TWO-DIMENSIONAL MORTALITY DATA: PATTERNS AND PROJECTIONS

By S. J. Richards, J. R. Ellam, J. Hubbard, J. L. C. Lu, S. J. Makin and K. A. Miller

[Presented to the Faculty of Actuaries, 19 March 2007]

ABSTRACT

Patterns and trends in late-life mortality are of growing financial importance. The growth in pension liabilities, both public and private, are of crucial interest to governments, insurers and companies with defined-benefit pension schemes. This paper explores the patterns in international mortality data, and draws important lessons for actuaries in the United Kingdom.

KEYWORDS

Mortality Improvements; Cohort Effects; Period Effects; Penalised Splines; GLMs; Survival Models; Cause of Death

CONTACT ADDRESS

Stephen Richards, 4 Caledonian Place, Edinburgh, EH11 4AS, U.K.
Email: stephen@richardsconsulting.co.uk
Web: www.richardsconsulting.co.uk; Tel: +44(0)131 315 4470

1. Introduction

1.1 This paper is the result of work by the Faculty of Actuaries Mortality Research Group. The group members are Stephen Richards (chair), John Ellam, Jennifer Hubbard, Joseph Lu, Stephen Makin and Keith Miller.

1.2 Mortality patterns in the United Kingdom have been increasingly studied by actuaries of late. The reasons are simple. Longevity risks represent a growing share of the risk profile of many insurance companies as their books mature and business mixes evolve. Increased focus on the size of pension-scheme deficits means that pensions actuaries also need to know more about the possible trends in future life span. Much attention in the U.K. has focused recently on the so-called 'cohort effect', with the generations born between 1925 and 1940 experiencing unprecedented rapid falls in mortality rates at older ages when compared with previous generations.

1.3 Chief amongst the questions asked by actuaries and those whom they advise are "Is the cohort effect unique to the U.K.?'' and "What further improvements can be expected?''. Using the techniques most recently developed by the Continuous Mortality Investigation Bureau (CMIB) in the U.K., this paper will compare and contrast the mortality patterns of various countries to gain insight into these questions. We will look at the question of whether cohort effects or period effects dominate different countries' mortality patterns, and we will demonstrate that even where one effect is dominant it does not mean that the other is not also significant. Finally, we will look at models which actuaries can apply to their portfolio experience data to separate time trends from cohort-based patterns.

2. Format and Source of Mortality Data

2.1 The data used in this paper are the number of deaths aged x last birthday during each calendar year y, split by gender for seven national adult populations. Corresponding mid-year population estimates are also given. The data therefore lend themselves to modelling the force of mortality, $\mu_{x+\frac{1}{2},y+\frac{1}{2}}$, without further adjustment.

Table 1. Summary details from Human Mortality Database (2007) at mortality.org. Selected countries in order of decreasing female life expectancy at age 65.

Country	Population type	Start year	End year	Population aged 40–100	mortality.org last modified
Japan	Total	1947	2004	67.6m	30-Aug-2006
France	Civilian	1899	2004	29.0m	07-Dec-2006
Canada	Total	1921	2003	14.7m	01-Nov-2006
Sweden	Total	1751	2005	4.5m	24-Oct-2006
Germany (West)	Total	1956	2002	34.2m	30-Aug-2006
U.S.A.	Total	1959	2003	127.6m	09-Nov-2006
England and Wales	Civilian	1841	2003	25.1m	29-Aug-2006

Population exposure is for the middle of the appropriate end year, and includes males and females between ages 40 and 100.

2.2 The data were taken from the Human Mortality Database (2007), and summary details for the seven countries selected are shown in Table 1. The data used are the age-period files, rather than the age-cohort files, i.e. we are using the deaths and exposures classified by age and gender for each calendar year, as described in 2.1. The data were then reformatted into a two-dimensional matrix by age and calendar year. These reformatted data files are available at http://www.richardsconsulting.co.uk/Faculty and are in a format suitable for direct use in the software accompanying CMIB (2005a).

2.3 Although the seven countries chosen have many similarities (all are G8 members apart from Sweden), they were chosen for their differences. For example, the United States of America were reputed not to have cohort effects, or at least to have very weak cohort effects compared with other countries. Canada was chosen to contrast with the U.S.A. in the North American continent, not least because it has a very different approach to healthcare. England and Wales were included as territories of particular interest to U.K. actuaries, of course, but also because of the strength of the well-documented cohort patterns in generational mortality — see Richards, Kirkby & Currie (2006). Japan was included both as having particularly long-lived citizens, but also because of a pronounced cohort effect — see Willets (2004). Sweden was chosen because of its long history of high-quality mortality data and its consequent frequent use in other mortality and demographic studies. France was selected as a nation with strong differences in male and female mortality, and Germany was chosen due to its starkly different social and economic history over the 20th century.

Table 2. Life expectancy at age 65. Selected countries in order of decreasing female life expectancy.

		Life expectancy at 65:		
Country	Year	(i) Males	(ii) Females	(ii) - (i)
Japan	2004	18.24	23.27	5.03
France	2004	17.69	22.12	4.43
Canada	2003	17.35	20.69	3.34
Sweden	2005	17.36	20.58	3.22
Germany (West)	2002	16.19	19.75	3.56
U.S.A.	2003	16.78	19.59	2.81
England and Wales	2003	16.41	19.23	2.82

Source: Human Mortality Database (2007) at mortality.org.

2.4 The available data was truncated above age 100, partly to avoid any peculiarities associated with low numbers of deaths at higher ages, but mainly because some countries do not collect

individual data at higher ages and so-called 'data' is sometimes created from a curve-fitting exercise. In order to focus on adult mortality, the available data was truncated below age 40. Using the most recently available year shown in Table 1, the life-expectancy figures from mortality.org are as given in Table 2. For England and Wales, whilst deaths data is generally available by single year of age above age 89, official population estimates are not. This is discussed in Annex G of GAD (2001). The Human Mortality Database (HMD) uses a version of the Kannisto-Thatcher method for estimating populations by single year of age at those ages where national statistics offices provide aggregated data (or even no data). This is similar to the method used to produce these for the U.K. and constituent countries; however, it does utilise the deaths data since this is generally available by single year of age. For consistency across countries, we have used data up to age 100, despite England and Wales not having real data at ages 90–100.

2.5 Although all countries listed in Table 2 belong to the so-called 'developed world', there are clearly wide disparities in life expectancy at retirement. Of particular interest to U.K. actuaries is that life expectancy in retirement in England and Wales is bottom of the table for females, and second-bottom for males. The subject of mortality improvements has been very topical of late, and a glance at Table 2 suggests that there is still plenty of room for future increases in life expectancy: U.K. males live up to two years less than their Japanese counterparts, while U.K. females live around four years less.

2.6 Also of interest is the fact that, despite rapid recent male mortality improvements in the likes of the U.K., females still have a clear lead in life expectancy, both at birth (not shown) and at retirement. Many sources of heterogeneity remain, however, some of which could be as significant as gender. For example, Doll *et al* (2004) estimated the number of years of lifetime lost by a male smoker to be around ten years. Similarly, in graduating the mortality of assured lives, CMIB (2006) found that classification by smoker status had a larger influence on mortality rates than gender.

3. Canada in the 20th Century

3.1 Canada is the second-largest country in the world by area, but it is also a heavily urbanised society, particularly when compared with the U.S.A. Approximately 80% of the population lives within 200km of the border with the U.S.A. The main social factor affecting the Canadian population during the 20th century was immigration. Between 1945 and 1990 immigration was about 140,000 people a year rising in the 1990s to about 230,000 a year, representing some 0.5% to 1% of the population p.a. The proportion of the population which is foreign-born has varied between 22% at the start of the 20th century to 16% at the end. The country of origin of immigration has also changed. Up until 1960 or so Canada encouraged immigration of people from similar cultures who would fit in easily. This led to a preponderance of immigrants from the U.K., the U.S.A. and northern Europe. The policy was changed in the 1960s to concentrate more on occupational skills and education levels, and, as a result, the levels of immigration from some ethnic groups increased, particularly those from Asia. The nature of this switch can be seen by comparing immigration statistics: in 1966, almost 75% of immigrants were from Europe, including 32% from the U.K., whereas by 1996 only 17% were from Europe and 63% were from Asia. Despite this recent influx, only 6% of the population is currently of Asian origin (Canada Census 2001).

3.2 As a result of these strong patterns of immigration, it is perhaps unsurprising that mortality improvement rates do not exhibit cohort properties: the experience at any time is more affected by the influx of new citizens from different backgrounds than by the change in living standards of existing citizens. Alongside the large-scale immigration, Canada has strong welfare and healthcare systems. During the first half of the 20th century, the approach to welfare moved from a private and charity-led moral one (look after only the deserving poor) to a public and government-funded all-inclusive one. The scope of the service gradually increased until by the mid-1970s it became a fully comprehensive welfare system.

3.3 The government department Health Canada is committed to improving the lives of all Canada's people and to making the country's population among the healthiest in the world as

measured by longevity, lifestyle and effective use of the public health care system. The level of success in this commitment can be seen in their life-expectancy figures, particularly in contrast with their larger, richer neighbour to the immediate south: over one extra year of life expectancy at age 65 for females, and over half a year for males. This interventionist approach can be seen particularly in their approach to smoking. Surveys of smoking habits have been conducted since 1966, leading to the Tobacco Act (1997) and the Tobacco Control Program (TCP) which started in 1999. Smoking prevalence has reduced significantly over the last 50 years for males and less dramatically over the last 40 years for females. During the late 1950s about 75% of adult males were smokers and during the late 1960s about 32% of adult females were smokers. By 2006 this had reduced to about 20% for males and 16% for females. Furthermore, the average number of cigarettes smoked has reduced, as has their typical strength. The success of the TCP can be seen in the overall reduction in smoking prevalence from about 25% of adults in 1999 to about 18% of adults in 2006.

4. England and Wales in the 20th Century

4.1 In the first half of the 20th century, most of the population of England and Wales could not afford proper healthcare. Hospitals charged patients, although free service was provided to some low-paid workers. Some charitable organisations and local governments provided limited free medical treatment to the public. In 1948, the National Health Service (NHS) was established to provide healthcare to all U.K. residents based on need and not the ability to pay. One early innovation was the setting up of community health centres providing facilities for family doctors and other healthcare professionals to serve the public. Family doctors would refer patients to hospitals for further treatment where necessary. Since its establishment, the NHS has provided the U.K. population with services including medicine, surgery, diagnosis of diseases and advice on prevention.

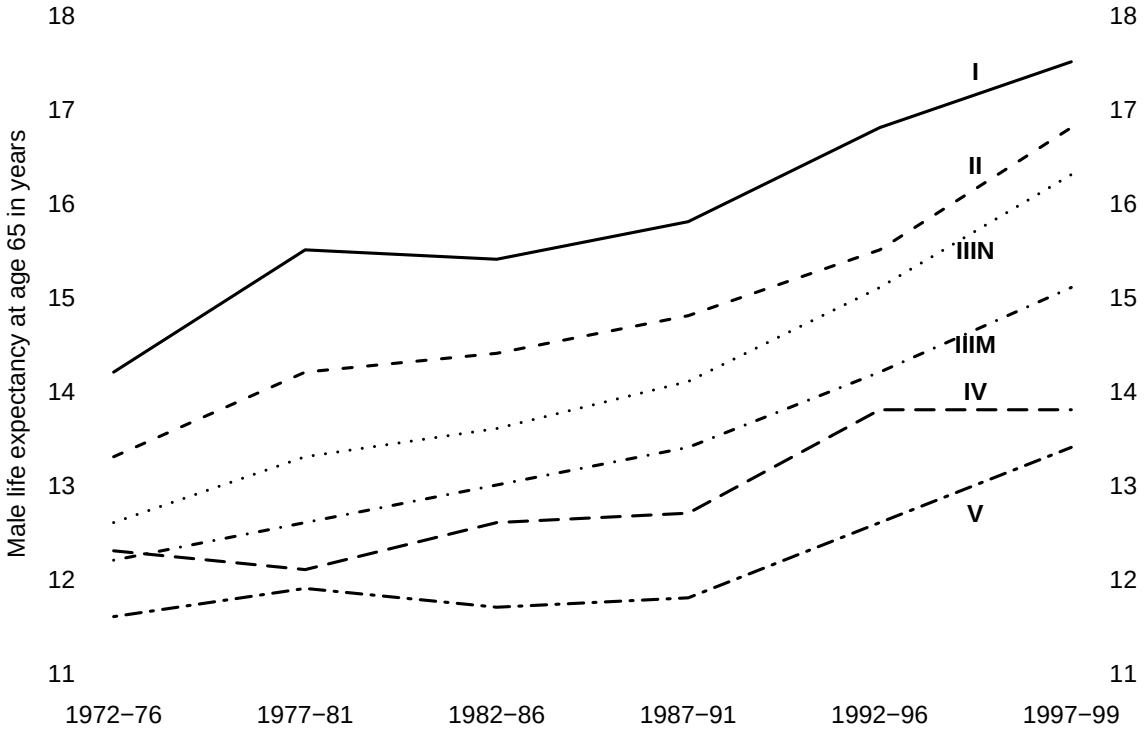

Figure 1. Life expectancy at age 65 for males in England and Wales by socio-economic group. Socio-economic group is defined by occupation type as follows: professional (I), managerial (II), skilled non-manual (IIIN), skilled manual (IIIM), semi-skilled manual (IV) and unskilled manual (V). Source: ONS Longitudinal survey.

4.2 Willets *et al* (2004) discuss the mortality trends in the U.K. in detail and this will not be repeated here. However, some key phenomena underlying the improvements in mortality at retirement age include (i) cohort effects: people born between 1925 and 1940 have experienced faster improvements in mortality than generations born before or after them; (ii) the 'ageing of mortality improvement': the acceleration of mortality improvement at older ages; (iii) a reduction in prevalence of cigarette smoking: whereas 44% of males above 60 smoked in 1974, the corresponding figure in 2000 was only 16%; and (iv) widening social class differentials: people in higher social classes have enjoyed a greater increase in life expectancies than their counterparts in lower social classes, as shown in Figure 1.

4.3 The First and Second World Wars had a major impact on later mortality experience in England and Wales, not least because servicemen and women were given free cigarettes. In the U.K., the highest smoking prevalence was a peak of about 75% for those born in the 1920s, and it has been reducing ever since.

5. France in the 20th Century

5.1 Over the course of the 20th century France went from being the fifth-most populous country in Europe behind Russia, Germany, U.K. and Italy, to being the third-most populous country, leapfrogging both the U.K. and Italy. Its demographic growth occurred principally between 1945 and 1974, known as '*les trente glorieuses*'. During this period, France was in a period of reconstruction with steady economic growth and labour immigration in the 1960s. The 1973 energy crisis prompted a change to the immigration law, thereafter limited to those rejoining family members already domiciled in France. France currently has a fertility rate second only to Ireland in Europe, and has an increasing population, due mainly to organic growth.

5.2 The data set used in this paper was the civilian population, and thus excluded military deaths. However, the civilian population was much affected by both the First and Second World Wars.

5.3 Influenza vaccination for older people was introduced in 1977, which helped stabilise variations in mortality at older ages. In particular, from 1985 the vaccination was made free for 'people at risk' which includes all people aged 65 and over. In the recent past, the heat wave of 2003 resulted in 15,000 extra deaths, and was followed in 2004 by a year with 41,500 fewer deaths. While some commentators would view this as evidence of a frailty effect — the frail in 2003 were killed off, leaving a less-frail population for 2004 — the sharp reduction in 2004 is also partly due to a lack of influenza virus in 2004.

5.4 In recent years, road-safety campaigns have contributed to large falls in the number of deaths from motor vehicle accidents. Motor vehicle accidents affect largely the age group 20–40, and, as overall death rates are otherwise low in this age band, an improvement of 36% in deaths from road accidents was observed between 2001 and 2005.

5.5 The prevalence of smoking in France has, like in most countries, reduced since the 1950s. For males the reduction has been from about 77% of population over 15 in 1951–55 to about 30% in 2002–05. The corresponding figures for females are 35% reducing to 21%. These results hide other features: in the case of males, the prevalence by age is relatively constant up to about age 60 and has reduced, with the exception of a period in the 1960s, every five years. The corresponding picture for females is that peak consumption is between ages 20–24 and then reduces by between 15% and 20% for every five-year band. Over the past 50 years there has been little variation in prevalence at the younger ages (under 35) with, if anything, an increase at the end of the last century. The WHO data would also suggest that there has been little or no change in female prevalence over the last decade. France is the latest country to introduce a smoking ban in workplaces (February 2007) and it has widespread public support even amongst smokers, but this will not be fully operational in restaurants and bars until 2008.

6. Germany in the 20th Century

6.1 Germany has had a welfare state for over a century, including a generous health-insurance system and a high level of state pension benefits. These date back to before the 20th century: Bismarck introduced a health-insurance law in 1883, followed by an accident-insurance law in 1884 and a law enacting pension and invalidity benefits in 1889.

6.2 German social history in the first half of the 20th century is one of extremes. The British naval blockade during the First World War meant that the German civil population suffered from food shortages: as early as February 1915, rationing was introduced for bread and wheat flour, followed later that year by milk and meat. After the war's end in 1919, the German civil population experienced further suffering in the depression following the 1929 Wall Street crash, with record unemployment. Hard circumstances returned once again during the Second World War, this time with mass civilian deaths, forced migrations and food shortages. Food rationing was introduced from the start of the Second World War in 1939, and some food products were still being rationed in the Soviet-occupied zone as late as 1951.

6.3 Germany also has a unique history regarding cigarette consumption. The earliest research providing a clear link between smoking and lung cancer was conducted in Germany, preceding the results in Doll *et al* (1954) by at least a decade: work by Schairer & Schöniger (1943) at the Institute for Research into Tobacco Risks at the University of Jena provided conclusive links, as did even earlier work by Lickint (1929). Partly as a result of this, Germany had some of the earliest government anti-smoking campaigns prior to and during the Second World War, including a ban on smoking in universities, post offices and military hospitals. However, this anti-smoking drive was a Nazi measure: a review of early German research into smoking and mortality — and its unfortunate sidelining because of Nazi involvement — is described by Proctor (2001). Perhaps as a reaction to this, post-war Germany has noticeably lagged other countries such as Italy, France, Ireland and Scotland in the banning of smoking in enclosed public spaces. Culturally, modern Germany has quite a different attitude towards smoking from the U.K.: non-smoking sections of restaurants and cafes are largely unheard of and, until January 2007, anyone with a few euro coins — including, doubtless, under-age children — could buy cigarettes from vending machines on most street corners. Germany's smoking prevalence — 37% for males and 30% for females — is relatively high for the countries surveyed here.

6.4 Germany has been subject to a considerable degree of turmoil during the 20th century, and the territory has changed considerably during this time. Following the Second World War, Germany was both split in two and also lost several territories to neighbouring countries, including East Prussia and Königsberg to the Soviet Union, and Danzig and Silesia to Poland. As a result there were large migrations to the newly reduced German territories of the Federal Republic (West Germany) and the GDR (East Germany). These migrations also included several million ethnic Germans from outside Germany, for example around 3m from the Sudetenland. Aggregated population data for the re-unified Germany is only available from 1991, and for the purposes of this paper it was decided to focus on the more stable Western population for continuity. Note, however, that the last census in the former territory of the Federal Republic of Germany was in 1987, a full 20 years prior to the publication of this paper. Since that census, the people, the society, and the economy in Germany have undergone historic changes: the Berlin Wall fell, unleashing significant internal migration from East to West. In addition to this, Germany has also had significant north-south migration, as well as hundreds of thousands of both immigrants and emigrants since the last census. Immigrants include not only asylum seekers and refugees, but also many immigrants of ethnic German origin. The low birth rate and a rising life expectancy have also changed the age structure of the population. The Federal Statistical Office itself describes the official population figures as 'probably much too high' and any findings in this paper need to bear this in mind.

7. Japan in the 20th Century

7.1 Japan is currently the country with the highest life expectancy in the world, both for males and females. Japan's current position is even more astonishing when one considers both that its life expectancy at birth in 1951 was well below the level enjoyed by most modern nations and that Japan's life expectancy continues to grow faster than that of other countries with high life expectancies. A major portion of the Japanese mortality decline over the last 100 years has been a matter of catching up with the low-mortality countries. During this period, life expectancy at birth has more than doubled. Even after 1950 it increased by more than 15 years for males and by nearly 20 years for females.

7.2 In the early 20th century infectious and parasitic diseases were the major causes of death and responsible for the high infant and childhood mortality. These causes continued to be important even after the Second World War: the decline in mortality between 1955 and 1960 due to infectious diseases made an important contribution to the increase in life expectancy over that period: 46% and 38% respectively of the increase in male and female life expectancy. Vigorous public health activities to control infectious diseases and the use of antibiotics played a major role in the control of these mortality causes. By the mid-1960s, infant and childhood mortality had declined so much that further increase in life expectancy from this source had become negligible.

7.3 The rate of increase in life expectancy decreased after 1965. However, while gradually approaching and then surpassing Western standards, a new pattern of age-specific mortality decline emerged. The greatest contribution to increased longevity has shifted from the younger to older age groups. After 1965 increases in life expectancy became predominantly the consequence of mortality trends at the middle and older ages. For the 1985-1990 period declines in childhood mortality were responsible for improvements of only 7% and 6% of male and female life expectancy. In contrast, over the same period declines in mortality above age 65 contributed respectively 50% and 69% of the improvement in life expectancy.

7.4 As the age pattern of mortality changed there was a remarkable change in the causes of death. The role of non-infectious disease has been gaining in importance in determining the length of life. In 1990 malignant neoplasms, followed by heart disease and cerebro-vascular disease, were the most important causes of death. An interesting case is the role of cerebro-vascular disease (stroke), which was traditionally considerably higher in Japan than in European countries except for Italy. It declined rapidly and substantially in the later years of the century, and this decline is one of the most important features in the evolution of mortality in post-war Japan.

7.5 The reason why cerebro-vascular disease in Japan was higher than in Europe and is now declining as a cause of mortality is that salty food was an important ingredient of the traditional Japanese diet. With increasing standards of living, diet shifted towards less salty and more protein-rich food. The introduction of the refrigerator to the Japanese household lessened the use of salt for food preservation, and has enabled the Japanese to eat more fresh food.

7.6 Further improvements in life span seem to be possible. Such a conclusion is warranted on the basis of several observations. One is seen on the island of Okinawa, which has the highest life expectancy in Japan. Although Okinawa has the lowest per-capita income among Japanese prefectures, female life expectancy is currently more than two-and-a-half years above that of other Japanese women. For males there is a smaller difference in the same direction. Differences in mortality from heart disease, cerebro-vascular disease and malignant neoplasms constitute 93% of the difference. Differences in dietary habits, the consequence of a strong Chinese influence, seem to be the most important factor in explaining the more favourable survival prospects of the inhabitants of Okinawa.

7.7 Another consideration confirming further possibilities for increase in life expectancy comes from the fact that mortality at older ages continues to decline at a steeper rate than in Western countries. Even though Japan's economy has grown rapidly from the 1960s and the variety of foodstuffs available has expanded greatly, the average daily food intake has scarcely risen from 2,000 calories a day. This pattern is very different from America and Europe, where increased

standards of living have led to growing obesity.

Table 3. Smoking prevalence by gender. Selected countries in order of decreasing female life expectancy in Table 2.

Country	Males	Females	Period
Japan	47	12	2000
France	30	21	2002–2005
Canada	24	20	2001
Sweden	14	19	2002–2005
Germany (West)	37	30	2002-2005
U.S.A.	24	18	2005
U.K.	25	23	2005

Source: CDC (2006), ONS (2006), WHO (2007).

7.8 As Table 3 shows, Japanese males have by far the highest smoking prevalence rate of any country considered in this paper, and their lung-cancer death rates are continuing to rise whereas for most other countries they are falling. Despite this, their expectations of life have continued to rise more quickly. In Japan, the highest prevalence rates for males are in post-war generations, with over 80% for all generations born up to 1970. The consequence of this smoking is possibly not yet feeding through to expectations of life at age 65. The large difference between male and female smoking prevalence in Japan will partly explain the relatively large difference in male and female life expectancies in Table 2.

7.9 As with the U.K., Japan appears to have a cohort effect, and this is considered in detail in Willets (2004).

8. Sweden in the 20th Century

8.1 Sweden's 20th century political history is characterised by neutrality. When the First World War broke out in 1914, Sweden declared itself neutral and remained so during the war. When neighbouring Denmark and Norway were occupied by Germany during the Second World War, many thought that Sweden would have no choice but to enter the war. However, Sweden was never attacked by Germany, largely because of the assistance it provided in the form of metal ores from its northern mines and in the form of troop transportation to Norway and Finland using its rail infrastructure.

8.2 Sweden's 20th century social history is characterised by near-socialism. Although Sweden has always had a solid market-driven economy, the social democratic governments, in power for the majority of the 20th century, borrowed many ideas from socialism, including a greater degree of wealth redistribution than any of the other countries surveyed here.

8.3 As early as 1913, Sweden's government broadened the range of social benefits, and in that year passed a national Pensions Act to provide security for the elderly. In 1918, a liberal coalition government passed a poor law, passing responsibility for helping the needy to local governments, albeit with central-government administrative support. This law remained the main pillar of Sweden's welfare programme for much of the inter-war period. In 1932, the social democratic government instigated a new policy providing the unemployed with state-funded employment, albeit at the expense of heavy increases in taxation.

8.4 When the Second World War ended, Sweden was largely unscathed and well-placed, with its industry intact, to produce the goods that the numerous devastated European countries needed. This led to an economic boom, making possible the rapid development of the Swedish welfare-state during the 1950s and 1960s. In this period, Sweden adopted comprehensive welfare legislation, covering — at the expense of the world's heaviest tax burden — schools, child care, health care,

social services, and pensions. This welfare state — often referred to as 'the Swedish Model' — has since been emulated in other countries.

8.5 In recent decades, from its peak at the time of the 1973 oil crisis, the Swedish welfare state has been under increasing (and often severe) pressure, yet the main features remain intact. Today, official unemployment in Sweden runs at around 5%, although estimates of the 'true' levels are closer to around 25%, were it not for the significant level of state-provided employment which still exists to this day.

8.6 While cigarette smoking has been the major form of tobacco consumption in virtually all developed countries over the 20th century, Sweden is a very notable exception to this. Whilst smoking prevalence amongst Swedish females is broadly consistent with those of other European countries, and indeed tobacco consumption in Sweden is as high as that of other Western countries, smoking rates amongst Swedish males have always been lower than those of their European counterparts. Instead of smoking, many Swedish men use moist snuff called 'snus' (pronounced 'snoos') which is placed in the mouth rather then being smoked. Despite studies which show that use of snus leads to no material increase in mortality rates (either from cardiovascular diseases, pulmonary diseases, or oral or other cancers), snus is prohibited in every other country in the European Union, where rates of smoking — and smoking-related deaths — are far higher. More details on Swedish tobacco consumption can be found in Rodu (2004) and Rodu & Cole (2004).

8.7 The World Wars led to a sharp increase in smoking prevalence rates, particularly amongst the countries directly affected, reaching 50% in many such countries in the 1950s and 1960s. Whilst Sweden was not entirely immune to the effects of the World Wars on smoking, Swedish men smoked many fewer cigarettes (with a preference for snus) than any other comparable society: it is believed that Sweden's neutrality in both World Wars was a key factor in this.

8.8 Other cultural influences may also have contributed to Sweden's pattern of tobacco usage. It is not a coincidence that a decline in snus consumption ended in the late 1960s, just as cigarette consumption started to reach a plateau, driven by both health concerns (the adverse health effects of smoking were becoming widely-known and accepted by the mid-1960s) and other cultural influences (e.g. the consumption of snus by Swedish football and ice hockey players). Uniquely amongst the seven countries surveyed here, males in Sweden have a lower smoking prevalence rate than females (Table 3). Cigarette consumption in Sweden has fallen sharply since 1990, a trend driven largely by a dramatic decline in male smoking prevalence. In 1999 Sweden became the first — and it remains the only — country to meet the WHO's target for a combined adult smoking prevalence of below 20%.

9. The U.S.A. in the 20th Century

9.1 The U.S.A. is not only the most populous of the surveyed countries, but it is also the most populous of the industrialised countries in the world. In common with the other countries surveyed, infant mortality and mortality rates at ages below 65 have declined dramatically over the last century. Currently, three-quarters of all deaths occur at ages 65 and above, contrasting with the position in 1900 where just less than one-fifth of all deaths were at ages 65 and older.

9.2 The average annual rate of mortality improvement for males below age 65 has been fairly consistent over the 20th century, at around 1.6% p.a. for the century as a whole (1.4% p.a. for the latter half of it, and 1.6% p.a. for the final two decades). The average annual rate of decline in mortality for men at 65 and older was also fairly consistent over the 20th century, the average rate of improvement being almost 0.6% p.a. for the whole century and about 0.7% p.a. over both the last 50 and 20 years (ERA, 2003).

9.3 For females, the position is slightly different. The annual rate of improvement in female mortality at ages below 65 was around 2.0% p.a. over the 20th century (around 1.5% p.a. over the latter half, and around 1.0% p.a. over the final 20 years). The annual rate of improvement in female mortality at age 65 and above was around 0.9% p.a. over the 20th century, and nearly 1.0% p.a. over the latter half. Over the final 20 years, however, improvement rates slowed to around 0.5% p.a. (ERA, 2003).

Table 4. Annual rates of mortality improvement assumed by the Social Security Administration Program.

	Aged under 65		**Aged over 65**	
Period	**Males**	**Females**	**Males**	**Females**
2000–2027	c1.1%	c1.0%	c0.7%	c0.7%
2027–2077	c0.9%	c0.8%	c0.7%	c0.7%

9.4 The future expectations of the Social Security Administration Program for rates of mortality improvement are very similar for males and for females, particularly at ages 65 and above, as shown in Table 4. The accomplishments of the 20th century, with the pure positive effects of improved sanitation, nutrition, and medical accessibility, will not be easy to match. For example, the obesity rate among adults in the U.S.A. is the highest of all OECD countries: 30.6% in 2002 (OECD, 2005), followed by Mexico (24.2% in 2000) and the U.K. (23% in 2003). Increased obesity rates are likely to lead to a rise in onset of related chronic diseases, such as diabetes and asthma. Nevertheless, a combination of the 'baby boom' and continued expectations of increased longevity mean that the Social Security Administration estimates that the number of beneficiaries is likely to more than double by 2050 (ERA, 2003).

9.5 Following the Second World War, there was a dramatic increase in fertility rates. Rates began to soar in 1946 and, although they peaked in 1957, their effect on annual birth rates persisted until 1964. The total fertility rate (the average number of children that women have) was about 3.3 children per woman during the baby-boom years of 1946–1965 (Goss, 2005). By 1972, however, the total fertility rate had dropped to around 2 children per woman and has stayed at about that level since then. The current assumption for the future is for an average total fertility rate of just less than 2 (Goss, 2005).

9.6 The trend toward an ever-widening gap in life expectancy between men and women ended in around 1980. The widening of the gap during most of the 20th century can be attributed, at least in part, to the fact that men smoked more than women. In recent decades, however, the prevalence of smoking amongst men has decreased more significantly than for women, to the extent that smoker prevalence rates are broadly similar. The proportion of smokers among adults has fallen from 33.5% in 1980 to 17.5% in 2003 (OECD, 2005), the lowest rate among OECD countries along with Canada and Sweden.

9.7 The U.S.A. is unique among the nations surveyed here in that it has the highest health spending as a proportion of GDP (15% in 2003). The equivalent figure for Germany is 11.5%, and Canada and France each spend around 10% of GDP on health (OECD figures, 2005). The U.S.A. also spends more per capita: \$5,635 (adjusted for purchasing power parity), compared to the OECD average of just \$2,307 in 2003. The U.S.A. is also unique in the proportion of health spending funded by government: just 44% compared to the OECD average of 72%. Despite all the extra spending, life expectancies in the U.S.A. are markedly lower than other nations which spend a lot less.

10. Penalised-Spline Regression

10.1 Penalised-spline regression was chosen to fit models to the past data. This P-spline regression is well documented in non-actuarial literature — see Currie, Durban & Eilers (2003, 2004a, 2004b). P-splines have recently found application within actuarial work — see CMIB (2005a) and Richards, Kirkby & Currie (2006) — and the content of this section is drawn from the latter paper.

10.2 The P-spline methods and models used in this paper were those described in CMIB (2005a). The software used was that made publicly available by the CMI from the profession's website (www.actuaries.org.uk) and is based on R (2004), which is also freely available (www.r-project.org).

10.3 The basis for a P-spline model is to let $D_{x,t}$ be the random variable denoting the number of deaths at age x and year t. Note that we can use t for either year of birth or year of observation,

and thus smooth over either. A model smoothing over age and year of birth is referred to here as an age-cohort model, while a model smoothing over age and year of observation is an age-period model.

10.4 We assume that $D_{x,t}$ has a Poisson distribution with mean $E^c_{x,t}\mu_{x,t}$ where $E^c_{x,t}$ is the central exposed to risk and $\mu_{x,t}$ is the force of mortality. If $d_{x,t}$ is the observed value of $D_{x,t}$ then the raw force of mortality is $\hat{\mu}_{x,t} = d_{x,t}/E^c_{x,t}$. We must smooth the $\hat{\mu}_{x,t}$. Once smooth values of $\hat{\mu}_{x,t}$ are obtained we can, if necessary, convert to smooth values of the mortality probability, $\hat{q}_{x,t}$, and hence to smooth mortality improvement rates.

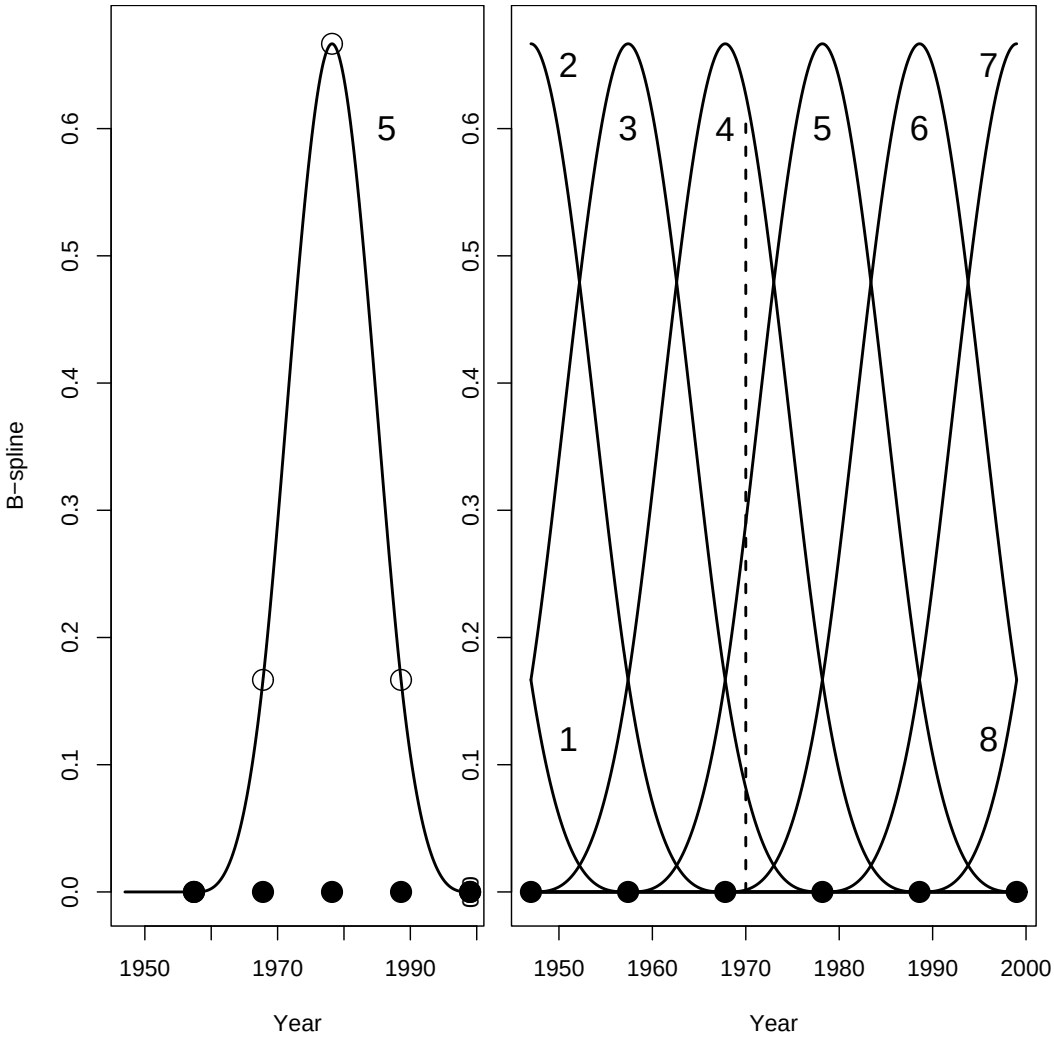

Figure 2. Left panel: a single cubic B-spline, $B_5(t)$, with knot positions, $\bullet$, and smooth joints, $\circ$; right panel: a basis of $K = 8$ cubic B-splines with knot positions. A set of equally spaced knots is used for all splines in the basis, with each B-spline being zero-valued except on a finite interval

10.5 We illustrate smoothing with penalised B-splines (shorthand P-spline) by considering some data on male assured lives at age 65; these data were supplied by the CMIB for years of observation from 1947 to 1999. It is convenient to model the force of mortality on the log scale, since the ratio of mortality at older ages to mortality at younger ages is very large. Our model is:

$$D_t \sim \text{Poisson}(E^c_t \mu_t), \quad \log \mu_t = \sum_{k=1}^{K} B_k(t)\theta_k \tag{1}$$

where the θ_k are the regression coefficients whose value we want to find and the $B_k(t)$ are splines (reference to age, x, has been omitted for ease of presentation). This is a regression-type model with the set of B-splines $\{B_1(t), \ldots, B_K(t)\}$ providing the regression basis (instead of the more familiar

powers of t in a traditional polynomial regression). The regression coefficients are denoted by $\theta_1, \ldots, \theta_K$. The left panel of Figure 2 shows a single cubic B-spline. A cubic B-spline consists of cubic polynomial pieces bolted together at points known as knots; in the diagram the knots are equally spaced in calendar time at 1957.4, 1967.8, 1978.2, 1988.6 and 1999.0 and the B-spline is zero to the left of 1957.4 and to the right of 1999.0. The B-spline pieces are continuous, and have continuous first and second derivatives at the join points, shown $\circ$ in the left panel of Figure 2. The right panel in Figure 2 shows a basis of B-splines with $K = 8$. See de Boor (2001) for some actuarial references to smoothing mortality data with splines.

Figure 3. Smooth assured lives mortality with regression coefficients, $\circ$. Left panel: B-spline regression; Right panel: P-spline regression (B-spline with structure penalty on regression coefficients)

10.6 This model can be fitted with standard software since the Poisson distribution together with the linear structure for $\log \mu_t$ defines a generalised linear model (see McCullagh & Nelder, 1989, and Renshaw, 1991); the regression coefficients θ_k are chosen by maximum likelihood. The left panel in Figure 3 shows the result of fitting the regression with a basis of $K = 23$ B-splines. It is clear that if we had been smoothing by eye then we would not have been satisfied with the fit, although it is likely to have passed any goodness-of-fit test. The problem is that we have too many B-splines in our basis and the resulting fit seems too flexible. In the 1970s and 1980s much effort went into the determination of the optimal number of B-splines, i.e. the number of B-splines that provides an optimal level of smoothing. Eilers & Marx (1996) proposed a different strategy. In Figure 3 the regression coefficients θ_k are plotted at the maximum value of $B_k(t)$. Eilers & Marx observed that the sort of under-smoothing evident in the left panel of Figure 3 (the saw-tooth effect in the $\circ$ plot) was a result of the erratic behaviour of the θ_k and they proposed penalising this erratic behaviour by placing a difference penalty on adjacent θ_k, as in:

$$P(\theta) = (\theta_1 - 2\theta_2 + \theta_3)^2 + \ldots + (\theta_{K-2} - 2\theta_{K-1} + \theta_K)^2 \tag{2}$$

10.7 This defines a quadratic penalty; linear and cubic penalty functions are also possible. The penalty function is incorporated into the log-likelihood function, $L(\theta)$, to give the penalised log-likelihood function, $PL(\theta)$:

$$PL(\theta) = L(\theta) - \frac{1}{2}\lambda P(\theta) \tag{3}$$

10.8 This method is known as penalised B-spline regression or P-splines for short. The parameter λ is the smoothing constant and plays exactly the same role as the bandwidth parameter of kernel smoothing (Richards, Kirkby & Currie, 2006). As with kernel smoothing, the larger the value of λ, the stronger the smoothing. For a given value of λ, the regression coefficients are chosen by maximising $PL(\theta)$. One advantage of our statistical model is that we can use some statistical criterion to select the tuning constant, λ; possibilities include the Akaike Information Criterion (AIC, Akaike, 1987), the Bayesian Information Criterion (BIC, Schwarz, 1978) or generalised cross-validation (GCV) (Craven & Wahba, 1979). These criteria balance (a) the closeness of fit of the observations to the fitted values with (b) the complexity of the fitted model. Just as one maximises a plain likelihood function to find the maximum-likelihood estimates (MLEs) of the parameters, one maximises a penalised likelihood function for a given value of λ to do the same but with smoothed parameters. However, λ is neither known nor part of the model, so a different approach is needed to estimate it as well as the parameters. Our approach is to minimise the BIC (or AIC) to find the best trade-off between model fit and smoothness, while also estimating the model parameters and the value of λ. Indeed, one can go further and use the BIC to not only do this, but also determine the optimum order of the B-splines, their optimum number and even the optimum penalty order. In practice, however, experience shows that one cannot go too far wrong with lots of knots, B-splines of degree 3 (i.e. cubic splines) and second-order penalties. In Equation 3 we have a single value of λ for smoothing in only one direction, whereas for the two-dimensional data sets here we will smooth in two dimensions and thus need two different values of λ. The values can be quite different, as different amounts of smoothing are typically required across ages than are required across time.

10.9 The right panel of Figure 3 shows the result of smoothing with $K = 23$ cubic B-splines in the regression basis but this time a quadratic penalty is used to smooth the regression coefficients; the tuning constant was chosen with the BIC ($\lambda = 3900$). We need less smoothing with fewer B-splines in the basis. For example, with $K = 13$ we find $\lambda = 310$; the resulting fit is indistinguishable from the right panel of Figure 3 and is omitted. Thus, the use of the BIC adjusts automatically for the number of B-splines in the basis by choosing the appropriate level of smoothing.

10.10 The method of P-splines has some similarities with moving averages and kernel smoothing. In view of Figure 3, we can interpret the coefficients θ_k as pseudo-observations; fitted values at year t are weighted averages of these pseudo-observations where the weights are equal to the values of the non-zero B-splines at year t. For example, with the basis in the right panel of Figure 2, the estimate of $\log \mu_{1970}$ is:

$$0.0817 \times \hat{\theta}_3 + 0.6267 \times \hat{\theta}_4 + 0.2901 \times \hat{\theta}_5 + 0.0016 \times \hat{\theta}_6 \tag{4}$$

where these weights are given by the intersection of the dashed line in Figure 2 with the labelled B-splines $B_3(t)$, $B_4(t)$, $B_5(t)$ and $B_6(t)$. Notice that at most four of the weights are non-zero at any year and that the weights in Equation 4 sum to unity. Thus the B-spline weights move across the years from 1947 to 1999 in much the same way as the weights in a moving average or in kernel smoothing.

11. Results from Applying P-Spline Regression to Various National Data

11.1 One way to assess the relative strength of cohort effects is to compare the BIC under age-cohort and age-period models. A model with a lower BIC value is better-fitting, and would indicate the dominance of that model's primary characteristic over that of an alternative model. Note that saying feature A dominates feature B does not mean that feature B is not present, merely that feature A is stronger in explaining observed patterns than feature B. Also, as we shall see, the flexible nature of P-splines is that one can still see cohort effects visually in the results of an age-period model.

11.2 One important misconception to dispel is that low values of λ are somehow invalid, in particular the common assumption among users of the CMIB (2005a) software that a value of 0.001 indicates a failed fitting procedure. As discussed in 10.9, the value of λ is closely linked to the number of splines in the basis, and a small value of λ simply means that less explicit smoothing is required on the values of θ. In this, the role of λ is directly analogous to the tuning constant used in kernel smoothing (Richards, Kirkby & Currie, 2006) and therefore any positive value is valid. Users of the CMIB software should set the minimum value of λ to be much lower (say 0.000001) to avoid any premature termination of the BIC optimisation algorithm. Another useful tip is to change the starting value of λ: sometimes starting from 1.0 will work whereas starting from the default value of 100.0 will not.

Table 5. BIC values for age-cohort (AC) and age-period (AP) P-spline models. Selected countries in order of decreasing female life expectancy in Table 2. A smaller BIC indicates a better-fitting model, i.e. a positive number in the AC-AP column indicates period effects dominate cohort effects, whereas a negative number indicates cohort effects are dominant (marked in bold type).

		BIC:		
Country	**Gender**	**(i) AC**	**(ii) AP**	**(iii) AC - AP**
Japan	Males	11,292.5	11,299.3	**−6.8**
	Females	10,871.3	10,782.2	89.1
France	Males	11,412.6	10,743.4	669.2
	Females	n/a	10,405.8	n/a
Canada	Males	4,118.4	3,977.6	140.8
	Females	4,030.4	3,866.4	164.0
Sweden	Males	3,518.6	3,506.0	12.5
	Females	3,893.0	3,906.7	**−13.7**
Germany (West)	Males	11,944.1	12,107.5	**−163.4**
	Females	12,061.1	12,128.2	**−67.1**
U.S.A.	Males	18,979.6	19,568.5	**−588.9**
	Females	19,776.7	19,830.9	**−54.2**
England and Wales	Males	10,666.9	10,493.4	173.5
	Females	10,623.5	10,290.3	333.2

Source: Own calculations using HMD data over age range 40–100, and using data from 1961 until the end-year shown in Table 1. Five-year knot spacing for each of age, cohort and period. The age-cohort figure for females in France could not be ascertained due to non-covergence of the model.

11.3 Table 5 shows wide variety: cohort effects dominate for males in Japan, Germany and U.S.A., and for females in Sweden, Germany and the U.S.A.; for everyone else, period effects dominate. The key thing to remember about the BIC (or any other information criterion) is that it is the relative difference which counts, not the absolute value: a lower BIC signifies a better-fitting model, balanced against the complexity of that model (as measured by the number of parameters). In this context, all of the differences in column (iii) of Table 5 are statistically significant.

11.4 One point of particular note is the result for males in England and Wales, as this flatly contradicts the result in Richards, Kirkby & Currie (2006), who found a smaller BIC for the age-cohort model and therefore concluded that cohort effects dominated period effects. One possible source of the discrepancy lies in the data set used: Richards, Kirkby & Currie (2006) used a data set for England and Wales which is not publicly available, and which contained population estimates which had been revised during 2004–2005. This paper has used the data from the HMD for consistency across all countries and, given that the HMD data for England and Wales stops at 2003 (see Table 1), there are differences in the population estimates. Also, the data set in Richards, Kirkby & Currie (2006) only goes up to age 89, whereas the data sets used here go up to age 100. As discussed in 2.4, the data above age 89 for England and Wales has been constructed, rather than estimated from census data, and this appears to have a very strong influence on the dominance of cohort effects.

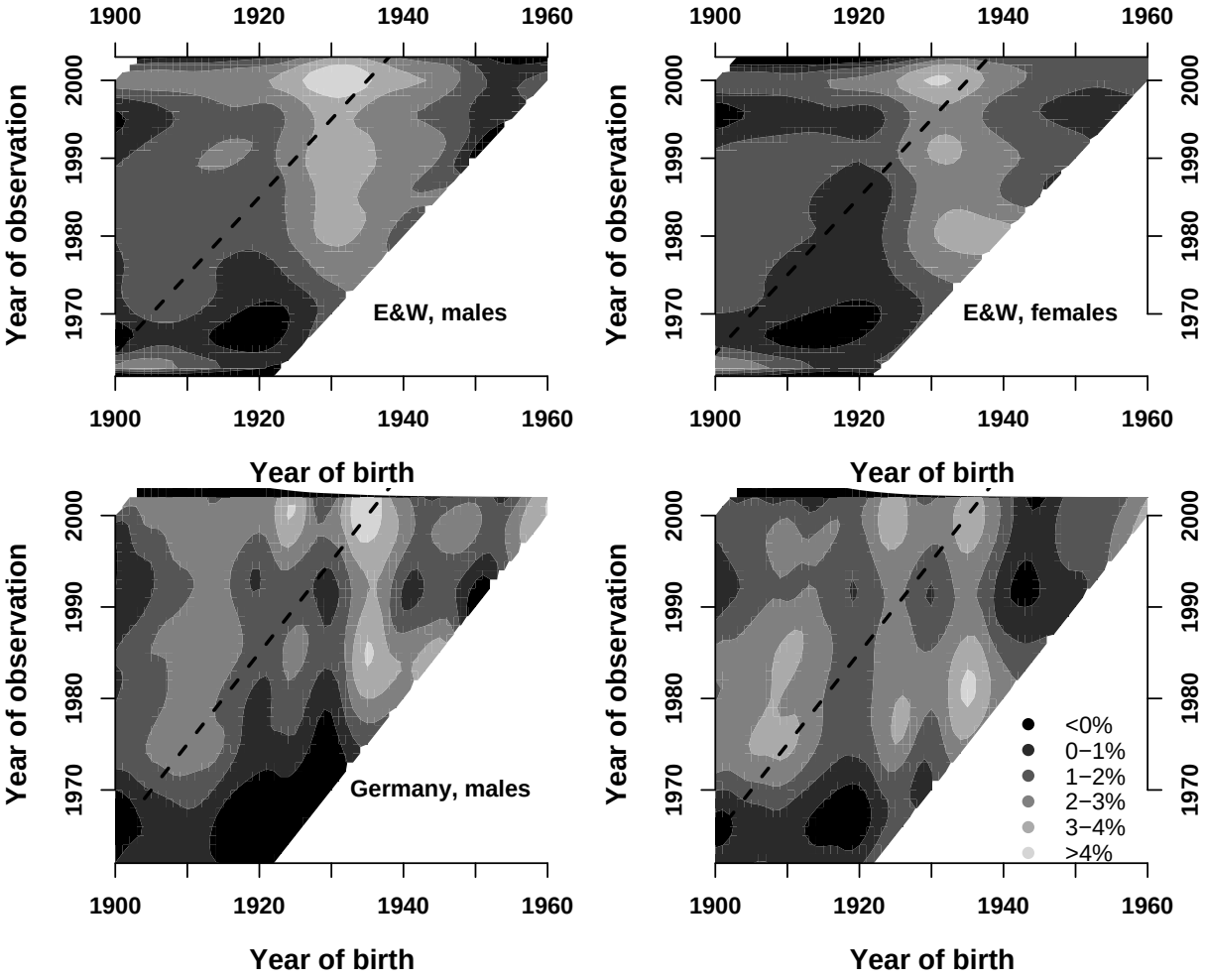

Figure 4. Improvements in smoothed force of mortality: P-spline regression with age-period penalties for England and Wales (top panels) and age-cohort penalties for Germany (bottom panels). Males in left panels, females in right panels. The dashed line shows age 65 for each cohort

11.5 Figure 4 shows the improvements in the force of mortality over the past forty years for Germany (bottom panels) and England and Wales (top panels). Despite using the age-period P-spline model to fit the model to the data, the 1925–1940 cohort effect is clearly visible as a vertical pattern. Also of note is the existence of cohort effects for females, as well as the expected (and well-documented) cohort effect for males. Cohort effects are also visible for both males and females in Germany, although cohort effects appear visually stronger for males than for females.

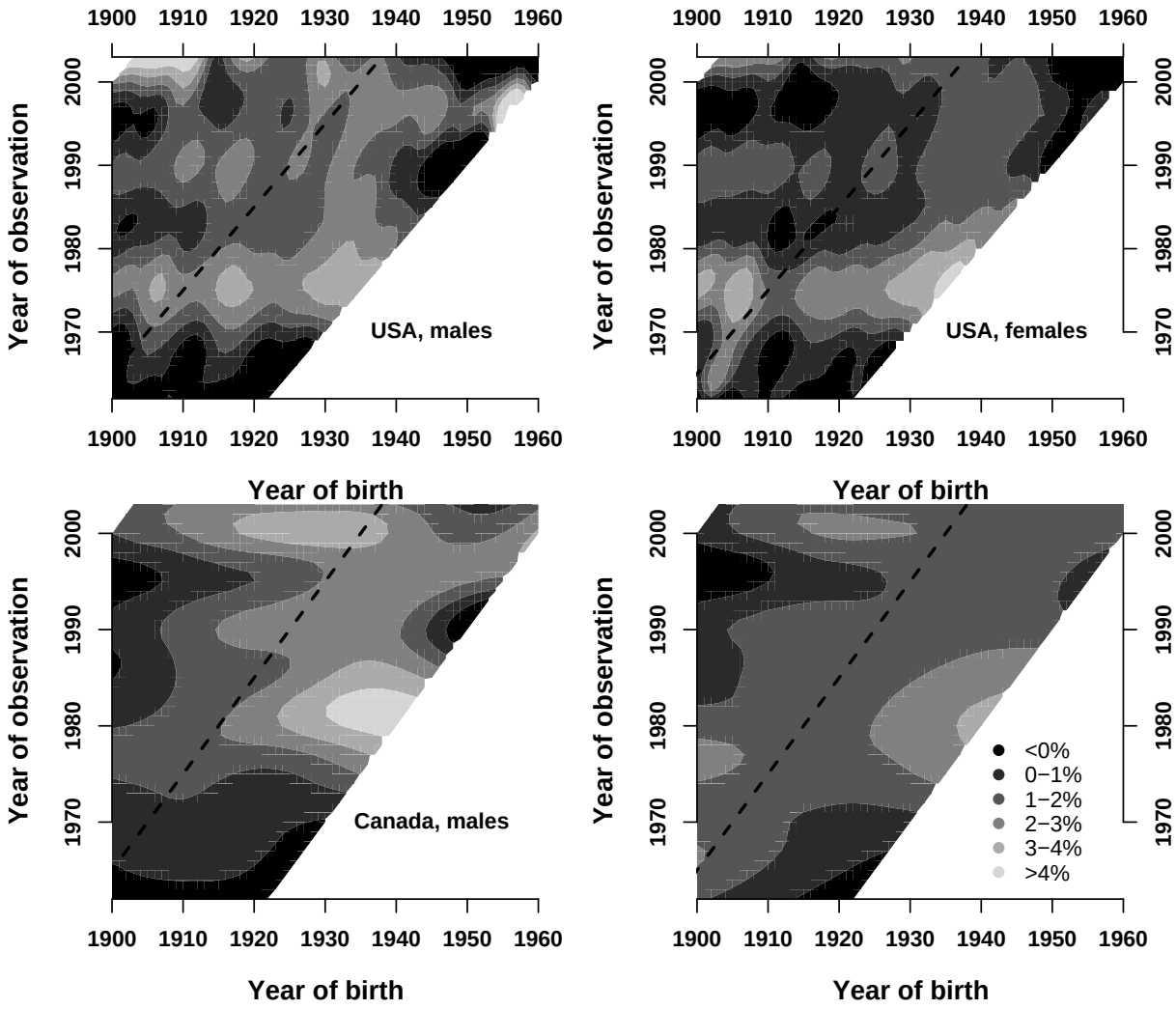

Figure 5. Improvements in smoothed force of mortality: *P*-spline regression with age-cohort penalties for U.S.A. (top panels) and age-period penalties for Canada (bottom panels). Males in left panels, females in right panels. The dashed line shows age 65 for each cohort

Table 6. λ_{age} values from the P-spline models in Table 5 (males only). Selected countries in order of increasing population size (both genders) in Table 2.

Country	Population aged 40–100	λ_{age}
Sweden	4.5m	1,907.4
Canada	14.7m	1,747.2
England and Wales	25.1m	246.9
France	29.0m	692.8
Germany (West)	34.2m	1.0
Japan	67.6m	47.1
U.S.A.	127.6m	2.8

Source: Own calculations using HMD data over age range 40–100, and using data from 1961 until the end-year shown in Table 1. Five-year knot spacing for each of age, cohort and period.

11.6 Figure 5 shows the improvements in the force of mortality over the past forty years for the U.S.A. (top panels) and Canada (bottom panels). Despite using the age-cohort P-spline model

to fit the model to the data for the U.S.A., period effects in the 1970s are clearly visible for both genders, emphasizing the local flexibility of P-splines whichever penalties are used. Although the patterns appear very different, this is in part due to the fact that the U.S.A. has a population over eight times as large as Canada. As a result, heavier smoothing has been applied: λ_{age} for the U.S.A. is 2.8, while for Canada it is 1747.2. As a general rule, less smoothing is required with larger populations, as Table 6 shows. The obvious outlier is Germany, and this might be because of the data problems outlined in 6.4. Alternatively, it might be because the underlying data have in fact been pre-smoothed before being made available to the public.

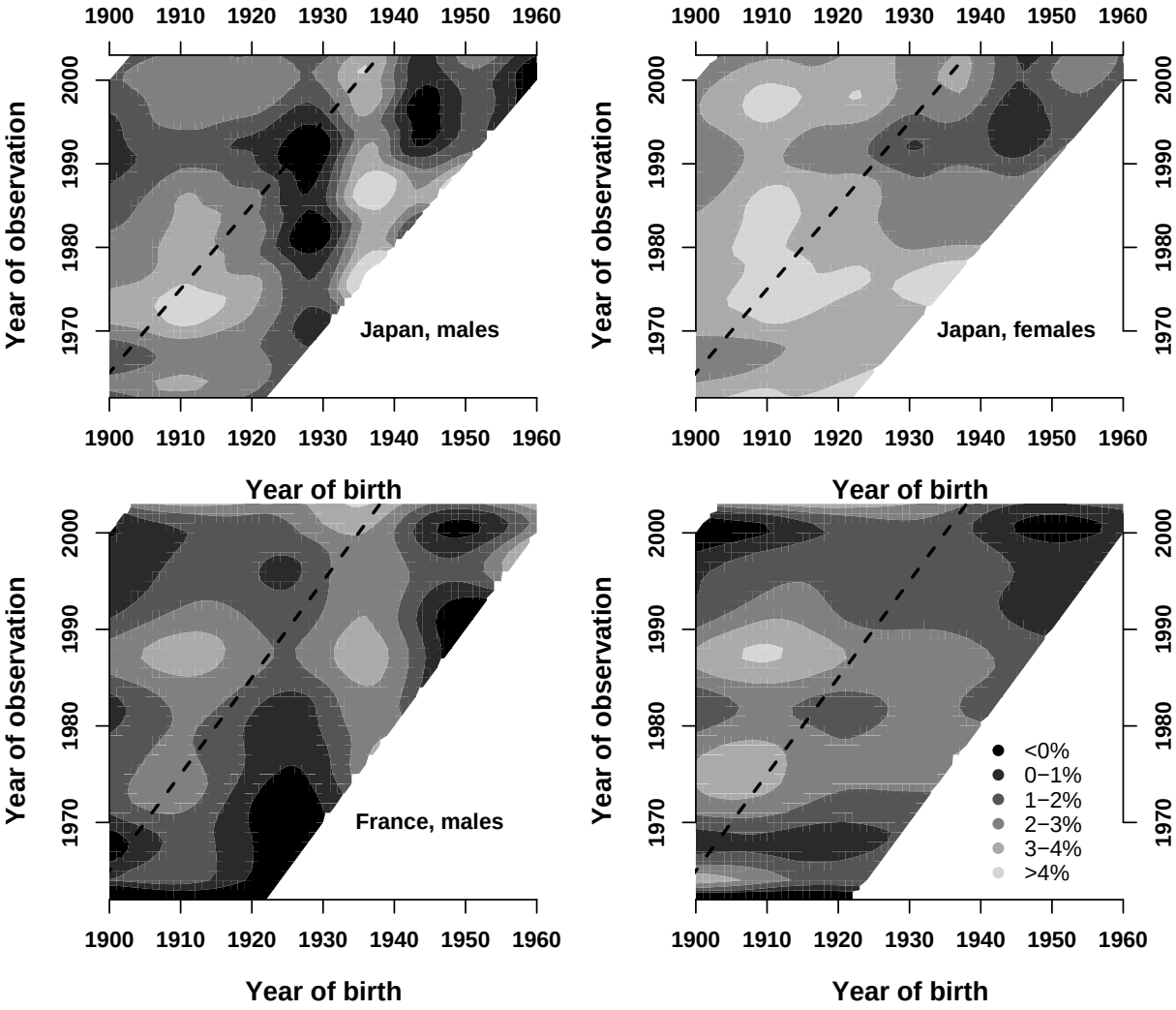

Figure 6. Improvements in smoothed force of mortality: *P*-spline regression with age-period penalties for Japan (top panels) and France (bottom panels). Males in left panels, females in right panels. The dashed line shows age 65 for each cohort

11.7 Figure 6 shows the improvements in the force of mortality over the past forty years for Japan (top panels) and France (bottom panels). Despite using the age-period P-spline model to fit the model to the data for males in both Japan and France, cohort effects are clearly visible as vertical patterns. Of all the heat maps in this paper, the one for Japanese females stands out for its general lightness, i.e. consistent high rates of improvement at all age, cohorts and times.

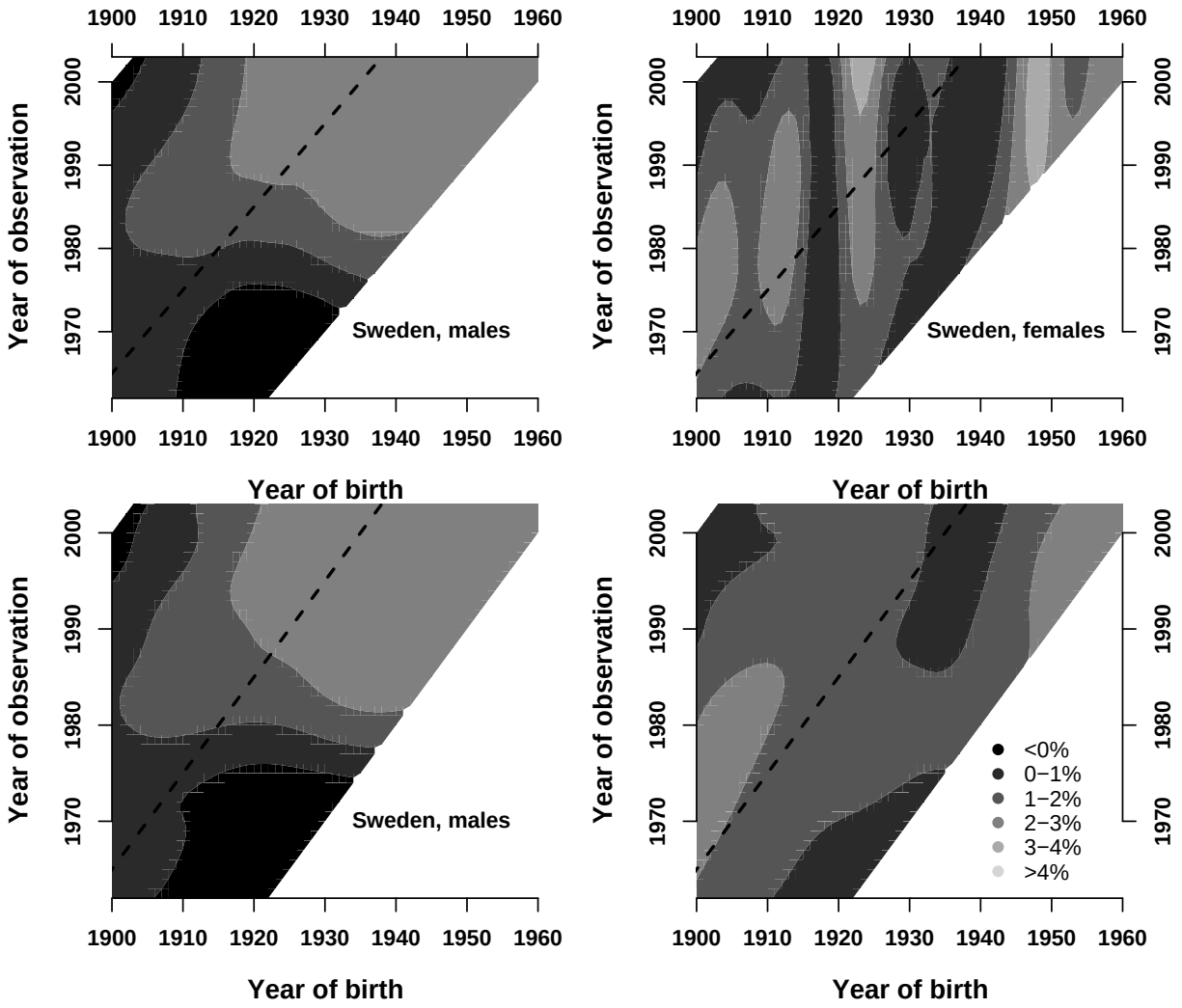

Figure 7. Improvements in smoothed force of mortality for Sweden: P-spline regression with age-cohort penalties (top panels) and age-period penalties (bottom panels). Males in left panels, females in right panels. The dashed line shows age 65 for each cohort

11.8 Figure 7 shows the improvements in the force of mortality over the past forty years for Sweden using both age-cohort penalties (top panels) and age-period penalties (bottom panels). Only females show clear evidence of cohort effects as vertical patterns, which echoes the finding of Table 5, where the BIC for the age-cohort model for females was lower than the BIC for the equivalent age-period model. The general lack of feature details in Figure 7 compared with Figures 4–6 is a by-product of the heavier smoothing applied to a small population (see Table 6).

12. International Trends by Cause of Death

12.1 This section examines the trends in cause of death between 1980 and 2000 of the seven selected countries. Causes of death are grouped into four broad categories, namely (i) diseases of the circulatory system, (ii) diseases of the respiratory system, (iii) malignant neoplasms and (iv) 'other'. This categorisation groups diseases that affect the same physiological system or else share similar characteristics. Diseases of the circulatory system include diseases of the heart and blood vessels, including heart attack and stroke. Respiratory diseases are diseases of the breathing apparatuses such as the windpipe, respiratory tracts and lungs, as well as influenza. Malignant neoplasms are diseases characterised by uncontrolled tissue growth, and for simplicity they will be referred to as cancers in this paper. These three categories accounted for some four-fifths of all deaths in most of the countries investigated here in 2000. The remaining causes of death are assigned as other causes.

12.2 Cause-of-death data for the seven selected countries were obtained from the WHO website. Data were analysed by gender and broad age band (65–74 and 75 plus). For Germany and Sweden, data were only available from 1990 onwards. For other countries, data from 1980 were available.

Table 7. Percentage of deaths falling into four broad categories in the indicated years for males aged 65–74. Bold figures indicate the leading cause of death. Where 1980 data are not available, 1990 data is used.

Country	Year	Circulatory	Cancers	Respiratory	Other	Total
Japan	1980	**42**	31	9	18	100
	2000	25	**46**	10	19	100
France	1980	**35**	33	6	26	100
	2000	26	**42**	6	25	100
Canada	1980	**50**	28	8	14	100
	2000	34	**39**	7	19	100
Sweden	1990	**53**	29	5	14	100
	2000	**44**	33	6	17	100
Germany	1990	**46**	30	7	16	100
	2000	**40**	35	7	18	100
U.S.A.	1980	**51**	27	8	14	100
	2000	**38**	34	10	19	100
U.K.	1980	**52**	28	13	8	100
	2000	**41**	35	13	11	100

Table 8. Percentage of deaths falling into four broad categories in the indicated years for females aged 65–74. Bold figures indicate the leading cause of death. Where 1980 data are not available, 1990 data is used.

Country	Year	Circulatory	Cancers	Respiratory	Other	Total
Japan	1980	**47**	27	6	20	100
	2000	28	**41**	8	22	100
France	1980	**37**	30	4	29	100
	2000	24	**41**	5	31	100
Canada	1980	**49**	31	5	16	100
	2000	28	**43**	8	21	100
Sweden	1990	**42**	37	5	16	100
	2000	33	**41**	7	18	100
Germany	1990	**44**	32	5	19	100
	2000	**39**	36	5	20	100
U.S.A.	1980	**50**	28	6	16	100
	2000	33	**35**	11	21	100
U.K.	1980	**51**	28	10	12	100
	2000	35	**38**	14	14	100

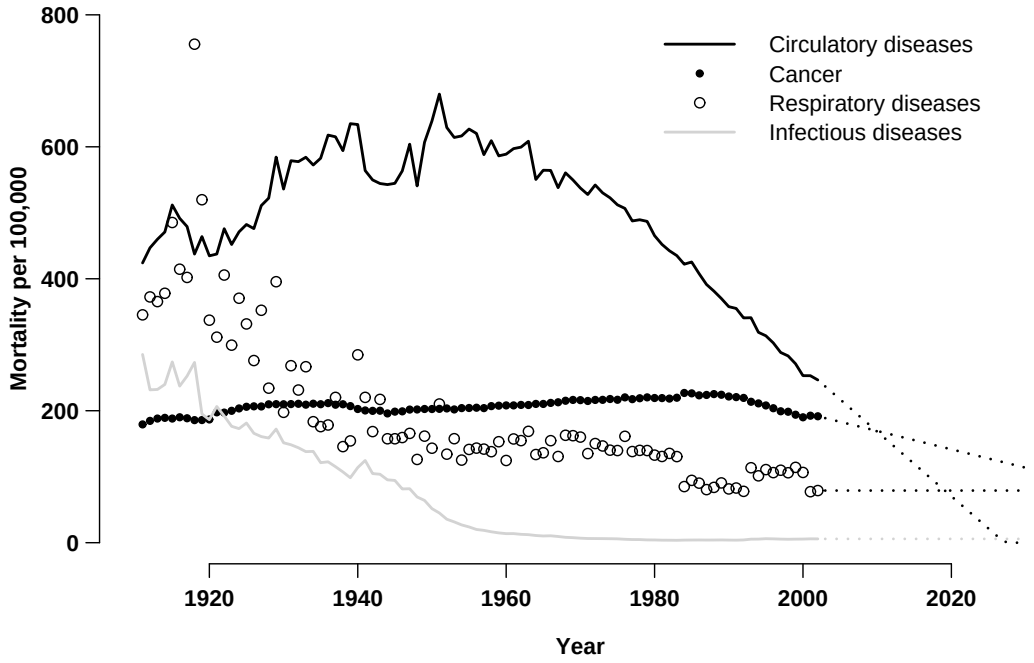

Figure 8. Mortality rates per 100,000 for England and Wales by main causes of death. Source: ONS data with own extrapolations beyond 2002. Taken from Richards, Kirkby and Currie (2006).

12.3 For people aged 65–74, the two leading causes of death in the countries analysed here are circulatory diseases and cancers. In 1980, circulatory diseases were the leading cause of death for males and females aged 65–74 in most of the seven countries shown in Tables 7 and 8. The proportion of people dying of circulatory diseases reduced over the two decades leading up to 2000. In 2000, cancers overtook circulatory diseases as the main killer for both genders in Canada, France and Japan, and for females in U.S.A. and Sweden. It would not be surprising if cancers were to replace circulatory diseases as the leading cause of death in the remaining four countries in the future for all people aged 65–74. The proportion of deaths due to cancers seems set to shortly overtake circulatory diseases as the leading category of cause of death, as shown for England and Wales in Figure 8.

Table 9. Percentage reduction in death rates by cause of death between 1980 and 1990 for ages 65–74. Bold figures indicate the cause of death making the largest single contribution to the overall reduction.

Gender	Country	Total	Circulatory	Cancers	Respiratory	Other
Males	Japan	23.7	**18.7**	1.0	–0.3	4.2
	France	26.9	**14.1**	2.3	1.9	8.6
	Canada	16.0	**16.6**	–1.9	0.7	0.5
	U.S.A.	14.5	**14.3**	–0.8	0.1	1.0
	U.K.	17.2	**11.5**	0.9	4.7	0.2
Females	Japan	30.8	**21.6**	3.7	0.4	5.1
	France	31.0	**17.2**	2.9	0.9	10.0
	Canada	13.6	**17.9**	–3.3	–1.3	0.2
	U.S.A.	6.8	**13.0**	–3.4	–2.6	–0.2
	U.K.	10.8	**11.7**	–3.0	2.2	–0.1

Table 10. Percentage reduction in death rates by cause of death between 1990 and 2000 for ages 65–74. Bold figures indicate the cause of death making the largest single contribution to the overall reduction.

Gender	Country	Total	Circulatory	Cancers	Respiratory	Other
Males	Japan	9.4	**8.3**	–2.3	2.3	1.2
	France	4.9	**4.1**	1.3	0.4	–0.9
	Canada	19.4	**12.9**	4.1	2.3	0.1
	Sweden	19.3	**16.8**	2.0	0.4	0.1
	Germany	17.8	**13.1**	1.5	1.6	1.5
	U.S.A.	14.7	**11.3**	3.5	0.9	–1.0
	U.K.	18.1	**13.9**	4.8	–1.3	0.7
Females	Japan	22.7	**14.1**	2.2	2.3	4.1
	France	3.6	**6.0**	–0.3	0.1	–2.2
	Canada	11.4	**10.6**	1.1	0.5	–0.8
	Sweden	13.3	**12.9**	1.0	–1.1	0.5
	Germany	19.5	**13.0**	2.7	0.4	3.5
	U.S.A.	3.5	**7.8**	0.7	–1.8	–3.1
	U.K.	11.9	**12.2**	2.4	–4.3	1.5

12.4 For ages above 75, the leading cause of death between 1980 and 2000 was circulatory diseases in every one of the seven countries and for both genders. The second major cause of death was either cancers or 'other', depending on the country. The proportion of deaths relating to circulatory diseases also reduced during this period. Indeed, one wonders if the transformation of mortality due to circulatory disease between ages 65–74 will not shortly be repeated at age 75 and over.

12.5 The population above age 65 in all seven countries experienced decreasing number of deaths relating to circulatory diseases per 100,000 people over the two decades leading to the year 2000. The trends seem likely to continue into the future. There are noticeable differences in the death rates relating to circulatory diseases between countries. For example, the death rates of the males aged 65–74 in the U.K., U.S.A. and Germany were still higher in 2000 than those in Japan in 1985. Death rates relating to circulatory diseases are higher for males compared with females within the same country. However, this is not true when comparing the death rates of males and females between countries. For example death rates relating to circulatory diseases of females above age 75 in Sweden, U.K., and Germany were higher than those of males in France and Japan in 2000.

Table 11. Percentage of deaths falling into four broad categories in the indicated years for males aged 75 and over. Bold figures indicated the leading cause of death. Where 1980 is not available, 1990 data is used.

Country	Year	Circulatory	Cancers	Respiratory	Other	Total
Japan	1980	**51**	16	13	21	100
	2000	**36**	24	22	19	100
France	1980	**43**	20	9	27	100
	2000	**35**	25	10	30	100
Canada	1980	**56**	19	11	14	100
	2000	**40**	25	12	23	100
Sweden	1990	**56**	18	11	16	100
	2000	**51**	20	9	20	100
Germany	1990	**56**	20	10	15	100
	2000	**52**	22	9	16	100
U.S.A.	1980	**59**	18	10	13	100
	2000	**45**	21	13	21	100
U.K.	1980	**51**	18	21	9	100
	2000	**42**	22	21	14	100

Table 12. Percentage of deaths falling into four broad categories in the indicated years for females aged 75 and over. Bold figures indicated the leading cause of death. Where 1980 data is not available, 1990 data is used.

Country	Year	Circulatory	Cancers	Respiratory	Other	Total
Japan	1980	**55**	11	9	25	100
	2000	**41**	19	17	24	100
France	1980	**49**	13	6	31	100
	2000	**39**	15	8	38	100
Canada	1980	**63**	15	7	15	100
	2000	**44**	19	9	29	100
Sweden	1990	**59**	14	8	19	100
	2000	**53**	14	8	25	100
Germany	1990	**63**	15	6	17	100
	2000	**60**	16	6	18	100
U.S.A.	1980	**67**	13	6	14	100
	2000	**49**	15	11	25	100
U.K.	1980	**58**	13	17	12	100
	2000	**43**	16	20	21	100

12.6 The trends of death rates relating to cancers of the various countries were less coherent, unlike the clear decreasing trends of circulatory deaths between 1980 and 2000. The shape of the trends depends on gender, age-group, country and year. Some populations exhibited decreasing death rates relating to cancers between 1980 and 2000. Examples of these are males aged 65–74 in the U.K. and France. Females aged 65–74 in Japan and France also saw reductions between 1980 and 2000. Some populations experienced increasing death rates relating to cancers over the two decades leading to 2000. Examples include females above 75 in Canada and U.S.A., and both genders above age 75 in Japan. There were also populations that showed 'humped' shapes in the trends of cancer death rates. These populations experienced increases in death rates relating to cancers from the 1980s to 1990s, followed by a fall in death rates in the 1990s. Examples of these

populations are males aged 65–74 in Canada and U.S.A.; females aged 65–74 in Canada, U.K. and U.S.A.; males above 75 in Canada, France, U.K. and U.S.A.; and females above 75 in U.K. and Germany.

12.7 The fall in death rates in the 65–74 age-group in most countries was of the order of 10-20% per decade between 1980 and 2000, with the largest fall occurring between 1980 and 1990. A large proportion of the improvement in mortality could be attributed to the decline in death rates due to circulatory diseases (see Tables 9 and 10). For example, between 1990 and 2000 the total reduction in death rates for U.K. males aged 65–74 was 18.1%, of which 13.9% could be attributed to the fall in death rates relating to circulatory diseases (Table 10). So, the fall in circulatory deaths contributed to some four-fifths of the total mortality improvement over the decade leading to 2000. Similarly for the other countries, reduction in death rates due to circulatory diseases has been the largest contributor by far to the total fall in death rates.

12.8 Of particular interest in Table 10 is the experience of Japanese females. Although they already had the lowest mortality rates of any group in 1990, they nevertheless experienced the largest reduction in mortality in the decade to 2000. Clearly, already-low mortality rates are no barrier to further large improvements.

Table 13. Percentage reduction in death rates by cause of death between 1980 and 1990 for ages 75 and over. Bold figures indicate the cause of death making the largest single contribution to the overall reduction.

Gender	Country	Total	Circulatory	Cancers	Respiratory	Other
Males	Japan	12.8	**15.3**	–2.2	–4.8	4.4
	France	10.3	**9.1**	–1.6	–0.2	2.8
	Canada	9.1	**15.1**	–1.3	–1.8	–3.0
	U.S.A.	8.4	**13.9**	–1.7	–2.0	–1.9
	U.K.	13.9	**10.3**	–1.2	7.7	–2.9
Females	Japan	20.2	**16.2**	–0.8	–2.0	6.8
	France	9.9	**9.8**	–0.4	–1.2	1.6
	Canada	7.3	**15.9**	–1.0	–2.1	–5.5
	U.S.A.	4.2	**13.3**	–1.6	–3.5	–4.0
	U.K.	10.1	**10.7**	–1.6	5.9	–4.9

Table 14. Percentage reduction in death rates by cause of death between 1990 and 2000 for ages 75 and over. Bold figures indicate the cause of death making the largest single contribution to the overall reduction.

Gender	Country	Total	Circulatory	Cancers	Respiratory	Other
Males	Japan	14.6	**12.9**	–2.2	1.0	2.9
	France	8.9	**5.8**	1.9	2.1	–0.8
	Canada	10.7	**8.9**	0.7	3.3	–2.1
	Sweden	8.8	**9.7**	–0.5	2.3	–2.7
	Germany	18	**12.6**	1.8	2.2	1.4
	U.S.A.	6.8	**8.1**	1.2	1.1	–3.5
	U.K.	10.4	**10.1**	2.4	–3.5	1.3
Females	Japan	18.8	**16.1**	–0.8	0.1	3.5
	France	4.9	**6.5**	0.7	1.1	–3.4
	Canada	3.9	**9.4**	–0.9	0.7	–5.3
	Sweden	3.0	**7.8**	0.1	1.1	–6.0
	Germany	10.5	**8.8**	0.5	0.5	0.8
	U.S.A.	–4.8	**5.1**	–0.9	–1.0	–8.0
	U.K.	2.2	**10.6**	0.4	–7.0	–1.8

12.9 The second-most important contributor to the fall in death rates was less clear-cut for the various countries. For example, the second largest contributor to the fall in death rates for U.S. males aged 65–74 between 1980 and 1990 was 'Other' deaths. However, between 1990 and 2000 it was cancer deaths (Table 9). Death rates relating to some causes of death had increased in certain countries. One notable example was the rise (negative improvement) in 'other causes' death rates in the U.S.A. for both genders above age 75 (Tables 13 and 14). This increase caused an actual rise in death rates above aged 75 for females in the U.S.A. between 1990 and 2000. We also note again the experience of Japanese females in Table 14: despite already having the lowest mortality rates in 1990, they still managed the largest proportionate fall in mortality in the decade to 2000.

12.10 For males aged 65–74, deaths relating to circulatory diseases have been reducing but maintained the position as the leading cause of death between 1980 and 2000 in the U.K. and U.S.A. For France and Japan, circulatory diseases were the leading cause of death in 1980 but fell below that of cancers from the mid-1980s. Pronounced declines in death rates relating to circulatory diseases, but not other causes of death, were observed in all countries. French males aged 65–74 in 1980 experienced higher death rates relating to 'other' causes than males in other countries. This could be explained, at least partly, by their higher death rates in 'accidents and adverse effects' and diseases of the digestive system especially 'chronic liver disease and cirrhosis'. For females, the general trends and features resemble that of the males discussed above. However, death rates of the various broad categories of causes of death were generally lower than those of males.

12.11 Males over the age of 75 in France enjoyed higher life expectancies than their counterparts in the U.K. between 1980 and 2000. However, males in these two countries experienced comparable death rates relating to causes other than circulatory diseases, ranging between about 1,000 and 3,000 deaths per 100,000 individuals from 1980 to 2000 (rates by cause of death are not shown). Death rates relating to circulatory diseases appear to be the distinguishing factor for the differences in life expectancies of males above age 75 in these two countries. For females aged over 75, the death rates relating to causes other than circulatory diseases were relatively stable between 1980 and 2000, at below 2,000 deaths per 100,000 individuals in the U.K. and U.S.A. However, deaths relating to 'Other' diseases increased in France but fell in Japan from 1980 to 2000. Circulatory deaths among women over 75 in all four countries fell over the same period.

12.12 In the two decades up to 2000, the key trends in causes of death for the countries inves-

tigated here could be summarised as: (i) death rates relating to circulatory diseases fell and seem set to fall further in the future; (ii) death rates relating to cancers fell or stabilised; (iii) death rates relating to respiratory diseases stabilised; (iv) death rates relating to 'Other', except in the U.S.A., had been stabilising; and (v) trends of deaths caused by the four broad categories appear to be relatively independent. This is consistent with the views of some researchers that competing mortality from other causes with either circulatory or cardiovascular deaths is minimal (Unal *et al*, 2005).

13. The Fall in Circulatory Disease

13.1 Circulatory disease plays a key role as a major killer of older people and changes in this were the biggest contributor to the improvement in mortality rates between 1980 and 2000. It is worth discussing some potential factors that have caused the decline in death rates relating to circulatory disease.

Table 15. Components of reduction in mortality due to coronary heart disease in Scotland (1975–1994), England and Wales (1981–2000) and New Zealand (1982–1993). Adapted from Capewell *et al* (1999), Capewell *et al* (2000) and Unal *et al* (2004).

Proportion of reduction due to changes in:	**Scotland**	**England and Wales**	**New Zealand**
Smoking	36	48	30
Population blood pressure	6	10	8
Cholesterol	6	10	12
Deprivation	3	3	n/a
Other factors	9	–13	4
Total risk-factor reduction (%)	60	58	54
Heart failure	8	13	6
Secondary prevention	6	11	7
Acute myocardial infarction	10	8	12
Hypertension treatment	9	3	7
Aspirin for angina	5	3	9
CABG surgery	2	4	5
Total treatment (%)	40	42	46
Overall (%)	100	100	100

13.2 The prevention or delay of deaths resulting from circulatory diseases can be attributed to two broad factors, namely medical treatment and changes in risk factors in the population: Bunker (1995), Unal *et al* (2005), Capewell *et al* (1999) and Capewell *et al* (2000). Treatment and changes in risk factors have been estimated to have contributed 40% and 60% respectively of the fall in death rates relating to coronary heart disease (CHD) in Scotland (Capewell *et al*, 1999). Very similar results were estimated for New Zealand and England and Wales. The percentage contribution of treatment and changes in risk factors to the total fall in CHD mortality based on the findings of Capewell and colleagues are summarised in Table 15.

13.3 Changes in risk factors and treatment could prevent or postpone deaths at two levels — the occurrence of the disease on the one hand, and the fatality from the disease on the other. The World Health Organisation Multinational Monitoring of Trends and Determinants in Cardiovascular Disease (WHO MONICA) Project has been set up to monitor the trends in determinants of mortality and morbidity due to cardiovascular disease. Studies on the MONICA data have shown that two-thirds of the fall in CHD deaths can be attributed to the fall in occurrence rates and one-third to

the fall in fatality rates (Tunstall-Pedoe *et al*, 2000). In other words, reductions in the causes of heart attacks have contributed more to mortality improvement than the factors that help people survive one, i.e. prevention is better than cure.

13.4 A study on 21 countries from the WHO MONICA project estimated that changes in the classic risk factors could explain 40% and 15% of the trends in the occurrence of CHD for men and women respectively (Kuulasmaa *et al*, 2000). The risk factors examined were smoking, blood pressure, cholesterol level and body mass index (BMI). In a parallel study on the effect of medical treatment on both CHD events and subsequent fatality rates, an association between medical care and reduction in coronary events or mortality was reported (Tunstall-Pedoe *et al*, 2000), albeit with a note of caution that there might have been other causal factors which were not measured in their analysis.

13.5 Changes in risk factors over the past few decades could be due to complex interactions of several factors. These factors include changes in lifestyle, changes in living and working conditions, and changes in life-long accumulation of exposure to health risks. Examples of changes in lifestyle include reduction in smoking and reduction in physical exercise in the U.K. population (Unal *et al*, 2005). An example of a change in living conditions is the improvement in heating systems, which could prevent high blood pressure as a result of coldness and thus reduce winter deaths. With the contraction of heavy industry and expansion of service industries, the workforce has moved from factories and mines to offices with better working conditions. Improvement in mortality rates among the elderly could be the result of the accumulation of benefits to health throughout their lives, including improved nutrition and reduced exposure to infections during childhood.

13.6 The impact of treatment depends on several factors such as the development of treatment, availability of infrastructure to deliver the treatment and behaviour of patients. Breakthroughs in science and technology can result in much more effective treatment, as is the case with statins, a cholesterol-lowering drug available since the 1990s. The availability of an effective health care service allows the delivery of appropriate intervention to patients (Nolte & McKee, 2004). The attitude of patients in seeking medical help and compliance with treatment regimes also affects the outcome of medical intervention. Improvement in any of the factors affecting treatment mentioned above could result in further improvements in morbidity and mortality due to circulatory disease.

14. Communicating Trends Using a Cause-of-Death Interpretation

14.1 Trends in causes of death have been widely studied or used for communication purposes (Bunker *et al*, 1994; Nolte & McKee, 2004; Hayflick, 2001). Although not all authors seek to project future mortality, their work often includes estimation of scope for improvement in mortality. For example, it has been suggested that preventative services for hypertension are not fully utilised in the U.S.A. Bunker *et al* (1994) estimated that if used to their fullest potential, such services could add a further $1\frac{1}{2}$–2 months to the life expectancy of the population of the U.S.A.

14.2 The concept of deaths which are preventable or amenable to health care is reviewed in detail by Nolte & McKee (2004). This concept suggests that there are deaths which are either untimely or unnecessary given the existing health-care facilities and knowledge. Examples of such deaths include most life-threatening diseases of the circulatory system and cancers for people below age 75 (Nolte & McKee, 2004).

14.3 Between 2000 and 2010 the Department of Health in the U.K. aims to reduce deaths due to CHD by 40% and cancer deaths by 20% for people aged below 75. Hayflick (2000), a leading cellular biologist, used the scenario of elimination of deaths due to cancers, cardiovascular diseases and stroke when discussing longevity issues. Thus, one means of communicating and expressing the future outlook for mortality is in terms of decline in specific causes of death.

Table 16. Reduction in cause of death: number of years to achieve the reduction from 2001 as projected using P-splines: 50th percentile of age-cohort projection using England and Wales data based on ages 20–100 between 1956 and 2002. Each scenario shows the fall in mortality due to circulatory diseases/cancer/other causes.

		Scenario:		
Gender	**Age**	**40/20/10**	**100/100/0**	**100/50/50**
Males	70	7	37	52
	80	8	29	46
	90	11	26	41
Females	70	8	45	61
	80	10	39	55
	90	14	29	47

14.4 For example, deaths relating to diseases of the circulatory system and cancers contributed to about 70% of total deaths of males aged 80-84 in England and Wales in 2001 (own calculation from WHO data). On the face of it, the elimination of deaths due to these two causes could result in a reduction of 70% in death rates for a male aged 80. A P-spline projection at 50th percentile using the age-cohort method and England and Wales data for ages 20–100 from 1956 to 2002 showed that this level of reduction could be achieved by 2030 (i.e. after 29 years — see Table 16). One could therefore express the results of a P-spline projection as equivalent to the eradication of 'Circulatory' and 'Cancer' deaths within 29 years from 2001 for males aged 80 in England and Wales, assuming that death rates due to other causes remain constant and are independent of each other, as shown in Table 16. The corresponding number of years required to achieve the various scenarios of reduction in causes of death as indicated by P-spline projection are shown in Table 16 for males and females at selected ages.

14.5 The approach in Table 16 ignores the problem of competing risks — the figures in Table 16 are overstatements under this suggested mechanism, as the elimination of one cause will result in other causes killing more people. This is because those who would have been removed by the eliminated cause will still be exposed to risk, even if remaining death rates are unchanged. This insight was first documented by Bernoulli (1766), who attempted to disentangle the risk of dying from smallpox from other risks. The method underlying Table 16 also ignores an issue with identifiability: we have observed cause A in conjunction with competing cause B. How do we know that cause A will not change in behaviour if cause B is eliminated? An interesting example is smoking, which causes deaths due to circulatory diseases and cancers. If medical interventions such as statins and ACE inhibitors save many smokers from death due to heart disease, this could result in an increase in certain cancer rates in the shorter term as more smokers in a pre-cancerous state will now be in the exposed-to-risk. It is perhaps mere coincidence, but Table 13 shows (i) sharp falls in mortality due to circulatory disease, and (ii) small increases in cancer mortality for people over age 75 between 1980 and 1990.

14.6 Table 16 could be used to communicate the results of actuarial projections in terms of the reduction of various causes of death. However, it would be up to the user to assess the reasonableness of the underlying hypothesis of elimination (or permanent reduction) of a cause of death, and to note that the approach assumes risks are independent and errors in recording cause of deaths are within acceptable range. Both P-spline projections — and others, such as the 'long cohort' (CMIB, 2002) — have been questioned anecdotally as requiring very large reductions in certain causes of death. This needs clarifying, not least because cause-of-death elimination is a widely used currency to describe changes in mortality. It is misleading to think that a particular P-spline projection (or, indeed, any other projection) requires the elimination of a major cause of death to come true, and that the P-spline projection is therefore unrealistic or unlikely. One view, which is entirely

consistent with falls in the incidence of particular causes of death in certain age bands, is that much of the improvement seen to date might simply be a delay in the onset of these causes of death. At the extreme, if everyone's ultimate cause of death remained the same but the age of onset was merely delayed by a few months each year, this will have the same effect as the P-spline mortality improvements while keeping the relative roles of the causes of death entirely unchanged.

14.7 We advocate viewing mortality improvements as having three components by cause of death: (i) delay in onset; (ii) reduction in incidence, for example lung cancer; and (iii) genuine elimination or near-elimination, for example smallpox. We also suggest that there are four potential 'brakes' acting in the opposite direction: (i) increase in incidence of an existing cause of death, for example alcohol-related deaths in England and Wales are rising; (ii) resurgence, for example tuberculosis in an antibiotic-resistant form; (iii) acceleration of onset, e.g. diabetes; and (iv) wholly new causes of death, for example SARS (Severe Acute Respiratory Syndrome).

14.8 Projection of the various causes of death could complement current methods of projecting mortality. Historical evidence from various countries (Nolte & McKee, 2004) illustrates that mortality rates due to different causes of death have experienced different trends of improvement. The possibility of projecting these trends deserves consideration. After all, many methods of projection — including P-splines — involve extrapolation of past trends.

14.9 The scope for improvement in mortality rates as a result of a better use of current health care service is different for different causes of death (Bunker *et al*, 1994). For example, hypertension is more easily treatable than a brain tumour. Future medical advances could also have a different impact on mortality rates for different causes of death. For example, the development of a vaccine for cancer-causing human papilloma virus would only affect cervical cancer deaths in women and not other causes of death.

14.10 The nature and causes of diseases are different, hence the complexity of methods of dealing with them must also be different. Some infectious agents mutate regularly making new interventions necessary: topical examples are the H5N1 influenza virus, MRSA, antibiotic-resistant tuberculosis and some strains of human immunodeficiency virus (HIV). Other diseases, such as heart attack, do not change. Changes in the constitution of the body would have different effects on different diseases. For example, improvement in health as a result of a healthier diet would reduce diet-related diseases, such as colo-rectal cancer.

14.11 A cause-of-death projection might be suitable for projection of period effects. This might suit populations or segments of populations where period effects predominate. However, one must note that cause-of-death assessments should be interpreted with care. Shojania & Burton (2004) conducted a meta-analysis of 53 studies of autopsy-detected errors in clinical diagnosis, including some which spanned a 40-year period (1959 to 1999). Discrepancies between clinical and autopsy diagnoses were defined as 'class I errors' when the patient might have survived to hospital discharge if antemortem diagnosis had occurred. The studies reported class I error rates of 9.0% (ranging from zero to 20.7%), which would be a close proxy to misdiagnoses between ICD classes such as respiratory, cardiovascular and cancers (Shojania *et al.*, 2002). For our analysis in section 12, we have grouped causes of death into fewer groups than ICD classes hence one could suggest that the error rate would be smaller. Even if all deaths were to be diagnosed by autopsy, the current 'gold standard' in identifying cause of death, there would still be an error of 1–5% uncertainty (Shojania *et al*, 2002). For narrower definitions of cause of death, such as at the level of pulmonary oedema and aorta rupture, the error rates could be as high as 23.5% (ranging from 4.1% to 49.8% for 'major errors'). However, the various error rates estimated by Shojania *et al* (2002, 2004) may well be higher than the true error rates because clinicians tend to send clinically uncertain cases for autopsy.

15. P-Splines for Projections

15.1 CMIB (2005a) introduced the idea of using the results of P-spline regression not just for fitting mortality data, but for mortality projections as well. The projected values are provided by the penalties, which essentially extrapolate from the recent past using the given penalty and λ values. The advantage of this is that standard errors can also be obtained from the model. Note that the CMIB only offered this methodology to the profession and remains fundamentally agnostic on the choice of projection models: at the time of writing the same working party behind CMIB (2005a) is about to publish a working paper on the Lee-Carter (1992) model (CMIB, 2007).

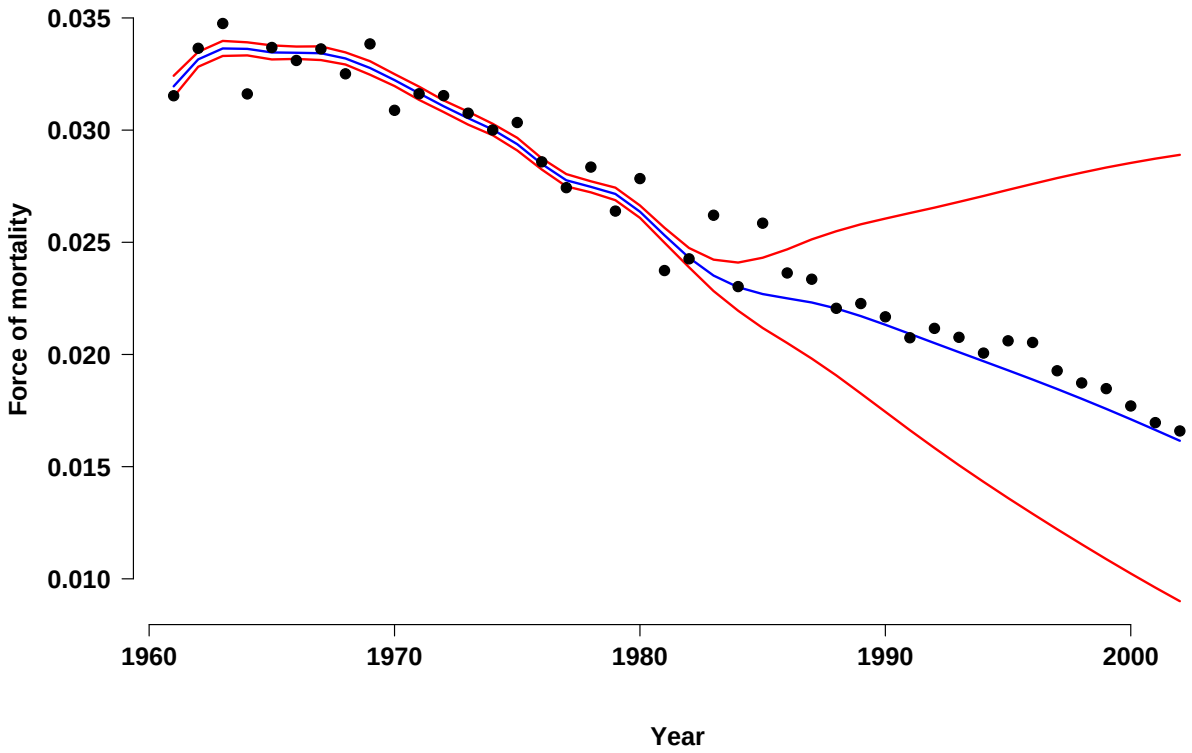

Figure 9. Observed force of mortality for French males at age $65\frac{1}{2}$ ($\bullet$), together with projected values and $2\frac{1}{2}\%$ and $97\frac{1}{2}\%$ percentiles based on a P-spline model using only the first half of the data to 1982 (age-cohort penalty). The projection from 1982 has proved remarkably prophetic

15.2 One way to test the appropriateness of a projection methodology is to conduct a back test, i.e. to fit the model to the first half of the data and then to compare the resulting projection with the subsequently observed values. Figure 9 shows the results of just such a back test applied to data for French males. Clearly the P-spline methodology, had it been available in the past, would have provided a very good projection of the future force of mortality. Of course, this does not mean that P-spline projections applied now will give as good a projection in the future, but it does at least give confidence that the methodology is not obviously flawed. Back-testing was used in CMIB (2005a) when validating the P-spline methodology.

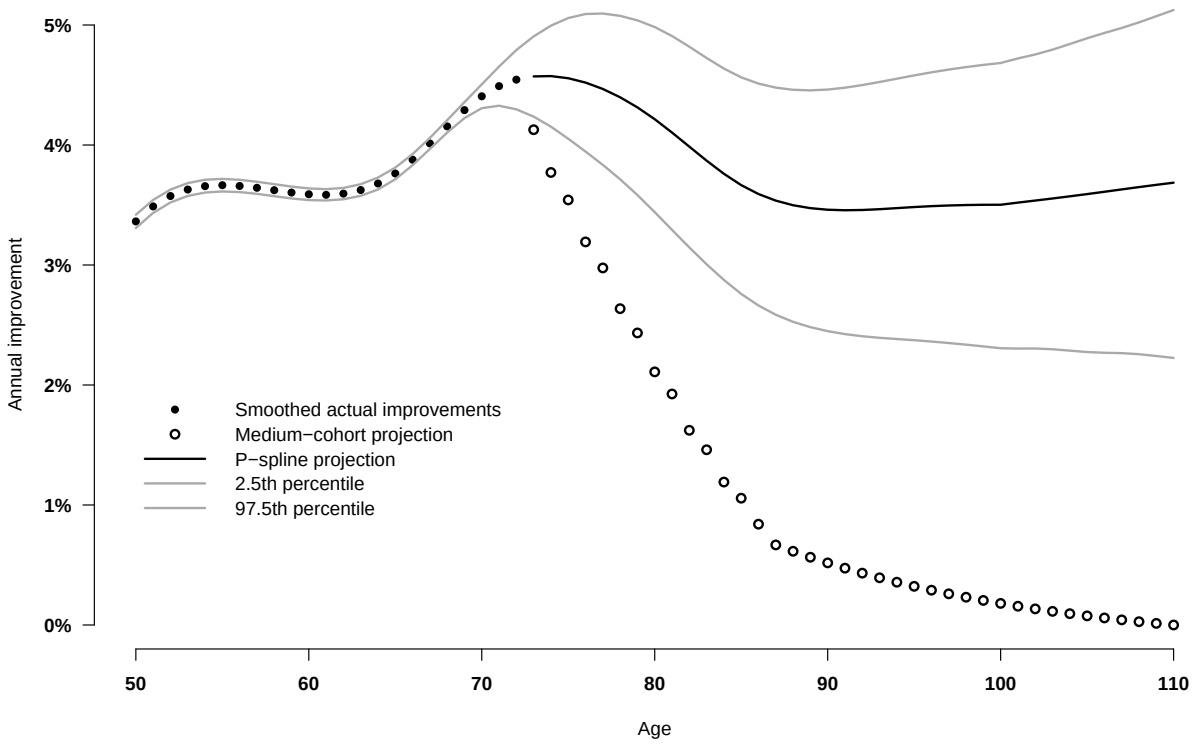

Figure 10. Observed smoothed improvement in force of mortality for males born in 1931 in England and Wales (•), together with projected improvements in q_x according to the medium cohort (◦) and the projected improvements in force of mortality according to a P-spline model using age-cohort penalties (solid black line). 2.5% and 97.5% percentiles for the fitted and projected P-spline values are indicated by the solid grey lines

15.3 Another test, albeit a subjective one, is whether a projection method produces a natural-looking extrapolation of actual data. Figure 10 shows the smoothed actual improvements observed for males in the 1931 birth cohort in England and Wales. The P-spline projections arguably show a more intuitive short-term extrapolation of past improvements than other projection bases, such as the medium-cohort projection (CMIB, 2002). Given that the past twenty years have seen a steady *acceleration* in improvement rate — from $3\frac{1}{2}\%$ to $4\frac{1}{2}\%$ p.a. — the sharp deceleration of the medium cohort does not look like a natural extrapolation. Furthermore, the low rate of improvement of under $\frac{1}{2}\%$ from age 91 onwards suggests that some kind of floor value should be used. Equally, however, one could argue that the P-spline projected improvement of around $3\frac{1}{2}\%$ above age 100 seems strong, but at least such projections come with percentile values to indicate the likeliness of possible alternative values under the same model.

16. Possible Criticisms of P-Splines for Projections

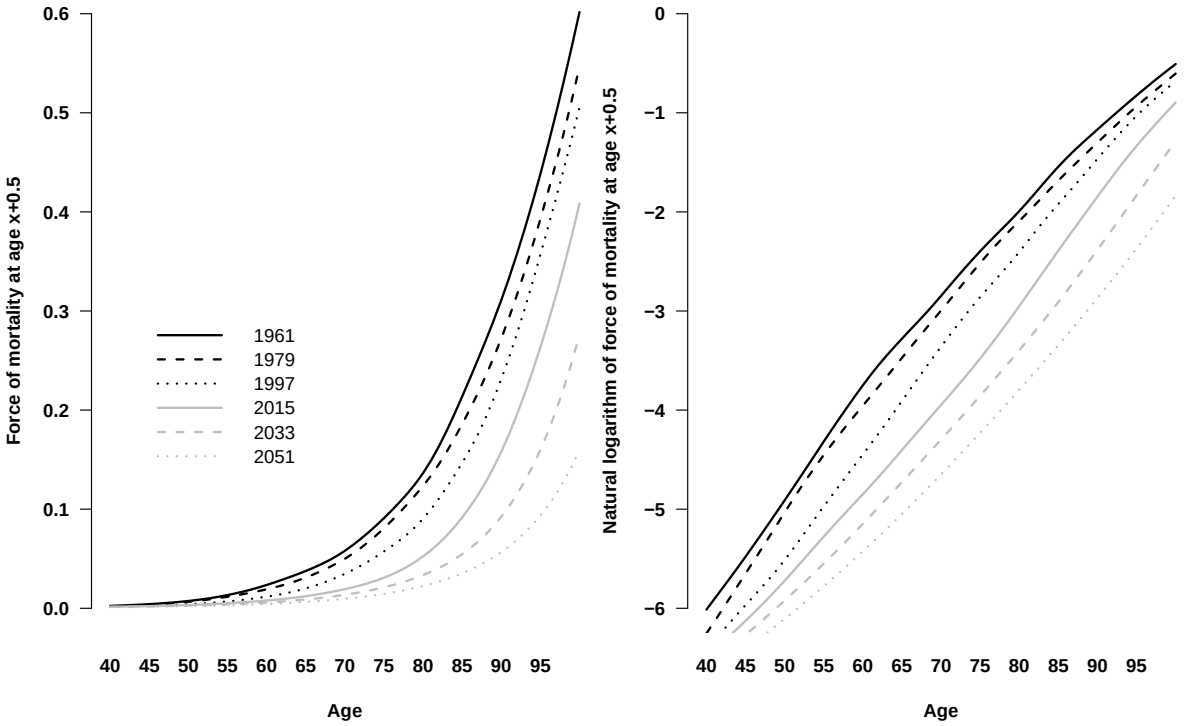

Figure 11. Fitted force of mortality for mid-point of calendar years 1961, 1979 and 1997 and projected force of mortality for calendar years 2015, 2033 and 2051. Left panel: original scale; Right panel: natural log scale; *P*-spline regression on data from mortality.org for males in England and Wales, 1961–2003 (with age-period penalties)

16.1 One possible criticism of P-spline projections is what they can do to the structure of mortality by age. This is demonstrated in Figure 11, which shows the age structure of the fitted force of mortality for male mortality in England and Wales over the period 1961-2002 at equally spaced intervals. The black lines are fitted values based on the actual data, the grey lines are the projected values. The left panel shows the very substantial improvements in actual mortality between 1961 and 1997, including the recent acceleration in improvements: the vertical gap between the 1979 and 1997 lines is often bigger than the gap between the 1961 and 1979 lines. The left panel also shows the continuation of this improvement trend for later years, although there appears to be a qualitative difference between the 2015 line and the later ones: the curvature of the age progression is preserved in the 2015 projection, but it appears to be being lost in the later ones. Indeed, long-term projections using P-splines can tend to flatten the mortality curve against age.

16.2 Gavrilov & Gavrilova (1991) called the convergence of mortality rates on a logarithmic scale the 'law of mortality convergence'. The right panel shows convergence between the first three lines, which illustrates the tendency for lower initial mortality to be compensated for with a faster rate of ageing as measured by the slope of the line on the logarithmic scale. It is here that we see the qualitative difference between the three projection lines. The 2015 projection is clearly part of the same pattern as the data lines, with convergence of the lines looking likely to be some time after age 95. One implication of this is a slowing down in the rate of improvement with age: if the lines are converging with higher age, the implied mortality improvements are progressively smaller. In contrast, the lines for 2033 and 2051 are either nearly parallel (which implies constant improvements) or even diverging (which implies ever-accelerating improvements). However, it is not clear if the projected mortality in Figure 11 won't simply converge at a higher age than the current data region. Furthermore, it is not axiomatic that the lines should have to converge at all.

Indeed, one must be careful of the underlying data — 6.4 describes how the 'data' for England and Wales above age 89 is not actually data but that the exposed-to-risk are artificial constructs.

16.3 One way of imposing convergence would be to enter a target pair at an advanced age at which no improvements are expected in the future. Although this target pair would be treated as part of the data for fitting purposes, in actual fact it would be part of the model as it would represent an a priori expectation of a limiting value for the force of mortality. The target pair would be made up of two components, namely the count of deaths and the exposed-to-risk. The pair would also have a Bayesian-like interpretation: the same value target of 0.5 could be represented as 1 death with an exposure of 2.0 (a weak prior expectation), or it could be represented as 1,000 deaths with an exposure of 2,000.0 (a strong prior expectation). With this target pair in the future at an advanced age, the penalty function in the P-spline regression will ensure that the fitted values smoothly progress from the region of the real data in such a manner that convergence is guaranteed. This procedure is analogous to the technique of exponential smoothing towards a target.

16.4 Another alternative being explored by Dr Iain Currie and James Kirkby is the use of a cross-product term in age and time, with an independent third penalty on this new term in the regression basis. However, while the projections are reported to be better, they do still have the tendency to flatten the mortality curve as before.

Table 17. Volatility in projected improvements under P-spline model with either the addition of an extra year of data, or in reducing the knot spacing.

Last year of data	**Knot spacing**	**Range of improvements**	**BIC**	**Annuity value**
2001	5	2–3%	2,818	14.35
2002	5	$\frac{1}{2}$–1%	2,901	13.42
2002	2	$\frac{1}{2}$–2%	2,928	13.79

16.5 There are some other potential criticisms of P-splines. For example, we were unable to fit an age-cohort model for French females in Table 5. Also, in performing a large number of runs of the P-spline model against the population data from all the countries we have encountered results which can vary greatly with the selection of data. Table 17 shows results from Canadian female data using an age-period model. The improvement-factor range covers ages 60–90 and the annuity value is calculated for year of use 2007 at 4.5%. The only observation that can be made from this is that care needs to be taken when using results for financial projections.

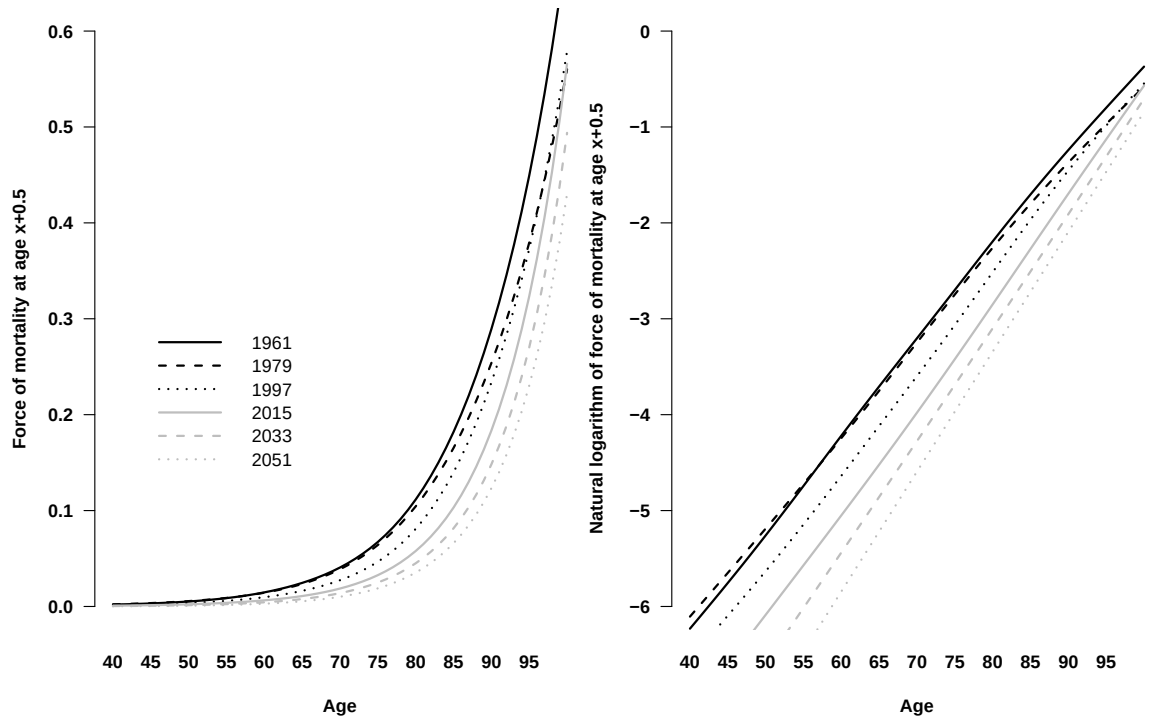

Figure 12. Fitted forces of mortality for the mid-points of calendar years 1961, 1979 and 1997 and projected forces of mortality for calendar years 2015, 2033 and 2051. Left panel: original scale; Right panel: natural log scale; P-spline regression on data from mortality.org for males in Sweden, 1961–2005 (with age-period penalties)

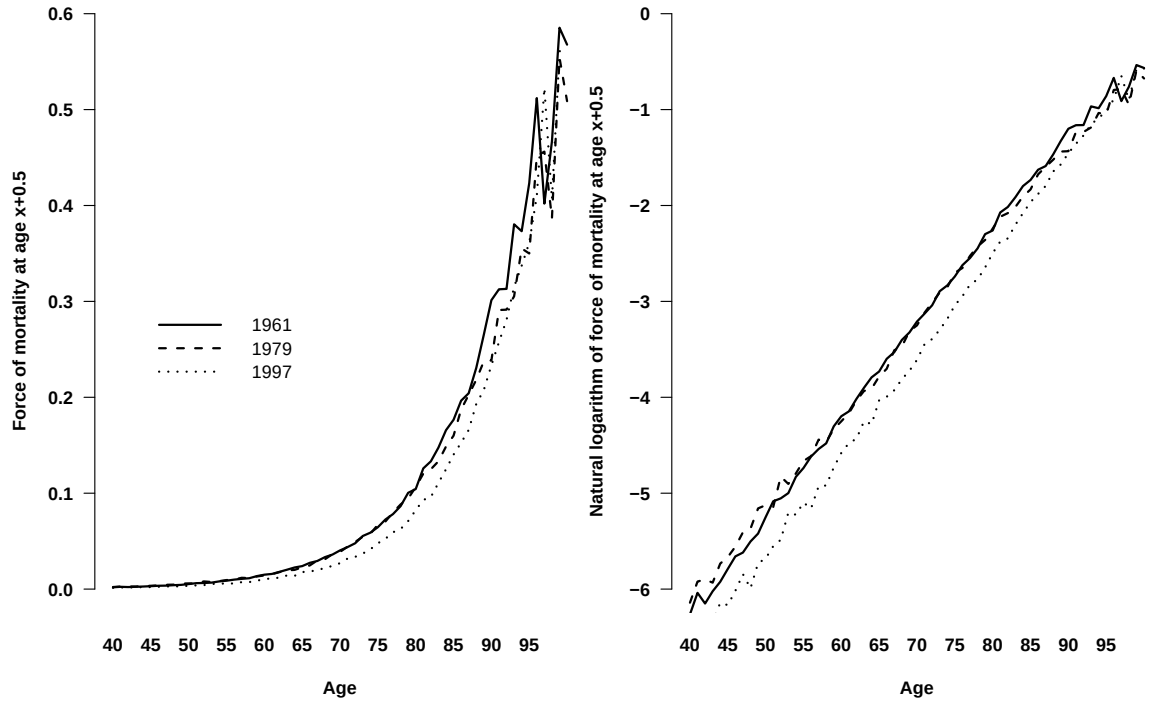

Figure 13. Crude forces of mortality for Sweden in calendar years 1961, 1979 and 1997. Left panel: original scale; Right panel: natural log scale

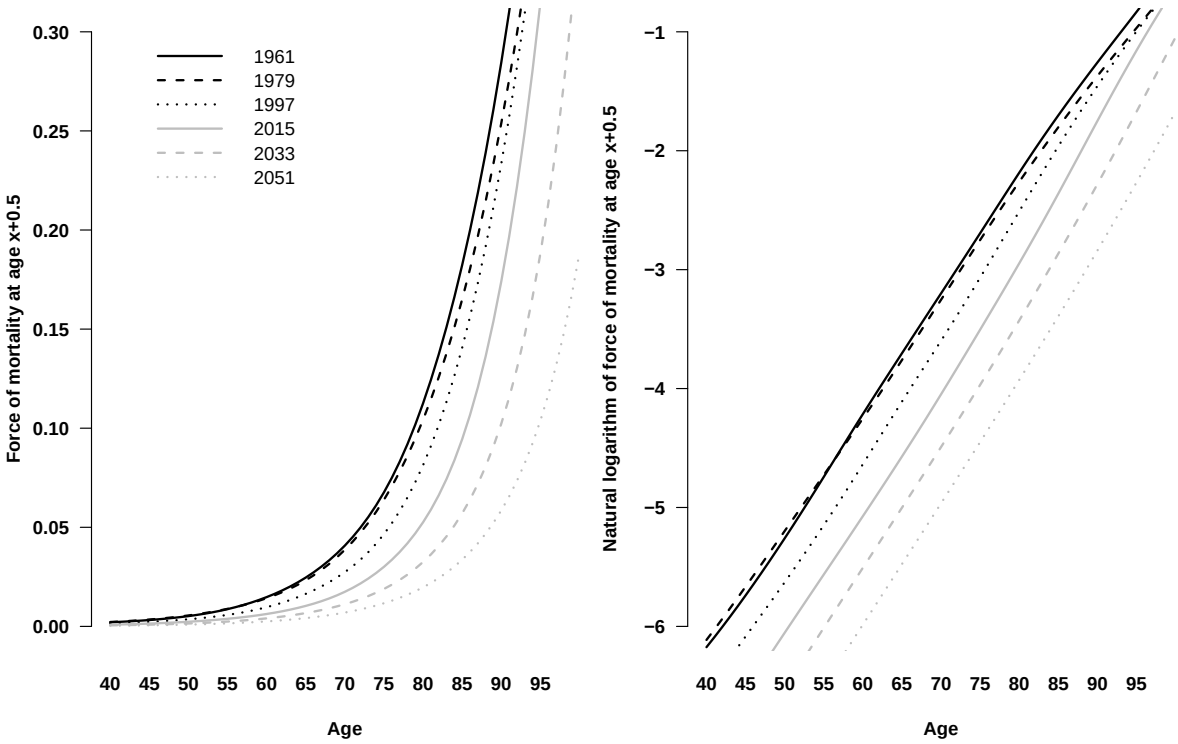

Figure 14. Fitted forces of mortality for the mid-points of calendar years 1961, 1979 and 1997 and projected forces of mortality for calendar years 2015, 2033 and 2051. Left panel: original scale; Right panel: natural log scale; P-spline regression on data from mortality.org for males in Sweden, 1961–2005 (with age-cohort penalties)

16.6 A potentially useful way of forming a view on whether the results from P-spline models are suitable basis for future mortality improvements, or informing one's thinking on such a basis, is to analyse what they do to the age structure of mortality. Whilst Figure 11 shows some flattening of the mortality curve against age towards the end of the projection period, the results nevertheless do not appear to be unreasonable. Similar comments may be made for projections on different data sets. Figure 12 shows fitted forces of mortality for males in Sweden. Whilst one wouldn't immediately expect the rates for 1979 and 1997 to cross over at older ages, as observed in the left panel of Figure 12, this phenomenon is not a criticism of P-splines: it occurs in the area of the data and it is entirely consistent with the observed crude death rates in the underlying data, as shown in Figure 13.

16.7 It is interesting to compare the effect on age structure if one instead used age-cohort penalties, as shown in Figure 14. The extent of the flattening of the mortality curve by age is significantly more pronounced, particularly for the projections in the later years. It is perhaps no coincidence that the age-period model seems to provide a better projection than the age-cohort model: we observed in Table 5 that the age-period model is a better-fitting model with a lower BIC.

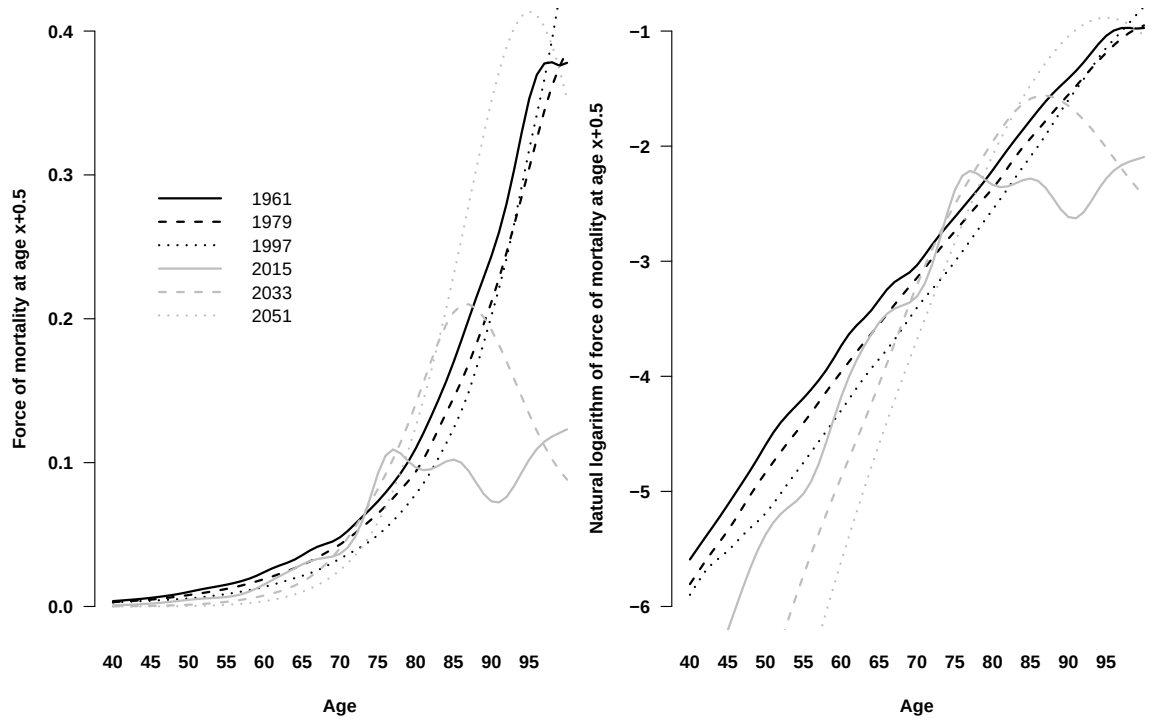

Figure 15. Fitted forces of mortality for the mid-points of calendar years 1961, 1979 and 1997 and projected forces of mortality for calendar years 2015, 2033 and 2051. Left panel: original scale; Right panel: natural log scale; P-spline regression on data from mortality.org for males in U.S.A., 1961–2003 (with age-cohort penalties)

16.8 Figure 15 shows fitted forces of mortality for males in the U.S.A. From Table 5, the age-cohort model is a better-fitting model than the corresponding age-period model. The fitted forces of mortality in the region of the data look slightly odd at the older ages, but this is a feature of the data (crude rates not shown) rather than one of P-splines. That notwithstanding, one cannot describe the P-spline projections for 2015 and 2033 as anything other than erratic. Whilst this does not necessarily imply that P-splines can't be used to project sensibly on all data sets (it is possible, for example, that different knot spacings may produce more sensible results), it very strongly reinforces the need to look critically at the outputs of the P-spline model before applying them in any practical context. In situations where strange effects are occurring at young ages, these effects would be carried carried through to older ages by the projection of the cohort effect. With the U.S. data, any increased mortality at pre-retirement ages, perhaps due to AIDS or even violent crime, would be carried through into the projections. One interesting feature of Figure 15 is that the P-splines have re-asserted a more sensible shape of mortality curve by 2051, albeit with higher mortality at ages above 75.

16.9 Notwithstanding these comments on the suitability of P-spline modelling for projecting future improvements in mortality rates, we do advocate their use in analysing past trends. By construction, P-spline models smooth data: one can therefore more readily apply them to gaining insights into the developing trends in historic mortality experience.

16.10 As described in CMIB (2004) and CMIB (2005b), P-spline models apply more smoothing to smaller data sets than they do larger data sets, even if they relate to the same age and period ranges. A manifestation of this in P-spline projection is that confidence intervals become narrower as the data sets become smaller. This is initially counter-intuitive: one naturally expects greater levels of uncertainty relating to estimates from smaller data sets. To show the extent of this effect, we took the data for males in Sweden and created two further data sets: one larger data set where the numbers of deaths at each age in each year, and the corresponding exposures, were increased by a factor of 10; and one smaller data set where the deaths and exposures were decreased by a factor

of 10. Although $D_{x,t}$ represent the number of deaths from a Poisson process, and must therefore strictly be whole numbers, the numbers of deaths at each age and in each year do not need to be rounded as the mathematics can deal equally well with non-integer death counts.

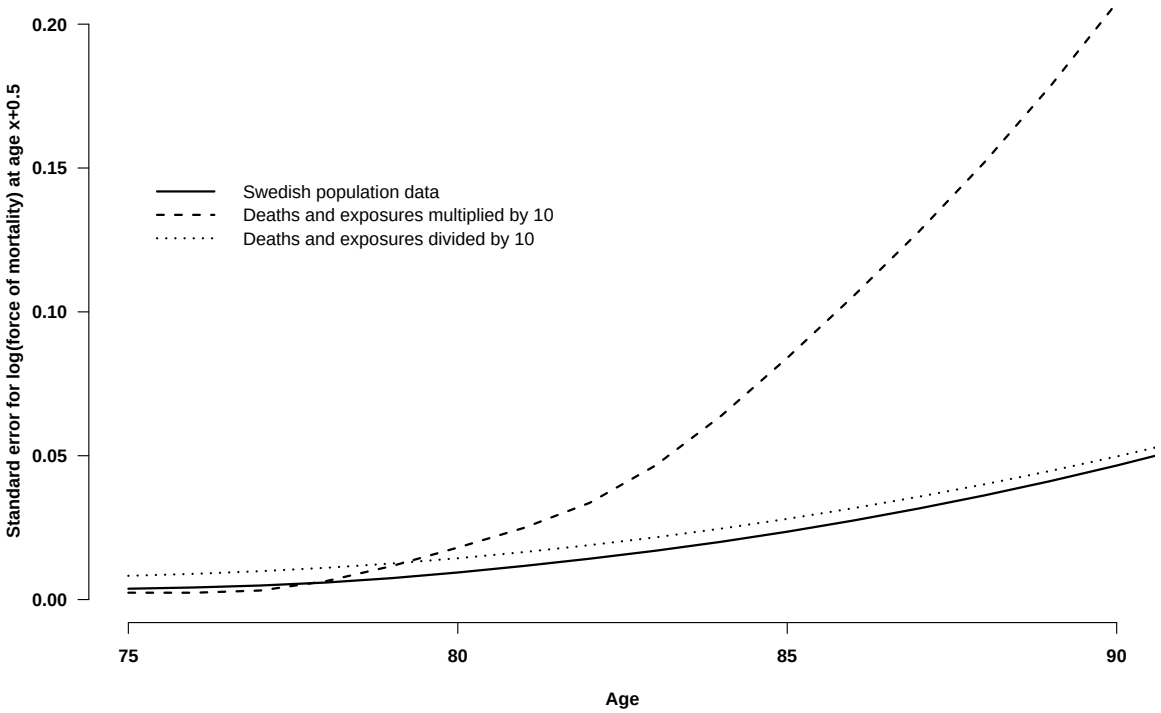

Figure 16. Standard errors for natural logarithm of fitted and projected forces of mortality for a male life initially aged 75 years in 2003; P-spline regression on adjusted data from mortality.org for males in Sweden, 1961–2005 (age-period penalty)

16.11 Figure 16 shows the standard errors from the fitted models on the 'central', large and small data sets, in this instance for a cohort-based view (i.e. year of birth) for a 75 year-old in 2003. It is interesting to observe two things: (i) that the standard errors in the early years of the projection follow the intuitive order s.e.(large data set) $\leq$ s.e.(central data set) $\leq$ s.e.(small data set); and (ii) that the standard errors in the later years of the projection follow the counter-intuitive intuitive order s.e.(central data set) $\leq$ s.e.(large data set).

16.12 It is important that any modelling or projection method is not used as a black-box. The model fit and properties of the projections should always be examined for suitability. Other models and projections should always be considered if at all possible. The two-dimensional P-spline method will generally give a good fit to data since it is a genuine two-dimensional local smoother, but projections depend on a number of other factors. Forecasting will work best when (a) the year knots are not too close together (b) when the time signal is strong relative to the age signal. By 'time signal' we mean either calendar time or birth cohort, depending on the model, and by 'strong signal' we mean a discernable trend instead of strong year-on-year variability. This happens with the CMIB's assured-lives data set, which is made available alongside the projection software in CMIB (2005a), and P-spline smoothing and projection works well there. However, there are national data sets where this is not the case, and, in such cases the projection can very quickly become linear across ages and this gives very odd projections. In such cases other methods should be used.

16.13 As we saw with the U.S. data, the projections can do odd things at high ages. This is not of itself an issue as the data do odd things at these ages as well. The P-spline model is primarily a local two-dimensional smoother, and it will pick up local behaviour and balance that with smoothing. If you have ideas about how the smooth values should behave in projection, then

you should put them into the model, such as the targetting approach outlined in 16.3. Without any such guidance within the model, the P-spline method will simply smooth — and project — the data as it finds it.

16.14 The scale paradox is important. Multiplying deaths and exposures by 10, say, leaves the data 'the same' — except that it has much more credibility. The P-spline model thus follows the data more readily since the balance of fit and smoothness has been fundamentally changed by the adjustment of the data. In the limit the smoothing function will do its best to interpolate. Reversing the process we find that reducing the data gives smoother functions because the method has less grounds for following the data. In the limit we get a straight line in one dimension. A straight line is a very strong modelling assumption, and it is this assumption that feeds into the confidence interval, resulting in a narrower spread. The explanation of the paradox is that the confidence interval is conditional on the choice of model. However, the model choice has not been taken into account in the computation of the confidence interval, and this paradox applies to the whole of classical confidence-interval calculation. If simple linear regression is the correct model, then the confidence interval is correct. However, if simple linear regression has been chosen from a larger family of models then the confidence interval is not correct. The modern approach is to use a model-averaging method, but it is not obvious how this applies in the smoothing context.

16.15 It is worth thinking a bit more about the paradox of wider confidence intervals when multiplying the data by a factor. Like all paradoxes, it is incomprehensible until the explanation is found. The argument is (i) smooth some Poisson data, (t_i, d_i, e_i), where t_i is year, d_i is deaths and e_i is exposure, and compute the confidence intervals; (ii) multiply d_i and e_i by 10, say, and repeat. The confidence intervals are now wider. This appears paradoxical, since with more data we would expect a narrower confidence interval. There is, however, an error in (i) and (ii): both cannot be correct. The problem is if the data in (i) are Poisson then the data in (ii) cannot be Poisson — the Poisson dispersion is only appropriate for the real data. There are two possible solutions. We can estimate the overdispersion parameter, ϕ, introduced in (ii) by comparing the usual variance $\sum \frac{\left(d_i - \hat{d}_i\right)^2}{n-p}$ with the Poisson variance $\sum \frac{d_i}{n}$ and adjust the BIC to:

$$BIC = \text{deviance} + \log n \times \phi \times p \tag{5}$$

where p is the trace of the model and the estimate of lambda should now go back to roughly what it was (using the terminology of Kirkby & Currie (2006), the confidence interval would be computed from $(B'WB+P)^{-1}$ so with larger W (since d_i is increased) the confidence interval should be narrower). The other possibility is to use a mixed-model approach, which introduces overdispersion automatically through the pure error variance. However, this is beyond the scope of this paper, and Dr Iain Currie is working on a paper using a mixed-model approach to smooth the APC model.

17. Results from Generalised Linear Models

17.1 An alternative approach to smoothing and projecting the force of mortality is to fit a generalised linear model (McCullagh & Nelder 1989). We could achieve automatic smoothing by using age as a direct variate: the right-hand panel in Figure 11 suggests that the full flexibility of P-splines is not necessary for such a regular and near-linear progression of the logarithm of the force of mortality with age. As an alternative to P-splines we could therefore fit a straight line on a log scale with age:

$$\log \mu_x = \alpha + \beta x \tag{6}$$

which we note is a generalised linear model with log link and Poisson error variable. It is also the law for the force of mortality proposed by Gompertz (1825):

$$\mu_x = e^{\alpha+\beta x} \tag{7}$$

17.2 We can thus model the observed number of deaths as a Poisson variable:

$$D_{x,t} \sim \text{Poisson}\left(\mu_{x+\frac{1}{2},t+\frac{1}{2}} \times E^c_{x+\frac{1}{2},t+\frac{1}{2}}\right) \tag{8}$$

where $E^c_{x+\frac{1}{2},t+\frac{1}{2}}$ is the central exposed-to-risk at age $x+\frac{1}{2}$ at time $t+\frac{1}{2}$, $\mu_{x+\frac{1}{2},t+\frac{1}{2}}$ is the force of mortality applying at that same age and time, and $D_{x,t}$ is the number of deaths aged x last birthday in year t. This is a simple generalised linear model, which we can fit in R, or even in a spreadsheet. We simply maximise the log-likelihood function (not shown) to get the maximum-likelihood estimates for the various parameters, and we compute the following statistic as an overall measure of badness-of-fit:

$$\mathrm{X}^2 = \sum \frac{\left(D_{x,t} - \mu_{x+\frac{1}{2},t+\frac{1}{2}} \times E^c_{x+\frac{1}{2},t+\frac{1}{2}}\right)^2}{\mu_{x+\frac{1}{2},t+\frac{1}{2}} \times E^c_{x+\frac{1}{2},t+\frac{1}{2}}} \tag{9}$$

17.3 The statistic X^2 measures the deviance of the model when applied to the data set, i.e. the size of the departure of the observed deaths from the fitted values under the model. Technically speaking, X^2 has a χ^2-squared distribution, but, as we shall see, it is unnecessary to calculate degrees of freedom and p-values to test various models and hypotheses. We can extend Equation 7 to cater for a number of alternative models, thus:

$$\mu_{x,t} = e^{\alpha+\beta x+\delta t} \tag{10}$$

where δ is the time-based direction of mortality and t denotes calendar time. We typically offset t by 2000.0 to keep the parameters well-scaled, i.e. t measures time since 1st January 2000 and can therefore be positive or negative. The parameter δ thus enables us to test time-based changes in the overall level of mortality and, clearly, gives us a simple model of projection as well as a simple model to explain past variation. Another extension is:

$$\mu_{x,c} = e^{\alpha_c+\beta_c x} \tag{11}$$

where c denotes the birth cohort, and α_c and β_c are cohort-specific values. We can use this model to explore the role of birth cohort in mortality levels. Again, this is also a model of projection: by fitting values of α and β for each cohort, we have in Equation 11 a formula which can be used to project future forces of mortality for each generation.

17.4 We can also combine Equations 10 and 11 to simultaneously test cohort effects *and* any separate time trend:

$$\mu_{x,c,t} = e^{\alpha_c+\beta_c x+\delta t} \tag{12}$$

which gives us a so-called Age-Period-Cohort (APC) model in which we can fit both cohort- and time-based parameters. We can use this model to give projected values for the force of mortality specific to each cohort and including an extrapolated time trend.

17.5 In a full Age-Period-Cohort (APC) model, a separate parameter would be fitted for each year of birth and for each year of observation. This typically leads to identifiability problems: with only a single observation of each cohort in each calendar year, further constraints are necessary for parameter estimation. In this model, however, we remove the need for identifiability constraints by grouping several years of birth into each cohort, and we model the broad time trend across many years, instead of fitting effects for every year. Note that with some simple models (possibly including this one), one can 'tilt' the values of β_c with consequent adjustments to the period and age effects. In other words, if there is a linear trend in the cohort effect then it can be difficult to disentangle this from a period effect. We will not explore this problem further here, however.

17.6 We fit these models by maximising the log-likelihood function (not shown), and we can compare models using Akaike's Information Criterion (AIC, Akaike, 1987). However, in this section of the paper we will just use the X^2 statistic to test quickly how well a particular model explains

the variation in observed death counts, and thus infer the broad relative significance of the features being explained by the model.

Table 18. Results of various models with time-based and cohort-based parameters. The best-fitting model (lowest X^2 statistic) has both cohort and time-trend effects.

Model	X^2	X^2 **relative to model 1**	**Number of parameters**
1. α, β constant, $\delta = 0$	24,365.3	0	2
2. α, β constant and $\delta = -0.01435$	6,668.9	−17,696.4	3
3. α and β varying by cohort, $\delta = 0$	2,880.4	−21,484.9	20
4. α and β varying by cohort, $\delta = -0.00824$	1,896.2	−22,469.1	21

Male population mortality data from mortality.org for England and Wales between ages 60 and 100 over the period 1961-2002. Where values for δ are non-zero, they are the maximum-likelihood estimates within the model. Models fitted using spreadsheet available at http://www.richardsconsulting.co.uk/Faculty

17.7 Table 18 shows just how much better a model including a time-trend effect fits compared to a model without: one extra parameter in Model 2 has led to a large drop in deviance. What is also clear is that allowing α and β to vary by five-year cohort bands is better-fitting still (Model 3). Perhaps most interesting of all is Model 4, which separates mortality improvements for this data-set into both cohort and time components and which fits best of all the four models. The corollary of this is that mortality improvements are primarily composed of cohort effects, but that a significant residual component is time-based. This vindicates the current life-office practice of using a 'floor percentage' or 'underpin' for mortality improvements in annuity reserving. The value of $\delta = -0.00824$ in the model appears to translate into a time-based improvement of around 0.82% p.a. However, it is not possible — and potentially misleading — to read across from this to the various life-office bases, as the value of δ can only be interpreted relative to the other parameters in the model. Furthermore, one must always remember that a valuation basis is about future improvements, for which the past experience may not be relevant.

17.8 Note that the assignment of experience data to cohort has not been done according to the usual Lexis diagram (Kirkby & Currie, 2006). This is because the cohort bands are relatively broad (each cohort is a five-year band) and any error in assigning a given year's experience to the correct cohort band is therefore relatively small. The benefit is that the spreadsheet model is kept much simpler as a result: see http://www.richardsconsulting.co.uk/Faculty

18. Results from Survival Models

18.1 Adopting the formulation in Equation 7 has a further advantage for life insurers and large pension funds: it means that future lifetime is being treated as a random variable and modelling can be done at the level of the individual instead of the group. Equation 7 is the basis for a *survival model*, and it can be shown that modelling μ_x is the same thing as modelling T_x, the future lifetime of the individual, as a random variable. In contrast to the Cox model (1972), we use the full likelihood instead of the partial likelihood. The full likelihood is based on the probability density function for T_x, which is:

$$f_{T_x}(s) = {}_sp_x\mu_{x+s} = \exp\left(-\frac{e^{\alpha+\beta x}}{\beta}\left(e^{\beta s}-1\right)\right)e^{\alpha+\beta(x+s)}, \qquad s \geq 0 \tag{13}$$

where the values of α and β can vary within the population, i.e. the i^{th} individual has values α_i and β_i.

18.2 This is a parametric regression model, similar to the Cox model (1972), but with t as a covariate, and with simultaneous estimation of the baseline hazard. Life-insurance and pension-scheme data is typically very well-suited to survival models: the exact entry age is known, as is the

exact age at death, and several relevant risk factors such as gender, birth cohort, mode of retirement and postcode. Thus, survival models not only analyse mortality differentials on an individual basis — which is useful for pricing — but they can also yield projections based on a portfolio's own mortality experience. Survival models are also not new: an actuary retiring at age 60 now might have qualified at age 25 in 1972, so the Cox model has been a standard tool among statisticians during his or her entire professional career.

18.3 Note that this is not to suggest that the P-spline approach is not also a survival model. Whenever P-splines are applied to the force of mortality (as is the case throughout this paper), then this is also a survival-type model. Thus, the density function under the P-spline model in Equation 12 for the future lifetime is:

$$f_{T_x}(s) = {}_sp_{x,t}\mu_{x+s,t+s} = \exp\left(-\int_0^t \mu_{x+s,t+s}ds\right)\mu_{x+s,t+s}, \qquad s \geq 0 \tag{14}$$

where the integral can be evaluated numerically.

18.4 There are important benefits of a parameteric formula for the force of mortality: (i) the explicit functional form for μ_x yields automatically smooth fitted values, and so no extra smoothing mechanism is needed like the penalty in P-splines; (ii) the model can be used to separate the components of improvement into cohort-driven parts and overall time-trends; and (iii) the model can be expanded to include other risk factors on an individual basis.

18.5 Equally, if we adopt the Equation 12 as our functional form for $\mu_{x,t}$, we get:

$$f_{T_x}(s) = {}_sp_x\mu_{x+s} = \exp\left(-\frac{e^{\alpha+\beta x+\delta t}}{\beta+\delta}\left(e^{(\beta+\delta)s}-1\right)\right)e^{\alpha+\beta(x+s)+\delta(t+s)}, \qquad s \geq 0 \tag{15}$$

18.6 We illustrate this approach by applying it to an insured data set comprising 92,890 males in receipt of a private pension. The pension records have been deduplicated, i.e. multiple records paid to the same person were identified and merged to form a single record. This is essential for statistical modelling at the individual level in order to preserve the independence assumption. The matching scheme used was based on a combination key of (i) date of birth, (ii) gender, (iii) surname, (iv) first initial, and (v) postcode. If all five data elements matched for two or more records, then they are assumed to be for the same person and they were replaced with a single merged record with the total pension. An algorithm called 'metaphone phonetic matching' (Phillips, 1990) was further used to catch common alternative spellings of surnames, e.g. Richie and Ritchie, and such records were also merged if the other four fields matched.

18.7 15,961 deaths were observed amongst the 92,890 males over a period of six calendar years. Exposure outside these dates was discarded, giving a total exposed-to-risk of 461,026 life-years, measured daily. This figure excludes any exposure or deaths either under age 60 or over age 95: the younger-age exposure is not modelled because it does not exhibit the same log-linear pattern of age-related mortality, while the exposure and deaths over age 95 are not felt to be wholly reliable. Note that not all 92,890 males were observed at the start of the period as some were new entrants to the portfolio during the period of observation. The elegance of this kind of survival modelling is that it can effortlessly handle fractional years of exposure.

18.8 Using Equation 15 as the form for $\mu_{x,t}$, the log-likelihood function, ℓ, to be maximised is:

$$\ell = \sum_{i=1}^{n} -\frac{e^{\alpha_i+\beta_i x_i+\delta t_i}}{\beta_i+\delta}\left(e^{(\beta_i+\delta)s_i}-1\right) + \sum_{i=1}^{n} d_i\left(\alpha+\beta_i(x_i+s_i)+\delta(t_i+s_i)\right) \tag{16}$$

where s_i is the time observed for life i, observed from age x_i at calendar time t_i to age $x_i + s_i$ at calendar time $t_i + s_i$, and where d_i is an indicator variable taking the value 1 on death or 0 on survival. In survival-modelling parlance, the data is both left-truncated and right-censored: left-truncated because each life is only observed from age x_i onwards, not age 0, and right-censored because because not every life has died before the end of the study.

Table 19. Results of various models with time-based and cohort-based parameters applied to an insured data set. The best-fitting model (lowest AIC) has both cohort and time-trend effects.

Model	AIC	AIC relative to model 1	Number of parameters
1. α, β constant, $\delta = 0$	128,770	0	2
2. α, β constant and $\delta = -0.032823$	128,703	−67	3
3. α and β varying by cohort, $\delta = 0$	128,421	−349	8
4. α and β varying by cohort, $\delta = -0.0210656$	128,402	−368	9

Male mortality data for ages 60–95 over a six–year period. Where values for δ are non-zero, they are the maximum-likelihood estimate within the model.

18.9 As with the population data, we fit similar model structures to this insured data set and the results are shown in Table 19, which is analogous to Table 18. Here we use Akaike's Information criterion (AIC, Akaike, 1987) to compare the four models, and the interpretation is similar to the χ^2 statistic in Table 18: the lower the AIC, the better the model. Unlike the χ^2 statistic, however, the AIC balances goodness of fit with the number of parameters in the model, so any difference in the AIC can be regarded as statistically significant.

18.10 Models 3 and 4 in Table 19 include cohort effects, and here we have used a function to find the optimum three breakpoints among the years of birth to create four broad cohorts: years of birth 1903–1909, 1910–1923, 1924–1932 and 1933–1944. Note that we have assumed four cohorts, but it is also possible to optimise not just the breakpoints, but also the number of cohorts as well. The optimal breakpoints between the years of birth are found by minimising the value of the AIC in the same way as the optimal smoothing constant λ is found by minimising the value of the BIC in the P-spline methodology.

18.11 We can see evidence of significant cohort effects as the AIC values for models 3 and 4 are much lower than Model 1. The time trend suggested in Model 2 is an improvement of around 3.23% p.a. ($3.23\% = 100\% - e^{-0.032823}$). Although the parameter values in different models are not strictly comparable, the suggestion in Model 4 is that this value of 3.23% is in part due to the cohort effect, as the not-quite-equivalent figure in Model 4 drops to 2.08% when the cohort effect is allowed for explicitly ($2.08\% = 100\% - e^{-0.0210656}$). In fact, we can see that the cohort effects are more significant than the time trend overall: adding the time trend on its own caused the AIC to fall by 67 ($= 128,770 - 128,703$), whereas adding the cohort parameters caused the AIC to fall by 349 ($= 128,770 - 128,421$). The two drops in the AIC can be compared, suggesting (i) that both cohort effects and a time trend are present, and (ii) that cohort effects are the stronger of the two.

18.12 One important difference between this insured data set and the wider population is that the former is a select sub-set. The apparently higher time-trend parameter could be due to the historic tendency of the insured population to outpace other groups in mortality improvements: Figure 1 suggests that life expectancies for socio-economic groups IV and V have risen proportionately less than for socio-economic groups I–IIIM, from which the insured population is disproportionately drawn. Another important difference is the time period: the population data stretches from 1961 to 2003, whereas this particular insured data set covers just six years from 1998 to 2004 inclusive. If there were a secular trend to a higher time-based rate of improvement, then this would give larger values of δ for the insured data set. To test this, the value of δ in the models in Table 18 could be allowed to vary piecemeal instead of a single value covering the entire period of study.

18.13 Thus, this U.K. insured data set echoes the same conclusion as the GLM for the England and Wales population data, namely that cohort effects are more significant than the time trend. The latter is nevertheless still a significant separate component of mortality improvements: the drop in AIC for a combined model with cohort effects and time trend was 368 ($= 128,770 - 128,402$),

which is better than just using time trend or cohort effects on their own. However, this apparent greater significance of cohort effects contradicts the P-spline findings in Table 5. This may be due to the use of a constant time-trend parameter: if it were free to vary, as is the case with the P-spline model, then we may find period effects dominant. One way to reconcile these two contradictions is that the two types of models are measuring slightly different things: the P-spline model copes with true period effects, i.e. years of elevated or reduced mortality, whereas the GLM and survival models here are measuring the broader trend. Arguably, an insurer with an annuity portfolio is more interested in long-term trends, rather than year-on-year fluctuations.

18.14 Note that the value of δ cannot be taken from a model in isolation and applied uncritically to a portfolio, such as for the floor value of improvements in a valuation basis. Indeed, it would be misleading to do so. The reason is that the value of δ can only be interpreted relative to the model which contains it. For example, the value of δ from Model 2 cannot be directly compared to the value of δ in Model 4 as they play subtlely different roles, and their values are defined only in relation to the other model parameters. By way of illustration, consider the value of δ in a refinement of this model. Richards and Jones (2004) identified duration since retirement as a significant rating factor, and, in the presence of duration as a variate, we have $\delta = -0.0227893$ (i.e. an improvement of 2.25% p.a., coming from $100\% - e^{-0.0227893}$). However, the coefficient of the duration variate is $+0.0202412$, i.e. mortality in this portfolio is increasing by 2.04% for each year after retirement. Thus, the net time-based change year-on-year for this portfolio is only 0.25% p.a. ($0.25\% = 100\% - e^{0.0202412 - 0.0227893}$). Note also the considerable uncertainty which also surrounds these parameters: the 95% confidence interval for the 2.25% value of δ is (1.41%, 3.09%)

18.15 This strong duration effect may in large part be due to anti-selection arising from the existence of the impaired-life and enhanced-annuity markets. The data used here come from a portfolio of non-underwritten annuities, and so business sourced externally has been pre-screened by financial advisers as qualifying for such annuities. The recently written business will therefore exhibit lighter mortality than can be explained by age, cohort and time-trend effects, thus giving rise to a duration effect. Thus, what was thought to be a pure time trend in Models 2 and 4 in Table 19 may in fact be due to a secular change in business mix (i.e. a healthier generation of non-underwritten annuitants). If we fitted a model explicitly contrasting externally sourced business (which is anti-selected) against internal vesting business (which is less anti-selected), then we would likely see further changes in δ. An alternative approach to stripping out the confounding effects of impaired-life anti-selection — which cannot be anything other than very significant — would be to fit a survival model to the business in-force up to, say, 1st January 2000, i.e. before the rapid development of the impaired-life market. This only serves to emphasise that particular parameters from a model for one dataset cannot be taken in isolation and applied to a possibly quite different portfolio.

19. Conclusions

19.1 Despite the faster improvements in male mortality over recent decades, the difference in life expectancy between males and females remains pronounced and there are no signs of it vanishing. Most recent commentary on the cohort effect in England and Wales has focused on males, but females also experience a similar cohort effect.

19.2 Although mortality improvements for males in England and Wales have been particularly strong, there seem few grounds for assuming that they will slow down or stop anytime soon. The low international ranking of both male and female life expectancies in England and Wales suggests that continued strong improvements are at least very possible. The example of Japanese females between 1990 and 2000 shows that relatively low mortality rates are no barrier to further dramatic improvements. This should be borne in mind by anyone tempted to argue that strong past improvements must somehow mean that future improvements will be less.

19.3 We have shown through back-testing that P-splines can provide good projections. However, we have also shown that there are circumstances where the projections are trickier. As with any

projection methodology, the P-spline toolkit provided alongside CMIB (2005a) must not be used as a 'black-box' forecasting tool.

19.4 The extent to which period or cohort effects dominate mortality patterns differs among countries, even to the extent that one can dominate for males and the other for females within the same country. Where one is dominant, however, this does not imply that the other is not significant: a simple statistical model shows how one can separate cohort effects from any long-term trend improvement using population data. A more sophisticated model describes how cohort effects and trends can be separated in an insured population. This model can also be used to allow for any complications caused by changes in business mix.

19.5 Finally, where an actuary chooses to apply a cohort-based projection of future mortality rates, evidence from both the general population and an insured data set in England and Wales suggests that a floor value of improvements may be required.

Acknowledgements

The authors thank Dr Iain Currie of Heriot-Watt University for helpful guidance on the use and interpretation of penalised splines, and for writing the original version of the section on penalised-spline regression. The authors also thank Professor Angus Macdonald, Professor Andrew Cairns and Nigel Knowles for their helpful suggestions and corrections. Any remaining errors are the sole responsibility of the authors.

References

Akaike, H. (1987). Factor analysis and AIC, *Psychometrica,* **52***, 317–333.*

Bernoulli, D. (1766). Essai d'une nouvelle analyse de la mortalite causee par la petite verole, *Mem Math Phy Acad Roy Sci, Paris.*

de Boor, C. (2001). A practical guide to splines, *Applied Mathematical Sciences* **27***, Springer-Verlag, New York.*

Bunker, J. P., Frazier, H. S. and Mosteller F. (1994). Improving health: measuring effects of medical care, *The Milbank Quarterly,* **72***, No. 2, 225–257.*

Capewell, S., Morrison C. E. and McMurray, J. J. (1999). Contribution of modern cardiovascular treatment and risk factor changes to the decline in coronary heart disease mortality in Scotland between 1975 and 1994, *Heart,* **81***, 380–386.*

Capewell S., Beaglehole, R., Sheddon, M. and McMurray, J. (2000). Explanation for the decline in coronary heart-disease mortality rates in Auckland, New Zealand, between 1982 and 1993, *Circulation,* **102***, 1511–1516.*

CDC (Center for Disease Control) (2006). Tobacco use among adults in the United States, 2005., *Morbidity and Mortality Weekly Report 2006,* **55(42)***, 1145–1148.*

CMIB (Mortality Sub-Committee) (2002). An interim basis for adjusting the 92 Series mortality projections for cohort effects, *Working Paper No. 1.*

CMIB (Mortality Sub-Committee) (2004). Projecting future mortality: a discussion paper, *Working Paper No. 3.*

CMIB (Mortality Sub-Committee) (2005a). Projecting future mortality: towards a proposal for a stochastic methodology, *Working Paper No. 15.*

CMIB (Mortality Sub-Committee) (2005b). Stochastic projection methodologies: further progress and P-Spline model features, example results and implications, *Working Paper No. 20.*

CMIB (Mortality Graduation Working Party) (2006). The graduation of the CMI 1999–2002 Mortality Experience: Final 00 Series Mortality Tables — Assured Lives, *Working Paper No. 21.*

CMIB (Mortality Projections Working Party) (2007). Stochastic projection methodologies: Lee-Carter model: features, example results and implications, *Working Paper No. 25.*

Cox, D. R. (1972). Regression models and life tables, *Journal of the Royal Statistical Society, Series B* **24***, 187–220 (with discussion).*

Craven, P. and Wahba, G. (1979). Smoothing noisy data with spline functions, *Numer. Math. 31, 377–403.*

Currie, I. D., Durban, M. and Eilers, P. H. C. (2003). Using P-splines to extrapolate two-dimensional Poisson data, *Proceedings of 18th International Workshop on Statistical Modelling, Leuven, Belgium, 97–102.*

Currie, I. D., Durban, M. and Eilers, P. H. C. (2004a). Array regression: an approach to smoothing data on arrays, *Unpublished paper.*

Currie, I. D., Durban, M. and Eilers, P. H. C. (2004b). Smoothing and forecasting mortality rates, *Statistical Modelling,* **4***, 279–298.*

Doll, R. and Hill, A. B. (1954). The mortality of doctors in relation to their smoking habits: a preliminary report, *British Medical Journal 1954* **ii** *1451–1455.*

Doll, R., Peto, R., Boreham, J. and Sutherland, I. (2004). Mortality in relation to smoking: 50 years' observations on male British doctors, *British Medical Journal 2004* **328***, 1519–.*

Durban, M., Currie, I. D. and Eilers, P. H. C. (2002). Using P-splines to smooth two-dimensional Poisson data, *Proceedings of 17th International Workshop on Statistical Modelling, Chania, Crete, 207–214.*

Eilers, P. H. C. and Marx, B. D. (1996). Flexible smoothing with *B*-splines and penalties, *Statist. Sci.,* **11***, 89-121.*

ERA (2003). Enrolled Actuaries Report, *Vol 28, No 4, Winter 2003.*

GAD (Government Actuary's Department) (2001). National population projections: review of methodology for projecting mortality, *NSQR Series No. 8.*

Gavrilov, L. A. and Gavrilova, N. S. (1991). Biology of life span: a quantitative approach, .

Gompertz, B. (1825). The nature of the function expressive of the law of human mortality, *Philosophical Transactions of the Royal Society.*

Goss, S. C. (2005). Testimony before the Subcommittee on Social Security of the House Committee on Ways and Means, 24 May 2005, *Statement of Stephen C. Goss, Chief Actuary, Social Security Administration.*

Hayflick, L. (2000). The future of ageing, *Nature,* **408** *(6809), 267–269.*

Human Mortality Database (2007). HMD, *University of California, Berkeley (U.S.A.), and Max Planck Institute for Demographic Research (Germany). URL www.mortality.org.*

Hunink, M. G. M., Goldman, L. and Tosteson, A. N. A. (1997). The recent decline in mortality from coronary heart disease, 1980–1990, *Journal of the American Medical Association,* **277** *(7), 535–541.*

Kirkby, J. G. and Currie, I. D. (2006). Modelling mortality data on the Lexis diagram, *Proceedings of 21st International Workshop on Statistical Modelling, Galway, 279–285.*

Kuulasmaa, K., Tunstall-Pedoe, H. and Dobson, A. (2000). Estimation of contribution of changes in classic risk factors to trends in coronary-event rates across the WHO MONICA Project populations, *The Lancet* **355***, 675–687.*

Lickint, F. (1929). Tabak und Tabakrauch als ätiologischer Factor des Carcinoms, *Zeitschrift für Krebsforschung 1929* **30***, 349–365.*

McCullagh, P. and Nelder, J. A. (1989). Generalized linear models, *2nd ed., Monographs on Statistics and Applied Probability 37, Chapman and Hall, London.*

Nolte, E. and McKee, M. (2004). Does health care save lives? Avoidable mortality revisited, *The Nuffield Trust.*

OECD (2005). OECD health data 2005: how does the United States compare?, *OECD.*

ONS (2006). Office for National Statistics website, *URL www.statistics.gov.uk.*

Philips, L. (1990). Hanging on the metaphone, *Computer Language, 1990,* **7** *(12), 39–43.*

Proctor, R. N. (2001). Commentary: Schairer and Schoeniger's forgotten tobacco epidemiology and the Nazi quest for racial purity, *International Journal of Epidemiology 2001* **30***, 31–34.*

R Development Core Team (2004). R: A language and environment for statistical computing, *R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL www.r-project.org.*

Renshaw, A. E. (1991). Actuarial graduation practice and generalised linear and non-linear models, *Journal of Institute of Actuaries 118, 295-312.*

Richards, S. J. and Jones, G. L. (2004). Financial aspects of longevity risk, *Staple Inn Actuarial Society.*

Richards, S. J., Kirkby, J. G. and Currie, I. D. (2006). The importance of year of birth in two-dimensional mortality data, *British Actuarial Journal,* **12***, I, 5–61.*

Rodu, B. (2004). Swedish tobacco use: smoking, smokeless, and history, *American Council on Science and Health.*

Rodu, B. and Cole, P. (2004). The burden of mortality from smoking: comparing Sweden with other countries in the European Union, *European Journal of Epidemiology,* **19***, 129-131.*

Schairer, E. and Schöniger, E. (1943). Lungenkrebs und Tabakverbrauch, *Zeitschrift für Krebsforschung, 1943* **54***, 261–269.*

Schwarz, G. (1978). Estimating the dimension of a model, *Annals of Statistics,* **6** *(2), 461–464.*

Shojania, K. G. and Burton, E. C. (2004). The persistent value of the autopsy, *American Family Physician.*

Shojania, K. G., Burton, E. C., McDonald, K. M. and Goldman, L. (2002). The autopsy as an outcome and performance measure, *Evidence Report/Technology Assessment. Number 58. Agency for Healthcare Research and Quality Publication No. 03-E002.*

Tunstall-Pedoe, H., Vanuzzo, D. and Hobbs, M. (2000). Estimation of contribution of changes in coronary care to improving survival, event rates and coronary heart-disease mortality across the WHO MONICA Project populations, *The Lancet,* **355***, 688–700.*

Unal, B., Critchley, J. A. and Capewell, S. (2004). Explaining the decline in coronary heart-disease mortality in England and Wales between 1981 and 2000, *Circulation,* **109***, 1101–1107.*

Unal, B., Critchley, J. A. and Capewell, S. (2005). Life-years gained from modern cardiological treatments and population risk-factor changes in England and Wales, 1981–2000, *American Journal of Public Health,* **95***, 1, 103–107.*

Willets, R. C., Gallop, A. P., Leandro, P. A., Lu, J. L. C., Macdonald, A. S., Miller, K. A., Richards, S. J., Robjohns, N., Ryan, J. P. and Waters, H. R. (2004). Longevity in the 21st century, *British Actuarial Journal,* **10***, IV, 695–898.*

Willets, R. C. (2004). The cohort effect: insights and explanations, *British Actuarial Journal,* **10***, IV, 833–877.*

WHO (2007). World Health Organisation website, *URL www.who.int.*