	15.8%	14.1%	17.4%	20.3%	22.1%	21.6%	24.0%	23.3%	18,229,893
2013	9.5%	15.6%	16.5%	20.5%	23.9%	21.8%	23.3%	23.8%	18,964,704
2015	6.4%	16.9%	16.6%	19.9%	23.4%	22.8%	24.9%	24.2%	19,689,593
2017	5.1%	14.2%	17.7%	18.7%	23.0%	25.4%	24.4%	24.7%	20,400,332
2019	4.4%	10.0%	17.8%	19.9%	22.5%	25.5%	24.5%	24.6%	21,083,429
2021	4.1%	7.0%	20.0%	19.3%	22.6%	24.9%	25.8%	24.9%	21,732,680
2023	3.7%	6.2%	13.9%	20.6%	21.4%	25.0%	27.3%	24.4%	22,304,919
2025	3.3%	5.8%	10.3%	21.2%	21.4%	24.7%	26.6%	25.9%	22,755,033
2027	3.0%	5.4%	8.4%	18.4%	22.2%	23.2%	26.6%	26.9%	23,129,460
2029	2.8%	5.0%	7.8%	14.0%	22.1%	24.1%	26.0%	26.4%	23,539,913
2031	2.5%	4.7%	7.4%	10.4%	24.0%	23.5%	26.2%	26.8%	24,029,232

The prevalences for those who have both morbidities of CVD and DHBSs are shown in Table 9-5. As would be expected, these rates are composites of those for the same two morbidities in isolation. The anomalous low prevalence for DHBS for the 80 to 84 year old age band in 2011 is replicated here. For CVD the pattern within each age band was for reductions in prevalence for future years whilst for DHBS these reductions were only seen in the 50 to 59 year old age band, here the long term pattern of the CVD prevalence trends appears to dominate. However, there are some moderating influences. Firstly the reductions over the 20 year period are not so dramatic, particularly for the old-old and there appears to be a peak in prevalences during the forecast periods 2013 to 2021 for this old-old band. The rate of this comorbidity in the Replenisher population is low.

Table 9-5 : Forecast CVD and diabetes or high blood sugar prevalences for England

	50-54	55-59	60-64	65-69	70-74	75-79	80-84	85 plus	Total
2011	1.6%	1.8%	1.8%	4.3%	5.4%	9.1%	7.4%	9.9%	18,229,893
2013	0.7%	1.5%	2.2%	3.5%	6.0%	6.4%	11.1%	9.8%	18,964,704
2015	0.3%	1.8%	2.2%	2.9%	5.8%	6.6%	10.8%	10.2%	19,689,593
2017	0.3%	1.5%	2.0%	2.7%	5.1%	6.7%	9.0%	11.0%	20,400,332
2019	0.2%	0.5%	2.2%	3.0%	4.0%	6.5%	8.6%	11.3%	21,083,429
2021	0.2%	0.5%	2.1%	3.0%	3.4%	6.3%	8.0%	10.5%	21,732,680
2023	0.1%	0.4%	1.3%	2.3%	3.7%	5.3%	7.8%	9.6%	22,304,919
2025	0.1%	0.3%	0.8%	2.5%	3.5%	4.5%	7.5%	9.5%	22,755,033
2027	0.1%	0.2%	0.7%	2.0%	3.1%	4.1%	6.5%	8.8%	23,129,460
2029	0.1%	0.2%	0.5%	1.1%	3.0%	4.3%	5.2%	8.5%	23,539,913
2031	0.0%	0.2%	0.4%	1.1%	2.8%	4.1%	4.6%	7.9%	24,029,232

Whilst morbidity prevalences are informative when it comes to making comparisons between time periods and areas, in terms of health care planning it is the number of individuals with the morbidity and their location that is important. This spatial aspect to

both the prevalence and the number of individuals with morbidities is discussed in the next section but to finish this picture for England as a whole, Figure 9-2 shows the number of individuals with each morbidity in England and the total 50 years and older population from 2011 to 2031 and, for context, the equivalent prevalences are provided.

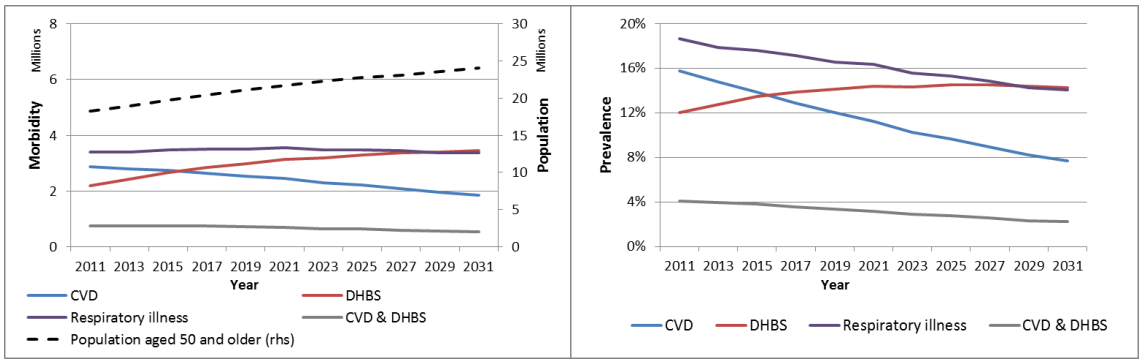


Figure 9-2 : Trends in the prevalence count (lhs) and rate (rhs) for morbidities, England, 2011 to 2031

There are three contrasting patterns here. With CVD and the comorbidity of CVD and DHBS both the prevalence counts and rates decrease. In contrast, the prevalence count for DHBS increases and this increase is greater than that seen in the general population, giving an increased prevalence rate. Finally the prevalence count for respiratory illness stays fairly constant, although given the rise in the population, this translates to a fall in the prevalence.

9.4 RESULTS FOR NHS LOCAL AREA TEAMS

Within England there are 25 NHS Local Area Teams (LATs) whose primary roles are to provide a strategic function that focuses on local priorities and also balances resources and strategies across a wide and coherent geographic area. Mostly these LATs are composed of a number of LADs, although some LADs do span more than one LAT. The best fit, by area, of these LADs to Local Area Teams is shown in Figure 14-2.

Table 9-6 gives the forecasts of the prevalence of CVD in each LAT across all ages. For each year (rather than the entire table) the cells are shaded so that for each forecast year the LAT with the highest prevalence has the darkest shade and the LAT with the lowest prevalence has the lightest shade. Intermediate prevalences will have a graduated shade. Using this shading, an LAD that has consistently high prevalences from 2011 to 2031 will feature as a dark band across the table.

Table 9-6 : Trends in CVD prevalence for NHS Local Area Teams

LAT	2011	2013	2015	2017	2019	2021	2023	2025	2027	2029	2031
Cheshire, Warrington & Wirral	15.6%	14.9%	14.2%	13.3%	12.6%	11.9%	11.0%	10.5%	9.8%	9.1%	8.6%
Durham, Darlington & Tees	16.3%	15.8%	15.4%	14.7%	14.0%	13.5%	12.5%	12.0%	11.4%	10.7%	10.1%
Greater Manchester	16.3%	15.0%	14.1%	12.9%	11.9%	11.1%	10.0%	9.5%	8.7%	8.0%	7.4%
Lancashire	16.5%	15.3%	14.3%	13.2%	12.2%	11.5%	10.4%	9.8%	9.1%	8.3%	7.8%
Merseyside	16.7%	15.9%	15.1%	14.2%	13.4%	12.8%	11.7%	11.1%	10.4%	9.6%	9.1%
Cumbria, Northumberland, Tyne & Wear	16.1%	15.4%	14.8%	13.9%	13.1%	12.5%	11.6%	11.0%	10.3%	9.6%	9.1%
North Yorkshire & Humber	15.7%	14.8%	14.0%	13.1%	12.3%	11.7%	10.8%	10.2%	9.5%	8.8%	8.3%
South Yorkshire & Bassetlaw	16.6%	15.4%	14.4%	13.3%	12.2%	11.4%	10.3%	9.7%	8.9%	8.2%	7.7%
West Yorkshire	15.9%	14.7%	13.7%	12.5%	11.5%	10.7%	9.6%	9.0%	8.2%	7.5%	7.0%
Arden, Herefordshire & Worcestershire	15.5%	14.8%	14.0%	13.0%	12.2%	11.5%	10.6%	10.1%	9.3%	8.7%	8.1%
Birmingham & the Black Country	17.2%	15.9%	14.8%	13.6%	12.4%	11.5%	10.3%	9.7%	8.9%	8.1%	7.5%
Derbyshire & Nottinghamshire	16.1%	15.0%	14.1%	13.1%	12.1%	11.4%	10.4%	9.8%	9.1%	8.4%	7.9%
East Anglia	15.7%	14.9%	14.1%	13.2%	12.4%	11.7%	10.7%	10.1%	9.5%	8.7%	8.2%
Essex	15.4%	14.6%	13.8%	12.8%	12.0%	11.3%	10.3%	9.8%	9.1%	8.4%	7.8%
Hertfordshire & the South Midlands	14.7%	13.8%	13.1%	12.2%	11.4%	10.7%	9.8%	9.2%	8.5%	7.9%	7.4%
Leicestershire & Lincolnshire	16.1%	15.0%	14.2%	13.2%	12.3%	11.6%	10.6%	10.1%	9.4%	8.7%	8.2%
Shropshire & Staffordshire	15.7%	15.0%	14.3%	13.4%	12.5%	11.9%	10.9%	10.4%	9.7%	9.0%	8.5%
London	15.4%	14.0%	12.7%	11.5%	10.4%	9.5%	8.3%	7.8%	7.0%	6.2%	5.7%
Bath, Gloucestershire, Swindon & Wiltshire	15.1%	14.3%	13.5%	12.6%	11.9%	11.2%	10.3%	9.7%	9.0%	8.4%	7.9%
Bristol, North Somerset, Somerset & South Gloucestershire	15.7%	14.7%	13.8%	12.8%	11.9%	11.2%	10.3%	9.7%	9.0%	8.3%	7.7%
Devon, Cornwall & Isles of Scilly	16.2%	15.3%	14.5%	13.6%	12.7%	12.0%	11.1%	10.5%	9.8%	9.1%	8.5%
Kent & Medway	15.5%	14.6%	13.8%	12.9%	12.1%	11.5%	10.5%	10.0%	9.3%	8.6%	8.0%
Surrey & Sussex	15.6%	14.8%	13.9%	13.0%	12.1%	11.4%	10.5%	9.9%	9.2%	8.5%	8.0%
Thames Valley	14.3%	13.6%	12.8%	12.0%	11.2%	10.5%	9.6%	9.2%	8.5%	7.9%	7.4%
Wessex	15.7%	14.8%	14.0%	13.1%	12.2%	11.5%	10.6%	10.1%	9.4%	8.7%	8.2%

Table 9-7 : Trends in diabetes or high blood sugar prevalence for NHS Local Area Teams

LAT	2011	2013	2015	2017	2019	2021	2023	2025	2027	2029	2031
Cheshire, Warrington & Wirral	11.5%	12.0%	12.5%	12.7%	12.8%	13.0%	12.8%	12.9%	12.9%	12.7%	12.7%
Durham, Darlington & Tees	12.1%	12.8%	13.5%	13.8%	14.1%	14.3%	14.3%	14.4%	14.4%	14.2%	14.1%
Greater Manchester	12.5%	13.5%	14.5%	15.0%	15.3%	15.7%	15.7%	16.0%	16.1%	16.0%	15.9%
Lancashire	12.2%	13.0%	13.7%	14.1%	14.3%	14.5%	14.4%	14.5%	14.6%	14.4%	14.3%
Merseyside	12.5%	13.0%	13.5%	13.8%	13.9%	14.1%	13.9%	14.0%	13.9%	13.7%	13.6%
Cumbria, Northumberland, Tyne & Wear	11.9%	12.4%	13.0%	13.2%	13.4%	13.5%	13.4%	13.5%	13.4%	13.2%	13.1%
North Yorkshire & Humber	11.4%	12.0%	12.6%	12.9%	13.1%	13.3%	13.1%	13.2%	13.2%	13.0%	12.9%
South Yorkshire & Bassetlaw	12.4%	13.2%	13.9%	14.3%	14.5%	14.8%	14.7%	14.8%	14.8%	14.7%	14.6%
West Yorkshire	12.4%	13.6%	14.6%	15.2%	15.6%	16.0%	16.0%	16.3%	16.4%	16.4%	16.3%
Arden, Herefordshire & Worcestershire	11.6%	12.5%	13.2%	13.6%	13.8%	14.1%	14.0%	14.2%	14.2%	14.1%	14.0%
Birmingham & the Black Country	13.8%	15.5%	16.9%	17.9%	18.5%	19.0%	19.2%	19.7%	19.9%	19.9%	19.8%
Derbyshire & Nottinghamshire	12.1%	13.0%	13.8%	14.3%	14.6%	14.9%	14.9%	15.1%	15.2%	15.1%	15.1%
East Anglia	11.4%	11.9%	12.4%	12.7%	12.8%	13.0%	12.9%	12.9%	12.9%	12.7%	12.6%
Essex	11.3%	11.9%	12.5%	12.8%	12.9%	13.1%	13.1%	13.2%	13.2%	13.1%	13.1%
Hertfordshire & the South Midlands	11.4%	12.1%	12.8%	13.1%	13.4%	13.6%	13.6%	13.8%	13.8%	13.7%	13.6%
Leicestershire & Lincolnshire	12.2%	12.9%	13.6%	13.9%	14.1%	14.4%	14.3%	14.5%	14.5%	14.3%	14.2%
Shropshire & Staffordshire	11.8%	12.4%	13.1%	13.5%	13.7%	13.9%	13.8%	13.9%	13.9%	13.8%	13.7%
London	13.7%	14.9%	15.9%	16.6%	17.0%	17.3%	17.2%	17.6%	17.7%	17.4%	17.2%
Bath, Gloucestershire, Swindon & Wiltshire	11.1%	11.7%	12.2%	12.4%	12.6%	12.8%	12.6%	12.7%	12.7%	12.6%	12.6%
Bristol, North Somerset, Somerset & South Gloucestershire	11.4%	11.9%	12.4%	12.6%	12.8%	12.9%	12.8%	12.8%	12.8%	12.6%	12.5%
Devon, Cornwall & Isles of Scilly	11.5%	12.0%	12.5%	12.8%	12.8%	13.0%	12.9%	13.0%	12.9%	12.7%	12.5%
Kent & Medway	11.4%	12.0%	12.6%	13.0%	13.2%	13.4%	13.3%	13.5%	13.5%	13.4%	13.3%
Surrey & Sussex	11.2%	11.8%	12.3%	12.6%	12.8%	12.9%	12.8%	13.0%	12.9%	12.8%	12.8%
Thames Valley	11.1%	11.7%	12.2%	12.5%	12.6%	12.8%	12.7%	12.9%	12.8%	12.7%	12.7%
Wessex	11.3%	11.8%	12.4%	12.6%	12.8%	13.0%	12.9%	13.0%	13.0%	12.9%	12.8%

All LATs show a reduction in the prevalence of CVD, echoing what was seen in Table 9-1, but the rate of reduction varies by LAT. The best performing LAT is London where the prevalence falls from one of the lowest in 2011 to a distinctly low prevalence, by some margin, by 2031 (its shading lightens across the table). The reason behind this good performance in London is attributable to the ethnic mix of the population, its younger age structure and the effect captured by the ONS area type of the London LADs (see Table 8-1 and Figure 14-2). Another LAT whose performance is good is Birmingham & the Black Country which moves from being the LAT with the highest prevalence in 2011 to a mid-table position in 2031, for much the same reasons as London, especially driven by its large Black ethnic population. The Durham, Darlington & Tees LAD shows the poorest improvement in the prevalence rate by 2031, its constituent LADs are mainly of the Industrial Hinterland type (see Figure 14-2). Here the impact attributable to the area type classification is probably driving this effect, where deprivation, lifestyle and the environment of the LADs that belong to this LAT are causing this adverse trend. Another LAT whose relative position worsens is the Thames Valley LAT, which moves from having the lowest prevalence in 2011 to having the fourth lowest in 2031.

Table 9-7 shows the prevalence rates for DHBS. Again, the trend seen in Table 9-2 for increased prevalence of this morbidity in England more generally is apparent here too. However, there appears to be less 'dynamism' in the changes in relative prevalence between the LATs. Birmingham and the Black Country remains the LAT with the highest prevalence rates and this position is further entrenched by 2031. This is in contrast to London which starts with a similar prevalence to Birmingham and the Black County in 2011 and whilst it remains one of the highest, it is has rates somewhat lower than Birmingham and the Black Country by 2031. Birmingham and the Black Country has a sizeable south Asian and Black ethnic communities and both these communities have higher prevalence rates for DHBS than the White ethnic group. Also the LADs within Birmingham and the Black Country are predominantly of the Centre with Industry type which has a positive effect on the probability of people living in such authorities to acquire DHBS (see Table 8-2). The London pattern is less clear. The same ethnic drivers as seen for Birmingham and the Black Country apply to London, i.e. a large BME population, but in London there are a mix of ONS LAD types, some of which increase the probability (London Centre) and some that decrease the probability (London Suburban and Cosmopolitan). The LATs that have consistently low prevalences for DHBS are mainly in the prosperous South of England

.

Table 9-8 : Trends in respiratory illness prevalence for NHS Local Area Teams

LAT	2011	2013	2015	2017	2019	2021	2023	2025	2027	2029	2031
Cheshire, Warrington & Wirral	18.7%	17.9%	17.6%	17.1%	16.5%	16.4%	15.6%	15.3%	14.9%	14.4%	14.2%
Durham, Darlington & Tees	19.6%	19.0%	18.9%	18.5%	18.0%	18.0%	17.2%	16.9%	16.6%	16.0%	15.7%
Greater Manchester	19.6%	18.6%	18.2%	17.6%	16.8%	16.6%	15.7%	15.4%	14.9%	14.2%	14.0%
Lancashire	19.3%	18.5%	18.3%	17.8%	17.2%	17.0%	16.2%	15.9%	15.4%	14.8%	14.6%
Merseyside	20.0%	19.2%	18.9%	18.4%	17.8%	17.8%	16.8%	16.5%	16.1%	15.5%	15.2%
Cumbria, Northumberland, Tyne & Wear	19.2%	18.6%	18.5%	18.2%	17.7%	17.7%	17.0%	16.8%	16.4%	15.8%	15.6%
North Yorkshire & Humber	18.4%	17.8%	17.6%	17.2%	16.7%	16.7%	16.0%	15.7%	15.3%	14.8%	14.5%
South Yorkshire & Bassetlaw	19.5%	18.7%	18.4%	17.8%	17.1%	16.9%	16.0%	15.7%	15.2%	14.6%	14.4%
West Yorkshire	19.3%	18.3%	18.0%	17.4%	16.8%	16.5%	15.6%	15.3%	14.8%	14.2%	13.9%
Arden, Herefordshire & Worcestershire	18.5%	17.8%	17.5%	17.0%	16.4%	16.2%	15.5%	15.2%	14.7%	14.2%	13.9%
Birmingham & the Black Country	20.2%	19.2%	19.0%	18.4%	17.7%	17.4%	16.5%	16.2%	15.7%	15.1%	14.9%
Derbyshire & Nottinghamshire	19.0%	18.3%	17.9%	17.4%	16.8%	16.6%	15.8%	15.5%	15.0%	14.5%	14.2%
East Anglia	18.2%	17.6%	17.4%	17.0%	16.5%	16.4%	15.7%	15.4%	15.0%	14.4%	14.2%
Essex	18.2%	17.5%	17.2%	16.8%	16.3%	16.2%	15.5%	15.2%	14.8%	14.3%	14.1%
Hertfordshire & the South Midlands	18.0%	17.0%	16.7%	16.2%	15.6%	15.4%	14.6%	14.3%	13.9%	13.4%	13.1%
Leicestershire & Lincolnshire	18.9%	18.1%	17.9%	17.5%	16.9%	16.8%	16.1%	15.8%	15.4%	14.8%	14.5%
Shropshire & Staffordshire	18.8%	18.2%	17.9%	17.5%	16.9%	16.8%	16.0%	15.7%	15.3%	14.7%	14.5%
London	18.8%	17.3%	16.6%	15.8%	14.8%	14.4%	13.3%	13.0%	12.4%	11.7%	11.3%
Bath, Gloucestershire, Swindon & Wiltshire	17.9%	17.1%	16.7%	16.2%	15.6%	15.4%	14.7%	14.3%	13.9%	13.4%	13.2%
Bristol, North Somerset, Somerset & South Gloucestershire	18.2%	17.5%	17.1%	16.6%	16.1%	16.0%	15.2%	14.9%	14.5%	14.0%	13.7%
Devon, Cornwall & Isles of Scilly	18.5%	18.5%	18.6%	18.6%	18.4%	18.5%	18.0%	17.9%	17.6%	17.2%	17.0%
Kent & Medway	18.3%	17.7%	17.5%	17.1%	16.6%	16.6%	15.8%	15.6%	15.2%	14.7%	14.5%
Surrey & Sussex	17.9%	17.4%	17.3%	17.0%	16.6%	16.6%	15.9%	15.7%	15.4%	15.0%	14.9%
Thames Valley	17.6%	17.0%	16.8%	16.4%	16.0%	15.9%	15.3%	15.1%	14.8%	14.3%	14.2%
Wessex	18.1%	17.5%	17.4%	17.1%	16.6%	16.6%	15.9%	15.7%	15.4%	14.9%	14.7%

Table 9-9 : Trends in CVD and diabetes or high blood sugar prevalence for NHS Local Area Teams

	2011	2013	2015	2017	2019	2021	2023	2025	2027	2029	2031
Cheshire, Warrington & Wirral	3.8%	3.7%	3.6%	3.4%	3.2%	3.1%	2.9%	2.7%	2.5%	2.3%	2.2%
Durham, Darlington & Tees	4.2%	4.0%	4.0%	3.9%	3.7%	3.6%	3.4%	3.2%	3.1%	2.9%	2.7%
Greater Manchester	4.4%	4.2%	4.1%	3.8%	3.5%	3.4%	3.0%	2.9%	2.7%	2.4%	2.3%
Lancashire	4.3%	4.1%	4.0%	3.7%	3.5%	3.3%	3.0%	2.8%	2.6%	2.4%	2.3%
Merseyside	4.3%	4.1%	4.0%	3.8%	3.6%	3.5%	3.1%	3.0%	2.8%	2.5%	2.4%
Cumbria, Northumberland, Tyne & Wear	4.1%	3.9%	3.8%	3.6%	3.4%	3.3%	3.0%	2.9%	2.7%	2.5%	2.3%
North Yorkshire & Humber	3.8%	3.7%	3.6%	3.4%	3.2%	3.0%	2.8%	2.6%	2.4%	2.2%	2.1%
South Yorkshire & Bassetlaw	4.3%	4.1%	4.0%	3.7%	3.5%	3.3%	3.0%	2.8%	2.6%	2.4%	2.2%
West Yorkshire	4.3%	4.1%	4.0%	3.7%	3.5%	3.3%	2.9%	2.8%	2.6%	2.4%	2.2%
Arden, Herefordshire & Worcestershire	3.9%	3.8%	3.7%	3.5%	3.3%	3.1%	2.9%	2.8%	2.6%	2.4%	2.2%
Birmingham & the Black Country	5.1%	4.9%	4.8%	4.5%	4.1%	3.9%	3.5%	3.4%	3.1%	2.9%	2.7%
Derbyshire & Nottinghamshire	4.2%	4.0%	3.8%	3.6%	3.4%	3.2%	2.9%	2.8%	2.6%	2.4%	2.3%
East Anglia	3.8%	3.6%	3.6%	3.4%	3.2%	3.0%	2.8%	2.6%	2.4%	2.2%	2.1%
Essex	3.7%	3.6%	3.5%	3.3%	3.1%	2.9%	2.7%	2.6%	2.4%	2.2%	2.1%
Hertfordshire & the South Midlands	3.8%	3.6%	3.5%	3.3%	3.1%	3.0%	2.7%	2.6%	2.4%	2.2%	2.1%
Leicestershire & Lincolnshire	4.3%	4.1%	4.0%	3.7%	3.5%	3.3%	3.0%	2.9%	2.7%	2.5%	2.4%
Shropshire & Staffordshire	4.0%	3.9%	3.8%	3.6%	3.4%	3.2%	3.0%	2.8%	2.6%	2.4%	2.3%
London	4.9%	4.5%	4.3%	4.0%	3.6%	3.4%	3.0%	2.9%	2.6%	2.3%	2.1%
Bath, Gloucestershire, Swindon & Wiltshire	3.7%	3.5%	3.4%	3.2%	3.1%	2.9%	2.7%	2.5%	2.4%	2.2%	2.1%
Bristol, North Somerset, Somerset & South Gloucestershire	3.8%	3.6%	3.5%	3.3%	3.1%	2.9%	2.7%	2.5%	2.3%	2.1%	2.0%
Devon, Cornwall & Isles of Scilly	3.9%	3.7%	3.6%	3.4%	3.2%	3.1%	2.9%	2.7%	2.5%	2.3%	2.1%
Kent & Medway	3.8%	3.7%	3.6%	3.4%	3.2%	3.1%	2.8%	2.7%	2.5%	2.3%	2.2%
Surrey & Sussex	3.8%	3.7%	3.6%	3.4%	3.2%	3.0%	2.8%	2.7%	2.5%	2.3%	2.2%
Thames Valley	3.6%	3.5%	3.4%	3.2%	3.0%	2.9%	2.6%	2.5%	2.4%	2.2%	2.1%
Wessex	3.8%	3.6%	3.5%	3.3%	3.1%	3.0%	2.7%	2.6%	2.4%	2.3%	2.2%

For the prevalences of respiratory disease in Table 9-8 the general trend for reductions in the illness is apparent, although to varying degrees amongst LATs. The London LAT shows a large reduction in the prevalence, from a mid-table position in 2011 to the LAT with the lowest prevalence in 2031. In Table 8-3 the hazard model parameter estimates shows that the incidence of respiratory illness amongst the BME community is greater than the White ethnic group, however this effect appears not to feed through to the prevalences here (as it did for DHBS). However, there are two other important influences at work for this morbidity. Firstly, smoking is a significant influence on the morbidity, but the smoking rates from the HSfE in Table 8-13 show lower smoking rates amongst the BME population. This implies that the smoking influence will have less impact on this population, particular those who are female. Secondly, the area type parameter estimates for the London type authorities predicts that there is a lower probability of acquiring a respiratory illness amongst the inhabitants of such authority types, all other things being equal. The converse trend to London appears to be taking place for the Devon, Cornwall and Isles of Scilly LAT. This LAT sees only a modest reduction in the prevalence. The ethnic population of these areas is predominately White and thus will have the higher smoking rates shown in Table 8-13, which will feed through to a higher prevalence. Also 8 of the 11 LADs within this LAT are Coastal and Countryside ONS area types and the parameter estimate for this type of authority is positive which means that people in such authorities have a higher incidence of respiratory illnesses. No other LAT has such a large concentration of Coastal and Countryside type authorities within its borders. Such Coastal and Countryside LADs are popular retirement destinations for the mobile elderly who may have lived and worked in more urban type LADs and on retirement moved to these less urban LADs (Rural Evidence Research Centre, 2005 and Raymer et al., 2007). A consequence of this is that their early life lifestyle behaviours are transplanted to these more rural LADs. The LAD area type estimate is capturing some of this contextual behaviour.

The prevalence for the comorbidity of CVD and DHBS is shown in Table 9-9. Whilst there is no explicit hazard model for this comorbidity in the modelling framework, it is clearly shaped by the pattern in Table 9-6 and Table 9-7. The Birmingham and the Black Country LAT has a consistently high prevalence for this comorbidity. For CVD this LAT has the highest prevalence in 2011 but by 2031 its prevalence is mid table. However, for DHBS the prevalence remains high for this LAT. London starts with a high prevalence for this comorbidity but by 2031 it has moved into the third of LATs with the lowest prevalence. Whilst the prevalence of DHBS remains high in London, the prevalence for CVD reduces considerably in London. This low prevalence for CVD accounts for the different trend in London. The position of Durham, Darlington & Tees

deteriorates by 2031, and this appears to be driven by its trend to relatively high CVD rates by 2031, rather than DHBS.

Recent trends in reported prevalences for a range of morbidities arising from the NHS Quality and Outcomes Framework (QOF) show some corroborative year on year changes in these morbidities (HSCIC, 2014b). Data are published at national, LATs, CCG level and GP practice level. Looking at the changes in the reported prevalence at LAT level of CVD related morbidities, for CHD the average yearly reduction is 1.39% whilst for strokes there is a 0.93% increase. The picture is similarly mixed for respiratory illnesses, with asthma showing an average yearly 1.16% reduction but COPD a much higher 2.47% increase. For diabetes there is a large yearly 3.14% average increase in its QOF prevalence. These are recent but very short term trends, 2012-13 to 2013-14, but if these trends were to persist then the longer term reduction in CVD prevalence and increase in diabetes prevalence seen here look plausible.

9.5 RESULTS FOR LOCAL AUTHORITIES

Whilst the tables in the previous sections provide a picture of trends by age and year for the whole of England and regionally, there will be local variations to these national trends. To illustrate the local picture a series of maps of English LADs are provided for the 2031 forecasts (as was done for 2011 in Figure 6-2). Durham in the North East of England has both a high prevalence rate of CVD (it is coloured red) and a high prevalence count (the circle is large). However, looking just south of Durham to Leeds a different picture emerges. Here the circle is almost as large as Durham's but the prevalence is much lower. This is due to the greater size of Leeds' population. Thus Leeds has a lower prevalence rate of the morbidity but in terms of treatments and health care resources for CVD, its need is as almost as large as Durham's. Copeland on the coast of the North West has a high prevalence (red) like Durham but low number of individuals (small circle) so its requirements for health care resources is less than Durham's even though it has a similar prevalence rate.

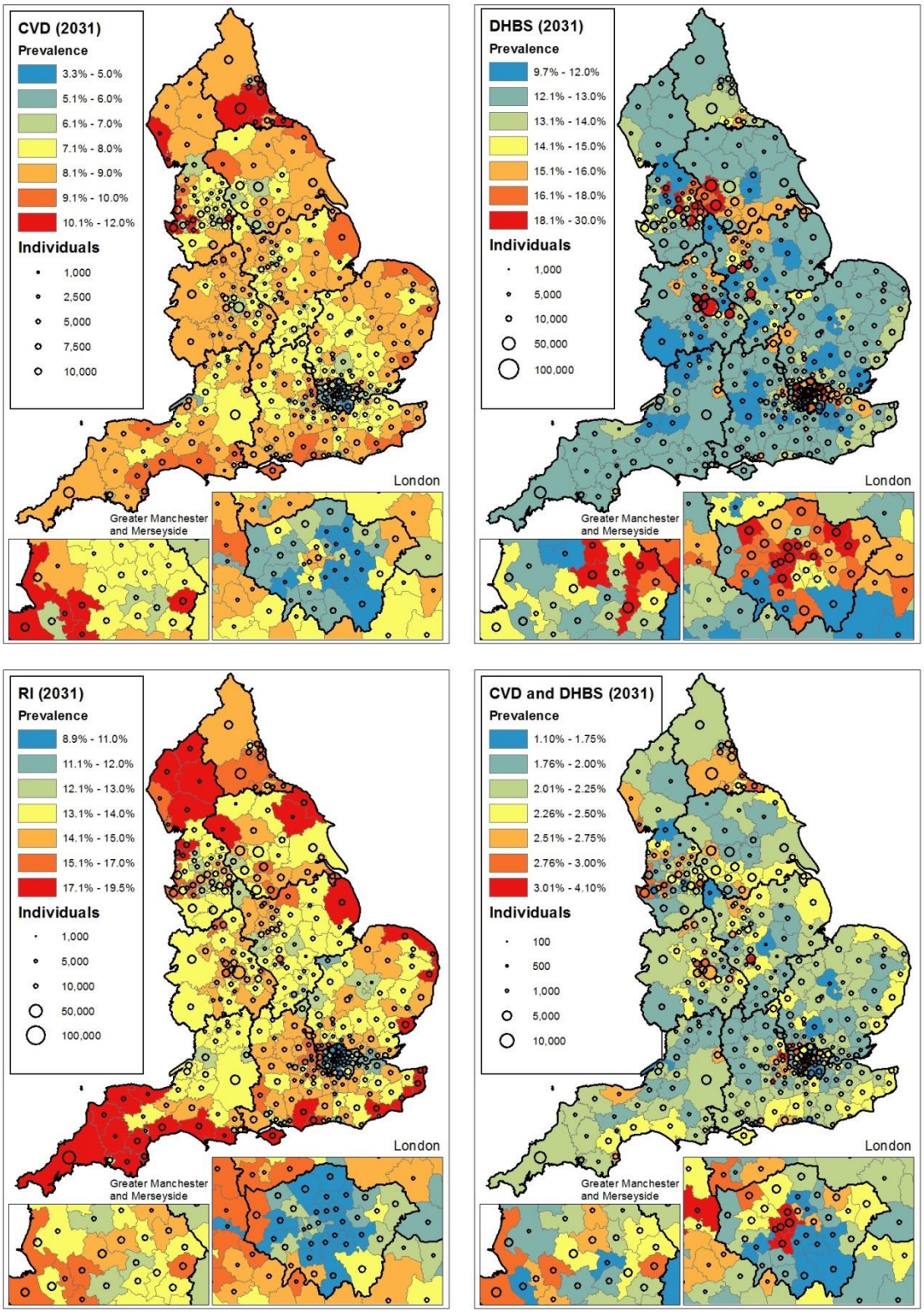


Figure 9-3 : Prevalences and number of individuals with morbidities, LADs, 2031, from top left, CVD, top right, diabetes or high blood sugar, bottom left, respiratory illness and botton right CVD and diabetes or high blood sugar

One distinct pattern in the DHBS map is high prevalence rates in LADs located in West Yorkshire, Greater Manchester, the West Midlands and London plus Derby, Nottingham and Leicester. These are all authorities with large BME communities in 2011 and the size of 50 and older population in these communities is projected to increase by 2031. The high prevalences for respiratory illnesses appear to be concentrated around the coast and in large rural LADs in the north of England (particularly the North West). However in terms of individuals with the morbidity these LADs have small numbers (with the exception of Cornwall which has both a high prevalence count and rate). Hot spots for the comorbidity of CVD and DHBS appear to be confined to LADs in and around London, and Leicester and Oadby and Wigston (which is located just south of the City of Leicester).

9.6 RESULTS FOR CASE STUDY AUTHORITIES

In this section trends in the prevalence rates of each morbidity within the seven case study authorities are illustrated and described. These trends are presented in terms of both the overall prevalence within the LADs and their ranking for the morbidity amongst all LADs (with a low ranking indicating a low prevalence).

9.6.1 Cardio Vascular Disease

The trend charts for CVD are shown in Figure 9-4. The trend in all LADs is for a decrease in the prevalence rate of CVD, although this rate of decline varies by LAD. Haringey has much the lower prevalence of all the case study LADs in 2011 and this difference is perpetuated and increased by 2031. The population of this LAD in 2011 has a sizeable black and Afro Caribbean population (see Figure 7-5) which is beginning to age into the 50 year and older age band, and by 2031 (see Figure 7-18) this ethnic group is sizeable. Both the literature and the hazard models presented in Table 8-1 shows that this ethnic group has a lower than average prevalence of this morbidity in comparison to the remaining ethnic groups. Also in Table 8-1, Haringey is a London Cosmopolitan type of LAD and people living in these types of authority have a lower probability of acquiring CVD, relative to most other area types. In terms of trends, Haringey remains one of the LADs with the lowest CVD prevalence. Also in the trend chart, the two LADs that show the greatest improvement in their relative prevalence rates for CVD by 2031 are Leeds and Redbridge.

The ethnic mix of Redbridge has a large south Asian population and the expectation is that this should actually increase the probability of acquiring CVD (See Table 8-1) relative to the remaining ethnic groups. However there is also the lifestyle element of smoking behaviour (as discussed in 3.3.1.1), with this south Asian group having lower rates of smoking (see Table 8-13) and therefore the contribution to the probability from

smoking is reduced in this population. Redbridge is also a London Suburban type of authority and this type of authority has one of the largest negative parameters in Table 8-1. The third element contributing the improved performance of this LAD is that it is located in London and as seen in Figure 2-5, retains its youthful 50 years and older population, ageing from 65.5 years of age in 2011 to just 66.1 years in 2031 and therefore the age contribution to the probability of acquiring CVD will be relatively small for this LAD. Some of these same effects are seen to influence the relatively good performance of Leeds, a diverse, Cities and Services type of LAD with a young population, although additionally, Leeds does retain a sizeable Black ethnic group, which has a lower incidence of CVD than the White and Asian groups.

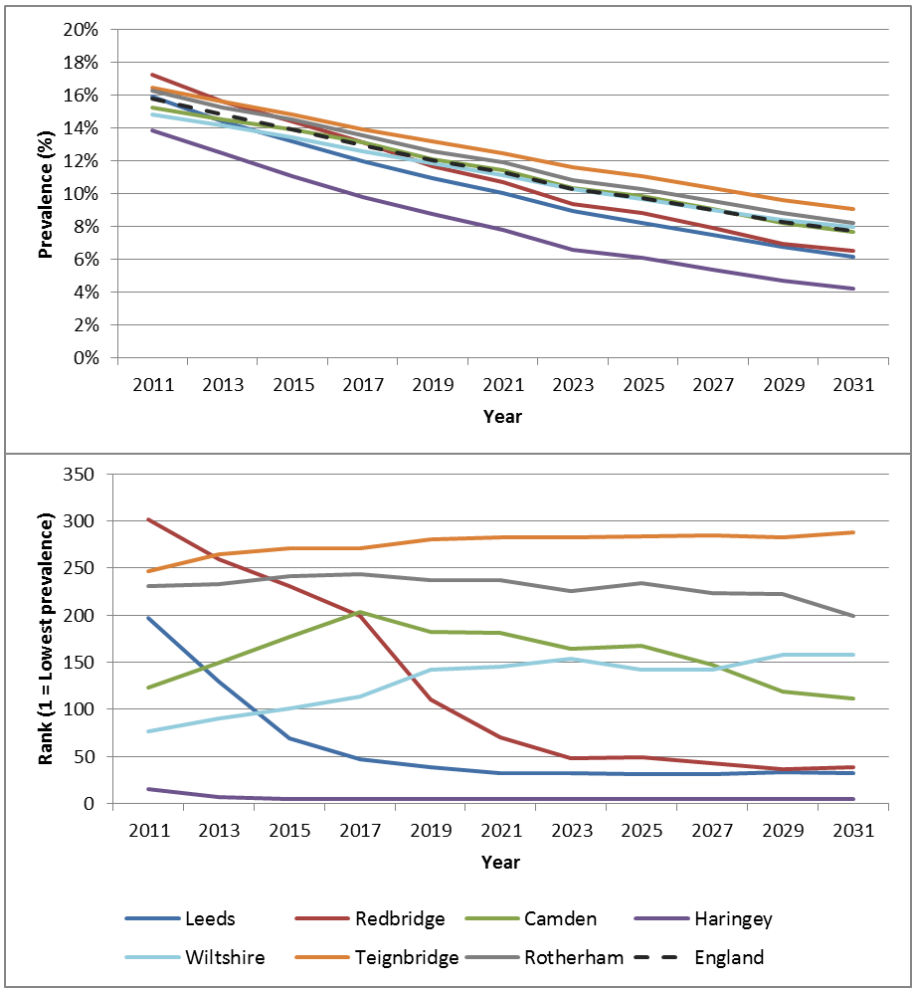


Figure 9-4 : Prevalence rates (top) and rankings (bottom) for CVD, case study LADs, 2011 to 2031

Over the 20 year time span the case study authority whose relative position deteriorates the most is Wiltshire. The ethnic mix of this LAD remains almost homogeneously White over the 20 year time period, but does, like other large rural LADs, age from 66.3 years of age in 2011 to 69.6 in 2031. The ONS area type of Wiltshire is Prospering

Smaller Towns which has a modest and insignificant negative impact on the probability of acquiring CVD relative to Manufacturing Towns.

9.6.2 Diabetes or high blood sugar

Looking at the trend charts in Figure 9-5 the general pattern is for increases in the prevalence of DHBS, the exception being Haringey which shows a slight reduction. For the ranks, the relative position for most of the LADs does not change much, except, again, for Haringey and possibly Leeds. The parameter estimates in Table 8-2, show that smoking is not a significant influence on acquiring DHBS making the differential smoking rates by ethnicity less of a factor here than for CVD.

The main driver for the stark increase in the prevalence for Camden is its ONS area classification of London Centre. This type of authority is one of only two that has a positive contribution to the probability of acquiring the morbidity relative to Manufacturing Towns, and it has the largest such contribution. In terms of ethnicity, Figure 7-17 would indicate that there is a sizeable cohort from the Black ethnic group ageing into the 50 years and older population, and this ethnic group has the largest positive influence on acquiring DHBS in Table 8-2. Camden ages only modestly, from 65.5 years of age in 2011 to 66.3 years in 2031.

Nearby Haringey is a London Cosmopolitan type of authority and so has the second most negative influence on the acquisition of this morbidity. It does however have a sizeable BME population and this would tend to increase the probability of acquiring the morbidity. As seen earlier, and like most London LADs, the population of Haringey remains young, making the contribution of the age related element of the hazard model minor.

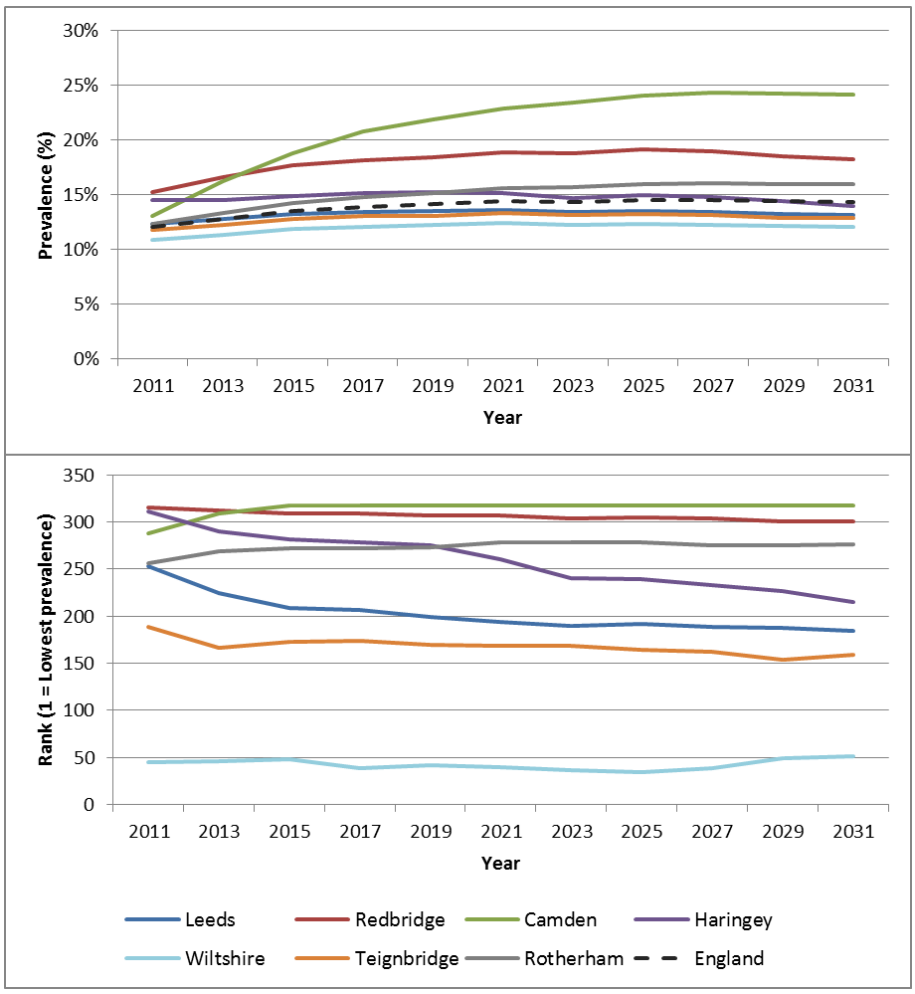


Figure 9-5 : Prevalence rates (top) and rankings (bottom) for diabetes or high blood sugar, case study LADs, 2011 to 2031

9.6.3 Respiratory Illness

All the case study authorities show a reduction in the prevalence of respiratory illnesses between 2011 and 2031 (even Teignbridge which shows a rather modest -0.1% reduction). The greatest relative reductions are in the London authorities of Redbridge, Camden and Haringey.

All the ONS area types specifically connected to London LADs have large negative parameter estimates in Table 8-3. The LAD estimate for London Suburban type LADs is particularly negative and large. This will drive the reductions in the incidence of respiratory illness in these types of authority. Here the role of ethnicity will again be confounded by the impact of smoking; whilst the BME ethnic group have a higher incidence of respiratory illness, they have lower rates of smoking and, unsurprisingly, smoking is an important influence on respiratory illness. Teignbridge is an authority with a predominately White ethnic population and as such will, following the probabilities in Table 8-16, have the highest smoking rates both in the original ELSA sample and the Replenishers. Also, the Coastal and Countryside area type has the most

positive estimate of all area types in Table 8-3 and therefore will increase the incidence of respiratory illness relative to LADs of other types.

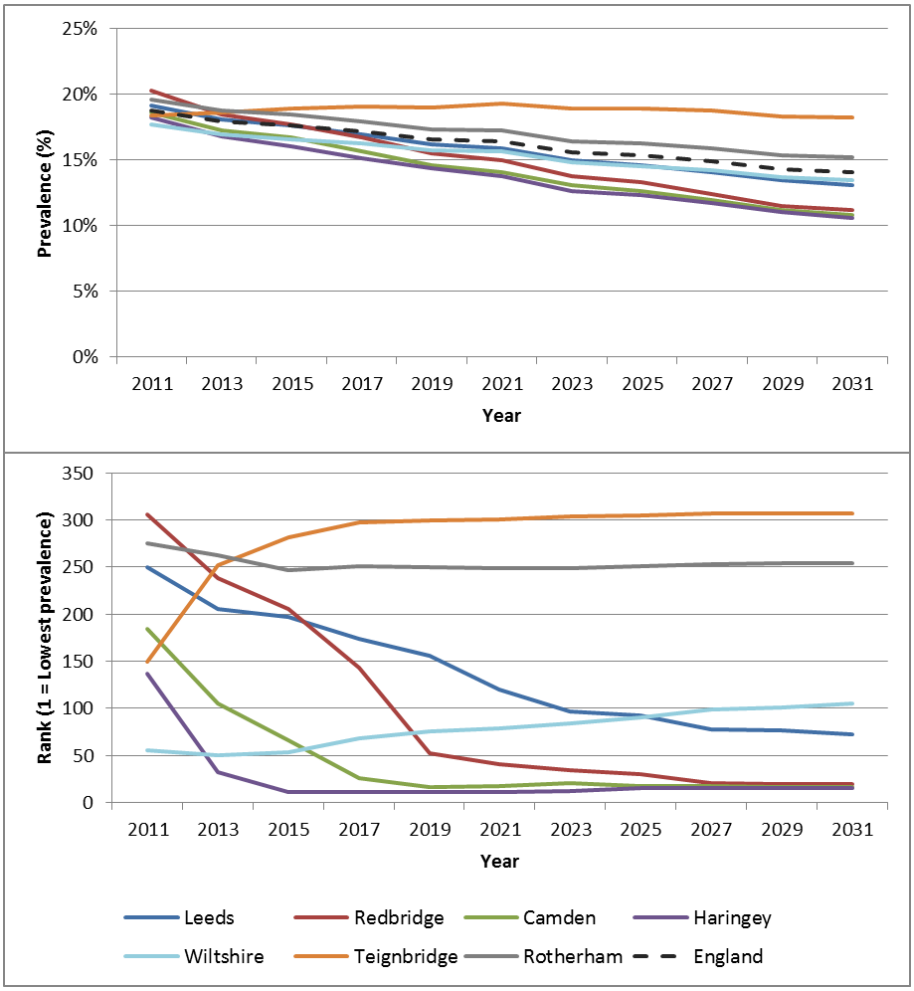


Figure 9-6 : Prevalence rates (top) and rankings (bottom) for respiratory illnesses, case study LADs, 2011 to 2031

9.6.4 CVD and diabetes or high blood sugar

As for the NHS LAT forecasts in Table 9-9 the interpretation of the trends for this comorbidity are influenced by the combined effects of the individual morbidities. All case study authorities show a reduction in the prevalence of this comorbidity. The reductions are the greatest, and almost equal, in Redbridge and in Haringey, although Redbridge starts with a much higher prevalence in 2011. The dynamic that drives the low comorbidity prevalence for Haringey is that the prevalence rates for both CVD and DHBS in Haringey fall over the 20 year period, more so for CVD than DHBS. This naturally reduces the rate for the comorbidity. With Leeds there is also a reduction in the CVD prevalence rate, although the DHBS rate increases slightly, but this combination still produces a reduction in the comorbidity prevalence rate, but not to the

same extent as seen in Haringey. The comorbidity rate for Camden is kept high by the increase in the prevalence of DHBS even in the context of a reduction in its CVD rate.

Looking at the relative performance, both Haringey and Leeds improve their position somewhat over the 20 year period, largely due to their relative improvements in CVD prevalence seen in Figure 9-4. Whilst Redbridge also improves its relative position for CVD in Figure 9-4, its relative position on DHBS in Figure 9-5 stays poor, which limits the potential for similar improvements for the comorbidity.

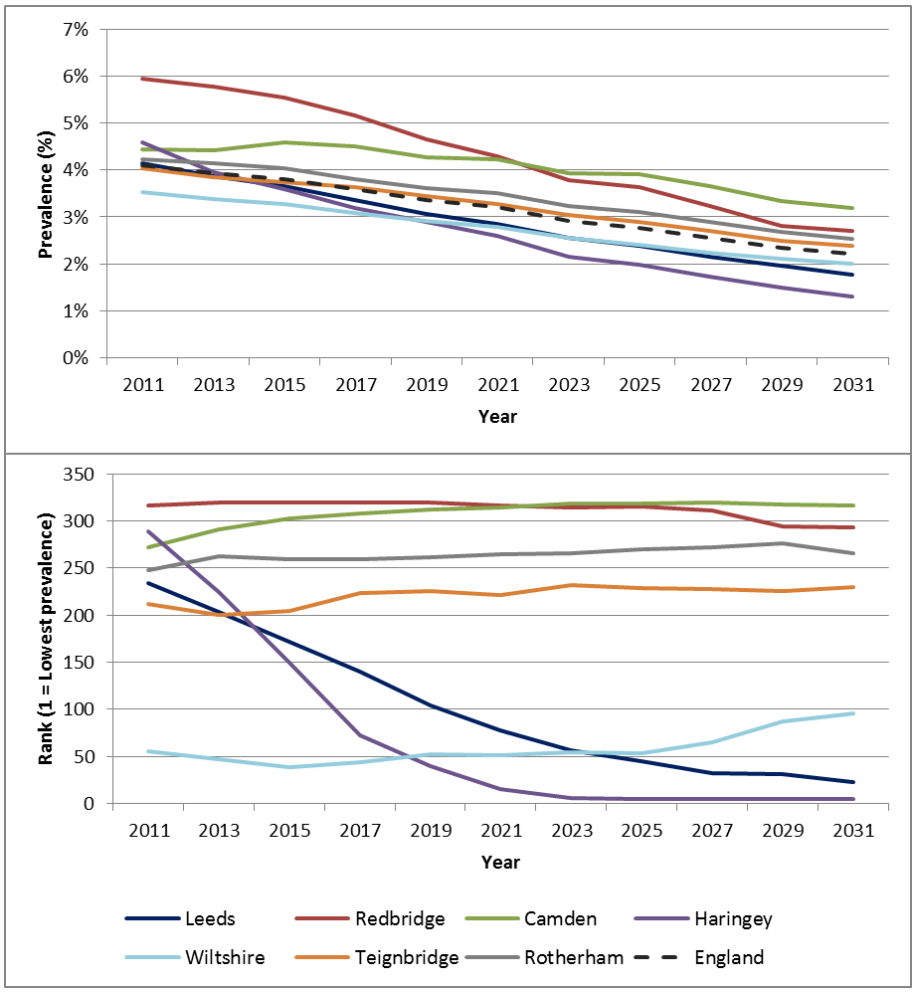


Figure 9-7 : Prevalence rates (top) and rankings (bottom) for comorbidity of CVD and diabetes or high blood sugar, case study LADs, 2011 to 2031

9.7 ROBUSTNESS OF HEALTH MICROSIMULATION

The approach adopted in the health microsimulation is stochastic in its use of pseudo random numbers for Monte Carlo draws. As with the spatial microsimulation in section 6.6, some assessment of how sensitive the model outputs are to the sequence of random numbers is required. This is achieved in a similar manner here by performing two additional model runs where the initial random seed for the random number generator is varied. The measure used to assess the correspondence between the outputs amongst all

three model runs is the prevalence rate of the morbidity in the LAD for the final forecast year of 2031. This is illustrated visually by a scatter plot of these prevalences and the points should ideally lie on a straight line with slope 1.0 and passing through the origin with minimal scatter around the straight line. These scatter plots are shown in Figure 9-8 for three health microsimulation runs with different sequences of random numbers along with the linear correlation coefficient r^2, whose value should ideally be close to 1.0.

As seen earlier in Figure 6-3 there is some degree of scatter in the plots (although to a lesser degree than seen in Figure 6-3) but the correspondence between the forecasts is good which means that the outputs from the health microsimulation appear to be insensitive to stochastic nature of the microsimulation.

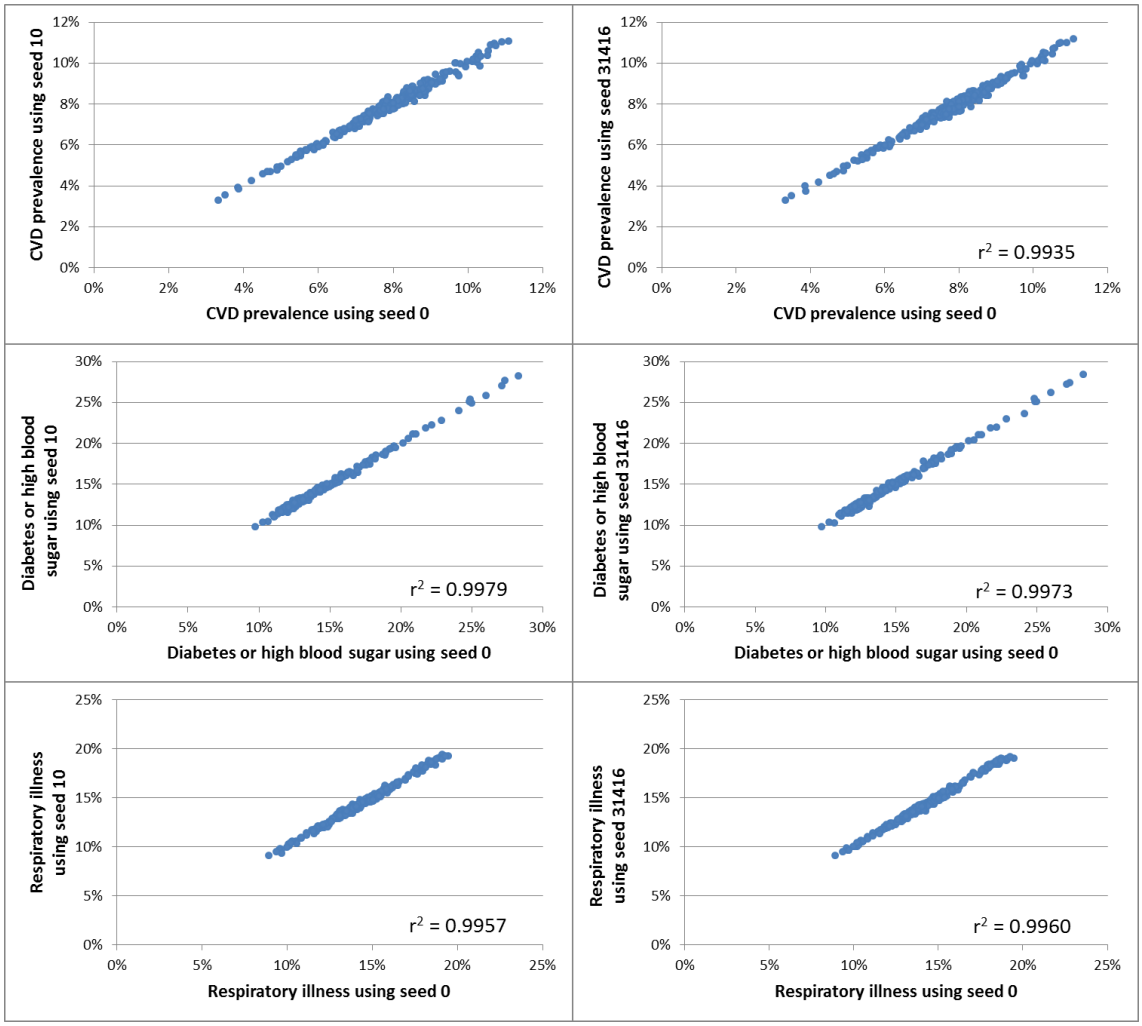


Figure 9-8 : Comparison of LAD prevalences in 2031 from three health microsimulations, top : CVD, middle : diabetes or high blood sugar, bottom : respiratory illness

9.8 COMPARISON WITH OTHER FORECASTS

There is the potential for some insight to be gained by comparing the forecasts derived here with alternative forecasts. Such comparative forecasts are, however, difficult to obtain. The POPPI project (Poppi User Guide, 2011) produces a series of updated forecasts for English LADs of a range of socio demographic and economic factors, including prevalence rates for a number of morbidities. The three relevant morbidities for this study are the number of people who have experienced a heart attack, a stroke and the number who have diabetes. The POPPI model uses gender and age band specific rates for morbidity prevalences and applies these to the latest ONS subnational population projections to arrive at the number of people with the morbidity. This can be converted back to a prevalence rate for the whole population by dividing by the size of the population. The current gender and age rates used by POPPI are shown in Table 9-10a to Table 9-10c.

Table 9-10a : POPPI stroke prevalence rates, Great Britain, 2007

Age range	Males	Females
65-74	2.8%	1.2%
75+	3.8%	1.9%

(source : POPPI, 2011)

Table 9-10b : POPPI heart attack prevalence rates, Great Britain, 2007 and ELSA CHD rates, England, wave 5

	POPPI		ELSA	
Age range	Males	Females	Males	Females
65-69	6.5%	2.6%	19.9%	10.6%
70-74			21.0%	14.5%
75-79	5.7%	5.1%	32.0%	21.4%
80+			34.2%	25.7%

(source : POPPI, 2011, and Institute for Fiscal Studies, 2012)

Table 9-10c : Diabetes prevalence rates from POPPI, 2006, and ELSA, England, wave 5

Age range	POPPI		ELSA	
	Males	Females	Males	Females
65-69	15.7%	10.4%	15.0%	10.0%
70-74			17.6%	11.3%
75-79	13.5%	10.6%	18.5%	15.8%
80+			15.7%	13.7%

(source : POPPI, 2011, and Institute for Fiscal Studies, 2012)

To ensure compatibility between the forecasts, the health microsimulation prevalence rates are re-calculated on the 65 and older population as opposed to the usual 50 and older population. There remain however some compatibility issues. Firstly CVD includes a diagnosis of angina but POPPI does not have a prevalence rate for this

morbidity. This will lead POPPI to under estimate the morbidity of CVD. Secondly, if the counts of individuals with a stroke or heart attack diagnosis from POPPI are added together, there will be a double counting of those who have had both a stroke and heart attack. In the health microsimulation, they will only be counted once. This will lead to an overestimate by POPPI. Finally POPPI rates are for diabetes alone and not DHBS, and this will lead to an under estimate by POPPI. The scatter plot of the health microsimulation forecasts and the equivalent POPPI rates for English LADs in 2031 are shown in Figure 9-9.

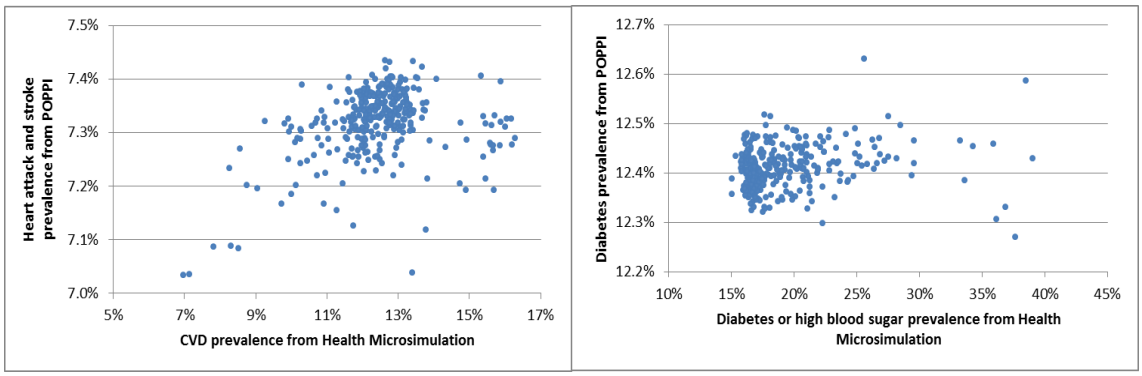


Figure 9-9 : Comparison of health microsimulation and POPPI prevalence rates, CVD (lhs) and Diabetes (rhs), LADs, 2031

The agreement between the two prevalences is best for CVD. Except for a sizeable cluster to the right hand side of the plot, there appears to be a positive relationship between the two forecasts. However, the level of the POPPI forecasts of strokes or heart attacks is much lower and has a more narrow a range than for CVD from the health microsimulation. The correspondence between the two forecasts for diabetes is poor, with only a slight hint of a positive relationship. Again the level and range of POPPI forecasts are lower. Mason et al. (2015) conclude that the use of simple age based models, like POPPI, will tend to under estimate the level of morbidity in a population, a result not found here. There are a number of explanations as to why this is so. Firstly, this is not a like for like comparison, as discussed above. Secondly, the cross sectional (and historic) POPPI rates for the morbidities are lower than those in ELSA, see Table 9-10c which provides a comparison of prevalence rates for diabetes used by POPPI versus DHBS rates published from ELSA. For CVD, unfortunately ELSA does not publish rates to compare directly with Table 9-10a and Table 9-10b, but the CHD rates are provided for some context in the case of heart attacks (heart attacks are a morbidity that forms part of CHD, strokes do not). Thirdly, the model forecasts in Table 9-2 shows that DHBS is a morbidity that is predicted to be more concentrated in the old-old population by 2031, a dynamic which the constant POPPI rates does not capture.

9.9 CONCLUSION

In this chapter the results of the health microsimulation have been presented. The microsimulation produces a population of individuals for each LAD that represents its projected gender, age and ethnic composition along with the smoking status of each individual. As the microsimulation progresses through each simulation time step the morbidity status of individuals is updated using a Monte Carlo approach. These outputs can be summarised in a number of ways, the two used here are prevalence counts of individuals with a morbidity and the prevalence rate of the morbidity within the population. These statistics can be aggregated to other higher geographies, to produce morbidity forecasts at the national or health region level.

An overview of the trends in the three morbidities and the comorbidity shows that the number of individuals with CVD and respiratory illness will decrease over the forecast 20 year time period. This decrease will occur in all LADs for CVD, and in all but 11 LADs for respiratory illness. These trends are driven by reductions in the population who have smoked and a general negative time trend in the incidence of these morbidities that reflects non-quantifiable causes, e.g. occupation, lifestyles and air quality. The trend for DHBS is the reverse, in all but 15 LADs there are increases in the prevalence of this morbidity from 2011 to 2031. This morbidity has a strong ethnic gradient, being more prevalent in the BME population and the demographic projections are for the relative size of the BME population within the 50 year and older population to increase. Also, this morbidity is not influenced by smoking so the same extent as the other morbidities thereby negating the improvements attributable to this lifestyle change. The comorbidity of CVD and DHBS is driven by the trends in both its component morbidities and the reduction in the CVD prevalence means that the prevalence of the comorbidity falls in all but one LAD.

These trends are also examined using both the sub national, health related, geography of NHS LATs and the public health geography of LADs in general. All the influences discussed above come to play in the dynamics that drive particular trends within LATs. The influence of having a large BME population in some LATS is seen to greatly increase their DHBS prevalences.

Looking at case study authorities, whilst there is some unanimity in the trends across LADs, as highlighted above, the actual level of the morbidity in each LAD varies. This is a result of different starting levels in 2011 and different gradients in the trends. These geographically disparities in both prevalence counts and rates are mapped for 2031. These maps illustrate that to gain a full picture of the health status and health demands of an area both these measures are insightful.

An examination of the robustness of the health microsimulation to the Monte Carlo aspect of its operation, as was done for the spatial microsimulation, shows the robustness to be good. The final section provides a comparison with some alternative forecasts of related morbidities. The task here is doubly complicated by the paucity of such alternatives and differences in the definition of each morbidity and its framing.

The general impact on the total burden of ill health associated with these three morbidities can inform the headline trends seen in Figure 2-10 which show modest improvements in disability free life expectancies over time. In this study, CVD and respiratory illness are seen to reduce over time helping to extend disability free life expectancy whilst increases in DHBS will reduce disability free life expectancy. Whilst these three morbidities capture a large proportion of the morbidity cases in the population, they do not capture all; and changes in the prevalences of other morbidities (e.g. cancer, arthritis and dementia) will similarly influence the disability free life expectancy trend. It is therefore difficult to infer how the trends seen with these three case study morbidities will drive future disability free life expectancies.

Whilst these results have provided a 'headline' picture of likely future local trends in the morbidities there is certainly further insight to be gained by decomposing these trends and attributing them to various demographic and health improvement factors. This are the second and third research questions and they are addressed in the following chapter which examines how the changing gender, ethnic, age mix of a LAD and the health improvements contribute towards this headline picture.

10 DECOMPOSITION OF HEALTH TRENDS

An Accounting System

10.1 INTRODUCTION

In Chapter 9 the morbidity outcomes from the health microsimulation are presented biennially from 2011 to 2031 at national, NHS LAT and LAD geographies. There are clearly a number of influences at work that are causing the observed trends. The most frequent question asked about a complex forecasting technique such as the health microsimulation is "*what factors are driving it?*" A method is therefore needed that quantifies how much influence is attributable to each factor.

Whilst there are explicit trends embodied in the probabilities of health outcomes captured by the hazard models there are also important demographic influences at work. Even if there were no changes to the health outcomes, these demographic influences have the potential to affect both the morbidity prevalence count and rate within the population. In this chapter the influence of five main demographic factors will be examined using an accounting system framework similar to that proposed by Bongaarts and Bulatao (1999) and Rees et al. (2013b). In addition the influence of the health improvements embodied in the health microsimulation will be examined to complete the picture.

The demographic factors are related back to the original population which is the 2011 Census population of the LAD. The first demographic factor is the size of the population, and, all other things being equal, a larger population will lead to a larger prevalence count but the same prevalence rates. The next demographic factor is the gender split in the population, and this outcome represents how the change in the size of the population and the relative sizes of the two genders influences prevalence counts and rates. The third additional demographic factor is the projected ethnic split of the population, whose influence will once again be felt on the prevalences. The next two demographic influences are age related. One captures the re-structuring of the population along size, gender, ethnic and age projections but with an assumption that there is no improvement in mortality over time. The final demographic influence allows for the projected influence of mortality improvement to take place. In all these scenarios the prevalence rates within the various sub-populations will remain constant over time. No one's morbidity status will be updated and any changes in the prevalence rate for the whole population occurs through a simple re-weighing of these prevalence rates for the sub-populations whose sizes change in line with the demographic projections. The final

outcome is the situation when these health outcomes change according to the hazard models in Chapter 8. This is the outcome described in Chapter 9. Here the prevalence rates over time will vary as the incidence probabilities predicted by the hazard models vary according to time, age, gender, ethnicity, smoking status, presence of comorbidities and type of LAD. There will also be further impacts through a recruitment of healthier 'Replenishers' into the future synthetic waves and the increased mortality of those who have a morbidity.

It should be made clear at this stage that only the final outcome is thought to be plausible. All the previous outcomes in terms of the demography within a LAD are purely artificial constructs whose purpose is to provide stepping stones from an original population to one that has evolved over time, both in its structure and its health outcomes. Also the sequence in which demographic changes are considered may be varied. For example rather than the sequence: size, gender, ethnicity, age, mortality, used here, alternative sequences are possible, e.g.: size, gender, age, ethnicity, mortality.

Section 10.2 of this chapter describes how these various decomposition scenarios have been implemented and the degree to which the necessary population structures have been achieved. The third section, 10.3, provides a picture of the impact of the decomposition at the national level, looking at the outcomes measured by the prevalence counts and rates and the ratios in equation 5-21a (page 132). This is followed by section 10.4 which discusses the distribution of these decomposition ratios amongst the English LADs. The penultimate section 10.5 looks in detail at the differences in these ratios amongst the seven case study authorities and the final section, 10.6, summaries the finding of this chapter.

10.2 IMPLEMENTATION

The implementation of each of these decomposition scenarios is achieved using the same C code that is used in Chapter 9 to produce the health microsimulation outcomes. The various demographic influences are represented by changing the input data used by the program and not performing the full range of population dynamics.

10.2.1 Adapting the input data

To capture each demographic influence the only data that needs to change is the data that define the size and shape of the population in each LAD by gender, age and ethnicity. The information contained in the ELSA population, e.g. gender, age in 2011, ethnicity, smoking and morbidity status in 2011, is required but does depend on the scenario. Information on the number of Replenishers at each future synthetic wave is not required since no one ages and, since there are no changes to individual's morbidity

status or a need to weight individuals for selection, no hazard model parameter estimates are required.

As described in 5.6 the structure of these artificial population projections is achieved in a number of ways. What is required initially is the size of each LAD population without mortality improvement ($PN^{LAD,y}_{+,+,+}$ in equation 5-12). For the structure where the original population increases to this size, equation 5-14 is used. This population has the same gender, age and ethnic structure as the original 2011 Census population, but the size of the population is that without mortality improvement in each forecast year. The next two structures, where the gender (equation 5-15) and then the gender and ethnic structure (equation 5-16) are in line with revised ETHPOP projections, is achieved in two stages. The first stage is to use simple multipliers (as shown in equations 5-15 and 5-16) as reasonable starting points to achieve these structures, followed by a second optimisation stage. The impact of this optimisation is measured in how accurately the required structure is achieved and what scale of changes are necessary to achieve it. It is these two projections that this section is most concerned with. The remaining two demographic structures are simply $PN^{LAD,y}_{g,a,e}$ (gender, age and ethnic structures in line with revised ETHPOP but the population size does not incorporate mortality improvement) and $P^{LAD,y}_{g,a,e}$ (the revised ETHPOP projections).

The results of this optimisation, like the revision of ETHPOP projections, produces non integer population counts for each gender, age and ethnic combination. By the way that the health microsimulation works, these fractional projections are integerised by rounding to the nearest integer.

10.2.1.1 Optimised for gender

Here the task is to minimise the objective function 5-17 that measures how far the counts obtained from applying the simple multipliers are from reproducing the gender split from the revised ETHPOP and the age by ethnicity splits from the 2011 Census counts. Figure 10-1 illustrates the outcome and the impact of the optimisation process from optimisations of the 324 LADs, with one optimisation for each forecast year. The value of the TAE_G at the end of the optimisation is shown in the upper left chart. The sum over the 2 + (8 x 5) = 42 marginal constraints in equation 5-17 is never greater than 5.0. The resultant reduction in TAE_G is shown in the upper right chart. These reductions are modest , suggesting that the initial count estimates from applying the multipliers were already close to those required. These improvements have been obtained by changing the 2 * 8 * 5 = 80 individual gender by age by ethnicity counts in each LAD. The absolute sum of these changes expressed as a percentage of the LAD population, are shown in the lower left hand chart. These percentages are small, with no LAD requiring more than a total of a 1.1% change over all its constituent counts. The

greatest adjustment that was required in each LAD for any forecast year is shown in the lower right chart. For only four years was there a need to make an adjustment to an age, gender and ethnic group category that was greater than 100.

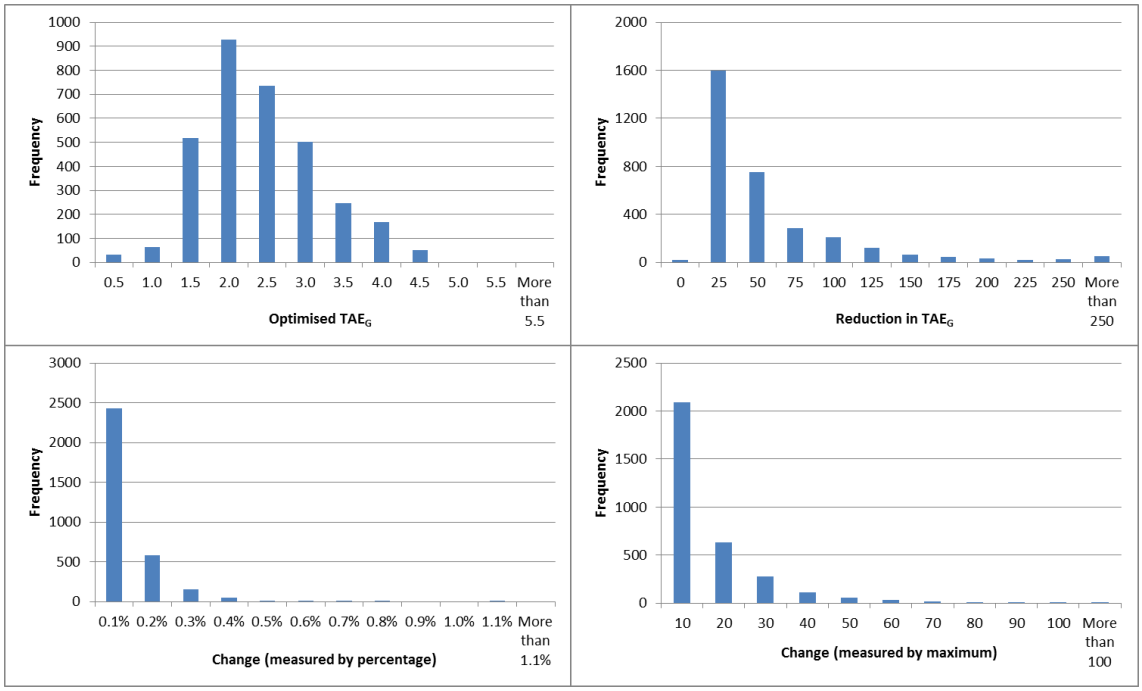


Figure 10-1 : Impact of optimisation to reproduce revised ETHPOP gender splits

The charts in Figure 10-1 show that the adjustment process was successful in achieving the stated objective and this was achieved by only conservative changes to the initial estimates obtained by applying the multipliers. This is consistent with the assumption that the gender splits with each LAD are stable over time.

10.2.1.2 Optimised for gender and ethnicity

Here the task is to minimise the objective function 5-18 that measures how far the counts obtained from applying the simple multipliers are from reproducing the gender by ethnicity splits from the revised ETHPOP projections and the age splits from the 2011 Census counts. As previously, the impact of this process is illustrated in a series of charts in Figure 10-2.

Once again the optimisation has achieved a set of projections that reproduce the required marginal totals well. In this case there are (2 x 5) + 8 = 18 such marginal totals. No year within a LAD has an optimised TAE_{GE} greater than 2.0 (upper left). However to achieve this optimisation there has been some large reductions from the initial TAE_{GE} (upper right). These large reductions have been obtained by some substantial changes to the initial gender by age by ethnicity counts both measured using the sum of the absolute changes expressed as a percentage of the LAD population (lower left) and the maximum adjustment made (lower right).

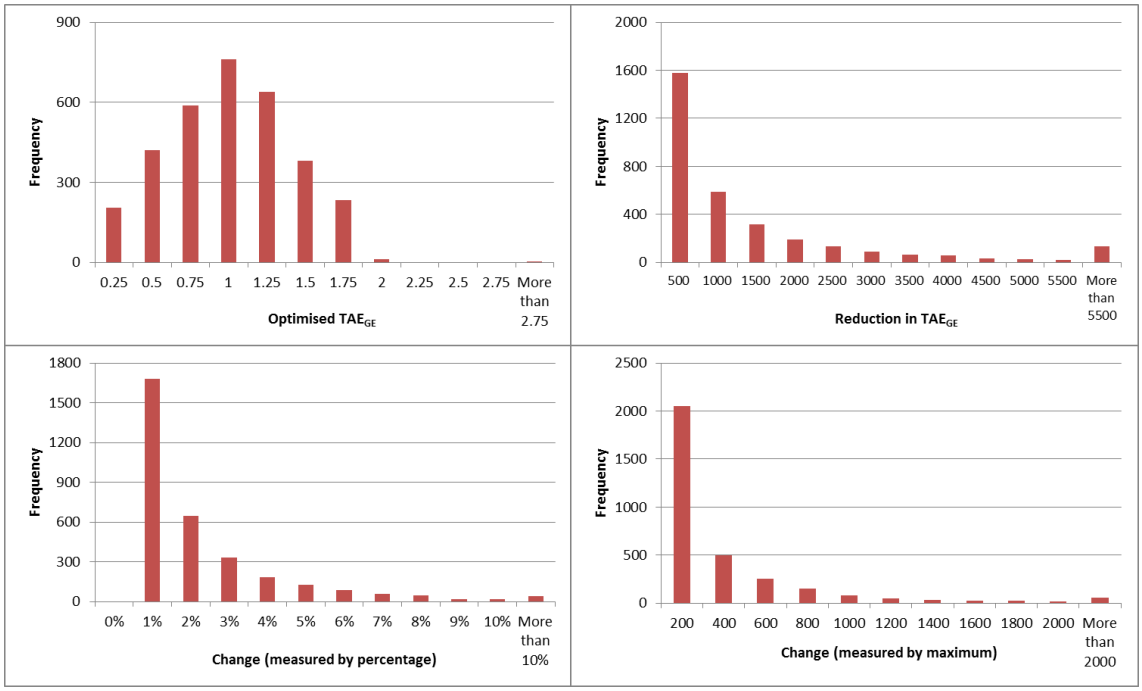


Figure 10-2 : Impact of optimisation to reproduce revised ETHPOP gender and ethnicity splits

To adjust for both gender and ethnicity is always a more complex task than just adjusting for gender alone, and consequently the changes illustrated in Figure 10-2 are much more substantial than in Figure 10-1. This implies that there is the potential for significant changes in the gender by ethnicity split within LADs over the forecast years that cannot be addressed by simply applying the multipliers.

10.2.2 Simplifying the software

Of critical importance to the implementation of the accounting system is that the morbidity rates within the spatially microsimulated population should not change. This is achieved by ensuring in the bespoke C program that a number of actions do not take place. Firstly, individuals do not age, so for example, a person who represents a 54 year old south Asian male in 2011, represents this same demographic in all forecast years. Since nobody ages then there is no need to introduce Replenishers into the sample. Indeed to do so would affect the prevalence rate at the younger age bands. Secondly, individuals are selected for cloning or removal within the population with equal weight. To allow this weight to vary, for example by morbidity status, would target certain individuals within the population and, since having a morbidity increases the chances of selection for removal, the prevalence rate would fall. Thirdly, and perhaps most importantly, the morbidity status of the individuals does not change - if they do not have the morbidity in 2011 then they cannot acquire the morbidity at any stage. This once

again ensures that the prevalence rates do not change, so if the 54 year old south Asian male in 2011 does not have the morbidity in 2011, they will not have it by 2031 as well. In practice, however, the prevalence rates are not constant since they are influenced by the Monte Carlo sampling that still takes place to re structure the population and if any shadow individuals are required then their morbidity status will affect the overall prevalence rate. In practice, however, these sampling effects are minimal.

10.3 TRENDS IN ENGLAND

Using the health microsimulation software and the range of input data produces a series of model outputs in a consistent output format. The size of the population at risk in each scenario is shown in Table 10-1. There are a number of features in this table. Firstly, for the original population, the size of the population is the same as the 2011 Census population in each forecast year. Re-sizing increases the size of the population, to the size of a population for the LAD without mortality improvements. For the next three scenarios: gender; gender and ethnicity; and gender, ethnicity and age the population sizes are very similar, influenced only by different patterns of integerisation. Incorporating mortality improvement into the projections increases the size of the population and this is the same population that is used in the health microsimulation.

Table 10-1 : Population at risk totals under the decomposition scenarios

Total	Original	Re-size	Gender	Gender and Ethnicity	Gender Ethnicity and Age	Mortality Improvement	Health Microsimulation
2011	18,229,893	18,229,893	18,229,893	18,229,893	18,229,893	18,229,893	18,229,893
2013	18,229,893	18,964,511	18,964,471	18,964,753	18,964,704	18,964,704	18,964,704
2015	18,229,893	19,612,338	19,612,286	19,612,399	19,612,348	19,689,593	19,689,593
2017	18,229,893	20,211,922	20,211,911	20,211,990	20,211,925	20,400,332	20,400,332
2019	18,229,893	20,754,507	20,754,524	20,754,526	20,754,480	21,083,429	21,083,429
2021	18,229,893	21,240,004	21,240,081	21,240,073	21,239,982	21,732,680	21,732,680
2023	18,229,893	21,628,697	21,628,633	21,628,713	21,628,634	22,304,919	22,304,919
2025	18,229,893	21,880,739	21,880,726	21,880,821	21,880,775	22,755,033	22,755,033
2027	18,229,893	22,046,432	22,046,369	22,046,350	22,046,344	23,129,460	23,129,460
2029	18,229,893	22,238,128	22,238,072	22,238,203	22,238,105	23,539,913	23,539,913
2031	18,229,893	22,502,369	22,502,421	22,502,405	22,502,352	24,029,232	24,029,232
Difference	0	4,272,476	4,272,528	4,272,512	4,272,459	5,799,339	5,799,339

10.3.1 Decomposition of CVD

The number of individuals with CVD in each forecast year under the various decomposition scenarios is given in Table 10-2.

Table 10-2 : Individuals with CVD under the decomposition scenarios

With CVD	Original	Re-size	Gender	Gender and Ethnicity	Gender Ethnicity and Age	Mortality Improvement	Health Microsimulation
2011	2,873,694	2,873,694	2,873,694	2,873,694	2,873,694	2,873,694	2,873,694
2013	2,873,694	2,992,904	2,994,784	3,002,632	3,027,879	3,027,330	2,803,706
2015	2,873,694	3,093,818	3,097,418	3,111,044	3,138,662	3,164,032	2,737,698
2017	2,873,694	3,187,593	3,192,244	3,212,994	3,240,308	3,303,070	2,634,641
2019	2,873,694	3,272,072	3,278,106	3,307,879	3,354,343	3,464,498	2,531,619
2021	2,873,694	3,347,778	3,354,152	3,395,654	3,455,898	3,622,299	2,451,228
2023	2,873,694	3,408,292	3,414,993	3,469,429	3,587,731	3,818,306	2,292,377
2025	2,873,694	3,447,070	3,454,076	3,523,048	3,705,034	4,005,625	2,207,095
2027	2,873,694	3,472,578	3,478,763	3,564,194	3,825,943	4,201,724	2,076,336
2029	2,873,694	3,502,031	3,507,667	3,611,815	3,946,325	4,403,298	1,943,832
2031	2,873,694	3,543,279	3,548,232	3,672,311	4,053,720	4,593,366	1,858,903
Difference	0	669,585	674,538	798,617	1,180,026	1,719,672	-1,014,791

Using the counts in Table 10-1 and Table 10-2 it is possible to calculate the prevalence rates for this morbidity which is given as Table 10-3. In this table the trend in the prevalence rates within each decomposition scenario over time is highlighted using red shading, from dark red for years with high prevalence to white for the lowest prevalence. Even though there is some gradation in the prevalences for the re-sized and gender scenarios, the actual prevalences are not much different and are mainly attributable to the Monte Carlo sampling that takes place to re structure the population. In this table, re-sizing make very little difference to the prevalence rate, whilst the impact of changing the gender split makes a similarly modest difference to the prevalence rates. Changing both the gender and ethnic splits to be in line with revised ETHPOP projections does cause the prevalence rate to increase over time, and this quantifies the impact of the changing ethnic structure of the English population on this morbidity. Changing the gender, ethnic and age structure (but having no mortality improvement) further increases the prevalence rate over time. The final demographic scenario of using the revised ETHPOP projections, which incorporates mortality improvements, forecasts the highest prevalence by 2031. Only in this case, when the health improvements indicated by the hazard models and the selective weight of mortality or migration in Chapter 8 are incorporated are improvements in the prevalence rate forecast (as reported in Chapter 9).

Table 10-3 : CVD prevalence rates under the decomposition scenarios

CVD	Original	Re-size	Gender	Gender and Ethnicity	Gender Ethnicity and Age	Mortality Improvement	Health Microsimulation
2011	15.76	15.76	15.76	15.76	15.76	15.76	15.8
2013	15.76	15.78	15.79	15.83	15.97	15.96	14.8
2015	15.76	15.77	15.79	15.86	16.00	16.07	13.9
2017	15.76	15.77	15.79	15.90	16.03	16.19	12.9
2019	15.76	15.77	15.79	15.94	16.16	16.43	12.0
2021	15.76	15.76	15.79	15.99	16.27	16.67	11.3
2023	15.76	15.76	15.79	16.04	16.59	17.12	10.3
2025	15.76	15.75	15.79	16.10	16.93	17.60	9.7
2027	15.76	15.75	15.78	16.17	17.35	18.17	9.0
2029	15.76	15.75	15.77	16.24	17.75	18.71	8.3
2031	15.76	15.75	15.77	16.32	18.01	19.12	7.7

Perhaps most revealing are the ratio multipliers as defined in equation 5-20a to 5-20f and which can be calculated from Table 10-2 and are provided here in Table 10-4 (with the column headings identifying the relevant ratio multiplier). These ratios are multiplicative, as illustrated in equation 5-21.

Table 10-4 : CVD decomposition ratio multipliers

CVD	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improvement (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.04	1.00	1.00	1.01	1.00	0.93	0.98
2015	1.08	1.00	1.00	1.01	1.01	0.87	0.96
2017	1.11	1.00	1.01	1.01	1.02	0.80	0.92
2019	1.14	1.00	1.01	1.01	1.03	0.73	0.87
2021	1.16	1.00	1.01	1.02	1.05	0.68	0.85
2023	1.19	1.00	1.02	1.03	1.06	0.60	0.80
2025	1.20	1.00	1.02	1.05	1.08	0.55	0.76
2027	1.21	1.00	1.02	1.07	1.10	0.49	0.71
2029	1.22	1.00	1.03	1.09	1.12	0.44	0.67
2031	1.23	1.00	1.03	1.10	1.13	0.40	0.63

Looking at Table 10-4, re-sizing the population, unsurprisingly increases the number of individuals with the morbidity. However, for the resizing scenario (RS, using the notation of equations 5-20 and 5-21) the increase in the number of individuals with CVD relative to the original population is in proportion to the increase in the at risk population (this ratio can be calculated using Table 10-1) and so does not affect the prevalence, as seen in Table 10-3. By 2031 the gender scenario (RG) has little impact on the number with CVD in comparison to the re-size scenario. The size, gender and ethnicity scenario (RE) increases those with CVD by a factor of 1.03 by in 2031. Further increases are seen for the size, gender, ethnicity and age (RN) and mortality

improvement ratios (RM) – these reflect a society that is changing its ethnic structure, ageing and living longer. All these impacts cause there to be greater numbers with CVD. This trend for an increase in the population with CVD continues, until the health improvements are incorporated using the health microsimulation, which dramatically lowers the ratio. By 2031 there are just 0.40 times as many individuals with CVD, compared with the previous scenario of mortality improvements. The 0.35 reduction from 2011 to 2031 in the number of individuals with CVD seen in the final column of Table 10-2 is composed of a 1.57 increase due to demographic changes in the English population which is then offset by a 0.60 reduction due to forecast health improvements and selective mortality and migration.

10.3.2 Decomposition of diabetes or high blood sugar

Table 10-5 provides the counts of individuals with DHBSs under the various decomposition scenarios (the population at risk is the same for all morbidities and is given in Table 10-1 and using these two tables the reader is able to calculate the prevalence of DHBS).

Table 10-5 : Individuals with diabetes or high blood sugar under the decomposition scenarios

Diabetes or high blood sugar	Original	Re-size	Gender	Gender and Ethnicity	Gender Ethnicity and Age	Mortality Improvement	Health Microsimulation
2011	2,189,938	2,189,938	2,189,938	2,189,938	2,189,938	2,189,938	2,189,938
2013	2,189,938	2,276,801	2,278,310	2,293,971	2,296,726	2,297,138	2,424,851
2015	2,189,938	2,354,167	2,357,184	2,386,839	2,382,834	2,396,726	2,656,947
2017	2,189,938	2,426,042	2,429,929	2,475,466	2,464,241	2,497,686	2,834,345
2019	2,189,938	2,491,301	2,495,686	2,560,065	2,550,691	2,608,150	2,979,780
2021	2,189,938	2,550,083	2,554,500	2,638,726	2,636,668	2,723,639	3,128,636
2023	2,189,938	2,597,227	2,601,630	2,707,846	2,731,366	2,851,741	3,191,523
2025	2,189,938	2,628,044	2,632,876	2,762,728	2,811,236	2,968,168	3,300,025
2027	2,189,938	2,648,832	2,653,463	2,809,809	2,883,514	3,077,822	3,358,974
2029	2,189,938	2,673,327	2,677,338	2,863,466	2,967,637	3,203,789	3,386,244
2031	2,189,938	2,706,720	2,710,272	2,928,029	3,060,763	3,340,327	3,436,486
Difference	0	516,782	520,334	738,091	870,825	1,150,389	1,246,548

Using the information in Table 10-5, the ratio multipliers for DHBS are calculated as shown in Table 10-6.

Table 10-6 : Diabetes or high blood sugar decomposition ratio multipliers

Diabetes or high blood sugar	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improvement (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.04	1.00	1.01	1.00	1.00	1.06	1.11
2015	1.07	1.00	1.01	1.00	1.01	1.11	1.21
2017	1.11	1.00	1.02	1.00	1.01	1.13	1.29
2019	1.14	1.00	1.03	1.00	1.02	1.14	1.37
2021	1.16	1.00	1.03	1.00	1.03	1.15	1.42
2023	1.19	1.00	1.04	1.01	1.04	1.12	1.46
2025	1.20	1.00	1.05	1.02	1.06	1.11	1.51
2027	1.21	1.00	1.06	1.03	1.07	1.09	1.54
2029	1.22	1.00	1.07	1.04	1.08	1.06	1.55
2031	1.24	1.00	1.08	1.05	1.09	1.03	1.58

Here the pattern of contribution for each scenario is different to that for CVD with all scenarios showing an increase in the number of individuals with DHBS by 2031. After the re-size scenario (RS), the greatest increase is from the mortality improvement scenario (RM), which itself is just larger than that for the gender and ethnicity scenario (RE). The pattern for health improvements implemented in the health microsimulation show a peak in its contribution, occurring around 2021, and the longer term trend in this column suggests that soon after 2031 this ratio should fall below 1.00. Overall, the trends in Table 10-6 indicate that the increased number of individuals with DHBS is mainly attributable to changes in the ethnic composition of England (RE) and general improvements in life expectancies (RM).

10.3.3 Decomposition of respiratory illnesses

In terms of the absolute number of individuals with a respiratory illness, Table 10-7 shows that under the five demographic decomposition scenarios the number of individuals with such an illness increases, with only the health microsimulation scenario showing a slight reduction in the numbers with a respiratory illness (however the prevalence in 2031 will be much lower than in 2011 since the population is much larger in 2031, see Table 10-1).

Table 10-7 : Individuals with a respiratory illness under the decomposition scenarios

Respiratory illness	Original	Re-size	Gender	Gender and Ethnicity	Gender Ethnicity and Age	Mortality Improvement	Health Microsimulation
2011	3,408,810	3,408,810	3,408,810	3,408,810	3,408,810	3,408,810	3,408,810
2013	3,408,810	3,549,744	3,549,489	3,551,121	3,555,658	3,555,688	3,396,242
2015	3,408,810	3,670,182	3,669,021	3,672,768	3,680,422	3,697,798	3,469,615
2017	3,408,810	3,782,064	3,779,818	3,787,692	3,797,642	3,840,417	3,495,865
2019	3,408,810	3,883,385	3,880,363	3,893,839	3,906,242	3,981,644	3,490,010
2021	3,408,810	3,973,542	3,971,271	3,992,476	4,003,948	4,116,560	3,562,378
2023	3,408,810	4,045,735	4,043,389	4,074,072	4,089,857	4,244,890	3,477,186
2025	3,408,810	4,092,263	4,090,041	4,131,986	4,161,921	4,363,131	3,482,338
2027	3,408,810	4,122,784	4,120,848	4,175,443	4,229,128	4,479,240	3,439,145
2029	3,408,810	4,158,318	4,156,446	4,224,881	4,305,979	4,609,816	3,366,595
2031	3,408,810	4,207,681	4,206,100	4,289,796	4,395,539	4,754,767	3,376,245
Difference	0	798,871	797,290	880,986	986,729	1,345,957	-32,565

Calculating the ratio multipliers in Table 10-8 shows a similar pattern to that for CVD in Table 10-4, with the demographic scenarios showing increases in the numbers with the morbidity which is moderated down by the health improvements in the health microsimulation (although not to the same extent as seen for CVD).

Table 10-8 : Respiratory illness decomposition ratio multipliers

Respiratory illness	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.04	1.00	1.00	1.00	1.00	0.96	1.00
2015	1.08	1.00	1.00	1.00	1.00	0.94	1.02
2017	1.11	1.00	1.00	1.00	1.01	0.91	1.02
2019	1.14	1.00	1.00	1.00	1.02	0.88	1.02
2021	1.17	1.00	1.01	1.00	1.03	0.87	1.06
2023	1.19	1.00	1.01	1.00	1.04	0.82	1.02
2025	1.20	1.00	1.01	1.01	1.05	0.80	1.03
2027	1.21	1.00	1.01	1.01	1.06	0.77	1.01
2029	1.22	1.00	1.02	1.02	1.07	0.73	0.99
2031	1.23	1.00	1.02	1.02	1.08	0.71	0.98

10.3.4 Decomposition of CVD and diabetes or high blood sugar

The trends in the table that show how many individuals have both CVD and DHBS under the various decomposition scenarios more closely resembles that for CVD (Table 10-2) rather than that for DHBS (Table 10-5). All the demographic decomposition scenarios show increasing numbers with the comorbidity whilst the incorporation of health improvements in the health microsimulation substantially reduces the numbers with the morbidity.

Table 10-9 : Individuals with CVD and diabetes or high blood sugar under the decomposition scenarios

CVD and diabetes or high blood sugar	Original	Re-size	Gender	Gender and Ethnicity	Gender Ethnicity and Age	Mortality Improvement	Health Microsimulation
2011	748,821	748,821	748,821	748,821	748,821	748,821	748,821
2013	748,821	779,985	780,548	789,555	796,235	796,221	746,541
2015	748,821	806,124	807,617	822,957	829,627	836,168	751,172
2017	748,821	830,644	832,746	855,450	859,343	875,039	733,176
2019	748,821	853,097	855,477	887,488	895,339	921,947	708,491
2021	748,821	873,242	875,844	917,963	931,018	971,756	696,057
2023	748,821	889,565	892,243	945,474	977,212	1,033,286	648,791
2025	748,821	900,200	903,049	968,380	1,015,126	1,088,451	630,625
2027	748,821	907,438	910,163	988,852	1,051,486	1,142,294	592,600
2029	748,821	916,056	918,502	1,011,360	1,092,203	1,202,593	552,603
2031	748,821	927,517	929,946	1,037,631	1,134,712	1,265,497	533,763
Difference	0	178,696	181,125	288,810	385,891	516,676	-215,058

These changes in the numbers with the comorbidity are also reflected in the multipliers shown in Table 10-10.

Table 10-10 : CVD and diabetes or high blood sugar decomposition ratio multipliers

CVD and diabetes or high blood sugar	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improvement (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.04	1.00	1.01	1.01	1.00	0.94	1.00
2015	1.08	1.00	1.02	1.01	1.01	0.90	1.01
2017	1.11	1.00	1.03	1.00	1.02	0.84	0.98
2019	1.14	1.00	1.04	1.01	1.03	0.77	0.95
2021	1.17	1.00	1.05	1.01	1.04	0.72	0.93
2023	1.19	1.00	1.06	1.03	1.06	0.63	0.87
2025	1.20	1.00	1.07	1.05	1.07	0.58	0.84
2027	1.21	1.00	1.09	1.06	1.09	0.52	0.79
2029	1.22	1.00	1.10	1.08	1.10	0.46	0.73
2031	1.24	1.00	1.12	1.09	1.12	0.42	0.71

10.3.5 Summary

The results presented in this section all clearly identify that there are significant demographic drivers that show, all other things being equal, the number of individuals in the population with a morbidity should increase significantly. This is primarily caused by changes in the ethnic structure of the population, its ageing due to cohort replacement and increases in life expectancy. However, in the case of CVD and respiratory illness, these demographic impacts are mitigated to a large degree by various health trends that affect the incidence of the morbidities. Firstly there are time trends

that take account of the impact of influences not explicitly included in the hazard models. The trends for these morbidities are negative which will lead to both reductions in future incidence for the synthetic population in each LAD and reduced prevalence in the Replenisher cohorts who join the sample at future synthetic waves. Secondly the health microsimulation incorporates a selection weight that is influenced by various factors, including the presence of a morbidity. As illustrated in Box 5-11, this will tend to reduce the prevalence rate by disproportionately reducing the numerator (those with the morbidity) in the prevalence rate calculation (equation 5-23). Without this selection weight, as is the case in all the demographic scenarios which all have equal weights, the prevalence is higher.

10.4 DECOMPOSITION FOR LOCAL AUTHORITY

Whilst section 10.3 provides a picture for the whole of England, there will be variation in the decomposition ratios by LADs. In this section this variation is examined using the distribution of the 2031 ratios within each scenario. A traditional box plot chat will be used to identify the distribution by its medians, inter quartile ranges, whiskers (1.5 x the interquartile range beyond the quartile) and outliers below and above the whiskers.

10.4.1 Cardio Vascular Disease

The distribution of the population re-sizing ratio (RS) in Figure 10-3 represents the pro-rata increase in the number of individuals with the morbidity relative to the size of the population at risk. This is to be anticipated and is consistent across all morbidities since they all experience the same growth in population. The scenario that changes the future population to reflect the LADs gender structure (RG) shows very little impact, with the ratio being close to 1.0 with very little deviation. Again, this feature is consistent across all the morbidities, suggesting that the relative sizes of the two genders in each LAD either remains static or that the differential in health outcomes associated with each gender are not critical.

With the scenario that reflects both projected gender and ethnic structures (RE), there are a range of ratios, with the distribution having a median near 1.0 that is positively skewed towards higher values for this ratio. This shows that future changes to the ethnic structure of LADs will, after size and gender changes, tend to increase the number of individuals with CVD. The ratios derived from changes in the gender, ethnic and age structure in LADs (RN) also shows a wide range of ratios, and the median is again greater than 1.0. This means that the changing age structure will tend to increase those with CVD, but for some authorities the change will reduce those with CVD. These latter LADs are primarily London area type authorities, whose population is not anticipated to age a great deal in the next 20 years (see Figure 2-5). The impact of mortality

improvement (RM) is always to cause an increase in those with CVD, although the impact is consistent and small. The health trends embodied in the health microsimulation (RH) cause a dramatic reduction in those with CVD, after having accounted for potential growths associated with the demographic transitions in each LAD. The median is 0.42 and the quartiles are very close to this median. A consequence is that there are a large percentage (15%) of LADs that are considered outliers. The lower outliers are mainly London authorities and Cities and Services centres, e.g. cities and large towns in southern England. The larger outliers are mainly LADs of type Mining and Manufacturing.

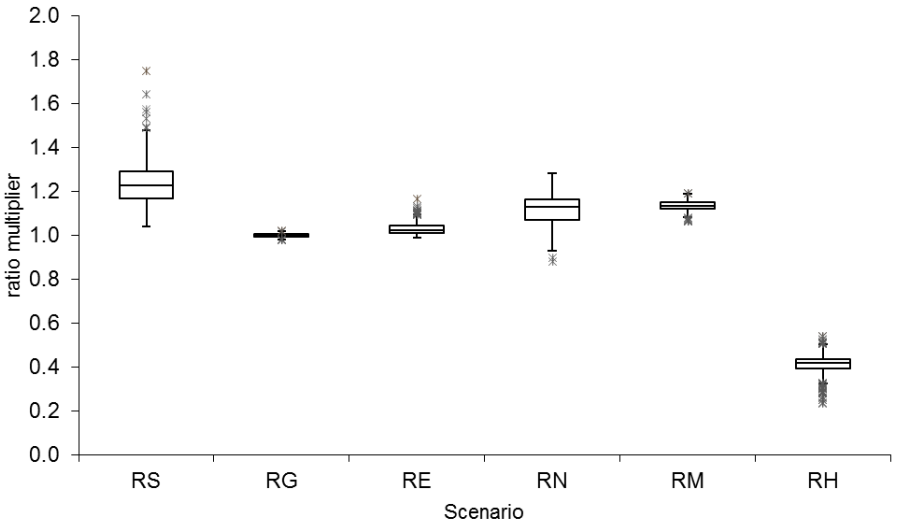


Figure 10-3 : Distribution of decomposition ratios in LADs for CVD, 2031

Key RS : Re-size; RG : Gender; RE : Gender and Ethnicity; RN Gender, Ethnicity and Age; RM : Mortality Improvement; RH : Health microsimulation

10.4.2 Diabetes or high blood sugar

The box plots for the distribution of the 2031 decomposition ratios for DHBS show some contrasts to those for CVD. Firstly the spread of these ratios for the scenario where the gender and ethnic structures are adapted (RE) is more skewed with a longer positive tail of high ratios. In contrast, the subsequent scenario where age is also adapted (RN), shows a smaller spread. The median ratio for the health microsimulation scenario (RH) is just less than 1.0 but contains a distinct cluster of high values. These values are all London Centre type authorities and the reasons for these outliers is discussed in the context of this area type case study authority (see section 10.5.3). Even discounting these outliers there appears to be considerable spread in these ratios.

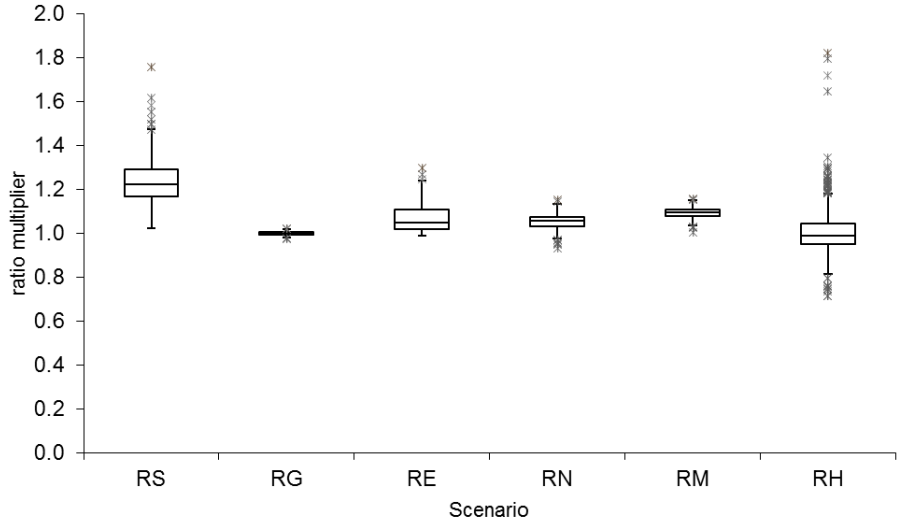


Figure 10-4 : Distribution of decomposition ratios in LADs for diabetes or high blood sugar, 2031

10.4.3 Respiratory Illness

Excepting the re-size and health microsimulation scenarios, the distribution of the ratios for respiratory illness show a tightness of distribution, with the difference in upper and lower quartiles and the upper and lower whiskers being small. The medians are also just slightly above 1.0, which indicates that these scenarios are fairly neutral relative to the previous scenarios. There are sizeable reductions in the number of individuals with this morbidity as a result of the health trends embodied in the health microsimulation (RH) scenario, but to a lesser degree than for CVD.

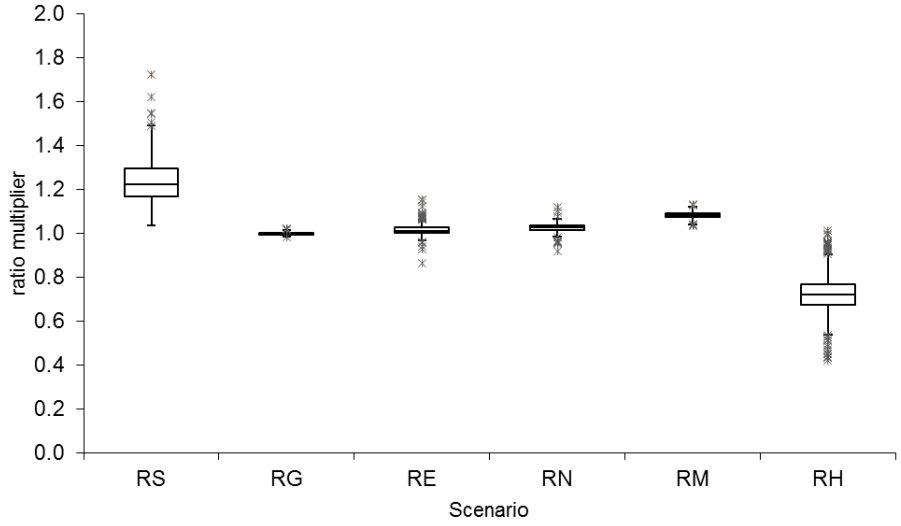


Figure 10-5 : Distribution of decomposition ratios in LADs for respiratory illness, 2031

10.4.4 CVD and diabetes or high blood sugar

The comorbidity shows sizeable spreads in the distribution of all its decomposition ratios (excepting gender, RG), which is in contrast to the distribution for the single morbidities of CVD and DHBS which contain a variety of spreads. This suggests that both the morbidities are playing a role in influencing the variety in the number of individuals with the comorbidity in each scenario. Only in the case of this comorbidity does the spread in ratios for the scenario with no morbidity increase (RM) appear large. The distribution of ratios for the final health microsimulation (RH) more closely resembles that for CVD than for DHBS.

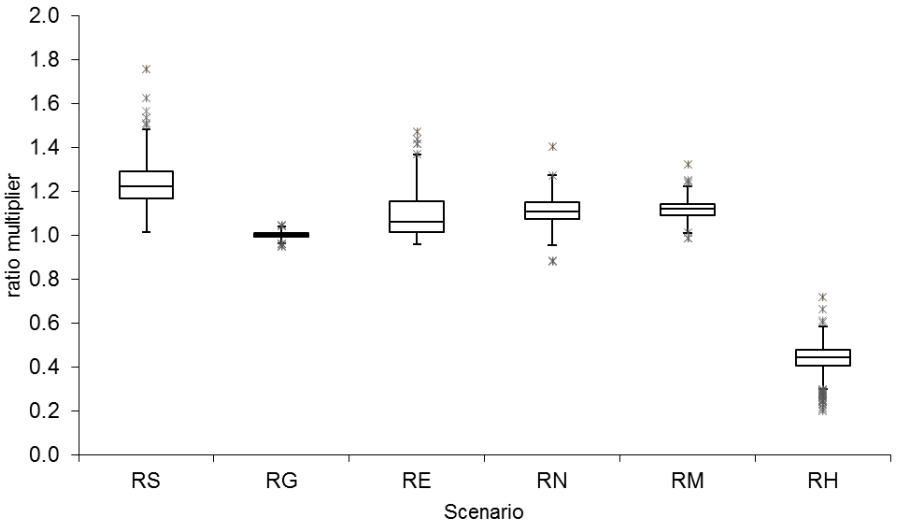


Figure 10-6 : Distribution of decomposition ratios in LADs for CVD and diabetes or high blood sugar, 2031

10.4.5 Summary

What the box plots in this section have shown is that there are important local variations in how the changing ethnic and age structure of individual LADs impacts on the morbidity prevalence counts. This most starkly reflected in the ranges of the middle three ratios (RE, RN and RM), which is some case can be large. The impact of these demographic factors is different depending on the morbidity. With CVD it is an ageing population that is important whilst for DHBS it is the ethnic composition of the LAD. Respiratory illness shows the most consistency in the ratios whilst the comorbidity shows least consistency, having greater variability. Excepting the re-sizing ratio, the health improvements associated with the health microsimulation (RH) embody some of the largest variation in impacts.

10.5 DECOMPOSITION FOR CASE STUDY AUTHORITIES

In this section potentially different trends in the decomposition scenarios for LADs are explored in detail using the ratios for each morbidity in the seven case study authorities. Insight into these trends is gained by making comparisons with the national trends shown in Table 10-4, Table 10-6, Table 10-8 and Table 10-10 and by highlighting some contrasting features amongst the LADs.

10.5.1 Leeds (Cities and Services)

In the tables for Leeds the four critical demographic scenario multipliers (RG, RE, RN and RM) are not dissimilar to those for England. The shift to a more diverse ethnic population increases those with a morbidity and the changing age structure does likewise (although there is a slight dip from 2017 to 2021 for CVD, Table 10-11, and for DHBS, Table 10-12, and the comorbidity, Table 10-14). As the life expectancy of the population increases, the number of individuals with a morbidity increases. The improvement in health through the mechanisms covered in the health microsimulation occurs for all the morbidities by 2031, and more so than for England. In particular for DHBS, there is a reduction in those with this morbidity which is in contrast to the situation for England as a whole, where it the ratio is still greater than 1.00 by 2031.

Table 10-11 : Decomposing multipliers for CVD, Leeds, 2011 to 2031

CVD	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.03	1.00	1.00	1.00	1.00	0.90	0.93
2015	1.05	1.00	1.00	1.00	1.01	0.82	0.87
2017	1.08	1.00	1.00	0.99	1.02	0.74	0.81
2019	1.10	1.00	1.01	0.99	1.03	0.67	0.76
2021	1.12	1.00	1.01	0.99	1.05	0.61	0.72
2023	1.14	1.00	1.01	1.00	1.06	0.54	0.66
2025	1.14	1.01	1.01	1.02	1.08	0.48	0.61
2027	1.15	1.01	1.02	1.04	1.09	0.42	0.56
2029	1.16	1.01	1.02	1.06	1.11	0.37	0.52
2031	1.17	1.01	1.02	1.07	1.13	0.33	0.48

Table 10-12 : Decomposing multipliers for DHBS, Leeds, 2011 to 2031

Diabetes or high blood sugar	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.03	1.00	1.01	1.00	1.00	1.02	1.06
2015	1.05	1.00	1.02	1.00	1.01	1.05	1.14
2017	1.08	1.00	1.03	0.99	1.01	1.07	1.19
2019	1.10	1.00	1.04	0.99	1.02	1.06	1.22
2021	1.12	1.00	1.05	0.99	1.03	1.06	1.27
2023	1.14	1.00	1.06	0.99	1.04	1.02	1.27
2025	1.14	1.00	1.07	1.00	1.05	1.01	1.29
2027	1.15	1.00	1.08	1.01	1.06	0.98	1.30
2029	1.16	1.00	1.09	1.02	1.08	0.95	1.32
2031	1.17	1.00	1.11	1.03	1.08	0.91	1.31

Table 10-13 : Decomposing multipliers for respiratory illness, Leeds, 2011 to 2031

Respiratory Illness	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.03	1.00	1.00	1.00	1.00	0.94	0.97
2015	1.05	1.00	1.00	1.00	1.01	0.92	0.98
2017	1.08	1.00	1.00	1.00	1.02	0.88	0.97
2019	1.10	1.00	1.00	1.00	1.02	0.84	0.94
2021	1.12	1.00	1.00	1.00	1.03	0.83	0.96
2023	1.14	1.00	1.00	1.00	1.03	0.78	0.92
2025	1.15	1.00	1.00	1.00	1.04	0.76	0.91
2027	1.15	1.00	1.00	1.01	1.05	0.72	0.88
2029	1.16	1.00	1.00	1.02	1.06	0.68	0.85
2031	1.17	1.00	1.00	1.03	1.07	0.66	0.85

Table 10-14 : Decomposing multipliers for comorbidity of CVD and DHBS, Leeds, 2011 to 2031

CVD and diabetes or high blood sugar	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.03	1.00	1.02	1.01	1.01	0.90	0.96
2015	1.05	1.00	1.03	1.00	1.02	0.84	0.93
2017	1.07	1.01	1.03	0.99	1.03	0.77	0.87
2019	1.10	1.01	1.05	0.98	1.04	0.70	0.83
2021	1.12	1.01	1.06	0.99	1.05	0.64	0.80
2023	1.13	1.01	1.07	1.01	1.07	0.55	0.73
2025	1.14	1.01	1.08	1.02	1.08	0.50	0.68
2027	1.15	1.01	1.10	1.03	1.09	0.43	0.62
2029	1.16	1.01	1.11	1.06	1.11	0.38	0.58
2031	1.17	1.01	1.12	1.07	1.12	0.33	0.52

10.5.2 Redbridge (London Suburb)

The trends in Redbridge show some contrasting features to those for England as a whole. Firstly the scenario that adapts to the projected gender and ethnic composition (RE) estimates a much greater impact for this scenario, and this is true for all three morbidities and the comorbidity. Clearly the changing ethnic makeup of this LAD is an important influence on the number of individuals with each morbidity. This is contrasted with sharp reductions in the number of individuals with a morbidity when the changes incorporated into the health microsimulation are implemented. In particular, like Leeds, the mechanisms within the health microsimulation that determine the prevalence counts of DHBS in Redbridge results in a reduction in those with such a morbidity in 2031 compared with 2011 (Table 10-16).

Table 10-15 : Decomposing multipliers for CVD, Redbridge, 2011 to 2031

CVD	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.05	1.00	1.02	0.99	1.01	0.88	0.94
2015	1.08	1.00	1.03	0.99	1.02	0.80	0.90
2017	1.12	1.00	1.04	0.98	1.02	0.73	0.85
2019	1.14	1.00	1.06	0.98	1.03	0.64	0.78
2021	1.18	1.00	1.07	0.97	1.05	0.58	0.75
2023	1.21	1.00	1.09	0.97	1.06	0.49	0.66
2025	1.25	1.01	1.11	0.98	1.07	0.44	0.65
2027	1.28	1.01	1.13	0.98	1.09	0.39	0.61
2029	1.33	1.01	1.15	0.99	1.10	0.33	0.56
2031	1.38	1.01	1.17	0.99	1.12	0.30	0.54

Table 10-16 : Decomposing multipliers for DHBS, Redbridge, 2011 to 2031

Diabetes or high blood sugar	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.03	1.00	1.03	0.99	1.01	1.07	1.14
2015	1.06	1.00	1.06	0.98	1.02	1.11	1.25
2017	1.10	1.00	1.09	0.97	1.02	1.12	1.33
2019	1.13	1.00	1.11	0.97	1.03	1.11	1.39
2021	1.16	1.00	1.14	0.97	1.04	1.11	1.48
2023	1.19	1.01	1.16	0.98	1.05	1.07	1.54
2025	1.23	1.01	1.19	0.99	1.06	1.05	1.63
2027	1.27	1.01	1.21	0.99	1.07	1.01	1.66
2029	1.31	1.00	1.23	1.01	1.08	0.96	1.69
2031	1.36	1.00	1.25	1.02	1.08	0.92	1.72

Table 10-17 : Decomposing multipliers for respiratory illness, Redbridge, 2011 to 2031

Respiratory Illness	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.04	1.01	1.01	1.01	1.00	0.89	0.95
2015	1.07	1.00	1.01	1.01	1.01	0.84	0.93
2017	1.11	1.00	1.03	1.02	1.02	0.78	0.93
2019	1.14	1.00	1.04	1.02	1.02	0.71	0.88
2021	1.18	1.00	1.05	1.02	1.04	0.67	0.88
2023	1.21	1.00	1.07	1.02	1.05	0.61	0.85
2025	1.24	1.00	1.09	1.02	1.06	0.57	0.83
2027	1.28	1.00	1.11	1.02	1.07	0.52	0.81
2029	1.33	1.00	1.13	1.02	1.08	0.47	0.78
2031	1.38	0.99	1.15	1.02	1.10	0.45	0.79

Table 10-18 : Decomposing multipliers for comorbidity of CVD and DHBS, Redbridge, 2011 to 2031

CVD and diabetes or high blood sugar	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.05	1.01	1.06	0.98	1.00	0.93	1.02
2015	1.08	1.01	1.09	0.97	1.02	0.85	1.00
2017	1.12	1.02	1.13	0.95	1.02	0.77	0.96
2019	1.15	1.01	1.17	0.96	1.03	0.67	0.90
2021	1.18	1.01	1.22	0.95	1.04	0.60	0.86
2023	1.21	1.01	1.25	0.98	1.05	0.50	0.79
2025	1.25	1.02	1.29	0.99	1.06	0.45	0.78
2027	1.29	1.02	1.34	0.99	1.07	0.39	0.73
2029	1.32	1.02	1.38	1.01	1.08	0.32	0.65
2031	1.38	1.02	1.42	1.01	1.09	0.30	0.66

10.5.3 Camden (London Centre)

For Camden perhaps the most critical impact is the large ratios for the health microsimulation with DHBS (Table 10-20). By 2031 the ratio is far larger than that seen for England as a whole, and this increase is seen to be most rapid in the early period 2011 to 2019. All the London Centre type LADs exhibit this feature in the health trend, implying large increases in those with this morbidity. A further consequence of this high ratio for DHBS is that the ratio for the comorbidity is also higher for Camden than that for England.

Table 10-19 : Decomposing multipliers for CVD, Camden, 2011 to 2031

CVD	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.05	1.00	1.01	1.01	1.00	0.94	1.01
2015	1.09	1.00	1.02	1.00	1.01	0.89	1.00
2017	1.13	1.01	1.02	0.99	1.02	0.84	0.99
2019	1.17	1.01	1.03	0.99	1.03	0.77	0.96
2021	1.21	1.01	1.03	0.99	1.04	0.72	0.93
2023	1.25	1.00	1.04	1.00	1.05	0.64	0.87
2025	1.29	1.01	1.05	1.00	1.06	0.60	0.87
2027	1.32	1.01	1.06	1.01	1.08	0.54	0.83
2029	1.36	1.00	1.06	1.02	1.09	0.48	0.77
2031	1.41	1.00	1.07	1.02	1.10	0.44	0.74

Table 10-20 : Decomposing multipliers for DHBS, Camden, 2011 to 2031

Diabetes or high blood sugar	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.05	1.00	1.02	0.98	1.00	1.24	1.30
2015	1.09	1.00	1.03	0.98	1.01	1.42	1.58
2017	1.14	1.00	1.04	0.97	1.02	1.56	1.83
2019	1.17	1.00	1.05	0.96	1.02	1.63	1.96
2021	1.21	1.00	1.07	0.97	1.03	1.68	2.17
2023	1.26	1.00	1.08	0.97	1.04	1.68	2.31
2025	1.30	1.00	1.10	0.97	1.05	1.69	2.46
2027	1.33	1.00	1.12	0.98	1.06	1.67	2.58
2029	1.36	1.00	1.13	0.99	1.06	1.63	2.63
2031	1.41	1.00	1.14	1.00	1.07	1.59	2.73

Table 10-21 : Decomposing multipliers for respiratory illness, Camden, 2011 to 2031

Respiratory Illness	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.04	1.00	1.00	1.00	1.01	0.92	0.97
2015	1.09	1.00	1.00	1.00	1.01	0.89	0.98
2017	1.13	1.00	1.01	1.00	1.02	0.83	0.97
2019	1.17	1.00	1.01	1.00	1.03	0.76	0.93
2021	1.21	1.00	1.02	1.00	1.04	0.73	0.94
2023	1.25	1.00	1.02	1.00	1.05	0.67	0.90
2025	1.29	1.01	1.03	0.99	1.05	0.65	0.91
2027	1.32	1.01	1.04	1.00	1.06	0.60	0.88
2029	1.35	1.01	1.05	1.00	1.07	0.56	0.86
2031	1.40	1.00	1.07	1.00	1.09	0.53	0.87

Table 10-22 : Decomposing multipliers for comorbidity of CVD and DHBS, Camden, 2011 to 2031

CVD and diabetes or high blood sugar	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.05	0.99	1.05	1.00	0.99	0.96	1.04
2015	1.09	0.99	1.07	0.99	1.01	0.97	1.12
2017	1.13	1.00	1.09	0.98	1.03	0.94	1.17
2019	1.17	1.00	1.10	0.98	1.02	0.88	1.13
2021	1.20	1.01	1.13	0.99	1.03	0.85	1.19
2023	1.25	1.01	1.16	1.00	1.03	0.76	1.15
2025	1.28	1.01	1.19	1.01	1.03	0.73	1.17
2027	1.31	1.00	1.22	1.02	1.04	0.66	1.12
2029	1.35	1.00	1.25	1.02	1.04	0.59	1.06
2031	1.40	1.00	1.27	1.04	1.05	0.55	1.07

10.5.4 Haringey (London Cosmopolitan)

Looking at the results in Chapter 9, Haringey is a LAD whose health outcomes are either good or forecast to improve. Looking at the gender, ethnicity and age ratios (RN) in Table 10-23 and Table 10-24 the impacts here are seen to be less than that for England as a whole, which shows that in this LAD there are no large increases in the numbers with these two morbidities as a result of changes to its age structure – Haringey is projected to retain a younger over 50 population relative to other LADs over the forecast period. However, it is clear in all these tables that these overall health improvements mostly derive from the impacts of the dynamics within the health microsimulation, which are consistently lower than those for England as a whole.

Table 10-23 : Decomposing multipliers for CVD, Haringey, 2011 to 2031

CVD	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.07	1.00	1.00	1.00	0.99	0.89	0.94
2015	1.14	1.00	1.01	0.99	1.00	0.79	0.90
2017	1.20	1.00	1.01	0.99	1.00	0.70	0.84
2019	1.25	1.00	1.02	0.99	1.01	0.62	0.79
2021	1.31	1.00	1.02	0.98	1.02	0.55	0.73
2023	1.36	1.00	1.03	0.99	1.03	0.45	0.64
2025	1.40	1.01	1.04	1.00	1.04	0.41	0.63
2027	1.45	1.01	1.05	1.01	1.05	0.35	0.57
2029	1.50	1.01	1.06	1.02	1.06	0.30	0.52
2031	1.56	1.00	1.07	1.03	1.07	0.26	0.48

Table 10-24 : Decomposing multipliers for DHBS, Haringey, 2011 to 2031

Diabetes or high blood sugar	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.07	1.00	1.01	0.96	1.00	1.03	1.07
2015	1.13	0.99	1.02	0.95	1.00	1.06	1.15
2017	1.19	0.99	1.04	0.94	1.01	1.08	1.26
2019	1.25	0.98	1.05	0.93	1.01	1.09	1.32
2021	1.30	0.98	1.06	0.94	1.01	1.07	1.37
2023	1.35	0.98	1.06	0.94	1.02	1.02	1.37
2025	1.39	0.98	1.07	0.95	1.03	1.01	1.44
2027	1.44	0.98	1.08	0.97	1.04	0.98	1.51
2029	1.49	0.98	1.09	0.98	1.05	0.93	1.52
2031	1.56	0.98	1.10	0.99	1.05	0.88	1.54

Table 10-25 : Decomposing multipliers for respiratory illness, Haringey, 2011 to 2031

Respiratory Illness	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.07	1.00	1.01	0.99	0.99	0.93	0.99
2015	1.13	1.00	1.00	0.99	1.00	0.89	1.00
2017	1.19	1.00	1.00	0.98	1.00	0.84	0.98
2019	1.24	1.00	1.01	0.98	1.01	0.80	0.99
2021	1.30	0.99	1.01	0.98	1.01	0.77	0.99
2023	1.35	0.99	1.02	0.98	1.02	0.70	0.95
2025	1.39	0.99	1.02	0.99	1.03	0.67	0.96
2027	1.43	0.99	1.03	1.00	1.03	0.63	0.95
2029	1.48	0.99	1.04	1.00	1.04	0.58	0.92
2031	1.55	0.99	1.05	1.01	1.04	0.55	0.93

Table 10-26 : Decomposing multipliers for comorbidity of CVD and DHBS, Haringey, 2011 to 2031

CVD and diabetes or high blood sugar	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.05	0.99	1.02	0.98	0.99	0.88	0.91
2015	1.11	0.98	1.05	0.97	1.00	0.79	0.88
2017	1.17	0.98	1.06	0.97	1.01	0.70	0.83
2019	1.22	0.98	1.08	0.98	1.00	0.62	0.78
2021	1.28	0.97	1.09	0.99	1.02	0.54	0.74
2023	1.33	0.97	1.10	1.01	1.03	0.43	0.63
2025	1.37	0.98	1.10	1.04	1.04	0.38	0.61
2027	1.42	0.98	1.11	1.06	1.04	0.33	0.56
2029	1.47	0.98	1.13	1.07	1.06	0.27	0.50
2031	1.53	0.97	1.14	1.09	1.07	0.23	0.45

10.5.5 Wiltshire (Prospering UK)

In contrast to the previous three London LADs the changing ethnic structure of Wiltshire hardly impacts on the number of individuals with a morbidity, the gender and ethnicity (RE) ratios are close to 1.00 and lower than those for England as a whole. The gender, ethnicity and age (RN) ratios are conversely larger than for England as a whole, which suggests that in contrast to the London LADs, the ageing of the 50 and older population in Wiltshire causes a disproportionate increase in those with morbidities. Again the health improvements implemented in the health microsimulation have a net effect of reducing those with each morbidity by 2031 (but only just so for DHBS, Table 10-28).

Table 10-27 : Decomposing multipliers for CVD, Wiltshire, 2011 to 2031

CVD	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.05	1.00	1.00	1.01	1.00	0.94	1.00
2015	1.10	1.00	1.01	1.02	1.01	0.88	1.01
2017	1.14	1.00	1.01	1.02	1.02	0.82	0.98
2019	1.18	1.00	1.01	1.03	1.03	0.75	0.95
2021	1.21	1.00	1.01	1.04	1.05	0.70	0.93
2023	1.24	0.99	1.02	1.06	1.07	0.62	0.88
2025	1.25	0.99	1.02	1.09	1.09	0.57	0.85
2027	1.27	0.99	1.02	1.12	1.11	0.51	0.81
2029	1.27	0.99	1.02	1.15	1.13	0.46	0.77
2031	1.28	0.99	1.02	1.17	1.15	0.42	0.73

Table 10-28 : Decomposing multipliers for DHBS, Wiltshire, 2011 to 2031

Diabetes or high blood sugar	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.05	1.00	1.00	1.00	1.00	1.04	1.09
2015	1.10	1.00	1.00	1.01	1.01	1.08	1.21
2017	1.14	1.00	1.00	1.00	1.02	1.10	1.28
2019	1.18	1.00	1.01	1.01	1.02	1.10	1.35
2021	1.21	1.00	1.01	1.02	1.03	1.10	1.41
2023	1.24	1.00	1.01	1.03	1.04	1.07	1.44
2025	1.26	1.00	1.01	1.04	1.06	1.06	1.49
2027	1.27	1.00	1.02	1.05	1.08	1.04	1.53
2029	1.27	0.99	1.02	1.06	1.09	1.01	1.50
2031	1.28	0.99	1.02	1.08	1.10	0.99	1.52

Table 10-29 : Decomposing multipliers for respiratory illness, Wiltshire, 2011 to 2031

Respiratory Illness	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.05	1.00	1.00	1.00	1.00	0.96	1.01
2015	1.10	1.00	1.00	1.00	1.00	0.93	1.02
2017	1.14	1.00	1.00	1.00	1.01	0.91	1.05
2019	1.18	1.00	1.00	1.01	1.02	0.88	1.07
2021	1.21	1.00	1.00	1.01	1.03	0.87	1.10
2023	1.24	1.00	1.00	1.01	1.04	0.83	1.08
2025	1.25	1.00	1.01	1.01	1.05	0.81	1.08
2027	1.26	1.00	1.01	1.02	1.06	0.78	1.07
2029	1.27	1.00	1.01	1.03	1.07	0.74	1.05
2031	1.28	1.00	1.01	1.03	1.08	0.73	1.05

Table 10-30 : Decomposing multipliers for comorbidity of CVD and DHBS, Wiltshire, 2011 to 2031

CVD and diabetes or high blood sugar	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.06	1.00	1.00	1.02	1.00	0.94	1.02
2015	1.10	1.00	1.01	1.02	1.02	0.89	1.03
2017	1.15	0.99	1.01	1.02	1.02	0.84	1.00
2019	1.18	1.00	1.01	1.03	1.03	0.79	1.00
2021	1.22	0.99	1.01	1.04	1.04	0.74	0.98
2023	1.24	0.99	1.01	1.07	1.06	0.65	0.91
2025	1.26	0.99	1.02	1.09	1.08	0.60	0.90
2027	1.27	0.99	1.02	1.11	1.10	0.54	0.85
2029	1.27	0.99	1.02	1.14	1.11	0.49	0.80
2031	1.28	0.99	1.02	1.15	1.14	0.46	0.78

10.5.6 Teignbridge (Coastal and Countryside)

Like Wiltshire, changing the gender and ethnic composition of Teignbridge has very little impact on the number of individuals with the morbidities, the gender and ethnicity ratio (RE) is consistently close to 1.00. It is the age and mortality related ratios (RN and RM) that show the largest ratios and these are mostly greater than those for England as a whole. Here only for CVD (Table 10-31) and the comorbidity (Table 10-34) does the health microsimulation reverse some of these demographic health trends. For DHBS the health microsimulation has a neutral impact by 2031 (Table 10-32) and the health improvement for respiratory illness is meagre (Table 10-33), certainly in comparison to England as a whole and the other case study authorities.

Table 10-31 : Decomposing multipliers for CVD, Teignbridge, 2011 to 2031

CVD	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.04	1.00	1.00	1.00	1.00	0.94	0.98
2015	1.07	1.01	1.00	1.00	1.01	0.89	0.97
2017	1.10	1.01	1.00	1.00	1.02	0.83	0.94
2019	1.12	1.01	1.00	1.01	1.04	0.77	0.91
2021	1.15	1.01	1.00	1.02	1.06	0.71	0.89
2023	1.16	1.01	1.00	1.04	1.08	0.65	0.86
2025	1.17	1.01	1.00	1.07	1.10	0.60	0.83
2027	1.17	1.01	1.00	1.09	1.12	0.54	0.78
2029	1.17	1.01	1.00	1.12	1.13	0.49	0.73
2031	1.18	1.01	1.00	1.14	1.15	0.45	0.70

Table 10-32 : Decomposing multipliers for DHBS, Teignbridge, 2011 to 2031

Diabetes or high blood sugar	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.04	1.00	1.00	1.00	1.00	1.03	1.07
2015	1.07	1.00	1.00	1.00	1.01	1.07	1.16
2017	1.11	1.00	1.00	1.00	1.03	1.09	1.25
2019	1.13	1.00	1.00	1.00	1.04	1.08	1.27
2021	1.15	1.00	1.00	1.01	1.05	1.09	1.33
2023	1.16	1.00	1.00	1.01	1.07	1.06	1.33
2025	1.17	1.00	1.00	1.03	1.08	1.05	1.37
2027	1.17	1.00	1.00	1.03	1.09	1.05	1.38
2029	1.17	1.00	1.00	1.05	1.10	1.02	1.38
2031	1.18	1.00	1.00	1.06	1.12	1.00	1.40

Table 10-33 : Decomposing multipliers for respiratory illness, Teignbridge, 2011 to 2031

Respiratory Illness	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.04	1.00	1.00	1.00	1.00	1.01	1.05
2015	1.08	1.00	1.00	1.00	1.01	1.03	1.12
2017	1.10	1.00	1.00	1.00	1.01	1.03	1.14
2019	1.13	1.00	1.00	1.00	1.02	1.02	1.18
2021	1.15	0.99	1.01	1.00	1.03	1.04	1.23
2023	1.16	0.99	1.01	1.01	1.04	1.02	1.24
2025	1.17	0.99	1.01	1.01	1.05	1.01	1.25
2027	1.17	0.99	1.01	1.01	1.07	1.00	1.26
2029	1.17	0.99	1.01	1.02	1.08	0.96	1.24
2031	1.18	0.99	1.01	1.02	1.10	0.96	1.27

Table 10-34 : Decomposing multipliers for comorbidity of CVD and DHBS, Teignbridge, 2011 to 2031

CVD and diabetes or high blood sugar	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.04	1.00	0.99	1.01	1.00	0.95	0.99
2015	1.08	1.01	0.99	1.00	1.02	0.92	1.01
2017	1.12	1.00	0.99	1.00	1.03	0.87	0.99
2019	1.14	1.01	0.98	1.01	1.05	0.82	0.98
2021	1.17	1.00	0.98	1.02	1.07	0.76	0.95
2023	1.18	1.01	0.98	1.03	1.10	0.68	0.90
2025	1.18	1.01	0.98	1.05	1.12	0.64	0.88
2027	1.19	1.01	0.98	1.07	1.13	0.58	0.83
2029	1.18	1.01	0.98	1.10	1.14	0.53	0.78
2031	1.20	1.00	0.97	1.12	1.16	0.50	0.76

10.5.7 Rotherham (Mining and Manufacturing)

Rotherham is perhaps the most typical of England as a whole of these seven LADs in terms of its outcome ratios. By 2031, and for many of the intervening years, the values for the ratios compare well with those for England. The only exception to this is the outcome for DHBS which peaks at a higher value for 2021 (Table 10-36), and consequently falls away to a value that is still higher than for England by 2031. This high mid period peak for DHBS also contributes to a lower 2031 reduction for the comorbidity (Table 10-38).

Table 10-35 : Decomposing multipliers for CVD, Rotherham, 2011 to 2031

CVD	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.03	1.00	1.00	1.01	1.00	0.93	0.97
2015	1.06	1.00	1.00	1.02	1.00	0.87	0.94
2017	1.08	1.00	1.00	1.02	1.01	0.81	0.90
2019	1.11	1.00	1.01	1.03	1.02	0.74	0.87
2021	1.13	1.01	1.01	1.04	1.04	0.69	0.86
2023	1.14	1.01	1.01	1.05	1.05	0.61	0.78
2025	1.14	1.01	1.01	1.08	1.07	0.56	0.75
2027	1.13	1.01	1.02	1.11	1.09	0.50	0.70
2029	1.13	1.01	1.02	1.13	1.11	0.45	0.66
2031	1.13	1.01	1.03	1.15	1.12	0.41	0.62

Table 10-36 : Decomposing multipliers for DHBS, Rotherham, 2011 to 2031

Diabetes or high blood sugar	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.03	1.00	1.00	1.01	1.00	1.07	1.11
2015	1.06	1.00	1.01	1.01	1.01	1.13	1.23
2017	1.09	1.00	1.01	1.01	1.02	1.16	1.32
2019	1.11	1.00	1.02	1.01	1.02	1.19	1.39
2021	1.14	1.00	1.02	1.01	1.03	1.21	1.46
2023	1.15	1.00	1.03	1.03	1.05	1.18	1.51
2025	1.15	1.00	1.04	1.04	1.06	1.18	1.56
2027	1.14	1.00	1.04	1.05	1.07	1.17	1.56
2029	1.13	1.00	1.05	1.06	1.08	1.13	1.53
2031	1.13	1.00	1.07	1.06	1.10	1.11	1.56

Table 10-37 : Decomposing multipliers for respiratory illness, Rotherham, 2011 to 2031

Respiratory Illness	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.03	1.00	1.00	1.00	1.00	0.95	0.98
2015	1.06	1.00	1.00	1.01	1.00	0.93	1.00
2017	1.09	1.00	1.00	1.01	1.00	0.91	1.00
2019	1.11	1.00	1.00	1.01	1.01	0.88	1.00
2021	1.14	1.00	1.00	1.01	1.01	0.87	1.01
2023	1.15	1.00	1.00	1.01	1.02	0.83	0.98
2025	1.15	1.00	1.00	1.02	1.03	0.82	0.99
2027	1.14	1.00	1.00	1.02	1.04	0.79	0.96
2029	1.14	1.00	1.01	1.03	1.05	0.76	0.95
2031	1.14	1.00	1.01	1.04	1.06	0.74	0.94

Table 10-38 : Decomposing multipliers comorbidity of CVD and DHBS, Rotherham, 2011 to 2031

CVD and diabetes or high blood sugar	Re-size (RS)	Gender (RG)	Gender and Ethnicity (RE)	Gender Ethnicity and Age (RN)	Mortality Improve-ment (RM)	Health Micro-simulation (RH)	Product
2011	1.00	1.00	1.00	1.00	1.00	1.00	1.00
2013	1.03	1.00	1.01	1.01	1.00	0.95	1.00
2015	1.06	1.00	1.02	1.01	1.01	0.92	1.01
2017	1.09	1.00	1.02	1.01	1.03	0.85	0.98
2019	1.11	1.00	1.03	1.01	1.04	0.80	0.96
2021	1.13	1.00	1.04	1.02	1.05	0.76	0.96
2023	1.14	1.00	1.05	1.03	1.07	0.67	0.88
2025	1.14	1.00	1.06	1.05	1.09	0.63	0.87
2027	1.13	1.01	1.07	1.06	1.11	0.57	0.82
2029	1.13	1.01	1.09	1.08	1.12	0.51	0.77
2031	1.13	1.01	1.11	1.09	1.13	0.46	0.72

10.5.8 Summary

The results discussed in this section show that it is a combination of the changing demographics of the population of England and the health trends over time associated with these demographics that determine the health outcomes. The demographic scenarios are incremental in their application. All other things being equal, little of the forecast changes in health outcomes can be attributed the changing gender structure of the population in each LAD. This is not however true for ethnicity, ageing or changes in life expectancies. For some more ethnically diverse LADs the changing ethnic structure of the LAD can indicate large increases in those with particular morbidities, although with these same LADs the increases associated with age and life expectancies are less pronounced. Conversely there are other less ethnically diverse LADs where the increase attributable to a changing ethnic mix is negligible but most of the change can be attributed to an ageing of the population and increased life expectancies.

Clearly the health microsimulation is a complex process, with influences on health outcomes taking place through the incidence models and the models that determine the selection weight for morbidity or migration. For CVD, respiratory illness and the comorbidity of CVD and DHBS, nearly all the health improvements can be attributed to the mechanisms inside the health microsimulation – although this assertion is borderline for respiratory illness in some Coastal and Countryside type authorities. In contrast, for DHBS, in most authorities by 2031 the health microsimulation forecasts an increase in the number of people with this morbidity, even after accounting for the demographic changes, but this influence appears to be on a downward trend by 2031.

10.6 CONCLUSION

In this chapter an investigation into the relative contribution of various demographic and health related factors in influencing the number of individuals with each morbidity is provided. This is achieved by utilising the same C program in all scenarios but varying the input data and the functionality within the program. The input data are synthesised using a combination of adjustment factors and an optimisation to meet a set of demographic constraints. These changes to the input data define the five demographic scenarios. Of critical importance to the implementation of these demographic scenarios is that the mechanisms associated with morbidity incidence, mortality selection and migration selection are not in place.

Generally the demographic factors are seen to increase the number of individuals with the morbidities, although the relative impact of each demographic factor varies by LAD. For those LADs with diverse ethnic communities it is ethnicity that is seen to have the

greatest influence whilst for those with less ethnic diversity, changing age structures and improvements in life expectancy have greatest influence.

The impact of the health microsimulation for CVD, respiratory illness and the comorbidity of CVD and DHBS reduces the numbers with the morbidity. The net impact of the demographic and health improvement scenarios for these morbidities is a reduction in those with the morbidity. For DHBS, the health microsimulation generally increases those with the morbidity and, when combined with the demographic factors, the number with this morbidity increases over time.

Just as the five demographic factors accumulate to potentially large impacts (i.e. RS x RG x RE x RN x RM) the health microsimulation in itself has a large impact. To an extent the drivers behind these impacts can be inferred using the parameter estimates in the morbidity, mortality and migration hazard models. Such an investigation is not however possible for the demographic factors (where there are no model parameters available) so the utility of this chapter remains in its consideration of the five demographic scenarios and the health microsimulation impact merely completes the picture.

This chapter is the final chapter of the five results chapters of this thesis. In the remaining chapter a summary of the work, its main conclusions, the policy implications of these findings and the recommendations for future work are covered.

11 CONCLUSION

Thesis Summary, Policy Implications and Further Work

11.1 INTRODUCTION

In this thesis a body of work is presented that forecasts the number of individuals in the 50 year and older population of each Local Authority District (LAD) in England who have one or more of three specific morbidities. These morbidities are CVD, DHBS and respiratory illness. Together they are cited as the main cause of nearly 50% of all deaths in the 50 and older population of England and more than 20% of hospital admissions. In addition the numbers of individuals with the commonly diagnosed comorbidity of CVD and DHBS are also provided. The significance of this work is driven by a realisation that in western societies, an ageing population is almost inevitable and that the resources available to care for these individuals are limited. By better understanding this phenomenon, policies and plans can be put in place to help meet the populations' health care needs.

In section 11.2 of this chapter the material presented in the earlier chapters is summarised and the findings from the work highlighted. The policy implications that arise from the work are covered in section 11.3 which addresses the final research question in section 1.4. Section 11.4 illustrates how the work may be improved and taken forward. The final section describes how the objectives set out in chapter one of this thesis have been met

11.2 SUMMARY OF WORK

11.2.1 Demographics

There is a consistent pattern amongst Western European nations of a projected increase in the size and proportion of their older population. However, the rate of this ageing varies by nation and the UK is one of those where this phenomenon is less pronounced. Using sub national population projections provided at the LAD level by ONS it is also possible to examine this phenomenon more locally. Unsurprisingly, the most populous LADs, either through their large geographic size or dense urban populations are those with the largest population aged 50 and older, and also show the greatest increases from 2011 to 2031. However, when measured as percentage changes, these LADs with larger 50 and older populations in 2011 show smaller percentage increases in comparison to other LADs. Thus the relative impact of an ageing population is more keenly felt on small to midsize LADs.

Coupled with the size of this demographic, the health status of the population will also have an influence on the total health care demand. Both life expectancies and disability free life expectancies show historic increases, and whilst people are living longer, the age to which that they are living without a disability is also increasing. These increases are not however occurring at the same rate, so that the difference between these two measures appears to be increasing. These points to an expansion in morbidity. Females tend to have higher life expectancies than males but similar disability free life expectancies, which means that females spend most of these extra years of life living with a disability. These life and disability free life expectancies can be estimated for local areas too. When mapped at the Clinical Commissioning Group (CCG) level they show distinct regional variation. Areas in the Midlands and Northern England (particularly the urban areas) have low life and disability free life expectancies, whilst areas in the South, outside London, (particularly rural areas) have high expectancies.

This concentration of an ageing population and populations with lower disability free life expectancies in urban areas is corroborated when looking at demand for hospital care, measured using FCEs. As well as a concentration of demand in the urban areas, the provision of hospital health care is also concentrated in the cities within such areas.

The general context of the aged population in England is a population which is getting larger, older and living longer but not necessarily healthier.

11.2.2 Literature

The health care demands associated with an elderly population and what determinants affect this demand is a well-studied field. Some studies examine the prevalence of specific morbidities, others look at either general health or presence of limiting activities and many (particularly those from the United States) model health care costs as a proxy for ill health. The general findings are that as people get older they are more prone to develop morbidities or be in generally worse health which means that an older population will tend to have worse health. However, age itself may not be the actual influence. Some studies have reported that it is the remaining years of life that are important, with an individual's health deteriorating in the final months or years before death. Others take the argument further and suggest that it is not age or the remaining years of life that are important but the presence of a disabling condition.

Gender and ethnic difference influence health status as well. As mentioned above, females live longer than males but these extra years are not necessarily spent in good health. There are some morbidities that are more associated with certain ethnic groups, e.g. DHBS in the south Asian population. A person's socio-economic status can be measured using income, wealth, education or employment indicators. These are all correlated and can all influence health.

A range of data exists to study the health outcomes of an aged population. Population Censuses provide extensive coverage of the characteristics of the population and provide rich detail on the many determinants often cited as important for health outcomes. However, the actual information collected on health is general and rarely touches on specific morbidities and the time interval between Censuses can be large. To try and overcome these issues, governments often commission national sample surveys that can either be generic or targeted on a particular public policy issue, such as health. Finally, bespoke, one-off surveys are commissioned by local government or academics that are designed to study a particular health outcome in a specific place. In an era where administrative systems are increasingly being implemented and coordinated, scope has arisen to use such data for researching health status.

The literature also contains details of the modelling approaches used in studies of elderly health care. The two main approaches are either statistical modelling or simulation. Statistical modelling is the most often used approach. The advantage of this approach is that it is grounded in statistical theory which allows for various interpretative and testing regimes to be followed. The disadvantage is that statistical modelling is rigid in both its outcome and the reliance on the assumptions that underlie the modelling technique. Just as are there are a wide variety of statistical techniques which can be used, there are also a wide variety of simulation techniques. These techniques attempt to replicate the composition or behaviour of a population, either real or hypothetical. They commonly use the technique of Monte Carlo simulation to replicate a decision making process between alternatives within the simulation system. The advantage of such an approach is that it can be built using simple rules that are informed by either an understanding of the dynamics of health care or informed by data. The drawback is that it is sometimes difficult to disentangle these dynamics, particularly when unexpected or little understood interactions occur.

Whilst most of the literature studied has focused on health care demand, there also related studies of health care delivery, with an increasing move to more pharmacological and technological treatments of morbidities delivered through GP clinics or at home. There are some concerns that such treatments many absent such cases from administrative recording mechanisms, although there are moves to capture such treatments e.g. the proposed release of GP patient data as part of the English NHS's care.data initiative, NHS England (2015).

11.2.3 Data

The review of the literature identified a number of useful data sources for this study. The two which emerged as having most utility were the 2011 UK population Census and ELSA. The 2011 Census is able to provide a comprehensive picture of the

population in England at a variety of geographic scales. The information captured by the Census includes most of those that are determinants identified in the literature as being important determinants. However, the range of actual health information in the Census is limited with no information on specific morbidities. In contrast, ELSA is able to provide the detailed information about individuals and their morbidities. It is also a longitudinal survey so it is possible to gain a contemporaneous understanding of their circumstances when an event of significance occurs, i.e. the acquisition of a morbidity. Whilst the sample size of ELSA is large, its coverage is still not sufficient to provide robust estimates of health status at geographies smaller than region.

Since both of these potentially useful data sources have their advantages and disadvantages, a modelling technique is required to best utilise the advantages of each data set. Before this is accomplished a consideration of the compatibility of information between the two sources is required. This compatibility found to be good in terms of the main socio-demographic measures, opening the way for using both data sets in the suggested modelling approach.

11.2.4 Methodology

The methodological approach adopted in this study utilises a variety of individual components. A spatial microsimulation model is used to create a synthetic base population of those aged 50 and older in each English LAD for 2011. The 2011 Census counts provide the constraint population and ELSA provides the sample population. This 2011 base population is then evolved through time, biennially, from 2011 to 2031. During this process the population is aged, is replenished at the younger ages of 50 and 51 and the morbidity status of individuals is updated. This updating is influenced by a range of hazard models that predict the probability of a morbidity, mortality or migration event occurring. Also information on the likely future smoking status of cohorts as they become part of future synthetic ELSA waves is derived. It is of critical importance that the size and structure of these future LAD populations evolve along the projected lines. This is achieved by constraining the populations for each forecast year to agree with an external population projection of gender by age by ethnicity. However, a comparison of these projections with 2011 Census counts highlighted the need to revise these projections in light of the mismatch between the Census counts and the 2011 projections. To gain further understanding of how the trends in morbidity status are influenced by the ‘natural’ demographic changes in the population, an accounting system to apportion changes in this status to various demographic factors is formulated and applied.

11.2.5 Spatial microsimulation results

The methodology requires that a synthetic base population for the year 2011 is established in each LAD. This is achieved by conducting a spatial microsimulation using 2011 Census data and wave 5 data from ELSA. A consideration of the commonality of variables between these two data sources identified a number of potentially useful constraint variables. To select those that were likely to be of use a multi-faceted strategy was adopted. Firstly, the variable needs to have a common definition and interpretation in the constraint and sampling population. This was achieved by consulting the relevant documentation for the variable and examining its coding. Secondly, the variable must have been identified in the literature to be an important determinant of elderly health status in aggregate. This is informed by the literature. Thirdly, this importance needed to be demonstrated at the individual level. Here this is tested using hazard models to explain morbidity incidence.

A set of initial constraint variables are identified, based on the literature, namely age, gender, disability status and ethnicity. Of these four, not surprisingly, disability status was the most consistently important, as measured by the reduction in deviance. The next most consistently important constraint is gender, which is important for the linked morbidities of CVD and DHBS but not respiratory illness. Age was important for CVD and finally, ethnicity is particularly important for DHBS, a finding which is strongly supported by the literature.

The next stage was to identify a number of variables to capture the socio-economic impact on morbidity incidence. Here there is a need to identify not only additional constraint variables but also validation variables to investigate how well the spatially microsimulated population reproduces counts for the validation variable. After examining the goodness of fit of a range of candidate variables, vehicle ownership was found to be the socio-economic variable that is most important in explaining morbidity incidence, and given the need to have a reliable validation outcome, it is used as a validation variable. The next most important variable is living arrangements, which is important for DHBS and respiratory illness. This used as an additional constraint variable. The third and final socio-economic variable to be used as an additional validation variable is social status (recorded through the NS-SEC).

The spatial microsimulation was performed using a combinatorial optimisation implementation with five constraint tables containing: age and gender, disability (household residents), disability (clients resident in institutional establishments), ethnicity and living arrangements. The optimisation performed well, quickly achieving low and stable values of the Total Absolute Error goodness of fit measure. The validation of the synthetically derived LAD population was seen to be generally good,

although the size of the ‘Other’ social class was consistently under estimated in the synthetic populations.

There is additional utility in the synthetic LAD populations from being able to estimate the prevalence of morbidities in the LAD for 2011, employing the information on whether each member of the ELSA sample has a morbidity at wave 5. This analysis revealed both ethnic and wealth gradients in the estimated prevalences. LADs with a large south Asian population that were also less affluent are seen to have high prevalences of both CVD and respiratory illness. Many of the LADs with high DHBS prevalence are located in central London and have a diverse ethnic composition. The more prosperous and affluent LADs in southern England have low prevalences for all three morbidities.

11.2.6 ETHPOP adjustment results

As part of the evolution of the base LAD population through time it is important that its size and structure accord with some likely population projection. Whilst the ONS regularly produce subnational population projections for LADs in England, they are only disaggregated by gender and age. These two demographic factors may not accurately reflect the changing nature of English LADs. A more credible outcome is possible if some additional demographic characteristic could also be incorporated. Probably the most useful and available is an ethnic dimension and this study has used such a projection that arose from a University of Leeds research project in the ESRC Understanding Population trends and Processes (UPTAP) initiative, branded ETHPOP.

An initial assessment of the performance of these projections for the 50 and older population against 2011 Census outcomes showed some differences. In the ethnically diverse LADs the ETHPOP projections tended to overestimate the White ethnic group and under estimate the BME group, in both absolute numbers and share of the LAD population. In LADs with a more homogeneous White population, the projections are much more accurate when looking at absolute numbers, but for the BME groups, which have small numbers, the percentage differences are large.

In light of these discrepancies between 2011 ETHPOP projections and 2011 Census counts, a methodology is deployed that attempts to revise projections beyond 2011 in light of this past performance. Firstly, any projections are re-based to have the 2011 Census counts as their starting point; secondly the ETHPOP projections beyond 2011 are then corrected by use of a linear adjustment factor and finally constrained to 2012 ONS sub national projections. The range of the actual adjustment factors used are limited by a distribution derived from corresponding factors in other similar types of authority. The impact of these revisions on 2021 and 2031 projections are discussed for a number of case study authorities.

11.2.7 Hazard models and probabilities

The health microsimulation is implemented using the Monte Carlo technique of randomly determining outcomes using draws from a probability distribution. In order to implement this technique a range of probabilities of certain events occurring need to be defined. The most obvious are the probabilities that measure whether an incidence of a morbidity will take place. These probabilities are estimated using a hazard model estimated on wave 1 to 6 ELSA data. These hazard models are also used to forecast the morbidity status of those aged 50 and 51 who replenish future artificial waves.

Hazard models are also use to estimate a set of weights to be used for selecting individuals for cloning or removal from the LAD population. It is this mechanism that re-structures the population to an external projection (here the revised ETHPOP projections). Using the hazard model probabilities allows for the chance of removal (due to mortality or out-migration) or inclusion (due to in-migration) to be influenced by a person's characteristics, including their morbidity status. Notwithstanding the utility of these weights in regards to determining who rather than how many individuals to select, a comparison with actual mortality and migration rates is insightful. This comparison is done with contemporary official mortality statistics and migration counts from the 2011 Census. The trends in the mortality predicted by the hazard models agree well with the official mortality probabilities but the level of morbidity is, for some types of individual, much less than these official probabilities. The migration probabilities are also less than those estimated from the 2011 Census and a compounding of probabilities is necessary to increase this agreement.

Not only do the Replenishers for the future artificial waves need to have a morbidity status on joining, they also need a smoking status. Using evidence from the 2011 HSfE, a tabulation of whether someone aged 30 to 50 had ever smoked revealed that there was little difference by birth cohort within this age range but there were gender and ethnic differences.

Together, these probability models and rates allow the health microsimulation to be dynamic in its health status and its composition over time.

11.2.8 Health microsimulation results

The outcome of the health microsimulation is an individual population that represents the projected population of each LAD biennially from 2011 to 2031. Each individual will have their own morbidity status and this information can be used to estimate both the prevalence count for morbidities, or a combination of morbidities, and also a prevalence rate. Aggregated national trends for England from the microsimulation show

a reduction in the prevalence and number of people with CVD and respiratory illness but an increase in the number and prevalence with DHBS.

To begin to explore how these prevalences and trends over time vary locally, the overall prevalence rates in each NHS Local Area Team (LAT) is provided. For CVD the national trends seen for reductions in the prevalence are apparent in all LATs although to varying degrees. The London LAT sees a substantial reduction in its prevalence rate and whilst the prevalence for Durham, Darlington and Tees also reduces by 2031, this reduction is more modest. For DHBS, there is much more stability in prevalence rates, all LATs show increases and, generally, LATs with high prevalences in 2011 also have high prevalences in 2031, and vice-versa, with Birmingham and the Black Country and the London LATs retaining high prevalences. The pattern with respiratory illness is similar to that for CVD. Whilst the reductions in prevalence are less pronounced, the London LAT again shows an improvement in both prevalence and its relative position amongst LATs whilst Devon, Cornwall and Isles of Scilly becomes the LAT with the highest prevalence. The comorbidity prevalences have elements from both the CVD and the DHBS trends, so whilst all the prevalences reduce, London improves its relative position, Birmingham and the Black Country remains one of the highest, and Durham, Darlington and Tees joins it as the LATs with the higher prevalences in 2031.

Inspection of the trends in the seven individual case study authorities corroborates some of the features reported for the LATs. Many of the case study authorities are based in London and so see the improvements in both their prevalence and position relative to all other LADs - as the LAT results would imply. Leeds also tends to improve its rankings for most morbidities, for reasons similar to those in London. The two more prosperous LADs, Wiltshire and Teignbridge, retain their relatively good position for most morbidities, but show deteriorations for respiratory illness, especially so for Teignbridge. The remaining LAD, Rotherham, exhibits changes in line with the national trends and a stability in terms of its relative performance amongst LADs.

These results predict that both nationally and locally the incidence and prevalence of CVD will fall up to 2031. For DHBS there will be increases in both the prevalence count and rate. The number of individuals with respiratory illness will remain largely constant. However, given the larger population, this will translate to a reduced prevalence rate. Within these national trends however there are geographic differences which are apparent at both the regional and local level.

11.2.9 Decomposition Results

Whilst the parameter estimates from the hazard models provide some insight into how the characteristics of the population in a LAD affect the health outcomes, the impact of relative shifts in demographic factors is not so clear. To gain this insight an accounting

system has been adopted to decompose the health trends attributable to these shifting demographics. On a national level these demographic impacts are seen to increase the number of individuals with morbidities. For CVD and respiratory illness and the comorbidity of CVD and DHBS these increases are reversed when the health trends in the health microsimulation are applied. For DHBS, however, over the forecast period 2011 to 2031, there is also an increase attributable to the health microsimulation.

Looking at the case study LADs, there is a distinct patterning amongst them. The more ethnically cosmopolitan LADs show the largest increases associated with changes in the ethnic composition of the LAD and not so much the changes in the age composition or life expectancy of their populations. For the more prosperous LADs the relative impacts are reversed, changing the ethnic composition of the LADs does not increase the number of individuals with a morbidity but changing the age composition and life expectancy does.

11.3 POLICY IMPLICATIONS

As discussed elsewhere in this thesis, for the foreseeable future, the English population is expected to age. As the House of Lords (2013) report says:

"The ageing of the population is inevitable, and affects us all."

The impact of this ageing has consequences for society, both financial and social. One sector where there is the potential for this ageing to be most significantly felt is health.

"The NHS is facing a major increase in demand and cost consequent on ageing and will have to transform to deal with this. Because of this rising demand, …, needs will remain unmet and cost pressures will rise inexorably." House of Lords (2013)

This research has demonstrated that there are indeed significant demographic drivers in society that would lead to the conclusion that there will be greater pressure on health care services linked to the three morbidities studied here. An older population which is more ethnically diverse, has an older structure and incorporates increases in life expectancies will mean that more people will have one or more of these morbidities. However all is not what it seems. There are other trends in the population that help to potentially mitigate or exaggerate these demographic influences. The shift away from types of employment associated with ill health (e.g. heavy industry, chemicals or textiles), better employment practices (health and safety regulations), reduced prevalence of smoking in the population and better air quality all have the potential to improve health at all ages but, given that many of these factors have their impacts later in life, more so at older ages. There are however trends working in the opposite direction that will lead to poorer health outcomes. The oft cited trends are related to diet and lifestyle. Increases in obesity seen at younger ages may affect later life health status,

especially if this obesity is carried forward into later life ("*Obesity is the new smoking*", Guardian, 2015).

The results presented in this thesis have corroborated these trends to a large degree. Of the two morbidities that show improvements in population health, CVD and respiratory illness, these are often cited as having links to those trends which are improving, e.g. occupation, housing, smoking, air quality and medications. In addition, medication with statins, first used in the UK in 1987 (Chatterjee, 2008), has been shown to reduce the level of LDL ('bad') cholesterol thereby avoiding occurrence of CVD (Law et al, 2003). The morbidity which shows an increase, both in those with the morbidity and its prevalence, is diabetes and this has a strong lifestyle link to obesity.

This suggests at least three public policy messages. Firstly, there are reassuring trends in health improvements with two of the largest causes of death and hospital admission showing substantial reductions. Public health and policy work in this area is clearly successful, having an impact and should continue. Secondly, more work is required with DHBS. On current trends the number of individuals with this morbidity and its prevalence are projected to increase. Clearly more attention is needed to try and identify and mitigate this trend. Factors that would be worthy of investigation include the targeting of public health messages at locations and communities where the growth in DHBS is anticipated to be greatest. To what extent the 'boat has sailed' on this issue (Singh et al, 2008) and that any public health intervention taken now is unlikely to yield significant impacts before 2031 is debatable but this should not stop work commencing in this area (Reilly, 2003). Thirdly, given the difference in forecast trends with these morbidities, in a period of constraints on resources available for public health, consideration should be given to tilting the balance of such resources towards DHBS.

There are distinct geographic trends in the future prevalence counts and rates. Examination of the dynamics of both prevalence counts and rates in NHS Local Area Teams (LAT) (Table 9-6 to Table 9-9) and case study authorities, together with the differences evident in the 2011 and 2031 LAD maps shown in Figure 6-2 and Figure 9-3, illustrate this point. The area covered by the Darlington, Durham and Tees LAT stands out as an area that performs badly in terms of health outcomes. Whilst there are some improvements in prevalence rates over time, the improvements are not as marked as in other LATs. The Birmingham and Black Country LAT is also an area whose health status deteriorates relative to others. The more prosperous LATs, primarily located in southern England outside London largely retain their relatively positive health status from 2011 to 2031. In London the situation is mixed but, excepting DHBS, its relative performance amongst LATs and LADs improves markedly. This is contrasted with the situation that many health care resources are actually centralised in

London, given its historic importance, attractiveness, prestige and population. To be fairer in the resourcing of such health care it makes sense to relocate some of these resources to areas such as the Midlands and North East where the relative need is seen to increase.

As well as the revealed need demonstrated in these findings, the findings also point to locations where there is a need for screening services to identify individuals potentially at risk of developing a morbidity. Such initiatives help to capture individuals at risk and provide appropriate treatment before potential problems from the morbidity's symptoms manifest and complications become apparent (see also section 11.4.4 for how undiagnosed cases in the population can be explicitly modelled). As an illustration of the importance of screening services, the work of Kanavos (2012) is reported by Diabetes UK (2015a) and they report that the annual inpatient care, to treat short and long term complications of diabetes, is estimated at between £1,800 and £2,500 per patient. This compares with annual outpatient costs, which include the cost of medications and monitoring supplies, estimated at between £300 and £370 per patient. Clearly, mitigating longer term complications can provide not only a better quality of life for the patient but also financial savings.

Here the medical need is quantified as prevalence counts and rates for each LAD. It is possible to transform these prevalences into alternative measure of health care need through the application of rates. One such calculation allows the costs associated with morbidities to be estimates. Research based on residents of Shropshire reports that a healthy patient costs the NHS about £293 a year, those with one long-term condition cost an estimated £795, those with two cost £1,655 and those with three cost £2,726 each. For patients with four or more such conditions, their need for frequent treatment and monitoring costs £4,549 a year, those with five conditions cost £5,841 and those with six conditions about £7,325 Those with seven or more conditions cost £11,233 per annum (Kasteridis et al., 2012). Where available, morbidity specific cost may also be used. For example the diabetes cost quoted above could be used to calculate a total cost to the NHS in each LAD for diabetes. Performing this calculation for all the morbidities and allowing differential costs by gender, age and location is possible to do but requires access to sensitive person-level hospital data. As well as costs it may be also possible to attach other resources costs to each patient, e.g. clinician or general practitioner time or specialist equipment .

11.4 FURTHER WORK

In this thesis a grounding has been provided in defining a methodological structure to forecast the number of individuals with specific morbidities in each LAD to 2031.

There are a number of ways in which this work may be taken forward. In this section some of these ways are introduced and guidance offered on how to achieve them.

11.4.1 Revised area types

In the constraining of the ETHPOP adjustment factors, the morbidity hazard models and the selection of the case study authorities, the 2001 ONS area classifications are used. The role of these classifications is to capture some element of the socio-economic background of the individuals who live in a LAD. There may, however, be questions around it being fit for purpose. Firstly, it is designed for general use in a variety of contexts and not specifically geared towards differentiating on health outcomes for the 50 and older population. Such health classifications do exist; Experian (CACI, 2007) produced a health outcome classification (Health Acorn) at unit postcode level for Great Britain which has some potential utility, although the geography is not compatible with that provided within ELSA. Secondly, some of the clustering in the ONS classification is unhelpful. For example, the small seven member Group of London Centre contains LADs as diverse as Camden, Kensington and Chelsea, Southwark and Tower Hamlets. The Coastal and Countryside Group is also diverse encompassing the economically challenged LAD of Blackpool and prosperous rural Ryedale. Whether the differentials in health outcomes are consistent in these types of LADs is debatable. Finally, these classifications are clearly becoming out of date. Since December 2014 the ONS has been promising the release of updated LAD area classifications based on 2011 Census data, but as of late Spring 2015 they are not yet available. It will be insightful to see how these area classifications differ, not only in their nomenclature or characteristics but also their membership. If this scrutiny is sustained then it will be worthwhile re-fitting the hazard morbidity models using these updated area type classifications and re-running the health microsimulation.

11.4.2 Health microsimulation decomposition

As mentioned in Chapter 10, the impacts of the health microsimulation can be discerned on an individual level through the interpretation of the parameters estimated in the hazard models. Further insight into aggregate impacts can, however, be obtained by attempting a similar accounting system for the health influences and how they decompose the total impact. This can be done by maintaining the full functionality of the health microsimulation and the input data but using more parsimonious hazard models. A possible regime would be to re-estimate and use models that initially contained just the wave and age influences, and then re-fitted models that successively introduced gender, ethnicity, smoking status, co-morbidities and LAD area types. Extensions of equations 5-18a-f, 5-19, 5-20a-f and 5-21 would enable these impacts to be measured.

The main concern with this proposal is whether the goodness of fit of these parsimonious models with so few explanatory terms would be adequate to the task. If the models were not of sufficient quality then the acquisition of a morbidity may be a random or even counter intuitive event. Also the use of the morbidity hazard models for other purposes would need to be taken account of. To retain consistency consideration would need to be given to the use of either the full models as described in section 8.2 or the more parsimonious morbidity hazard models to calculate the accumulated incidence from age 30 to 49 to ensure that the Replenishers aged 50 and 51 joined the future synthetic wave with an appropriate level of morbidity. Also an assessment would need to be made of whether to use the full models or similarly parsimonious hazard models for the estimation of the mortality and migration selection weights.

11.4.3 Other morbidities

Both CVD and respiratory illness account for a larger number of causes of death and hospital admissions in the 50 and older population. Also DHBS, whilst not an important cause of death or hospital admission (except through complications), is often cited as a comorbidity with CVD and may still affect individual's quality of life and perception of their disability status. These rationales influence the choice of morbidities to study in this thesis. However there are still other morbidities that are also important and could be modelling using this framework. Firstly, there are the cancers or neoplasms. These are differentiated by ELSA which collects information on the organ or part of body in which the cancer started (lung, breast, prostate, colon, bowel or rectum, lymphoma, leukaemia, melanoma or other skin cancer or somewhere else). Clearly cancer is a complex morbidity arising at a number of sites in the body and predicated by a variety of factors, which gives the view that cancer is almost a collection of comorbidities with a challenging and diverse range of determinants. Other potential morbidities that are worthy of consideration if the work is taken forward are those which are not life threatening but can severely affect the quality of life and incur significant health care resources, such as arthritis, Parkinson disease, dementia or depression. Of the six morbidities reported in the ELSA wave reports, arthritis has the highest prevalence and is particularly prevalent amongst the female population (Institute for Fiscal Studies, 2012). Dementia is a little understood morbidity with few treatments available meaning that the quality of life for suffers and their carers is poor with many resorting to expensive institutionalise care. This makes it worthy of further study.

11.4.4 Undiagnosed cases

Not only can the health microsimulation be extended to include other morbidities but it can also be used to estimate or forecast the extent of undiagnosed or potential prevalence within each LAD. Such knowledge will help to identify those locations at

which to target screening services. Using biomarker data collected at a nurse visit it is possible to identify those individuals with a potentially undiagnosed morbidity or a high risk of developing a morbidity. For CVD these indicators include the level of fibrinogen, the systolic and diastolic blood pressure (linked to hypertension), and the balance of HDL-C ('good' cholesterol) and LDL-C ('bad' cholesterol). The risk for the stroke component of CVD can be tested for by measuring triglycerides in the blood. The potential presence of diabetes can be picked up using levels of glucose in the urine or blood, an oral glucose test or Hemaglobin Alc (HBAlc) levels. Finally, the ratio of the Forced Vital Capacity to the Forced Expiratory Volume (FVC/FEV1) or a peak flow test can indicate reduced lung capacity due to respiratory illness. Levels of these biomarkers that indicate a potential medical issue are shown in Table 11-1.

Table 11-1 : Levels of biomarkers to indicate morbidity risk or undiagnosed

Morbidity	Measure	Criteria for concern
CVD	Fibrinogen[1]	≥ 3.43g/l
	Systolic and diastolic blood pressure[2]	Systolic ≥140 and diastolic ≥ 90mmHg
	HDL-C[3]	≤ 1mmol/l (men) ≤ 1.2 mmol/l (females)
	LDL-C[4]	≥ 4.15mmol/l
Diabetes	Blood glucose[5]	≥ 7.0 mmol/l
	Oral glucose[5]	≥ 11.1 mmol/l
	HbAlc[5]	≥ 48.0mmol/mol or ≥ 6.5%
Respiratory illness	FEV1/FVC[6]	< 70%
	Peak flow[7]	Depends on device

[1] Kannel et al, (1987), [2] Rapsomaniki et al, (2015), [3] Heart UK (2015a), [4] Heart UK (2015b), [5] Diabetes UK (2015b), [6] National Institute for Health and Care Excellence (2015), [7] Clement Clarke International (2015).

The simplest way to assess this degree of undiagnosis or risk prevalence is to assume that the morbidity profile for the population with the morbidity is not dis-similar to this 'unknown' population. This allows the population in each LAD to be defined by the same constraints as those that have the morbidity. Using the biomarker information for the ELSA participants those at risk or with an undiagnosed morbidity can be identified and this information can be attached to their morbidity profile and, just as with diagnosed cases, the prevalence rate estimated. A more complex approach would be to treat these undiagnosed or at risk morbidities as separate incidences and examine how to include them as outcomes in their own right. This would require further hazard modelling as described in section 8.2.

There are however a number of issues with this approach to identifying undiagnosed or at risk prevalences. Firstly, the sample population is composed of wave 5 ELSA participants but there was no biomarker data collected at wave 5. Instead information from such data collected at waves 4 and 6 may be used to infer their undiagnosed or at risk status at wave 5. Secondly, the biomarker data are only collected for core members (i.e. C1CM, C3CM, C3CM and C6CM) whilst the sampling population used in the spatial microsimulation is composed of all individuals at wave 5 with the required information, irrespective of their participation status. Thus there will be some members of the sampling population whose undiagnosed or potential prevalence is unknown. Thirdly, not all core members consented to a nurse visit to collect the biomarker data (the response rate is about 85% at the three nurse visit waves, NATCEN, 2014) so there will be sample members who are core members without this biomarker data. Clearly it would be possible to redefine the spatial microsimulation sample population to only include those participants who had the required biomarker data. However, this would be at the cost of a deterioration in the heterogeneity and coverage of the sample population.

11.4.5 Capturing uncertainty

The results from the health microsimulation are based on information that is, to a degree, uncertain. The population projection used to re-structure future populations is a projection and subject to the uncertainties in the assumptions that underlie the projection. The hazard models are estimated using a survey sample of data. The spatial microsimulation and health microsimulation involve a stochastic element and whilst this is seen to have little impact on the outcomes, it is still a source of uncertainty.

These uncertainties are not necessarily fatal to any attempt at modelling, in fact they can be exploited to provide some measure of uncertainty around any outcomes. If the range of this uncertainly is credible it can then be used to lend confidence to any observed trends and differentials found in the outcomes. Perhaps the easiest way to incorporate the uncertainty is by exploiting the stochastic nature of methods within the health microsimulation. As demonstrated, this can be achieved by the use of different random number seeds that create different streams of pseudo random numbers to be used in the Monte Carlo sampling. Based on the available evidence the amount of variance introduced by this approach is not great.

The next easiest mechanism to incorporate uncertainty recognises the fact that the hazard models are based on a sample of survey results. The parameter estimates from the hazard models are point estimates and have associated standard errors and also equivalent variance and covariances. These variance-covariance matrices are trivially provided in most statistical estimation software. If it is assumed that the parameter estimates to use in the health microsimulation are actually sample draws from some

multivariate distribution with its mean set to the point estimates and its variance set to the variance-covariance matrix then it can be run a number of times using different draws for the parameter estimates. Those estimates with low standard errors whose point estimates are therefore more reliable will be consistent amongst these draws, whilst those with high standard errors will vary considerably, capturing the uncertainty associated with their estimates.

A multivariate normal distribution is the most credible such sampling distribution and one of these draws can be efficiently made at the start of each health microsimulation using the parameter estimates and an estimated variance covariance matrix. The requirement to draw such a sample will be computationally trivial since there are typically just 10-15 parameters in each hazard model. The health microsimulation takes about 30 minutes to run on a reasonable desktop PC and, if there are multiple processing cores available, then more than one run can take place at a time. The impact of this modification on the variation in outcomes is likely to be greater than the use of different random number seeds.

The third source of uncertainty is perhaps the easiest to implement but the most difficult to define. It simply involves using different population projections as input to the health microsimulation. However the provenance of such projections is a challenge. Whilst variants exist for national ONS projections they do not exist for the ONS sub national projections. But even if they did, a question arises as to whether they represent sufficient credible uncertainty in the projections. Many of the variants are designed to produce 'extreme' outcomes and to use these would give far too much weight to unlikely outcomes. What are required in reality are probabilistic forecasts that follow some distributional outcomes, where moderate outcomes are more likely to occur than extreme outcomes. Such models are not unknown in the literature (Alkema, 2015) but at the moment the projections from models are not available at the geography required here. It is difficult to assess the likely impact of using probabilistic projections since the variation in the health microsimulation outcomes will be related to the variability of the projections. However, it is likely to be greater than the use of random number seeds.

The three methods for introducing uncertainly here are all credible sources of variation within the model and can be implemented with just superficial changes to the code and input data. They can also be incorporated incrementally and variously to assess each method's impact on estimation uncertainty.

11.4.6 Composition of the ELSA population

The ELSA survey is a large and complex survey that captures well the dynamics associated with ageing in the general population. Whilst the size of the survey is large and stringent efforts are deployed to reduce attrition within the survey, there are issues

for this study in the size of the BME population. Proportionately ELSA may reflect well the size of the aged 50 and older BME population, but what this means in practice is that the number of BME individuals in the sample is small. This is evidenced in Table 5-3 where some merging of age bands is required to ensure that there is at least, some representation of all BME populations in each gender and age band. However, the question arises as to whether these small numbers are truly representative of the specific ethnic group, where a bias towards a few individuals having (or not having) a morbidity could sway the prevalence estimates for those LADs with high BME populations. Also the standard errors for the BME parameters in Table 8-1, Table 8-2 and Table 8-3 are large. At the moment it is difficult to deal with this issue with these data. Ideally what should have been given some consideration is an over sampling of the BME population at an early stage of ELSA or during one of the refresher cohorts.

11.5 CONCLUSION

In the first chapter of this thesis a number of research questions are posed and the objectives that go towards answering these research questions are highlighted. The first and substantive research question concerns a desire to forecast the likely number of people aged 50 and older, who will have one or more of the specific morbidities in each LAD, to 2031. This has been achieved by the using the outlined modelling approach described in the Methodology chapter, Chapter 5, and reported in the Health Microsimulation chapter, Chapter 9.

Of critical importance to the health microsimulation is that its operation is informed by knowledge about the socio-demographic and socio-economic makeup of each LAD. This has been gained through the contextual information and literature review provided in Chapters 2 and 3 and the modelling work in Chapters 6 and 8. These influences have been incorporated through the selection of constraints that define the 2011 base population in each LAD and the hazard models that predict morbidity incidence, mortality and migration. This allows the production of forecasts that answer the second research question which is how these morbidities vary by the gender, age, ethnicity and other socio-demographic characteristics of the population in the area.

To gain a deeper understanding of these local area trends an accounting system approach has been adopted to show how are these forecasts are influenced by both changes in the demographic composition of an area and by changing trends in health outcomes. This work fulfils the third research question and is covered in Chapter 10.

Whilst the first research question is the most substantive, the final question around what are the public policy implications of the work is perhaps the most important. This issue

has been addressed at points throughout this thesis but are drawn into focus in this chapter.

12 GLOSSARY

BHPS	British Household Panel Survey
BME	Black and Minority Ethnic
C1CM	Cohort 1 Core Member
C1CP	Cohort 1 Core member Partner
C1NP1	Cohort 1 New Partner at wave 1
C1NP2	Cohort 1 New Partner at wave 2
C1NP3	Cohort 1 New Partner at wave 3
C1NP4	Cohort 1 New Partner at wave 4
C1NP5	Cohort 1 New Partner at wave 5
C1NP6	Cohort 1 New Partner at wave 6
C1OP	Cohort 1 Older Partner
C1SM	Cohort 1 Sample member
C1YP	Cohort 1 Younger Partner
C3CM	Cohort 3 Core Member
C3CP	Cohort 3 Core member Partner
C3NP3	Cohort 3 New Partner at wave 3
C3NP4	Cohort 3 New Partner at wave 4
C3NP5	Cohort 3 New Partner at wave 5
C3NP6	Cohort 3 New Partner at wave 6
C3OP	Cohort 3 Older Partner
C3SM	Cohort 3 Sample Member
C3YP	Cohort 3 Younger Partner
C4CM	Cohort 4 Core Member
C4CP	Cohort 4 Core member Partner
C4NP4	Cohort 4 New Partner at wave 4
C4NP5	Cohort 4 New Partner at wave 5
C4NP6	Cohort 4 New Partner at wave 6
C4OP	Cohort 4 Older Partner
C4SM	Cohort 3 Sample Member
C4YP	Cohort 4 Younger Partner
C6CM	Cohort 6 Core Member
C6CP	Cohort 6 Core member Partner
C6NP6	Cohort 6 New Partner at wave 6
C6OP	Cohort 6 Older Partner

C6SM	Cohort 6 Sample Member
C6YP	Cohort 6 Younger Partner
CCG	Clinical Commissioning Groups
CHD	Coronary Heart Disease
CO	Combinatorial Optimisation
COPD	Chronic Obstructive Pulmonary Disease
CT	Computerised Tomography
CVD	Cardio Vascular Disease
DC	Detailed Characteristics
DES	Discrete Event Simulation
DHBS	Diabetes or high blood sugar
ELSA	English Longitudinal Survey of Ageing
ELOB	Expected years Of Life at Birth
ESRC	Economic and Social Science Research Council
ETHPOP	ETHnic POPulation projections
EU	European Union
FCE	Finished Consultant Episode
FMF	Flexible Modelling Framework
GCSE	General Certificate of Secondary Education
GDP	Gross Domestic Product
GHS	General Household Survey
GP	General Practitioner doctor
HDL	High-Density Lipoprotein
HES	Hospital Episode Statistics
HSCIC	Health and Social Care Information Centre
HSfE	Health Survey for England
ICD-10	International Statistical Classification of Diseases 10th Revision
IHD	Ischaemic Heart Disease
IPF	Iterative Proportional Fitting
IQR	Inter-Quartile Range
KS	Key Statistics
LAD	Local Authority District
LAT	NHS Local Area Teams
LC	Local Characteristics
LDL	Low-Density Lipoprotein
LSOA	Lower Super Output Areas
MRI	Magnetic resonance imaging
MSOA	Middle Super Output Areas

NHS	National Health Service
NS-SEC	National Statistics socio-economic classification
OA	Output Areas
OECD	Organisation for Economic Co-operation and Development
ONS	Office for National Statistics
PCO	Primary Care Organisations
QS	Quick Statistics
RI	Respiratory Illness
SD	Systems Dynamic
SHA	Strategic Health Authorities
SNPP	Sub national Population Projections
SPSS	Statistical Package for the Social Sciences
TAE	Total Absolute Error
TREND	UPTAP trend forecast
UK	United Kingdom of Great Britain and Northern Ireland
UKHLS	United Kingdom Household Longitudinal Study
UPTAP	Understanding Population Trends and Processes
UPTAP-ER	UPTAP Emigration Ratio forecast
WHO	World Health Organisation

13 APPENDIX

Structure of Constraint and Sampling Population Tables for the Spatial Microsimulation

Table 13-1 : Gender by age (DC1117)

Usual Residents	Male	Female
Aged 50 to 54	M50	F50
Aged 55 to 59	M55	F55
Aged 60 to 64	M60	F60
Aged 65 to 69	M65	F65
Aged 70 to 74	M70	F70
Aged 75 to 79	M75	F75
Aged 80 to 84	M80	F80
Aged 85 and older	M85	F85

Table 13-2 : Gender by age by ethnicity (DC2101)

Usual residents		White	Mixed	Asian	Black	Other
Male	Aged 50 to 64	M50W	M50M	M50A	M50B	M50O
	Aged 65 and older	M65W	M65M	M65A	M65B	M65O
Female	Aged 50 to 64	F50W	F50M	F50A	F50B	F50O
	Aged 65 and older	F65W	F65M	F65A	F65B	F65O

Table 13-3 : Gender by age by disability (household residents) (LC3101)

Household residents plus institutional residents who are not clients		No LLTI	LLTI
Male	Aged 50 to 64	M50N	M50L
	Aged 65 to 74	M65N	M65L
	Aged 75 to 84	M75N	M75L
	Aged 85 and older	M85N	M85L
Female	Aged 50 to 64	F50N	F50L
	Aged 65 to 74	F65N	F65L
	Aged 75 to 84	F75N	F75L
	Aged 85 and older	F85N	F85L
Institutional residents who are clients		INSTITUTION	

Table 13-4 : gender by age by disability (institutional residents) (DC3402)

Institutional residents who are clients		No LLTI	LLTI
Male	Aged 50 to 64	M50N	M50L
	Aged 65 to 74		M65L
	Aged 75 to 84		M75L
	Aged 85 and older		M85L
Female	Aged 50 to 64	F50N	F50L
	Aged 65 to 74		F65L
	Aged 75 to 84		F75L
	Aged 85 and older		F85L
Household residents plus institutional residents who are not clients		HOUSEHOLD	

Table 13-5 : Gender by age by provision of unpaid care (LC3301)

Household residents plus institutional residents who are not clients		No care	Between 1 and 19 hours per week	Between 20 and 49 hours per week	Fifty or more hours per week
Male	Aged 50 to 64	M50C0	M51C1	M520C20	M550C50
	Aged 65 and older	M65C0	M65C1	M65C20	M65C50
Female	Aged 50 to 64	F50C0	F51C1	F520C20	F550C50
	Aged 65 and older	F65C0	F65C1	F65C20	F65C50
Institutional residents who are clients		INSTITUTION			

Table 13-6 : Gender by age by vehicle availability (DC4109)

Household residents plus institutional residents who are not clients		No vehicles	One vehicle	Two or more vehicles
Male	Aged 50 to 64	M50V0	M51V1	M52V2
	Aged 65 to 74	M65V0	M65V1	M65V2
	Aged 75 to 84	M75V0	M75V1	M75V2
	Aged 85 and older	M85V0	M85V1	M85V2
Female	Aged 50 to 64	F50V0	F51V1	F52V2
	Aged 65 to 74	F65V0	F65V1	F65V2
	Aged 75 to 84	F75V0	F75V1	F75V2
	Aged 85 and older	F85V0	F85V1	F85V2
Institutional residents who are clients		INSTITUTION		

Table 13-7 : Gender by age by tenure (DC3408)

Household residents plus institutional residents who are not clients		Owned outright	Mortgaged	Social rented	Private rented
Male	Aged 50 to 64	M50O	M50M	M50S	M50P
	Aged 65 and older	M65O	M65M	M65S	M65P
Female	Aged 50 to 64	F50O	F50M	F50S	F50P
	Aged 65 and older	F65O	F65M	F65S	F65P
Institutional residents who are clients		INSTITUTION			

Table 13-8 : Gender by age by highest qualification (DC5107)

Usual residents		No qualifications	Qualification below degree	Degree	Other or foreign
Male	Aged 50 to 64	M50LN	M50L123A	M50L4	M50LO
	Aged 65 to 74	M65LN	M65L123A	M65L4	M65LO
	Aged 75 to 84	M75LN	M75L123A	M75L4	M75LO
	Aged 85 and older	M85LN	M85L123A	M85L4	M85LO
Female	Aged 50 to 64	F50LN	F50L123A	F50L4	F50LO
	Aged 65 to 74	F65LN	F65L123A	F65L4	F65LO
	Aged 75 to 84	F75LN	F75L123A	F75L4	F75LO
	Aged 85 and older	F85LN	F85L123A	F85L4	F85LO

Table 13-9 : Gender by age by NS-SEC (DC6114)

Usual residents		Managerial and Professional	Intermediate occupations	Small employer and own account	Lower supervisory	Semi-routine and routine	Other (never worked and long term unemployed)
Male	Aged 50 to 64	M50S1	M50S2	M50S3	M50S4	M50S5	M50SO
	Aged 65 and older	M65S1	M65S2	M65S3	M65S4	M65S5	M65SO
Female	Aged 50 to 64	F50S1	F50S2	F50S3	F50S4	F50S5	F50SO
	Aged 65 and older	F65S1	F65S2	F65S3	F65S4	F65S5	F65SO

Table 13-10 : Gender by age by living arrangement (DC1108)

Household residents plus institutional residents who are not clients		Married	Co-habiting	Single	Married (not in couple)	Separated	Widowed
Male	Aged 50 to 64	M50M	M50C	M50A	M50N	M50S	M50B
	Aged 65 to 74	M50M	M50C	M50A	M50N	M50S	M50B
	Aged 75 and older	M50M	M50C	M50A	M50N	M50S	M50B
Female	Aged 50 to 64	F50M	F50C	F50A	F50N	F50S	F50B
	Aged 65 to 74	F65M	F65C	F65A	F65N	F65S	F65B
	Aged 75 and older	F75M	F75C	F75A	F75N	F75S	F75B
Institutional residents who are clients		INSTITUTION					

Table 13-11 : Sample of LC1117 (Gender and age) constraint table (usual residents)

CODE	M50	M55	M60	M65	M70	M75	M80	M85	F50	F55	F60	F65	F70	F75	F80	F85
Hartlepool	3,229	2,762	2,878	2,095	1,799	1,469	940	588	3,235	2,833	2,959	2,141	2,139	1,858	1,301	1,268
Middlesbrough	4,556	3,845	3,739	2,794	2,402	1,856	1,269	766	4,677	4,088	3,792	2,886	2,751	2,422	1,846	1,699
Redcar and Cleveland	4,656	4,094	4,606	3,798	3,159	2,292	1,539	1,010	4,773	4,364	4,905	3,978	3,483	2,725	2,165	2,103
Stockton-on-Tees	6,533	5,760	5,733	4,278	3,456	2,718	1,730	1,119	6,866	5,849	5,915	4,467	3,901	3,371	2,549	2,362
Plymouth	8,310	6,782	7,442	6,000	4,598	3,628	2,491	1,757	8,057	7,173	7,698	6,156	5,245	4,492	3,487	3,861

Table 13-12 : Sample of LC3301 (disability) constraint table (household residents)

CODE	M50N	M50L	M65N	M65L	M75N	M75L	M85N	M85L	F50N	F50L	F65N	F65L	F75N	F75L	F85N	F85L	INSTITUTION
Hartlepool	6,049	2,768	1,827	2,023	718	1,607	79	436	5,916	3,074	2,084	2,162	834	2,169	124	829	795
Middlesbrough	8,461	3,572	2,681	2,466	1,050	2,003	125	560	8,431	4,075	2,881	2,712	1,295	2,820	192	1,198	866
Redcar and Cleveland	9,621	3,657	3,748	3,147	1,274	2,461	168	757	9,877	4,113	3,959	3,435	1,470	3,229	212	1,533	989
Stockton-on-Tees	13,556	4,291	4,402	3,251	1,636	2,653	151	837	13,628	4,930	4,633	3,658	1,863	3,770	251	1,649	1,448
Plymouth	16,347	6,048	5,928	4,586	2,198	3,758	297	1298	16,344	6,485	6,403	4,876	2,650	5,043	490	2,653	1,773

Table 13-13 : Sample of DC3402 (disability) constraint table (institutional residents)

CODE	M50N	M50L	M65L	M75L	M85L	F50N	F50L	F65L	F75L	F85L	HOUSEHOLD
Hartlepool	14	46	44	79	70	18	37	33	149	305	32,699
Middlesbrough	17	97	45	70	80	17	50	43	147	300	44,522
Redcar and Cleveland	5	75	60	96	85	7	49	66	191	355	52,661
Stockton-on-Tees	97	97	76	151	129	46	69	75	271	437	65,159
Plymouth	41	115	74	157	161	55	88	114	275	693	85,404

Table 13-14 : Extract of the sample population derived from ELSA wave 5

idauniq	DC1117	LC3301	DC2101	DC1108	DC4109	DC6114	LC3301	DC3408	DC3402
103717	M85	M85L	M65W	M75B	M65V1	M65S4	M65C0	M65O	HOUSEHOLD
103722	M85	M85L	M65W	M75B	M65V0	M65S1	M65C0	M65M	HOUSEHOLD
103724	F75	F75L	F65W	F75D	F65V1	F65S2	F65C0	F65M	HOUSEHOLD
103725	M60	M50N	M50B	M50C	M50V2	M50S3	M50C0	M50S	HOUSEHOLD
103726	M60	M50N	M50W	M50D	M50V1	M50S5	M50C20	M50S	HOUSEHOLD
103727	F60	F50N	F50W	F50M	F50V2	F50S4	F50C0	F50S	HOUSEHOLD
103729	F60	F50L	F50W	F50B	F50V1	F50S5	F50C0	F50M	HOUSEHOLD
103730	M85	INSTITUTION	M65W	INSTITUTION	INSTITUTION	M65S1	INSTITUTION	INSTITUTION	M85L
103731	F60	F50L	F50W	F50M	F50V1	F50S3	F50C0	F50P	HOUSEHOLD
103734	F65	F65L	F65W	F65B	F65V1	F65S5	F65C0	F65S	HOUSEHOLD
103943	F60	F50N	F50W	F50M	F50V2	F50S5	F50C1	F50O	HOUSEHOLD
103945	M65	M65N	M65W	M65M	M65V1	M65S3	M65C0	M65O	HOUSEHOLD
103946	M65	M65N	M65A	M65M	M65V1	M65S4	M65C0	M65M	HOUSEHOLD
103947	M70	M65L	M65W	M65D	M65V1	M65S5	M65C0	M65M	HOUSEHOLD
103949	F65	F65N	F65W	F65M	F65V2	F65S2	F65C0	F65M	HOUSEHOLD
103955	F75	INSTITUTION	F65W	INSTITUTION	INSTITUTION	F65S5	INSTITUTION	INSTITUTION	F75L
103956	M60	M50N	M50W	M50M	M50V1	M50S3	M50C0	M50O	HOUSEHOLD
103957	M70	M65L	M65W	M65M	M65V1	M65S4	M65C50	M65S	HOUSEHOLD
103960	F70	F65N	F65W	F65M	F65V2	F65S5	F65C0	F65P	HOUSEHOLD

14 MAPS

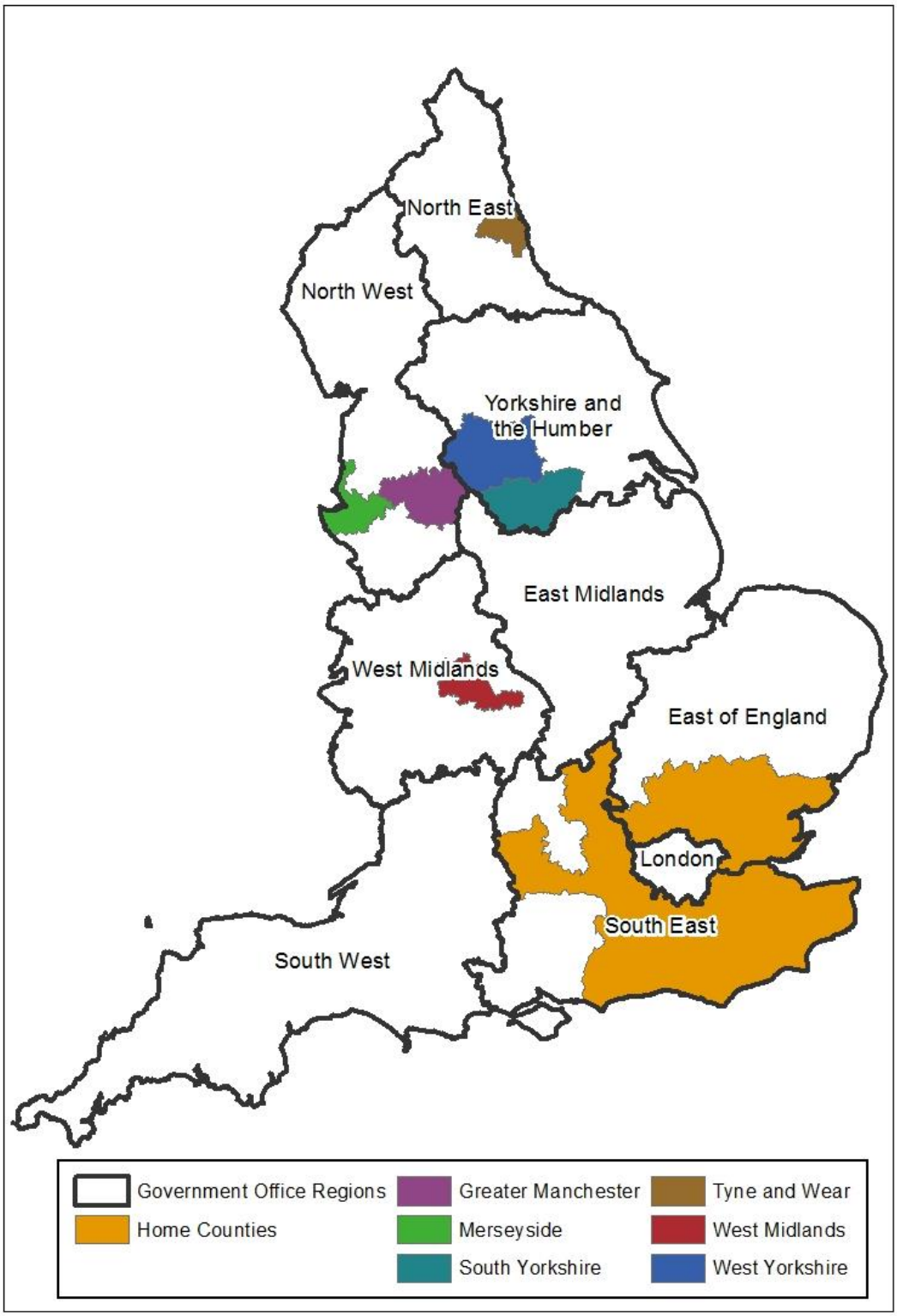


Figure 14-1 : Former English Government Regions, Metropolitan Counties and Home Counties

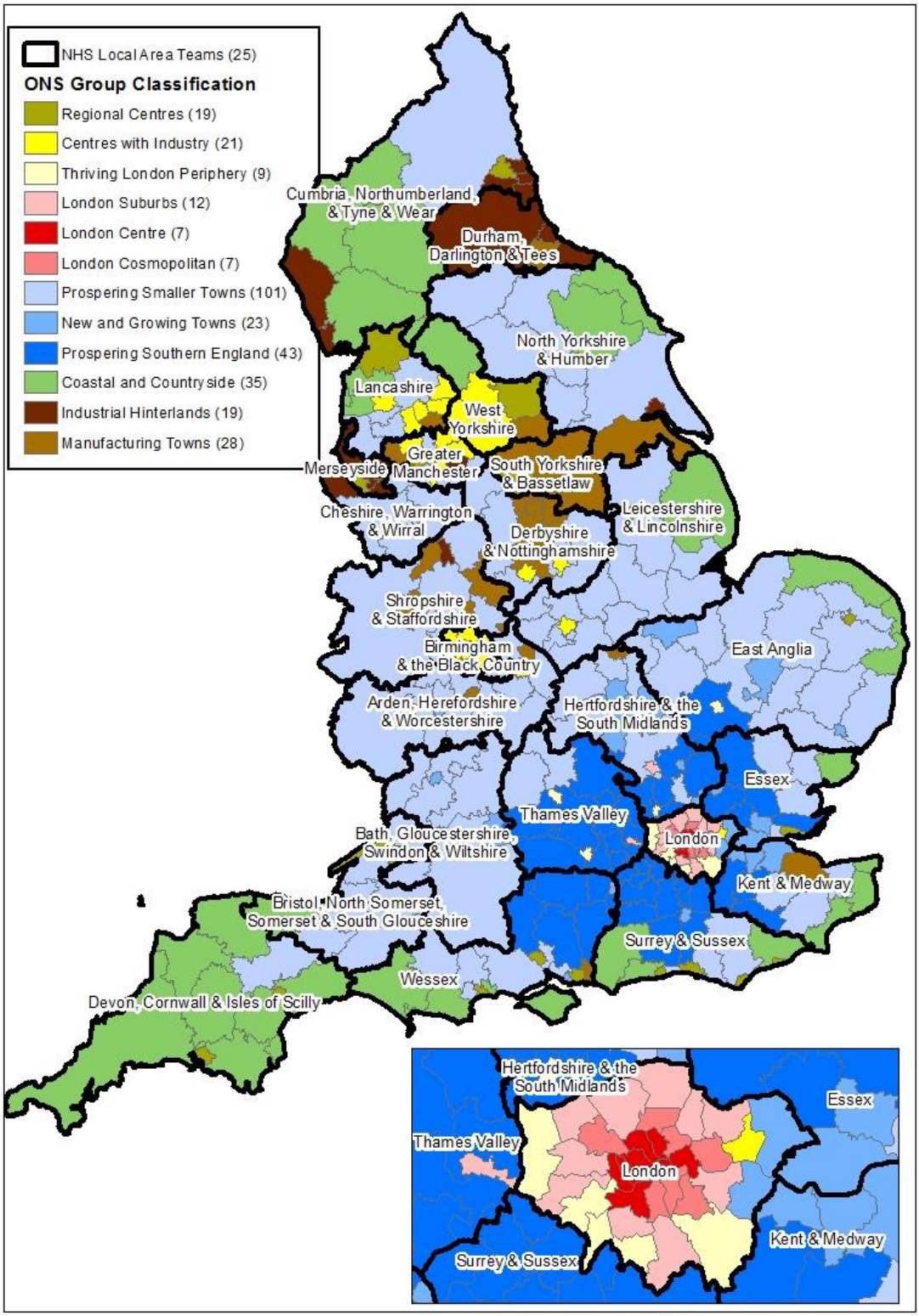


Figure 14-2 : ONS area types and NHS Local Area Teams

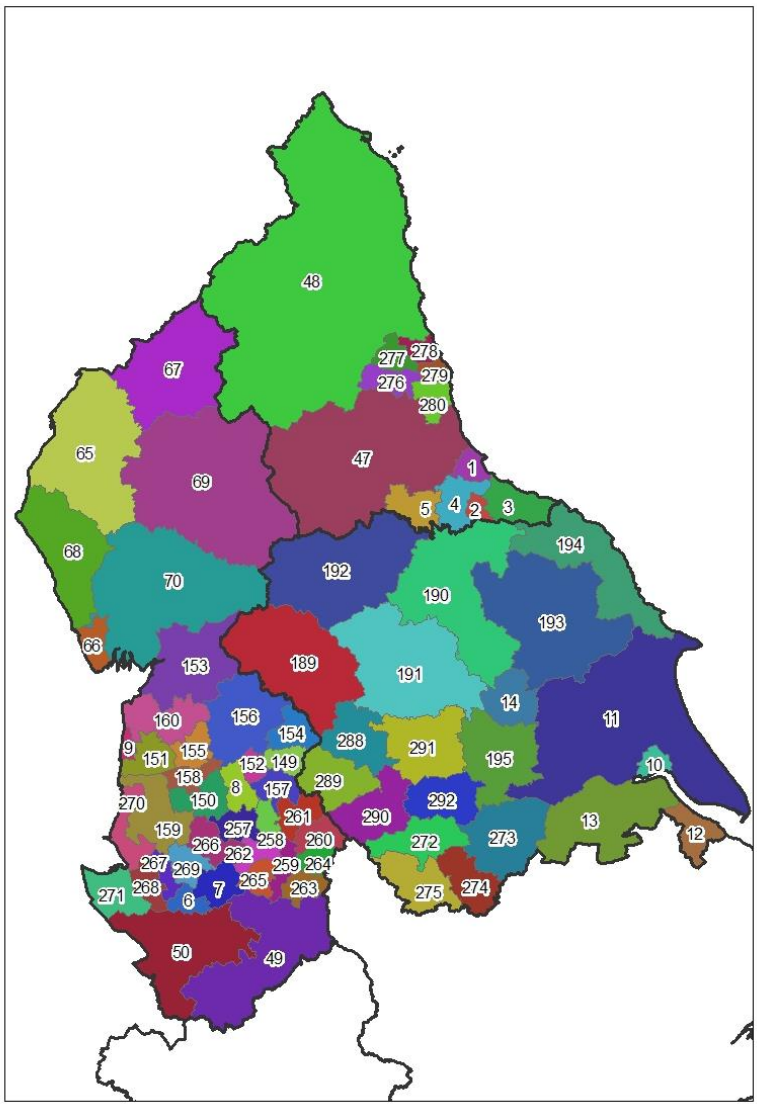


Figure 14-3 : Northern English LADs

Name	Region	ONS Code	Label No
Hartlepool	North East	E41000001	1
Middlesbrough	North East	E41000002	2
Redcar and Cleveland	North East	E41000003	3
Stockton on Tees	North East	E41000004	4
Darlington	North East	E41000005	5
Halton	North West	E41000006	6
Warrington	North West	E41000007	7
Blackburn with Darwen	North West	E41000008	8
Blackpool	North West	E41000009	9
Kingston upon Hull	Yorkshire and The Humber	E41000010	10
East Riding of Yorkshire	Yorkshire and The Humber	E41000011	11
North East Lincolnshire	Yorkshire and The Humber	E41000012	12
North Lincolnshire	Yorkshire and The Humber	E41000013	13
York	Yorkshire and The Humber	E41000014	14
County Durham UA	North East	E41000047	47
Northumberland UA	North East	E41000048	48
Cheshire East UA	North West	E41000049	49
Cheshire West and Chester UA	North West	E41000050	50
Allerdale	North West	E41000065	65

Name	Region	ONS Code	Label No
Barrow-in-Furness	North West	E41000066	66
Carlisle	North West	E41000067	67
Copeland	North West	E41000068	68
Eden	North West	E41000069	69
South Lakeland	North West	E41000070	70
Burnley	North West	E41000149	149
Chorley	North West	E41000150	150
Fylde	North West	E41000151	151
Hyndburn	North West	E41000152	152
Lancaster	North West	E41000153	153
Pendle	North West	E41000154	154
Preston	North West	E41000155	155
Ribble Valley	North West	E41000156	156
Rossendale	North West	E41000157	157
South Ribble	North West	E41000158	158
West Lancashire	North West	E41000159	159
Wyre	North West	E41000160	160
Craven	Yorkshire and The Humber	E41000189	189
Hambleton	Yorkshire and The Humber	E41000190	190
Harrogate	Yorkshire and The Humber	E41000191	191
Richmondshire	Yorkshire and The Humber	E41000192	192
Ryedale	Yorkshire and The Humber	E41000193	193
Scarborough	Yorkshire and The Humber	E41000194	194
Selby	Yorkshire and The Humber	E41000195	195
Bolton	North West	E41000257	257
Bury	North West	E41000258	258
Manchester	North West	E41000259	259
Oldham	North West	E41000260	260
Rochdale	North West	E41000261	261
Salford	North West	E41000262	262
Stockport	North West	E41000263	263
Tameside	North West	E41000264	264
Trafford	North West	E41000265	265
Wigan	North West	E41000266	266
Knowsley	North West	E41000267	267
Liverpool	North West	E41000268	268
St Helens	North West	E41000269	269
Sefton	North West	E41000270	270
Wirral	North West	E41000271	271
Barnsley	Yorkshire and The Humber	E41000272	272
Doncaster	Yorkshire and The Humber	E41000273	273
Rotherham	Yorkshire and The Humber	E41000274	274
Sheffield	Yorkshire and The Humber	E41000275	275
Gateshead	North East	E41000276	276
Newcastle upon Tyne	North East	E41000277	277
North Tyneside	North East	E41000278	278
South Tyneside	North East	E41000279	279
Sunderland	North East	E41000280	280
Bradford	Yorkshire and The Humber	E41000288	288
Calderdale	Yorkshire and The Humber	E41000289	289
Kirklees	Yorkshire and The Humber	E41000290	290
Leeds	Yorkshire and The Humber	E41000291	291
Wakefield	Yorkshire and The Humber	E41000292	292

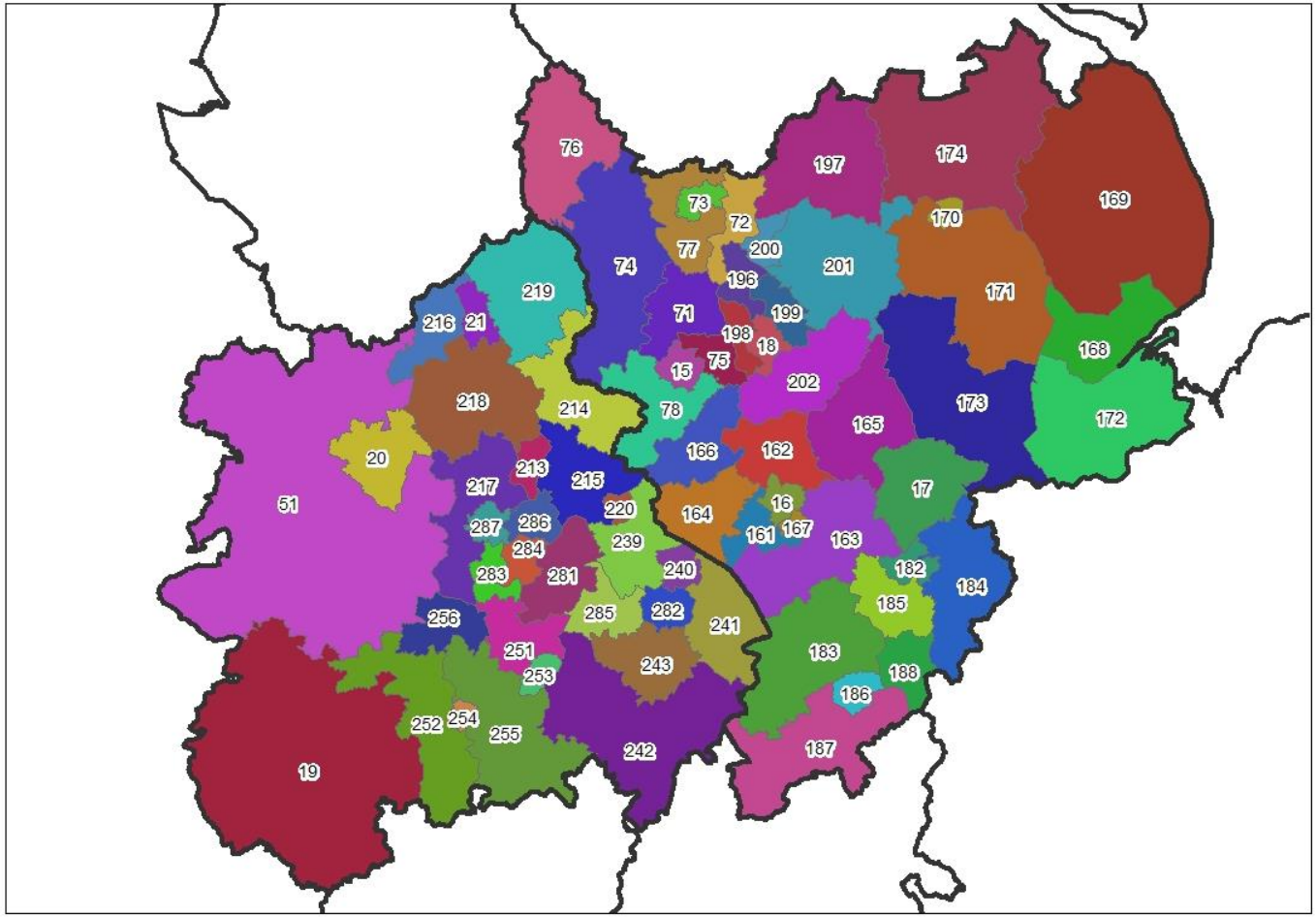


Figure 14-4 : Midlands LADs

Name	Region	ONS Code	Label No
Derby	East Midlands	E41000015	15
Leicester	East Midlands	E41000016	16
Rutland	East Midlands	E41000017	17
Nottingham	East Midlands	E41000018	18
Herefordshire	West Midlands	E41000019	19
Telford and Wrekin	West Midlands	E41000020	20
Stoke on Trent	West Midlands	E41000021	21
Shropshire UA	West Midlands	E41000051	51
Amber Valley	East Midlands	E41000071	71
Bolsover	East Midlands	E41000072	72
Chesterfield	East Midlands	E41000073	73
Derbyshire Dales	East Midlands	E41000074	74
Erewash	East Midlands	E41000075	75
High Peak	East Midlands	E41000076	76
North East Derbyshire	East Midlands	E41000077	77
South Derbyshire	East Midlands	E41000078	78
Blaby	East Midlands	E41000161	161
Charnwood	East Midlands	E41000162	162
Harborough	East Midlands	E41000163	163
Hinckley and Bosworth	East Midlands	E41000164	164
Melton	East Midlands	E41000165	165
North West Leicestershire	East Midlands	E41000166	166
Oadby and Wigston	East Midlands	E41000167	167
Boston	East Midlands	E41000168	168
East Lindsey	East Midlands	E41000169	169
Lincoln	East Midlands	E41000170	170
North Kesteven	East Midlands	E41000171	171
South Holland	East Midlands	E41000172	172
South Kesteven	East Midlands	E41000173	173
West Lindsey	East Midlands	E41000174	174
Corby	East Midlands	E41000182	182
Daventry	East Midlands	E41000183	183
East Northamptonshire	East Midlands	E41000184	184
Kettering	East Midlands	E41000185	185
Northampton	East Midlands	E41000186	186

Name	Region	ONS Code	Label No
South Northamptonshire	East Midlands	E41000187	187
Wellingborough	East Midlands	E41000188	188
Ashfield	East Midlands	E41000196	196
Bassetlaw	East Midlands	E41000197	197
Broxtowe	East Midlands	E41000198	198
Gedling	East Midlands	E41000199	199
Mansfield	East Midlands	E41000200	200
Newark and Sherwood	East Midlands	E41000201	201
Rushcliffe	East Midlands	E41000202	202
Cannock Chase	West Midlands	E41000213	213
East Staffordshire	West Midlands	E41000214	214
Lichfield	West Midlands	E41000215	215
Newcastle-under-Lyme	West Midlands	E41000216	216
South Staffordshire	West Midlands	E41000217	217
Stafford	West Midlands	E41000218	218
Staffordshire Moorlands	West Midlands	E41000219	219
Tamworth	West Midlands	E41000220	220
North Warwickshire	West Midlands	E41000239	239
Nuneaton and Bedworth	West Midlands	E41000240	240
Rugby	West Midlands	E41000241	241
Stratford-on-Avon	West Midlands	E41000242	242
Warwick	West Midlands	E41000243	243
Bromsgrove	West Midlands	E41000251	251
Malvern Hills	West Midlands	E41000252	252
Redditch	West Midlands	E41000253	253
Worcester	West Midlands	E41000254	254
Wychavon	West Midlands	E41000255	255
Wyre Forest	West Midlands	E41000256	256
Birmingham	West Midlands	E41000281	281
Coventry	West Midlands	E41000282	282
Dudley	West Midlands	E41000283	283
Sandwell	West Midlands	E41000284	284
Solihull	West Midlands	E41000285	285
Walsall	West Midlands	E41000286	286
Wolverhampton	West Midlands	E41000287	287

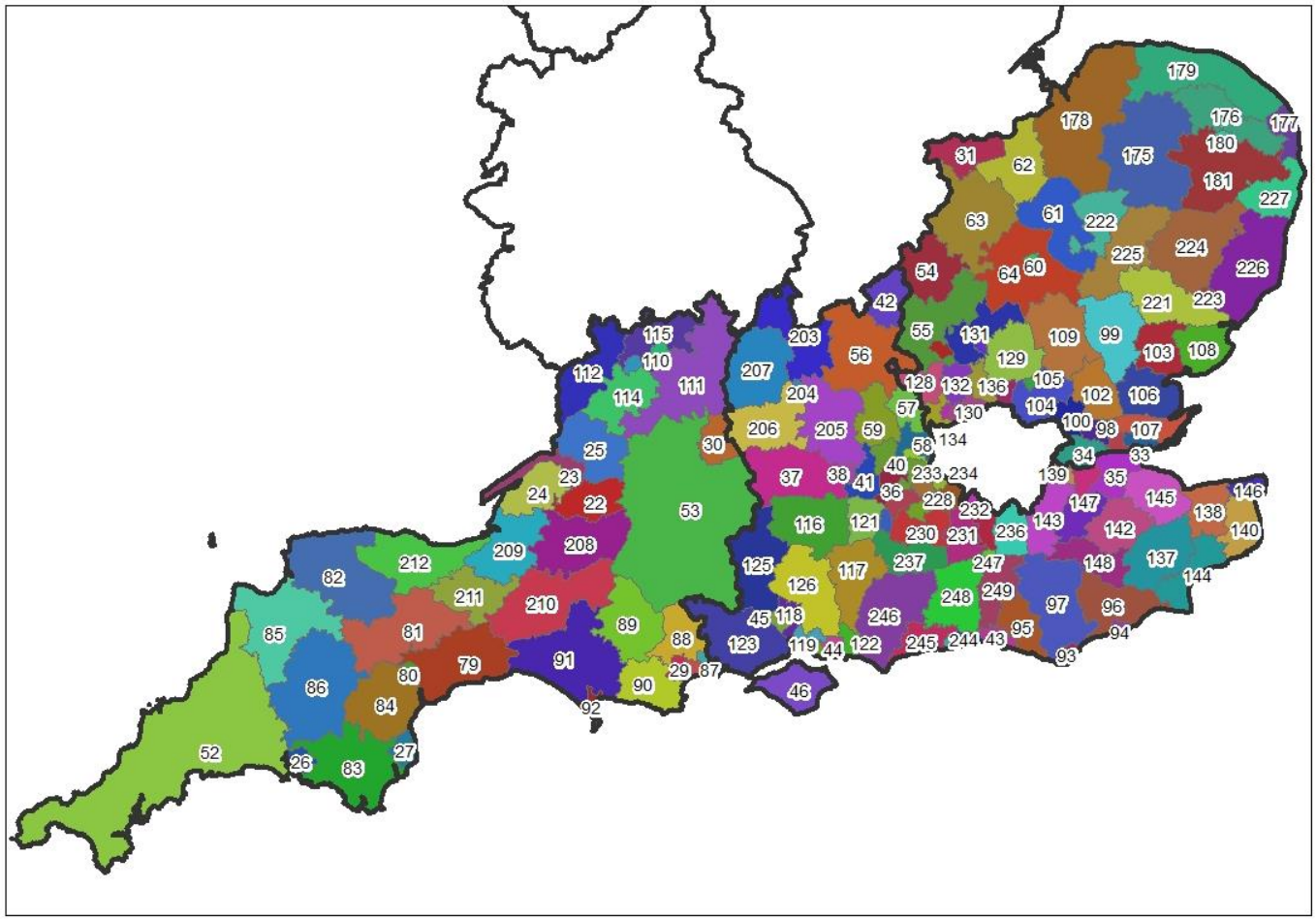


Figure 14-5 : Southern English LADs

Name	Region	ONS Code	Label No
Bath and North East Somerset	South West	E41000022	22
Bristol	South West	E41000023	23
North Somerset	South West	E41000024	24
South Gloucestershire	South West	E41000025	25
Plymouth	South West	E41000026	26
Torbay	South West	E41000027	27
Bournemouth	South West	E41000028	28
Poole	South West	E41000029	29
Swindon	South West	E41000030	30
Peterborough	East of England	E41000031	31
Luton	East of England	E41000032	32
Southend-on-Sea	East of England	E41000033	33
Thurrock	East of England	E41000034	34
Medway	South East	E41000035	35
Bracknell Forest	South East	E41000036	36
West Berkshire	South East	E41000037	37
Reading	South East	E41000038	38
Slough	South East	E41000039	39
Windsor and Maidenhead	South East	E41000040	40
Wokingham	South East	E41000041	41
Milton Keynes	South East	E41000042	42
Brighton and Hove	South East	E41000043	43
Portsmouth	South East	E41000044	44
Southampton	South East	E41000045	45
Isle of Wight	South East	E41000046	46
Cornwall UA	South West	E41000052	52
Wiltshire UA	South West	E41000053	53
Bedford UA	East of England	E41000054	54
Central Bedfordshire UA	East of England	E41000055	55
Aylesbury Vale	South East	E41000056	56
Chiltern	South East	E41000057	57
South Buckinghamshire	South East	E41000058	58
Wycombe	South East	E41000059	59
Cambridge	East of England	E41000060	60
East Cambridgeshire	East of England	E41000061	61
Fenland	East of England	E41000062	62
Huntingdonshire	East of England	E41000063	63
South Cambridgeshire	East of England	E41000064	64

Name	Region	ONS Code	Label No
East Devon	South West	E41000079	79
Exeter	South West	E41000080	80
Mid Devon	South West	E41000081	81
North Devon	South West	E41000082	82
South Hams	South West	E41000083	83
Teignbridge	South West	E41000084	84
Torridge	South West	E41000085	85
West Devon	South West	E41000086	86
Christchurch	South West	E41000087	87
East Dorset	South West	E41000088	88
North Dorset	South West	E41000089	89
Purbeck	South West	E41000090	90
West Dorset	South West	E41000091	91
Weymouth and Portland	South West	E41000092	92
Eastbourne	South East	E41000093	93
Hastings	South East	E41000094	94
Lewes	South East	E41000095	95
Rother	South East	E41000096	96
Wealden	South East	E41000097	97
Basildon	East of England	E41000098	98
Braintree	East of England	E41000099	99
Brentwood	East of England	E41000100	100
Castle Point	East of England	E41000101	101
Chelmsford	East of England	E41000102	102
Colchester	East of England	E41000103	103
Epping Forest	East of England	E41000104	104
Harlow	East of England	E41000105	105
Maldon	East of England	E41000106	106
Rochford	East of England	E41000107	107
Tendring	East of England	E41000108	108
Uttlesford	East of England	E41000109	109
Cheltenham	South West	E41000110	110
Cotswold	South West	E41000111	111
Forest of Dean	South West	E41000112	112
Gloucester	South West	E41000113	113
Stroud	South West	E41000114	114
Tewkesbury	South West	E41000115	115
Basingstoke and Deane	South East	E41000116	116

Name	Region	ONS Code	Label No
East Hampshire	South East	E41000117	117
Eastleigh	South East	E41000118	118
Fareham	South East	E41000119	119
Gosport	South East	E41000120	120
Hart	South East	E41000121	121
Havant	South East	E41000122	122
New Forest	South East	E41000123	123
Rushmoor	South East	E41000124	124
Test Valley	South East	E41000125	125
Winchester	South East	E41000126	126
Broxbourne	East of England	E41000127	127
Dacorum	East of England	E41000128	128
East Hertfordshire	East of England	E41000129	129
Hertsmere	East of England	E41000130	130
North Hertfordshire	East of England	E41000131	131
St Albans	East of England	E41000132	132
Stevenage	East of England	E41000133	133
Three Rivers	East of England	E41000134	134
Watford	East of England	E41000135	135
Welwyn Hatfield	East of England	E41000136	136
Ashford	South East	E41000137	137
Canterbury	South East	E41000138	138
Dartford	South East	E41000139	139
Dover	South East	E41000140	140
Gravesham	South East	E41000141	141
Maidstone	South East	E41000142	142
Sevenoaks	South East	E41000143	143
Shepway	South East	E41000144	144
Swale	South East	E41000145	145
Thanet	South East	E41000146	146
Tonbridge and Malling	South East	E41000147	147
Tunbridge Wells	South East	E41000148	148
Breckland	East of England	E41000175	175
Broadland	East of England	E41000176	176
Great Yarmouth	East of England	E41000177	177
Kings Lynn and West Norfolk	East of England	E41000178	178
North Norfolk	East of England	E41000179	179

Name	Region	ONS Code	Label No
Norwich	East of England	E41000180	180
South Norfolk	East of England	E41000181	181
Cherwell	South East	E41000203	203
Oxford	South East	E41000204	204
South Oxfordshire	South East	E41000205	205
Vale of White Horse	South East	E41000206	206
West Oxfordshire	South East	E41000207	207
Mendip	South West	E41000208	208
Sedgemoor	South West	E41000209	209
South Somerset	South West	E41000210	210
Taunton Deane	South West	E41000211	211
West Somerset	South West	E41000212	212
Babergh	East of England	E41000221	221
Forest Heath	East of England	E41000222	222
Ipswich	East of England	E41000223	223
Mid Suffolk	East of England	E41000224	224
St Edmundsbury	East of England	E41000225	225
Suffolk Coastal	East of England	E41000226	226
Waveney	East of England	E41000227	227
Elmbridge	South East	E41000228	228
Epsom and Ewell	South East	E41000229	229
Guildford	South East	E41000230	230
Mole Valley	South East	E41000231	231
Reigate and Banstead	South East	E41000232	232
Runnymede	South East	E41000233	233
Spelthorne	South East	E41000234	234
Surrey Heath	South East	E41000235	235
Tandridge	South East	E41000236	236
Waverley	South East	E41000237	237
Woking	South East	E41000238	238
Adur	South East	E41000244	244
Arun	South East	E41000245	245
Chichester	South East	E41000246	246
Crawley	South East	E41000247	247
Horsham	South East	E41000248	248
Mid Sussex	South East	E41000249	249
Worthing	South East	E41000250	250

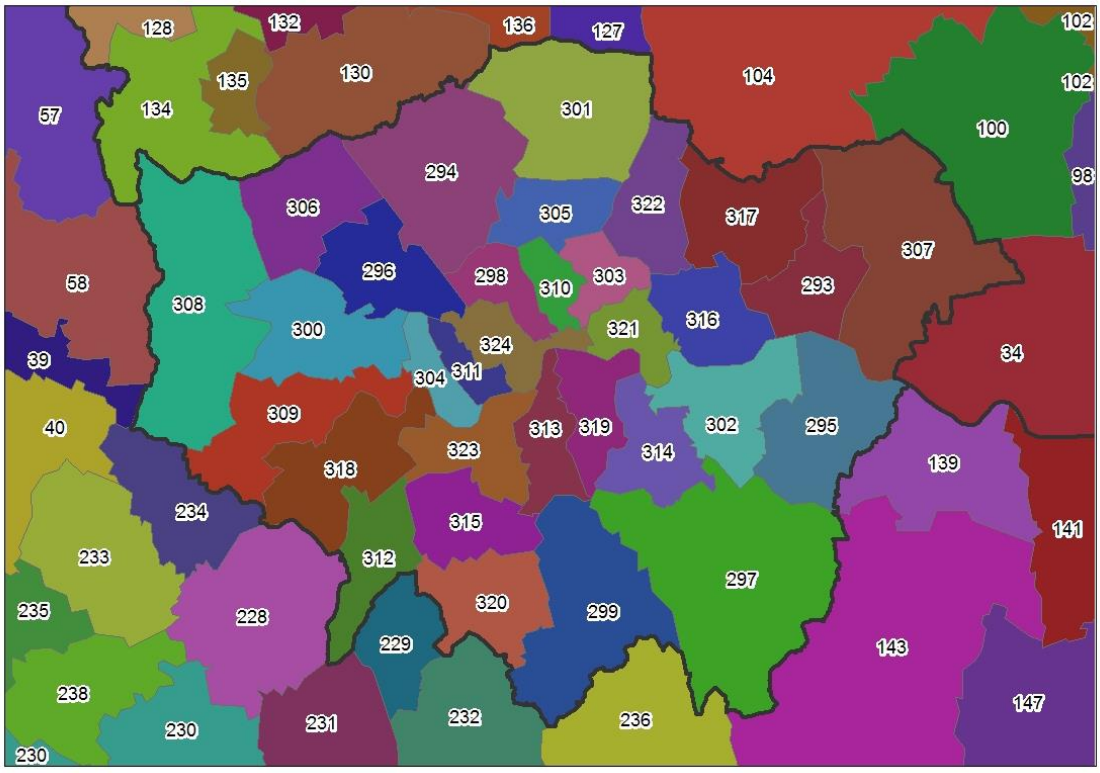


Figure 14-6 : London LADs

Name	ONS Code	Label No
Barking and Dagenham	E41000293	293
Barnet	E41000294	294
Bexley	E41000295	295
Brent	E41000296	296
Bromley	E41000297	297
Camden	E41000298	298
Croydon	E41000299	299
Ealing	E41000300	300
Enfield	E41000301	301
Greenwich	E41000302	302
Hackney	E41000303	303
Hammersmith and Fulham	E41000304	304
Haringey	E41000305	305
Harrow	E41000306	306
Havering	E41000307	307
Hillingdon	E41000308	308

Name	ONS Code	Label No
Hounslow	E41000309	309
Islington	E41000310	310
Kensington and Chelsea	E41000311	311
Kingston upon Thames	E41000312	312
Lambeth	E41000313	313
Lewisham	E41000314	314
Merton	E41000315	315
Newham	E41000316	316
Redbridge	E41000317	317
Richmond upon Thames	E41000318	318
Southwark	E41000319	319
Sutton	E41000320	320
Tower Hamlets	E41000321	321
Waltham Forest	E41000322	322
Wandsworth	E41000323	323
Westminster & City of London	E41000324	324

15 REFERENCES

Alemayehu, B and Warner, KE (2004). The Lifetime Distribution of Health Care Costs. Health Service Research. Vol 39, No 3, pp627–642. http://dx.doi.org/10.1111/j.1475-6773.2004.00248.x

Alkema, L, Gerland, P, Raftery, A and Wilmoth, J (2015). The United Nations Probabilistic Population Projections: An Introduction to Demographic Forecasting with Uncertainty. Foresight: The International Journal of Applied Forecasting, No 37, pp19-24.

Allin, S, Masseria, C and Mossialos, E (2006). Inequality in health care use among older people in the United Kingdom: an analysis of panel data. Working Paper No. 1/2006. Published by LSE Health, London School of Economics, London. http://dx.doi.org/10.1596/1813-9450-3975

Anderson, B (2007). Creating Small Area Income Estimates for England: spatial microsimulation modelling. Chimera Working Paper Number: 2007-07, Chimera, University of Essex. Available at: http://opendepot.org/166/1/CWP-2007-07-Income-Deprivation-England.pdf

Anderson, HR, Gupta, R, Strachan, DP and Limb, ES (2007). 50 years of asthma: UK trends from 1955 to 2004. Thorax, Vol 62, pp85-90. http://dx.doi.org/10.1136/thx.2006.066407

Anderson, RE and Hicks, C (2011). Highlights of Contemporary Microsimulation. Social Science Computer Review, Vol 29, No 1, pp3-8. http://dx.doi.org/10.1177/089443938500100101

Appleby, J (2013). Spending on health and social care over the next 50 years Why think long term? The Kings Fund. Available at : http://www.kingsfund.org.uk/time-to-think-differently/publications/spending-health-and-social-care-over-next-50-years

Atella, V and Conti, V (2013). The effect of age and time to death on health care expenditures: the Italian experience. CEIS Tor Vergata, Research Paper Series, Vol 11, No 4. Available at : http://www.ceistorvergata.it/area.asp?a=415&oc=457&d=500. http://dx.doi.org/10.2139/ssrn.2230487

Auchincloss, AH and Diez Roux, AV (2008). A New Tool for Epidemiology: The Usefulness of Dynamic-Agent Models in Understanding Place Effects on Health. American Journal of Epidemiology. Vol 168, No 1. http://dx.doi.org/10.1093/aje/kwn118

Audit Commission (2004). Assistive Technology: Independence and Well-Being 4. Belmont Press, London. Available at : http://archive.audit-commission.gov.uk/auditcommission/SiteCollectionDocuments/AuditCommissionReports/NationalStudies/NationalReport_FINAL.pdf

Bains, N (2009). Standardization of rates. Association of Public Health Epidemiologists in Ontario. Available at : http://www.apheo.ca/resources/indicators/Standardization%20report_NamBains_FINALMarch16.pdf

Ballas, D, Clarke, G, Dorling, D and Rossiter, D (2007).Using SimBritain to Model the Geographical Impact of National Government Policies. Geographical Analysis, Vol 39, pp44–77. http://dx.doi.org/10.1111/j.1538-4632.2006.00695.x

Ballas, D, Clarke, G, Dorling, D, Eyre, H, Thomas, B and Rossiter, D (2005). SimBritain: A Spatial Microsimulation Approach to Population Dynamics. Population Space Place, Vol 11, pp13–34. http://dx.doi.org/10.1002/psp.351

Ballas, D, Clarke, G, Dorling, D, Rigby, J and Wheeler, B (2006). Using geographical information systems and spatial microsimulation for the analysis of health inequalities. Health Informatics Journal, Vol 12, No 1, pp65–79. http://dx.doi.org/10.1177/1460458206061217

Baltes PB and Smith J (2003). New frontiers in the future of aging: from successful aging of the young old to the dilemmas of the fourth age. Gerontology, Vol 49, No 2, pp123-35. http://dx.doi.org/10.1159/000067946

Bardsley, M, Georghiou, T and Dixon, J (2010). Social care and hospital use at the end of life. The Nuffield Trust for Available at : http://www.nuffieldtrust.org.uk/sites/files/nuffield/social_care_and_hospital_use-full_report_081210.pdf

Barr, B and Bambra, C and Whitehead, M (2014). The impact of NHS resource allocation policy on health inequalities in England 2001-11: longitudinal ecological study. British Medical Journal, Vol 348. http://dx.doi.org/10.1136/bmj.g3231

Barr, PJ, McElnay, JC and Hughes, CM (2012). Connected health care: the future of health care and the role of the pharmacist. Journal of Evaluation in Clinical Practice, Vol 18, No 1, pp56–62. http://dx.doi.org/10.1111/j.1365-2753.2010.01522.x

Bayliss, J and Sly, F (2010) Ageing across the UK - Regional Trends 2010. Office for National Statistics. Available at : http://www.ons.gov.uk/ons/rel/regional-trends/regional-trends/no--42--2010-edition/ageing-across-the-uk--.pdf

Bennett, JE, Li, G, Foreman, K, Best, N, Kontis, V, Pearson, C, Hambly, P and Ezzati, M (2015). The future of life expectancy and life expectancy inequalities in England and Wales; Bayesian spatiotemportal forecasting. The Lancet, on-line first. http://dx.doi.org/10.1016/S0140-6736(15)60296-3

Bilge, U and Saka O (2006). Agent based simulations in healthcare. Studies in Health Technology Information, Vol 124, pp699-704. http://dx.doi.org/10.3233/978-1-58603-647-8-699

Billings, J, Dixon J, Mijanovich T and Wennberg D (2006). Case finding for patients at risk of readmission to hospital: development of algorithm to identify high risk patients. British Medical Journal, Vol 333, No 7563, pp327-332. http://dx.doi.org/10.1136/bmj.38870.657917.AE

Birkin, M and Clarke, M (1989). The Generation of Individual and Household Incomes at the Small Area Level using Synthesis. Regional Studies, Vol 23, No 6, pp535-548. http://dx.doi.org/10.1080/00343408912331345702

Birrell, F, Johnell, O and Silman, A (1999). Projecting the need for hip replacement over the next three decades: influence of changing demography and threshold for surgery. Annals of Rheumatic Diseases, Vol 58, pp569–572. http://dx.doi.org/10.1136/ard.58.9.569

Bodenheimer, T, Lorig, K, Holman, H and Grumbach, K (2002). Patient self-management of chronic disease in primary care. Journal of the American Medical Association, Vol 288, No 19, pp2469–2475. http://dx.doi.org/10.1001/jama.288.19.2469

Bongaarts, J and Bulatao, RA (1999). Completing the Demographic Transition. Population and Development review, Vol 25, No 3, pp515-529. http://dx.doi.org/10.1111/j.1728-4457.1999.00515.x

Boyle, S (2011). Health Systems in Transition. United Kingdom (England). Health System Review. Available at : http://www.euro.who.int/__data/assets/pdf_file/0004/135148/e94836.pdf

Brailsford, SC (2007). Tutorial: Advances and Challenges in Healthcare Simulation Modeling. Proceedings of the 2007 Winter Simulation Conference. Henderson, SG, Biller, B, Hsieh, MH, Shortle, J, Tew, JD and Barton, RR, (eds). http://dx.doi.org/10.1109/WSC.2007.4419754

Brailsford, SC, Harper, PR, Patel, B and Pitt, M (2009). An analysis of the academic literature on simulation and modelling in health care. Journal of Simulation, Vol 3, pp130–140. http://dx.doi.org/10.1057/jos.2009.10

Brice, J, Buck, N and Prentice-Lane, E (2010). British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices. Taylor, MF (ed). Available at : http://www.esds.ac.uk/doc/5151%5Cmrdoc%5Cpdf%5C5151userguide_vola.pdf

British Thoracic Society (2006). The Burden of Lung Disease. A Statistics Report from the British Thoracic Society. Available at : https://www.brit-thoracic.org.uk/document-library/delivery-of-respiratory-care/burden-of-lung-disease/burden-of-lung-disease-2006/

Brown, L (2011). Editorial Special Issue on 'Health and Microsimulation'. International Journal of Microsimulation, Vol 4, No 3 pp1-2.

Brown, L, Harris, A, Picton, M, Thurecht, L, Yap, M, Harding, A, Dixon, P and Richardson, J (2009). Linking Microsimulation and Macro-Economic Models to Estimate the Economic Impact of Chronic Disease Prevention NATSEM, University of Canberra, Canberra, Australia. Available at : https://guard.canberra.edu.au/natsem/index.php?mode=download&file_id=780

Bruggink, JW (2011). Men die, women suffer? Gender differences in health expectancy. Presentation at Réseau Espérance de Vie en Santé (REVES) 2011. Available at : https://reves.site.ined.fr/fichier/s_rubrique/20054/bruggink.fr.pdf

Buck, N and McFall, S (2012). Understanding Society: design overview. Longitudinal and Life Course Studies, Vol 3, No 1, pp5-17. Available at : http://www.llcsjournal.org/index.php/llcs/article/view/159

CACI (2007). HealthACORN Postcode-Level Directory for the United Kingdom. UK Data Archive Study Number 6070. Available at : http://www.esds.ac.uk/doc/6070%5Cmrdoc%5Cpdf%5C6070_HealthAcorn_userguide.pdf

Caley, M and Sidhu, K (2010). Estimating the future healthcare costs of an aging population in the UK: expansion of morbidity and the need for preventative care. Journal of Public Health, Vol 33, No 1, pp117–122. http://dx.doi.org/10.1093/pubmed/fdq044

Calot, G and Sardon, JP (2003). Methodology for the calculation of Eurostat's demographic indicators. Detailed report by the European Demographic Observatory. Population and social conditions 3/2003/F/no 26. Available at : http://www.cedefop.europa.eu/files/Methodology_for_the_calculation_of_Eurostats_demographic_indicators.pdf

Charfeddine, M and Montreuil, B (2010). Integrated Agent-Oriented Modeling and Simulation of Population and Healthcare Delivery Network: Application To COPD Chronic Disease in a Canadian Region. Proceedings of the 2010 Winter Simulation Conference. Johansson, B, Jain, S, Montoya-Torres, J, Hugan, J and Yücesan, E, (eds). http://dx.doi.org/10.1109/WSC.2010.5678930

Charlton, J (1998). Use of the Census Samples of Anonymised Records (SARs) and survey data in combination to obtain estimates at local authority level. Environment and Planning A, Vol 30, No 5, pp775-784. http://dx.doi.org/10.1068/a300775

Chatterjee, T, MacDonald, AS & Waters, HR (2008). A model for ischaemic heart disease and stroke III: Applications. Annals of Actuarial Science, Vol 3, pp105-119. http://dx.doi.org/10.1017/S1748499500000476

Chaturvedi, N (2003). Ethnic differences in cardiovascular disease. Heart, Vol 89, pp681–686. http://dx.doi.org/10.1136/heart.89.6.681

Cheshire, H, Hussey, D, Medina, J, Pickering, K, Wood, N, Ward, K, Taylor, K and Carli Lessof, C (2012). Financial circumstances, health and well-being of the older population in England: The 2008 English Longitudinal Study of Ageing Technical Report. National Centre for Social Research. Available at : http://www.ifs.org.uk/elsa/report10/w4_tech.pdf

Christensen, K, Doblhammer, G, Rau, R and Vaupel, JW (2009). Ageing populations: the challenges ahead. Lancet, Vol 374, pp1196–1208. http://dx.doi.org/10.1016/S0140-6736(09)61460-4

Clark, RL, Burkhauser, RV, Moon, M, Quinn, JF and Smeeding, TM (2004). The Economics of an Ageing Society. Part IV Health and Long-Term Care for Older Persons. Blackwell Publishing

Clark, S.D., Birkin, M. & Heppenstall, A.J. (2014). Sub regional estimates of morbidities in the English elderly population. Health and Place, Vol 27, pp176-185. http://dx.doi.org/10.1016/j.healthplace.2014.02.01

Clement Clarke International (2015). Predictive Normal Values (Nomogram, EU scale). Available at : http://www.peakflow.com/top_nav/normal_values/

Cockings, S, Harfoot, A, Martin, D and Hornby, D (2011). Maintaining existing zoning systems using automated zone-design techniques: methods for creating the 2011 Census output geographies for England and Wales. Environment and Planning A, Vol 43, pp2399-2418. http://dx.doi.org/10.1068/a43601

Coleman (2010). Projections of the Ethnic Minority populations of the United Kingdom, 2006-2056. Population and Development Review, Vol 36, No 3, pp441 – 486. http://dx.doi.org/10.1111/j.1728-4457.2010.00342.x

Cowper, A (2004). A New Data Dawns. Health Service Journal, Vol 114, No 5935, pp12-13

Craig, R and Mindell, J (2005). Health Survey for England. The health of older people Summary of key findings. Available at : http://www.ic.nhs.uk/webfiles/publications/hseolder/HSESummary.pdf

Crooks, A, Castle, C and Batty, M (2008). Key challenges in agent-based modelling for geo-spatial simulation. Computers, Environment and Urban Systems, Vol 32, pp417–430. http://dx.doi.org/10.1016/j.compenvurbsys.2008.09.004

Crooks, AT and Heppenstall, AJ (2012). Introduction to Agent-Based Modelling. In Agent-Based Modelling of Geographical Systems, Heppenstall, AJ, Crooks, AT, See, LM and Batty, M (eds), Springer. http://dx.doi.org/10.1007/978-90-481-8927-4_5

Darkins, A, Ryan, P, Kobb, R, Foster, L, Edmonson, E, Wakefield, B and Lancaster, AE (2008). Care coordination/home telehealth: the systematic implementation of health informatics, home telehealth and disease management to support the care of veteran patients with chronic conditions. Telemedicine and E-Health, Vol 14, No 10, pp1118–1126. http://dx.doi.org/10.1089/tmj.2008.0021

de Meijer C, Koopmanschapa, M, d'Uvac, TB and van Doorslaera, E (2011). Determinants of long-term care spending: Age, time to death or disability? Journal of Health Economics 30 (2011), Vol 30, No 2, pp425–438. http://dx.doi.org/10.1016/j.jhealeco.2010.12.010

Deming, WE and Stephan, FF (1940). On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known. The Annals of Mathematical Statistics, Vol 11, No 4, pp427-444. http://dx.doi.org/10.1214/aoms/1177731829

Denton, FT, Gafni, A and Spencer, BG (2002). Exploring the effects of population change on the costs of physician services. Journal of Health Economics, Vol 21, pp781–803. http://dx.doi.org/10.1016/S0167-6296(02)00013-9

Department of Business, Innovation and Skills (2010a). Is business ready for an ageing nation? Economic opportunities and challenges of ageing: Discussion paper. Ageing Population Team, Department of Business, Innovation and Skills, Victoria Street, London. Available at : http://storage.globalcitizen.net/data/topic/knowledge/uploads/20120203124321302.pdf

Department of Business, Innovation and Skills (2010b). Is business ready for an ageing nation? Economic opportunities and challenges of ageing: Analytical paper. Ageing Population Team, Department of Business, Innovation and Skills, Victoria Street, London. Available at : http://www.researchonline.org.uk/sds/search/download.do?ref=B15528

Department of Health (2011). Whole System Demonstrator, Programme Headline Findings – December 2011. Available at : http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_131689.pdf

Department of Health (2012). The new public health role of local authorities. Gateway reference: 17876. Available at : https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/213009/Public-health-role-of-local-authorities-factsheet.pdf

Devereux, G (2006). ABC of chronic obstructive pulmonary disease Definition, epidemiology, and risk factors. British Medical Journal, Vol 332, No 7550, pp1142–1144. Available at : http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1459603/. http://dx.doi.org/10.1136/bmj.332.7550.1142

Diabetes UK (2012). Diabetes in the UK 2012. Available at : http://www.diabetes.org.uk/Documents/Reports/Diabetes-in-the-UK-2012.pdf

Diabetes UK (2015a).Cost of Diabetes. Available at : http://www.diabetes.co.uk/cost-of-diabetes.html

Diabetes UK (2015b). Diagnosis and Tests. Available at : http://www.diabetes.co.uk/diabetes_care/blood-sugar-level-ranges.html ; http://www.diabetes.co.uk/oral-glucose-tolerance-test.html ; http://www.diabetes.co.uk/what-is-hba1c.html

Dixon, T, Shaw, M, Frankel, S and Ebrahim, S (2004). Hospital admissions, age and death: retrospective cohort study. British Medical Journal, Vol 328, pp1-4. http://dx.doi.org/10.1136/bmj.38072.481933.EE

Dormont, B, Grignonc, M and Huberd, H (2006). Health expenditure growth: reassessing the threat of ageing. Health Economics, Vol 15, pp947–963. http://dx.doi.org/10.1002/hec.1165

Edwards, KL and Clarke, GP (2009). The design and validation of a spatial microsimulation model of obesogenic environments for children in Leeds, UK: SimObesity. Social Science and Medicine, Vol 69, pp1127–1134. http://dx.doi.org/10.1016/j.socscimed.2009.07.037

Edwards, KL, Clarke, GP, Thomas, J and Forman, D (2011). Internal and External Validation of Spatial Microsimulation Models: Small Area Estimates of Adult Obesity. Applied Spatial Analysis and Policy, Vol 4, No 4, pp281-300. http://dx.doi.org/10.1007/s12061-010-9056-2

Erler, A, Bodenheimer, T, Baker, R, Goodwin, N, Spreeuwenberg, C, Vrijhoef, HJM, Nolte, E and Ferdinand M (2011). Preparing primary care for the future – perspectives from the Netherlands, England and USA. Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen (Journal of Evidence, Education and Quality in Healthcare), Vol 105, No 8, pp571–580. http://dx.doi.org/10.1016/j.zefq.2011.09.029

European Commission (2014). The 2015 Ageing Report. Underlying Assumptions and Projection Methodologies. European Economy 8|2014. Available at http://ec.europa.eu/economy_finance/publications/european_economy/2014/pdf/ee8_en.pdf

Evandrou, M, Falkingham, J and Green, M (2010). Migration in later life: evidence from the British Household Panel Study. Population Trends, No 141, pp77-94. http://dx.doi.org/10.1057/pt.2010.22

Experian (2003). Mosaic United Kingdom. The Consumer Classification for the UK. Available at : http://www.experian.co.uk/assets/business-strategies/brochures/Mosaic_UK_2003_brochure%5B1%5D.pdf

Faiz, O, Warusavitarne, J, Bottle, A, Tekkis, PP, Clark, SK, Darzi, AW and Aylin, P (2010). Nonelective Excisional Colorectal Surgery in English National Health Service Trusts: A Study of Outcomes from Hospital Episode Statistics Data between 1996 and 2007. Journal of the American College of Surgeons, Vol 210, No 4, pp390-401. http://dx.doi.org/10.1016/j.jamcollsurg.2009.11.017

Fries, JF (1980). Aging, natural death and the compression of morbidity. New England Journal of Medicine, Vol 570, pp130–5. http://dx.doi.org/10.1056/NEJM198007173030304

FrontlineSolver(2014). Developers of the EXCEL solver. Available at : http://www.solver.com/

Fukawa, T (2011). Household Projection and its Application to Health/Long-Term Care Expenditures in Japan Using INAHSIM-II. Social Science Computer Review, Vol 29, No 1, pp52-66. http://dx.doi.org/10.1177/0894439310370097

Garcia-Retamero, R and Galesic, M (2014). On defensive decision making: how doctors make decisions for their patients. Health Expectations, Vol 17, No 5, pp664–669. http://dx.doi.org/ 10.1111/j.1369-7625.2012.00791.x

Garratt, E, Barnes, H and Dibben, C (2010). Health administrative data : Exploring the potential for academic research. St Andrews: Administrative Data Liaison Service. Available at : http://www.adls.ac.uk/wp-content/files_flutter/1306411298ADLSHealthResearchPaperprint.pdf

Gilbert, R, Todd, C, M. May, M, Yardley, L and Ben-Shlomo, Y (2009). Socio-demographic factors predict the likelihood of not returning home after hospital admission following a fall. Journal of Public Health, Vol 32, No 1, pp117–124. http://dx.doi.org/10.1093/pubmed/fdp077

Glaser, K, Stuchbury, R, Tomassini, C and Askham, J (2008). The long-term consequences of partnership dissolution for support in later life in the United Kingdom. Ageing and Society, Vol 28, pp329–351. http://dx.doi.org/10.1017/S0144686X07006642

Glover, G (2008). Estimating the future number of cases of dementia in PCTs and upper tier local authorities in England. Mental Health Observatory Briefing No 3, May. Available at : http://www.nepho.org.uk/uploads/doc262_1_brief003_10.pdf

Goldeing, S, Henretty, N, Mills, K, Johnson, K and Smallwood, S (2011). Mortality of the 'Golden generation': What can cause the ONS Longitudinal Study tell us? Population Trends, No 145, pp1-8.

Goldman, DP, Shekelle, PG, Bhattacharya, J, Hurd, M, Joyce, GF, Lakdawalla, DN, Matsui, DH, Newberry, SJ, Panis, CWA and Shang, B (2004). Health Status and Medical Treatment of the Future Elderly Final Report. Available at : http://www.rand.org/pubs/technical_reports/2005/RAND_TR169.pdf

Gordon, AL, Franklin, M, Bradshaw, L, Logan, P, Elliott, R and Gladman, JRF (2014). Health status of UK care home residents: a cohort study. Age and Ageing, Vol 43, No 1, pp97-103. http://dx.doi.org/10.1093/ageing/aft077

Gray, A (2009). The social capital of older people. Ageing and Society, Vol 29, pp5–31. http://dx.doi.org/10.1017/S0144686X08007617

Gruenberg, EM (1977). The failures of success. Milbank Memorial Foundation Quarterly/Health and Society, Vol 55, pp3-24. http://dx.doi.org/10.2307/3349592

Grundy, E and Sloggett, A (2003). Health inequalities in the older population: the role of personal capital, social resources and socio-economic circumstances. Social Science and Medicine, Vol 56, pp935–947. http://dx.doi.org/10.1016/S0277-9536(02)00093-X

Gupta, A and Harding, A (eds) (2007). Modelling Our Future : Population Ageing, Health and Aged Care. International Symposia in Economic Theory and Econometrics. Elsevier

Guradian (2015). Obesity is 'the new smoking' says NHS England's chief executive. Available at : http://www.theguardian.com/society/2015/may/31/obesity-new-smoking-nhs-england-chief-executive

Hanratty, B, Jacoby, A and Whitehead, M (2008). Socio-economic differences in service use, payment and receipt of illness-related benefits in the last year of life: findings from the British Household Panel Survey. Source: Palliative Medicine, Vol 22, No 3, pp248-255. http://dx.doi.org/10.1177/0269216307087140

Harding, A, Vidyattama, Y and Tanton, R (2011). Demographic change and the needs-based planning of government services: projecting small area populations using spatial microsimulation. Journal of Population Research, Vol 28, pp203–224. http://dx.doi.org/10.1007/s12546-011-9061-6

Harker, R (2009). NHS bed availability in England. Standard Note: SN/SG/1359, House of Common Library, London

Harland, K (2013). Microsimulation model user guide. Flexible Modelling Framework. National Centre for Research Methods Working Paper 06/13. Available at : http://eprints.ncrm.ac.uk/3177/2/microsimulation_model.pdf

Harper, S, Howse, K and Baxter, S (2011). Living longer and prospering? Designing an adequate, sustainable and equitable UK state pension system. Club Vita LLP and Oxford Institute of Ageing. Available at : http://www.ageing.ox.ac.uk/system/files/110110%20Living%20Longer%20and%20prospering%20Final.pdf

Health and Social Care Information Centre (2005a). Health Survey for England 2004: The Health of Minority Ethnic Groups – headline tables. Available at : http://www.ic.nhs.uk/webfiles/publications/hlthsvyeng2004ethnic/HealthSurveyForEngland161205_PDF%20.pdf

Health and Social Care Information Centre (2005b). The health of minority ethnic groups, Volume 1. Edited by Sproston, K and Mindell, J. Available at : http://www.ic.nhs.uk/pubs/hse04ethnic

Health and Social Care Information Centre (2014a). The Health and Care of Older People. Available at http://www.hscic.gov.uk/catalogue/PUB14369/focu-on-hac-op-main-pub-doc%201.1.pdf

Health and Social Care Information Centre (2014b). Quality and Outcomes Framework – Prevalence, Achievements and Exceptions Report, England, 2013-14. Available at : http://www.hscic.gov.uk/catalogue/PUB15751/qof-1314-report.pdf

Health and Social Care Information Centre (2015). Hospital Episode Statistics, Admitted Patient Care, England, 2013-14 . Available at : http://www.hscic.gov.uk/article/2021/Website-Search?productid=17192&q=FCE+primary+diagnosis&sort=Relevance&size=10&page=1&area=both#top Data available at : http://www.hscic.gov.uk/catalogue/PUB16719/hosp-epis-stat-admi-diag-2013-14-tab.xlsx

Health in Europe: Information and Data Interface (2013). Health Status Indicators. http://epp.eurostat.ec.europa.eu/cache/ITY_SDDS/EN/proj_13n_esms.htm. Data available at : http://ec.europa.eu/health/indicators/echi/list/echi_40.html

Heart UK (2015a). High-density lipoprotein. Available at : http://heartuk.org.uk/files/uploads/huk_fs_mfsR_HDL_02b.pdf

Heart UK (2015b). Cholesterol tests – know your number. Available at : http://heartuk.org.uk/health-and-high-cholesterol/cholesterol-tests---know-your-number

Hermes, K and Poulsen, M (2012a). A review of current methods to generate synthetic spatial microdata using reweighting and future directions. Computers, Environment and Urban Systems, Vol 36, pp281–290. http://dx.doi.org/10.1016/j.compenvurbsys.2012.03.005

Hermes, K and Poulsen, M (2012b). Small area estimates of smoking prevalence in London. Testing the effect of input data. Health and Place, Vol 18, No 3, pp630-638. http://dx.doi.org/10.1016/j.healthplace.2011.12.010

Herrett, E, Gallagher, AM, Bhaskaran, K, Forbes, H, Mathur, R, van Staa, T and Smeeth, L (2015). Data Resource Profile: Clinical Practice Research Datalink (CPRD). International Journal of Epidemiology, pp 827–836. http://dx.doi.org/10.1093/ije/dyv098

Herrett, E, Thomas, SL, Schoonen, WM, Smeeth, L and Hall, AJ (2009). Validation and validity of diagnoses in the General Practice Research Database: a systematic review. British Journal of Clinical Pharmacology, Vol 69, No 1, pp4-14. http://dx.doi.org/10.1111/j.1365-2125.2009.03537.x

Hippisley-Cox, J and Vinogradova, Y (2008). Trends in Consultation Rates in General Practice 1995 to 2008: Analysis of the QResearch® database Final Report to the NHS Information Centre and Department of Health. Available at : http://www.ic.nhs.uk/webfiles/publications/gp/Trends_in_Consultation_Rates_in_General_Practice_1995_2008.pdf

Hirsch, GB and Immediato, CS (1999). Microworlds and generic structures as resources for integrating care and improving health. System Dynamics Review Vol. 15, No 3, pp315-330. http://dx.doi.org/10.1002/(SICI)1099-1727(199923)15:3<315::AID-SDR168>3.0.CO;2-3

Hirst, M (2004). Health Inequalities and Informal Care: End of Project Report. DH 1992 - January 2004 Social Policy Research Unit, University of York, York

Hoffmann, E and Nachtmann, J (2010).Old Age, the Need of Long-term Care and Healthy Life Expectancy. In Ageing, Care Need and Quality of Life : The Perspective of Care Givers and People in Need of Care, Doblhammer, G and Scholz, R (eds). VS Research. http://dx.doi.org/10.1007/978-3-531-92335-2_9

Homer, JB and Hirsch, GB (2006). System Dynamics Modeling for Public Health: Background and Opportunities. American Journal of Public Health. Vol 96, No3, pp452-8. http://dx.doi.org/10.2105/AJPH.2005.062059

Hosmer DW and Lemeshow S (1980). A goodness-of-fit test for the multiple logistic regression model. Communications in Statistics, Vol 9, No 10, pp1043-1069. http://dx.doi.org/10.1080/03610928008827941

House of Lords (2013). Ready for Ageing? House of Lords Select Committee Report. Available at : http://www.parliament.uk/business/committees/committees-a-z/lords-select/public-services-committee/report-ready-for-ageing/

Howse, K (2012). Healthy ageing: the role of health care services. Perspectives in Public Health, Vol 132, No 4, pp171-177. http://dx.doi.org/10.1177/1757913912444805

Institute for Fiscal Studies (2012). The Dynamics of Ageing: Evidence from the English Longitudinal Study of Ageing 2002-2010 Wave 5. Edited by Banks, J, Nazroo, J and Steptoe, A. Available at : http://www.elsa-project.ac.uk/publicationDetails/id/6367

Institute for Fiscal Studies (2015a). ELSA Main Reports. Available at : http://www.elsa-project.ac.uk/publications

Institute for Fiscal Studies (2015b). ELSA Technical Reports. Available at : http://www.elsa-project.ac.uk/publications/case/technical

Jagger, C, Cox, B and Le Roy, S (2007). Health Expectancy Calculation by the Sullivan Method: A Practical Guide. EHEMU Technical report 2006_3 June 2007. European Health Expectancy Monitoring Unit

Jagger, C, Matthews, R, Spiers, N, Brayne, C, Comas-Herrera, A, Robinson, T, Lindesay, R and Croft, P (2006). Compression or expansion of disability? Forecasting Future Disability Levels under Changing Patterns of Diseases. (Background paper for the Wanless Social Care Review). King's Fund, London. Available at : http://www.kingsfund.org.uk/document.rm?id=8290

Jeyarajah, S, Faiz, O, Bottle, A, Ayli, P, Bjarnason, I, Tekkis, PP and Papagrigoriadis, S (2009) Diverticular disease hospital admissions are increasing, with poor outcomes in the elderly and emergency admissions. Alimentary Pharmacology and Therapeutics, Vol 30, pp1171–1182. http://dx.doi.org/10.1111/j.1365-2036.2009.04098.x

Johnson, B (2011). Deriving trends in life expectancy by the National Statistics Socio-economic Classification using the ONS Longitudinal Study. Health Statistics Quarterly Vol 49. Available at : http://www.ons.gov.uk/ons/rel/hsq/health-statistics-quarterly/spring-2011/deriving-trends-in-life-expectancy-by-the-national-statistics-socio-economic-classification-using-the-ons-longitudinal-study.pdf. http://dx.doi.org/10.1057/hsq.2011.2

Joint Health Surveys Unit (2012). Health Survey for England – Trend Tables. Available at : https://catalogue.ic.nhs.uk/publications/public-health/surveys/heal-survey-eng-2011-tren-tabl/HSE2011-Trend-commentary.pdf

Jordan, R, Birkin, M and Evans, A (2010). An agent-based model of residential mobility: Assessing the impacts of urban regeneration policy in the EASEL district. Computers, Environment and Urban Systems, vol 48, pp49–63. http://dx.doi.org/10.1016/j.compenvurbsys.2014.06.006

Judge, A, Welton, NJ, Sandhu, J and Ben-Shlomo, Y (2010). Equity in access to total joint replacement of the hip and knee in England: cross sectional study. BMJ Research, pp1-10. http://dx.doi.org/10.1136/bmj.c4092

Kanagarajah, AK, Lindsay, P, Miller, A and Parker, D (2006). An Exploration into the Uses of Agent-Based Modeling to Improve Quality of Health Care. 6th International Conference on Complex Systems. Boston, MA, 25-30 June, 2006

Kanavos, P, van den Aardweg, S and Schurer, W (2012). Diabetes expenditure, burden of disease and management in 5 EU countries. LSE Health, London School of Economics, January 2012. Available at : http://www.lse.ac.uk/LSEHealthAndSocialCare/research/LSEHealth/MTRG/LSEDiabetesReport26Jan2012.pdf

Kannel, WB, Wolf, PA, Castelli, WP, D'Agostino, RB (1987). Fibrinogen and Risk of Cardiovascular Disease The Framingham Study. Journal of the American Medical Association. Vol 258, No 9, pp1183-1186. http://dx.doi.org/10.1001/jama.1987.03400090067035

Karlsson, M, Mayhew, L and Rickayzen, B (2009). Individualised Life Tables Investigating Dynamics of Health, Work and Cohabitation in the UK. Journal of Population Ageing, 2009, Vol 1, No 2-4, pp153-191. Also available at: http://www.ageing.ox.ac.uk/files/workingpaper_108.pdf . http://dx.doi.org/10.1007/s12062-009-9008-2

Karlsson, M, Mayhew, L, Plumb, R and Rickayzen, B (2007). Long term care financing in four OECD countries: Fiscal burden and distributive effects. Health Policy, Vol 80, No 1, pp107–134. http://dx.doi.org/10.1016/j.healthpol.2006.02.002

Kasteridis, P, Street, A, Dolman, M, Gallier, L, Hudson, K. Martin, J, Wyer, I (2012). The importance of multimorbidity in explaining utilisation and costs across health and social care settings: evidence from South Somerset's Symphony Project. Available at : http://www.york.ac.uk/media/che/documents/papers/researchpapers/CHERP96_multimorbidity_utilisation_costs_health_social%20care.pdf

Khan, NF, Harrison, SE and Rose, PW (2010). Validity of diagnostic coding within the General Practice Research Database : a systematic review. British Journal of General Practice, March, pp e128-e136. http://dx.doi.org/10.3399/bjgp10X483562

Khunti, K, Kumar, S and Brodie, J (2009). Diabetes UK and South Asian Health Foundation recommendations on diabetes research priorities for British South Asians. Available at : https://www.diabetes.org.uk/upload/Reports/South_Asian_report.pdf

Klodawski, E (2009). GLA 2008 Round Ethnic Group Population Projections. DMAG Briefing 2009-08. Available at :

http://www.lho.org.uk/Download/Public/15148/1/DMAG%20Briefing%202009-08%20R2008%20Low%20Ethnic%20Group%20Population%20Projections.pdf

Kolappan, C and Gopi, PG (2002). Tobacco smoking and pulmonary tuberculosis. Thorax Vol 57, pp964-966. http://dx.doi.org/10.1136/thorax.57.11.964

Kramer, M (1980). The rising pandemic of mental disorders and associated chronic diseases and disabilities. Acta Psychiatrica Scandinavica, No 62, pp382–97. http://dx.doi.org/10.1111/j.1600-0447.1980.tb07714.x

Kruskal, WH and Wallis, WA (1952). Use of ranks in one-criterion variance analysis. Journal of the American Statistical Association, Vol 47, No 260, pp583–62. http://dx.doi.org/10.1080/01621459.1952.10483441

Lam, P, Elliott, RA and George, J (2011). Impact of a self-administration of medications programme on elderly inpatients' competence to manage medications: a pilot study. Journal of Clinical Pharmacy and Therapeutics, Vol 36, pp80–86. http://dx.doi.org/10.1111/j.1365-2710.2009.01157.x

Laurence, J (2002). Why an ageing population is the greatest threat to society. The Independent, Wednesday 10th April, 2002. Available at : http://www.independent.co.uk/news/uk/home-news/why-an-ageing-population-is-the-greatest-threat-to-society-656997.html

Law, MR, Wald, NJ and Rudnicka, AR (2003). Quantifying effect of statins on low density lipoprotein cholesterol, ischaemic heart disease, and stroke: systematic review and meta-analysis. British Medical Journal, Vol 326, pp1423-1429. http://dx.doi.org/10.1136/bmj.326.7404.1423

Légaré, J (2011). Using Statistics Canada LifePaths Microsimulation Model to Project the Disability Status of Canadian Elderly. International Journal of Microsimulation, Vol 4, No 3, pp48-56.

Lievesley, N (2010). The future ageing of the ethnic minority population of England and Wales. Runnymede and The Centre for Policy on Ageing, London. Available at : http://www.cpa.org.uk/information/reviews/thefutureageingoftheethnicminoritypopulationofenglandandwales.pdf

Lindeman, D, Wang, A, Steinmetz, V and Redington, L (2011a). Editorial: Introduction to Special Issue on Technology Adoption and Diffusion for Older Adults (Issue 1 of 2). Ageing International Vol 36, No 1, pp1–4. http://dx.doi.org/10.1007/s12126-010-9078-8

Lindeman, D, Wang, A, Steinmetz, V and Redington, L (2011b). Editorial: Introduction to Special Issue on Technology Adoption and Diffusion for Older Adults (Issue 2 of 2). Ageing International Vol 36, No 2, pp155-158. http://dx.doi.org/10.1007/s12126-010-9084-x

Local Government Association (2014). Public health in local government– one year on. Available at : http://www.local.gov.uk/documents/10180/5854661/L14-150_Public+Health+in+LG_08.pdf

Lubitz, J, Cai, L, Kramarow, E and Lentzner, H (2003). Health, Life Expectancy and Health Care Spending among the Elderly. New England Journal of Medicine, Vol 349, No 11, pp1048-1055. http://dx.doi.org/10.1056/NEJMsa020614

Lymer, S, Brown, L, Harding, A and Yap, M (2009). Predicting The Need For Aged Care Services At The Small Area Level: The CAREMOD Spatial Microsimulation Model. International Journal of Microsimulation, Vol 2, No 2, pp27-42

Lymer, S, Brown, L, Yap, M and Harding, A (2008). 2001 Regional Disability Estimates for New South Wales, Australia, Using Spatial Microsimulation. Applied Spatial Analysis and Policy, Vol 1, No 2, pp99-116. http://dx.doi.org/10.1007/s12061-008-9006-4

Ma, J, Heppenstall, A, Harland, K and Mitchell, G (2014). Synthesising carbon emission for mega-cities: A static spatial microsimulation of transport CO2 from urban travel in Beijing. Computers, Environment and Urban Systems, vol 45, pp78–88. http://dx.doi.org/10.1016/j.compenvurbsys.2014.02.006

Macal, CM and North, MJ (2010). Tutorial on agent-based modelling and simulation. Journal of Simulation, Vol 4, pp151–162. http://dx.doi.org/10.1057/jos.2010.3

Manton, KG (1982). Changing concepts of morbidity and mortality in the elderly population. Milbank Memorial Fund Quarterly, Vol 60, pp183–244. http://dx.doi.org/10.2307/3349767

Marshall, A (2009). Developing a methodology for the estimation and projection of limiting long term illness and disability, PhD Thesis, School of Social Sciences, University of Manchester

Martin, D, Nolan, A and Tranmer, M (2001). The application of zone design methodology to the 2001 UK Census. Environment and Planning A, Vol 33, pp1949-1962. http://dx.doi.org/10.1068/a3497

Mason, T, Sutton, M, Whittaker, W and Birch, S (2015). Exploring the limitations of age-based models for health care planning. Social Science and Medicine, Vol 132, pp11–19. http://dx.doi.org/10.1016/j.socscimed.2015.03.005

Mathur, R, Bhaskaran, K, Chaturvedi, N, Leon, DA, vanStaa, T, Grundy, E, and Smeeth, L (2014). Completeness and usability of ethnicity data in UK-based primary care and hospital databases. Journal of Public Health, Vol 36, No 4, pp684-692. http://dx.doi.org/10.1093/pubmed/fdt116

Matsumoto, M and Kurita, Y (1994). Twisted GFSR Generators II. ACM Transactions on Modelling and Computer Simulation, Vol 4, No 3, pps 254-266. http://dx.doi.org/10.1145/189443.189445

Maurya V, Vijayan VK and Shah, A (2002). Smoking and tuberculosis: an association overlooked. International Journal of Tuberculosis and Lung Disease, Vol 6, No 11, pp942-51

Mayhew, L (2005). Active Ageing in the UK-Issues, Barriers, Policy Directions, Innovation. The European Journal of Social Science Research, 2005, Vol 18, No 4, 455-477. http://dx.doi.org/10.1080/13511610500384202

Mayhew, L (2009). Increasing longevity and the economic value of healthy ageing and working longer. Cass Business School February 2009. Available at: http://www.pensions-institute.org/reports/increasinglongevity_report.pdf

Mayhew, L and Rickayzen, B (2012). The Ageing Population : Crunch Time for Government Reforms. Economic Affairs, Vol 32, No 2, pp96–100. http://dx.doi.org/10.1111/j.1468-0270.2012.02165.x

Mayhew, L and Smith, D (2012). Gender Convergence in Human Survival and the Postponement of Death. Actuarial research Paper No 200, faculty of Actuarial Science and Insurance, Cass Business School, City University London. Available at : http://www.cass.city.ac.uk/__data/assets/pdf_file/0019/136243/200ARP.pdf

McCulloch, A (2011). Health Service funding. Available at : http://www.significancemagazine.org/details/webexclusive/1385441/Health-Service-funding.html

McCulloch, A (2012a). Population ageing and mortality. Available at : http://www.significancemagazine.org/details/webexclusive/1664457/Population-ageing-and-mortality.html

McCulloch, A (2012b). Standardising death. Available at : http://www.significancemagazine.org/details/webexclusive/1501135/Standardising-Death.html

Melzer, D. McWilliams, B., Brayne, C., Johnson, T and Bond, J (2000). Socio-economic status and the expectation of disability in old age: estimates for England. The Medical Research Council Cognitive Function and Ageing Study. Journal of Epidemiology and Community Health, Vol 54, pp286–292. http://dx.doi.org/10.1136/jech.54.4.286

Meng, Y, Davies, R, Hardy, K and Hawkey, P (2010). An application of agent-based simulation to the management of hospital-acquired infection. Journal of Simulation, Vol 4, pp60–67. http://dx.doi.org/10.1057/jos.2009.17

Michaud, PC, Goldman, D, Lakdawalla, D, Gailey, A and Zheng, Y (2009). International Differences in Longevity and Health and Their Economic Consequences. National Bureau of Economic Research. Working Paper Series, Working Paper 15235. Available at : http://www.nber.org/papers/w15235. http://dx.doi.org/10.2139/ssrn.1509175

Moore, DC and Hanratty, B (2013). Out of sight, out of mind? A review of data available on the health of care home residents in longitudinal and nationally representative cross-sectional studies in the UK and Ireland. Age and Ageing, Vol 42, No 6, pp798-803. http://dx.doi.org/10.1093/ageing/aft125

Morrissey, K, Clarke, G, Ballas, D, Hynes, S and O'Donoghue, C (2008). Examining access to GP services in rural Ireland using microsimulation. AREA, Vol 40, No 3, pp354-364. http://dx.doi.org/10.1111/j.1475-4762.2008.00844.x

Morrissey, K, Clarke, G, Williamson, P, Daly, A and O'Donoghue, C (2015). Mental illness in Ireland: simulating its geographical prevalence and the role of access to services. Environment and Planning B: Planning and Design, Vol 42, No 2, pp338 – 353. http://dx.doi.org/10.1068/b130054p

Morrissey, K, Espuny, F and Williamson, p (2014). A Multinomial Model for Comorbidity in England of Longstanding CVD, Diabetes, and Obesity. Available at : http://arxiv.org/ftp/arxiv/papers/1411/1411.2514.pdf

Murphy, M and Di Cesare, M (2012). Use of an age-period-cohort model to reveal the impact of cigarette smoking on trends in Twentieth century adult cohort mortality in England and Wales. Population Studies: A Journal of Demography, Vol 66, No 3, pp259-277. http://dx.doi.org/10.1080/00324728.2012.678881

Murphy, M and Martikainen, P (2010). Demand for Long-term Residential Care and Acute Health Care by Older People in the Context of the Ageing Population of Finland. In Ageing, Care Need and Quality of Life : The Perspective of Care Givers and People in Need of Care, Doblhammer, G and Scholz, R (eds). VS Research. http://dx.doi.org/10.1007/978-3-531-92335-2_8

NATCEN (2012). English Longitudinal Study of Ageing Wave Five Interview Questionnaire – 2010 to 2011. Available at : http://www.ifs.org.uk/elsa/docs_w5/questionnaire_main.pdf

NATCEN (2014). English Longitudinal Study of Ageing (ELSA) Wave 2, 4 and 6 User Guide to the nurse datasets. Available at : http://doc.ukdataservice.ac.uk/doc/5050/mrdoc/pdf/5050_ELSA_Nurse_User_Guide_Waves_2_4_6_v1-0.pdf

National Institute for Health and Care Excellence (2015). Chronic obstructive pulmonary disease: Management of chronic obstructive pulmonary disease in adults in primary and secondary care (partial update). http://www.nice.org.uk/guidance/cg101/chapter/Working-definition-of-COPD

NHS England (2012). Clinical Commissioning groups: proposed configurations agreed. Available at : http://www.england.nhs.uk/2012/05/31/ccg-configs-agreed/

NHS England (2015). The care.data programme – collecting information for the health of the nation. Available at : http://www.england.nhs.uk/ourwork/tsd/care-data/

Noble D, Smith D, Mathur, R, Robson, J and Greenhalgh, T (2012). Feasibility study of geospatial mapping of chronic disease risk to inform public health commissioning. BMJ Open, Vol 2. http://dx.doi.org/10.1136/bmjopen-2011-000711

Norman, P, Rees, P, Wohland, P and Boden, P (2010). Ethnic Group Populations: The Components for Projection, Demographic Rates and Trends. In Ethnicity and Integration Understanding Population Trends and Processes, 2010, Vol 3, pp289-315. http://dx.doi.org/10.1007/978-90-481-9103-1_14

Nuorti JP, Butler JC, Farley MM, Harrison LH, McGeer A, Kolczak MS, Breiman RF. Kolappan C and Gopi P (2000). Cigarette smoking and invasive pneumococcal disease. Active Bacterial Core Surveillance Team. New England Journal of Medicine, Vol 342, No 10, pp681-9. http://dx.doi.org/10.1056/nejm200003093421002

O'Donoghue, C (2001). Dynamic Microsimulation: A Methodological Survey. Brazilian Electronic Journal of Economics, Vol 4, No 2.

OECD (2007), Age-Dependency Ratios. Society at a Glance 2006: OECD Social Indicators, OECD Publishing. http://dx.doi.org/10.1787/soc_glance-2006-4-en

Oeppen, J and Vaupel, JW (2002). Broken Limits to Life Expectancy. Science, Vol 296, pp1029-131. http://dx.doi.org/10.1126/science.1069675

Office for National Statistics (2003). National Statistics 2001 Area Classification. Available at : http://www.ons.gov.uk/ons/guide-method/geography/products/area-classifications/ns-area-classifications/index/index.html

Office for National Statistics (2008). Report: Health expectancies in the UK, 2004. Health Statistics Quarterly 37 Spring 2008 National Statistics 48

Office for National Statistics (2010a). Module 1: Research Methods, Data Collection Methods and Questionnaire Design. Statistical Training Unit. Available at : http://www.ons.gov.uk/ons/about-ons/what-we-do/training/courses/rss-ordinary-student-homepage/teaching-resources/module-1--research-methods--data-collection-methods-and-questionnaire-design.pdf

Office for National Statistics (2010b). Final recommended questions for the 2011 Census in England and Wales Health. Available at : http://www.ons.gov.uk/ons/guide-method/census/2011/the-2011-census/2011-census-questionnaire-content/question-and-content-recommendations/final-recommended-questions-2011---health.pdf

Office for National Statistics (2011). General Household Survey, 1971 — 2009. Social and Vital Statistics Division. Available at : http://www.ons.gov.uk/ons/rel/ghs/general-lifestyle-survey/2009-report/overview-report-2009.pdf

Office for National Statistics (2012a). Annual Mid-year Population Estimates for England and Wales, Mid 2011. Available at : http://www.ons.gov.uk/ons/rel/pop-estimate/population-estimates-for-england-and-wales/mid-2011--2011-census-based-/index.html

Office for National Statistics (2012b). Interim 2011-based sub national population projections for England. Available at : http://www.ons.gov.uk/ons/rel/snpp/sub-national-population-projections/Interim-2011-based/index.html

Office for National Statistics (2012c). Integrated Household Survey User Guide – Volume 1: IHS Background and Methodology 2012. Available at : http://www.esds.ac.uk/doc/6933%5Cmrdoc%5Cpdf%5Cihs_user_guide_volume_1_2012.pdf

Office for National Statistics (2013a). National Population Projections, 2012-based projections. Available at : http://www.ons.gov.uk/ons/rel/npp/national-population-projections/2012-based-projections/index.html

Office for National Statistics (2013b). National Population Projections, 2012-based Extra Variants Report. Available at : http://www.ons.gov.uk/ons/dcp171776_338113.pdf

Office for National Statistics (2014a). Methodology: 2012-based Sub national Population Projection. Available at : http://www.ons.gov.uk/ons/rel/snpp/sub-national-population-projections/2012-based-projections/rpt-snpp-2012-based-methodology-report.html

Office for National Statistics (2014b). Health Expectancies at Birth and at Age 65 in the United Kingdom, 2009–11. Available at : http://www.ons.gov.uk/ons/rel/disability-and-health-measurement/health-expectancies-at-birth-and-age-65-in-the-united-kingdom/2009-11/index.html

Office for National Statistics (2014c). Disability-free Life Expectancy at Birth and at Ages 50 and 65 by Clinical Commissioning Groups, England, 2010-12.

http://www.ons.gov.uk/ons/rel/census/2011-census-analysis/disability-free-life-expectancy-at-birth--at-age-50-and-at-age-65--clinical-commissioning-groups--ccgs--2010-12/rpt.html

Office for National Statistics (2014d). Deaths Registered in England and Wales (Series DR), 2013. Statistical Bulletin. Available at : http://www.ons.gov.uk/ons/dcp171778_381807.pdf and see http://www.ons.gov.uk/ons/rel/vsob1/mortality-statistics--deaths-registered-in-england-and-wales--series-dr-/2013/stb-deaths-registered-in-england-and-wales-in-2013-by-cause.html for the associated data

Office for National Statistics (2014e). National Life Tables, United Kingdom, 2010-2012. Available at : http://www.ons.gov.uk/ons/rel/lifetables/national-life-tables/2010---2012/stb-uk-2010-2012.html

Office for National Statistics (2014f). The ONS Longitudinal Study. Available at : http://www.ons.gov.uk/ons/guide-method/user-guidance/longitudinal-study/index.html

Office for National Statistics (2014g). Microdata. Available at : http://www.ons.gov.uk/ons/guide-method/census/2011/census-data/2011-census-prospectus/release-plans-for-2011-census-statistics/subsequent-releases-of-specialist-products/microdata/index.html

Office for National Statistics (2014h). Quality of Long-Term International Migration Estimates from 2001 to 2011. Available from http://www.ons.gov.uk/ons/guide-method/method-quality/specific/population-and-migration/international-migration-methodology/quality-of-long-term-international-migration-estimates-from-2001-to-2011---full-report.pdf

Office for National Statistics (2015). Open Geography Portal. Available at : https://geoportal.statistics.gov.uk/geoportal/catalog/main/home.page

Oliver, D, Foot, C and Humphries, R (2014). Making our health and care systems fit for an ageing population. Kings Fund. Available at : http://www.lifestorynetwork.org.uk/wp-content/uploads/downloads/2014/03/making-health-care-systems-fit-ageing-population-oliver-foot-humphries-mar14.pdf

Olshansky, SJ, Passaro, DJ, Hershow, RC, Layden, J, Carnes,BA, Brody, J, Hayflick, L, Butler, RN, Allison, DB and Ludwig, DS (2005). A Potential Decline in Life Expectancy in the United States in the 21st Century. New England Journal of Medicine, Vol 35, No 11, pp1138-1144. http://dx.doi.org/10.1056/NEJMsr043743

Parliamentary Office for Science and Technology (2006). Healthy Life Expectancy. No 257. Available at : http://www.parliament.uk/documents/post/postpn257.pdf

Parliamentary Office for Science and Technology (2007). Ethnicity and Health. No 276. Available at : http://www.parliament.uk/documents/post/postpn276.pdf

Payne, G, Laporte, A, Deber, R and Coyte, PC (2007). Counting backward to health care's future: using time-to-death modelling to identify changes in end-of-life morbidity and the impact of aging on health care expenditures. Milbank Quarterly, Vol 85, pp213–257. http://dx.doi.org/10.1111/j.1468-0009.2007.00485.x

Pearson, J, Lay-Yee, R, Davis, P, O'Sullivan, D, von Randow, M, Kerse, N and Pradhan, S (2011). Primary Care in an Aging Society: Building and Testing a

Microsimulation Model for Policy Purposes. Computer Review, Vol 29, No 1, pp21-36. http://dx.doi.org/10.1177/0894439310370087

Pevalin, DJ and Rose, D (2004). Social capital for health Investigating the links between social capital and health using the British Household Panel Survey. Health Development Agency

Philips, I (2015). Development and application of a two stage hybrid spatial microsimulation technique to provide inputs to a model of capacity to walk and cycle. Available at : http://leeds.gisruk.org/abstracts/GISRUK2015_submission_36.pdf

Pollock, AM, Macfarlane, AJ and Godden, S (2012). Dismantling the signposts to public health? NHS data under the Health and Social Care Act 2012. British Medical Journal, Vol 344. Available at : http://www.bmj.com/highwire/filestream/581438/field_highwire_article_pdf/0.pdf. http://dx.doi.org/10.1136/bmj.e2364

Poppi User Guide (2011). Available at : http://www.poppi.org.uk/POPPI%20User%20Guide.pdf

Procter, KL, Clarke, GP, Ransley, JK and Cade, J (2008). Micro-level analysis of childhood obesity, diet, physical activity, residential socio-demographic and social capital variables: where are the obesogenic environments in Leeds? AREA, Vol 40, No 3, pp323-340. http://dx.doi.org/10.1111/j.1475-4762.2008.00822.x

Public Health England (2013a). Key Fact Sheets : Cardiovascular disease. The National Cardiovascular Intelligence Network. Available at : http://www.yhpho.org.uk/default.aspx?RID=185796

Public Health England (2013b).Cardiovascular disease key facts, Diabetes. Fact sheet number 8. The National Cardiovascular Intelligence Network. Available at : www.yhpho.org.uk/resource/view.aspx?RID=188006

Rahman, A, Harding, A, Tanton, R and Liu, S (2010). Methodological Issues in Spatial Microsimulation Modelling for Small Area Estimation. International Journal of Microsimulation, Vol 3, No 2, pp3-22.

Rapsomaniki, E, Timmis, A, George, J, Pujades-Rodriguez, M, Shah, AD, Denaxas, S, White, IR, Caulfield, MJ, Deanfield, JE, Smeeth, L, Williams, B, Hingorani, A, Hemingway, H (2015). Blood pressure and incidence of twelve cardiovascular diseases: lifetime risks, healthy life-years lost, and age-specific associations in 1·25 million people. Lancet, Vol 383, No 9932. http://dx.doi.org/10.1016/S0140-6736(14)60685-1

Rasulo, D, Mayhew, L and Rickayzen, B (2009). The Decomposition of Disease and Disability Life Expectancies in England, 1992-2004. Faculty of Actuarial Science and Insurance, Actuarial Research Paper No. 191. Cass Business School, London. Available at : http://www.cass.city.ac.uk/__data/assets/pdf_file/0020/37181/191ARP.pdf

Raymer, J, Abel, G and Smith, PWF (2007). Combining census and registration data to estimate detailed elderly migration flows in England and Wales. Journal of the Royal Statistical Society, Series A, Vol 170, No 4, pp891–908. http://dx.doi.org/10.1111/j.1467-985X.2007.00490.x

Rechel, B, Doyle, Y, Grundy, E and McKee, M (2009). How can health systems respond to population ageing? Health Systems and Policy Analysis. World Health Organization, on behalf of the European Observatory on Health Systems and Policies.

Available at :
http://www.observatorysummerschool.org/materiale2011/113/policy_brief_10_aging_web.pdf

Rees, P (2011). The Dynamics of Populations Large and Small: Processes, Models and Futures. In J. Stillwell and M. Clarke (eds) (2011) Population Dynamics and Projection Methods. Springer. http://dx.doi.org/10.1007/978-90-481-8930-4_1

Rees, P and Clark, S (2014). The projection of ethnic group populations aged 18 and over for Westminster Parliamentary Constituencies in Great Britain for election years 2015, 2020, 2025, 2030 and 2035. A report to the Policy Exchange, 10 Storey Gate, London, SW1P 3AY

Rees, P, Wohland, P and Norman, P (2013b). The demographic drivers of future ethnic group populations for UK local Areas 2001-2051. The Geographical journal, Vol 179, No 1, pp44-60. http://dx.doi.org/10.1111/j.1475-4959.2012.00471.x

Rees, P, Wohland, P and Norman, P (2013c). Using 2011 census data to evaluate and update ethnic group projections. Presentation at the Census Research User Conference, September 2013, Birkbeck College, London. Available at :
http://ukdataservice.ac.uk/media/333637/rees.pdf

Rees, P, Wohland, P and Norman, P (2015). Learning lessons from Population projections: How well did we forecast the ethnic transition. GISRUK Conference 2015, University of Leeds.

Rees, P, Wohland, P, Norman, P and Boden, p (2011). A local analysis of ethnic group population treds and projections for the UK. Journal of Population Research, Vol 28, No 2-3, pp149-183. http://dx.doi.org/10.1007/s12546-011-9047-4

Rees, P, Zuo, C, Wohland, P, Jagger, C, Norman, P, Boden, P and Jasinska, M (2013a). The Implications of Ageing and Migration for the Future Population, Health, Labour Force and Households of Northern England. Applied Spatial Analysis and Policy, Vol 6, No 2, pp93-122. http://dx.doi.org/10.1007/s12061-013-9086-7

Reilly, JJ, Methven, E, McDowell, ZC, Hacking, B, Alexander, D, Stewart, L, Kelnar, CJH (2003). Health consequences of obesity. Archives of Disease in Childhood, Vol 88, pp748–752. http://dx.doi.org/10.1136/adc.88.9.748

Roland, M, Dusheiko, M, Gravelle, H and Parker, S (2005). Follow up of people aged 65 and over with a history of emergency admissions: analysis of outline admission data. British Medical Journal, Vol 330, pp289-292.
http://dx.doi.org/10.1136/bmj.330.7486.289

Roland, M, Guthrie, B and Thom, DC (2012). Primary Medical Care in the United Kingdom. Journal of the American Board of Family Medicine. Vol 25 Supplement, ppS6-S11. http://dx.doi.org/10.3122/jabfm.2012.02.110200

Rural Evidence Research Centre (2005). Rural England: Demographic Change and Projections 1991 - 2028 a source document. Department of Geography, Environment and Development Studies, Birkbeck College, University of London.

Rutherford, T (2011). Population ageing: statistics. House of Commons Library Standard Note: SN/SG/3228. Available at : www.parliament.uk/briefing-papers/SN03228.pdf

Rutter, CM, Zaslavsky, AM and Feuer, EJ (2010). Dynamic Microsimulation Models for Health Outcomes: A Review. Medical Decision Making, Vol 31, pp10-18. http://dx.doi.org/10.1177/0272989X10369005

Sato, RC and Zouain, DM (2010). Markov Models in health care. Einstein (São Paulo), Vol 8, No 3, pp376-379. Available at : http://www.academia.edu/download/30869160/Markov_Models.pdf

Schellhorn, M, Stuck, AE, Minder, CE and Beck, JC (2000) Health Services Utilization of Elderly Swiss: Evidence from Panel data. Econometrics and Health Economics, Vol 9, pp533–545. http://dx.doi.org/10.1002/1099-1050(200009)9:6<533::AID-HEC532>3.0.CO;2-V

Schofield, DJ, Shrestha, RN and Callander, EJ (2012). Access to general practitioner services amongst underserved Australians: a microsimulation study. Human Resources for Health, Vol 10, No 1, pp1-6. http://dx.doi.org/10.1186/1478-4491-10-1

Scholes, S, Medina, J, Cheshire, H, Cox, K, Hacker, E and Lessof, C (2009). Living in the 21st century: older people in England. The 2006 English Longitudinal Study of Ageing Technical Report, June 2009

Scholes, S, Pickering, K and Rayat, P (2007). Healthy Lifestyle Behaviours: Model Based Estimates for Middle Layer Super Output Areas and Local Authorities in England, 2003-2005: User Guide. The National Centre for Social Research. Available at : http://www.ic.nhs.uk/webfiles/Popgeog/MSOA%20and%20LA%20User%20Guide.pdf

Scuffham, P, Chaplin, S and Legood, R (2003). Incidence and Costs of Unintentional Falls in Older People in the United Kingdom. Journal of Epidemiology and Community Health, Vol 57, No 9, pp740-744. http://dx.doi.org/10.1136/jech.57.9.740

Seshamani, M and Gray, A (2004a). Ageing and health-care expenditure: the red herring argument revisited. Health Economics, Vol 13, pp303–314. http://dx.doi.org/10.1002/hec.826

Seshamani, M and Gray, A (2004b). Time to death and health expenditure: an improved model for the impact of demographic change on health care costs. Age and Ageing, Vol 33, pp556-561. http://dx.doi.org/10.1093/ageing/afh187

Sessions, JG, Yu, G and Wall, M (2011). Social Capital and Health: A Longitudinal Analysis from the British Household Panel Survey. Bath Economics Research Papers No. 6/11. Department of Economics, University of Bath, Bath

Shankar, A, McMunn, A and Steptoe, A (2010). Health-Related Behaviors in Older Adults Relationships with Socio-economic Status. American Journal of Preventative Medicine, Vol 38, No 1, pp39–46. http://dx.doi.org/10.1016/j.amepre.2009.08.026

Shapiro, J (2013). The NHS in 2013: plus ça change…. Clinical Medicine, Vol. 13, No. 4, pp374-377. http://dx.doi.org/10.7861/clinmedicine.13-4-374

Shaw, C (2007) 'Fifty years of United Kingdom national population projections: how accurate have they been?' Population Trends, No. 128, pp. 8–23. Available at : http://www.gad.gov.uk/Documents/Demography/Population%20Trends/Population_Trends_128.pdf

Shi, L, van Meijgaard, J and Fielding, J (2011). Forecasting Diabetes Prevalence in California: A Microsimulation. Preventing Chronic Disease, Vol 8, No 4, pp1-6

Shlomo, N (2007). Assessing the Impact of SDC Methods on Census Frequency Tables. Joint UNECE/Eurostat work session on statistical data confidentiality, Manchester, United Kingdom. Available at : http://www.unece.org/fileadmin/DAM/stats/documents/2007/12/confidentiality/wp.18.e.pdf

Shulman, H, Birkin, M, Clarke, GP (2015) A comparison of small-area hospitalisation rates, estimated morbidity and hospital access., Health & Place, Vol 36, pp134-144. http://dx.doi.org/10.1016/j.healthplace.2015.10.001

Simeonova, E (2013). Marriage, bereavement and mortality: The role of health care utilization. Journal of Health Economics, Vol 32, No 1, pp33–50. http://dx.doi.org/10.1016/j.jhealeco.2012.10.010

Simpson, CR and Sheikh, A (2010).Trends in the epidemiology of asthma in England: a national study of 333,294 patients. Journal of the Royal Society of Medicine, Vol 103, No 3, pp98–106. http://dx.doi.org/10.1258/jrsm.2009.090348

Simpson, L (2004). Integrating estimates and targets within a population projection. ESRC Research Methods Programme. Working Paper No 18

Simpson, L, Jivraj, S, Warren, J (2014). The stability of ethnic group and religion in the Censuses of England and Wales 2001-2011. CCSR/CoDE Working Paper 2014 University of Manchester. Available from : http://www.ethnicity.ac.uk/medialibrary/briefingsupdated/how-have-people's-ethnic-identities-changed-in-england-and-wales.pdf

Singer, JD and Willett, JB (2003). Applied Longitudinal Data Analysis. Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195152968.001.0001

Singh AS, Mulder C, Twisk JWR, van Mechelen W, Chinapaw MJM (2008). Tracking of childhood overweight into adulthood: a systematic review of the literature. Obesity reviews. Vol 9, No 5, pp474 – 488. http://dx.doi.org/ 10.1111/j.1467-789X.2008.00475.x

Smith, D, Pearce, JR and Harland, K (2011b). Can a deterministic spatial microsimulation model provide reliable small-area estimates of health behaviours? An example of smoking prevalence in New Zealand Health and Place, Vol 17, pp618–624.

Smith, MP and White, C (2009). An investigation into the impact of question change on estimates of General Health Status and Healthy Life Expectancy. Health Statistics Quarterly 41. Office for National Statistics. Available at : http://www.ons.gov.uk/ons/rel/hsq/health-statistics-quarterly/no--41--spring-2009/an-investigation-into-the-impact-of-question-change-on-estimates-of-general-health-status-and-healthy-life-expectancy.pdf. http://dx.doi.org/10.1057/hsq.2009.6

Smith, MP, Olatunde, O and White, C (2011a). Disability-free life expectancy: comparison of sources and small area estimates in England, 2006–08. Health Statistics Quarterly 50, Office for National Statistics. Available at : http://www.ons.gov.uk/ons/rel/hsq/health-statistics-quarterly/no--50--summer-2011/disability-free-life-expectancy--comparison-of-sources-and-small-area-estimates-in-england--2006-08.pdf

Social and Spatial Inequalities Research Group (2004). Constructing tracts 1981-2001 Version 2. Geography Department, University of Sheffield. Available at : http://sasi.group.shef.ac.uk/tracts/

Sood, S, Mbarika, V, Jugoo, S, Dookhy, R, Doarn, CR, Prakash, N and Merrell, R C (2007). What is telemedicine? A collection of 104 peer-reviewed perspectives and theoretical underpinnings Telemedicine and E-Health, Vol 13, No 5, pp573–590. http://dx.doi.org/10.1089/tmj.2006.0073

Spielauer, M (2007). Dynamic Microsimulation of Health Care Demand, Health Care Finance and the Economic Impact of Health Behaviours: Survey and Review. International Journal of Microsimulation, Vol 1, No 1, pp35-53

Spillman, BC and Lubitz, J (2000). The Effect of Longevity on Spending for Acute and Long-term Care. New England Journal of Medicine, Vol 342, No 19. http://dx.doi.org/10.1056/nejm200005113421906

Sprague, TB (1880). Explanation of a New Formula for Interpolation, Journal of the Institute of Actuaries, Vol 22, No 4, pp22-270.

Steel, N, Maisey, S and Cox, K (2006). Quality of healthcare. In National Centre for Social Research Retirement, health and relationships of the older population in England. The 2004 English Longitudinal Study of Ageing (Wave 2). Banks, J, Breeze, E, Lessof, C and Nazroo, J (eds). The Institute for Fiscal Studies, 7 Ridgmount Street, London

Stillwell, J and Duke-Williams, O (2007). Understanding the 2001 UK census migration and commuting data: the effect of small cell adjustment and problems of comparison with 1991. Journal of the Royal Statistical Society: Series A, Vol 170, No 2, pp425–445. http://dx.doi.org/10.1111/j.1467-985X.2006.00458.x

Stockdale, A (2011). A Review of Demographic Ageing in the UK: Opportunities for Rural Research. Population, Space and Place, Vol 17, pp204–221. http://dx.doi.org/10.1002/psp.591

Strachan, D (2000). Family size, infection and atopy: the first decade of the 'hygiene hypothesis'. Thorax, Vol 55, Supplement , ppS2–S10. http://dx.doi.org/10.1136/thorax.55.suppl_1.s2

Tanaka, T, Gjonc, E and Gulliford, MC (2011). Income, wealth and risk of diabetes among older adults: cohort study using the English longitudinal study of ageing. The European Journal of Public Health, Vol 22, No 3, pp1-7. http://dx.doi.org/ 10.1093/eurpub/ckr050

The British Thoracic Society (2006). The Burden of Lung Disease, 2nd Edition. Available at : https://www.brit-thoracic.org.uk/document-library/delivery-of-respiratory-care/burden-of-lung-disease/burden-of-lung-disease-2006/

Tomintz, MN, Clarke, GP and Rigby, JE (2008). The geography of smoking in Leeds: estimating the individual smoking rates and the implications for the location of stop smoking services. AREA, Vol 40, No 3, pp341-353. http://dx.doi.org/10.1111/j.1475-4762.2008.00837.x

Townend, P, Xu, J, Birkin, M, Turner, A and Wu, B (2009). MoSeS: Modelling and Simulation for e-Social Science. Philosophical Transactions of the Royal Society A, Vol 367, pp2781–2792. http://dx.doi.org/10.1098/rsta.2009.0041

University of Leeds (2011). ETHPOP Database. http://www.ethpop.org/

Voas, D and Williamson, P (2000). An evaluation of the combinatorial optimisation approach to the creation of synthetic microdata. International Journal of Population Geography, Vol 6, No 5, pp349-366. http://dx.doi.org/10.1002/1099-1220(200009/10)6:5<349::AID-IJPG196>3.0.CO;2-5

Wadsworth, J and Vaitilingam, R (2014). Immigration, the European Union and the UK Labour Market. Centre for Economic Performance Policy Analysis. Available at : http://cep.lse.ac.uk/pubs/download/pa015.pdf

Walker, A (2007). Health Status, Health Inequalities and the ability of Older Australians to Stay in the Labour Force. In Modelling Our Future : Population Ageing, Health and Aged Care. Gupta, A and Harding, A (eds) (2007). International Symposia in Economic Theory and Econometrics. Elsevier

Wanless, D (2004). Securing our Future Health: Taking a Long-Term View. Final Report. HM Treasury, London. Available at : http://webarchive.nationalarchives.gov.uk/+/http://www.hm-treasury.gov.uk/consult_wanless_final.htm

Webber (2004). Neighbourhood Inequalities in the Patterns of Hospital Admissions and their Application to the Targeting of Health Promotion Campaigns. Working Paper 90, CASA Working Paper Series. Available at: http://www.bartlett.ucl.ac.uk/casa/pdf/paper90.pdf

Werblow, A, Felder, S and Zweifel, P (2007). Population Ageing and Health Care Expenditure: a School of 'Red Herrings'? Health Economics, Vol 16, pp1109–1126. http://dx.doi.org/10.1002/hec.1213

Westerhout, E (2014). Population Ageing and Health care Expenditure Growth. International Handbook on Ageing and Public Policy, Handbooks of Research on Public Policy series, pp178-190. Edited by Harper, S and Hamblin, K, Oxford Institute of Population Ageing, University of Oxford, UK.

White, I (2009). The 2011 Census taking shape: methodological and technological developments. Population Trends, 136. Available at : http://www.ons.gov.uk/ons/rel/population-trends-rd/population-trends/no--136--summer-2009/the-2011-census-taking-shape--methodological.pdf. http://dx.doi.org/10.1057/pt.2009.19

Wilmoth, JR L. J. Deegan, LJ, Lundstro, H S and Horiuchi, S (2000).Increase of Maximum Life-Span in Sweden, 1861-1999. Science, Vol 289, pp2366-2368. http://dx.doi.org/10.1126/science.289.5488.2366

Wittenberg, R, Hu,B, Comas-Herrera, A and Fernandez, JL (2013). Care for older people. Projected expenditure to 2022 on social care and continuing health care for England's older population. Available at : http://www.nuffieldtrust.org.uk/publications/care-for-older-people-projected-expenditure-2022

Wood, N (2011). ELSA Life History Interview. Available at : http://www.ifs.org.uk/elsa/user_day2011/wood.pdf

Word Bank (2012). Health expenditure, total (% of GDP), indicator SH.XPD.TOTL.ZS. Available at : http://data.worldbank.org/indicator/SH.XPD.TOTL.ZS

World Health Organisation (1946). Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, New York, 19-22 June 1946, and entered into force on 7 April 1948.

Zaidi, A and Rake, K (2001). Dynamic Microsimulation Models: A Review and Some Lessons for Simulating Social Policy in an Ageing Society. SAGE Discussion Paper no. 2. SAGEDP/02. London School of Economics and King's College, London

Zaloznik, M (2011). Iterative Proportional Fitting Theoretical Synthesis and Practical Limitations. PhD Thesis. Available at : http://www.researchgate.net/profile/Maja_Zaloznik/publication/262258986_Iterative_Proportional_Fitting_-_Theoretical_Synthesis_and_Practical_Limitations/links/0f317537217407187200000.pdf

Ziegler, U and Doblhammer, G (2010). Projections of the Number of People with Dementia in Germany 2002 Through 2047. In Ageing, Care Need and Quality of Life : The Perspective of Care Givers and People in Need of Care. Doblhammer, G and Scholz, R (eds). VS Research. http://dx.doi.org/10.1007/978-3-531-92335-2_6

Zucchelli, E, Jones, AM and Rice, N (2010). The evaluation of health policies through microsimulation methods. HEDG Working Paper 10/03, January 2010. Available at : http://york.ac.uk/res/herc/hedgwp

Zweifel, P, Felder, S and Meiers, M (1999). Ageing of Population and Health Care Expenditure: A Red Herring? Health Economics, Vol 8, pp485–496. http://dx.doi.org/10.1002/(SICI)1099-1050(199909)8:6<485::AID-HEC461>3.0.CO;2-4