



WORKING FOR A HEALTHY FUTURE

July 2011

IOMLIFET version 2011 Spreadsheets for life-table calculations

Dr Brian G Miller



WORLD HEALTH ORGANISATION
COLLABORATING CENTRE
FOR OCCUPATIONAL HEALTH

RESEARCH CONSULTING SERVICES

Multi-disciplinary specialists in Occupational and Environmental Health and Hygiene

www.iom-world.org

IOMLIFET VERSION 2011

SPREADSHEETS FOR LIFE-TABLE CALCULATIONS

The package IOMLIFET contains MS Excel spreadsheets set up to perform life-table calculations to quantify the predicted impacts of altering future mortality hazard rates. The need for these calculations arose from considerations of the impacts on mortality of proposed changes in the airborne concentrations of particulate air pollution, e.g. PM_{2.5} or PM₁₀, but the method is not specific to air pollution, and is applicable in any situation where it is believed that an action will impact on mortality rates.

The spreadsheets have been set up with example mortality data for England and Wales in the year 2008, separately by sex. Neonatal deaths (i.e. those in the first month after birth) have been excluded. We anticipate that users in other countries or areas will wish to use similar data, for chosen target populations, from their own authorities.

Prediction for other populations will need first to populate the sheets with numbers of deaths and mid-year populations (or just hazard rates) from your own local sources. It is often necessary to deal with the situation where the last years of age are grouped: there are a number of ways you can deal with this; all, of course, approximations. For simple life expectancy, there is a standard solution for dealing with the last, open-ended age group e.g. 90+. For the full 2-dimensional IOMLIFET sheets, our version assumes the hazards are the same for ages 90-105 inclusive, so as to spread the deaths out in time rather than assigning them to one calendar year; we set survival after age 105 to zero.

All the sheets are set up to work in 1-year age groups and time periods. The sheets could be adapted to work in 5-year groups, but an easy alternative is to calculate the average mortality rate over the 5-year group and use it for each 1-year group it contains.

The spreadsheets for calculating life expectancy and for the full IOMLIFET life-table calculations come in two versions, depending on whether you want to model impacts on all-cause hazards or on specific causes. For the cause-specific approach, you can define the cause-groups in any way you like, provided they're mutually exclusive and exhaustive, i.e. add up to all deaths. If you're only interested in effects on all-cause mortality, then of course that's all you need to obtain, and you can use the all-cause versions.

More information on what the spreadsheets do, and how and why, is available in an IOM Research Report TM/06/01; this note should be read in conjunction with the report, which is distributed as a PDF with the spreadsheets and may also be downloaded free from

http://www.iom-world.org/pubs/IOM_TM0601.pdf

Simple life expectancy from all-cause mortality

The spreadsheets in the folder **Life Expect All-Cause** are used to carry out simple life-table calculations of life expectancy for a cohort, starting from all-cause baseline hazard rates; and the change in life expectancy following an impact on those hazards. I have supplied two, one each for 'Males' and 'Females', because mortality rates differ, and are customarily reported, by age and sex. For whole-population rates, you could use either.

The worksheet 'Baseline' does life-table calculations from the hazards as supplied. The principal outputs are in the last two columns; the second last is the expected remaining life given each starting age: for age 0 this is the life expectancy for a birth cohort experiencing these hazard rates as its members age. The last column adds this to the achieved age, so giving the expected total life length given an achieved age.

Some textbooks show these calculations applied to a notional birth cohort of say 100,000 imaginary persons. In the days of hand calculation this avoided a lot of problems keeping track of decimal places, but these days it's not necessary. Life expectancy is a function only of the hazard rates (or equivalently, of the survival probabilities), not of the cohort size, so I've left it out. (If you want, you could multiply the cumulative survival column by 100,000 to show numbers per 100,000 surviving out of each age band.)

The 'Impacted' worksheet performs the same calculations after applying impacts to the hazard rates. The impact factors are multipliers. If, for example, you are interested a 1% reduction in hazards, the impact factor is $(100-1)/100 = 0.99$. There is complete flexibility in how the impact factors can differ by age. It is common in HIAs for air pollution to assume changes to hazards for those aged 30 and above, because the existing cohort studies have no data on effects in children and younger people. The supplied example assumes a risk ratio coefficient of 1.06 per $10\mu\text{g}/\text{m}^3$ change in annual population-weighted average airborne concentration of $\text{PM}_{2.5}$, and a log-linear response curve. Calculations are done for a $1\mu\text{g}/\text{m}^3$ reduction in pollution, leading to an impact factor f given by

$$f = 1.06^{(-1/10)} = 0.99419$$

and this factor is applied to the over-30s only. This sheet also calculates the change in the life expectancies predicted by the baseline and impacted rates, expressed both in years and in days. These are shown at the foot of the columns. The difference in the predicted time patterns of deaths is shown in a graph. Since the whole cohort is extinguished in either case, the area enclosed by the deaths graph has a zero sum.

Life expectancy given cause-specific changes

Extending the life-expectancy calculations to allow for effects on only some cause(s) of death is straightforward, because mortality risks, as represented by hazard rates, are additive. Given hazard rates for different causes, these can be impacted separately and recombined.

The spreadsheets in the folder **Life Expect Multi-Cause** can be used to carry out calculations of life expectancy for a cohort, starting from cause-specific hazard rates; these should be mutually exclusive (i.e. no double-counting) and exhaustive (i.e. include all causes of death). The supplied spreadsheets allow for up to 6 separate cause groups; to work with fewer, set the unused columns to 0.

Our example uses data for 5 specific causes believed to be affected by air pollution, plus the remainder, that is

- lung cancer
- ischaemic heart disease
- cerebrovascular causes
- other cardiovascular causes
- non-malignant respiratory disease
- all other causes

Again, there are two versions, one each for 'Males' and 'Females'. For whole-population rates, you could use either. The Multi-Cause version has a separate worksheet for setting the impact factors for each cause. Once this is done, the spreadsheets recombine the impacted hazards and calculate the impacted life expectancy results in the same way as for the all-cause case.

IOMLIFET 2008 ALL-CAUSE

The spreadsheets in this folder extend the life-table calculations to allow for the fact that policy makers are interested in effects on whole populations, with their range of achieved ages at the time changes are made. In order to calculate impacts, it is necessary to separate the dimensions of age and calendar time, and life-table calculations are performed down diagonals of 2-dimensional Excel worksheets. There are again separate workbooks for males and females; here the baseline and impacted scenarios are also separate.

NB: If you are using rates combining the sexes, you should program these in the Male versions of the spreadsheets, so that the Impacts workbook picks up the correct parameters.

The first step is to obtain and organise all-cause hazard rates; here I've used those from England and Wales, for 2008. Note that inserting the year 2008 in cell A1 of the Inputs sheet is necessary, because the other sheets use this to set up the range of years.

I have truncated the calculations after age 105. Where rates are available only up to an open-ended interval such as ages 90+, I recommend filling the rates from 90 to 105 with the pooled rate for 90+.

The next step is to decide how to populate the baseline hazards for future years (here 2009 and above). In this example I've simply replicated the 2008 hazards forward into all future years, but lots of alternative assumptions are possible, e.g. if you assume that

death rates would reduce year-on-year anyway (as they have been doing in Britain and many other countries).

An additional input assumption concerns the sizes of initial age-specific cohorts and new birth cohorts in future years, in the sheets labelled 'start popns'. For the initial populations, we have reused the given mid-year populations, but this is not necessary, e.g. if rates from one population are being applied to a different population. We have also assumed constant future birth rates, but it might be more realistic to make the birth rate dynamic, e.g. as a function of the number of women of child-bearing age in the previous year.

Once the hazard and population data are in place, all the calculations will follow through the remaining sheets, calculating expected numbers of deaths and numbers of life-years in each combination of age and calendar year. Row 3 of the last sheet shows the life expectancy of each cohort (inaccurately for those cohorts not fully followed up at the end of the matrix). Then save the file to record the results.

The Male_Impacted and Female_Impacted workbooks take the baseline hazards from the appropriate workbook, and apply impact factors to them. As before, the impact factor for a 1% reduction will be $(100-1)/100 = 0.99$. The simple example included uses the same factor as before for a $1\mu\text{g}/\text{m}^3$ reduction in $\text{PM}_{2.5}$, i.e. 0.99419, at ages 30+.

We often envisage the impacts as the result of a change, and define the 'current year' as that in which the change is made. Obviously, the change cannot have any effect in years before the 'current', but we cannot assume that the impact will be full and immediate in that year. Considerations of lag or latency may mean that the impact factors lag behind the change, or that the impacts phase in gradually over several years. In addition, it may be desired to predict impacts for a sequential series of changes, and any or all of these could have impacts that are not immediate. The transition from changes to impacts thus requires careful thought, but the two-dimensional layout of the Impact Factors sheet can accommodate any pattern of age- and/or year-specific (e.g. lagged) impacts.

When the Male_Impacted and Female_Impacted workbooks are opened, you will need to allow the inputs to be updated from the Baseline workbooks. Once the impact factors are in place, the remaining life-table calculations are carried out automatically. Then save each file to record and store the results.

The workbook 'Impacts' is designed to summarise various aspects of the difference in survival patterns between the baseline and impacted scenarios. The various sheets within it contain the differences between the scenarios by age and calendar year for each sex (and all these values are available for summary in ways additional to the defaults provided; this workbook and the 'Summary' one – see below – could be adapted if different summaries are needed.)

When opened, 'Impacts' will ask whether the file is to be updated, to which the answer is 'Yes' to incorporate the results of a recent run of the life-tables. The Summary worksheet contains a lot of information, controlled by the input variables coloured green. The values in the first four rows are descriptive. These should be entered

carefully and explicitly, so that it is possible to see later what this summary represents. The remaining values control various functions of the sheet: they define the 'current' year and the range of 'follow-up' years over which the results are summarised. Results are shown separately for follow-up in the population existing in the 'current' year, and for that population augmented by new birth cohorts in subsequent years.

The summaries can also produce valuations of the impacts, with or without discounting. LY base value is the value of a life-year, followed by the percentage discount rate, which can be set to 0 for undiscounted results.

The various graphs show the time-pattern of the changes in numbers of deaths and life-years over the follow-up period, with and without discounting.

Saving results

Because the IOMLIFET spreadsheets retain all the intermediate calculations in a matrix broken down by age and calendar year, they are large files. If you are doing lots of different scenarios, they'll build up. One option is to save the Impacts files in a ZIP file, which typically reduces them to about 25% of their original size. It will be necessary to use 'save as' to rename each one with some mnemonic or reminder of which run was which, or else number or label them and keep a catalogue.

If the summary sheet gives all the information needed, and you don't need to keep the age*year breakdown, the Summary file can capture just that sheet. Open and update it, and again use 'save as' to give the file a new name.

***** A WARNING *****

Since the Impacts and Summary files are designed to update from previous sheets, when you open them you'll be invited to update the contents. If you do, and the other sheets have recently run different scenarios, the results of the original scenario will be overwritten.

TO AVOID THIS, you must navigate to Edit/Links and break the links before using 'Save As'. This freezes the current values in the file, and saving it will save the results, but not the links.

IOMLIFET 2008 MULTI-CAUSE

For your own application, you may wish to apply cause-specific impacts; for example, you may want to assess a scenario in which air pollution may impact upon cardio-respiratory hazards, but not on others.

For cause specific impacts, our example uses data for 5 specific causes plus the remainder, that is

- lung cancer
- ischaemic heart disease

- cerebrovascular causes
- other cardiovascular causes
- non-malignant respiratory disease
- all other causes

The multi-cause version stores the arrays of hazard rates for these separately, and allows each to be impacted by a different set of factors (or 1.0 for no impact) before being recombined into an all-cause hazard. This is permissible because hazards for different causes are essentially additive.

If you have fewer than 6 cause groups, then you can set the hazards in the unused sheets to 0, and the impact factors for those phantom causes to 1.

In all other respects the spreadsheets operate exactly as for the all-cause version. If calculating without distinguishing the sexes, you should use Male versions.

QALYs and DALYS

This implementation does not include QALY or DALY weightings, but I envisage that step as a further sheet of age- and year-specific weights that would be applied to the YOL values and then summarised, similarly to what is already implemented for valuation.

Disclaimer

I hope you can get some value from these sheets. I can't guarantee that they are free from errors. (Please let me know if you spot any.) But they and others like them have been useful to us and our collaborators in the past. You're welcome to contact me if any of this is not clear. I'll be happy to give the odd bit of advice on how to use them, or to discuss possible collaborative working if your needs are extensive.

Acknowledgement

These spreadsheets have been developed with funding from the UK Department of Health, and are distributed free of charge. If you use them in your work, I'd appreciate an acknowledgement, preferably by referencing the Research Report. If you're not doing this sort of work, please feel free to pass the files on to one of your colleagues who is interested. The most up-to-date files can be downloaded, in ZIP files, from

<http://www.iom-world.org/research/iomlifet.php>

Dr Brian G Miller
July 2011
Principal Epidemiologist
Institute of Occupational Medicine
Edinburgh EH14 4AP
Tel: +44 131 449 8044
Email: brian.miller@iom-world.org

HEAD OFFICE:

Research Avenue North,
Riccarton,
Edinburgh, EH14 4AP,
United Kingdom
Telephone: +44 (0)131 449 8000
Facsimile: +44 (0)131 449 8084

Email: iom@iom-world.org

Tapton Park Innovation Centre,
Brimington Road, Tapton,
Chesterfield, Derbyshire, S41 0TZ,
United Kingdom
Telephone: +44 (0)1246 557866
Facsimile: +44 (0)1246 551212

Research House Business Centre,
Fraser Road,
Perivale, Middlesex, UB6 7AQ,
United Kingdom
Telephone: +44 (0)208 537 3491/2
Facsimile: +44 (0)208 537 3493

Brookside Business Park,
Cold Meece,
Stone, Staffs, ST15 0RZ,
United Kingdom
Telephone: +44 (0)1785 764810
Facsimile: +44 (0)1785 764811