MESOTHELIOMA MORTALITY IN GREAT BRITAIN: ESTIMATING THE FUTURE BURDEN

December 2003

Mesothelioma mortality in Great Britain: estimating the future burden

Summary

Mesothelioma deaths in Great Britain continue to increase annually both for males and females. Statistical modelling of male mesothelioma deaths can be used to produce an estimate of the future peak number of mesothelioma deaths to males and females. Based on this methodology, the annual total number of mesothelioma deaths in Great Britain is estimated to peak at around 1950 to 2450 deaths some time between 2011 and 2015

Contents

1. Introduction

Mesothelioma is a rare form of cancer that principally affects the external lining of the lungs (pleura) and lower digestive tract (peritoneum). It has a strong association with exposure to asbestos dust, and the long latency period between first exposure to asbestos and the development and diagnosis of mesothelioma is seldom less than 15 years and can be as long as 60 years. Mesothelioma is almost always fatal; most affected die within 1 or 2 years of diagnosis.

Numbers of deaths and other statistics on mesothelioma in Great Britain are published annually by HSEs Epidemiology and Medical Statistics Unit (EMSU). These statistics are derived from a register maintained by EMSU (the mesothelioma register) which comprises all deaths where the cause of death on the death certificate mentioned the word 'mesothelioma'. Further information regarding HSEs asbestos related disease registers is available on the HSE statistics web pages.

This factsheet presents the results of the latest statistical modelling to predict the future burden of mesothelioma in Great Britain. Predictions of the scale and timing of the peak number of annual deaths are reported. The factsheet also describes the development of the modelling methodology used over the last few years.

Overview of Mesothelioma in Great Britain

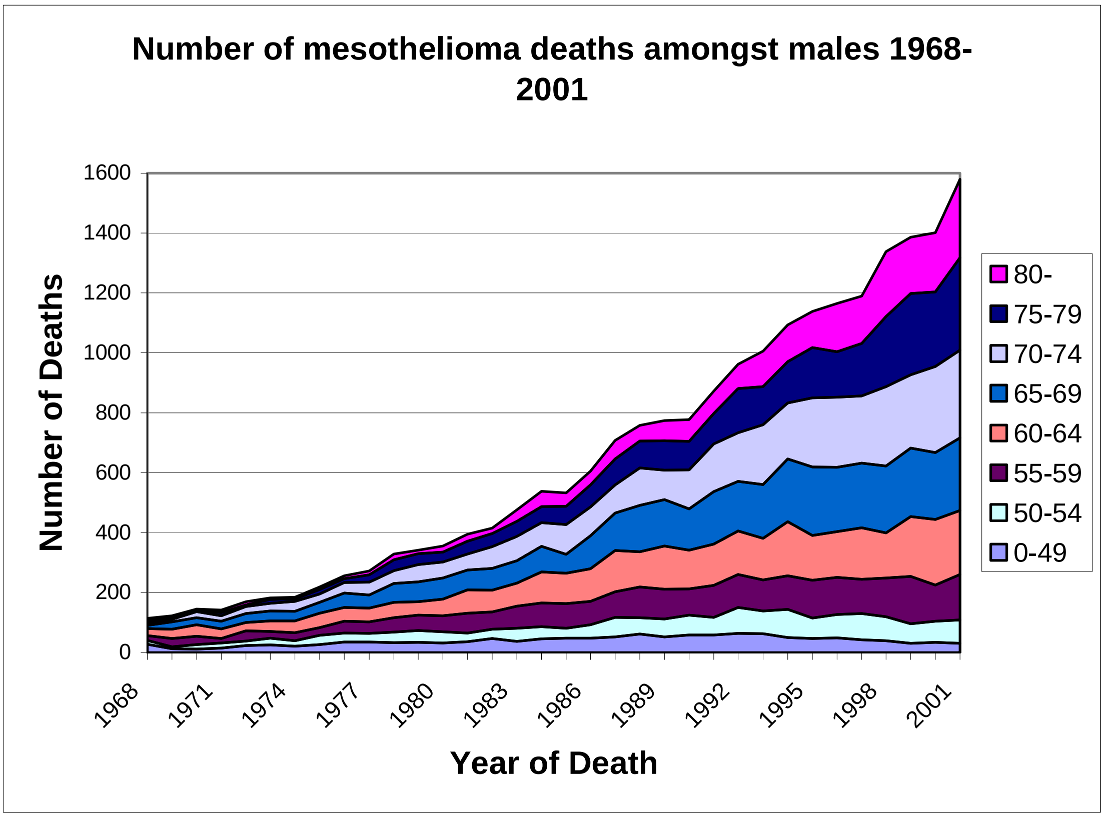
An overview of statistics from the mesothelioma register is available on the HSE statistics web pages. A brief summary is reproduced here to provide the context of the statistical modelling.

The annual number of mesothelioma deaths in Great Britain has risen fairly constantly over time from 153 in 1968 (the first complete year of data after the register of mesothelioma deaths was set up) to 1848 in 2001 (the latest year for which data are available). Because of the long latency period, much of the current burden of mesothelioma deaths is a result of heavy asbestos exposures in the past.

Although nearly all mesothelioma cases are caused by exposure to asbestos, a small number of deaths each year occur in people with no history of exposure. There is evidence (discussed later in this factsheet) to suggest that there are likely to be at least 50 of these so called spontaneous mesotheliomas each year in Great Britain.

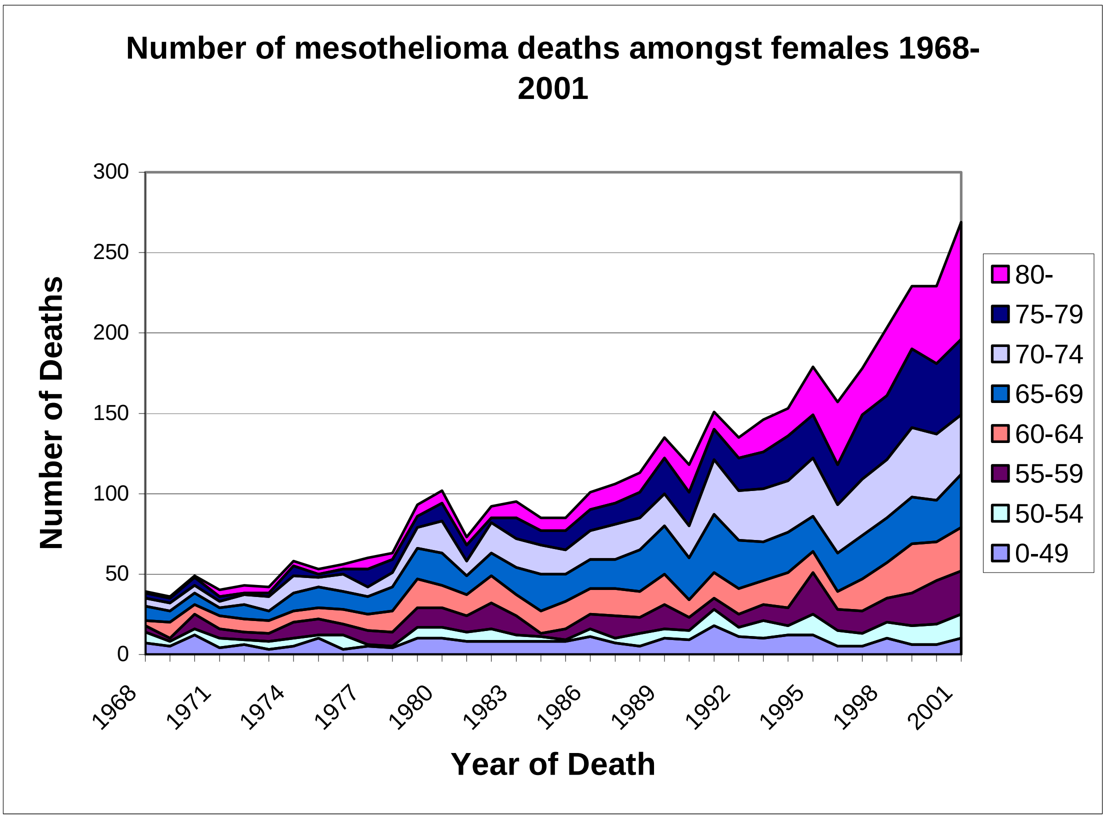
Most of those who die from mesothelioma each year are male: in 2001 there were 1579 male deaths, 85% of the total. Figure 1 shows the number of mesothelioma deaths amongst males by age group and year, since 1968. The highest concentration of deaths is amongst people who are in their 60s and 70s. There are very few deaths recorded among people who are under 50. After reaching a peak in the early 1990s, the number of deaths in this age group has been gradually falling since. Of all mesothelioma deaths amongst males, 99% are between the ages 20 and 89.

Figure 1



15% of deaths in 2001 were in females. The age distribution of deaths is similar amongst females to that for males and again 99% of deaths are between the ages of 20-89. There is more variability in the increasing trend due to the fewer number of deaths.

Figure 2



2. Methods

Statistical modelling of mesothelioma deaths

In order to predict the future course of the mesothelioma epidemic, some form of statistical model is required. Such models generally comprise two components. At the heart of a model is a mathematical expression to describe the relationship between the quantity of interest (here it is the annual number of male mesothelioma deaths), and certain variables which are considered important in determining its value – for example, the level of asbestos exposure, the age or year at which exposure occurred. Secondly, there is an error component in the model to take account of the statistical (random) variation in the data. The models described in this factsheet are known as Poisson regression models, since we are modelling annual counts of deaths – which are assumed follow a Poisson distribution. The modelling process also involves assessment of how well a particular mathematical expression predicts the observed number of deaths, taking into account this statistical variation. The calculation of projected future deaths then involves extrapolation by making assumptions about future values of the variables in the model.

The modelling of mesothelioma deaths in this way has only been carried out for males in Great Britain. The lower numbers of mesothelioma deaths in females do not provide an adequate basis for modelling because of the much larger degree of uncertainty that would be associated with any derived model.

Results of projections of future mesothelioma deaths in Great Britain produced by HSE and the Institute of Cancer Research were first published in 1995 [1]. These projections were based on a simple multiplicative birth cohort model in which mesothelioma risk is related to age and date of birth. The model included mesothelioma deaths in males aged 25-89 from 1968-1991.

An underlying assumption in this model was that the ratio of death rates at different ages is identical across all cohorts (equivalently that the ratio of rates between birth cohorts is the same at all ages). The mesothelioma data up to the end of the 1980s fitted this pattern quite closely, but as data accumulated through the 1990s it became clear that this was no longer the case – especially for the most recent birth cohorts. More details of this model are given in Appendix 1.

The inadequacy of the original model in the light of the most recent mesothelioma death data motivated HSE to develope the current model using a different approach. This model was first developed for the Regulatory Impact Assessment for the revised Control of Asbestos at Work Regulations [2] and is described in the next section.

The current model - based on inference of past collective dose

The current model is based on a number of assumptions.

Firstly, it is assumed that the population's total exposure to asbestos can be summarised in each year by a single estimate and that men's exposure in any year depends on their age. Then, that the relationship between the summarised exposure index and future deaths from mesothelioma takes the same form as is widely assumed for the relationship between asbestos exposure and mesothelioma risk over time at the individual level. This states that mesothelioma risk for an individual exposed at a given age is proportional to their cumulative exposure multiplied by approximately the second or third power of time since the start of their exposure [3].

Finally, terms to model a possible trend in the completeness of mesothelioma diagnosis, and the clearance of asbestos fibres from the lung were also included. Full details of the model are given in Appendix 2.

Given these assumptions, a theoretical exposure profile representing the past collective dose of asbestos can be determined which, when fed into the model, gives good agreement with the numbers of mesothelioma deaths actually observed to date. Future projections of mortality can then be made by making assumptions about the trend in the collective dose and by applying the predicted mesothelioma future rates to estimates of the future population of Britain.

Model Fitting

Fitting the model involves estimating the value of the various model parameters – ie the growth and decline rates to determine the shape of the theoretical exposure distribution, the relative exposure potentials at different ages, the power of time since first exposure, and the rate of clearance of asbestos fibres from the lung. The parameters were estimated using an iterative approach with the aim of minimising the model deviance, a measure of how well the fitted model compares to the observed number of mesothelioma deaths (assuming that the observed number follows a Poisson distribution).

The following table shows the values of the parameters in the fitted model.

Table 1: Parameter estimates for the current model

Power of time since start of exposure (k)	2.6	**Diagnostic trend (decrease in cases missed - % per year)**	5
Maximum exposure year	1967	**Lung clearance half-life (years)**	1000
Change in exposure index (%per year) in...		**Relative exposure potential by age group**	
1922	29	5 to 15	0.03
1932	6	16 to 19	0.21
1942	11	20 to 29	1.00
1952	9	30 to 39	1.24
1962	5	40 to 49	1.11
1967	0 (by definition)	50 to 59	0
1972	-14	60 to 69	0
1982	-39	70 to 79	0

Further assessment of the adequacy of the model to describe the observed pattern of mesothelioma deaths was carried out by comparing plots of observed and fitted (ie predicted by the model) numbers of mesothelioma deaths by age, year of death, and year of birth. Residuals plots (again for comparing observed and fitted numbers of deaths) by age and year of birth were also examined. These plots are give in Appendix 3.

3. Results

The estimate of the peak annual total number of mesothelioma deaths in Great Britain is composed of an estimate for males aged 20-89, based on the updated statistical model, and estimates for the numbers of females and individuals aged 90 or over based on the proportion of total mesothelioma deaths observed to date that these groups constitute:

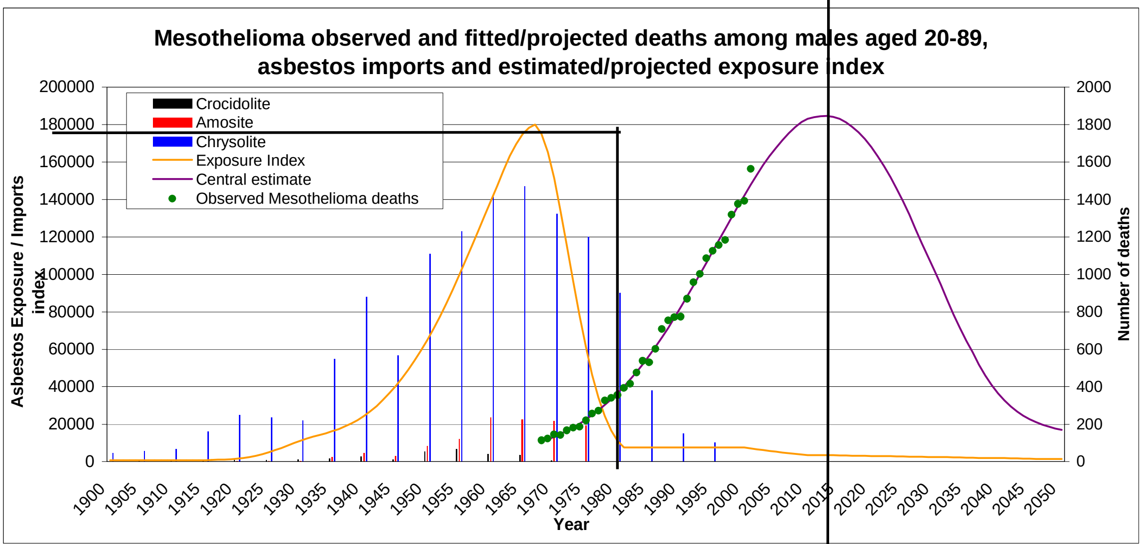
- The annual total number of mesothelioma deaths in Great Britain is predicted to peak at a level of 1950 to 2450 deaths during the period 2011 to 2015.

- The updated modelling suggests that annual total number of mesothelioma deaths to males aged 20-89 in Great Britain will peak at a level of 1650 to 2100 deaths during the period 2011 to 2015.

- The annual total number of mesothelioma deaths to females in Great Britain is predicted to peak at a level of 250 to 310 deaths, though this is more uncertain than the result for males based on the statistical model. Around 10 mesothelioma deaths among males aged 90 or over are expected to occur annually during the period of the peak.

Figure 3 shows observed mesothelioma deaths to males, asbestos imports into Great Britain together with the inferred past collective dose (exposure index), and the central estimate of projected course of the mesothelioma epidemic in males to 2050. The central estimate of peak mesothelioma deaths in males is 1850 deaths in year 2013.

Figure 3



4. Discussion

Model fit and parameter estimates

The parameter values obtained from the fitted model are given in Table 1 above. The fit is surprisingly good, given the simplifications inherent in this model. The model has an overall deviance of 225.80 on 182 degrees of freedom. Plots of fitted and observed deaths by age; year of death and year of birth show close agreement (see Appendix 3. The two-dimensional display of deviance residuals by five-year age and year of birth groupings also shows a reasonably uniform pattern with no strong clustering of residuals of the same sign, which would indicate inadequacies in the model. However, two birth cohorts (births in 1923-27 and 1938-42) show levels consistently above predicted levels.

The power of time from first exposure to asbestos (parameter k) is 2.6. This is in the middle of the range expected *a priori*. The year in which exposure was estimated to be a maximum was 1967, with a very steep (but poorly determined) reduction in exposure after this date.

The estimates of relative exposure potential at different ages imply that exposure is concentrated on the age group 20 to 49. Higher age groups are estimated by the model to have a zero exposure weight, but this is due to the lag from exposure to effect, which means that these parameters have only a small influence on the fit. Of more interest is the fact that the weight for exposure between the ages of 16 and 19 is estimated at less than a quarter of exposure at adult ages, and that exposure from ages 5 to 15 is estimated very much lower still (3% of adult level). These estimates support the conclusion drawn from other evidence (e.g. the comparison between male and female rates) that it is occupations, especially male occupations, which provide the main source of exposures.

The diagnostic trend term in the model is estimated at 5% - ie the number of undiagnosed cases occurring decreases by 5% per year. We have assumed that diagnosis was effectively complete at an arbitrarily chosen 98% in 1997. The estimate of 5% trend implies that in 1968, the start of our data series, diagnosis was about 90% complete. This seems a reasonably plausible conclusion. It has only a minor effect on the model.

The value given to the half-life of asbestos fibres in the lungs in the model is 1000 years. Thus the model effectively suggests that there is in fact no clearance of asbestos fibres from the lungs of exposed individuals. There is some evidence from other studies [4] that this may not be the case. A variant model, which assumes that asbestos fibres have a half-life of 16 years, had a best fit deviance of 246. If the model is fitted using data mesothelioma deaths to 2000 rather than to 2001, this second model gave a much closer fit to the data and was only slightly worse than the "no-clearance" model.

Exposure index

The model suggests that the total exposure fell rapidly in the ten years or so following the peak in 1967. However, the statistical uncertainty about the course of exposure grows rapidly from around the mid 1970s. Since there is a lag between asbestos exposure and the occurrence of mesothelioma, our observations of mortality up to 2001 tell us nothing (directly) about exposure levels since 1980. This means the last few terms in the model regarding the speed of decrease in asbestos exposure do not have any appreciable effect on the fit of the model and therefore cannot be assessed in relation to observed deaths.

There is no real basis for assuming this rate of decline in exposure will have continued beyond the late 1970s (resulting in trivial exposure well before 2000). The main driver for the decrease will have been the rapid reduction in initial processing of imported fibre into asbestos products and their installation. Once exposure has fallen to the level generated by continued routine building maintenance and demolition (and asbestos removal), the rate of total population exposure would be expected to be fairly constant. We have no good measurement-based evidence for knowing what this level is.

Any contribution of asbestos removal to the mesothelioma epidemic cannot yet be assessed. Asbestos removal did not develop as a specialised industry until 1980, and the latency is still too short for the development of mesothelioma amongst these workers to be measured.

Developing assumptions about exposure levels since 1980

In addition to the statistical uncertainty about exposures from the late 1970s onwards, there is other credible evidence to suggest that the assumption that asbestos exposure in GB has reduced to zero is unreasonable. The RIA for the revised Control of Asbestos at Work Regulations suggested that the current exposure to asbestos is approximately 4% of the peak. We can assess how far we can be reasonably confident (on the basis of this direct evidence) that exposure has fallen by examining the impact on the model fit of leveling off the assumed exposure at different points on its downward track. Table 2 records the effect on the overall fit and on the fit for birth cohorts from 1948, of levelling off exposure from different years in the 1970s and 1980s. It also shows the effect this has on the timing and size of the peak of the epidemic. The reason for looking separately at more recent cohorts is that they constitute the most critical group for future deaths. If the model fits these cohorts poorly, it is unlikely to be a reliable guide to future levels of mortality.

The deviance of the model is significantly worse if the minimum exposure year is set before 1978. In 1977 the difference in fit is not quite significant at the 1% level. Earlier than this and the difference in fit starts to become very substantial. Any sensible alterations in exposure after 1983 would have very little effect on the fit of the model. If the model is restricted to looking at the 1948 and later cohorts, the deviance is significantly worse if the minimum exposure year is set before 1980.

Table 2: Model fit and projections for levelling off exposure from different years in the 1970s and 1980s

Exposure levelled off in year	Deviance	p-value* for equivalence to best fit	p-value* for equivalence to best fit for cohorts from 1948 onwards	Predicted deaths in individuals born since 1948 (observed = 424)	Exposure index from levelling off year (% of 1967 peak value)	Year for peak deaths	Peak number of deaths
1973	274.57	P<0.001	P<0.001	562.60	54%	2031	2621
1974	253.22	P<0.001	P<0.001	529.50	43%	2023	2329
1975	239.78	P<0.001	P<0.001	503.53	34%	2019	2145
1976	232.15	0.012	0.003	484.44	26%	2017	2026
1977	228.28	0.12	0.014	471.27	19%	2015	1951
1978	226.54	0.41	0.036	462.74	14%	2014	1904
1979	225.86	P>0.5	0.060	457.58	9%	2013	1874
1980	225.68	P>0.5	0.077	454.66	6%	2013	1856
1981	225.68	P>0.5	0.088	453.14	4%	2013	1845
1982	225.73	P>0.5	0.094	452.40	2%	2012	1838

*Note: these are approximate P-values given that residuals show some departure from normality.

These results show that there is not necessarily any conflict between the fitted model (in which the exposure is predicted to fall to trivial levels well before the year 2000), and the estimated current exposure of 4% of the 1967 peak value from the RIA.

For the purposes of estimating the level and timing of the peak number of mesothelioma deaths, the model in which the exposure levels off at 4% in 1981 was chosen. Of the results in the above table that are consistent with the RIA current exposure assumption (ie that exposure levels off at 4% of the peak some time before 2000), this is the scenario in which the deviance is least affected. It may be that exposure was higher than this during the 1980s and 1990s, in which case we could have fixed an assumed exposure track between some higher exposure level (which would have occurred earlier than 1981) and the assumed 4% level for 2000. However, because of the lack of information about what these levels might have been, and given that this approach would have a larger effect on the deviance, this method was not used.

Estimating total males deaths to 2050

Table 1 illustrates that estimates of the timing and level of the peak number of mesothelioma deaths (from models where the fit is not significantly affected) are not sensitive to assumptions about exposure from around 1980 onwards. However, estimates of the overall numbers of deaths beyond the peak number are sensitive to such assumptions. In particular, estimates of the number of deaths beyond about 2020 start to become highly dependent on assumptions about exposure levels beyond the present day. Although this is the case we have estimated the overall burden of male deaths aged 20-89 up to the year 2050. However, the resulting figure should be treated as a rough guideline only.

The RIA for the amendment to the Control of Asbestos at Work Regulations 1987 estimated the number of deaths over the next 50 years due to past exposure and exposure beyond 2000. The analysis assumed that if the regulations were not amended, then exposure would decrease from the 2000 level in proportion to the predicted demolition rate of the generation of buildings most likely to contain asbestos materials. Based on the fact that in the future as these buildings get older the demolition rate is expected to increase, the estimated future exposure was also estimated to decrease more and more quickly with time. In this scenario the exposure would fall to below 2% of the peak 1967 level by 2050.

Given that amended regulations will now be implemented (the Control of Asbestos at Work Regulations 2002 will come into force in May 2004), future exposures are likely to be lower than the levels suggested above. However, it is impossible to determine how much lower these exposures will be. Full compliance with the amended regulations could theoretically reduce exposure to zero, however, this is unlikely to happen in practice. We have therefore assumed arbitrarily that the exposure will be reduced from 4.2% in 2000 so that from 2010 onwards the exposure will be half the level expected if the regulations had not been amended. This implies an exposure of 2% of the peak 1967 level in 2010 reducing to below 1% by 2050.

Under these assumptions the model predicts that following the peak number, annual male deaths will fall to around 170 deaths in 2050. The total number of male deaths to 2050 from the beginning of the epidemic would be around 77,000 under this scenario, with over 55,000 of those occurring from 2002 onwards (ie beyond the period for which we currently have observations).

Uncertainties

Observed deaths in 2001

The observed number of deaths in 2001 is greater than predicted by the model, and this result is statistically significant at the 5% level (p=0.014). However, the probability of the result being this far removed is not excessively low and may simply be due to chance. Alternatively, it could be a sign that the annual number of mesothelioma deaths is rising faster than predicted. This will be assessed in the future when numbers of mesothelioma deaths for succeeding years have been observed. The high 2001 value for observed deaths could have been accounted for in the model by introducing a massive increase in the exposure level in 1982. Although this resulted in a lower overall deviance, the model was effectively being overfitted and the scenario was considered to be very unrealistic. The result was therefore disregarded.

Coding of deaths to Revision 10 of the International Classification of Diseases (ICD10) was introduced in 2001 in England and Wales and 2000 in Scotland. Because of the way mesothelioma deaths are compiled it is unlikely that this will have affected the number of observed deaths in these years (ie particularly for 2001), though it remains a possibility.

Mesothelioma at ages greater than 80

A large (and increasing) proportion of the predicted future deaths are at ages 80 and above. This is driven both by the form of the model, and by the increasing survival to older ages in the population. Although the mesothelioma model used here fits observed mortality in occupational cohort studies quite well, it can reasonably be doubted whether the risk of mesothelioma increases indefinitely with time after exposure. The few occupational cohorts with very long follow-up all show eventual falls in mesothelioma rate. For this reason previous risk assessments have truncated their predictions at age 80 (although 80 is an arbitrary figure). Because the population at ages over 80 years is growing, any error in the model at these ages will be amplified when projecting into the future. Observed and expected deaths at ages 80-89 were examined to specifically check for any systematic over-prediction of the numbers of deaths in this age group. There was no evidence of any problem with the model.

Uncertainty range

The model gives a single value for the estimated peak number of mesothelioma deaths. Standard practice would be to quote a 95% confidence interval for the estimated peak value to give an indication of the uncertainty due to the statistical variation in the data. However, due to the iterative approach used for fitting the model, confidence intervals can only be found from an informal numerical search rather than analytically. We therefore adopted an alternative approach for devising an interval to represent the likely uncertainty range of results. In practice this range is likely to be narrower than a true 95% confidence interval. More details of this method are given in Appendix 4.

Assessments of the adequacy of the model

Residuals are a measure of the difference between observed deaths and numbers predicted by the model - known as “fitted values”. A model which fits the observed data well will have residuals which exhibit certain statistical properties. Analysis of the residuals from the final model in this case indicated some departure from these properties. Technical details of this are given in Appendix 4.

This fact does not necessarily put into doubt the validity of the model. However, a consequence is that judgments about the model made on the basis of the deviance – which includes the estimation of the uncertainty range described in the previous section – are less robust.

Model fit for individuals born most recently

There is some evidence that the fit of the model is less satisfactory for more recent birth cohorts. There is increased uncertainty in this area because the numbers of deaths are small, however, the fact that the model predicts fewer deaths than observed for those born most recently is fairly clear – see Figure 5a, Appendix 3.

One reason why this might be the case is that the model does not account for the fact that there are a small number of mesothelioma deaths each year where the individuals have no history of asbestos exposure – so called “spontaneous” or “background” mesotheliomas. The annual number of background cases is thought to remain fairly constant. Because the total number of mesothelioma deaths in the most recent birth cohorts is small, background cases may account for a substantial proportion of the total, whereas in earlier cohorts the number caused by asbestos will far exceed the background cases.

Subtracting an estimate of the number of background cases from the total observed deaths for each birth cohort and age group and then comparing the adjusted observed year totals with the original fitted values indicates that the model no longer consistently underestimates the values of the latest birth cohorts. This suggests that the fit of the model could be improved by taking into account the number of background cases. Such a modified model could potentially account for the lack of fit in the most recent birth cohorts. However, there is currently insufficient data to determine whether background cases do provide an adequate explanation or whether the beginnings of a departure from fit of the current model are being witnessed.

Scaling estimates of future deaths to include groups not covered by the model

Females are not included in the model; neither are males aged 90 and over. Thus, the estimate of the peak number of mesothelioma deaths from the model also excludes these groups. In order to estimate the total number of deaths among males and females of all ages we have rescaled the estimate as follows.

Firstly, of the 25800 mesothelioma deaths which have occurred during 1968-2001, in 175 cases the age at death was 90 or over. The peak number of male deaths aged 20-89 predicted by the model was multiplied by 25800/(25800-175) to estimate the peak number of deaths to males of all ages. This gives 1859 males deaths in year 2013.

Secondly, a simple linear regression analysis of annual female mesothelioma deaths against annual male deaths for 1968-2001 gave the following relationship:

$$F = 22.9 + 0.136M$$

Where, F = annual female mesothelioma deaths
M = annual male mesothelioma deaths.

Using this relationship to estimate the peak number of female mesothelioma deaths (at ages 20-89) from the peak number of males gives 276 female deaths in 2013.

Combining these estimates gives an estimated peak number of deaths of 2135 in year 2013. The lower and upper limits of the uncertainty range described earlier were rescaled in the same way. However, in reality these adjustments introduce additional uncertainty and so the range should be wider than stated.

Other models with alternative parameter values

During the course of model validation two alternative sets of parameter values with a similar deviance to the adopted model were identified. The parameter values, deviance and predicted peak for these models are given in Appendix 5. It is possible that one of these, or indeed some other model as yet unidentified, describes the mesothelioma epidemic more accurately than the one chosen. However, the alternative models both give similar projections to our central estimate.

5. References

1. Peto. J, Hodgson. J, Matthews. J, Jones. J. Continuing increase in mesothelioma mortality in Britain. The Lancet 1995; 345: 535-39.
2. Amendment to The Control of Asbestos at Work Regulations 1987 and ACOP; Regulatory Impact Assessment
3. Health Effects Institute (1991) Asbestos in Public and Commercial Buildings: A Literature Review and Synthesis of Current Knowledge. Health Effects Institute - Asbestos Research, Cambridge, MA.
4. Berry. G. Models for mesothelioma incidence following exposure to fibers in terms of timing and duration of exposure and biopersistence of the fibers. Inhalation Toxicology, 11:111-130, 1999.

Appendix 1 –Simple multiplicative age/birth cohort model

The simple multiplicative birth cohort model assumes that the annual mesothelioma rate for a particular age is given by the overall mesothelioma death rate for that age group multiplied by mesothelioma risk in the appropriate birth cohort:

Annual age specific mesothelioma death rate, $r_{ab}= k_a c_b$

Where, k_a = predicted age specific death rates (a=1 for age group 25-29, 2 for 30-34,…,13 for 85-89);
c_b = birth cohort specific risks relative to the 1943-48 cohort (b=1 for 1893-98, 2 for 1898-1903, … 13 for 1953-58). Thus, c_{11}=1.

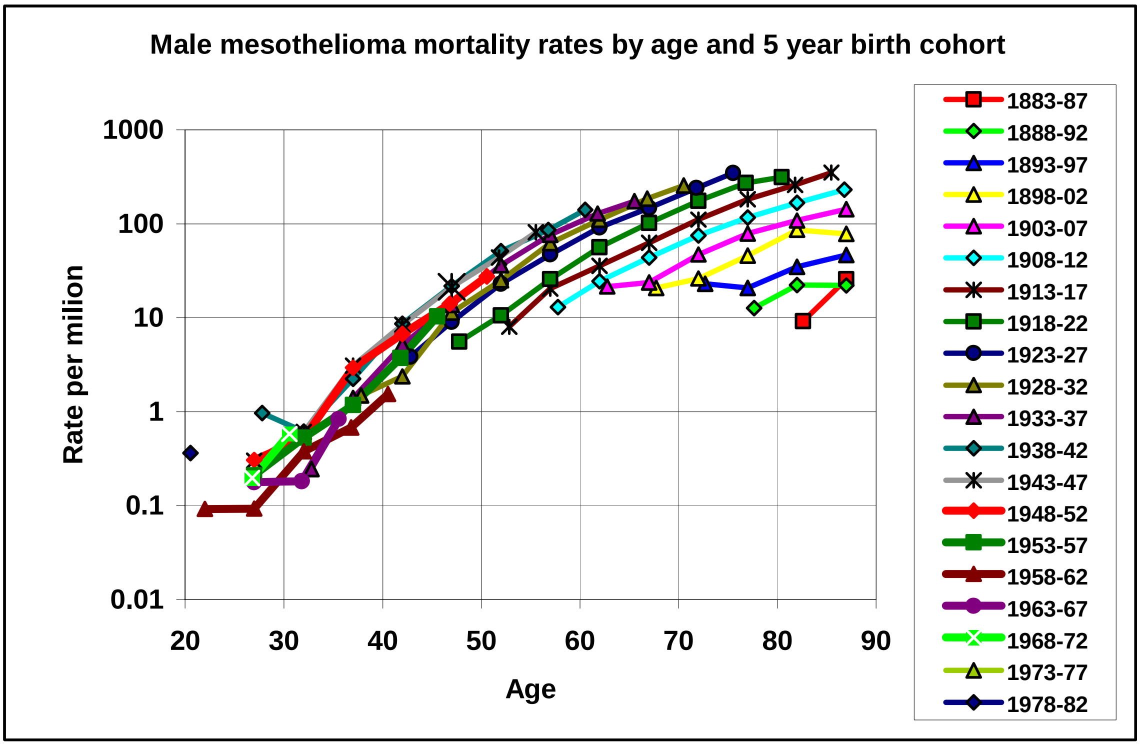
Given a good fit of this model to observed mesothelioma rates, projections of future rates can be made simply using age, and date of birth. These can then be translated into predictions of annual future numbers of deaths using projected population data.

The model was first fitted to data for observed mesothelioma deaths in males aged 25-89 from 1968 to 1991. The fit of the model suggested this was a sound basis for future predictions of the extent of the mesothelioma epidemic. Multiplying the death rates by the appropriate observed and projected population data allowed estimation of the number of mesothelioma deaths. Combining the deaths for all cohorts gave a peak of annual male mesothelioma deaths of between 2700 and 3300 deaths, in around the year 2020.

However, as stated when the model was first published, long-term projections are dependent on the assumption that the ratio of death rates at different ages is identical across all cohorts (or equivalently that the ratio of rates between birth cohorts is the same at all ages). This means that plots of mesothelioma rate by age for each birth cohort should all show the same pattern. However, if asbestos exposure fell sharply after around 1980, we would expect this not to be the case – since for men born after 1940 exposure will have largely ceased by age 40. At the time of publication, the fact that exposure had indeed fallen was supported by lower death rates in the latest birth cohorts, but there was no direct evidence of any change in the pattern of age dependence across cohorts. Figure 4 shows that incorporating death data to 2001 now suggests the pattern is not consistent across cohorts. The lines for some cohorts (notably the 1948-52 cohort) increase relatively less steeply with increasing age than for others. This suggests that the simple multiplicative model cannot accommodate the most recent data. In particular it suggests that the peak will be lower, and occur earlier, than predicted by this model.

A further reservation cited at the time of publication was the possibility that part of the increase in recorded death rates was an artefact of misdiagnosis. The existence of an increasing rate of accurate mesothelioma diagnosis would be statistically indistinguishable from a real increase in the actual number of deaths, but would cease as soon as reasonably complete diagnosis was achieved.

Figure 4


Male mesothelioma mortality rates by age and 5 year birth cohort
1000
100
10
1
0.1
0.01
20
30
40
50
60
70
80
90
Rate per milion
Age
1883-87
1888-92
1893-97
1898-02
1903-07
1908-12
1913-17
1918-22
1923-27
1928-32
1933-37
1938-42
1943-47
1948-52
1953-57
1958-62
1963-67
1968-72
1973-77
1978-82

Appendix 2 – The current model based on inference of past collective dose

The model attempts to explain the observed mesothelioma mortality in males (by year and by single year of age) in terms of a range of (unobserved) inputs. The exposure of men of a given age in a given year is assumed to be proportional to the product of two factors: a year dependent factor representing the total use of asbestos in that year, and an age factor representing the probability of contact with asbestos for a male of a given age.

Independent epidemiological evidence suggests that after a brief exposure to asbestos risk increases in proportion with a power of time probably in the range 2 to 3. This is possibly modified by the clearance of fibres from the lung, though the evidence for this is much more open.

A reasonable case can also be made that mesothelioma may have been under-diagnosed when records were first systematically kept in the late 1960s. A term representing increasing diagnostic completeness (as a function of time) is also included in the model.

Putting all these terms together we arrive at the following formulae for the fitted/predicted number of mesotheliomas at age A, in year T ($F_{A,T}$)

$$F_{A,T} = \left[\sum_{l=0}^{A+1} W_{A-l} \, D_{T-l} \, \{l+1-L\}^k 0.5^{\frac{l}{H}} \right] Dx_T \, P_{A,T} \, \frac{M}{F}$$

Where

$P_{A,T}$ person years for age A in year T;

D_T overall population exposure in year T;

Dx_T proportion of occurring mesotheliomas diagnosed in year T;

W_A age specific exposure potential at age A;

L lag period (in years) before effect starts;

H half life (in years) for clearance of asbestos from lungs;

k exponent of time modelling increase of risk with increasing time from exposure;

M total observed mesotheliomas;

The content of the { } is set to zero when negative; and

$$F = \sum_{A,T} \left[\sum_{l=0}^{A+1} W_{A-l} \, D_{T-l} \, \{l+1-L\}^k 0.5^{\frac{l}{H}} \right] Dx_T \, P_{A,T}$$

The outer summation being taken over all values of A and T for which there are observations.

The summations indexed by *l* represent the cumulated effects at age A of the hypothesised exposures at each earlier age. *l* indexes years lagged from the risk year.

The overall goodness of fit is assessed by comparing observed and expected numbers of deaths aggregated into cells defined by five-year groups for age, year of birth and year of death. The observed number is assumed to be Poisson distributed, so the deviance is the appropriate measure of discrepancy.

The parameterisation of the input factors is implemented as follows:

Age-specific exposure potential

Arbitrarily set to a baseline of 1 between age 20 and29. Other values set for ages 0 to 4 (pre-school); 5 to 15 (school age); 16 to 19 (school/work transition); 30 to 39, 40 to 49, 50 to 64 (work/retirement transition); and 65 plus.

General exposure level (conceptually perhaps best thought of as total asbestos fibres inhaled in year)

Parameterised by choice of a maximum year with an arbitrarily chosen value and a series of growth rates prior to that maximum and of declined rates after the maximum. Growth rates are chosen at maximum +/- 5 years and subsequently at 10-year intervals moving away from the maximum. Growth rates for intermediate years are fixed by linear interpolation between these chosen values. The spreadsheet has the facility to substitute the actual track of asbestos imports (either total or of specific fibre types) in place of this constructed index of hypothetical exposure. In models fitted so far these give much worse overall fits than the best hypothetical exposure tracks.

Diagnostic trend

Diagnosis assumed to be quasi-complete in 1997 at 98%. The diagnostic trend is parameterised as the annual percentage increase in the number of missed cases as one goes back in time from this point. Once the proportion of cases diagnosed has fallen to 50%, the parameterisation switches to (the same) percentage decrease in cases diagnosed. In fact, we believe it is implausible that the overall diagnostic efficiency was this low within relevant period.

Clearance of fibres from the lung

This is assumed to be an exponential decline and is parameterised by its half-life (ie the number of years it takes for the lung content to fall to half its initial level). A value of 1,000 has been taken which represents no clearance.

Appendix 3 - Graphical representations of the current model

Figure 5(a)

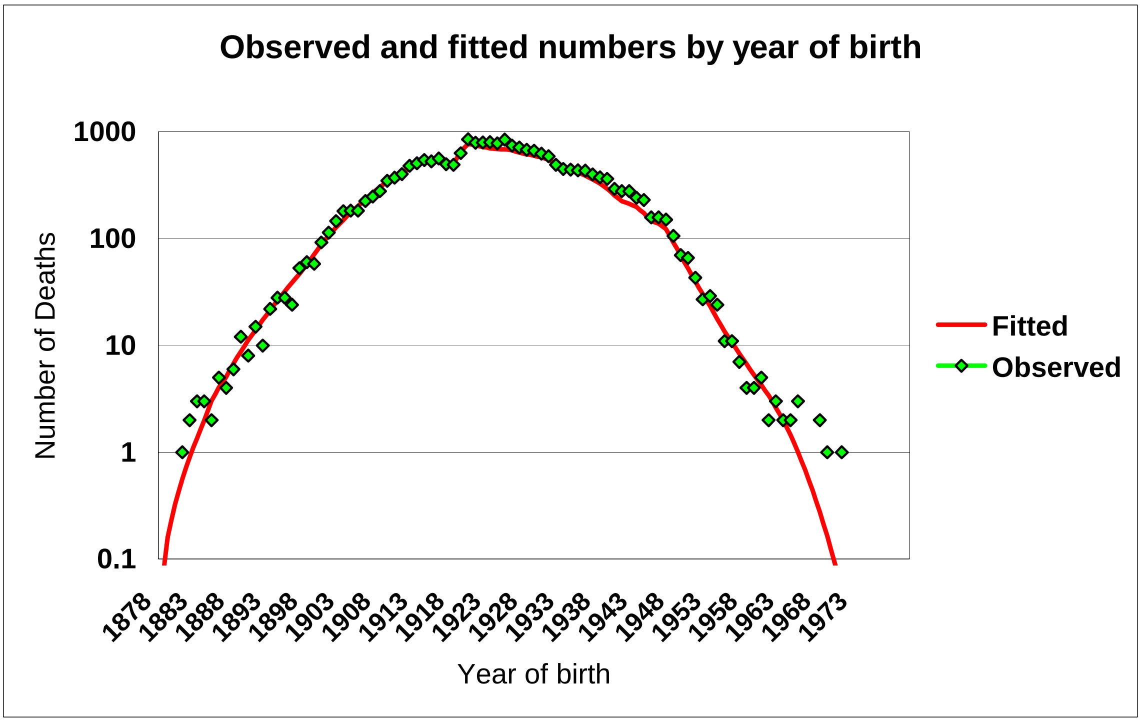


Figure 5(b)

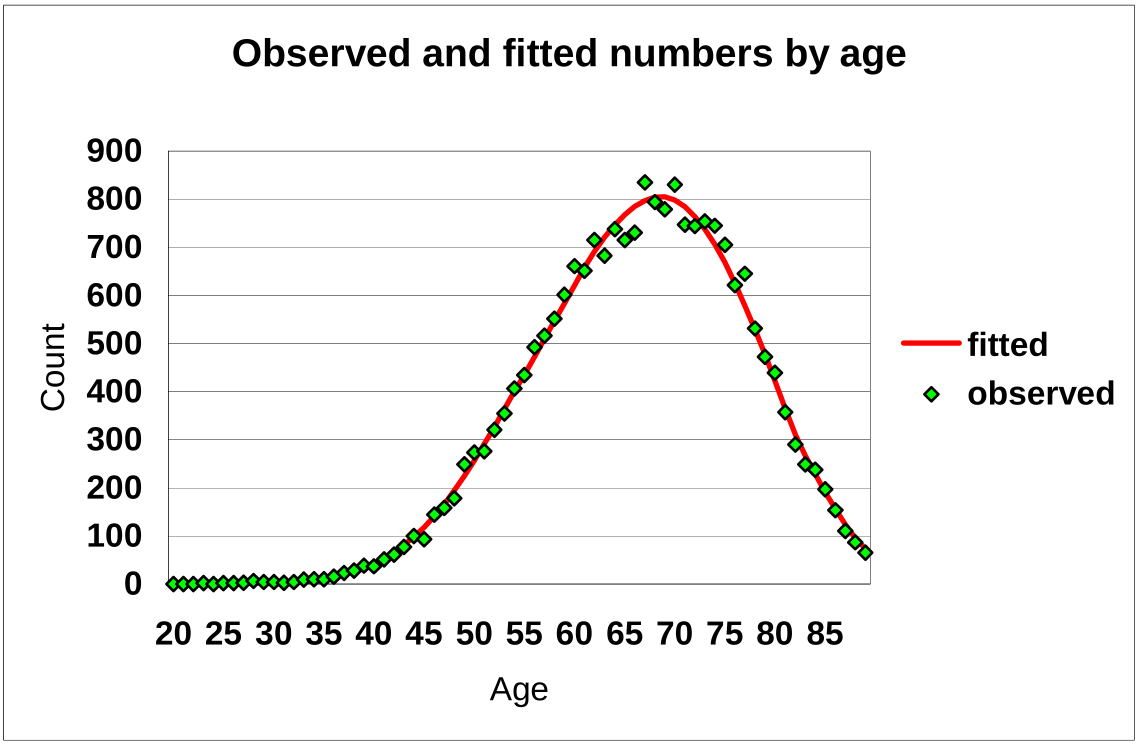


Figure 6(a)

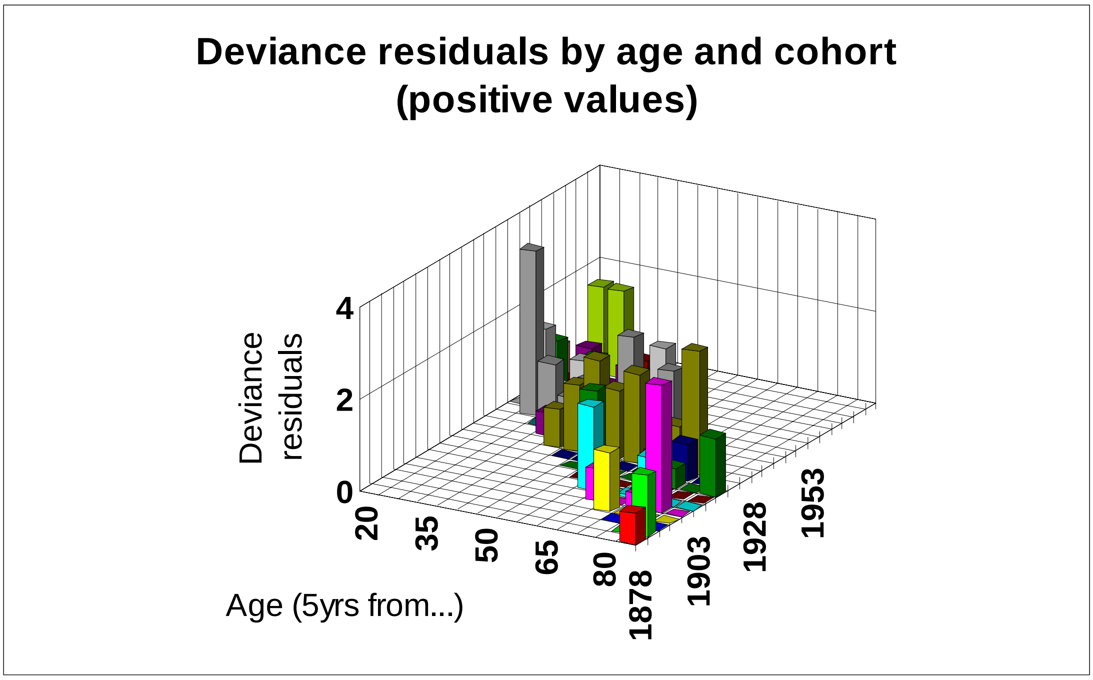
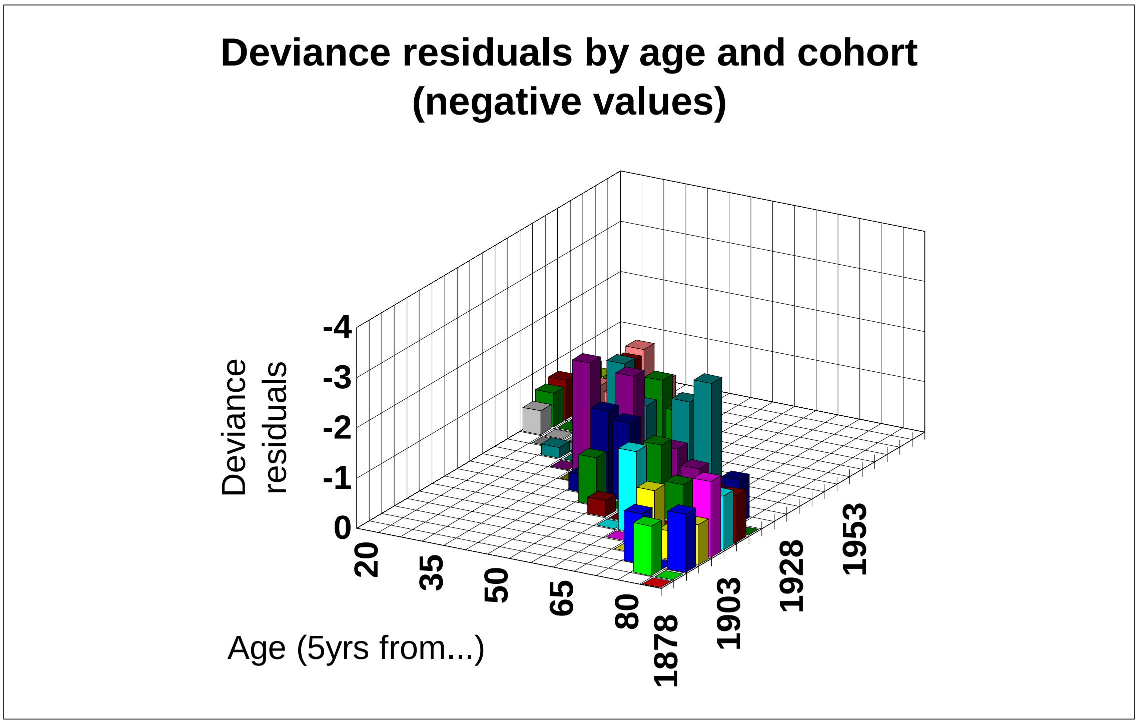


Figure 6(b)



Note: Deviance residuals measure the difference between observed and expected values in a standardised form. For a satisfactory fit, not more than one in 20 values should fall outside the range (-2, 2). For this model slightly more (8 of the 112) residuals lie outside this range.

Appendix 4 - Technical notes on the current model

Construction of uncertainty range

Confidence intervals for the predicted future course of the mesothelioma epidemic were not calculated due to practical difficulties in the way the model optimisation was carried out. However, since it is important to give an indication of the statistical uncertainty particularly when projecting into the future, a range based on a less formal approach was constructed.

The upper uncertainty limit for the peak number of deaths was determined by finding an alternative model that predicted a higher peak and where the increase in deviance over the standard model corresponded to the critical value for the appropriate chi-square goodness-of-fit test. In practice this process was started by changing k to the point were the model would give a significantly worse fit. Next, other parameters were changed to reduce the deviance and at the same time increase the value of the peak projection. The process was then repeated – first adjusting k to increase the deviance to the critical level once again, and then adjusting the other parameters – until no further adjustments could be made. The lower uncertainty limit was determined in the same way – however, in this case the second step involved minimising the deviance whilst at the same time reducing the value of the peak projection. These estimates effectively define a minimum size for the confidence interval.

Residual analysis

The deviance residuals for the model appear to be reasonably well behaved. The two-dimensional display of deviance residuals by five-year age and year of birth groupings shows a reasonably uniform pattern with no strong clustering of residuals of the same sign (see Figure 6, appendix 3). However, a histogram of the residuals indicates some departure from normality: there is increased clustering of residuals about the mean and the tails of the distribution are wider than expected. However, there is no appreciable skew.

Estimating the number of background cases

In order to investigate whether the number of "background" mesothelioma deaths could account for the consistent under prediction by the model of the number of deaths in the most recent birth cohorts, the total annual background deaths for males (b) was required. An estimate of b was calculated by first carrying out a linear regression analysis of the annual number of female mesothelioma deaths against the annual male deaths for 1968-2001. This analysis yielded the following relationship:

$$F = 22.9 + 0.136M \qquad [1]$$

Where, F = annual female mesothelioma deaths
M = annual male mesothelioma deaths.

An R^2 value of 0.952 suggest that the fit is good – although the residuals do exhibit some heteroscedasticity.

Assuming that there are the same number of background cases annually in females as in males (ie b cases for each), and assuming that there is no difference in mesothelioma risk due to asbestos exposure between the sexes, then the following relationship should hold:

$F - b = 0.136(M - b)$ [2].

Solving equations [1] and [2] gives b=26.5.

Having calculated the annual number of background mesothelioma deaths in males, these were distributed amongst the different ages in birth cohorts assuming that the mesothelioma rate is proportional to $(\text{Age} - 10)^{k+1}$ where k is the fitted value of the exponent of time in the current model (k=2.6).

Appendix 5 – Alternative models

Table 3

Overall model fit:			
Deviance	223.22	Degrees of freedom	182
Parameter estimates:			
k	3.0	Diagnostic trend (decrease in cases missed - % per year)	1
Maximum exposure year	1966	Lung clearance half-life (years)	102
Change in exposure index (%per year) in...		Relative exposure potential by age group	
1921	26	5 to 15	0.04
1931	5	16 to 19	0.43
1941	14	20 to 29	1.04
1951	7	30 to 39	1.16
1961	6	40 to 49	0.92
1966	0 (by definition)	50 to 59	0
1971	-21	60 to 69	0
1981	-9	70 to 79	0
Projections of future mesothelioma deaths in males aged 20-89:			
Peak level	1895	Peak year	2013

Table 4

Overall model fit:			
Deviance	224.68	Degrees of freedom	182
Parameter estimates:			
k	3.0	Diagnostic trend (decrease in cases missed - % per year)	1
Maximum exposure year	1967	Lung clearance half-life (years)	92
Change in exposure index (%per year) in...		Relative exposure potential by age group	
1922	26	5 to 15	0.04
1932	6	16 to 19	0.37
1942	13	20 to 29	1.07
1952	7	30 to 39	1.16
1962	5	40 to 49	0.94
1967	0 (by definition)	50 to 59	0
1972	-23	60 to 69	0
1982	-14	70 to 79	0
Projections of future mesothelioma deaths in males aged 20-89:			
Peak level	1900	Peak year	2013

Appendix 6 – Glossary of statistical terms

Deviance is a measure for judging the overall adequacy of a statistical model. A good model will account for most of the variation in the observed data and will have lower deviance value than a worse fitting model.

Residuals are the individual measures of how close each value predicted by the model (the so called ‘fitted values’) is to the observed data values. As well as having a low overall deviance, a good model should have residuals which exihibit certain statistical properties. Residuals thus provide an important way of assessing the validity of a particular model.

P-value (Probability-value) - A measure of the statistical fit between an observation (or set of observations) and the value (or values) predicted by a model. The P-value is the probability of observing a more extreme value than the observed value under the assumption that the model is true. A small P-value therefore indicates that an observation is unlikely to have taken the value it did if the model were true, and therefore suggests some inadequacy (lack of fit) in the model.

Statistical Significance – This concept is closely related to that of P-values. Observed numbers may be different to those predicted by a model by chance alone. Statistical significance levels are used to determine cut-off points for the difference between observed and fitted values which if exceed indicate that the observed difference is unlikely to be because of chance alone. There are 2 significance levels that are normally used, 5% (correspondeing to a P-value of 0.05), and 1% (corresponding to a P-value of 0.01). Thus, a P-value equal to or lower than 0.05 indicates statistical significance at the 5% level, and a P-value equal to or lower than 0.01 statistical significantce at the 1% level.

Enquiry contact point:

Epidemiology and Medical Statistics Unit
Health and Safety Executive
Room 244
Magdalen House
Stanley Precinct
Bootle
Merseyside L20 3QZ

Tel: 0151 951 3479 / 3051

HSE publications:

Latest mesothelioma statistics summary:
http://www.hse.gov.uk/statistics/causdis/mesothelioma/index.htm

The following leaflets are also available - free from HSE Books at:

HSE Books
PO BOX 1999
Sudbury
Suffolk
CO10 2WA
(Tel: 01787 881165).

“Asbestos Dust**:** The Hidden Killer. Are you at risk? Essential advice for building maintenance, repair and refurbishment workers”, INDG 187L

“Asbestos alert for building maintenance, repair and refurbishment workers”, INDG 188P- a pocket card for workers.

“Asbestos dust kills. Keep your mask on.” INDG 255

“Working with asbestos in buildings” INDG 289

“A short guide to managing asbestos in premises” INDG 223 (rev 3)