Supporting paper to COMEAP 2010 report:
"The Mortality Effects of Long Term Exposure to Particulate Air Pollution in the United Kingdom"

TECHNICAL ASPECTS OF LIFE TABLE ANALYSES

Brian G Miller and J Fintan Hurley

TECHNICAL ASPECTS OF LIFE TABLE ANALYSES

1 Core concepts of life table analysis

To perform a complete life-table prediction of impacts on mortality, we start with data comprising age-specific populations e_i and age-specific all-cause mortality hazard rates h_i. As we have often done, for the baseline we assume that the same hazards will apply in all future years, and that the numbers of viable new births in future years will equal that in 2008, which is taken as e_0. These assumptions allow us to populate a baseline input table such as shown in Figure 2.3 (all Figure references refer to those in the main report).

Using the formulae of the next section, survival curves are calculated down each distinct diagonal in the matrix, and these dictate how many deaths d_{ij} are expected in each cell, and how many life-years y_{ij} each cell will contribute, as in Figure 2.5.

The scenario for the impact of the envisaged policy can be constructed by altering the $h_{i(j)}$ to $h'_{ij} = f_{ij}*h_{i(j)}$. We have referred to the f_{ij} as 'impact factors'; they can differ by age (e.g. restricted to age 30+, so that $f_{ij} = 1$ for all i<30), and/or by calendar year (e.g. to allow for cessation lag), and by any combination of the two axes. The changes can be made permanent or acting for a finite period only. In the examples in the main report, we are quantifying the impact of a change in hazard rates taking place in 2008. The life-table calculations are be carried out on the diagonals of the matrix of h'_{ij}, to give corresponding matrices of d'_{ii} and y'_{ij}.

The pattern of the impacts is then calculated as the difference between the baseline and impacted versions of Figure 2.5, in terms either of differences in deaths (d'_{ij} - d_{ij}) or life years (y'_{ij} - y_{ij}) and these can be summed over any combinations of age and calendar years to produce summary impacts.

2 Formulae for life table calculations

All quantitative HIAs involve predicting sets of future rates under a policy-impacted scenario and comparing the outcomes with those predicted in the absence of the policy. In the case of mortality outcomes, data on current or recent rates of mortality are used. Typically these will consist of sex- and age-specific mid-year population estimates, and corresponding numbers of deaths; the ratios of deaths to population numbers are considered estimates of average annual hazard rates.

The examples in this report use population and death data from 2008 for England and Wales, Scotland and Northern Ireland by sex and by 1-year age groups (excluding neonatal deaths). These are shown in Figure 2.1. Their pattern is typical: perinatal mortality showing at age 0, followed by lower rates in childhood; a minor explosion in the late teens, greater in males, largely due to accidental causes; then an approximately exponential rise (linear on the log scale) in adulthood, with hazards for males consistently higher than for females.

It is common in life tables for the last age group to be open-ended, as in the 90+ groups here: the overall hazard rate is thus a weighted average of the rates for the different ages, which are likely still to be increasing with age. Here the pattern of weights will be different between the sexes, because relatively more females survive past the 90^{th} birthday.

These are the raw ingredients for calculations of life expectancy, which are conventionally laid out in a life table. Given a table of age-specific hazard rates h_i in one-year age groups (i=0, 1, …), the probability of survival from the i^{th} birthday to the $(i+1)^{th}$ is estimated by

$$s_{i+1} = \frac{(2-h_i)}{(2+h_i)}$$

Cumulative survival from birth to each birthday $k + 1$, denoted by ${}_0S_{k+1}$, is then calculated (denoting by Π the operator that multiplies together a set of values) as

$${}_0S_{k+1} = \prod_{i=0}^{k} s_i = \prod_{i=0}^{k} \frac{(2-h_i)}{(2+h_i)}$$

Joining these values describes a fine approximation to the underlying continuous survival function. Life expectancy $E(L)$, in units of 'life-years', is the area under this function, and we obtain (using Σ to denote addition)

$$E(L) = \sum_{j=0}^{A} 0.5 \times ({}_0S_j + {}_0S_{j+1})$$

where ${}_0S_0$=1.0, and A is the highest age achieved in the population so that ${}_0S_{A+1}$=0 (equivalent here to h_A=2.0). We therefore see that the survival function and the life expectancy it implies are defined uniquely by the set of hazard rates.

This equation estimates the life expectancy from birth, but policy measures will affect all members of a population, with a complete range of achieved ages, and we need to estimate the effects on all of those affected. ${}_0S_a$ is the proportion of the original population surviving to the a^{th} birthday, and we may estimate remaining expected life given an achieved age a, $E(L ¦ a)$, by adapting the above formula to

$$E(L ¦ a) = \frac{1}{{}_0S_a} \left\{ \sum_{j=a}^{A} 0.5 \times ({}_0S_j + {}_0S_{j+1}) \right\}$$

If we denote by ${}_aS_k$ the proportion of those achieving age a who then survive to their k^{th} birthday, we see that ${}_aS_k = {}_0S_k/{}_0S_a$, and the above can be written as

$$E(L ¦ a) = \sum_{j=a}^{A} 0.5 \times ({}_aS_j + {}_aS_{j+1})$$

The above formulae are appropriate when every age-group is one year wide, but not in the common situation of an open-ended last interval. In that case, however, the estimated contribution of life-years from the open-ended interval beginning at age a is given by ${}_0S_A/h_A$, the proportion surviving to the start of the interval divided by their subsequent hazard rate.

Thus we have

$$E(L) = \sum_{j=0}^{A-1} 0.5 \times ({}_0S_j + {}_0S_{j+1}) + {}_0S_A / h_A$$

and, conditional on reaching age a,

$$E(L ¦ a) = \frac{1}{{}_0S_a} \left\{ \sum_{j=a}^{A-1} 0.5 \times ({}_0S_j + {}_0S_{j+1}) + {}_0S_A / h_A \right\}$$

or

$$E(L \mid a) = \sum_{j=a}^{A-1} 0.5 \times ({}_aS_j + {}_aS_{j+1}) + {}_aS_A / h_A$$

We have implemented these calculations (including up to 6 groups of causes of death) in a suite of spreadsheets, IOMLIFET. More details are in Miller and Hurley (2006). The approach to the one-ended interval in IOMLIFET is slightly different from that described above, where the objective is simply to estimate life expectancy. Because we are particularly interested in the time course of the impacts, we have chosen simply to use the pooled hazard rate, at all ages 90 to 105 inclusive, so that the mortality at ages above 90 is spread in an approximate manner over 16 years.