How can a cause-of-death reduction be compensated for in the presence of heterogeneity? A population dynamics approach based on English data by deprivation

Sarah Kaakai, Héloïse Labit-Hardy, Séverine Arnold (-Gaille), Nicole El Karoui

How can a cause-of-death reduction be compensated for in the presence of heterogeneity? A population dynamics approach based on English data by deprivation

Sarah Kaakai[1], Héloïse Labit Hardy[2], Séverine Arnold (-Gaille)[3],
Nicole El Karoui[4]

April 10, 2018

Abstract

A growing number of studies indicate a widening of socioeconomic inequalities in mortality over the past decades. It has therefore become crucially important to understand the impact of heterogeneity and its evolution on the future mortality of heterogeneous populations. In particular, recent developments in multi-population mortality have raised a number of questions, among which are the issue of consistency between subnational and national forecasts, and the evaluation of cause-of-death reduction targets set by national and international institutions.

The aim of this paper is to show how the study of the population data and the population dynamics framework contribute to addressing these issues, by providing a new viewpoint on the evolution of aggregate mortality indicators in the presence of heterogeneity. Our findings rely on two unique datasets on the English population and cause-specific number of deaths by socioeconomic circumstances, over the period 1981-2015.

The analysis of the data first highlights the complexity of recent demographic developments, characterized by significant composition changes in the population, with considerable variations according to the age-class or cohort, along with a widening of socioeconomic inequalities. In a second part, we introduce a dynamic framework for studying the impact of composition changes on the mortality of the global population. In particular, we are interested in quantifying the impacts of cause-of-death mortality reduction in comparison with changes of composition in a heterogeneous population. We show how a cause of death reduction could be compensated for in the presence of heterogeneity, which could lead to misinterpretations when assessing public policies impacts and/or for the forecasting of future trends.

JEL: G22, J11, C60

Keywords: Population Dynamics, Deprivation, Heterogeneity, Cause-of-Death Mortality, Cohort Effect

[1]Corresponding author, Centre de Mathématiques Appliquées (CMAP), Ecole Polytechnique, Route de Saclay, 91128 Palaiseau, France. Email: sarah.kaakai@polytechnique.edu.

[2]ARC Centre of Excellence in Population Ageing Research (CEPAR), UNSW Australia, Sydney NSW 2052, Australia. Email: h.labithardy@unsw.edu.au.

[3]Département de Sciences Actuarielles (DSA), Faculté des HEC, Extranef, Université de Lausanne, 1015 Lausanne, Suisse. Email: severine.arnold@unil.ch.

[4]Laboratoire de Probabilités, Statistique et Modélisation (LPSM), Université Pierre et Marie Curie, 4 Place Jussieu, 75005 Paris, France. Email: nicole.el_karoui@upmc.fr.

1 Introduction

Large populations such as national populations usually present some heterogeneity, in the sense that individuals with different characteristics (gender, social characteristics, neighborhood, etc.) exhibit different demographic behaviors. Whenever possible, taking into account these characteristics can provide useful information, but at the same time, modeling the population in the presence of heterogeneity is much more complex.

Research on the relationship between socioeconomic status and mortality can be traced back as far as the nineteenth century (see e.g. Villermé (1830), or reports of the General Registrar Office in England). Since then, an important body of work has investigated the links between socioeconomic status (SES) or neigborhood, and mortality and causes of death(Pamuk (1985), Marmot et al. (1991), Mackenbach et al. (1997)). More recently, a growing number of studies have shown that the socioeconomic gradient in mortality appears to be increasing in a number of countries including England (Elo (2009), National Research Council (2011), Office for National Statistics (2015), El Karoui et al. (2018)), with different socioeconomic subgroups experiencing mortality rates rather different from those estimated from national data at the aggregated level. These gaps have even been forecasted to increase further (Villegas and Haberman (2014)).

The widening of these socioeconomic gaps has led a number of pension funds and insurance companies to rethink their models in order to tackle this heterogeneity issue and to understand the potential impact of socioeconomic inequalities. For instance, the Life & Longevity Markets Association (LLMA) and the Institute and Faculty of Actuaries (IFoA) have recently commissioned a series of reports on longevity basis risk (Haberman et al. (2014)). Indeed, the non-consideration of socioeconomic differences can have substantial impacts for insurance companies or governments, by leading, for instance, to errors in funding of annuity and pension obligations (see e.g. Meyricke and Sherris (2013); Villegas and Haberman (2014)) or increasing the basis risk (variation between sample and population mortality).

Consequently, a growing literature, facilitated by the recent release of data at a finer level, has recently taken an interest in the joint modeling and forecasting of the mortality of socioeconomic subgroups (Jarner and Kryger (2011); Villegas and Haberman (2014); Li et al. (2015); Cairns et al. (2016)). However, there are still many open questions regarding the impact of heterogeneity on mortality modeling. For instance, the issue of consistency of sub-national and national estimates and forecasts has been raised recently by Shang and Hyndman (2017) and also Shang and Haberman (2017), who proposed an approach based on recent developments in grouped functional time series methods. In order to evaluate the consistency of mortality forecasts made directly at the national and then at the subnational level, one needs to aggregate the mortality forecasts of the different subgroups composing the population, by taking into account the composition of each age class. Thus, in order to address this issue, it is necessary to understand future mortality developments in each subgroup, but also the population age-structure and composition, as well as its evolution over time: namely, the population dynamics.

The population heterogeneity also raises issues concerning the evaluation of public health

policies. Indeed, a number of institutions have defined public health goals in terms of cause-of-death mortality reduction (Department of Health (2003), World Health Organization (2013)). However, at the national level, the interpretation of standard indicators such as the period life expectancy can be complex in the presence of heterogeneity, since individuals with different socioeconomic status are affected differently by diseases (Bajekal et al. (2013b), National Research Council (2011), Villegas (2015)). For instance, by studying recommendations from the World Health Organization, Alai et al. (2017) have shown that these recommendations could actually increase life expectancy gaps in England, despite an increase of the national life expectancy. More generally, it is difficult to evaluate if changes in cause-of-death mortality have occurred by only analyzing data aggregated at the national level, since the population age-structure and socioeconomic composition change at the same time.

The aim of this paper is to show how the study of the population data and the population dynamics framework shed new light on the evolution of aggregate mortality patterns and longevity indicators in the presence of heterogeneity. The classical approach developed in actuarial science for the modeling and forecasting of mortality rates focuses on mortality data only. However, as the preceding examples illustrate, the information contained in the population age-structure is crucial in order to understand the effects of heterogeneity on aggregate mortality trends.

Our first goal is not to provide a new multi-population or cause-of-death mortality model, but rather to use the population dynamics point of view in order to represent the data differently than what is usually done. The study is based on two unique datasets obtained from the UK Office for National Statistics (ONS) and the Department of Applied Health Research (DAHR), University College London[1], containing data on the English population and cause-specific number of deaths by age, gender and socioeconomic circumstance over the period 1981-2015. In particular, data for the period 2001-2015 have only been recently released by the ONS. Our analysis reveals significant variations of the population's socioeconomic composition, across the different age classes and over time. This heterogeneity is combined with an increase in mortality differences, confirming observations of Lu et al. (2014) and Villegas and Haberman (2014) over the period 1981-2007, which magnifies the impact of population composition changes on aggregate mortality.

Our second goal is to provide a *dynamic* framework in order to study these composition changes and their effect on aggregate mortality over time. In particular, we are interested in studying cause-of-death mortality reduction in the presence of heterogeneity. In the demographic literature, Shkolnikov et al. (2006) and Jasilionis et al. (2011) have studied the contribution of compositional changes to mortality evolution, but rather with a static approach based on decomposition methods. By introducing an heterogeneous McKendrick-Von Foerster population dynamics model, we propose a dynamic approach in order to test the impact of cause-of-death mortality reduction under different demographic scenarios. This framework allows us to isolate demographic changes of different natures (cause-of-death reduction, reverse cohort effect) and to quantify how they inter-

[1]The authors thank Madhavi Bejekal, Senior Research Fellow at the Department of Applied Health Research (DAHR), University College London, for her assistance in obtaining the dataset from the DAHR.

act when they are combined. In particular, we show that the reduction of a cause of death may not necessarily result in an improvement in aggregate mortality rates or life expectancy, if the composition of the population changes at the same time. Thus, the effect of public health policies could be misinterpreted if only aggregated data are studied. The remainder of this paper is organized as follows.
In Section 2, we introduce the data used to carry out our study. Particular emphasis is being placed on presenting the main features of the age structures of the subpopulations grouped by socioeconmic circumstances, and their evolution over time. In Section 3, we present the deterministic population dynamics model used in Section 4. Section 4 presents our numerical results. We first show how different socioeconomic composition of the age classes can impact the life expectancy and mortality improvement rates. Then, we show how a cause of death reduction could be compensated for by a reverse cohort effect, induced by heterogeneity in fertility rates.

2 What can be learned from the data

In this section, we present the two datasets used in this paper. Particular emphasis is made on the evolution of the age structure of the subpopulations grouped by level deprivation, which we consider to be an important contribution of the study.

2.1 Datasets

The data we use provide mid-year population estimates in England by age class and socioeconomic circumstances for the years 1981-2015, combined with the number of deaths by age, cause and socioeconomic circumstance. Our study is based on two data sources:
(i) The first dataset was provided to us by the Department of Applied Health Research (DAHR) of the University College London in UK, and is based on the Index of Multiple Deprivation 2007 (IMD 2007) for the 1981-2006 period.
(ii) The second dataset was released in 2017 by the Office for National Statistics in UK (ONS)[2], and is based on the Index of Multiple Deprivation 2015 (IMD 2015) for the 2001-2015 period.

Deprivation Criterion In both datasets, socioeconomic circumstances are measured by the Index of Multiple Deprivation (IMD). The IMD is a geographically based index, created in order to provide an official measure of multiple deprivation dimensions at the level of small areas called LSOAs[3], each composed of about 1500 individuals (see Noble et al. (2007) and Department for Communities and Local Government (2015) for more details). The concept of deprivation has been defined by Townsend as the lack of *"types of diet, clothing, housing, household facilities and fuel and environmental, educational, working and social conditions, activities and facilities which are customary"* (Townsend

[2]Publicly available on the ONS website (www.ons.gov.uk) under the reference number 006925.

[3]In 2007, there were 32,482 Lower Layer Super Outupt areas (LSOAs) in England (34,753 in 2011) (Office for National Statistics (2012)).

(1979), cited in Noble et al. (2007)). More precisely, the IMD is based on the measure of seven broad socioeconomic factors: income, employment, health, education, barriers to housing and services, living environment and crime. It is computed for each small living area in England. The IMD score of a small area is used as a SES proxy for individuals living in the LSOA. Furthemore, the index also includes information on the physical and social environment of individuals (by including for instance the road distance to a GP surgery and supermarket, or crime statistsics), which can have a significant influence on health outcomes (Diez Roux and Mair (2010); Nandi and Kawachi (2011)).
One limitation of using area based measurements is to apply the same deprivation to all individuals living in the same area. However, LSOAs are rather small areas and geographical data are often more available on a large scale than multiple individual socioeconomic measurements. We also note that the IMD is computed at fixed dates, while being applied to a longer time period (e.g. IMD 2007 applied to the period 1981-2007). The implications of using a fixed IMD quintile allocation have been discussed comprehensively in Bajekal et al. (2013a) and in Appendix D of Lu et al. (2014), based on the period 1981-2001. Our comparison of data computed with the IMD 2007 and IMD 2015 for the overlapping period 2001-2006 also gives very similar results. See Appendix A for a more detailed discussion on this issue.

Structure of data Our two data sources are based on a *relative* measurement of deprivation. Small living areas are ranked by their IMD scores and grouped into five deprivation quintiles numbered from 1 to 5: IMD quintile 1 (IMD Q1) for the least deprived quintile, to IMD quintile 5 (IMD Q5) for the most deprived quintile. It is worth noting that for each year, the five deprivation quintiles have approximately the same number of individuals. Population data are structured by age-class and deprivation quintile, and deaths data are structured by age-class, deprivation quintile and cause of deaths. Specific features of each datasets are summarized in Table 1 and 2[4].

	Dataset 1	Dataset 2
Deprivation index	IMD 2007	IMD 2015
Time period	1981-2006	2001-2015
Ages	25-85+	0-90+
Population age group	5 years	1 year
Deaths age group	5 years	1 year 5 years per cause

Table 1: Datasets

	Datasets	
	1	2
Cardiovascular diseases	×	×
Neoplasms	×	×
Respiratory diseases	×	×
External causes	×	×
Diabetes		×
Digestives diseases	×	
Mental diseases	×	
Neonatal deaths		×

Table 2: Causes of death in each dataset

The compilation method of Dataset 1 is described in Lu et al. (2014). See also Labit Hardy

[4]In Dataset 2, cardiovascular diseases are divided in ischemic heart diseases, strokes and other cardiovascular diseases; external causes in intentional and unintentional injuries.

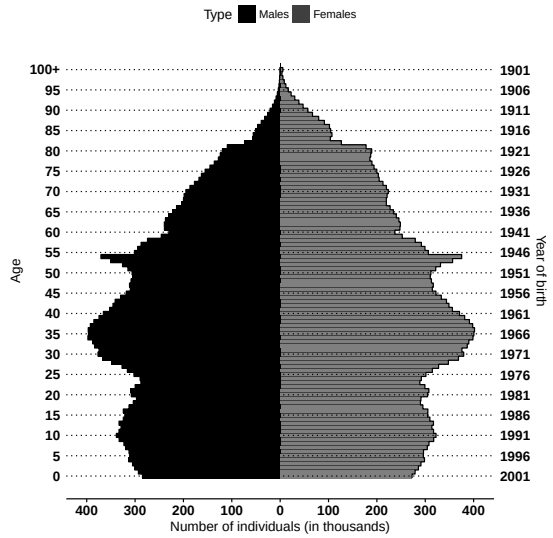
(2016) for a detailed description of the dataset. Those data have also been used in papers focusing on the study of mortality improvements and healthy life expectancy by deprivation (e.g. Bajekal (2005), Lu et al. (2014)) or in mortality modeling (e.g. Villegas and Haberman (2014)).
Dataset 2 has been only very recently released by ONS. In addition to updating the data for years 2007-2015, this second dataset provides disaggregated data by single year of age, as well as data for young ages below age 25. This constitutes an important contribution to our paper, by allowing a more precise analysis of the population by deprivation.

2.2 Data analysis

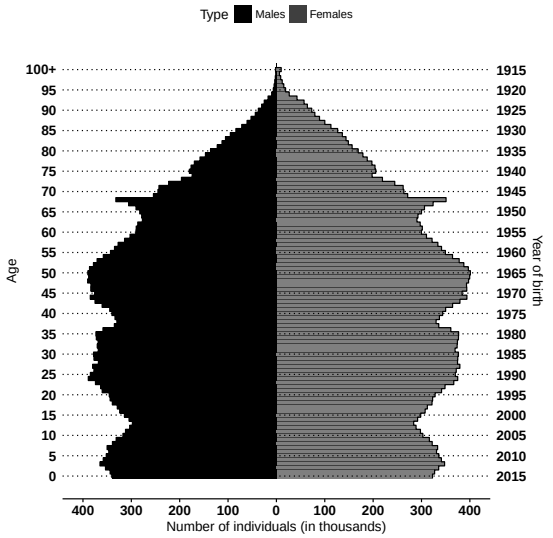
2.2.1 Population composition

Age Pyramids In order to illustrate differences in the composition of populations per deprivation quintile, age pyramids of the least deprived quintile (IMD Q1) and most deprived quintile (IMD Q5) are represented on Figure 1 for years 2001 and 2015, along with the age pyramid of England.
By reading Figure 1 vertically, we can see that for each year, the form of the age pyramids are very different between the different deprivation quintiles and the English population. IMD Q5 represented on Figure 1e and 1f, is much younger on average than IMD Q1, which is represented on Figure 1c and 1d. For instance, the median age in 2015 was 33 years (35.5 for the mean age) in IMD Q5, while the median in IMD Q1 was 44.2 (42.6) and 39 years (39.7) in England. One factor explaning these differences could be the natural life course-trajectory of individuals, with young adults (around 20-35) typically living in rented housing in inner-city areas and with older household (older than 35) moving out to less deprived neighborhoods.
Some cohorts are also more represented among a particular subpopulation. For instance, the English baby-boom generation (born in the years after World War II) and their children are strikingly more represented among IMD Q1.
The horizontal reading of Figure 1 shows that in addition to this heterogeneity in age, significant temporal changes in the age pyramids occurred from 2001 to 2015. These changes are caused by population ageing, but probably also by changes in birth patterns, coupled with internal and external migrations[5]. Furthermore, changes over time in populations age-structure are quite different according to the level of deprivation. For instance, the median age in IMD Q5 has dropped over 1%, from 33.4 to 33 years, while it has increased more than 9% in IMD Q1, from 40 to 44, and about 5% in the general population, from 37.1 to 39. Thus, IMD Q5 has become more youthful from 2001 the 2015, despite a general population ageing.

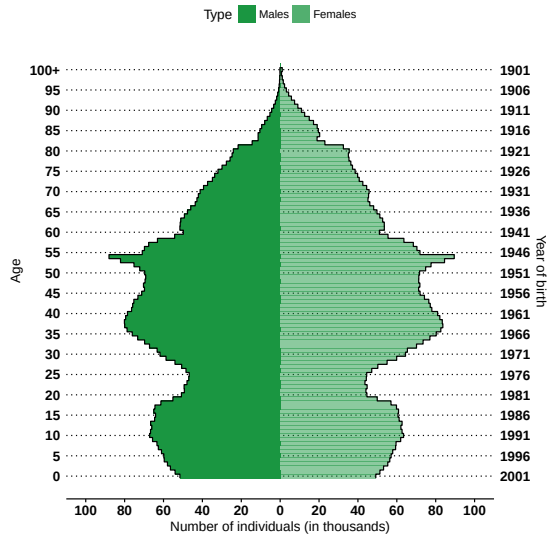
[5]Here, we refer to internal migrations as the migration of individuals in between IMD quintiles, whereas external migrations correspond to the migration of individuals from/to places outside of England.



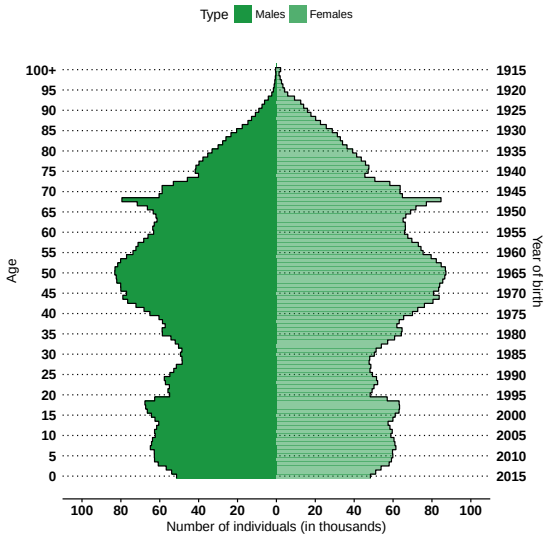
(a) English population, 2001



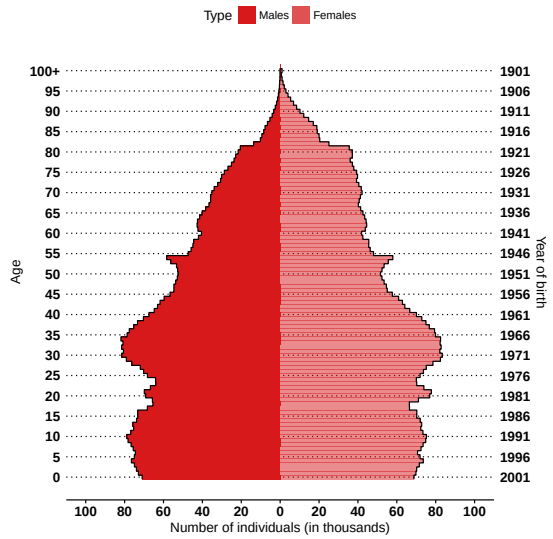
(b) English population, 2015



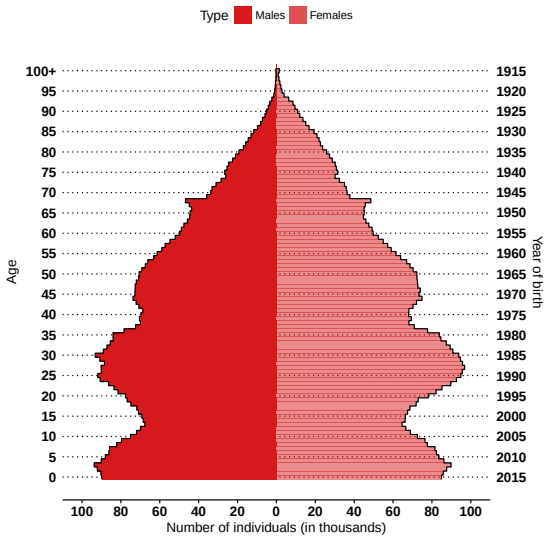
(c) Least deprived quintile (IMD Q1), 2001



(d) Least deprived quintile (IMD Q1), 2015



(e) Most deprived quintile (IMD Q5), 2001



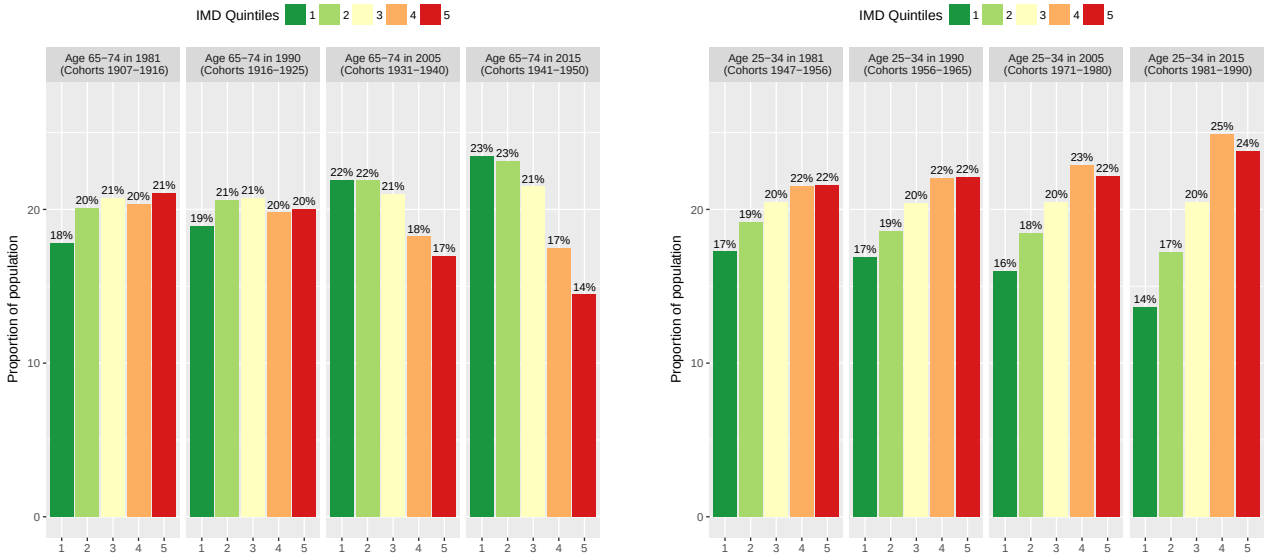
(f) Most deprived quintile (IMD Q5), 2015

Figure 1: Age pyramids in 2001 and 2015

There is an inherent difficulty in representing the time evolution of data structured in age. In order to better understand the evolution of the population composition, we also represent the evolution of the composition of specific *age classes* (the same age class is represented at different dates), and of specific *cohorts* (the same cohort is represented at different dates), over the period 1981-2015. In the following, we display data for males. When results are different for women, plots are also displayed.

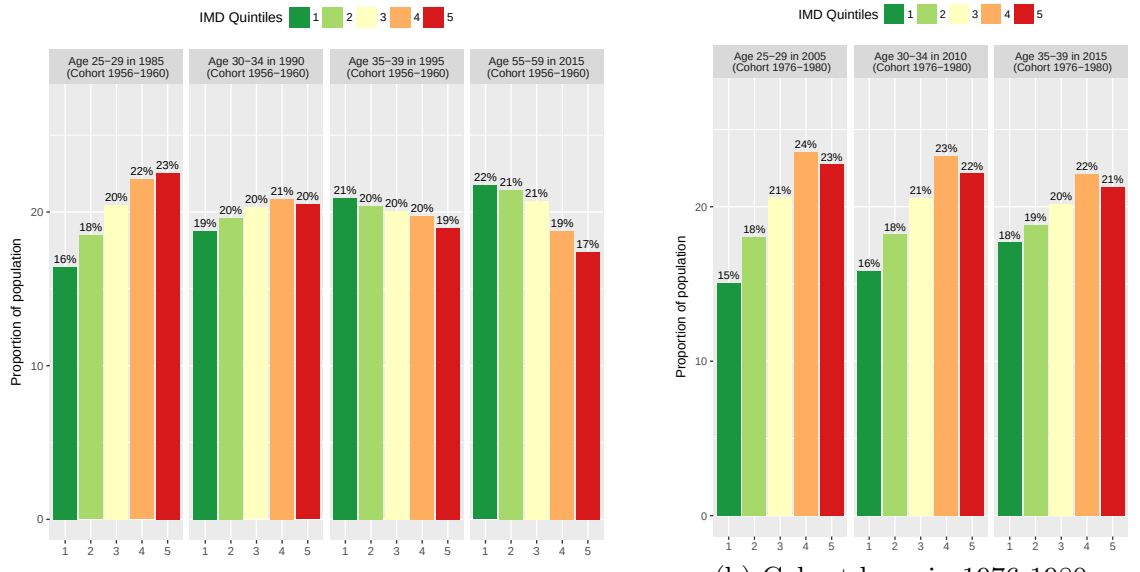
Fixed age classes Figure 2 represents the evolution of the distribution of males in each IMD quintile, for the years 1981, 1990, 2005 and 2015 and for two fixed ages classes: 65-74 (Figure 2a) and 25-34 years (Figure 2b).
The composition of the age class 65-74 significantly varied from 1981 to 2015, to the benefit of the least deprived populations. Thus, the proportion of males in the two least deprived quintiles (IMD Q1 and Q2) increased from 38% in 1981 to 46% percent in 2015; on the contrary, the proportion of males in the two most deprived quintiles (IMD Q4 and Q5) decreased from 41% to 32%. This could be explained by an improvement over time of living conditions for older individuals, but also, as noted above, by a baby-boom cohort effect. Indeed, individuals born during the English baby-boom are less deprived than the immediately preceding and following cohorts, regardless of the global improving trend.
As already observed in Figure 1, the level of deprivation is more important in younger age classes, this being true for the whole period 1981-2015. However, as shown in Figure 2b, the composition of the age class 25-34 also varied from 1981 to 2015, and the relative deprivation of this age class increased over time. For instance, the proportion of males in the age class 25-34 for IMD Q1 and Q2 has decreased from 36% to 31%, while the proportion of males in IMD Q4 and Q5 has increased from 43% to 49%.

Fixed cohorts Data can also be represented in the cohort dimension. Figures 3a and 3b represent the evolution of the proportion of males in each IMD quintile for the cohorts with ages 25-29 in 1985 (born in 1956-1960) and in 2005 (born in 1976-1980). The 1976-1980 cohort could only be represented up to ages 35-39.
The average level of deprivation of both cohorts decreased over time, due to internal and external migrations (internal migrations correspond to residential mobility) and of different level of mortality rates. However, the improvement of the deprivation for the older cohort, Figure 3a, is much more important than that of the younger cohort, Figure 3b, which confirms observations of Figure 2b. Thus, the proportion of males in the 1956-1960 cohort at age 35-39, in year 1995, for the two most deprived quintiles was 39%, against 43% in the 1976-1980 cohort at the same age (in 2015). Similarly, the proportion in the two least deprived quintiles was about 41% in the oldest cohort (1956-1960) against 37% in the youngest cohort (1976-1980).
At old ages, a classical selection effect can be observed, as cohorts become less and less deprived due to differences in the mortality rates, see also Figure 18 in Appendix B.1.



(a) Age class 65-74 (b) Age class 25-34

Figure 2: Proportion of males by age class and IMD quintile



(a) Cohort born in 1956-1960 (b) Cohort born in 1976-1980

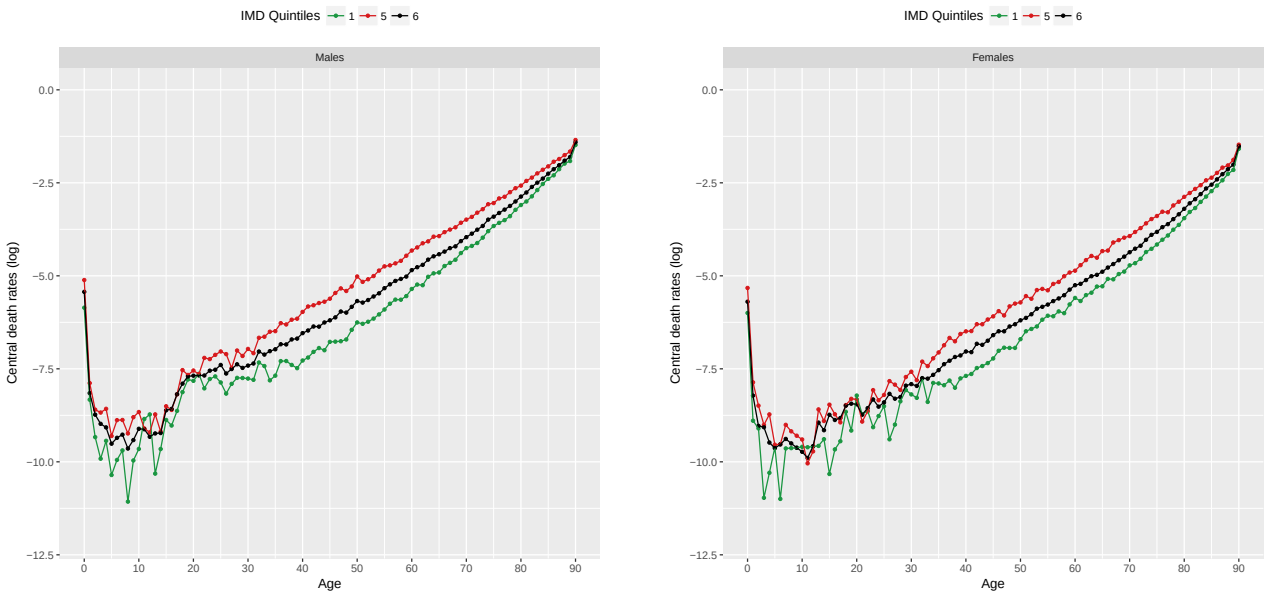
Figure 3: Proportion of males by cohort and IMD quintile

Thus, the population data show an important heterogeneity in the composition of the different age classes, combined with significant temporal changes in age classes composition, such as the striking evolution of the composition of the 65-74 age class between 1981 and 2015. This varying heterogeneity impacts mortality rates differently according to age or time, and generates additional complexity in the study of aggregated death rates. In particular, one might wonder how the increase in deprivation observed among younger cohorts will impact future mortality in England.

2.2.2 Mortality

Let us now give some insights on the mortality per deprivation in England, over the period 1981-2015. As mortality data are more commonly studied, we only give here a brief overview of the main stylized facts, with a particular focus on Dataset 2. For more details on the mortality data of Dataset 1, we refer to Bajekal (2005), Lu et al. (2014) or Villegas (2015).

Mortality rates Central death rates can be computed by age, gender, deprivation quintile and cause of death. They are estimated from our data by taking the number of deaths over the mid-year population, which is assumed to be an accurate estimate of the exposure to risk. Central death rates by single year of age in 2015 are represented on a log-scale in Figure 4 for the English population (noted 6) and the least and the most deprived IMD quintiles (Q1 and Q5). Both levels and shapes of central death rates vary with the level of deprivation, with a mortality higher at all ages for IMD Q5.
Before age 35-40 (with the exception of age 0), differences are less pronounced and central death rates in all IMD quintiles are lower than 1‰. When infant mortality is not taken into account, central death rates first attain the level of 10‰ at age 58 for males in IMD Q5, while this value is only attained at age 68 for males in IMD Q1. Similar differences can be observed for females. In IMD Q5, central death rates attain 10‰ at age 62, in comparison with 72 in IMD Q1.



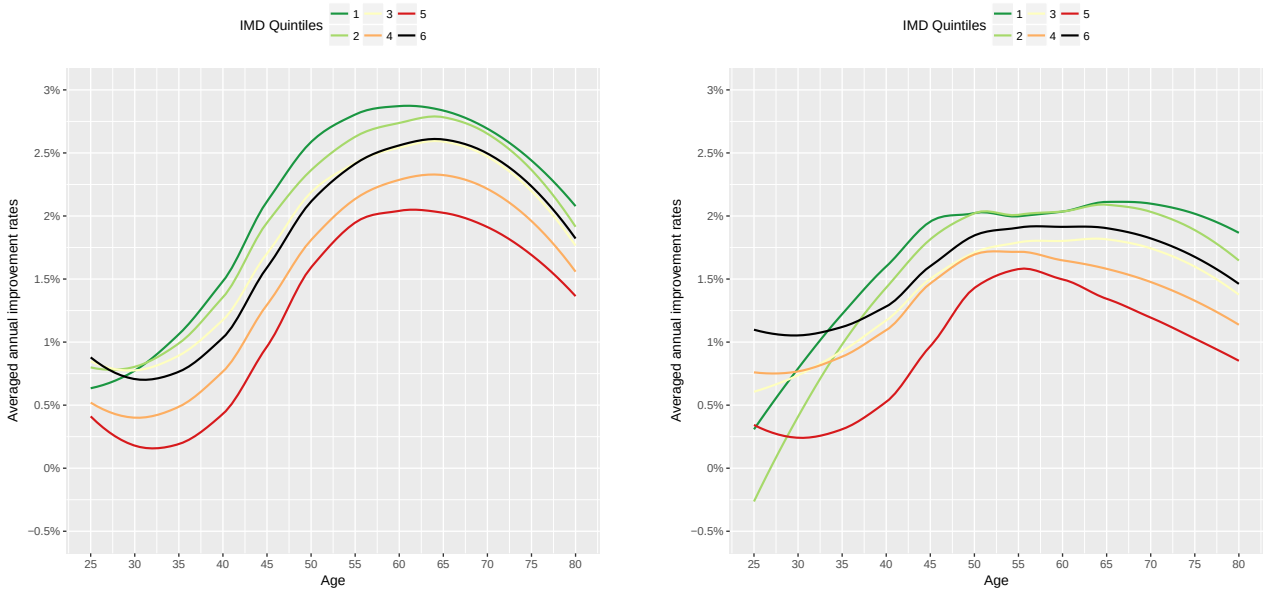
(a) Males (b) Females

Figure 4: Central death rates per single year of age and IMD quintile in 2015

Improvement rates Average annual rates of improvement in mortality over the 1981-2015 period are represented in Figures 5a and 5b[6], for age classes 25 and older. In these graphs, improvement rates are smoothed for visualization purpose only. Figure 5 shows that at all ages, IMD Q5 have experienced lower rates of improvement in mortality than IMD Q1 (with the exception of ages 25-34 for females). Males experienced overall higher improvement in mortality than females, with the highest differentials in deprivation being at ages 40-44 and 45-50. Females experienced the highest differentials in the rate of mortality improvement at ages 35-39 and 40-44.

Figures 6a and 6b represent the average annual mortality improvement rates for males over two distinct periods: 1981-1995 and 2001-2015. The two figures illustrate a clear widening of the gap in annual improvement in mortality at older ages, which is consistent with the observations of Lu et al. (2014) and Villegas and Haberman (2014) over the period 1981-2007.

Over the period 1981-1995, males in the two most deprived quintiles (IMD Q4 and Q5) actually experienced a deterioration of mortality at ages under 40 (negative average annual rate of improvement in mortality), while improvements in mortality over the period 2001-2015 are positive (and higher for IMD Q5 under age 30). Over this period, at ages above 60, the gap in mortality improvement rates increased significantly, with the highest differentials being at ages 75-79 and 80-84 (in comparison with 30-34 and 35-39 over the period 1981-1995). Improvement rates in mortality at younger ages have also changed significantly.

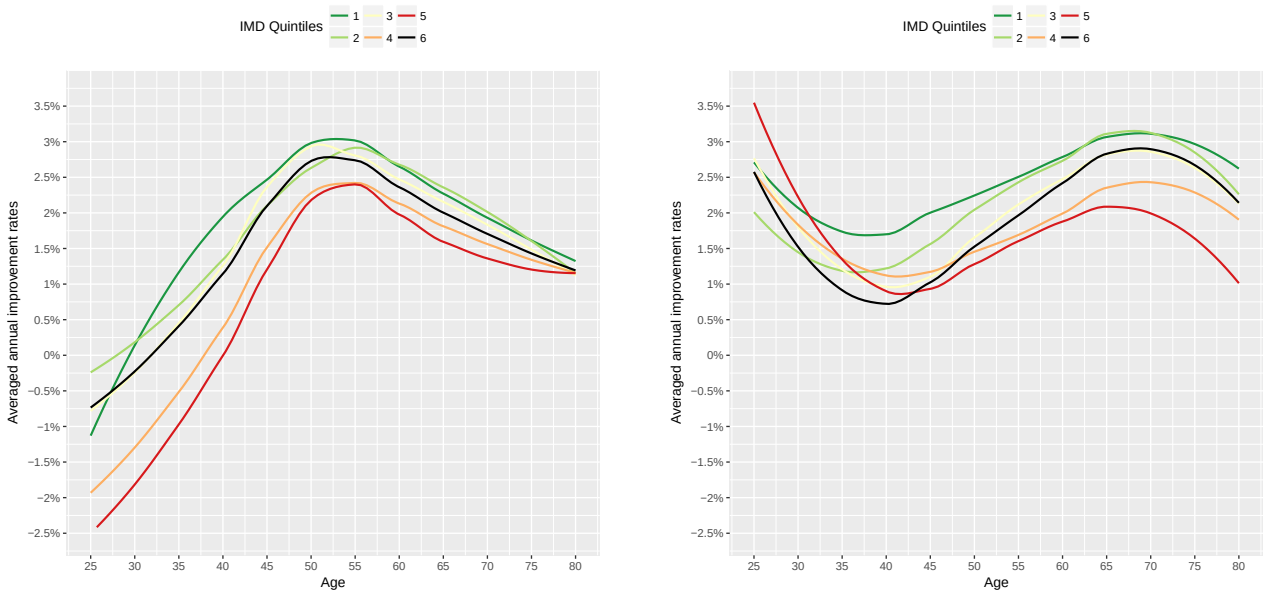


(a) Males

(b) Females

Figure 5: Average annual rates of improvement in mortality, 1981-2015

[6]Improvement rates are computed as the yearly improvement rates of central death rates over five-year age classes.



(a) 1981-1995

(b) 2001-2015

Figure 6: Average annual rates of improvement in mortality, males

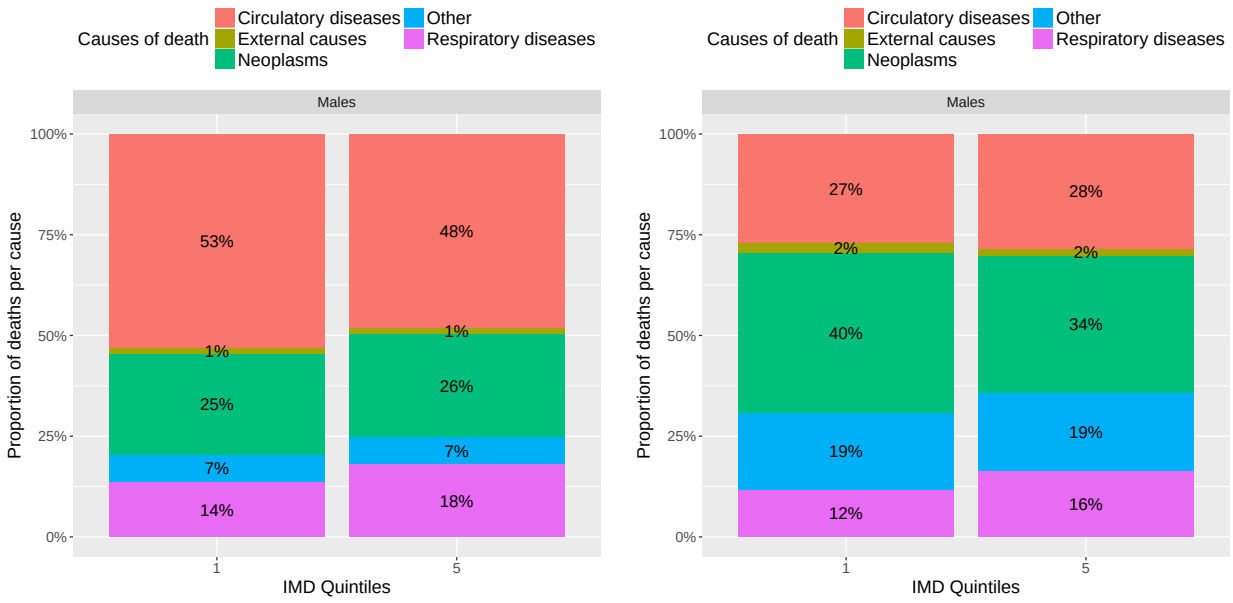
Plots for females are presented in Appendix B.2, Figure 19. It is worth noting that over the period 2001-2015 and for females of age above 65, the gap in mortality improvement rates between IMD Q1 and Q5 has been higher than for males. This can be explained by a sharp deceleration of mortality improvement for females in IMD Q5, already reported by Villegas and Haberman (2014) for the period 1981-2007.
A discussion on the potential drivers of these widening socioeconomic gaps can be found in El Karoui et al. (2018) or Lu et al. (2014).

Causes of death Figures 7a and 7b represent the proportion of deaths by cause of death for males at ages 65-85, in IMD Q1 and Q5 for years 1981 and 2015. Plots for females are available in Appendix B.3.
Over the whole period 1981-2015, cardiovascular diseases (CVD) constituted the leading cause of death in England for ages above 25, followed by cancers (neoplasms) and respiratory diseases. However, recent changes in cause of death trends have been observed since the early 2000s, with neoplasms becoming the leading cause of death, ahead of CVD. The speed of this evolution depended on the deprivation degree, gender and age class. For example, for males of age 25-85 in IMD Q1, neoplasms became the leading cause of death in 2005, while it only became the leading cause of death in 2010 for males in the same age class in IMD Q5.
Differences in cause-of-death mortality per deprivation can be observed for neoplasms, CVD and respiratory diseases. In 1981, differences between IMD Q1 and Q5 were mainly on CVD (53% of all deaths for IMD Q1 and 48% for Q5), while differences were mostly on neoplasms in 2015 (40% of all deaths for IMD Q1 and 34% for Q5). During the whole period, differences in the proportion of death from respiratory diseases remained rather stable.

It is interesting to note as well that at young ages, the most deprived quintiles are more affected by neonatal deaths and accidents (see e.g. Oakley et al. (2009); Guildea et al. (2001) for more details on mortality at younger ages). We aslo refer to Villegas (2015) for more details on trends in cause-of-death mortality per deprivation over the period 1981-2007.

Period life expectancy The period life expectancy at age 65 per IMD quintile is represented over the period 1981-2015 in Figures 8a and 8b, for males and females. Despite a common improvement for all IMD quintiles, the gap in life expectancy between IMD quintiles appears to have widened over time. For instance, the gap in life expectancy between IMD Q1 and Q5 has grown from 2.2 years for females and 2.9 years for males in 1981, to respectively 4.2 and 3.9 years in 2015.

To summarize, the analysis of the data shows that gaps in mortality by deprivation have widened, while significant changes of composition occurred in the population. On one hand, the increase in mortality differences magnifies the impact of compositional changes on the aggregated mortality. On the other hand, significant variations in the population composition can be observed across several dimensions: age-class (Figure 1), cohort (Figure 3), and over time (Figure 2).



(a) 1981 (b) 2015

Figure 7: Male deaths by cause and IMD quintile for ages 65-85

(a) Males

(b) Females

Figure 8: Life expectancy at age 65 over 1981-2015

3 Population dynamics model

Based on the data analysis, modeling the evolution of the population dynamics appears to be instrumental in order to better understand the impact of heterogeneity and its evolution on the aggregated mortality. Age-structured population dynamics models enable the information contained in the population age-structure to be taken into account and to model its evolution. The literature on two-sex age structured models is extensive, and a wide variety of models, both stochastic and deterministic, have been proposed for different purpose in biology, demography or ecology. Nevertheless, also taking into account the population heterogeneity is often a much more difficult exercise.

The joint evolution of the subpopulations (IMD quintiles in our case) is modeled by a linear and deterministic McKendrick-Von Foerster (McKendrick (1926); Von Foerster (1959)) with time-dependent fertility and mortality rates, which we describe in the next subsection.

Due to the lack of data availability, internal (and external) migrations are not taken into account in this first modeling. Although plausible for older age classes, the previous assumption can be unrealistic for young adults. Thus, the model does not account for the full complexity of the population evolution. However, even in this streamlined framework, we are able to reproduce composition changes and capture non-trivial effects of the heterogeneous population dynamics. Moreover, this framework allows the derivation of quasi-explicit formula and asymptotic results for a number of indicators, allowing for easier interpretation.

Note that this framework can actually be extended without difficulty to the broader scope of stochastic rates, depending for instance on a random environment. Furthermore,

deterministic models are linked to stochastic population models recently developed by Tran (2008) in the field of mathematical ecology, and extended by Bensusan (2010) and Boumezoued (2016) to the framework of heterogeneous human populations, possibly in a random environment. Indeed, deterministic models can be seen as limit models when the size of the population becomes very large. However, the notion of limit model should be applied cautiously, since it does not allow the computation of "average" indicators when the studied quantities are non linear (see Boumezoued (2016), Boumezoued et al. (2018), Ferriere and Tran (2009) for further discussions on these matters).
In this section, we first recall briefly the McKendrick-Von Foerster model for a two-sex population with time dependent demographic rates (fertility and mortality rates), and then move on to the description of the joint evolution of the subpopulations and of the aggregated population.

3.1 McKendrick-Von Foerster population dynamics model

The McKendrik-Von Foerster model is a classical age-structured deterministic population model, which can be easily adapted into a two-sex model with time dependent parameters. The model is continuous in age and time, in which the population is described at time t by the function of gender and age $(g(\epsilon, a, t))$, for $a \in [0, a^\dagger[$ and $\epsilon = f$ or m respectively for females and males. $g(\epsilon, a, t)$ should be understood as the number of individuals of gender ϵ between age a and $a + \mathrm{d}a$ at time t. It follows that $\int_0^{a^\dagger} g(\epsilon, a, t)\mathrm{d}a$ is the total number of individuals of gender ϵ in the population at time t.
The demographic rates are considered as parameters of the model. Coupled to an initial age pyramid, they are the determinants of the future shape of the age pyramid.

3.1.1 Demographic rates

(i) *Mortality rate:* For each gender ϵ, the mortality rate (or force of mortality) at age a and time t is denoted by $\mu(\epsilon, a, t)$. Two types of mortality indicators are usually computed, cohort indicators and period indicators:
- The period survival function is denoted $S(\epsilon, a, t) = e^{-\int_0^a \mu(\epsilon,x,t)\mathrm{d}x}$ and represents the probability to survive to age a, in the mortality conditions of time t. The period life expectancy at age a and time t is given by:

$$e(\epsilon, a, t) = \int_a^{a^\dagger} e^{-\int_a^x \mu(\epsilon,s,t)\mathrm{d}s}\mathrm{d}x = \frac{1}{S(\epsilon, a, t)} \int_a^{a^\dagger} S(\epsilon, x, t)\mathrm{d}x.$$

- The cohort survival function is denoted by $\mathcal{S}^c(\epsilon, a, t) = e^{-\int_0^a \mu_j(\epsilon,s,t-a+s)\mathrm{d}s}$, and represents the probability for an individual born at time $t - a$ to survive until age a. The cohort life expectancy, which is the average time that individuals born at time $t - a$ will live after age a, conditional to surviving to this age, is denoted by:

$$\mathcal{E}^c(\epsilon, a, t) = \int_a^{a^\dagger} e^{-\int_a^x \mu(\epsilon,s,t-a+s)\mathrm{d}s}\mathrm{d}x = \frac{1}{\mathcal{S}^c(\epsilon, a, t)} \int_a^{a^\dagger} \mathcal{S}^c(\epsilon, x, t)\mathrm{d}x.$$

We also denote by $S(\epsilon, a - x, a, t)$ and $\mathcal{S}^c(\epsilon, a - x, a, t)$ the respective period and cohort survival probabilities from age $a - x$ to age a at time t, with $S(\epsilon, 0, a, t) = S(\epsilon, a, t)$ (the

same holds for $\mathcal{S}^c$).

Observe that the period and cohort indicators do not provide the same information. The cohort life expectancy is "real", in the sense that it gives information on particular individuals living in the population. On the contrary, the period life expectancy is only an indicator which aggregates information on all individuals living in a given population at a specific date t. The period life expectancy can be interpreted as "the average duration of life of a representative individual living in the mortality conditions of time t".

(ii) *Fertility rate*: The fertility rate for an individual with gender ϵ and age a at time t is denoted by $b(\epsilon, a, t)$. In a two-sex population, modelling births can be quite complex (Iannelli et al. (2005), Boumezoued et al. (2018)). In this article, we adopt the usual assumption that only women give births (the number of births does not depend on the number of males in the population), so that $b(m, a, t) = 0$. Females are assumed to give birth to a female with a probability $p^f = p$ and to a male with probability $p^m = 1 - p$. For sake of simplicty, the female birth rate $b(f, a, t)$ is now denoted by $b(a, t)$.

3.1.2 McKendrick-Von Foerster transport equation

The evolution of the population is given by the solution of the following transport partial differential equation:

$$(\partial_t + \partial_a) g(\epsilon, a, t) = -\mu(\epsilon, a, t) g(\epsilon, a, t), \ \forall a, t > 0 \qquad \text{(balance law)} \quad (1)$$

$$g(\epsilon, 0, t) = p^\epsilon \int_0^{a^\dagger} b(a, t) g(f, a, t) \mathrm{d}a \qquad \text{(birth law)} \quad (2)$$

$$g(\epsilon, a, 0) = g_0(\epsilon, a) \qquad \text{(initial population)}$$

Intuitively, a proportion $\mu(\epsilon, a, t)\mathrm{d}t$ of individuals of age a and gender ϵ dies time t and $t + \mathrm{d}t$, and women of age in $[a, a + \mathrm{d}a[$ give birth to $b(a, t)g(f, a, t)\mathrm{d}a$ individuals at t. Equations (1)-(2) are usually solved along its characteristic curves, or cohorts lines. The resolution method can be interpreted as counting the number of survivors at time t in each cohort of individuals, either taken from the initial population or born after the initial time. Two regimes can thus be distinguished in the evolution of the population:

(i) *Individuals present in the initial population* ($a \geqslant t$): At time t, individuals who were already present in the initial population are individuals of age $a \geqslant t$. Their number at time t is the number of individuals of age $a - t$ in the initial population who survived until time t:

$$g(\epsilon, a, t) = g_0(\epsilon, a - t)\mathcal{S}^c(\epsilon, a - t, a, t), \quad a \geqslant t. \tag{3}$$

where we recall that $\mathcal{S}^c(\epsilon, a - t, a, t) = e^{-\int_{a-t}^{a} \mu(\epsilon, s, t-a+s)\mathrm{d}s}$ is the cohort survival function from age $a - t$ to a, for individuals of age a at time t or equivalently born at time $t - a$.

(ii) *Individuals born after the initial time* ($a < t$): At time t, individuals born after $t = 0$ are individuals of age $a < t$. Their number at time t is thus the number of individuals born at time $t - a$ and who survived until time t:

$$g(\epsilon, a, t) = p^\epsilon B(t - a)\mathcal{S}^c(\epsilon, a, t), \quad a < t, \tag{4}$$

where $B(t)$ is the birth function at time t which corresponds to the number of individuals (males and females) born at time t:

$$B(t) = \int_0^{a^\dagger} b(a,t)g(f,a,t)\mathrm{d}a. \tag{5}$$

Thus, if we look at the population at a small time t, the age pyramid will be mostly shaped by the time translated initial age pyramid, and will follow the idea that "today's youths give us most of the information on tomorrow's seniors". On a longer term, the initial population is naturally erased and the shape of the future age pyramid is only characterized by the birth and survival functions.

3.1.3 Stable age profile

The stable theory provides useful indicators of the demographic rates of a population (see Keyfitz and Caswell (2005), Inaba (2017) for more details, or Webb (1985) for a more general framework). The stable theory defines a *stable age profile*, given a fixed regime of time-independent age-specific demographic rates. As for the period life expectancy, the stable age profile gives information on a fictive population, living *in the mortality and fertility conditions of a given time*. For instance its comparison with the real age profile of the population allows us to observe if strong changes in birth or mortality rates have occurred in the past.
For now, the demographic rates are assumed to depend only on age and are denoted by $\mu(\epsilon, a)$ and $b(a)$. For simplicity of notation the variable ϵ is omitted when considering the female population only.

Stable solution of the McKendrick-Von Foerster equation A stable solution of the McKendrick-Von Foerster evolution Equation (1)-(2) (for females) is a solution which can be expressed as product function $\phi(a)T(t)$. A solution of this type is called stable because its age profile, or age distribution, remains constant over time, in the sense that the proportion of individuals in each age class remains constant. By replacing $g(a,t)$ by $\phi(a)T(t)$ in (1)-(2), it follows that $\phi(a) = e^{-\lambda^* a - \int_0^a \mu(x)\mathrm{d}x}$, where λ^* is the unique solution of:

$$1 = p^f \int_0^{a^\dagger} b(a) e^{-\lambda^* a - \int_0^a \mu(x)\mathrm{d}x}\mathrm{d}a. \tag{6}$$

The previous equation is called the characteristic equation and λ is called the intrinsic growth rate of the population.

Asymptotic exponential growth Actually, a much deeper property is that the solution of (1-2) with initial population g_0 and time independent demographic rates $\mu(\epsilon, a)$ and $b(\epsilon, a)$ behaves asymptotically as a stable solution. After a long period of time, the population increases or decreases exponentially at rate λ^*, justifying the term intrisic growth rate. This means that the birth function B satisfies the following asymptotic exponential growth property:

$$B(t) \underset{t\to+\infty}{\sim} C(\lambda^*, g_0)e^{\lambda^* t}.$$

Recalling that the population evolution is defined on the long term by equation (4), the previous equation gives an asymptotic stable equivalent for the age pyramid:

$$g(\epsilon, a, t) \underset{t\to\infty}{\sim} p^\epsilon C(\lambda^*, g_0)e^{\lambda^*(t-a)}S(\epsilon, a). \tag{7}$$

To come back to the solution of the McKendrick-Von Foerster equation with time dependent rates, Equation (7) can be interpreted as follows: at a given time t_0, the right hand side of (7) is the shape that the age pyramid at t_0 would have if the demographic rates had been constant in the past, equal to $\mu(a, t_0)$ and $b(a, t_0)$. Thus, the constant λ^* summarizes the combined effects on the population of birth and death rates at time t_0.

3.2 Joint evolution of the subpopulations

3.2.1 Subpopulations evolution

In the sequel, we consider the evolution of p socioeconomic subpopulations (for instance IMD quintiles). For each $j = 1...n$, the population j is described by the solution g_j of the McKendrick-Von Foerster Equations (1)-(2) with initial population g_0^j and demographic rates $\mu_j(a, t)$ and $b_j(a, t)$, where $g_j(\epsilon, a, t)$ is the number of individuals at time t in population j, gender ϵ and between age a and $a + \mathrm{d}a$.

3.2.2 Aggregated population

We call aggregated population the global population composed of all subpopulations, denoted by $g(\epsilon, a, t)$ with:

$$g(\epsilon, a, t) = \sum_{j=1}^{n} g_j(\epsilon, a, t).$$

The aggregated population dynamics is thus defined by

$$(\partial_t + \partial_a)g(\epsilon, a, t) = -\sum_1^n \mu_j(\epsilon, a, t)g_j(\epsilon, a, t), \quad g(\epsilon, 0, t) = p^\epsilon \int_0^{a^\dagger} (\sum_1^p b_j(a, t)g_j(f, a, t))\mathrm{d}a.$$

Aggregated mortality The mortality rate at age a in the aggregated population corresponds to the intensity of individuals aged a (for each gender) dying between a short interval of time $\mathrm{d}t$. Here, the previous partial differential equation can be rewritten as:

$$(\partial_t + \partial_a)g(\epsilon, a, t) = -\left(\sum_1^n \mu_j(\epsilon, a, t)\frac{g_j(\epsilon, a, t)}{g(\epsilon, a, t)}\right) g(\epsilon, a, t).$$

The mortality rate of the aggregated population is thus:

$$d(\epsilon, a, t) = \sum_{j=1}^{n} \mu_j(\epsilon, a, t)w_j(\epsilon, a, t), \text{ with } \quad \omega_j(\epsilon, a, t) = \frac{g_j(\epsilon, a, t)}{g(\epsilon, a, t)}. \tag{8}$$

Actually, the mortality rate in the aggregated population should be denoted by $d(\epsilon, a, t, (g_j)_{j=1..p})$, since it depends on the age pyramids of all subpopulations. The time dependence of the aggregate mortality rate is caused not only by the time-dependence of the specific mortality rates in each subpopulation, but also by the evolution of the proportions $w_j(\epsilon, a, t)$

of individuals in each subpopulation and age class.
In order to better understand how these weights could impact the aggregated population, let us again make the distinction between the two cases $a \geqslant t$ and $a < t$. In order to simplify notations, we consider for the remainder of this section that the aggregated population is composed of two subpopulations ($p = 2$). When there is no ambiguity, we also omit the gender variable ϵ.

Aggregated mortality over the short term For individuals present in the initial population (of age $a \geqslant t$ at time t), the age pyramid of the subpopulations are mostly shaped by the initial subpopulations and Equation ((3)) yields for $a \geqslant t$:

$$d(a,t) = \frac{g_0^1(a-t)\mathcal{S}_1^c(a-t,a,t)\mu_1(a,t) + g_0^2(a-t)\mathcal{S}_2^c(a-t,a,t)\mu_2(a,t)}{g_0^1(a-t)\mathcal{S}_1^c(a-t,a,t) + g_0^2(a-t)\mathcal{S}_2^c(a-t,a,t).} \tag{9}$$

For small times t, the previous equation holds for most ages and the aggregated mortality depends on three factors:
- The subpopulations mortality rates μ_1 and μ_2.
- The initial subpopulations g_0^1 and g_0^2: the aggregated mortality rate at age a depends on the initial composition of the age class $a - t$, since individuals are assumed to stay in the same subpopulation. In particular, if the initial age pyramid is very heterogeneous in age, i.e. if the age classes are composed very differently, aggregate death rates could experience significant changes (for instance if younger individuals are more deprived than older ones, this could lead to an increase in aggregated mortality rates).
- Cohorts survival: if the initial age pyramids in each subpopulation are equal ($g_0^1 = g_0^2$) Equation ((9)) becomes:

$$d(a,t) = \frac{\mathcal{S}_1^c(a-t,a,t)\mu_1(a,t) + \mathcal{S}_1^c(a-t,a,t)\mu_2(a,t)}{\mathcal{S}_1^c(a-t,a,t) + \mathcal{S}_2^c(a-t,a,t)} \quad a \geqslant t.$$

This illustrates a well known "selection" effect which is that if a subpopulation, say subpopulation 2, experiences a higher overall mortality, subpopulation 1 will have more and more weight at older ages and the aggregated mortality will tend to the mortality rate of subpopulation 1.

Aggregated mortality over the long-term On a longer term, the subpopulations evolution are mainly governed by the birth functions B_1 and B_2 and for $t > a$:

$$d(a,t) = p^{\epsilon}\frac{B_1(t-a)\mathcal{S}_1^c(a,t)\mu_1(a,t) + B_2(t-a)\mathcal{S}_2^c(a,t)\mu_2(a,t)}{B_1(t-a)\mathcal{S}_1^c(a,t) + B_2(t-a)\mathcal{S}_2^c(a,t)}. \tag{10}$$

Thus, if the subpopulations experience heterogeneity in birth patterns for a certain period of time, this can induce a temporary variation in the aggregated mortality rates and generate a so-called *cohort effect*.

In the case of time-independent demographic rates, we can obtain a stable equivalent for the aggregated death rate by replacing the birth functions with their asymptotic expression as in equation (7):

$$d(a,t) \sim p^{\epsilon}\frac{C(\lambda_1^*, g_0^1)e^{\lambda_1^*(t-a)}S_1(a)\mu_1(a) + C(\lambda_2^*, g_0^2)e^{\lambda_2^*(t-a)}S_2(a)\mu_2(a)}{C(\lambda_1^*, g_0^1)e^{\lambda_1^*(t-a)}S_1(a) + C(\lambda_2^*, g_0^2)e^{\lambda_2^*(t-a)}S_2(a)} \tag{11}$$

4 Numerical Results

We now apply the model presented in Section 3 in order to illustrate different potential impacts of heterogeneity on the mortality experienced by the aggregated population.
We first show in Section 4.1 how heterogeneity in age in the initial pyramid can affect the aggregated life expectancy and improvement rates on the short term. Then, in 4.2, we show how the effects of a cause of death reduction could be compensated for in the presence of changes to the population composition induced by a cohort effect, and thus misinterpreted when the population heterogeneity is not taken into account.

Numerical implementation The general model is implemented in C++ by discretization of the transport partial differential equation (1), using a first order implicit Euler scheme. More details on the convergence and stability of the scheme, as well as other numerical methods, can be found in the review of Pelovska and Iannelli (2006).
In the different applications presented in this section, the parameters of the model (mortality, fertility rates and initial age-pyramids) are estimated from the data presented in Section 2. The inputs and outputs of the model are processed using R, and interfaced with the C++ code using the package Rcpp. When only central death rates by five-year age classes are given, mortality rates are estimated based on the fitting procedure described in Appendix C. Furthermore, individuals in each IMD quintile of age above 85 are grouped in the age class "85 and older" in dataset 1[7]. To overcome this difficulty, we assume that individuals are distibuted in age classes until age 110 as in the UK population[8]. This assumption is consistent with the observation that mortality rates in all IMD quintiles converge at old ages (see e.g. Figure 4).
For illustrative purposes, we consider in the remainder of this section the evolution of a synthetic heterogeneous population composed of two subpopulations: the most and least deprived IMD quintiles (denoted by 1 and 5 in the following).

4.1 Impact of heterogeneity in the initial age pyramid

As seen in Section 2.2, the English population shows a strong *heterogeneity in age*, meaning that the population composition can vary a lot according to the age class. Furthermore, the composition of the population has also varied significantly over time. In this first subsection, we illustrate how the heterogeneity in age of the age-pyramid can impact two mortality indicators in the aggregated population: the period life expectancy at age 65 and the average annual mortality improvement rates at ages above 65.

Demographic scenarios In order to isolate the influence of changes in the population composition, we assume in this application that mortality rates in each subpopulation do not depend on time. Thus, the mortality rates in each subpopulation $j = 1, 5$ are

[7] 90 and older in dataset 2.

[8] The English age pyramid after age 85 is provided by the Human Mortality Database (The Human Mortality Database (2016)).

now denoted by $\mu_j(\epsilon, a)$. The aggregated mortality rate defined in Equation (8) is now equal to $d(\epsilon, a, t) = \sum_{j=1}^{p} \mu_j(\epsilon, a) w_j(\epsilon, a, t)$, and is only modified through changes in the composition of the synthetic population. The population is simulated on the "short term" (30 and 40 years) according to two different scenarios. The initial age pyramids and mortality rates of the population are based on the population and mortality data of year 1981 in the first scenario presented below, and year 2015 in the second scenario, so that:

$$d(\epsilon, a, t) = \sum_{j=1,5} \hat{\mu}_j(\epsilon, a, y) w_j(\epsilon, a, t), \quad g_0(\epsilon, a) = \sum_{j=1,5} \hat{g}_j(\epsilon, a, y), \tag{12}$$

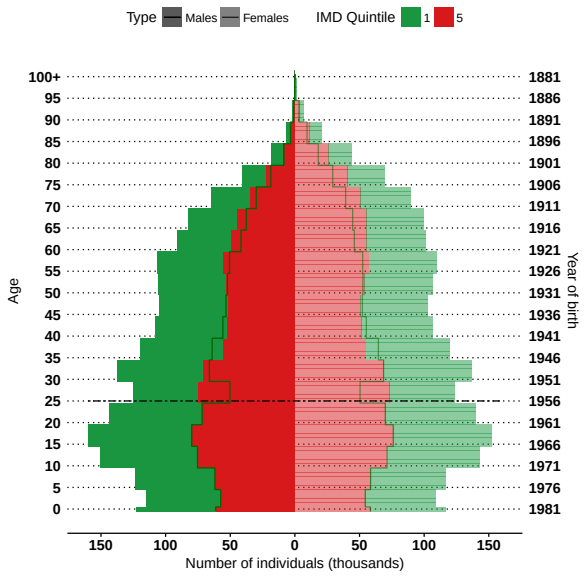
where $\hat{\mu}_j(\epsilon, \cdot, y)$ and $\hat{g}_j(\epsilon, \cdot, y)$ are the mortality rates and age-pyramid of subpopulation j fitted from the data of year $y = 1981$ and 2015.
Note that the age pyramids are structured differently in dataset 1 (years 1981-2007), structured by 5 year age classes, and in dataset 2 (2001-2015) which is structured by single year of age. Moreover, there are no data for the age class 0-25 in dataset 1. Hence, when $y = 1981$, individuals aged in 0-25 are assumed to be distributed identically in each subpopulation, based on the English age-pyramid. However, this hypothesis has no influence in this subsection since only indicators at ages above 65 are considered, over a period of 40 years. Thus, only individuals who were initially more than 25 years old are taken into account in the computation of the aggregated indicators, and therefore fertility rates have no influence on the results either. This justifies our terminology "short term".

Evolution of the aggregated population The evolution of the aggregated population is presented in Figure 9 for the population based on the 1981 data ($y = 1981$), and in Figure 10 for the population based on the 2015 data ($y = 2015$). The initial age pyramids are displayed in Figure 9a and Figure 10a. Each age class is represented by the addition of individuals in the most deprived subpopulation (5, in red) and in the least deprived IMD subpopulation (1, in green). The green line in each graph represents the shape of the least deprived subpopulation age-pyramid.
The differences in the age-structure of the two pyramids makes the comparison more difficult. However, we can see that there are more individuals from the baby boom cohort (of age around 30-35 in 1981 and 65-70 in 2015) in subpopulation 1 than in subpopulation 5. We can also see that older individuals are more deprived in the 1981 synthetic age pyramid than in the 2015 age pyramid, as already noted in section 2.2.
The age pyramids after 30 years are represented in Figures 9b and 10b. The two regimes in the population evolution described in 3.1 are distinguished in each figure by the black dashed line. Individuals of age over 30 were initially present in the population, and the age pyramid is defined by Equation (3). Individuals of age under 30 were born after the initial time, and the age pyramid is defined by Equation (4).
For the population based on the 1981 data (Figure 9), a decrease in deprivation can be observed for individuals over 60, due to the fact that younger cohorts were initially less deprived than older ones. Despite the important differences in mortality rates between the most and the least deprived subpopulations, there were initially more individuals of age over 60 in the most deprived subpopulation.

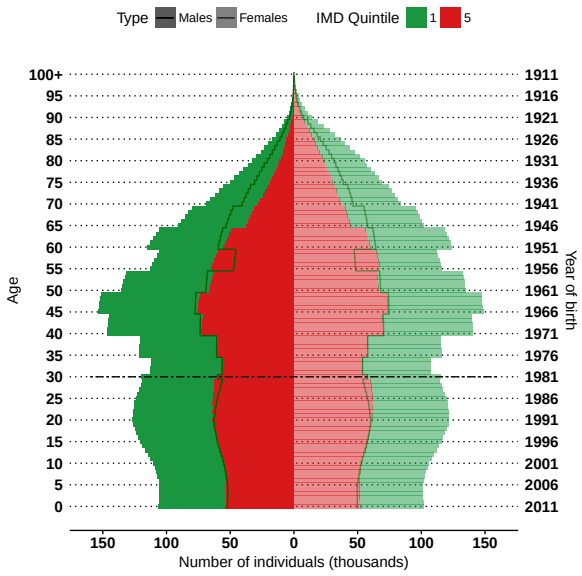
The population based on the 2015 data (Figure 10) presents important differences in cohorts sizes. The evolution of the larger cohort of individuals born during the 60s and initially of age 45-55 years, induces a significant increase in the size of the age class 75-85. However, the proportion of individuals in each subpopulation appears to be rather stable, with around 60% of individuals in subpopulation 1. On the other hand, the composition of the age class 55-70 changes significantly from Figure 10a to 10b. Initially in this age class, individuals are deprived, with more individuals in subpopulation 1. However, the situation is reversed after 30 years, with an increase in the number of individuals in subpopulation 5.

Impact on the aggregated mortality Since mortality rates in each subpopulation are assumed to be time-independent, the aggregated mortality only evolves due to changes in the subpopulations age pyramids. In order to quantify the impact of these composition changes, we first compute the period life expectancy at age 65 over time. Only individuals who were initially of age over 25 are taken into account, and the computation of the index is thus based on Equation (9) for the aggregated death rate ("short-term" regime).
The evolution of the male period life expectancy at age 65 is represented in Figure 11, and the female period life expectancy is presented in Appendix D.1. The axis on top of each graph represents the age at initial time of individuals who are 65 years old at time t. The results show that the indicator moves in *opposite directions*, according to the year chosen for the initial pyramid, Figure 11a for initial year 1981 and Figure 11b for initial year 2015. The evolution of the life expectancy at age 65 for the population based on the 1981 data is represented on Figure 11a. The evolution of life expectancy between $t = 30$ and $t = 40$ should be interpreted with caution. Indeed, we observe significant internal and external migrations in the data at ages 25-35, due to individuals moving (see e.g. Figure 3). Thus, the decrease of the period life expectancy between $t = 30$ and $t = 40$ in Figure 11a is also caused by the fact that the model does not take these changes into account.
The evolution of life expectancy at age 65 for the population based on the 2015 data is represented on Figure 11b. Contrary to Figure 11a, the life expectancy *decreases* of about 6 months in this case, after a short initial increase of the life expectancy. Thus, the evolution of the age pyramids of 2015 in the least and most deprived subpopulations contributes *negatively* in this model to the evolution of the life expectancy.

To supplement the analysis, average annual mortality improvement rates at ages above 65 are represented in Figure 12 for males and in Appendix D.1 for females. The study of average annual mortality improvement rates confirms the observations made on the evolution of life expectancy. In the conditions of 1981, average mortality improvement rates are *positive* at all ages, meaning that the evolution of the population composition contributes positively to the reduction of mortality rates. On the contrary, in the conditions of 2015, average improvement rates are *negative* at all ages below 80.

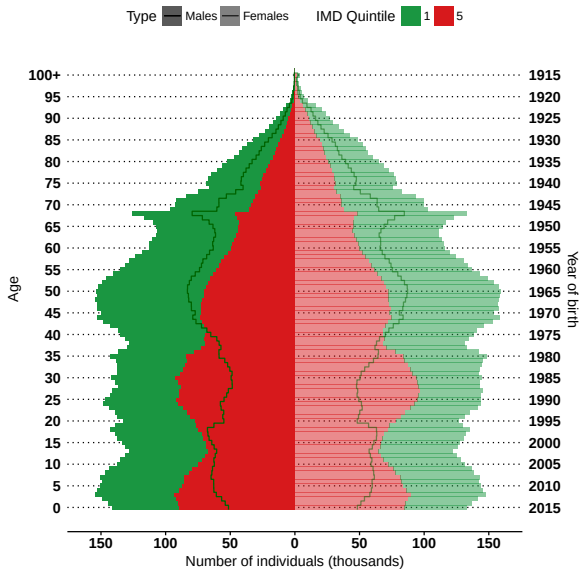


(a) Initial age pyramid

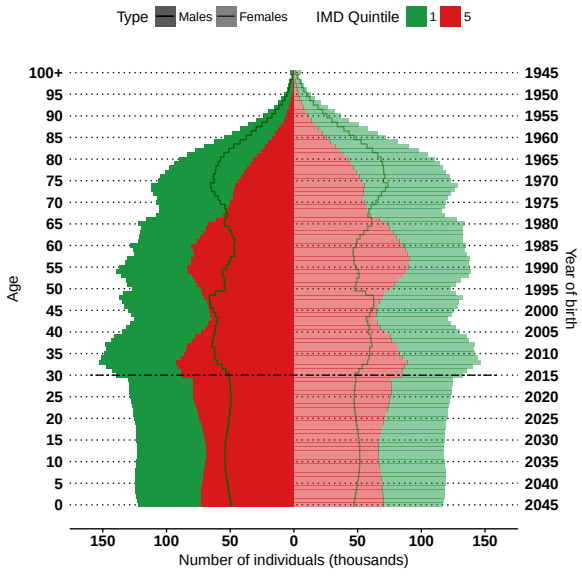


(b) Age pyramid at t=30

Figure 9: Evolution of the aggregated age-pyramid, $y = 1981$

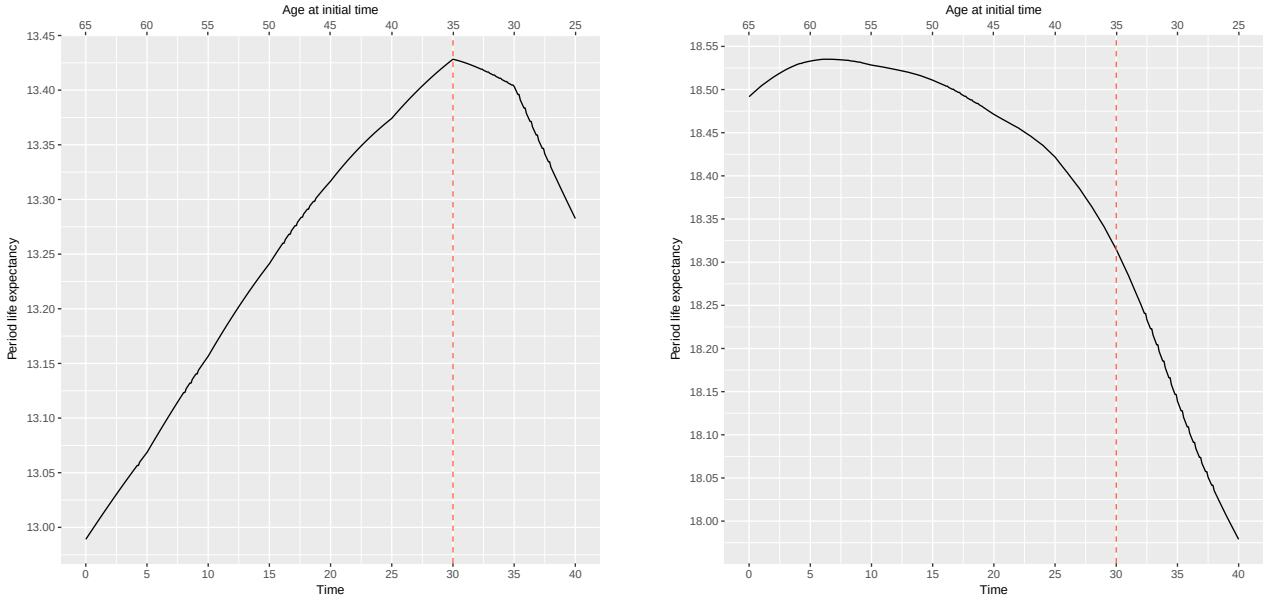


(a) Initial age pyramid



(b) Age pyramid at t=30

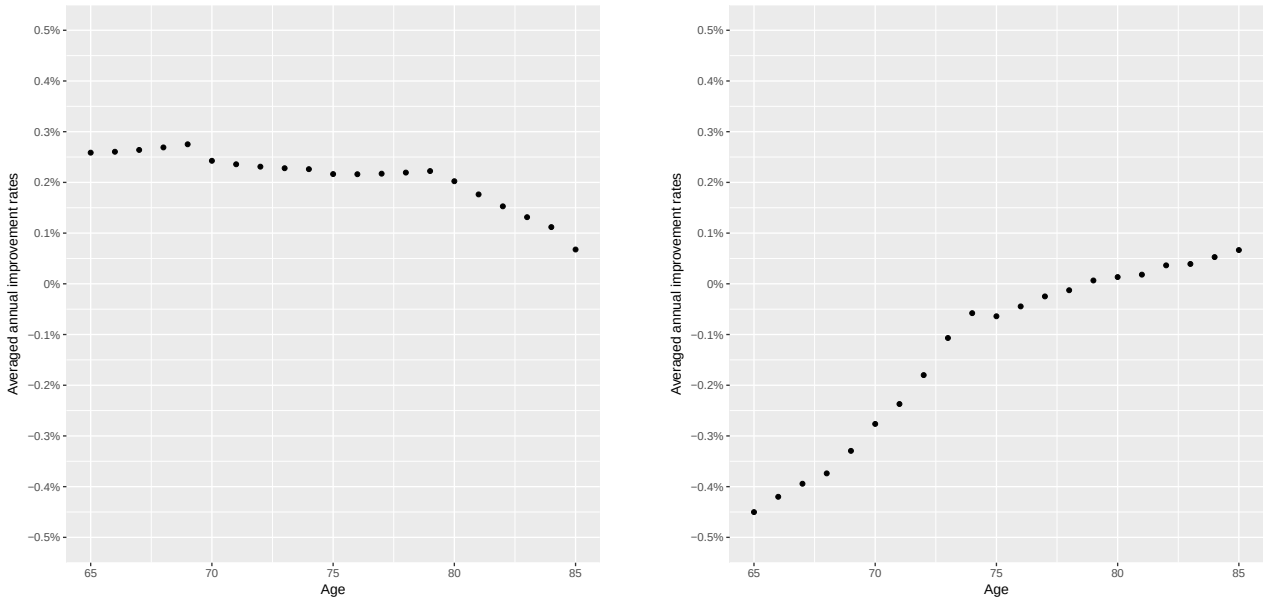
Figure 10: Evolution of the aggregated age-pyramid, $y = 2015$



(a) 1981 inputs

(b) 2015 inputs

Figure 11: Evolution of male life expectancy at age 65



(a) 1981 data

(b) 2015 data

Figure 12: Average male annual mortality improvement rates

To summarize, the two examples presented above demonstrate the importance of the heterogeneity of the age pyramid. The results show that the favorable changes in the composition of the population since 1981 probably contributed to the increase in the aggregated life expectancy. On the other hand, the model shows that based on the population of 2015, changes in the composition of the population might contribute negatively

to the evolution of the aggregated life expectancy, and might offset improvements in the subpopulations mortality rates.

4.2 Cause of death reduction and compensation effect

The numerical results of the previous subsection show how changes in composition of the heterogeneous population can impact the aggregated mortality, even when mortality rates in each subpopulation do not change over time. We are now interested in studying the evolution of aggregated mortality indicators when changes of composition are coupled with a cause-of-death mortality reduction. In the combined presence of changes in population composition and cause-specific mortality reduction, the analysis of aggregated data can be quite complex. This second numerical application illustrates the difficulty of interpreting the evolution of mortality indicators at the aggregated level when these changes of different nature are coupled.
In the following, changes in the population composition are quantified by birth patterns, and coupled with cause-specific reduction of mortality. The classical assumption of independence between causes of death[9] is assumed. By first studying the evolution of the population dynamics when demographic rates are changed separately, we show that period and cohort indicators computed on the aggregated population are able to capture changes in mortality of different natures. Finally, the results presented at the end of the section show how a cause of death reduction could be compensated and thus misinterpreted in presence of heterogeneity, due to structural changes in the population composition.

Baseline scenario (0) We first define a baseline scenario serving as a reference scenario, and for which the computed indicators are nearly constant. To that end, demographic rates are taken as time-independent in this reference scenario. The mortality rates in each subpopulation are the fitted mortality rates of year 2015: $\mu_j(\epsilon, a) = \hat{\mu}_j(\epsilon, a, 2015)$. Under this scenario, the period and cohort life expectancies in each subpopulation are thus fixed and equal to the 2015 period life expectancies. For example, the male life expectancy at age 25 is 58.1 in the least deprived subpopulation and 50.8 the most deprived subpopulation (60.8 and 55.1 for females).
Fertility rates are assumed to be the same in each population and estimated from English fertility rates in 2015[10]. Over the reporting period 1981-2015, the probability p^f of giving birth to a female is estimated to be between 0.4843 and 0.4886 (depending on the year).
We also have to chose initial age pyramids for each subpopulation whose evolution will not impact aggregated mortality indicators. As we have seen in the previous results, the heterogeneity of the age pyramid can lead to significant changes in the aggregated mortality even when mortality rates are time independent, and its choice should be made carefully. In order to limit the influence of the initial age pyramids, a natural choice could be to consider the same initial pyramid for each subpopulation, corresponding to an aggregated

[9] We refer to Chiang (1968) or Boumezoued et al. (2018) for a discussion on this assumption.

[10] English female fertility rates are estimated from Office for National Statistics (2015), for 5-year age classes and for ages 15 to 44 (see Boumezoued et al. (2018) for more details on the estimation).

population in which each age class is composed of the same number of individuals of each subpopulation. However, this choice of initial population leads to important variations in the aggregated mortality, due to the modification of cohorts composition over time induced by differences in mortality rates between the two subopulations. Since individuals in subpopulation 5 have higher mortality, the composition of cohorts is modified as individuals grow older, generating a change of composition in older age classes from the initial "50/50" distribution to a distribution composed of more individuals in subpopulation 1. The solution to this issue is to choose initial pyramids which "corresponds" to the specific mortality rates of each subpopulation, that is the stable pyramids corresponding to the specific demographic rates of each supopulation, as defined in Section 3.1.3. Since fertility rates of each subpopulation are the same in the baseline scenario, the intrinsic growth rates of each subpopulation are very close. This guarantees that the composition of the aggregated population stays almost constant over time, with cohorts composed at birth of approximately 52% of individuals in the most deprived subpopulation. In particular, the aggregated male period and cohort life expectancies at age 25 under the baseline scenario are 54.3 (57.8 for females).

Scenario 1a: cause of death mortality reduction In the first scenario, we consider a progressive reduction of mortality rates from Cardiovascular Diseases (CVD) (cause 2), which could be the result for instance of a targeted public policy or improvements in medical care.

The initial mortality rates (at $t = 0$) in each subpopulation $j = 1, 5$ are equal to the fitted 2015 mortality rates, as in the baseline scenario, and can be written as the sum of the fitted mortality rates $\hat{\mu}^{\epsilon}_{ij}$ for the m causes of death available in the datasets:

$$\mu^{\epsilon}_{j}(a, 0) = \sum_{i=1}^{m} \hat{\mu}^{\epsilon}_{ij}(a, 2015).$$

Under scenario 1a, mortality rates from CVD are reduced linearly over a period of $h_r = 30$ years starting from the year $t_r = 40$, in order to attain a reduction of $\alpha\%$ of CVD mortality rates at the end of the period. We will discuss the effects of the choice of t_r and h_r at the end of the section.

More formally, the mortality rate at age a and year t in subpopulation j is defined as follow:

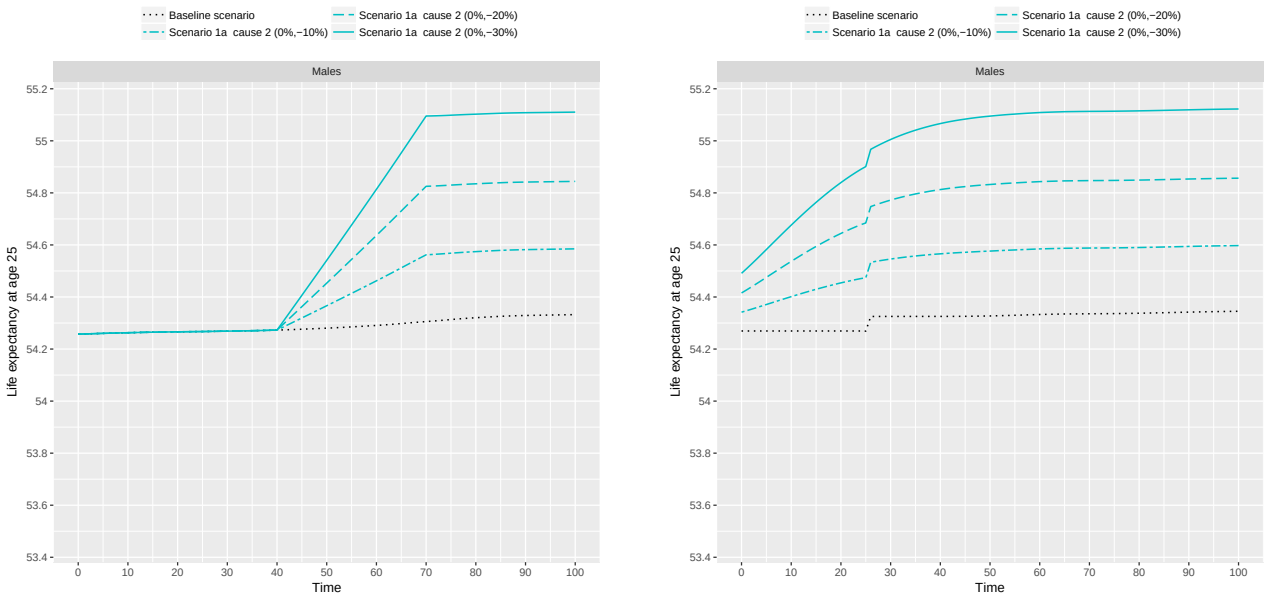
$$\mu^{\epsilon}_{j}(a, t) = \sum_{i \neq 2} \mu^{\epsilon}_{ij}(a, 0) + (1 - \alpha(t))\mu^{\epsilon}_{2j}(a, 0), \quad \alpha(t) = \mathbb{1}_{[t_r, t_r + h_r[}(t) \frac{\alpha}{h_r}(t - t_r). \tag{13}$$

The evolution of the aggregated period and cohort male life expectancies at age 25 is represented in Figure 13 under scenario 1a, for $\alpha = 10\%, 20\%$ and 30%. Plots for females are available in Appendix D.2.

When $\alpha = 30\%$, the reduction of CVD mortality rates generates an increase of respectively 0.7 and 0.9 years for the male period life expectancy in the least and more deprived subpopulations. As a rough order of magnitude, this represents approximately 10% and 16% of the historic increase in male life expectancy over the period 1985-2015.

Due to the reduction in mortality from CVD, the aggregated male period life expectancy at 25, represented in Figure 13a, also experiences an increase ranging from 0.3 to 0.8

years, with a middle value of 0.6 years when the cause is reduced by 20%. It is interesting to note that the period life expectancy is more adapted than the cohort life expectancy to capture this type of mortality changes. Indeed, the period life expectancy is a period index based on the mortality of individuals living at time t, and is therefore able to capture the starting time and the period of the mortality reduction, and to some extent the magnitude of the mortality reduction. On the other hand, the cohort life expectancy, by including future mortality rates in its computation at a given time, has a smoother evolution, which makes the interpretation of underlying mechanisms more difficult.
These remarks are coherent with the classical approach to mortality modeling and forecasting, in which mortality rates are often represented by an age dependent function, whose evolution over time is described using time-series (see e.g. the short model review in Ludkovski et al. (2016)). However, if the period dimension appears to be well-suited to capture mortality changes such as a cause-of-death mortality reduction, we show in the following that variations in mortality caused by changes in population composition are not necessarily well-captured by period indexes.



(a) Period life expectancy

(b) Cohort life expectancy

Figure 13: Aggregated life expectancy: cause-of-death mortality reduction (scenario 1a)

Scenario 1b: 'reverse' cohort effect In the second scenario, changes in the aggregated mortality are caused by mechanisms of a different nature: changes to the population composition. These changes in the population composition are generated by a 'reverse' cohort effect, modeled by fertility rates which are different in each subpopulation. Over a period h_b, the fertility rates of the most deprived subpopulation ($j = 5$) are increased of $\beta\%$, so that:

$$b_5(a,t) = \hat{b}(a,2015)(1 + \beta \mathbb{1}_{[0,h_b[}(t)). \tag{14}$$

The evolution of the aggregated period and cohort male life expectancies at age 25 under this scenario is represented in Figure 14, for $h_b = 20$ years and $\beta = 20\%, 40\%$ and 60%.

Plots for females are available in Appendix D.2.
Due to the higher fertility rates in the most deprived subpopulation over the period $[0, h_\beta[$, the composition of the population changes and the weight of the most deprived subpopulation becomes more important in the aggregated population/mortality. Cohorts born during this period are more deprived than in the baseline scenario. Thus, in our example, the cohorts born during the first 20 years of the simulation are respectively composed of 56%, 60% and 63% of individuals in the most deprived subpopulation.
Before time $t = 25$, the period life expectancy at age 25 is based on individuals present in the initial population, and since mortality rates are the same as in the baseline scenario, the two scenarios do not differ up until this time. From time $t = 25$, the computation of the period and cohort life expectancy at age 25 includes individuals born during the reverse cohort effect. Due to the increase in deprivation among the cohorts born over this period, the most deprived subpopulation has more weight in scenario 1b, which leads to a degradation of the aggregated mortality.
The cohort dimension is now more adequate to capture mechanisms underlying aggregated mortality changes under this scenario. At time $t = 25$, the cohort life expectancy (Figure 14b) experiences a negative jump, which captures the sudden increase in deprivation of the cohort of individuals aged 25 at t (born at the beginning of the reverse cohort effect). The last cohort born during the reverse cohort effect attain age 25 at time $25 + h_b = 45$, corresponding to the sudden increase in the cohort life expectancy. After time t=45, the cohort life expectancy progressively decreases over 20 years, corresponding to the introduction of generations born from these more deprived cohorts.
Changes of composition in these cohorts alone (which have the same mortality as in the baseline scenario) generates a decrease in the cohort life expectancy ranging from 0.3 to 0.8 years as β increases. Thus, changes in the life expectancy due to changes in the population composition are of the *same order of magnitude* as the changes caused by the cause-of-death mortality reduction of scenario 1a.
On the other hand, the period life expectancy, which is smoother, progressively integrates individuals from the more deprived cohorts over time. It is thus difficult to understand the underlying factors responsible for the life expectancy decrease, based only on the period index. The population dynamics have to be observed on a longer term in order to see the impact of composition changes on the period life expectancy, as the changes are due to differences in fertility rates.

Scenario 2 In this last scenario, we study the combined effects of cause-of-death mortality reductions and reverse cohort effects and illustrate how the impact of a cause of death reduction could be missed or misinterpreted in presence of heterogeneity, when the composition of the population changes. As in scenario 1a, we consider progressive cause-of-death mortality reductions for different causes of death, with $\alpha = 20\%$, but now coupled with the 'reverse' cohort effect presented in scenario 1b, with $\beta = 20\%, 40\%$ and 60%. The evolution of the period and cohort life expectancies at age 25 over time under this last scenario is represented in Figures 15 and 16 for males, and in Appendix D.2 for females.

Figure 15a represents the evolution of the period life expectancy for a progressive reduction of mortality from CVD under this scenario. As in scenario 1a, the CVD mortality is progressively reduced from $t_r = 40$, and the period life expectancy increases over the reduction period ($h_r = 30$ years). However, the increase in the aggregated life expectancy is now compensated for by the higher mortality of cohorts born during the reverse cohort effect, which have a higher level of deprivation. Firstly, the life expectancy progressively decreases up to time t=40, due to the 'reverse' cohort effect. Indeed, the most deprived cohorts are taken into account in the computation of the period life expectancy while the reduction of mortality has not occurred yet. At the end of the reduction period ($t = t_r + h_r$), individuals with higher mortality are in the age class 50-70 and increase the compensation of the mortality reduction by having more weight in the period life expectancy.

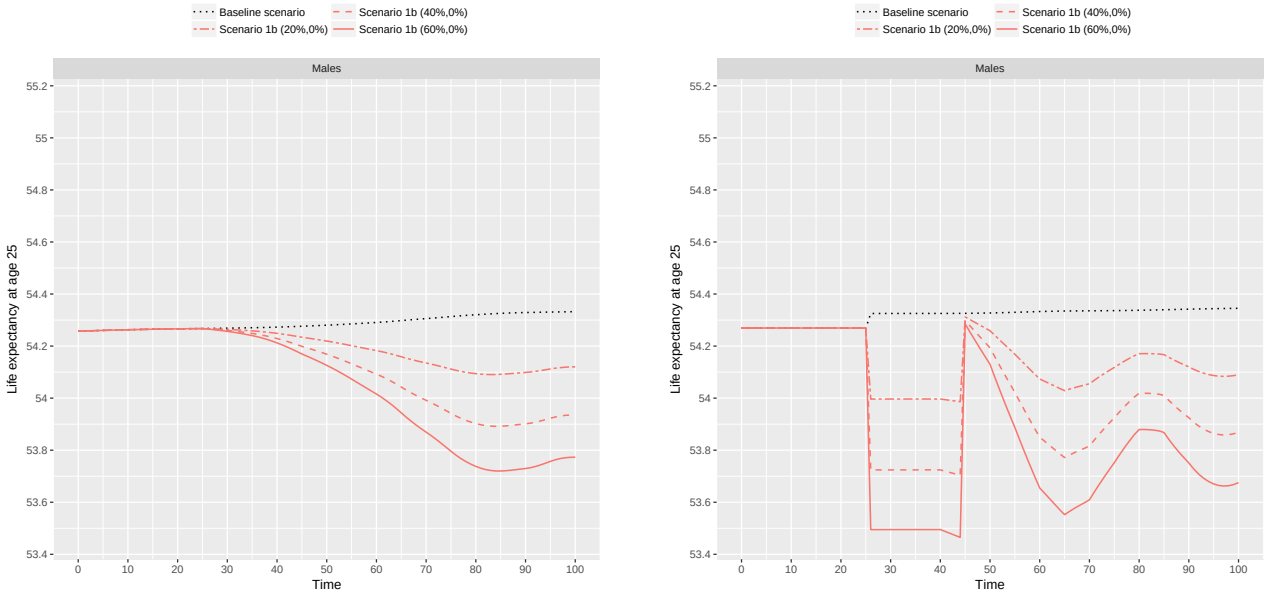
Due to this compensation effect, the period life expectancy is lower in scenario 2 than in scenario 1a (Figure 15a). When the CVD mortality reduction attains $\alpha = 20\%$, the male period life expectancy at age 25 is respectively 54.6, 54.5, 54.4 in scenario 2 with $\beta = 20\%, 40\%$ and 60%, in comparison with 54.8 where there are no changes of composition in the population (scenario 1a). Depending on the value of β, this corresponds to a compensation of 32%, 58% and 81% of the increase in the period life expectancy caused by the CVD mortality reduction. When individuals born during the reverse cohort effect grow older, the cause-of-death mortality reduction can even be counterbalanced by the cohort effect. This is the case when $\beta = 60\%$[11]: the increase in life expectancy due to the cause specific reduction is compensated for when the oldest individuals in the more deprived cohorts attain age 76.

The evolution of the period life expectancy under scenario 2 for different causes of death and with $\alpha = 20\%$ and $\beta = 60\%$ is represented in Figure 15b. The causes considered are numbered from 1 to 4 per order of importance: neoplasms (1), CVD (2), respiratory diseases (3) and external causes (4). With these parameters, all cause-of-death mortality reduction, except neoplasms, are eventually compensated for by the cohort effect. Naturally, the direction and magnitude of the impact of the cause reductions depend on the cause of death that is reduced, since different causes do not impact all age groups and socioeconomic categories in the same way. In particular, for causes 3 and 4, which are less important, the aggregate period life expectancy even decreases over time due to the compensation effect, and that could be incorrectly interpreted as an increase of individual mortality.

The evolution of the cohort life expectancy at 25 under scenario 2 is represented in Figures 16. Similar conclusions hold as well for the aggregated cohort life expectancy in term of cause-specific mortality reduction compensation. Nevertheless, as in scenario 1b, the impact of the reverse cohort effect is more clear-cut when looking at the cohort index. Figure 16b shows that for all causes of death, the cohort life expectancy at 25 of the more deprived cohorts (corresponding to the cohort life expectancy from time $t = 25$ to $t = 25 + h_b = 45$) is lower than in the baseline scenario. This means that the increase in

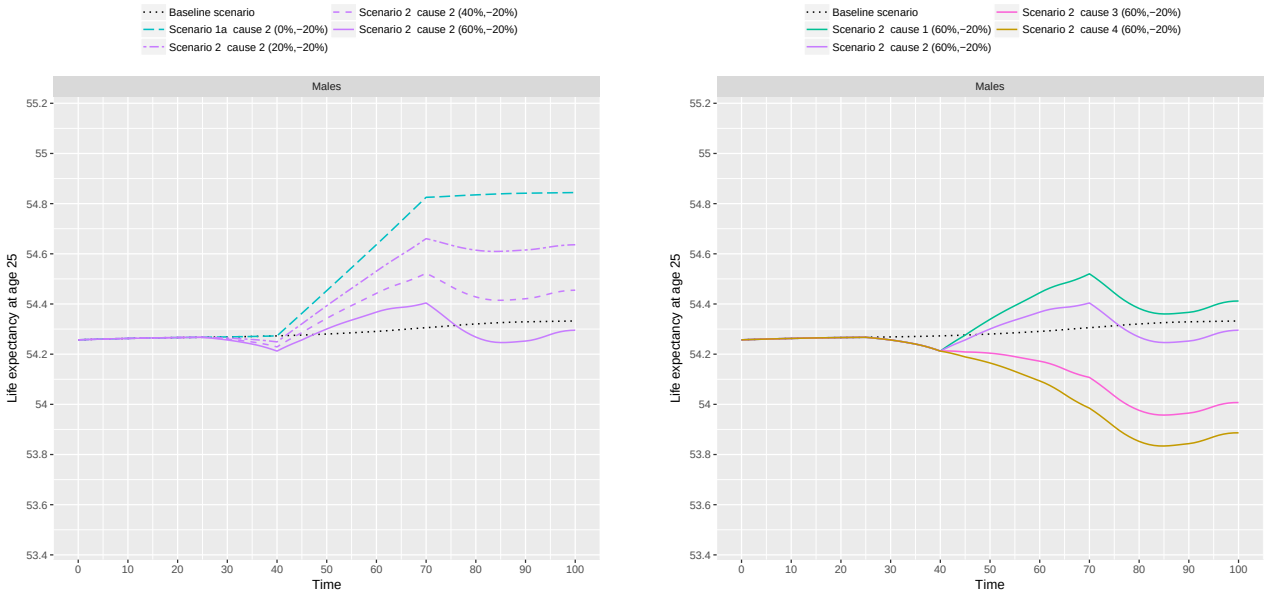
[11] Cohorts born during the period $[0, 20[$ composed of 63% of individuals in most deprived subpopulation.

deprivation of these cohorts has more impact than the cause-of-death mortality reduction.



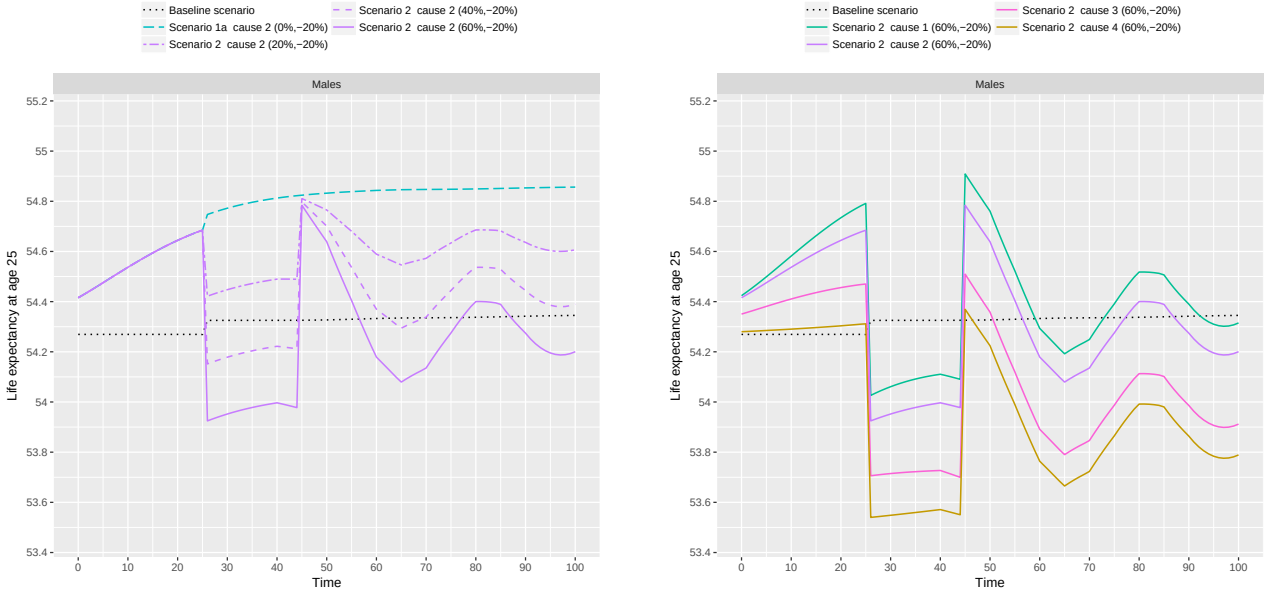
(a) Period life expectancy (b) Cohort life expectancy

Figure 14: Aggregated life expectancy: reverse cohort effect (scenario 1b)



(a) Cardiovascular diseases (b) All causes

Figure 15: Aggregated period life expectancy over time (scenario 2)



(a) Cardiovascular diseases

(b) All causes

Figure 16: Aggregated cohort life expectancy over time (scenario 2)

Discussion The demographic scenarios presented in this subsection illustrate the complexity induced by the presence of heterogeneity in the population. Figures 15 and 16 show that when the socioeconomic composition of the population changes over time, the sole study of aggregated period indexes can lead to misinterpretations of mortality trends. For instance, a "true" mortality reduction driven by a cause-of-death reduction could be minimized, or even not accounted for, due to structural changes in the population composition.

It is also important to note that the demographic changes in scenario 1a and 1b (cause-of-death mortality reduction and reverse cohort effect) operate on very different timescales, which add to the complexity of understanding aggregated indicators. In scenario 2, the change of composition due to the reverse cohort effect has a delayed impact on the aggregated mortality, and compensates for the increase in life expectancy because individuals in the more deprived cohorts (born during the period $[0, h_b]$) are old enough when the progressive cause-of-death mortality reduction starts at time t_r. When the mortality reduction starts at an earlier time t_r, the life expectancy first increases (individuals in the more deprived are too young to weight in the period life expectancy), but then starts to decrease faster at the end of the mortality reduction period (after time $t_r + h_r$), as individuals in the more deprived cohorts grow older and have more weight in the aggregated period life expectancy. On the other hand, when the mortality reduction starts at a later time t_r, the adverse mortality of individuals born during the reverse cohort effect compensates for the mortality reduction earlier and more strongly.

In scenarios 1b and 2, the population composition is changed using fertility rates, in order to compare the different demographic rates. However, compositional changes could be caused by other underlying mechanisms such as internal or external migrations.

Finally, causes of death are assumed to be independent in the model. However, this assumption could be relaxed by using a model allowing for dependence between causes of death such as that proposed in Dimitrova et al. (2013) or Alai et al. (2015). Under our assumption of independence between causes of death, a decrease in CVD mortality decreases the life expectancy gap between the least and the most deprived subpopulation, while Alai et al. (2017), using the model proposed in Alai et al. (2015), found that the cause elimination of CVD would actually have increased the life expectancy gap, under the mortality conditions of most years over the period 1981-2007. However, an increase in the mortality gap between the most and least deprived would actually magnify the effect of composition changes, and one could expect that a model allowing for dependence would strengthen the previous results.

5 Concluding Remarks

We have illustrated in this paper the complexity of understanding drivers of the aggregated mortality in the presence of heterogeneity. To that matter, we analyzed empirical evidence from two unique datasets, containing information on the English population and cause-specific number of deaths by age, gender and deprivation level over the period 1981-2015. The data analysis and numerical applications presented reveal the complex interactions between population dynamics and mortality indicators. In the context of increasing differences in mortality between different socioeconomic subgroups, changes in the population composition have more impact on the aggregated population mortality. At the same time, significant variations of deprivation levels have occurred across age groups and over time. Those notable changes in the population composition make studying the effect of composition changes even more complicated.

Numerical results presented in Section 4.1 show that favorable changes in population composition since 1981 might have contributed positively to the increase in aggregated life expectancy, while the model shows that based on the 2015 population, composition changes might offset future mortality improvements to a certain extent.

Furthermore, the reduction of a cause of death may not necessarily result in an improvement of aggregated mortality rates or life expectancy, if the composition of the population changes at the same time. In particular, a "reverse" cohort effect, or any changes in the population composition, could compensate for a cause-of-death reduction. In this case, the effect of public health policies could be misinterpreted if only aggregated data are studied. This raises the question of the assessment of mortality reduction targets set by national or international organization in the presence of heterogeneity. Indeed, it seems rather difficult, by only looking at aggregated indicators, to determine if a target is not met due to the failure of a public health policy or a change in population composition.

The results also stress the importance of studying not only mortality rates at older ages but also the whole population dynamics in the presence of heterogeneity, even when only interested in longevity indicators. The study of the whole population can give precious information complementary to mortality models which may not be able to capture population composition changes.

Depending on the nature of demographic changes and on the timing at which they impact the aggregated mortality, different changes are captured by different types of indicators, i.e. period or cohort indicators. In particular, we showed that composition changes, which can have a delayed effect on mortality, seem to be better captured by cohort indicators. Obviously, cohort indicators require to forecast future mortality rates in order to be computed. However, due to these delayed effects, taking into account the population composition and age structure gives insight into how future composition changes might impact mortality rates. Furthermore, modeling the population dynamics can help to better understand the impact of these changes, as well as to experiment and test the consistency of subpopulation mortality forecasts under different demographic scenarios.
Finally, the population dynamics framework is very flexible. The simplicity of the model used in this article allows us to derive closed formulas, and therefore to isolate and better identify specific mechanisms. However, the model could be extended to a more complex framework such as a stochastic individual-based model. In the numerical applications presented in Section 4.2, composition changes were modeled by differences in fertility rates, in order to compare these variations with a cause-of-death mortality reduction. Nevertheless, modification of the population composition could be extended to other underlying mechanisms. Thus, a more realistic modeling procedure could integrate internal migrations among the subpopulations, as well as external migrations. This constitutes a challenge due to data availability, but the population dynamics framework can operate as a tool for experimenting and simulating different scenarios. Another perspective would be to consider dependence between causes of death in the modeling, along with the study of consistency between subnational and national mortality forecasts.
To conclude, population dynamics theory is a promising and complementary field of research when modeling mortality rates in the presence of heterogeneity. The richness of available models should allow them to be used as simulation and validation tools. Additionally, they can capture dynamically effects of a different nature with respect to classical mortality modeling, such as composition changes.

Acknowledgements

The authors benefited from the financial support of the "Germaine de Staël" program *Modélisation des causes de décès dans la dynamique des populations* (Gds 2014-14). The authors also thank Madhavi Bejekal and Andres Villegas for their great help for obtaining the first dataset and their explanations.

References

Alai, D., Arnold, S., Bajekal, M., and Villegas, A. (2017). Causal mortality by socioeconomic circumstances: A model to assess the impact of policy options on inequalities in life expectancy. In Conference proceedings of the Society of Actuaries 2017 Living to 100 Symposium.

Alai, D., Arnold(-Gaille), S., and Sherris, M. (2015). Modelling Cause-of-Death Mortality and the Impact of Cause-Elimination. Annals of Actuarial Science, 9(1):167–186.

Bajekal, M. (2005). Healthy life expectancy by area deprivation: magnitude and trends in England, 1994-1999. Health Statistics Quarterly, 25:18.

Bajekal, M., Scholes, S., O'Flaherty, M., Raine, R., Norman, P., and Capewell, S. (2013a). Implications of using a fixed IMD quintile allocation for small areas in England from 1981 to 2007. PLoS One, 8(3).

Bajekal, M., Scholes, S., O'Flaherty, M., Raine, R., Norman, P., and Capewell, S. (2013b). Unequal trends in coronary heart disease mortality by socioeconomic circumstances, England 1982–2006: an analytical study. PLoS One, 8(3):e59608.

Bensusan, H. (2010). Risques de taux et de longévité: Modélisation dynamique et applications aux produits dérivés et à l'assurance vie. PhD thesis, École Polytechnique.

Boumezoued, A. (2016). Approches micro-macro des dynamiques de populations hétérogènes structurées par âge. Application aux processus auto-excitants et à la démographie. PhD thesis, Université Pierre et Marie Curie.

Boumezoued, A., Labit Hardy, H., El Karoui, N., and Arnold, S. (2018). Cause-of-death mortality: What can be learned from population dynamics? Insurance: Mathematics and Economics, 78:301 – 315. Longevity risk and capital markets: The 201516 update.

Cairns, A., Kallestrup-Lamb, M., Rosenskjold, C., Blake, D., and Dowd, K. (2016). Modelling socio-economic differences in the mortality of Danish males using a new affluence index. Working paper.

Chiang, C. (1968). Introduction to stochastic processes in biostatistics. John Wiley and Sons, New York.

Department for Communities and Local Government (2015). The English Indices of Deprivation 2015 - Statistical release. Technical report, Department for Communities and Local Government.

Department of Health (2003). Tackling Health Inequalities: A Programme for Action. Technical report, Department of Health.

Diez Roux, A. V. and Mair, C. (2010). Neighborhoods and health. Annals of the New York Academy of Sciences, 1186(1):125–145.

Dimitrova, D., Haberman, S., and Kaishev, V. (2013). Dependent competing risks: Cause elimination and its impact on survival. Insurance: Mathematics and Economics, 53(2):464–477.

El Karoui, N., Hadji, K., and Kaakai, S. (2018). Inextricable complexity of two centuries of worlwide demographic transition: a fascinating modeling challenge. Hal-01745901.

Elo, I. T. (2009). Social class differentials in health and mortality: Patterns and explanations in comparative perspective. Annual Review of Sociology, 35:553–572.

Ferriere, R. and Tran, V. (2009). Stochastic and deterministic models for age-structured populations with genetically variable traits. In ESAIM: Proceedings, volume 27, pages 289–310.

Guildea, Z., Fone, D., Dunstan, F., Sibert, J., and Cartlidge, P. (2001). Social deprivation and the causes of stillbirth and infant mortality. Archives of Disease in Childhood, 84(4):307–310.

Haberman, S., Kaishev, V., Millossovich, P., and Villegas, A. (2014). Longevity Basis Risk A

methodology for assessing basis risk. Technical report, Institute and Faculty of Actuaries (IFA), Life and Longevity Markets Association (LLMA).

Hautphenne, S. and Latouche, G. (2012). The Markovian binary tree applied to demography. Journal of Mathematical Biology, 64(7):1109–1135.

Iannelli, M., Martcheva, M., and Milner, F. A. (2005). Gender-structured Population Modeling: Mathematical methods, Numerics, and Simulations, volume 31. SIAM.

Inaba, H. (2017). Age-Structured Population Dynamics in Demography and Epidemiology. Springer Singapore.

Jarner, S. F. and Kryger, E. M. (2011). Modelling adult mortality in small populations: The SAINT model. ASTIN Bulletin: The Journal of the IAA, 41(2):377–418.

Jasilionis, D., Andreev, E. M., Kharkova, T. L., and Ward Kingkade, W. (2011). Change in marital status structure as an obstacle for health improvement: evidence from six developed countries. The European Journal of Public Health, 22(4):602–604.

Keyfitz, N. and Caswell, H. (2005). Applied mathematical demography, volume 47. Springer.

Labit Hardy, H. (2016). Impacts of cause-of-death mortality changes: A population dynamic approach. PhD thesis, Université de Lausanne.

Li, J. S.-H., Zhou, R., and Hardy, M. (2015). A step-by-step guide to building two-population stochastic mortality models. Insurance: Mathematics and Economics, 63:121 – 134. Special Issue: Longevity Nine - the Ninth International Longevity Risk and Capital Markets Solutions Conference.

Lu, J., Wong, W., and Bajekal, M. (2014). Mortality improvement by socio-economic circumstances in England (1982 to 2006). British Actuarial Journal, 19(01):1–35.

Ludkovski, M., Risk, J., and Zail, H. (2016). Gaussian process models for mortality rates and improvement factors. arXiv:1608.08291.

Mackenbach, J. P., Kunst, A. E., Cavelaars, A. E., Groenhof, F., and Geurts, J. J. (1997). Socioeconomic inequalities in morbidity and mortality in western Europe. The Lancet, 349(9066):1655 – 1659.

Marmot, M., Stansfeld, S., Patel, C., North, F., Head, J., White, I., Brunner, E., Feeney, A., and Smith, G. D. (1991). Health inequalities among British civil servants: the Whitehall II study. The Lancet, 337(8754):1387 – 1393. Originally published as Volume 1, Issue 8754.

McKendrick, A. (1926). Application of mathematics to medical problems. Proc. Edin. Math. Soc., 54:98–130.

Meyricke, R. and Sherris, M. (2013). The determinants of mortality heterogeneity and implications for pricing annuities. Insurance: Mathematics and Economics, 53(2):379–387.

Nandi, A. and Kawachi, I. (2011). Neighborhood effects on mortality. In International Handbook of Adult Mortality, pages 413–439. Springer.

National Research Council (2011). Explaining divergent levels of longevity in high-income countries. National Academies Press.

Noble, M., Mclennan, D., Wilkinson, K., Whitworth, A., Barnes, H., and Dibben, C. (2007). The English indices of Deprivation 2007. Technical report, Department of Communities and Local Government.

Norman, P. and Darlington-Pollock, F. (2017). The Changing Geography of Deprivation in Great Britain: Exploiting Small Area Census Data, 1971 to 2011. In Stillwell, J., editor,

The Routledge Handbook of Census Resources, Methods and Applications Unlocking the UK 2011 Census, pages 404–420. Routledge.

Oakley, L., Maconochie, N., Doyle, P., Dattani, N., and Moser, K. (2009). Multivariate analysis of infant death in England and Wales in 2005-06, with focus on socio-economic status and deprivation. Health Statistics Quarterly, 42(1):22–39.

Office for National Statistics (2012). 2011 Census: Population and Household Estimates for Small Areas in England and Wales, March 2011. Office for National Statistics.

Office for National Statistics (2015). Live births, stillbirths, and the intensity of childbearing measured by the total fertility rate. Access date: May 2016.

Office for National Statistics (2015). Trend in life expectancy at birth and at age 65 by socio-economic position based on the National Statistics Socio-economic Classification, England and Wales: 1982-1986 to 2007-2011. Technical report, Office for National Statistics.

Pamuk, E. R. (1985). Social Class Inequality in Mortality from 1921 to 1972 in England and Wales. Population Studies, 39(1):17–31. PMID: 11611750.

Pelovska, G. and Iannelli, M. (2006). Numerical methods for the Lotka-Mckendrick's equation. Journal of Computational and Applied Mathematics, 197(2):534–557.

Shang, H. L. and Haberman, S. (2017). Grouped multivariate and functional time series forecasting: An application to annuity pricing. Insurance: Mathematics and Economics.

Shang, H. L. and Hyndman, R. J. (2017). Grouped functional time series forecasting: An application to age-specific mortality rates. Journal of Computational and Graphical Statistics, 26(2):330–343.

Shkolnikov, V. M., Andreev, E. M., Jasilionis, D., Leinsalu, M., Antonova, O. I., and McKee, M. (2006). The changing relation between education and life expectancy in central and eastern europe in the 1990s. Journal of Epidemiology and Community Health, 60(10):875–881.

The Human Mortality Database (2016). University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). English data, access date: February 2016.

Townsend, P. (1979). Poverty in the United Kingdom. Allen Lane and Penguin Books.

Tran, V. C. (2008). Large population limit and time behaviour of a stochastic particle model describing an age-structured population. ESAIM: Probability and Statistics, 12(1):345–386.

Villegas, A. M. (2015). Mortality: Modelling, Socio-Economic Differences and Basis Risk. PhD thesis, Cass Business School.

Villegas, A. M. and Haberman, S. (2014). On the modeling and forecasting of socioeconomic mortality differentials: An application to deprivation and mortality in England. North American Actuarial Journal, 18(1):168–193.

Villermé, L. (1830). De la mortalité dans les divers quartiers de la ville de Paris, et des causes qui la rendent très différentes dans plusieurs d'entre eux, ainsi que dans les divers quartiers de beaucoup de grandes villes. Annales d'hygiène publique et de médecine légale, 3:294–321.

Von Foerster, H. (1959). The Kinetics of Cellular Proliferation. Grune & Stratton.

Webb, G. F. (1985). Theory of nonlinear age-dependent population dynamics. CRC Press.

World Health Organization (2013). Global action plan for the prevention and control of non-communicable diseases 2013-2020. Technical report, World Health Organization.

Appendix A IMD over time

The IMD is computed at fixed times, year 2007 and year 2015, and applied to larger time periods, see Figure 17. Therefore, the socioeconomic evolution of living areas over the reporting periods is not taken into account. However, the aggregation of small living areas into deprivation quintiles might reduce changes over reasonable periods of time. For the 1981-2006 period, previous studies have shown that the majority of small areas have stayed in the same deprivation quintile (se e.g. Bajekal et al. (2013a), Appendix D of Lu et al. (2014) or Norman and Darlington-Pollock (2017) for the 1970-2011 period).

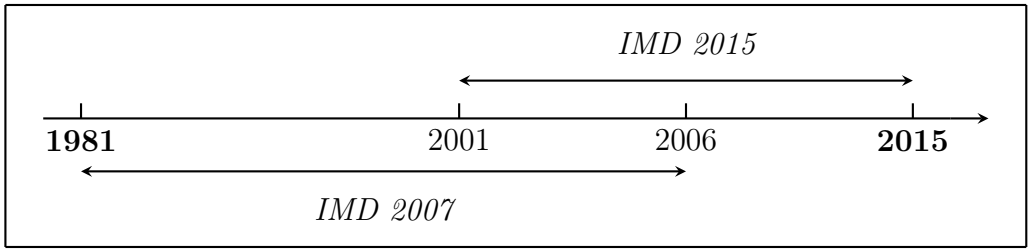
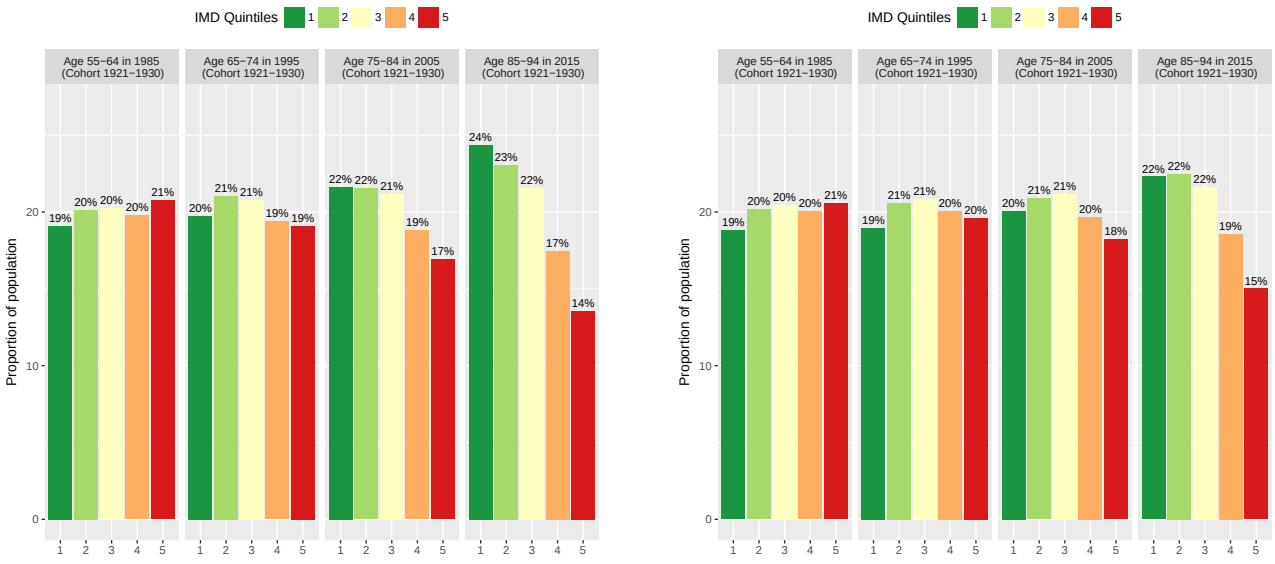


Figure 17: IMD2007-IMD2015

We have also compared data over the 2001-2006 period, computed with the IMD 2007 in the one hand (dataset 1) and with the IMD 2015 in the other hand (dataset 2). For the overlapping period 2001-2006, over the five quintiles, the relative difference in life expectancy at age 25, *respectively* at age 65, using IMD 2007 or IMD 2015 is less than $0,27\%$, *resp.* $0,78\%$, for males and less than $0,18\%$, *resp.* $0,46\%$, for females.

Appendix B Miscellaneous information on data

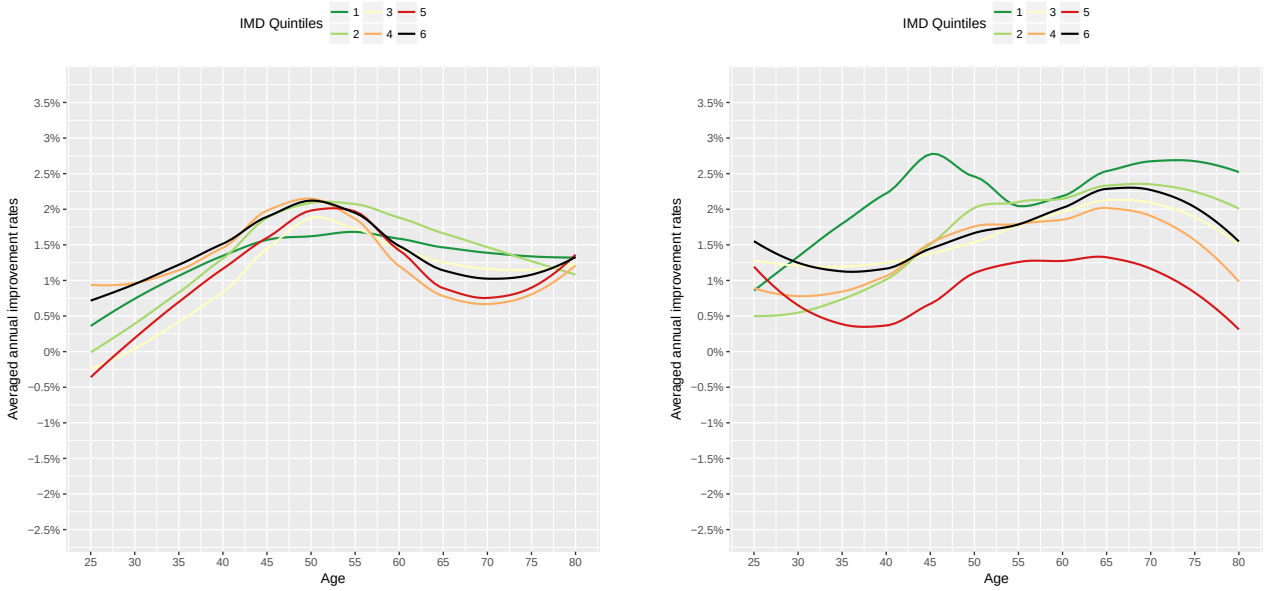
B.1 Fixed cohorts



(a) Males (b) Females

Figure 18: Proportion of individuals by IMD quintile for cohort 1921-1930

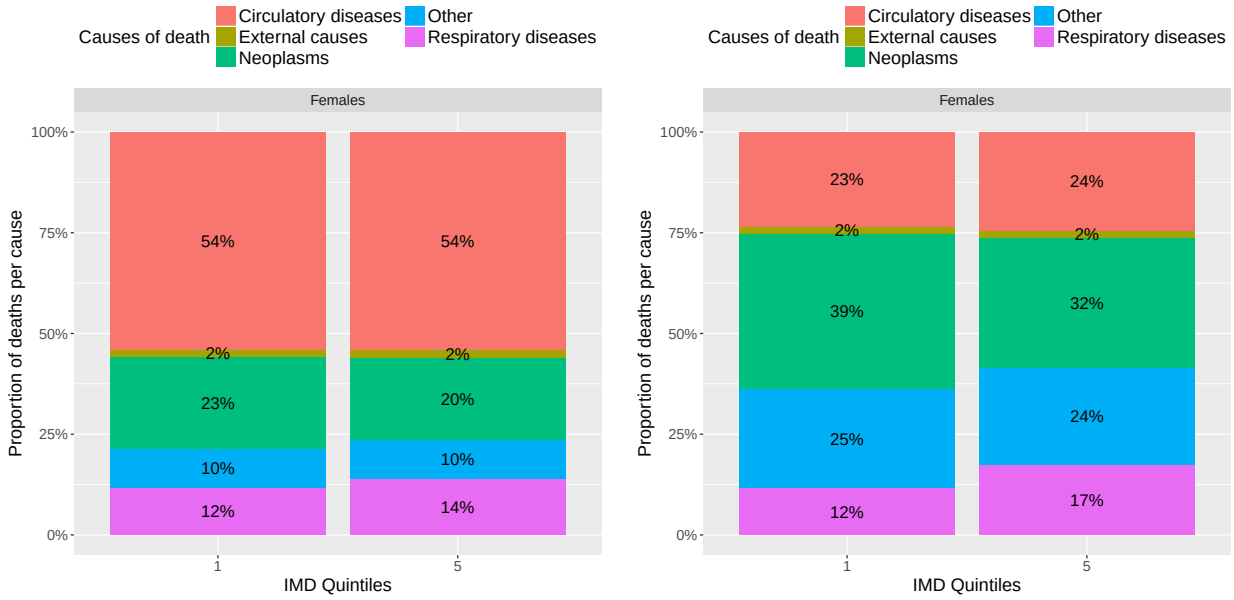
B.2 Central death rates



(a) 1981-1995 (b) 2001-2015

Figure 19: Average annual rates of improvement in mortality, females

B.3 Causes of death



(a) 1981 (b) 2015

Figure 20: Females deaths per cause and IMD quintile for ages 65-85

Appendix C Fitting procedure of mortality rates

In this paragraph, we consider the mortality of individuals of gender ϵ in a given IMD quintile j. For simplicity of notation, we omit these variables when there is no ambiguity. In both datasets, deaths by causes are given for each calendar year by 5 year age-classes (with the exception of the age class 0-5 divided in two classes, 0-1 and 1-5). Central death rates at age a and year t for 5 year age-classes can been estimated by ${}_5\hat{m}(a,t) = \frac{{}_5D(a,t)}{{}_5\hat{E}(a,t)}$; where ${}_5D(a,t)$ is the number of individuals who died during year t (in $[t, t+1[$) at an age in $[a, a+5[$, and where ${}_5\hat{E}(a,t)$ is the estimated exposure. Recall that the real exposure ${}_5E(a,t)$ is the cumulative time lived during year t by individuals in the age class[12].
In our model, we need to estimate the force of mortality $\mu(a,t)$ which is linked to the theoretical death rate by the following formula:

$$ {}_5m(a,t) = \int_t^{t+1} \int_a^{a+5} \mu(x,s) \frac{g(x,s)}{\int_t^{t+1}\int_a^{a+5} g(u,h)\mathrm{d}u\mathrm{d}h} \mathrm{d}x\mathrm{d}s \tag{15} $$

Equation (15) can be interpreted as follow: the central death rate is the average force of mortality on $[t, t+1[\times[x, x+5[$, weighted against the population distribution on this interval.
In the sequel, we make the assumption that the force of mortality is constant over each 1 year period, so that for all calendar years t, $\mu(a, t+s) = \mu(a,t) \quad \forall s \in [0,1[$. In this case,

[12] An individual attaining age a at time $t+s$ and who died at time $t+h+s < t+1$ will have weight h in the exposure.

Equation (15) can be rewritten as:

$$ {}_5m(a,t) = \int_a^{a+5} \mu(x,t) \frac{\int_t^{t+1} g(x,s)ds}{\int_a^{a+5}(\int_t^{t+1} g(u,s)du)ds} \mathrm{d}x $$

When data is structured by single year of age, the force of mortality is usually also assumed to stay constant over each age class, so that $\mu(x,t) = {}_1m(a,t)$ for $a \leqslant x < a+1$. However, this assumption seems quite unrealistic when data is aggregated over 5 year age-classes. The next simplest parametric assumption is to assume that the force of mortality is piecewise linear in age over the age classes:

$$ \mu(x,s) = \alpha(a,t)x + \beta(a,t) \quad \forall (x,s) \in [a, a+5[\times [t, t+1[. \tag{16} $$

When information on the population by single years of age is available, the distribution of the population within the age class can be approximated by a discrete distribution defined for $0 \leqslant k \leqslant 4$ by:

$$ \frac{\int_t^{t+1} g(x,s)ds}{\int_a^{a+5} \left(\int_t^{t+1} g(u,s)ds\right) \mathrm{d}u} = \hat{E}_k, \quad \forall x \in [a+k, a+k+1[, $$

where $\hat{E}_k$ is the estimated proportion of individuals of age in $[a+k, a+k+1[$ in the the age class. By replacing $\mu(a,t)$ and $g(x,s)$ in (15) with the new assumptions, we obtain that Equation (16) should be a line passing through ${}_5m(x,t)$ at the mean age of individuals in the age class. Therefore, an inductive procedure can be defined in order to fit the force of mortality for year t:

(i) *Initialization:* choose $\mu(0,t)$.

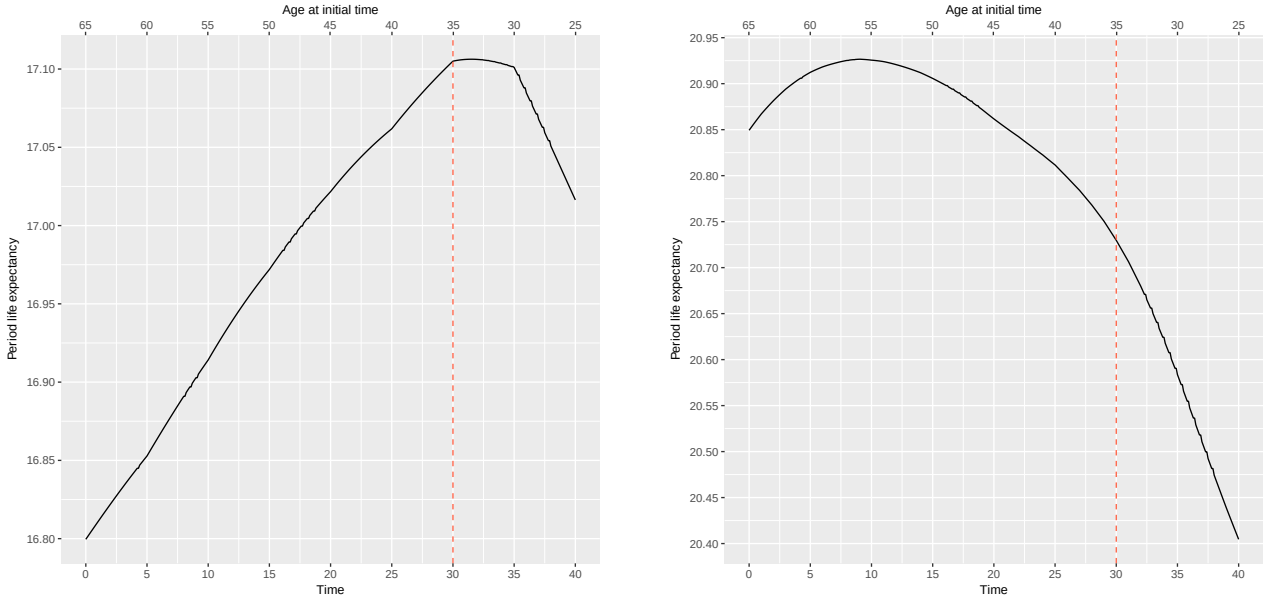
(ii) *Induction:* Assume that the mortality rate has been fitted until the ith age class $[a_i, a_{i+1}[$. The mortality rate on the next age class $[a_{i+1}, a_{i+2}[$ is the line passing through $\mu(a_{i+1}, t)^-$ at a_{i+1} and ${}_5m(x_{i+1}, t)$ at $\bar{x}_{i+1}$.

(iii) Reiterate the second step on the next age class.

The main advantage of the piecewise linear approximation is to be consistent with the theoretical definition of the aggregated and specific central death rates in our heterogeneous population model. However, the degree of liberty in the choice of the initial point $\mu(0,t)$ is a drawback of the method, and the fitting is not guaranteed to obtain positive death rates. In the numerical applications the initial point $\mu(0,t)$ is found by an optimization procedure. A similar fitting approach has been proposed by Hautphenne and Latouche (2012), with possible discontinuities in the death rates. See also Villegas and Haberman (2014) for an alternative method.

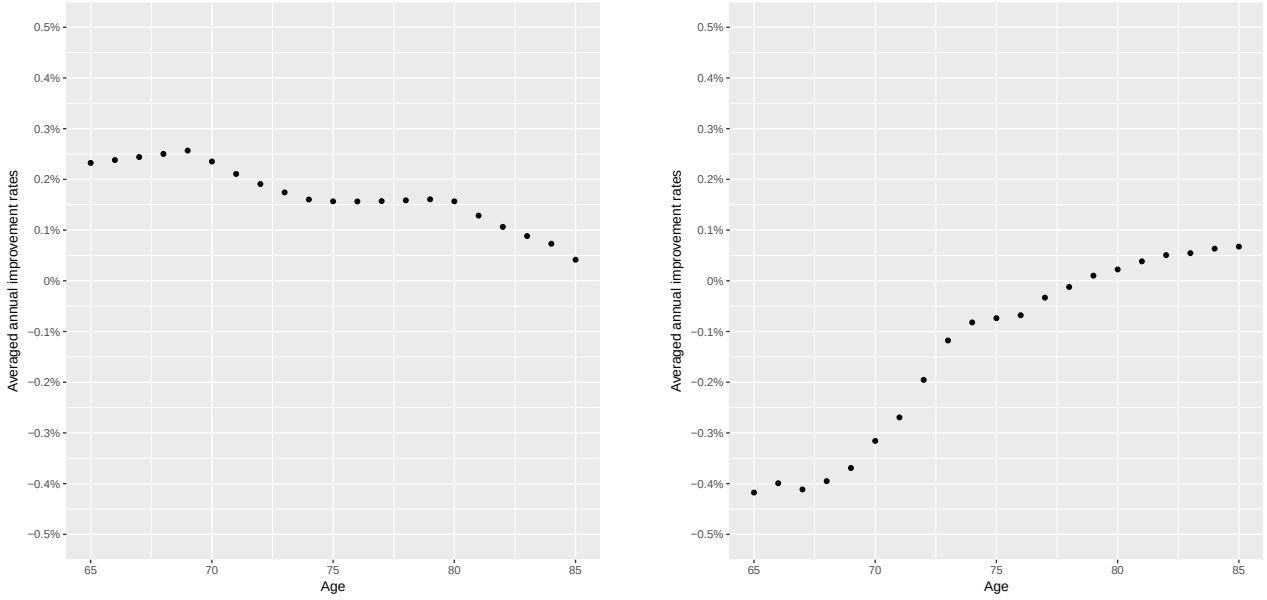
Appendix D Additional results

D.1 Impact of heterogeneity in the initial age pyramid



(a) 1981 data (b) 2015 data

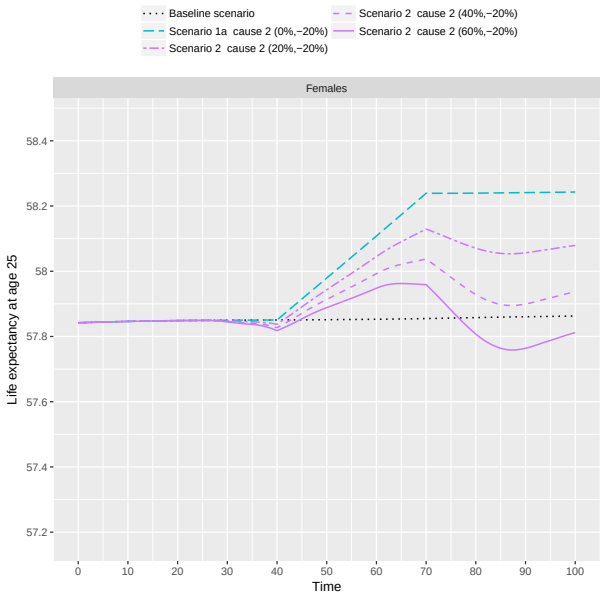
Figure 21: Evolution of female life expectancy at age 65



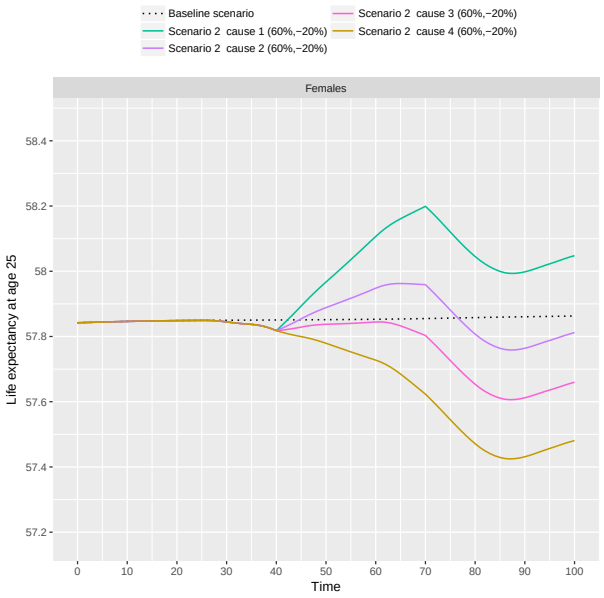
(a) 1981 data (b) 2015 data

Figure 22: Female averaged annual improvement rates

D.2 Cause of death reduction and compensation effect

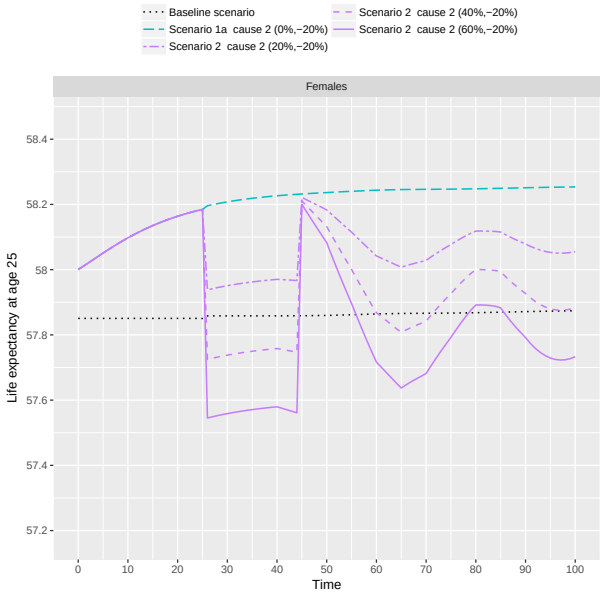


(a) Cardiovascular diseases

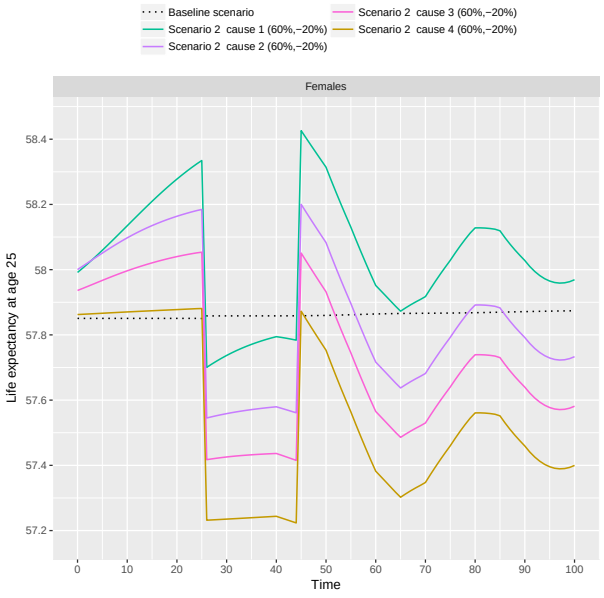


(b) All causes

Figure 23: Aggregated period life expectancy (females): scenario 2



(a) Cardiovascular diseases



(b) All causes

Figure 24: Aggregated cohort life expectancy (females): scenario 2