HSE

Health and Safety Executive

Mesothelioma mortality in Great Britain

The revised risk and two-stage clonal expansion models

Prepared by the **Health and Safety Laboratory**
for the Health and Safety Executive 2011

RR876
Research Report

HSE

Health and Safety Executive

Mesothelioma mortality in Great Britain

The revised risk and two-stage clonal expansion models

Emma Tan & Nick Warren
Harpur Hill
Buxton
Derbyshire
SK17 9JN

Asbestos is a known carcinogen that is the cause of the majority of mesothelioma cases worldwide. Various models have been used to describe the increase and likely future pattern of mesothelioma rates seen in many western countries – a legacy of past heavy industrial asbestos use. Following on from previous work (Tan and Warren, 2009), we analysed female mesothelioma mortality using the same risk model that was assumed for males. We also analysed mesothelioma mortality in males in Great Britain using two alternative risk models; the first is based on asbestos import data where the population is categorised into low and high exposure groups, with the calculation of risk based on the cumulative lung burden of the individual; the second is a two-stage clonal expansion model (TSCE), a biologically-based carcinogenesis model that assumes that the development of a malignant cell is the result of two critical and irreversible events, with asbestos lung burden as the measure of dose that enters the dose-response component of the TSCE model. We use Markov Chain Monte Carlo within a Bayesian framework to fit the models presented in this report.

Though considerably uncertain, peak mortality in females is predicted to occur over a decade later than in males, but with a substantially lower annual number of deaths. The updated models provide a reasonable basis for making relatively short-term projections of mesothelioma mortality in Britain. However, longer-term predictions comprise additional uncertainty not captured within the prediction intervals for the annual mortality rates. Taking this into account, 2100 deaths in 2016 represents our current best estimate of the upper limit for the male projections.

This report and the work it describes were funded by the Health and Safety Executive (HSE). Its contents, including any opinions and/or conclusions expressed, are those of the authors alone and do not necessarily reflect HSE policy.

HSE Books

First published 2011

ACKNOWLEDGEMENTS:

The authors would like to thank Andrew Darnton and John Hodgson (CSEAD Epidemiology, HSE) for their valuable contribution to this report.

EXECUTIVE SUMMARY

Aims

This report presents Bayesian statistical analyses of mesothelioma mortality in Great Britain between the years 1968 and 2007. This report updates previous work carried out in HSE Research Report RR728.

The aims of the statistical analysis were:

- To fit the previous statistical model to female mortality data;
- To consider alternative approaches to modelling mesothelioma mortality;
- To investigate whether the risk of mesothelioma increases indefinitely with time since first exposure to asbestos;
- To allow for varying asbestos exposure profiles for different subgroups of the population;
- To fit the alternative models to male mortality data;
- To produce and subsequently compare updated estimated annual mesothelioma deaths to 2050 with confidence and prediction intervals in order to determine whether the previous projections were specific to the previous model.

Main Findings

Males:

- Two alternative approaches have been developed that share a common exposure framework based upon asbestos imports data, the revised risk model and the two-stage clonal expansion (TSCE) model.
- Using the optimal revised risk and TSCE models considered in this report, mesothelioma mortality is predicted to peak at around 1860 deaths in 2012 and 1780 deaths in 2010 respectively.
- The estimated number of background cases in 2007 using the revised risk and the TSCE model were 19 and 9 cases respectively.
- Within the framework of the TSCE model, the cumulative hazard at age 89 for males who were exposed from age 25 for 5 years (the largest exposure category) was highest amongst the 1930 to 1940 birth cohort.
- The updated models provide a reasonable basis for making relatively short-term projections of mesothelioma mortality in Britain, including the extent and timing of the peak number of deaths. However, longer-term predictions comprise two additional sources of uncertainty which are not captured within the prediction intervals for the annual number of deaths: 1) whether the form of the model is valid for more recent and future exposure contexts; and 2) if the model is valid in such

contexts, the uncertainty arising from the particular choice of exposure model or model parameters;

- Given these uncertainties, a range spanning the lowest and highest confidence bands of the Tan and Warren (2009) model, the optimal revised risk and the TSCE model is likely to give a better reflection of the true uncertainty in the projections than the range based on any one of the three models. On this basis, the upper limit from Tan and Warren (2009), that is, 2100 cases in 2016 is represents our current best estimate of the upper limit for the male projections.

Females:

- There was a sharp increase in the implied exposure amongst females around the year 1948 with a rapid decline following; the implied exposure subsequently increased to a global peak around 1965, however there was greater uncertainty in the exposure levels after 1980.

- The background rate was estimated at approximately 1.3 cases per million, suggesting that there are a small number of cases (about 30 per year) that are not caused by exposure to asbestos.

- Although there was considerable uncertainty regarding current and recent exposure levels, a consistent finding was that the peak year of mortality is predicted to occur over a decade later in females than in males, with a lower number of peak deaths for females than males.

Recommendations

- Make comparisons of the projections with the latest mortality data as it becomes available in order to further assess the fit and the adequacy of the existing models. The models might also be refitted to obtain updated model parameters and model projections.

- Investigate the effects of asbestos exposure on malignant transformation rates within the two-stage clonal expansion modelling framework.

- Investigate what evidence exists for the determination of stock removal parameters.

.

CONTENTS

1 INTRODUCTION

Mesothelioma is a rapidly fatal form of cancer that is almost always caused by exposure to asbestos. The majority of those who develop mesothelioma have had occupations with significant exposure to asbestos fibres (Rake *et al.*, 2009). Mesothelioma has a long latency period; symptoms usually emerge between 15 and 60 years after exposure to asbestos.

Projections of the future burden of mortality in Great Britain have been published by the Health and Safety Executive and have been widely used both within HSE and externally. Hodgson *et al.* (2005) developed a statistical model based on the dose-response model for mesothelioma (Heath Effects Institute, 1991), where an individual's exposure to asbestos is assumed to be dependent on calendar year and the age of the individual in that calendar year. Using this model, mesothelioma mortality in Great Britain amongst males aged under 90 was predicted to reach a peak at around 1,650 to 2,100 deaths per year some time between 2011 and 2015, followed by a rapid decline.

Tan and Warren (2009) presented a more refined statistical analysis of mesothelioma mortality amongst males in Great Britain based on Markov Chain Monte Carlo (MCMC) methods using a modified form of the model formulated by Hodgson *et al.* (2005). The use of MCMC allowed the calculation of credible intervals for model parameters and prediction intervals for mesothelioma mortality to be made. Mortality amongst all males was predicted to peak at around 2,040 deaths in the year 2016, with a rapid decline following. The recommendation in Tan and Warren (2009) to carry out further investigations of fitting the model used for males to data for females, in particular whether to assume common parameter values for males and females in some cases, has been considered and the results presented in Section 2.

Although the Tan and Warren (2009) model fits the data well, it is not clear whether the model form is valid for more recent and future exposures. In particular, the fact that the model does not fit well when deriving exposure from imports and the relatively high impact of exposure at higher ages gives us cause to have some doubts about the model, as well as to question whether the risk should eventually level off with time since exposure. In Section 3, we attempt to address these issues by moving to a more empirically based exposure index; a statistical analysis of mesothelioma mortality in Great Britain from 1968 to 2007 is presented, using a revised risk model based on asbestos import data in which the male population is classified into low and high exposure categories. Those in the high exposure category are then subclassified according to age and duration of exposure, with the calculation of the risk based on the cumulative lung burden of the individual.

The analyses carried out by Hodgson *et al.* (2005) and Tan and Warren (2009) have been based on statistical models where it was assumed that the increase in subsequent mesothelioma risk caused by each period of asbestos exposure is proportional to the asbestos exposure during that period, and to a power of time since exposure. Section 4 presents a statistical analysis of mesothelioma mortality within the framework of the two-stage clonal expansion (TSCE) model (Moolgavkar and Knudson, 1981), a carcinogenesis model which takes into account biological considerations. The TSCE model allows us to incorporate information about exposure patterns to toxic carcinogens in mesothelioma risk assessment and has previously been used to model the effects of asbestos on lung cancer risk (Richardson, 2009) and the effects of tobacco smoke on lung cancer mortality (Hazelton *et al.*, 2005).

1.1 ASBESTOS

Asbestos is a mineral that is extremely flexible, durable and non-flammable at high temperatures. It first became popular in Great Britain in the late 1800s due to its highly desirable properties and was widely used in the rail and shipyard industries, and later in the building industry which represented the biggest use of asbestos. Asbestos use was at its highest around the 1940s to the 1970s when thousands of asbestos-containing products were made. Although the majority of asbestos fibres may lie intact in buildings for several years or decades, those that are disturbed may become airborne and pose a risk when inhaled and retained in the lungs. Asbestos is classed as a category 1 carcinogen and inhalation of fibres may lead to serious diseases such as lung cancer and mesothelioma.

There are three main types of asbestos fibres that have been commercially used in Great Britain: crocidolite (blue), amosite (brown) and chrysotile (white). Chrysotile is the most common and abundant form of asbestos with shorter fibre lengths than amosite and crocidolite. Several studies have shown that chrysotile has a half life in the lungs of around 15 days and does not pose as great a risk of cancer as crocidolite and amosite, which have longer fibre lengths and a longer half life of several decades, thus remaining in the lungs for a much longer period (Berry, 1999).

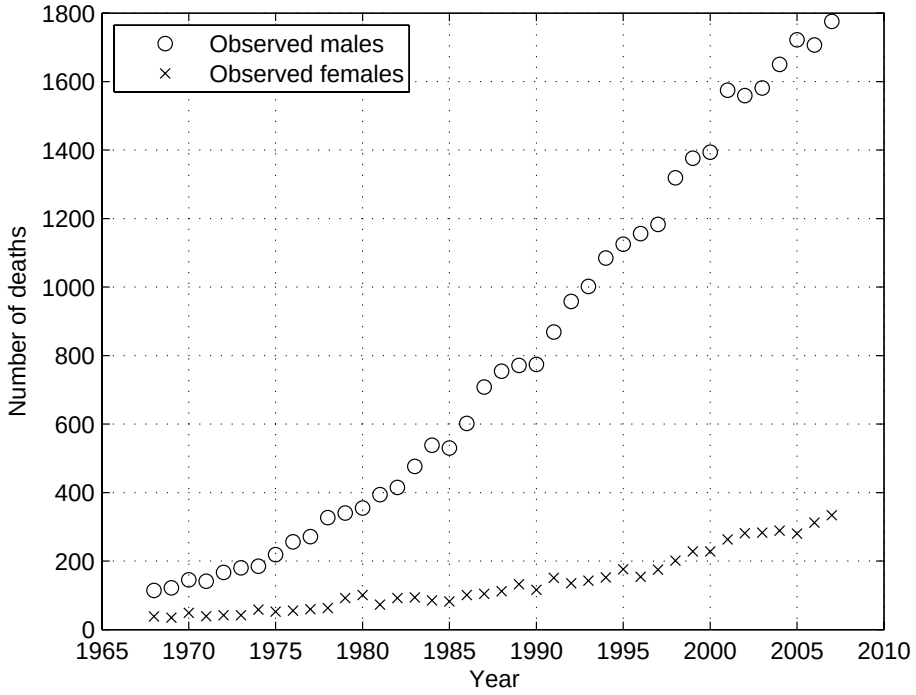

Figure 1 Male and female mesothelioma deaths (aged 20 to 89) from 1968 to 2007

1.2 MESOTHELIOMA DEATHS AND ASBESTOS IMPORTS DATA

The number of deaths due to mesothelioma in Great Britain (where mesothelioma was mentioned on the death certificate) is published annually by the Health and Safety Executive. In both males and females, 99% of all these deaths have been amongst those between the ages of 20 and 89. The data used in this report are based on deaths of males aged between 20 and 89 between the years 1968 and 2007. Figure 1 shows the observed deaths amongst males and females aged 20 to 89 between the years 1968 and 2007. The population data that was used in the analyses were the ONS mid-year population estimates for 1968 to 2007 and GAD population projections for 2008 to 2050.

The UK imports data for crocidolite, amosite and chrysotile from 1880 to 2007 were obtained from reports and submitted evidence from the Advisory Committee on Asbestos. From 1978 to 1995, data were obtained from the Asbestos Information Council with the exception of years 1984 to 1989 which were obtained from Eurostat. Crocidolite and amosite imports were formally banned in the UK in the mid-1980s; very little crocidolite was imported after 1970 whereas amosite continued to be imported and widely used until about 1980. Chrysotile imports and use declined during the 1980s and a total ban on asbestos came into effect in 1999. Linear interpolation was used to estimate imports for individual years where data were unavailable. Figure 2 shows the annual UK asbestos imports from 1880 to 1999.

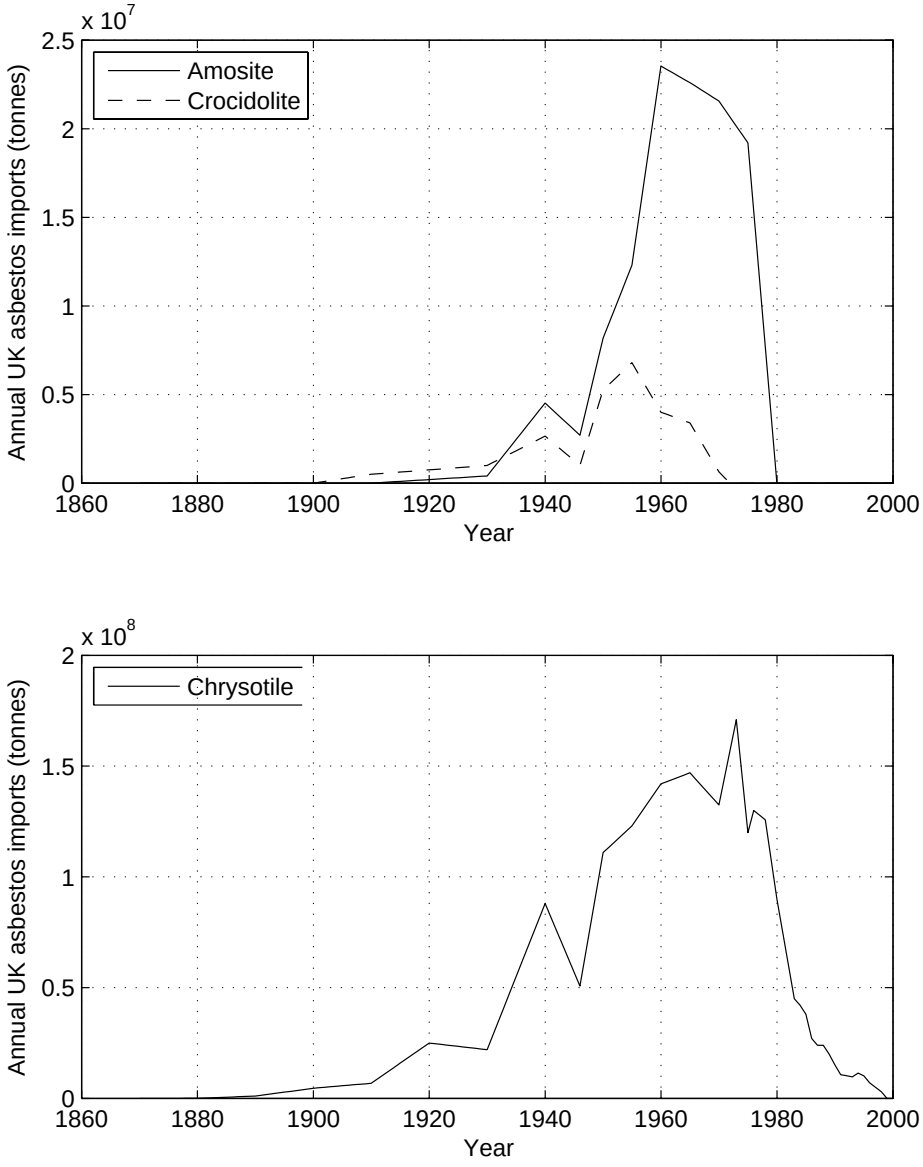

Figure 2 Annual UK asbestos imports

2 FEMALES

Initial exploration of fitting the Tan and Warren (2009) model to data for females showed that the smaller number of deaths in comparison to males led to greater uncertainty in the estimated model parameters. A simple substitution of the estimated parameters obtained for males did not result in a satisfactory estimation of female deaths, suggesting that some of the parameter values may not be common to both males and females and that a set of separate parameter estimates are required to make reliable inferences on female mortality and model parameters. This led to the recommendation in research report RR728 to carry out further work on female mesothelioma mortality data.

The number of mesothelioma deaths amongst females is much lower than amongst males so assuming a common background rate of mesotheliomas not caused by asbestos across both genders implies that background cases account for a higher proportion of female deaths. The data on females are thus important in their own right as they potentially allow more reliable estimation of background rates to be made.

2.1 THE MODEL

The model for males that was used in Tan and Warren (2009) has been used to model female deaths between the ages of 20 and 89 however, as the diagnostic trend parameter was found to be insignificant when analysing male data, the diagnostic trend component has been omitted from the female model. The female model takes the following form:

$$\lambda_{A,T} = \frac{[\sum_{l=1}^{A-1} W_{A-l} D_{T-l} I\,(l+1-L)^k\, 0.5^{l/H}] P_{A,T} (M - \sum_{A=20}^{89} \sum_{T=1968}^{2007} B_{A,T})}{\sum_{A=20}^{89} \sum_{T=1968}^{2007} [\sum_{l=1}^{A-1} W_{A-l} D_{T-l}\,(l+1-L)^k\, 0.5^{l/H}] P_{A,T}} + B_{A,T} \tag{1}$$

where

- $\lambda_{A,T}$ is the number of deaths at age A in year T;
- W_A is the overall age-specific exposure potential at age A;
- D_T is the overall population exposure in year T;
- L is the lag period in years between exposure and its contribution to the risk of mesothelioma and is fixed at 10 years;
- H is the half-life in years for asbestos clearance from the lungs;
- k is the power of time representing the increase in risk with time since exposure;
- $P_{A,T}$ is the person-years at risk for age A in year T;
- M is the total observed mesothelioma deaths from 1968 to 2007;
- I is an indicator variable where $I = 0$ if $l < L - 1$ and $I = 1$ otherwise;
- l indexes years lagged from the risk year;
- $B_{A,T}$ is the number of background cases for age A at year T.

The age-specific exposure potential, W_A, allowed the exposure of a female to differ by age. Nine parameters were assigned to W_A, representing the exposure weighting for the age groups (in years) 0 to 4, 5 to 15, 16 to 19, 20 to 29, 30 to 39, 40 to 49, 50 to 59, 60 to 64 and $65+$, with the age group 20 to 29 years chosen as the baseline category.

The overall population exposure, D_T, represents the average 'effective carcinogenic dose' in the breathing zone of females aged 20 to 89 years and is included as a unit-free parameter vector in the model. In this framework it is the shape of the exposure profile D_T over time, rather than the scale in any given year (which is set arbitrarily) that determines future mesothelioma risk. D_T was defined by growth and decline rates for years in multiples of 10 before and after the maximum exposure year, '$Peakyear$' (at which the gradient of the exposure curve is zero). The growth rates for intermediate years were determined by linear interpolation. The growth rates at $Peakyear - 65$ (D_1), $Peakyear - 55$ (D_2), $Peakyear - 45$ (D_3), $Peakyear - 35$ (D_4), $Peakyear - 25$ (D_5), $Peakyear - 15$ (D_6), $Peakyear - 5$ (D_7), $Peakyear + 5$ (D_8) and $Peakyear + 15$ (D_9) were included as parameters in the model. The proportion of the peak exposure in 2000, $Prop$, was also included as a parameter; this value was fixed at 4% for males in Tan and Warren (2009). Exposures for years between $Peakyear + 15$ and 2000 were calculated by linear interpolation.

The background rate ($Rate$) is represented by the number of background cases per million in the female population. The age distribution of the background cases in each year is assumed to be $(A - L)^{k_b}$ where k_b is the power of time representing the increase in risk with age. The proportion of background cases at age A in each year is therefore assumed to be $\frac{(A-L)^{k_b}}{\sum_A (A-L)^{k_b}}$.

2.2 MODELS FITTED

In the female model, the power of time k_b associated with background cases has been allowed to differ from the power of time k associated with asbestos exposure; both k and k_b have been estimated. As it was found for males by Tan and Warren (2009) that there was no optimal value of the clearance half-life H, H has been fixed at 1,000,000 years for females. Although the change in exposure index at $Peakyear - 65$, $Peakyear - 55$ and $Peakyear - 45$ were fixed in Tan and Warren (2009), they have been estimated at a common value for females for the following reasons: (i) allowing the parameters to vary potentially allows us to infer asbestos exposure from mortality data as opposed to fixing the parameters based upon little knowledge of exposure during that period, (ii) preliminary analysis suggested that assuming a common value provided almost an equally good fit as assuming three distinct values, and (iii) fewer estimated parameters result in a more parsimonious model. The effects of changing assumptions about exposure post-1980 were also explored. In particular, the assumptions used for males that the exposure levels were 4% of the peak level in 2000, 2% in 2010 and 0.75% in 2050 were changed to allow for the fact that the differential between the past peak exposure and recent/future exposures is likely to be lower for females. Historically, a lower proportion of females would have worked in high risk occupations compared to males, thus resulting in a lower peak level of exposure at a population level. In more recent years (around 2000 onwards), the exposure levels of males are likely to have reduced to similar levels as females, as the levels of exposure begin to level off to a background/threshold level. To assess this we fitted two different models:

- Model F1: the exposure in 2000 as a proportion of the peak, $Prop$, was fixed at 20%;
- Model F2: $Prop$ was estimated.

In both cases the exposure in 2010 and 2050 were set at $Prop/2$ and $Prop/4$ respectively, and exposures for other years calculated by linear interpolation.

2.3 STATISTICAL METHODOLOGY

Preliminary analyses were carried out by fitting the model using the *fminsearch* function in Matlab (The Mathworks, Inc., 2009) by minimising the model deviance, a measure of how well the model fits the observed data. The Poisson deviance can be expressed as

$$deviance = 2\sum_{A,T}\left[Y_{A,T}\log\left(\frac{Y_{A,T}}{\hat{F}_{A,T}}\right) - (Y_{A,T} - \hat{F}_{A,T})\right] \tag{2}$$

where $Y_{A,T}$ are the observations and $\hat{F}_{A,T}$ are the fitted values. Although *fminsearch* allows the data to be fitted quickly and easily, a disadvantage is that confidence intervals are not provided. *fminsearch* has thus been used to provide initial point estimates of the parameters, which in turn have been used as approximate starting values for the Metropolis-Hastings algorithm (Hastings, 1970), a Markov Chain Monte Carlo (MCMC) technique. This allowed not only model parameters to be estimated, but also allowed credible intervals to be easily obtained using formal statistical methods.

Markov Chain Monte Carlo

From a Bayesian perspective, the parameters of a statistical model are considered random quantities. Bayesian inference can usually be summarised by random draws from the posterior distributions of the model parameters. Let $Lik(Y|\theta)$ be the likelihood function of the data Y, θ be the vector of model parameters and $\phi(\theta)$ be the prior distribution of the parameters, which represents the prior information we have on θ. The posterior distribution $\pi(\theta)$ of θ is

$$\pi(\theta) \propto Lik(Y|\theta)\phi(\theta). \tag{3}$$

Assuming that the observations follow a Poisson distribution, the likelihood function is

$$Lik(Y|\theta) \;=\; \prod_{A,T}\left(\frac{e^{-\hat{\lambda}_{A,T}}\hat{\lambda}_{A,T}^{Y_{A,T}}}{Y_{A,T}!}\right) \tag{4}$$

which is the product of the individual likelihood contributions for each observation over all ages and years of death. Unfortunately, evaluation of the posterior distribution is normally extremely difficult and numerical techniques, particularly MCMC, are required. MCMC techniques require simulation to generate random samples from a complex posterior distribution. A large number of random draws from the posterior distribution is generated. After a burn-in period (where an initial portion of samples are discarded to minimise the effect of initial values on posterior inference), the empirical distribution should eventually closely approximate the true shape of the posterior distribution. The MCMC chain is thinned in order to reduce autocorrelation. The process of thinning records samples periodically and discards the remaining samples. Point estimates and credible intervals are then calculated.

In the Metropolis-Hastings algorithm, given θ_t at time point t, the next state θ_{t+1} in the chain is chosen by sampling a candidate point θ^* from a proposal distribution $q(\cdot|\theta_t)$. The candidate point θ^* is then accepted with probability p where

$$p = \min\left[1, \frac{\pi(\theta^*)q(\theta_t|\theta^*)}{\pi(\theta_t)q(\theta^*|\theta_t)}\right].$$

If the candidate point is accepted, the next state $\theta_{t+1} = \theta^*$. If the point is rejected, the chain does not move, i.e. $\theta_{t+1} = \theta_t$. The process is then repeated for state θ_t at every time point t to obtain a sequence of values $\theta_1, \theta_2, \ldots$. The approximate distributions at each step in the simulation converge to the target distribution of interest, $\pi(\theta)$. As θ is a vector of model parameters, each component will be individually updated for convenience. Further details on the choice of the prior and proposal distributions can be found in Appendix 1.

2.4 RESULTS

Models F1 and F2 were fitted to the dataset using the Metropolis-Hastings algorithm. The results from fitting Models F1 and F2 are displayed in Tables 1 and 2 respectively.

The posterior medians of k in both models were larger than the posterior medians of k_b, at around 2.7 and 2.4 respectively, suggesting that the power of time may be greater for the risk associated with asbestos exposure than with background cases. However, there was not a significant difference between the values k and k_b, with the posterior medians of k lying within the 90% credible intervals of k_b, and vice versa.

The background rate was estimated to be around 1.3 cases per million in both models, corresponding to around 30 cases in 2007 amongst females aged 20 to 89. The background rate amongst males estimated by Tan and Warren (2009) was 1.08 cases per million, corresponding to around 23 cases in 2007 amongst males aged 20 to 89. However, assuming an equal proportion of male and female background cases, the female data potentially allow a more reliable estimate of background rate due to the much lower number of all-cause female cases.

The posterior medians of the age-specific exposure potential parameters were highest for females between the ages of 30 and 39, indicating a higher exposure contribution from the 30 to 39 age group (where the exposure contribution encapsulates both the exposure levels and proportion of females exposed associated with that age group). The lowest were for females aged below 20 and above 50 years. However, due to the lag period, there was high uncertainty in the estimates of the relative exposure potential for females aged 60 and above.

Figures 3A to 3D show plots of the fitted and observed deaths in F1 for females aged 20 to 89 by year of birth, age and year of death. Figures 4 and 5 shows the observed and fitted deaths (with associated prediction intervals) for females aged 20 to 89 and estimated exposure profiles (with associated credible intervals) for F1 and F2 respectively.

The posterior median of the peak year of mortality ranged from 2025 (Model F1; 444 deaths) to 2027 (Model F2; 477 deaths). Credible intervals for the peak level of mortality were calculated, however due to the high levels of uncertainty in the exposure levels for females from 1980 onwards in F2, the upper bounds for the prediction intervals for the peak number of deaths increased for several decades after the expected peak year of mortality, whereas the upper bounds under F1 decreased. Similarly, the lower bounds for the prediction intervals under F2 decreased to much lower levels than that in F1. These patterns have been observed due to the fact that $Prop$ has been estimated in F2 (introducing a degree of uncertainty), but artificially fixed in F1.

The estimated exposure curve in both models indicated a local peak of exposure just prior to 1950, and a global peak around 1965. Between 1965 and 1980, the estimated exposure levels decreased rapidly. There was great uncertainty in the exposure levels after 1980 in F2, as any effects of exposure on the risk of mesothelioma from then onwards may not be observed for several decades. For F1 however, $Prop$ has been fixed at 0.2; this explains the apparent lack of uncertainty in the exposure profile from the year 2000 onwards.

The posterior median for the deviance was lowest in F1 at 279.3 (90% C.I. [263.7,290.0]), indicating that a better fit was achieved using F1. A plot of the deviance residuals by age group and birth cohort can be found in Figure A1 in Appendix A. Over 90% of the residuals lie within the range [2,2] and no obvious patterns can be seen, suggesting an adequate fit of the model to the data.

Table 1 Metropolis-Hastings: Posterior median and 90% credible intervals for Model F1

Posterior median (90% credible interval)			
k	2.63 (2.41,2.90)	Background rate	1.31 (1.05,1.58)
k_b	2.41 (2.04,2.77)	% of peak exposure in 2000	20 (fixed)
Maximum exposure year	1965	Half-life (years)	1000000 (fixed)
Change in exposure index (% per year) in...		Relative exposure potential by age group	
1900 ($D(1)$)	-13.8 (-91.4,145.4)	0 to 4	0.00 (0.00,0.02)
1910 ($D(2)$)	-13.8 (-91.4,145.4)	5 to 15	0.01 (0.00,0.03)
1920 ($D(3)$)	-13.8 (-91.4,145.4)	16 to 19	0.06 (0.00,0.23)
1930 ($D(4)$)	119.9 (-4.5,191.5)	20 to 29	1.00 (baseline)
1940 ($D(5)$)	122.8 (39.6,184.8)	30 to 39	1.20 (0.93,1.44)
1950 ($D(6)$)	-56.4 (-71.0,-26.3)	40 to 49	0.71 (0.25,1.15)
1960 ($D(7)$)	74.3 (42.7,102.5)	50 to 59	0.03 (0.00,0.12)
1965	0 (by definition)	60 to 64	0.09 (0.01,0.42)
1970 ($D(8)$)	-17.5 (-31.2,-3.77)	65+	0.27 (0.02,1.17)
1980 ($D(9)$)	5.2 (-22.1,33.4)		
Projections of future mesothelioma deaths in females aged 20-89			
Peak level	444 (380,527)	Peak year	2025 (2023,2027)
Deviance	279.3 (263.7,290.0)	Degrees of freedom	206

Table 2 Metropolis-Hastings: Posterior median and 90% credible intervals for Model F2

Posterior median (90% credible interval)			
k	2.67 (2.42,2.95)	Background rate	1.34 (1.03,1.64)
k_b	2.45 (2.07,2.80)	% of peak exposure in 2000	41 (4,93)
Maximum exposure year	1965	Half-life (years)	1000000 (fixed)
Change in exposure index (% per year) in...		Relative exposure potential by age group	
1900 ($D(1)$)	98.3 (-43.1,175.0)	0 to 4	0.00 (0.00,0.02)
1910 ($D(2)$)	98.3 (-43.1,175.0)	5 to 15	0.01 (0.00,0.03)
1920 ($D(3)$)	98.3 (-43.1,175.0)	16 to 19	0.07 (0.00,0.26)
1930 ($D(4)$)	100.2 (-9.9,188.1)	20 to 29	1.00 (baseline)
1940 ($D(5)$)	98.4 (49.1,169.8)	30 to 39	1.17 (0.92,1.42)
1950 ($D(6)$)	-43.5 (-64.8,-30.2)	40 to 49	0.62 (0.17,1.12)
1960 ($D(7)$)	61.1 (41.3,91.2)	50 to 59	0.03 (0.00,0.13)
1965	0 (by definition)	60 to 64	0.10 (0.01,0.39)
1970 ($D(8)$)	-15.1 (-32.6,-0.94)	65+	0.29 (0.02,1.29)
1980 ($D(9)$)	-1.2 (-31.8,37.5)		
Projections of future mesothelioma deaths in males aged 20-89			
Peak level	477 (371,-)	Peak year	2027 (2024,-)
Deviance	282.4 (265.8,294.2)	Degrees of freedom	206

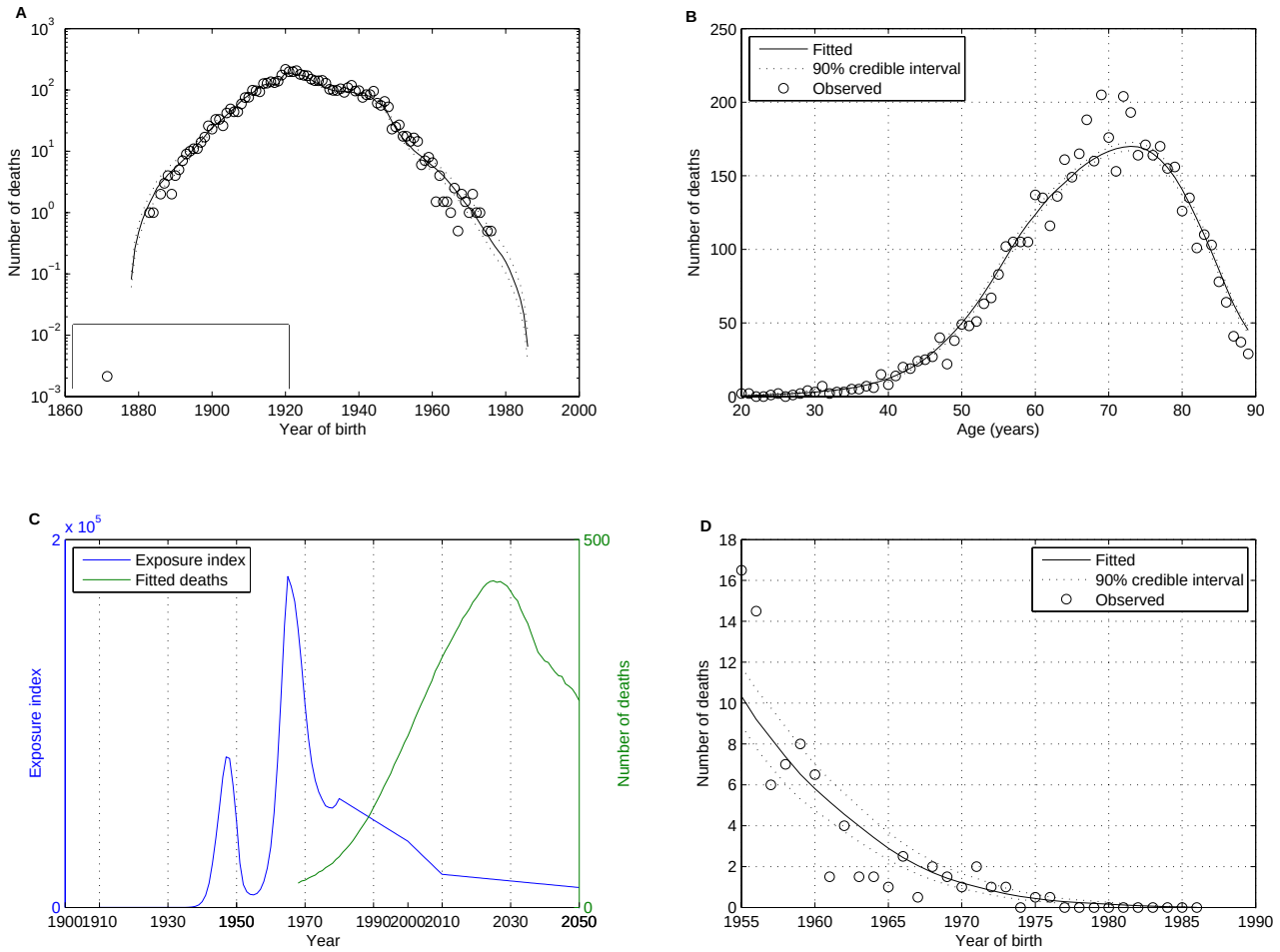

Figure 3 Model F1: (A) Observed and fitted deaths by year of birth. (B) Observed and fitted deaths by age. (C) Observed and fitted deaths by year of death, with derived exposure index. (D) Observed and fitted deaths for 1955-1985 birth cohorts.

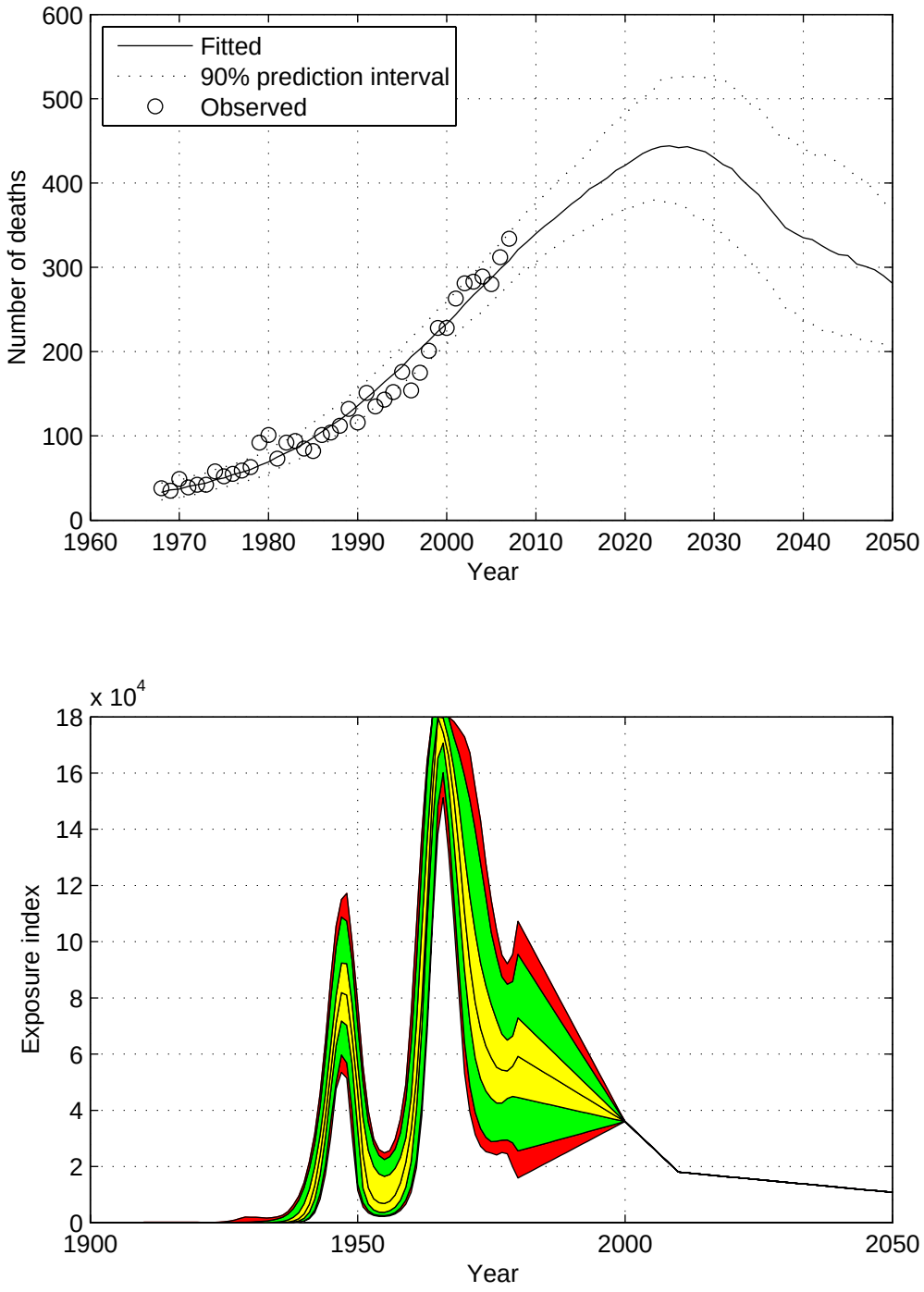

Figure 4 Model F1 (Top) Observed deaths with 50th percentile curve and 90% prediction interval (Bottom) Estimated exposure profile with 95% (red), 90% (green) and 50% (yellow) C.I.

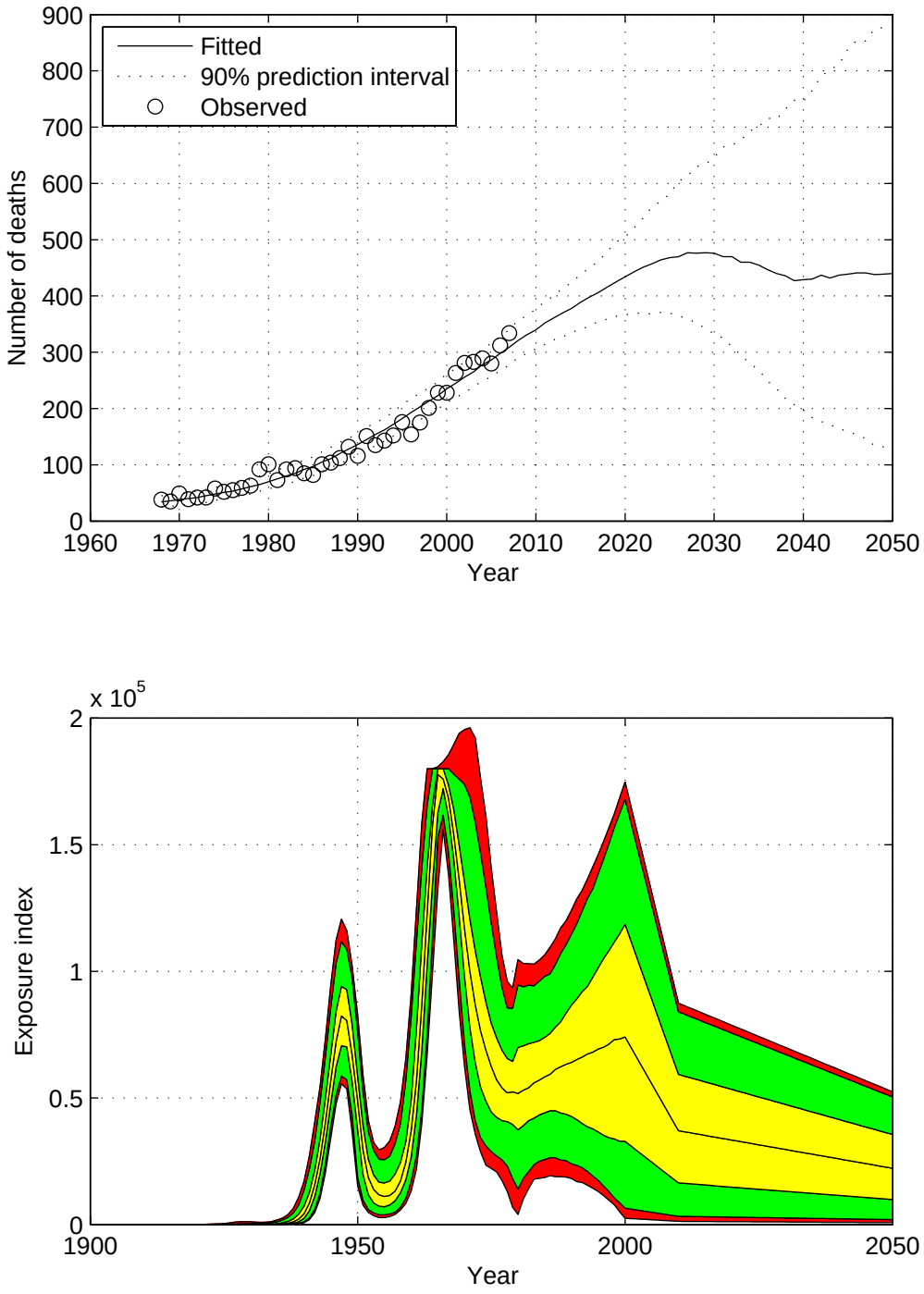

Figure 5 Model F2 (Top) Observed deaths with 50th percentile curve and 90% prediction interval (Bottom) Estimated exposure profile with 95% (red), 90% (green) and 50% (yellow) C.I.

2.5 DISCUSSION

This section has presented a statistical analysis of female mesothelioma mortality using a model based on that formulated by Tan and Warren (2009), with some additional modifications and assumptions. One of the modifications is that a distinction has been made between the power of time associated with asbestos-related risk and that associated with background risk. The reason for this is that the underlying factors or processes that contribute to risk may differ between the two; both powers of time have been estimated in the female analysis. The effect of these causes over time may differ from the effects attributed to occupational exposure; it is thus appropriate to allow for a distinction between the two power parameters. Although a difference was seen, it is not significant.

Another modification that has been made in the female analysis relates to the proportions of the peak level of exposure in 2000, 2010 and 2050. In the male analysis, these proportions have been fixed at 4%, 2% and 0.75% respectively. Historically however, a smaller proportion of females have had occupations in industries involving asbestos. It is therefore expected that the average exposure levels of females during periods of high industrial asbestos use were lower than that of males. Two different models have been fitted to the female data; one where the proportions of the peak level of exposure in 2000, 2010 and 2050 have been fixed at 20%, 10% and 5% respectively, and one where the proportions have been estimated. The final modification that has been made is the assumption of a common estimated parameter for the change in exposure index in $Peakyear-65$, $Peakyear-55$ and $Peakyear - 45$. These values had previously been fixed in the male analysis carried out by Tan and Warren (2009).

As there is a latency period of several decades between exposure to asbestos and the onset of mesothelioma, the majority of those who were exposed to high occupational levels of asbestos during the high exposure years around the 1960s are expected to die of mesothelioma within the next few decades. As with the male projections, the annual number of female deaths is expected to continue increasing until reaching a peak within the next few years, eventually decreasing due to the post-1970 decrease in levels of exposure, however this only holds for F1. A much wider range of projections are consistent with F2 due to the very high level of uncertainty in the estimate of exposure in 2000 (41% of the peak, 90% C.I. [4,93]). In F1 this value was fixed at 20% on the basis that both occupational and environmental exposures are likely to be substantially lower than those of the 1960s. A consistent finding of both models is that the year of the peak number of deaths is predicted to be over a decade later in women than men.

The number of annual female deaths have historically been much smaller than the number of male deaths. Long range forecasts of both male and female deaths are highly uncertain, however it is likely that the numbers will eventually converge once the effect of occupational exposures - especially those of the past that have predominated in men - cease to have a large impact on the population. The fact that these uncertainty ranges of the long term predictions for men and women overlap implies that such a scenario is consistent with these models. A comparison of the male projections (Tan and Warren, 2009) and female projections (based on F1) can be seen in Figure 6, which shows a convergence to less than 400 deaths annually by 2050. Additionally the background cases make up a much larger proportion of total cases in females, thus the female data arguably provide a more reliable estimate of confidence intervals for male and female background cases.

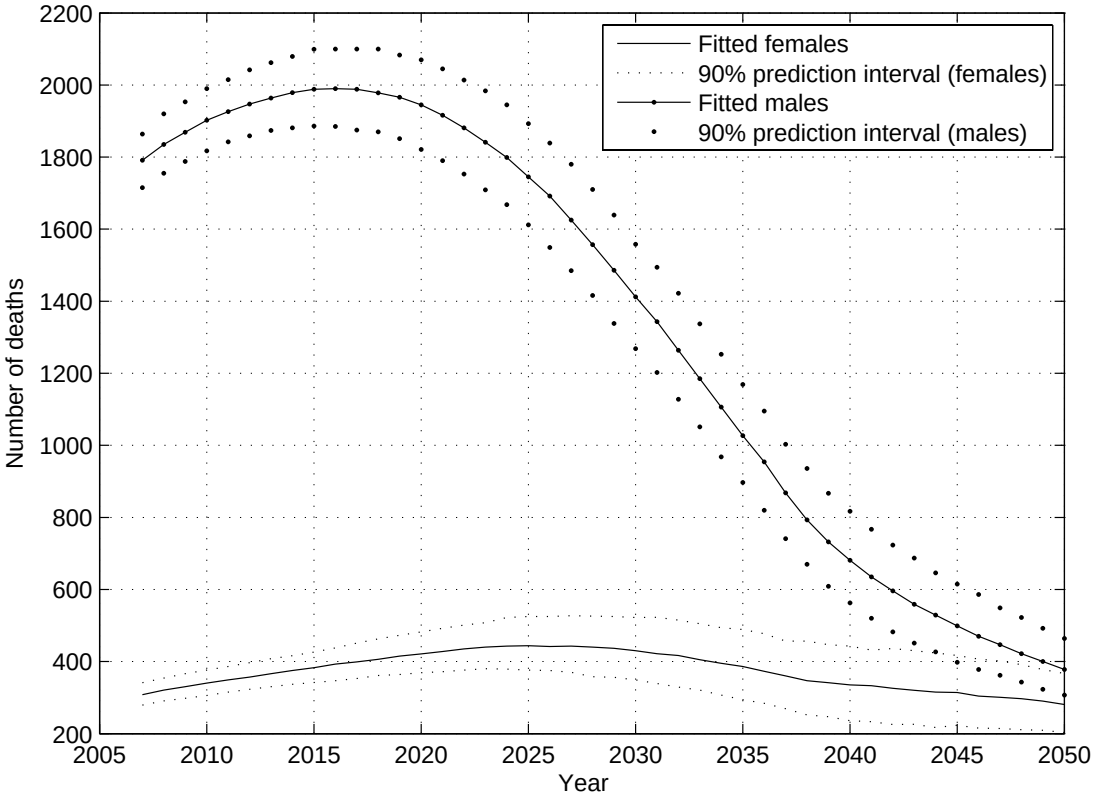

Figure 6 Comparison of male and female projections

3 REVISED RISK MODEL

3.1 REPRESENTATION OF EXPOSURE

In the models in previous work by Hodgson *et al.* (2005) and Tan and Warren (2009), asbestos exposure was inferred from the pattern of mortality data. It was not clear whether the model form was valid for more recent and future exposures, raising the question of whether the risk should eventually level off with time since exposure. In this section we move to a more empirically based exposure index, using actual asbestos imports data to infer asbestos exposure and lung burden of males. We also allow the risk to eventually level off with time since exposure, by including a parameter to represent the duration after exposure of increasing risk.

We assume that the male population can be sub-classified into low and high exposure categories, where males whose exposures were primarily non-occupational (e.g. due to asbestos in the household, environmental exposure) are classified as low exposed, and males whose exposures were primarily occupational (e.g. carpenters, plumbers) are classified as high exposed. A parameter has been included in the model to represent the proportion of males in the high exposure category. In addition, two parameters in the model represent the different exposure factors associated with the low and high exposure groups. Males in the high exposure category are then sub-classified according to the age at which high exposure started, and the duration of high exposure. Table 3 shows the sub-classification of the high exposure group. We assumed that high exposures could commence at one of three ages (15, 25 or 35 years) and for each, the distribution of exposure duration was estimated based on data for controls in a recent population based case-control study of mesothelioma in Britain (Rake *et al.*, 2009).

Table 3 Sub-classification of the high exposure group

Start age of high exposure	Proportion of males who were exposed for the following number of years								
	5	10	14	20	25	30	35	40	45
15	0.67	0.21	0.0764	0.0278	0.0101	0.0037	0.0013	0.0005	0.0001
25	0.45	0.24	0.09	0.07	0.08	0.05	0.02		
35	0.37	0.22	0.1	0.11	0.2				

We use lung burden as the measure of dose that enters the dose-response component of the model. The lung burden of an individual in a given year is assumed to be proportional to the sum of the effective amphibole exposure in all previous years, taking into account amphibole clearance from the lungs.

In year T, the imported tonnage of each of the three asbestos fibres is given by $Imp_{T,fib}$ where $fib = croc$, $amos$ or $chry$, representing crocidolite, amosite and chrysotile fibres. The cumulative amounts of each type of asbestos in year T are then given by

$$Cum_{T,fib} = Imp_{T,fib} + (1 - \delta)Cum_{T-1,fib} \tag{5}$$

where $Cum_{1870,fib} = Imp_{1870,fib}$ and δ is the proportion of the previous year's stock removed in the UK.

The annual exposure increments of each type in year T are then defined as

$$Inc_{T,fib} = Imp_{T,fib} + Cum_{T,fib} \times StockRel_T \tag{6}$$

where $StockRel_T = a_1$ if $T < 1980$ and a_2 otherwise.

The annual increment of amphibole exposure in year T is then defined as

$$D_T = Inc_{T,croc} \times CrocRatio + Inc_{T,amos} \tag{7}$$

where $CrocRatio$ is the risk factor of crocidolite relative to amosite. Lung burden is then derived from the sum of the effective amphibole exposure in all previous years, taking into account amphibole clearance from the lungs. For a low exposed male aged A in year T, the lung burden at time $T - i$ is

$$d^{low}_{A,T,i} = \beta_1 \sum_{j=i}^{A-10} D_{T-j} 0.5^{j/H} + \beta_3 Inc_{T,chry} \tag{8}$$

where β_1 is the low exposure factor corresponding to the amphiboles, β_3 is the risk factor of chrysotile relative to the amphiboles and H is the half-life (in years) for asbestos clearance from the lungs. For a male aged A in year T who was high exposed from age a for $5j$ years,

$$d^{high}_{A,T,a,i,j} = (\beta_1 + \beta_2 I_{a,i,j}) \quad \sum_{m=0}^{i} D_{T-A+m} 0.5^{\frac{A-a-m}{H}} + \beta_3 Inc_{T,chry} \tag{9}$$

is the lung burden at age i, where

$$I_{a,i,j} = \begin{cases} 1 & \text{if } i \in [a, a+5j] \\ 0 & \text{otherwise} \end{cases}$$

where β_1 and β_2 are the low and high exposure factors corresponding to the amphiboles.

3.2 RISK FUNCTION

In addition to the alternative representation of exposure, the model presented in this section utilises an alternative risk function. In the revised risk model, the risk of mesothelioma is assumed to be the sum of the risks due to exposure in all previous years of an individual's lifetime, excluding the most recent 10 years (the lag period). For each individual year, the contribution to the overall risk is assumed to be the product of the lung burden in that year, and a power of time since exposure (lagged by 10 years). A parameter s has been included in the model to represent the number of years after exposure during which the associated contribution to the risk increases, and after which the associated contribution to the risk levels off.

The revised model can be represented as follows:

$$\begin{aligned}\lambda_{A,T} &= P^*_{A,T}\Bigg\{(1-\sigma)\sum_{i=1}^{A} d^{low}_{A,T,i}(min[s, A-L-i])^k I(A-L-i) \\ &\quad +\sigma\Bigg[High_{15}\sum_{j=1}^{9} T_{15,5j} \quad \sum_{i=15}^{A} d^{high}_{A,T,15,i,j}(min[s, A-L-i])^k I(A-L-i) \\ &\quad +High_{25}\sum_{j=1}^{7} T_{25,5j} \quad \sum_{i=25}^{A} d^{high}_{A,T,25,i,j}(min[s, A-L-i])^k I(A-L-i) \\ &\quad +High_{35}\sum_{j=1}^{5} T_{35,5j} \quad \sum_{i=35}^{A} d^{high}_{A,T,35,i,j}(min[s, A-L-i])^k I(A-L-i) \Bigg]\Bigg\} \\ &\quad +Background_{A,T} \end{aligned} \tag{10}$$

where

- $\lambda_{A,T}$ is the number of deaths at age A in year T;
- $P^*_{A,T} = P_{A,T} \times 10^{-11}$ where $P_{A,T}$ is the person-years at risk at age A in year T;
- $Background_{A,T}$ is the number of background cases of males age A in year T.
- σ is the proportion of males who are high exposed (independent of age);
- k is the exponent of time associated with asbestos exposure;
- L is the lag period in years between exposure and its contribution to the risk of mesothelioma and is fixed at 10 years;
- $$I(A) = \begin{array}{ll} 1 & \text{if } A > 0; \\ 0 & \text{if } A \leq 0. \end{array}$$
- $High_{15}$ is the proportion of high exposed who were exposed from age 15;
  $High_{25}$ is the proportion of high exposed who were exposed from age 25;
  $High_{35}$ is the proportion of high exposed who were exposed from age 35;
  ($High_{15} = 1 - High_{25} - High_{35}$);
- $T_{15,5j}$ is the proportion of high exposed from age 15 who were exposed for $5j$ years;
  $T_{25,5j}$ is the proportion of high exposed from age 25 who were exposed for $5j$ years;
  $T_{35,5j}$ is the proportion of high exposed from age 35 who were exposed for $5j$ years;
- The background rate $Rate$, is the number of background cases per million population; the age distribution of the background cases is assumed to be $(A-10)^{k_b}$ where k_b is the exponent of time associated with the age distribution of background cases;
- s is the number of years after exposure ends, during which the associated contribution to the risk increases and after which the associated contribution to the risk levels off.

Models fitted

Exploratory analysis was initially carried out in MATLAB (The Mathworks, Inc., 2009) using *fminsearch*. Various sets of initial estimates were tested, however convergence to a single set of parameter estimates proved to be difficult. One explanation for this is that several combinations of parameter estimates may produce similar fits. In light of this, the stock removal parameter, δ, was fixed at 0.02; three different models were fitted where the proportion of the population in the high exposure group, σ, was fixed at 0.05, 0.10 and 0.15, and the effects on the parameter estimates and projections compared. It was found that s, H and k were correlated and preliminary analysis indicated that setting $H = 15$, $s = 100$ (effectively meaning that the risk does not level off) and estimating k provided a good fit to the data; in light of this, H and k were fixed at 15 and 100 respectively and k estimated in all three models. It was also found that β_1 converged to values close to zero, appearing very low in comparison to the estimates obtained for β_2. This suggests that the low exposure factor, and thus asbestos exposure for the low exposed group, is much smaller relative to the high exposed group, and that the majority of deaths in the low exposed group are background cases. Similarly, it was found that β_3 converged to values close to zero, suggesting that the risk of chrysotile relative to the amphiboles is very low. In light of these preliminary findings, β_1 and β_3 have been fixed at zero in all three models fitted. The estimate of a_1 also converged to values very close to zero in preliminary analysis and had minimal impact on the fit of the model. In light of this, a_1 has been fixed at zero in all three models. The three models fitted are denoted:

- Model R1: $\sigma = 0.05$;
- Model R2: $\sigma = 0.10$;
- Model R3: $\sigma = 0.15$;

3.3 STATISTICAL METHODOLOGY

Preliminary analyses were carried out by fitting the model using the *fminsearch* function in MATLAB (The Mathworks, Inc., 2009) by minimising the model deviance. The Metropolis-Hastings algorithm was then used to provide posterior medians and 90% credible intervals of the model parameters. The *fminsearch* estimates were used as approximate starting values for the Metropolis-Hastings algorithm. Details of *fminsearch* and the Metropolis-Hastings algorithm can be found in Section 2.3.

3.4 RESULTS

The revised risk model was fitted to the dataset using the Metropolis-Hastings algorithm. The results from fitting the different revised risk models can be found in Table 4.

The estimate of the high exposure factor corresponding to amphiboles ranged from around 57 (when $\sigma = 0.15$) to 180 (when $\sigma = 0.05$); the larger the proportion of high exposed, the lower the estimated levels of high-level exposure. This correlation presents a difficulty when fitting the model, as it suggests that when changing the value of one parameter, a similar fit can be obtained by compensating against another parameter. However, the best fit was obtained under Model R1, with a median deviance of 351.5.

The posterior medians of k for all values of σ were larger than the posterior medians of k_b, at around 2.80 and 2.00 respectively with no overlap in the credible intervals. This suggests that the

increase in the asbestos-associated risk with time since first exposure is greater than the increase in the spontaneous mesothelioma risk with time.

The background rate was estimated to be around 0.9 cases per million in the models, corresponding to around 20 cases in 2007 amongst males aged 20 to 89. This estimate is slightly lower than the 23 background cases in males estimated by Tan and Warren (2009).

The proportions of high exposed males who were exposed from the ages of 25 and 35 were estimated to be around 0.50 and 0.36 respectively; this equates to a proportion of around 0.14 who were exposed from the age of 15.

The estimate of the risk factor associated with crocidolite ranged from 0.11 to 0.16, and the estimate of the proportion of stock released from 1980 was around 0.0004.

Table 4 Revised risk model parameter estimates for male data using MCMC

	Posterior median (90% C.I.)		
Parameter	Model R1	Model R2	Model R3
σ (fixed)	0.05	0.1	0.15
$Rate$	0.94 (0.64,1.34)	0.87 (0.55,1.33)	0.85 (0.54,1.28)
k	2.79 (2.74,2.84)	2.77 (2.74,2.81)	2.81 (2.75,2.85)
k_b	2.00 (1.63,2.40)	1.92 (1.47,2.37)	1.89 (1.44,2.35)
β_2	183.2 (154.3,216.7)	98.8 (87.4,110.3)	56.9 (49.8,69.8)
$High_{25}$	0.49 (0.47,0.52)	0.49 (0.47,0.52)	0.50 (0.47,0.53)
$High_{35}$	0.36,0.34,0.39)	0.37 (0.34,0.39)	0.36 (0.33,0.38)
H (fixed)	15	15	15
s (fixed)	100	100	100
$CrocRatio$	0.11 (0.03,0.21)	0.12 (0.03,0.23)	0.16 (0.04,0.28)
a_1 (fixed)	0	0	0
a_2	0.0003(0.00004,0.0006)	0.0005 (0.00006,0.0008)	0.0004(0.00003,0.0007)
Deviance	351.5 (338.1,395.1)	353.2 (338.2,415.9)	357.0 (339.0,418.3)
Degrees of freedom	215	215	215

3.5 Discussion

This section has presented a statistical analysis of male mesothelioma mortality using a revised risk model. The underlying risk function differs from that used in Tan and Warren (2009) in that contributions from previous years' lung burden rather than exposure are summated in the revised risk model. In addition, actual asbestos imports data have been used to obtain estimates of male exposure levels. The male population has also been classified into low and high exposure categories according to age and duration of exposure. Incorporating this exposure classification has allowed us to characterise exposure in terms of lung burden based on empirical data from imports, rather than constructing exposure according to various assumptions in the original Tan and Warren (2009) model. It has also allowed us to estimate mortality rates of low exposed and high exposed males, although from the mortality data itself, it is not possible to distinguish between deaths occurring in low and high exposed categories due to the lack of information on individual exposure histories.

However, the revised risk model has presented us with some challenges. One of the problems that arose was that different sets of parameter values resulted in similar fits, making it difficult to obtain an understanding of the true relationships between exposure and risk. In particular, the value of σ was somewhat arbitrary as there is a whole distribution of exposures which we are representing by just two groups. Clearly the relative difference in the low and high exposure factors will be greater if high exposure related to the highest 5% rather than the highest 10%. In Section 3.4, we found that the best fit was obtained when the proportion of males in the high exposed category was set to 0.05.

Another problem encountered was that a_2, the proportion of stock released after 1980, was difficult to estimate as, due to the lag period between exposure and diagnosis of mesothelioma, the effects of the asbestos stock released after 1980 on mesothelioma mortality in individuals may not be seen for several decades. Although the MCMC analysis provided a median value of 0.0003, it is of interest to investigate the sensitivity of the projections to a_2. The impact of a_2 on the projections is discussed in Section 5.

Tan and Warren (2009) and Section 3 have presented risk models for mesothelioma that have been shown to fit historical mortality data well. However, an alternative approach to constructing a model for mesothelioma is via the likely biological framework for carcinogenesis. Carcinogenesis models have been shown to fit cancer data well and have the advantage of being more closely related to the underlying processes involved in the development of tumours. The TSCE model is one such model and is discussed further in Section 4.

4 TWO-STAGE CLONAL EXPANSION MODEL

4.1 BACKGROUND

The two-stage clonal expansion (TSCE) model was proposed by Moolgavkar and Knudson (1981) for modelling time to tumours and was motivated by biological considerations, in the context of carcinogenesis. The model assumes a susceptible stem cell population of fixed size X and that the transformation of any one of these normal cells to a malignant cell is the result of two critical and irreversible events. The first event is the transformation of a normal cell into one normal cell and one intermediate cell (a cell that has sustained the first mutation) and is referred to as the *initiation* process. An intermediate cell may subsequently undergo a *promotion* process whereby it divides into two intermediate cells, or dies/differentiates (the process during which the cell becomes a specialised cell). The intermediate cell may otherwise undergo the second event whereby a mutation into one intermediate cell and one malignant cell occurs, referred to as the *malignant transformation* process. Once a malignant cell is generated, a malignant tumour develops. In the case of mesothelioma, death occurs within a few years after diagnosis of a tumour. Although several mutations or stages may occur in the process of malignant transformation, it has been shown that the TSCE model with two stages is consistent with many experimental data (Moolgavkar *et al.*, 1988; Hazelton *et al.*, 2005; Moolgavkar *et al.*, 2009). One of the key features of the TSCE model is that it allows for the risk of mesothelioma to increase, reach a plateau, or decrease after exposure to a carcinogen ends. The shape of the risk function is determined by which of the initiation, promotion and conversion steps depend upon exposure, the form of the dose-response and the associated parameter estimates. This makes the TSCE model an attractive and appropriate alternative to the revised risk model fitted in Section 3, which is unable to accommodate an eventual decline in risk.

Certain environmental agents may increase the probability of transition from a normal stem cell to an intermediate cell, the transformation from an intermediate cell to a malignant cell, or the rate of proliferation of intermediate cells. The intensity of exposure to these agents may vary throughout the course of a study. Heidenreich *et al.* (1997) proposed a TSCE model with time-dependent parameters. Exact hazard functions were derived for the case where the exposures were assumed to be piecewise constant. It is well known that mesothelioma is caused mainly by asbestos exposure and there is concern about the carcinogenic risk posed by inhalation of asbestos fibres. In this report, the role of asbestos as an initiator and a promoter in the development of mesothelioma is explored.

TSCE parameters

Figure 7 is a representation of the TSCE model. The key parameters are as follows:

- ν, the rate of initiation;
- α, the rate of cell division;
- β, the rate of cell death and differentiation;
- μ, the rate of malignant transformation;
- The net change in subpopulation of initiated cells may be represented by $\gamma = \alpha - \beta - \mu$, the rate of promotion.

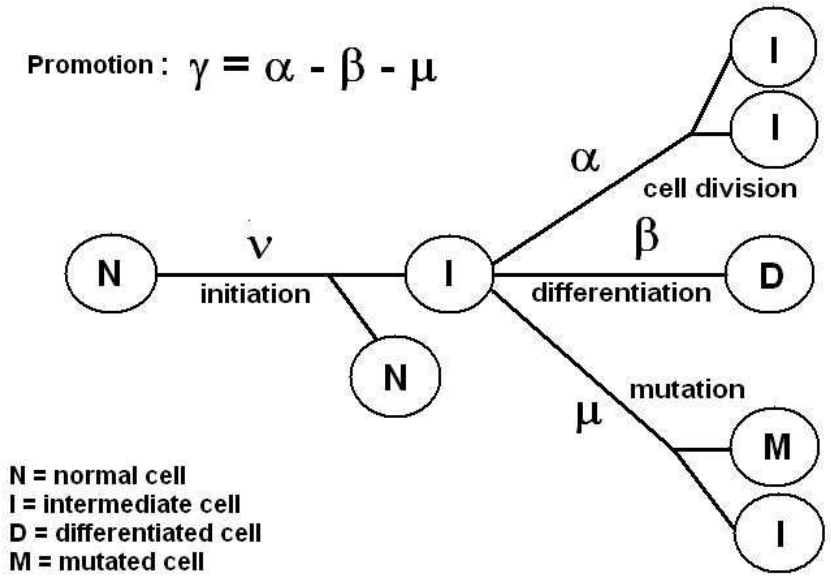

Figure 7 Two-stage clonal expansion model

Exposure to asbestos may affect several parameters in the TSCE model for mesothelioma. Moolgavkar *et al.* (2009) derived hazard functions from the TSCE model in application to pleural and peritoneal mesothelioma, estimating coefficients that adjust the hazard function for birth cohort and calendar year, however a dose-response term was not incorporated in any of the stages of carcinogenesis. Richardson (2009) analysed lung cancer mortality in chrysotile asbestos workers; the best model fit was obtained when a term describing a linear effect of asbestos on the promotion rate γ was included. Hazelton *et al.* (2005) analysed the effects of tobacco smoke on lung cancer mortality in three cohorts, incorporating dose-response in the initiation, promotion and division rates. In the context of mesothelioma in our analysis, we assume that exposure to asbestos may affect the initiation and promotion stages of carcinogenesis. The initiation and promotion rates are parameterised as follows:

$$\nu(d) = \nu_0 + \nu_1 d^{\nu_2} \quad (11)$$

$$\gamma(d) = \gamma_0 + \gamma_1 d^{\gamma_2} \quad (12)$$

where ν_0 is the background initiation rate, γ_0 is the background promotion rate, ν_1 and ν_2 are the asbestos initiation rate coefficient and power, γ_1 and γ_2 are the asbestos promotion rate coefficient and power, and d is the lung burden of an individual. We assume that d is piecewise constant between single year periods.

4.2 REPRESENTATION OF EXPOSURE

As in Section 3.1 the male population has been sub-classified into low and high exposure categories, where males whose exposures were primarily non-occupational (e.g. due to asbestos in the household, environmental exposure) are classified as low exposed, and males whose exposures were primarily occupational (e.g. carpenters, plumbers) are classified as high exposed. A parameter has been included in the model to represent the proportion of males in the high exposure category. Males in the high exposure category are then sub-classified according to the age at which high exposure

started, and the duration of high exposure. The sub-classification of the high exposure group can be found in Table 5. This differs from that in Section 3 in that the longer exposure duration categories have been combined to reduce computational runtime; this is unlikely to have a significant effect on the parameter estimates as the contributions to the total risk will be dominated by the higher proportion of males with shorter exposure durations of 5, 10 and 15 years. We assumed that high exposures could commence at one of three ages (15, 25 or 35 years) and for each, the distribution of exposure duration was estimated based on data for controls in a recent population based case-control study of mesothelioma in Britain (Rake *et al.*, 2009).

Table 5 Sub-classification of the high exposure group

Start age of high exposure	Proportions of males who were exposed for ... years					
	5	10	15	20 to 25	30 to 35	40 to 45
15	0.67	0.21	0.0764	0.0379	0.005	0.0006
25	0.45	0.24	0.09	0.15	0.07	
35	0.37	0.22	0.1	0.31		

We assume $X = 10^5$ susceptible stem cells in the mesothelium. Exposure to asbestos and inhalation of fibres contribute to the lung burden of an individual. We use lung burden as the measure of dose that feeds into the dose-response model. Individuals with different exposure histories will have different dose profiles. We assume that exposure is negligible or below the threshold (below which there is background risk but no exposure-related contribution to the risk of mesothelioma) during the years of low exposure. The lung burden $d(A,T)$ at age A in year T is thus calculated as

$$d(A,T) = \sum_{m \in E} D_{T-A+m} 0.5^{\frac{A-m}{H}} + \beta_3 Inc_{T,chry} \tag{13}$$

where H (half-life), D_T (annual increment of amphibole exposure in year T), β_3 (risk factor of chrysotile relative to the amphiboles) and $Inc_{T,chry}$ (annual exposure increment of chrysotile in year T) were defined in Section 3.1 and the summation is over all years of high exposure.

4.3 THE MODEL

The fitted number of deaths at age A in year T is expressed as

$$\begin{aligned} \lambda_{A,T} \;=\; & P_{A,T} \times \Bigg\{ (1-\sigma) p_{low}(A,T) + \\ & \sigma \left[High_{15} \sum_{j=1}^{9} T_{15,5j} p_{15,j}(A,T) + High_{25} \sum_{j=1}^{7} T_{25,5j} p_{25,j}(A,T) + High_{35} \sum_{j=1}^{5} T_{35,5j} p_{35,j}(A,T) \right] \Bigg\} \end{aligned} \tag{14}$$

where

- $\lambda_{A,T}$ is the number of deaths at age A in year T;
- $P_{A,T}$ is the person-years at risk at age A in year T;
- $p_{low}(A,T)$ is the probability of death at age A in year T of a low exposed male, $p_{15,j}(A,T)$,

$p_{25,j}(A,T)$ and $p_{35,j}(A,T)$ are the probabilities of death at age A in year T of a high exposed male who was exposed for $5j$ years from ages 15, 25 and 35 respectively;

- σ is the proportion of males who are high exposed;
- $High_{15}$ is the proportion of high exposed males who were exposed from age 15;
  $High_{25}$ is the proportion of high exposed males who were exposed from age 25;
  $High_{35}$ is the proportion of high exposed males who were exposed from age 35;
  ($High_{15} = 1 - High_{25} - High_{35}$);
- $T_{15,5j}$ is the proportion of high exposed from age 15 who were exposed for $5j$ years;
  $T_{25,5j}$ is the proportion of high exposed from age 25 who were exposed for $5j$ years;
  $T_{35,5j}$ is the proportion of high exposed from age 35 who were exposed for $5j$ years.

The final stage of carcinogenesis in the TSCE model is the malignant transformation of a cell. The hazard function associated with the probability of tumour (assuming that a malignant cell develops into a tumour and disregarding the lag time between malignant transformation and diagnosis of tumour) has been derived by Heidenreich *et al.* (1997). In a study where individual subjects are being followed, and under the assumption of no lag period between diagnosis of a tumour and death, $p_{low}(A,T) = h_{low}(A,T)s_{low}(A,T)$ where $h_{low}(A,T)$ is the hazard (death rate at age A in year T given that the individual has survived up to T) in the low exposed group and $s_{low}(A,T)$ is the corresponding survival function (probability of survival beyond year T). However we assume a lag period of 10 years between malignant transformation and death, thus the probability becomes $p_{low}(A,T) = h_{low}(A-10,T)s_{low}(A,T)$. When we are apply these probabilities to the actual (or projected) population alive in year T rather than following up individual subjects over time, as is the case in our analysis, $p_{low}(A,T) = h_{low}(A-10,T)$. p_{15}, p_{25} and p_{35} are similarly calculated. The parameters ν_0, α, β, μ and γ_0 enter the model via h_{low} and h_{15}, h_{25} and h_{35}; the dose-response parameters ν_1, ν_2, γ_1 and γ_2 enter the model via h_{15}, h_{25} and h_{35} only.

Explicit calculations for the hazard and survival functions have been described in detail elsewhere (Heidenreich *et al.*, 1997).

Models fitted

The TSCE model contains several unknown parameters and non-identifiability arises; a necessary condition for identifiability is to put a constraint on the parameters. Following Heidenreich *et al.* (1997), we make the assumption that $\nu_0 = \mu$, that is that the background initiation rate is equal to the malignant transformation rate. This assumption has also been made by Moolgavkar *et al.* (2001) when analysing the effects of long man-made fibres on lung cancer risk and Richardson (2009) when analysing the lung cancer in chrysotile asbestos workers. It was found through exploratory analysis that several different combinations of parameters resulted in similar fits. In light of this, δ was fixed at 0.02, H was fixed at 15, $CrocRatio$ was fixed at one, a_1 was assumed to equal a_2 and various values of σ were investigated.

Although an asbestos initiation rate power ν_2 was initially included in the model, it was found to be close to one. When compared to the reduced models with just an initiation rate coefficient, this power term did not greatly improve the model fit. We subsequently removed ν_2 from the model for reasons of parsimony. Three models were subsequently fitted, T1, T2 and T3 where $\sigma = 0.05, 0.10$ and 0.15 respectively.

4.4 STATISTICAL METHODOLOGY

Preliminary analyses were carried out by fitting the model using the *fminsearch* function in MATLAB (The Mathworks, Inc., 2009) by minimising the model deviance. The Metropolis-Hastings algorithm was then used to provide posterior medians and 90% credible intervals for the model parameters. The *fminsearch* estimates were used as approximate starting values for the Metropolis-Hastings algorithm. Details of *fminsearch* and the Metropolis-Hastings algorithm can be found in Section 2.3.

4.5 RESULTS

Dose-response on initiation and promotion

The posterior medians of the TSCE parameters ν_0, α and γ_0 did not differ greatly between the three models, with estimates (per cell per year) of around 2×10^{-7}, 400 and 0.04 respectively. Assuming 10^5 stem cells in the mesothelium, the estimate of ν_0 suggests a background number of 0.02 initiated cells per year (equivalent to an average period of around 50 years for a stem cell to undergo the initiation process during periods of no exposure). The estimate of α suggests around 400 cell divisions per cell per year; the estimate of γ_0 suggests that an initiated stem cell will, on average, make a faulty division into two initiated stem cells once every 25 years during periods of no exposure. The parameters ν_0, α and γ_0 alone determine the relative background risk, resulting in an estimate of about 9 background cases in 2007. The estimate of γ_1 was largest for T3 and smallest for T2, at 1.37 and 1.00 respectively. However a pattern emerges in the posterior medians of two of the dose-response parameters; both γ_2 and ν_1 decrease as σ increases. The reason for this is likely to be due to the dose-response parameters compensating for the increase in the proportion of high-exposed males by decreasing the contribution of the dose in the development of a tumour. The posterior medians of the parameters associated with the structure of the exposure groups β_3, a_2, $High_{25}$ and $High_{35}$ also did not differ greatly between the three models, with estimates of around 0.001, 0.002, 0.45 and 0.35 respectively. Posterior medians and 90% credible intervals for the three models can be found in Table 9. Projections of male mortality from 2008 to 2050 under the T2 model can be found in Section 5.

Table 6 MCMC: Parameter estimates

Parameter	Posterior median (90% credible interval) Model T1	Model T2	Model T3	Parameter estimates Model T4
σ (fixed)	0.05	0.10	0.15	0.10
α	326.1 (303.3,361.2)	426.7 (398.6, 478.1)	417.4 (383.4,461.5)	3428.4
γ_0	0.040 (0.039,0.040)	0.042 (0.040, 0.043)	0.038 (0.037,0.039)	0.165
γ_1	1.07 (1.03,1.14)	1.00 (0.96, 1.10)	1.37 (1.32,1.44)	-
γ_2	0.100 (0.095,0.102)	0.097 (0.091, 0.100)	0.081 (0.077,0.085)	-
ν_0	2.64 $\times 10^{-7}$ (2.45 $\times 10^{-7}$, 2.75 $\times 10^{-7}$)	2.0 $\times 10^{-7}$ (1.9×10^{-7}, 2.1×10^{-7})	2.0×10^{-7} (1.9×10^{-7},2.1×10^{-7})	3.3075×10^{-8}
ν_1	0.0082 (0.0077,0.0093)	0.0070 (0.0065, 0.0074)	0.0046 (0.0043,0.0050)	0.035
ν_2	-	-	-	1.19
$High_{25}$	0.45 (0.43,0.47)	0.45 (0.43, 0.47)	0.45 (0.43,0.48)	0.44
$High_{35}$	0.35 (0.33,0.37)	0.35 (0.33, 0.37)	0.34 (0.32,0.36)	0.38
$HighProp$ (fixed)	0.05	0.10	0.15	0.10
H (fixed)	15	15	15	15
$CrocRatio$ (fixed)	1	1	1	1
β_3	0.0009 (0.00008,0.0027)	0.0012 (0.0001,0.0029)	0.0010 (0.00003,0.0028)	0.0016
$a_1 = a_2$	0.002 (0.0003,0.004)	0.003 (0.0005, 0.006)	0.002 (0.0002,0.006)	0.006
Deviance	349.8 (342.5,356.0)	341.2 (336.4,446.1)	342.7 (338.2,347.4)	345.7
Degrees of freedom	214	214	214	215

Dose-response on initiation only

Although the parameter estimates for T1, T2 and T3 indicated the significance of dose-response on both the initiation and promotion rates, the effects of fitting a reduced model with dose-response on initiation only was also evaluated. Model T4 which incorporates an initiation rate coefficient and power in the absence of promotion rate parameters was fitted using *fminsearch*. The parameter estimates can be found in Table 6.

The posterior medians of the background rate parameters ν_0, α and γ_0 differed from those estimated by T1, T2 and T3. ν_0 was estimated at around 3×10^{-8} per cell per year, suggesting a background number of 0.003 initiated cells per year (equivalent to an average period of over 300 years for at least one stem cell to undergo the initiation process during periods of no exposure); this is much lower than the rates estimated by T1, T2 and T3. α was estimated at around 3400 cell divisions per cell per year. γ_0 was estimated at 0.17 per cell per year, suggesting that an intiated stem cell will, on average, make a faulty division into two intitiated stem cells approximately once every 6 years during periods of no exposure; this estimate is higher than those estimated by T1, T2 and T3, suggesting that although there is a lower background initiation rate for T4, once stem cells undergo the initiation process, there is a higher net rate of promotion.

The posterior medians of the parameters associated with the structure of the exposure groups $ChryRatio$, a_2, $High_{25}$ and $High_{35}$ also did not differ greatly from T1, T2 and T3, with estimates of 0.0016, 0.006, 0.44 and 0.38 respectively.

A comparison of the projections of male mortality from 2008 to 2050 under model T4 can be found in Section 5. A comparison of the background rates reveals another difference between the two dose-response models; whereas the number of estimated background cases in 2007 using a dose-response model on initiation and promotion was 9, only 2 background cases were predicted when initiation only depended on exposure.

4.6 HAZARD FUNCTION

In the context of carcinogenesis where time-to-tumour is being analysed, the hazard function is the rate of developing a tumour in the next short period of time conditional upon having survived. A lag period usually exists between development of a cancerous tumour and death; in our analysis, we have assumed a lag of 10 years.

We have also assumed that males in Great Britain fall into low and high exposure groups with the high exposed males being sub-classified according to age and duration of exposure, and as such, different hazard functions will apply to the different subcategories. In addition, as the asbestos exposure profile varies from year to year, different birth cohorts with the same starting age and duration of exposure will have been exposed to different levels of asbestos over their lifetime. We have assumed that the exposure levels of those in the low exposure group, and in the low exposure years of those in the high exposure group, are negligible.

Figures 8 and 9 show the lung burden for a selection of high exposed males: those who were born in 1930 and 1950 exposed from age 15 for 10 years, age 15 for 40 years, age 25 for 10 years and age 35 for 10 years. It can be seen that the lung burden for males born in 1930 increases for the duration of exposure, subsequently falling after exposure ends. The exception is for those males who were exposed from age 15 for 40 years, for whom the lung burden falls from the age of about 45 (after about 30 years of exposure) due to exposure levels rapidly decreasing from the early 1970s onwards. The lung burden in males born in 1950 was lower than that of their 1930s counterparts (with the exception of males exposed from age 15 for 10 years) due to the lower levels of post-1970s exposure; a male born in 1930 who was exposed from the age of 35 for 10 years would have been subject to a lung burden almost 20 times that of one born in 1950.

Figure 10 shows the hazard functions for low exposed males (assuming negligible exposure levels) using the best-fitting dose-response models for initiation and promotion, and initiation only (the optimal parameter estimates have been obtained using Matlab's *fminsearch* function). There is a difference in the shape of the hazard functions, with the dose-response model on initiation only resulting in a much greater rate of change from the ages of 20 to 40 before levelling off around age 50. In contrast, the dose-response model on initiation and promotion reveals a more gradual increase in the rate which continues to increase into older ages.

Figures 11 and 12 show the hazard functions (for T2) for males with two different exposure histories born in 1900, 1920, 1940, 1950 and 1960. It can be seen that the shape of the hazard function depends on both the year of birth and the exposure history of the individual. From the figures, it can be seen that out of the eight different categories of males considered, the highest probability of tumour at a single age occurs in males born in 1920 who were exposed at age 15 for 40 years, at the age of about 83; the associated probability is about 0.017. Males born in 1940 also have a relatively high probability (0.015) of developing a tumour around the age of 70.

Figures 13 and 14 show the cumulative hazard functions (for T2) for males with two different exposure histories, for males born in 1900, 1920, 1940, 1950 and 1960. It can be seen that over the eight different exposure categories of males considered, the highest cumulative hazards (over all ages) occur in males in the 1940 birth cohort who were high-exposed from age 15 for 40 years.

Figures 15 and 16 show the cumulative hazards at age 89 (for T2) for males with various exposure histories born between 1900 and 1980. Out of the eight exposure categories considered, the highest cumulative hazards at age 89 occur in males in the 1930 to 1940 birth cohorts who were exposed from age 15 for 40 years. It can be seen that for the lower starting age of exposure of 15 years, the highest cumulative hazards at age 89 occur in the 1930 to 1940 and the 1940 to 1950 birth cohorts for exposure durations of 40 and 10 years respectively. For the higher starting age of exposure of 35 years, it is the earlier 1920 to 1930 birth cohorts who have the highest cumulative hazards at age 89 for all exposure durations. Table A4 in the Appendix presents the cumulative hazard at age 89 for males in 1900 to 1980 birth cohorts with various exposure histories.

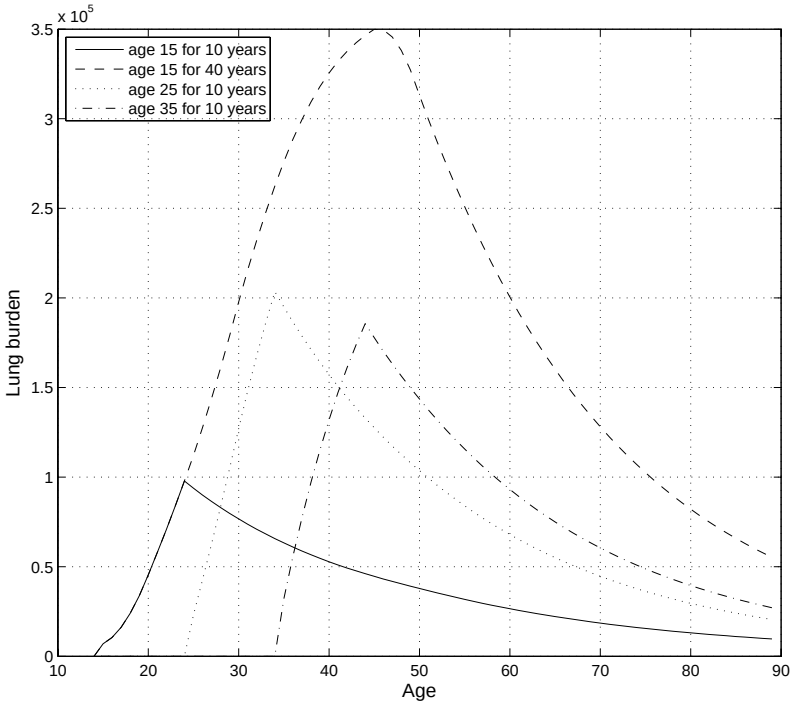

Figure 8 Lung burden of high-exposed males born in 1930 with different exposure histories

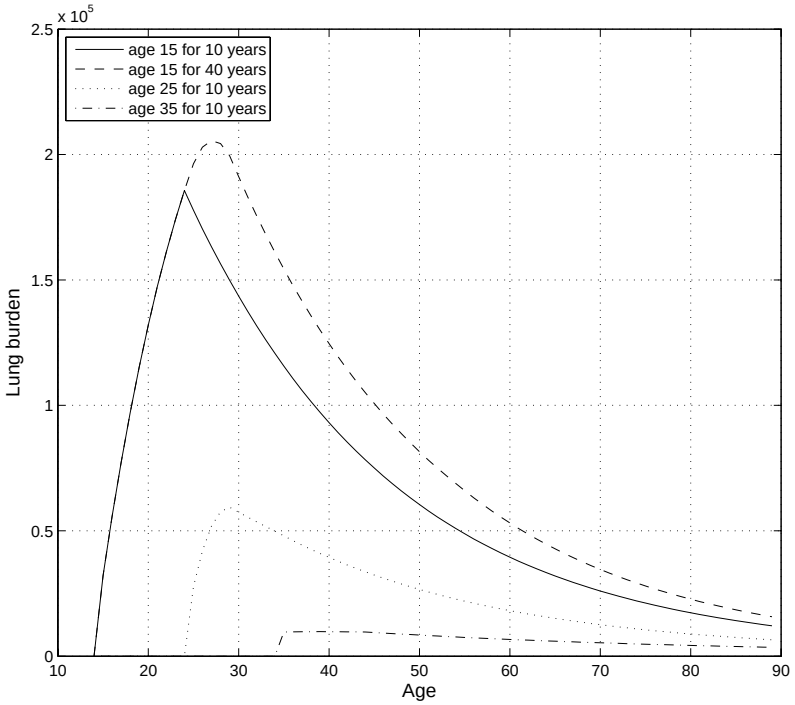

Figure 9 Lung burden of high-exposed males born in 1950 with different exposure histories

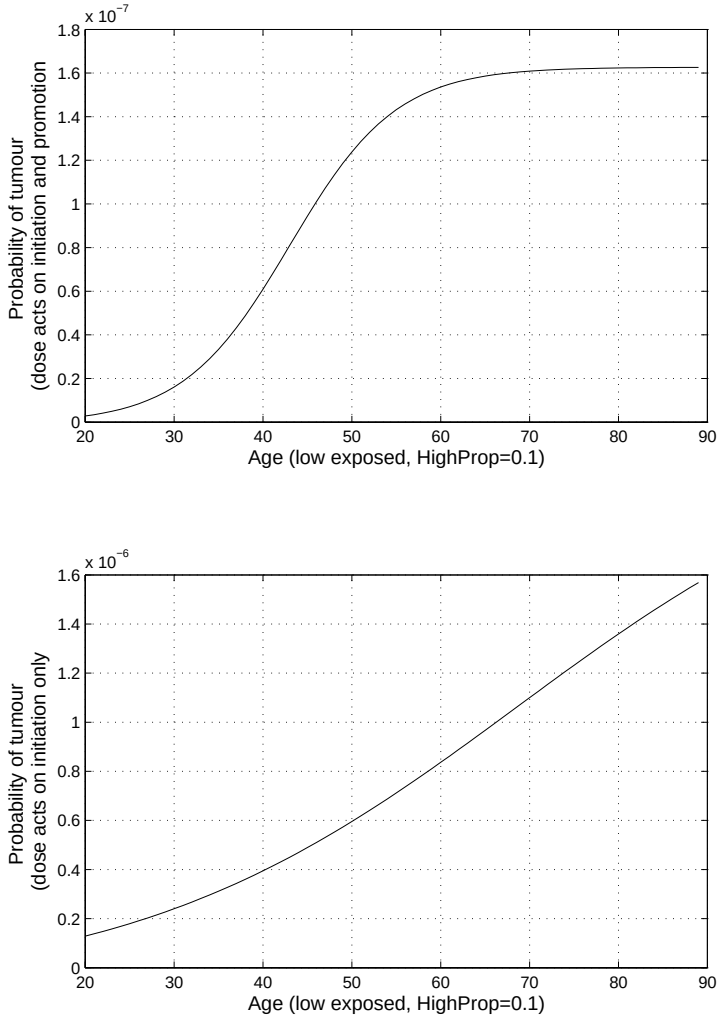

Figure 10 Probability of tumour for a low exposed male (top) dose-response on initiation only (bottom) dose response on initiation and promotion

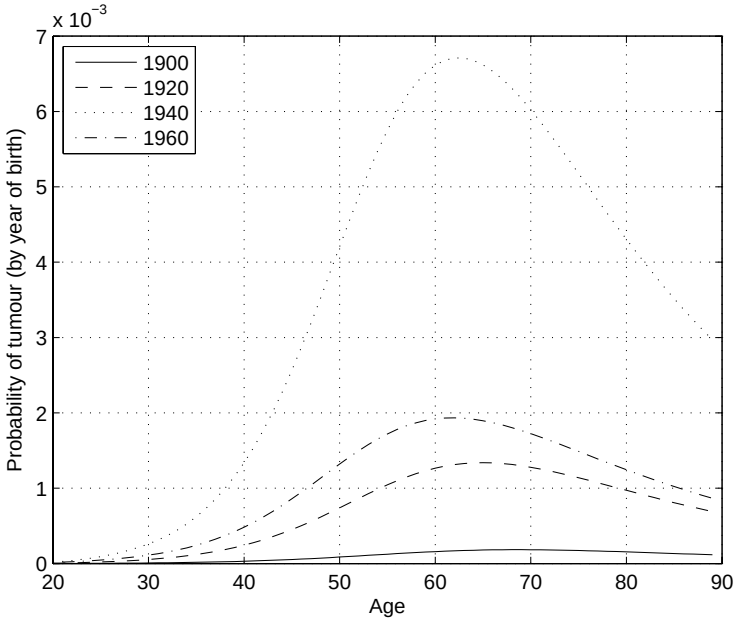

Figure 11 Probability of tumour by year of birth for a male high exposed from age 15 for 10 years

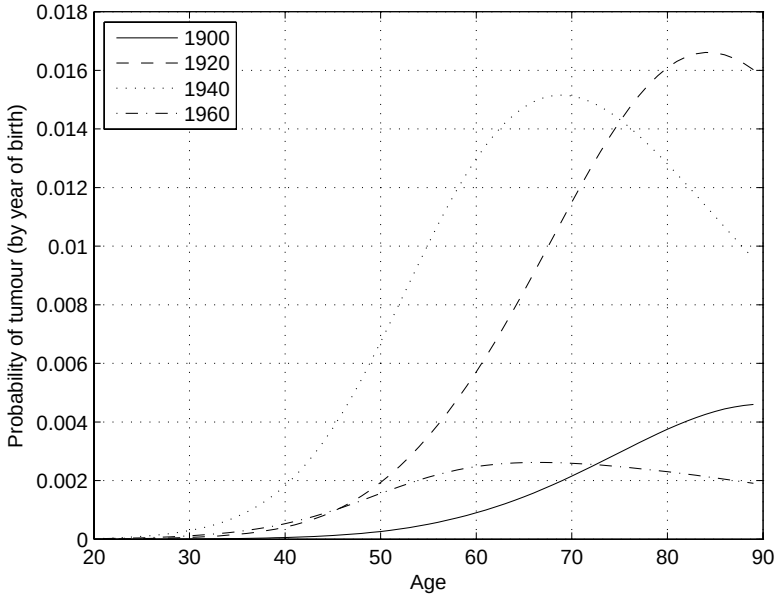

Figure 12 Probability of tumour by year of birth for a male high exposed from age 15 for 40 years

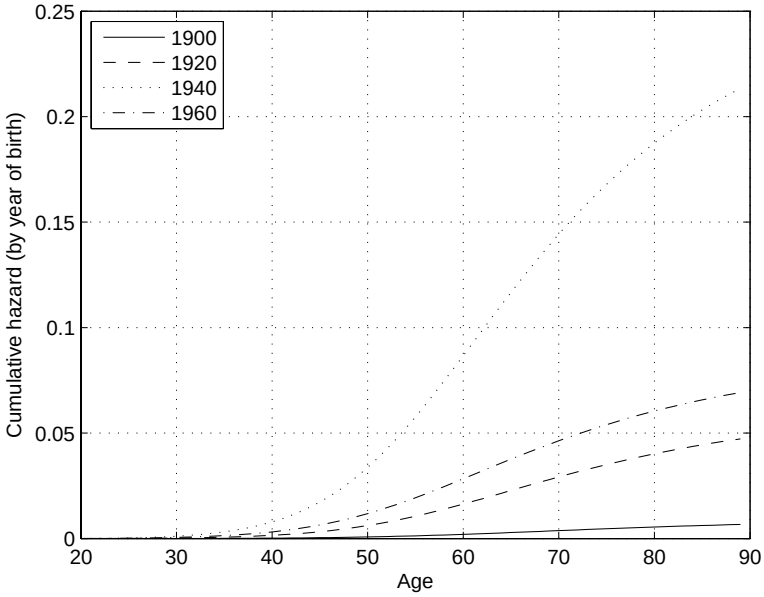

Figure 13 Cumulative hazard by year of birth for a male high exposed from age 15 for 10 years

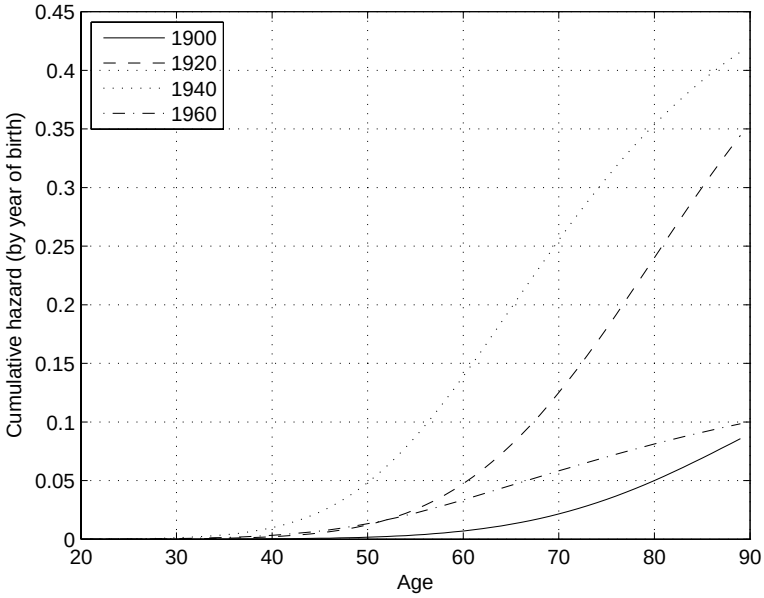

Figure 14 Cumulative hazard by year of birth for a male high exposed from age 15 for 40 years

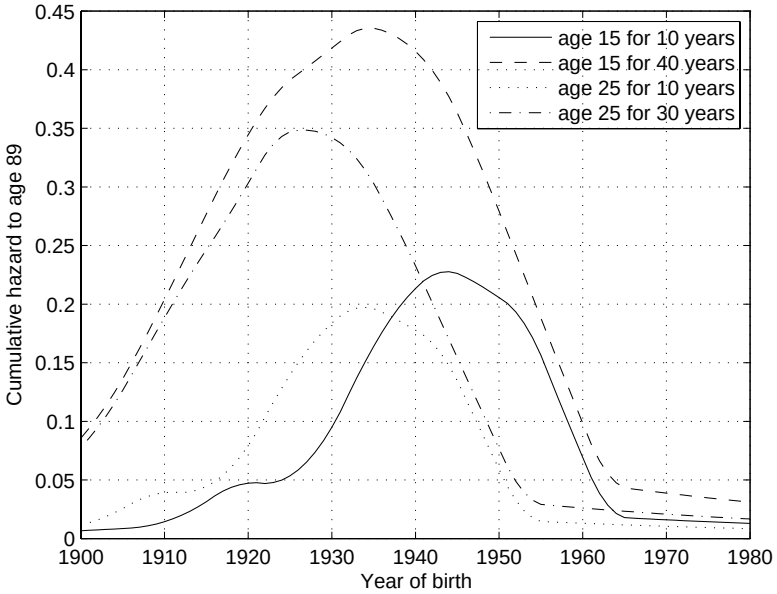

Figure 15 Cumulative hazard at age 89 for males born between 1900 and 1980 with various exposure histories

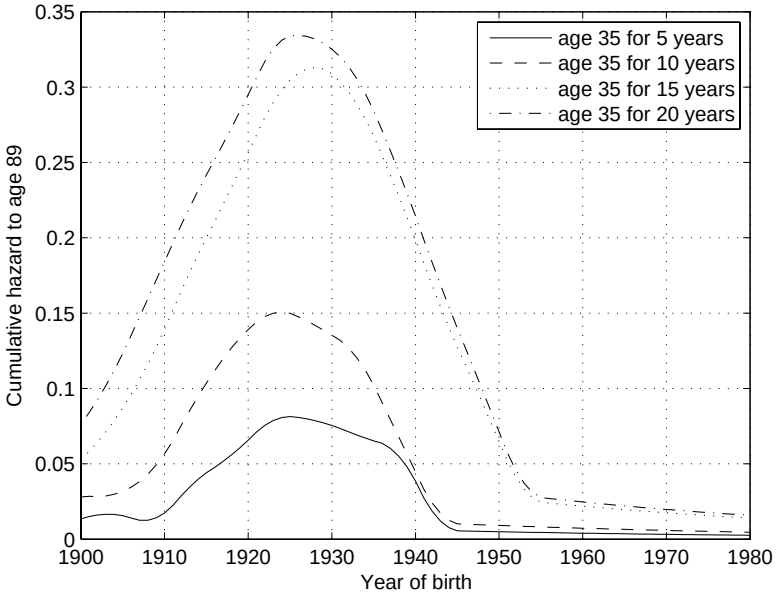

Figure 16 Cumulative hazard at age 89 for males born between 1900 and 1980 with various exposure histories

5 MALE PROJECTIONS

Comparison of projections

Sections 3 and 4 have presented two quite different models for mesothelioma mortality in males; the revised risk model and the TSCE model. The former assumed a risk function that increases with time since exposure to asbestos; the latter was motivated by biological carcinogenesis and assumed that a cell undergoes two mutations prior to development of a tumour. Both models allowed for time-varying exposures, and the analysis was carried out on the population of Great Britain as a whole. It was found that both the revised risk model and the TSCE model fit the data well, with a deviance of 352 obtained under the optimal revised risk model and 341 under the optimal TSCE model.

As well as providing a good fit to historical mortality data, the models should be able to provide us with predictions of future mortality rates. It is obviously difficult to establish how reliable the predictions are under either model until future mortality data is available. In addition, several assumptions have been made on historical exposure patterns; for example, assumptions have been made on exposure levels post-1980, and the proportion of high-exposed males has been fixed at 0.1 under model T2. We have therefore tested the the impact of changing these parameters on the future projections.

Figures 17A to 17D and 18A to 18D show plots of the fitted and observed deaths for males aged 20 to 89 by year of birth, age and year of death for models R1 and T2 respectively; these figures show similar fits under the two models. Figure 19 shows the median fitted deaths up to 2050 (with 90% prediction intervals) for R1; the number of deaths reaches a peak of 1858 in 2012. Figure 20 shows a comparison of the median projections under R1, R2 and R3; the projections under R1 and R2 are similar, with a median projection of just under 200 deaths in 2050, however R3 results in a slightly lower peak and much lower long-term projections.

Figure 21 shows the median fitted deaths (with 90% prediction intervals) for T2; the number of deaths reaches a peak of 1775 in the year 2010. Figure 22 shows a comparison of the median projections under T2 (dose-response on initiation and promotion) and T4 (dose-response on initiation only), where it can be seen that both models result in a similar peak number and year of deaths.

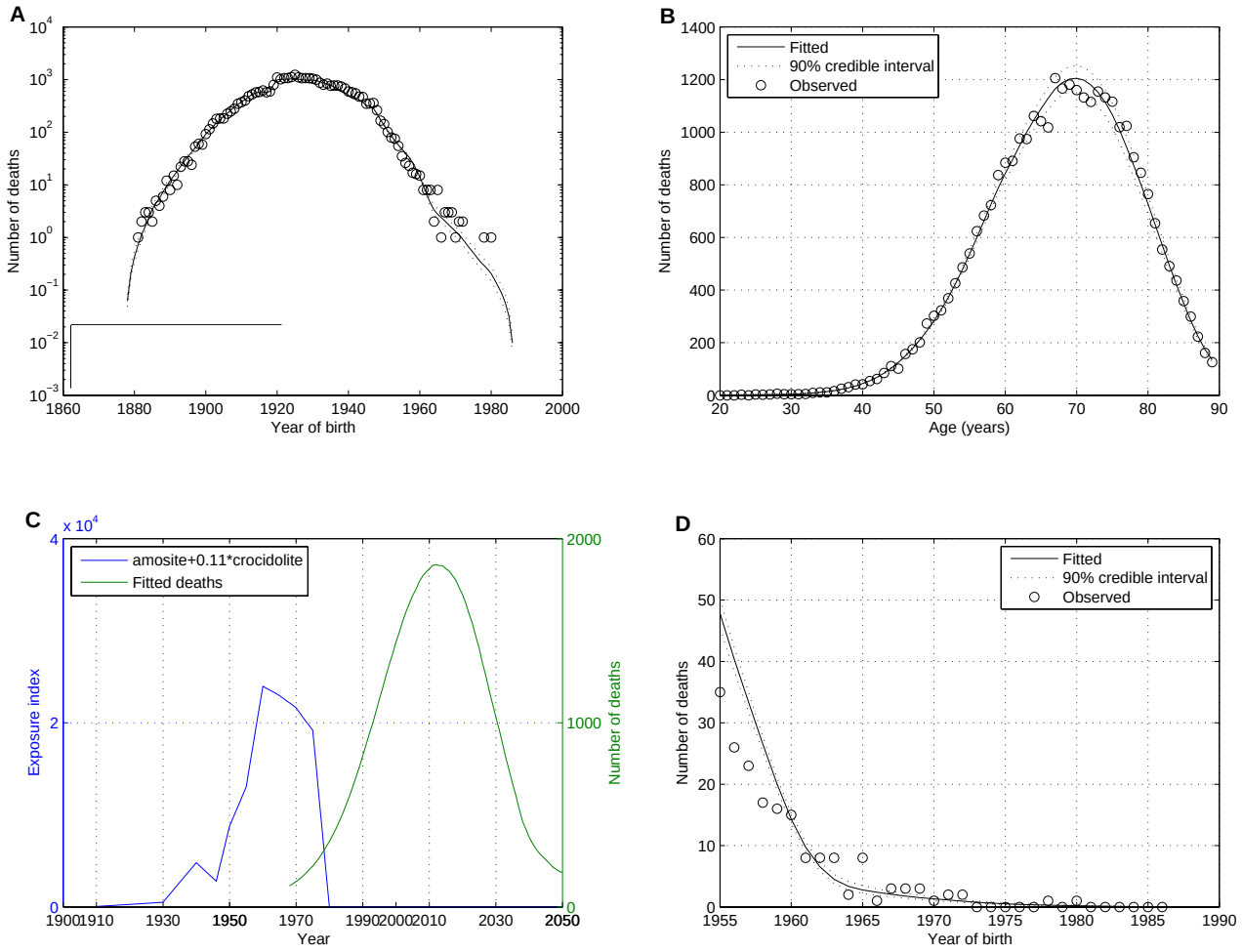

Figure 17 Revised risk model, $\alpha = 0.05$ (A) Observed and fitted deaths by year of birth. (B) Observed and fitted deaths by age. (C) Fitted deaths and asbestos imports. (D) Observed and fitted deaths for 1955-1985 birth cohorts.

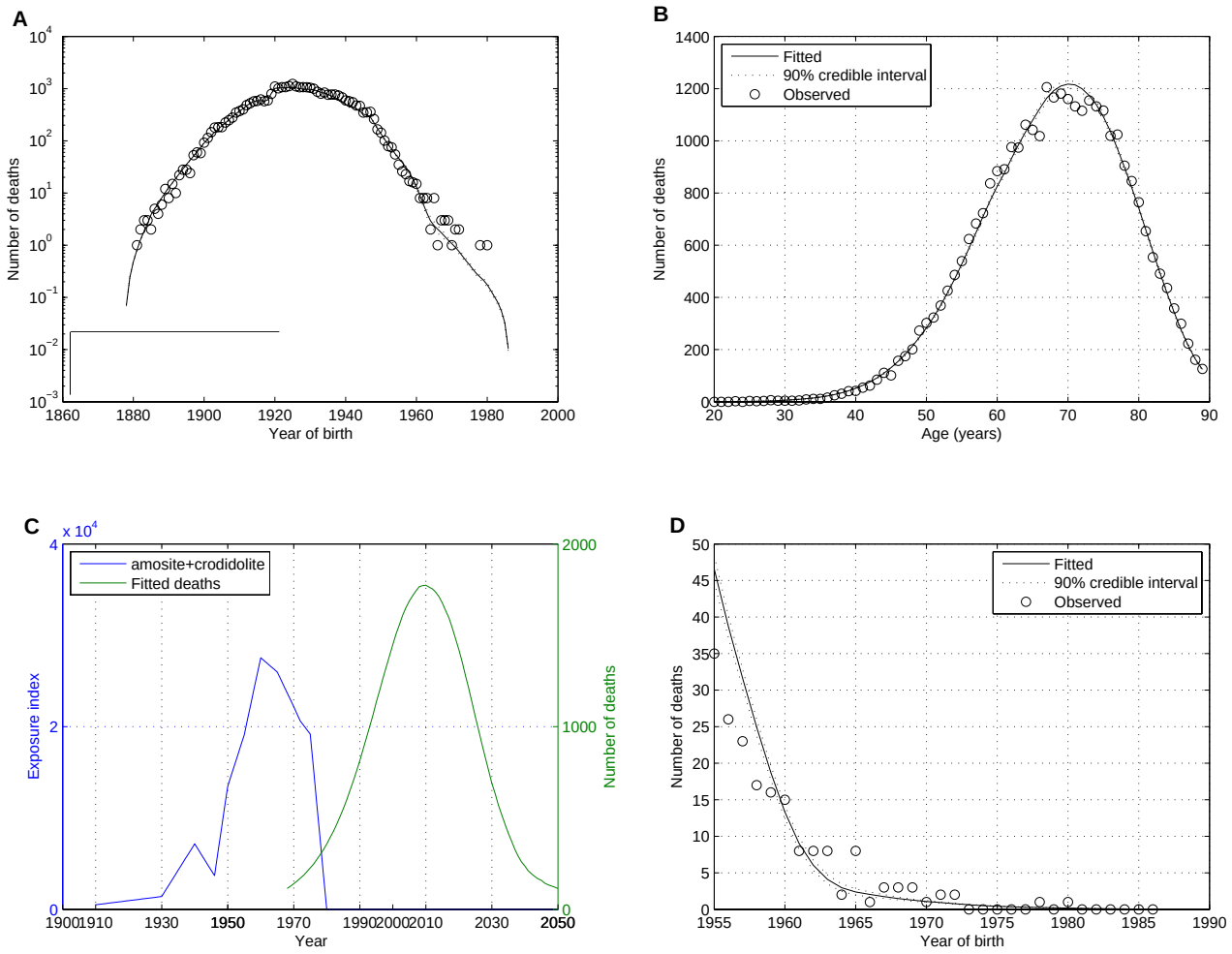

Figure 18 Dose-response on initiation and promotion with $\alpha = 0.1$: (A) Observed and fitted deaths by year of birth. (B) Observed and fitted deaths by age. (C) Fitted deaths and asbestos imports. (D) Observed and fitted deaths for 1955-1985 birth cohorts.

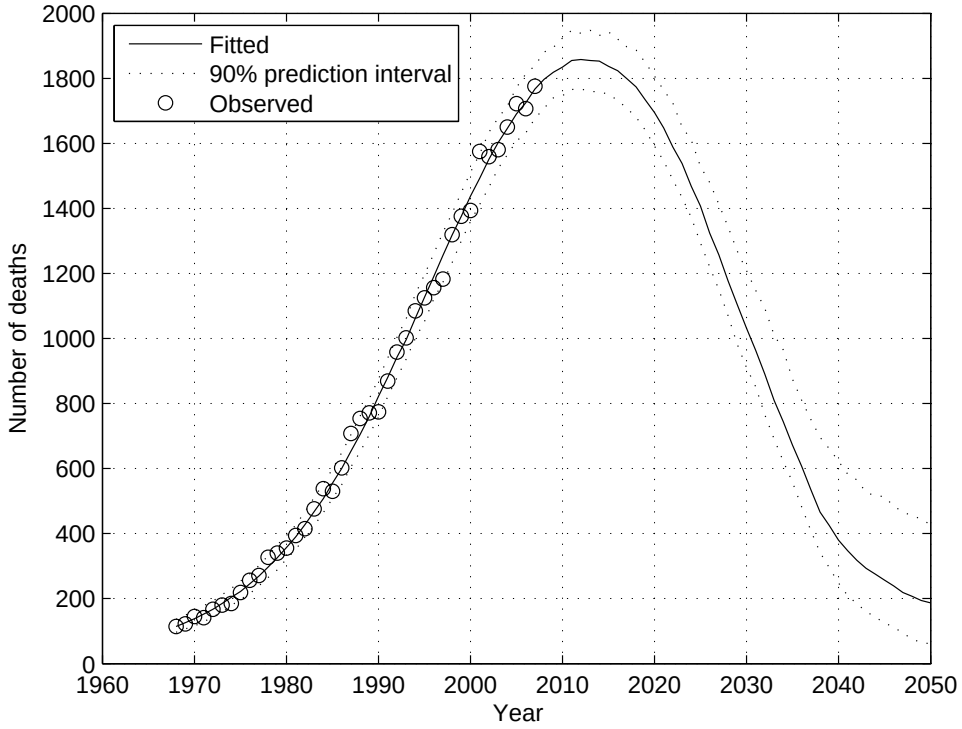

Figure 19 Observed and fitted deaths for R1, $\alpha = 0.05$

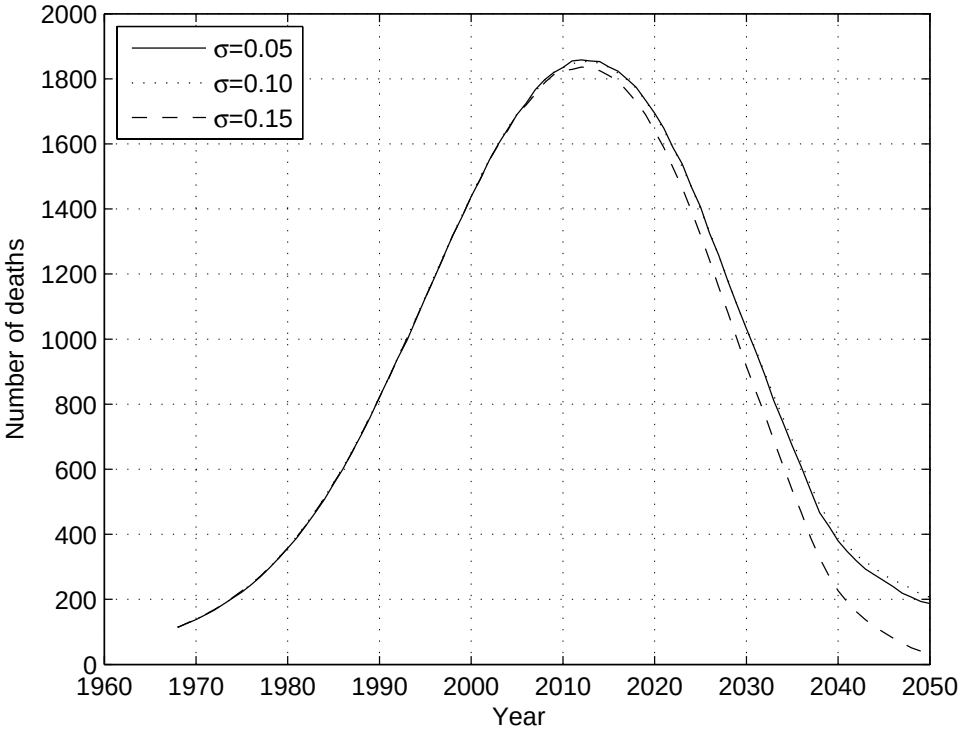

Figure 20 Fitted deaths for R1, R2, R3

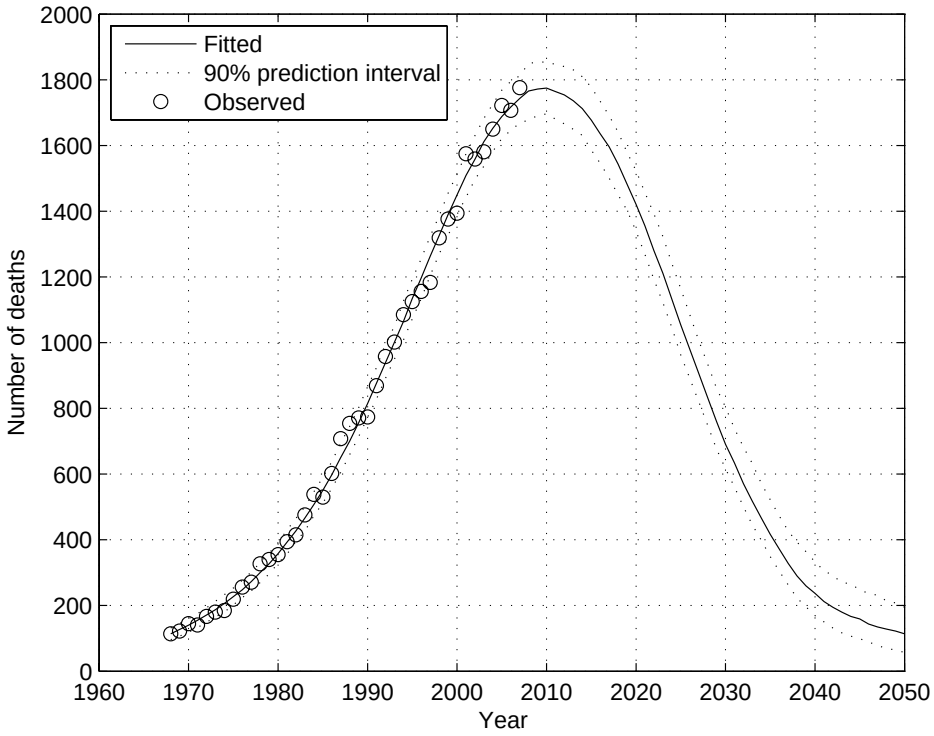

Figure 21 Observed and fitted deaths for T2 (dose-response on initiation and promotion)

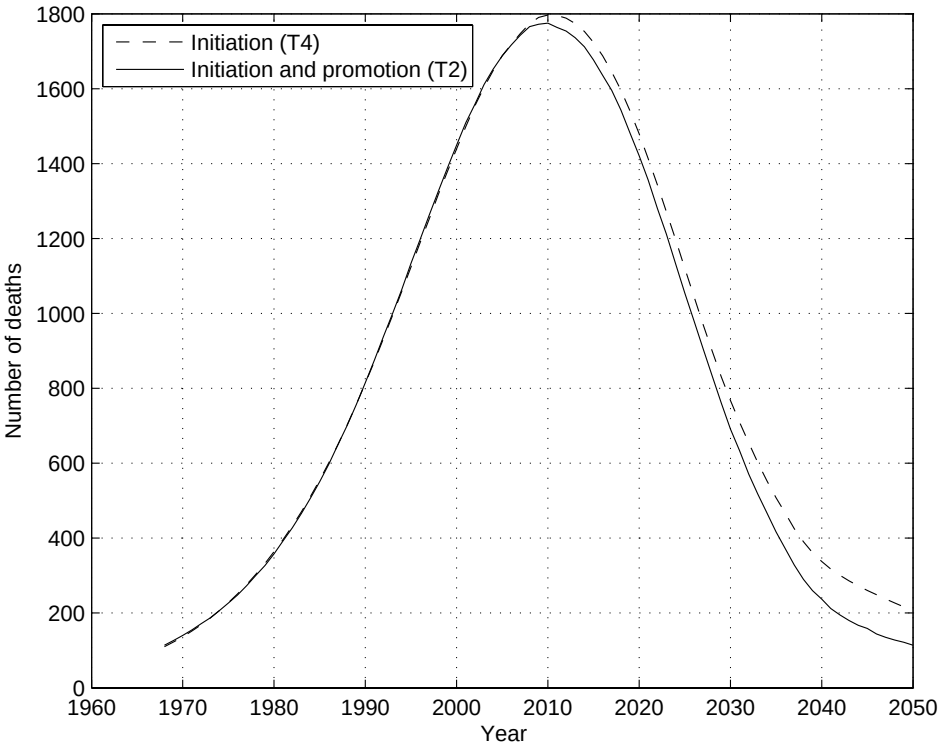

Figure 22 Comparison of the fitted deaths for T2 and T4

Background rates

The estimates of the background rates differ between R1 and T2; around 20 and 9 background cases in 2007 were estimated by the R1 and T2 models respectively. These figures are lower than the 23 cases estimated by Tan and Warren (2009) and the 30 cases estimated in Section 2 for females.

Impact of lag assumptions on projections under the TSCE model

The lag has been assumed to be 10 years in Tan and Warren (2009) and in the models investigated in this report. In the Tan and Warren (2009) and R1 models, the lag is the time period between exposure and its contribution to the risk of mesothelioma; in the TSCE model, the lag is the time period between malignant transformation and death. These assumptions mean that no deaths occur in males aged under 10 years. The concept of lag is different in the TSCE framework and it is thus less obvious that a lag of 10 is appropriate. The assumption of a 10 year lag has been widely assumed for mesothelioma (e.g. Hodgson *et al.*, 2005; Peto *et al.*, 1995) however it is of interest to investigate the sensitivity of the projections on the lag assumption; we have thus tested the impact of the lag on projections. Figure A2 and Table A5 in the Appendix show a comparison of the projections under the assumption of a 10, 12 and 15 year lag for the TSCE model; projections for the latter two were obtained by refitting the TSCE model using *fminsearch* and using the optimised set of parameters as point estimates.

We can see that when we increase the lag, the optimised set of parameters produces an increase in the peak number of deaths with 1811 and 1830 deaths projected under a 12 and 15 year lag respectively, bringing the short-term projections closer to those predicted by Tan and Warren (2009) and R1; the peak year remains unchanged at 2010. The projections under the three lag assumptions begin to converge around 2035, after which those for the 15 year lag decline more rapidly to just 40 deaths in 2050. Long term projections are thus sensitive to the lag assumption under the TSCE modelling framework, however the peak numbers and peak years of death projected do not differ greatly.

Impact of exposure assumptions on projections

Tan and Warren (2009) estimated the population-level exposures in their model, whereas the exposures in our R1 and T2 models are based on a combination of the imports of the different fibre types. Figure 23 shows a comparison of the normalised exposure indices for the three models. The profiles between the years 1900 and 1980 follow a similar pattern with a global peak exposure period between around 1950 and 1970, however the Tan and Warren model had two local peaks before 1950, compared to just one under R1 and T2. In the Tan and Warren model the exposure in 2000 was fixed at 4% of the peak value, based on an assessment of the current distribution of exposures set out in a recent HSE Regulatory Impact Assessment (HSE, 2002). In the R1 and T2 models, estimated exposure was based on annual asbestos imports data; an annual stock removal rate of 2% was assumed and the resultant exposures in the high risk group at 2000 (using the posterior medians as point estimates of the exposure parameters) were 0.4% and 3.4% of the peak respectively. This difference is due to the presence of a crocidolite risk factor and a stock released parameter in the models - the crocidolite risk factor is estimated at 0.11 in R1 and fixed at 1 in T2; the stock released parameter is estimated at 0.0003 (post-1980) in R1 and 0.003 in T2, which has resulted in significant differences in exposures post-1980 between the two models. Long-term projections are highly dependent on the proportion of stock removed as well as the proportion of stock released; increasing the value of annual stock removed will lower long-term projection whereas increasing

the value of stock released will increase the projections. There is thus some uncertainty in both of these parameters which in turn give rise to uncertainty in the long-term forecasts. We have therefore tested the impact of changing these parameters on the future projections.

Figure 24 shows a comparison of the median projections under the Tan and Warren model, R1 and T2. It can be seen that under T2, the peak year occurs earliest (in year 2010) and at the lowest peak value (1775 deaths); under the Tan and Warren (2009) model, the peak level occurs latest (in year 2016) and at the highest peak value (1990 deaths). Table 7 shows projections of male mortality from 2007 to 2050 for the Tan and Warren (2009) model, R1 and T2.

Figure 25 shows the effects of assuming that no stock was released after 1980 under the Tan and Warren, R1 and T2 models, with the other model parameters fixed at their posterior median values. A large difference between the projections still exists even with the assumption of no stock released post-1980, with the largest numbers projected for the Tan and Warren model and smallest for T2; however they eventually converge to under 50 deaths in 2050.

Impact of stock released assumptions on projections under the TSCE model

The estimates of a_2 under R1 and T2 presented in Sections 3 and 4 were quite small, with confidence intervals of (0.00004,0.0006) and (0.0005,0.006) respectively, suggesting that only a small proportion of the cumulative stock is removed each year under the model assumptions. However, it is likely that these confidence intervals will change with the model assumptions. The sensitivity of the projections to the post-1980 stock released parameter (a_2) under the revised risk model has been investigated; Figure 26 shows the projections under the median values of the parameters obtained under Model R1, with a_2 set to 0.0003 (the posterior median obtained under Model R1), 0.002, 0.006 and 0.01. It can be seen that increasing a_2 led to a higher number and delayed peak year of deaths. However, the largest impact was on the long-term forecasts, and these should therefore be regarded as very uncertain. It was found that increasing the parameter moved forward the peak year and increased the peak number of deaths. A similar investigation has been carried out for the TSCE model. Figure 27 and Table A5 in the Appendix show projections for different proportions of stock released post-1980.

It can be seen that the larger the proportion, the larger the peak number of deaths from 2007 onwards, and, in the case of $a_2 = 0.02$, a later peak year of deaths of 2013. Unlike the patterns seen in Figure A2 (where the curves are seen to converge around the year 2035), the curves in Figure 28 diverge, with over 700 deaths in 2050 predicted for $a_2 = 0.02$, compared to under 20 for $a_2 = 0$. Both short and long term projections are thus sensitive to the stock released post-1980 assumptions under the TSCE modelling framework.

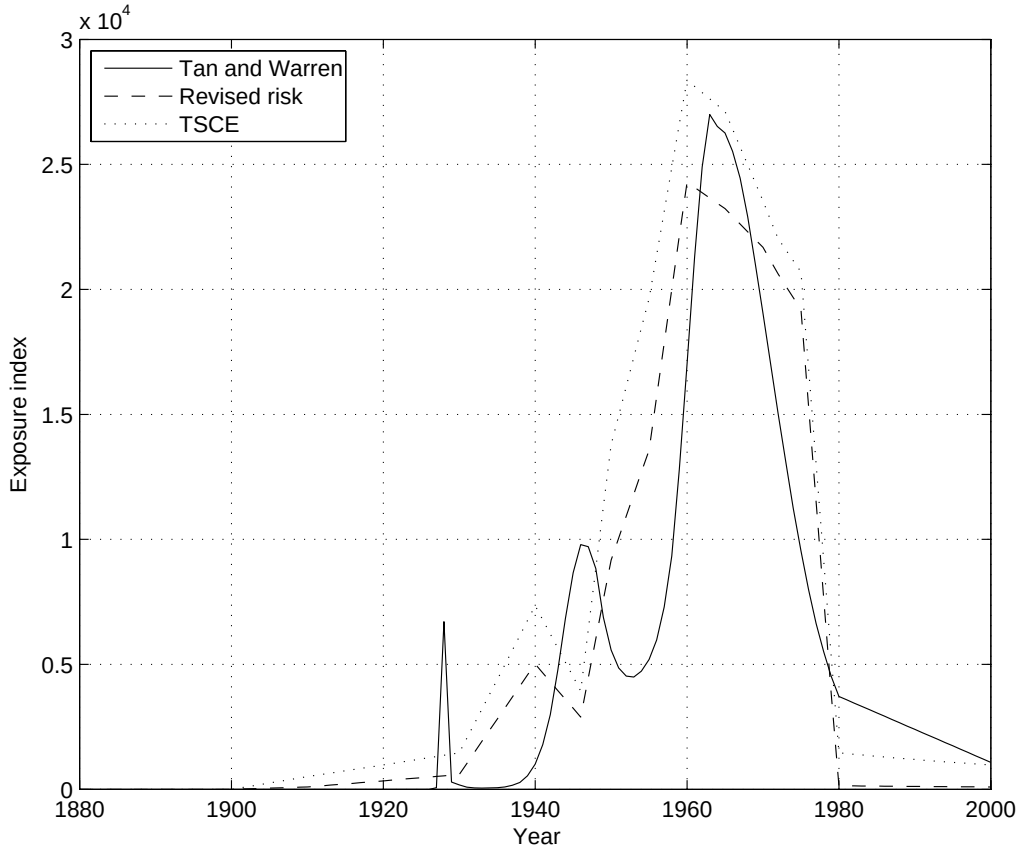

Figure 23 Comparison of the normalised exposure index under the three models

Table 7 Projections of male mesothelioma deaths

Year	Projection [90% prediction interval]		
	Tan and Warren (2009)	R1	T2
2007	1791 [1715,1864]	1767 [1677,1839]	1743 [1668,1812]
2008	1835 [1755,1920]	1797 [1708,1881]	1766 [1678,1850]
2009	1869 [1788,1953]	1819 [1740,1895]	1772 [1691,1854]
2010	1902 [1817,1990]	1835 [1747,1924]	**1775** [1696,1854]
2011	1926 [1842,2015]	1855 [1766,1948]	1764 [1680,1847]
2012	1947 [1859,2042]	**1858** [1768,1939]	1754 [1666,1840]
2013	1964 [1874,2062]	1855 [1760,1948]	1736 [1649,1830]
2014	1979 [1881,2079]	1853 [1759,1932]	1712 [1631,1809]
2015	1988 [1886,2099]	1837 [1751,1943]	1677 [1586,1777]
2016	**1990** [1885,2100]	1824 [1733,1921]	1636 [1546,1735]
2017	1988 [1875,2100]	1799 [1708,1897]	1596 [1501,1691]
2018	1978 [1870,2100]	1774 [1681,1887]	1543 [1450,1641]
2019	1966 [1851,2083]	1734 [1641,1849]	1482 [1402,1579]
2020	1945 [1821,2070]	1695 [1592,1807]	1421 [1339,1523]
2021	1916 [1790,2045]	1648 [1546,1764]	1356 [1269,1457]
2022	1881 [1753,2014]	1589 [1488,1724]	1280 [1193,1392]
2023	1841 [1709,1984]	1538 [1434,1651]	1211 [1122,1323]
2024	1799 [1668,1945]	1468 [1372,1610]	1132 [1046,1241]
2025	1745 [1612,1893]	1407 [1294,1529]	1055 [963,1165]
2026	1692 [1549,1839]	1324 [1220,1480]	981 [892,1086]
2027	1625 [1485,1780]	1257 [1150,1413]	907 [821,1007]
2028	1557 [1416,1710]	1177 [1070,1343]	833 [749, 935]
2029	1486 [1338,1639]	1104 [982,1269]	761 [677, 862]
2030	1412 [1268,1558]	1032 [911,1209]	691 [617, 804]
2031	1343 [1202,1494]	964 [852,1148]	634 [561,739]
2032	1264 [1128,1422]	890 [765,1092]	571 [501,682]
2033	1185 [1051,1337]	810 [691,1011]	518 [446,633]
2034	1106 [968,1253]	744 [622, 965]	467 [399,572]
2035	1027 [897,1169]	672 [558, 875]	417 [349,522]
2036	954 [820,1095]	607 [482, 810]	372 [305,473]
2037	868 [741,1003]	535 [409, 751]	328 [263,427]
2038	793 [670,936]	466 [342, 696]	290 [223,393]
2039	732 [609,867]	425 [295, 661]	260 [196,364]
2040	681 [563,817]	380 [248, 623]	237 [169,333]
2041	635 [520,767]	347 [209, 583]	212 [150,306]
2042	596 [482,723]	318 [184, 565]	195 [132,291]
2043	559 [451,687]	293 [160, 527]	180 [117,270]
2044	529 [427,646]	275 [143, 516]	167 [107,266]
2045	499 [398,615]	257 [125, 514]	159 [99,247]
2046	470 [378,586]	239 [112, 484]	144 [89,237]
2047	447 [362,549]	219 [94, 467]	135 [79,231]
2048	422 [343,522]	207 [76, 454]	128 [69,214]
2049	400 [323,492]	194 [64, 446]	122 [65,208]
2050	378 [307,464]	187 [61, 426]	114 [57,195]

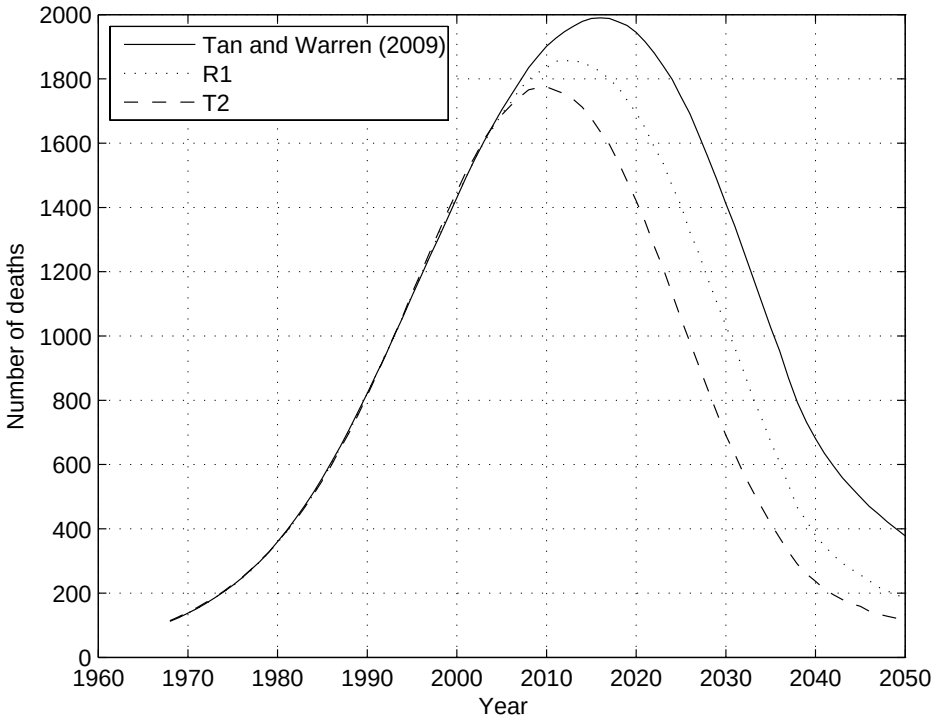

Figure 24 Male projections under three models

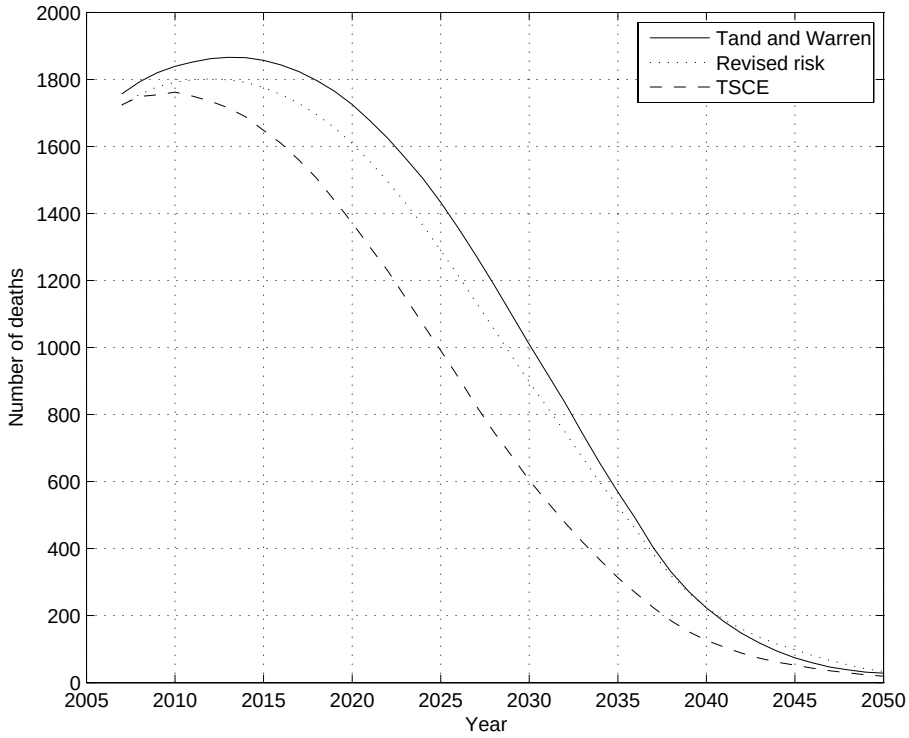

Figure 25 Male projections under the assumption of no stock released post-1980

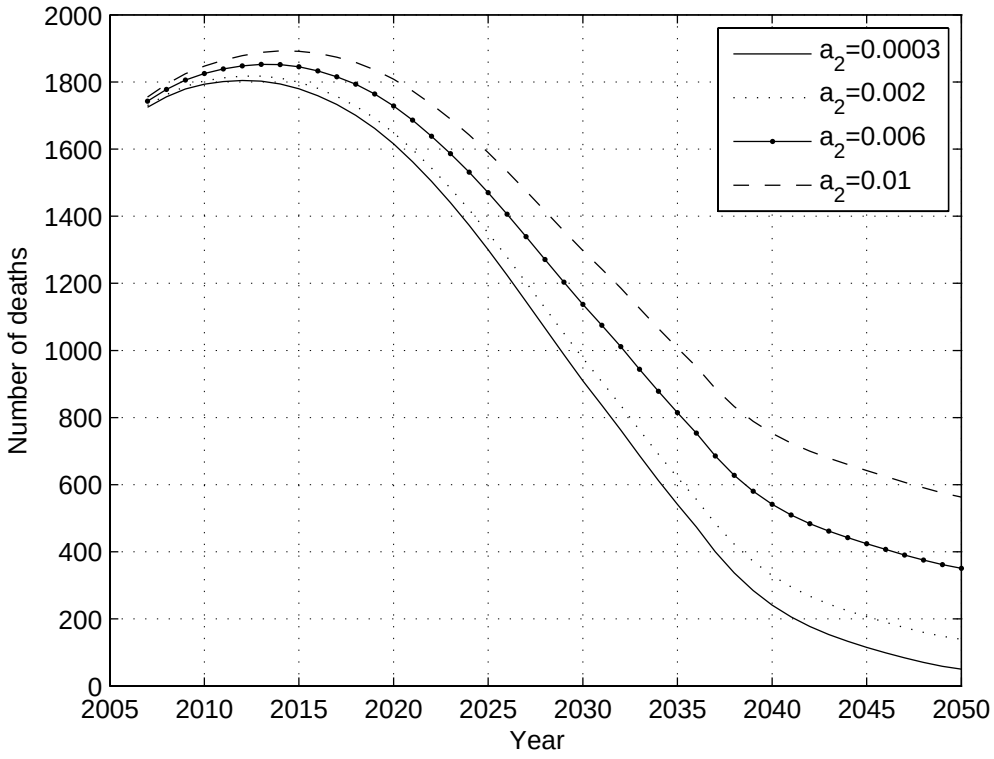

Figure 26 Male projections for different values of a_2 under the revised risk model

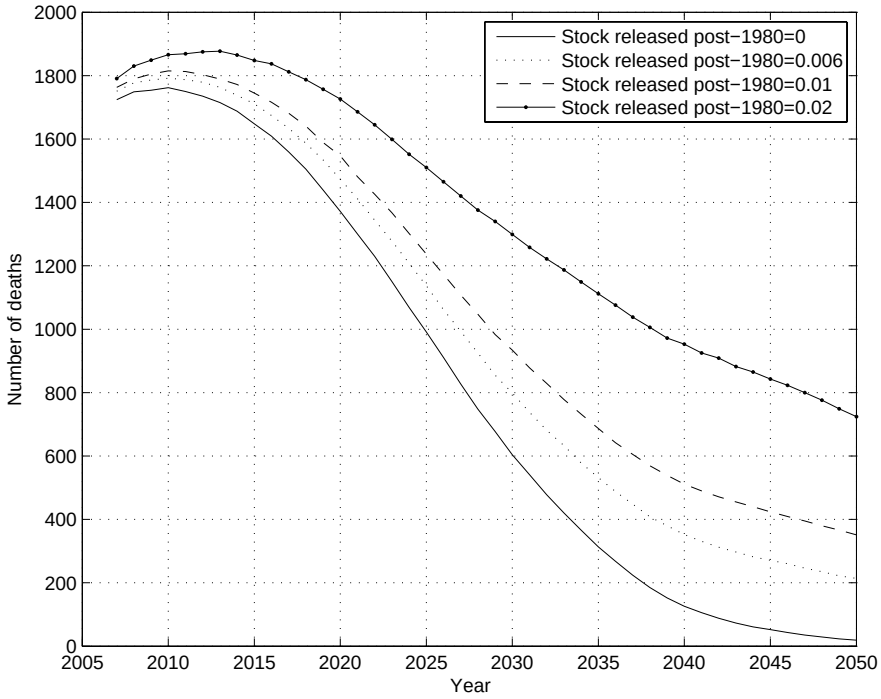

Figure 27 Male projections for different values of stock released post-1980 under TSCE model

6 DISCUSSION

Although the best-fitting TSCE model resulted in a lower deviance than the optimal revised risk model, it is difficult to establish the better of the two modelling frameworks as each has its merits. The form of the risk function in the revised risk model, where the risk increases with a power of time since exposure, may arise if a cancer cell was the result of a series of cellular changes, each change occurring with a low probability (Armitage & Doll, 1957); a lag period was incorporated by Pike (1966) to allow for the time for the tumour to develop. This form of the risk function has been used by several authors for modelling mesothelioma mortality (Berry, 1999; Hodgson *et al.*, 2005) and has been found to fit mesothelioma data well.

The biologically-based TSCE model allows for the investigation of the effects of asbestos exposure on different stages of carcinogenesis. We found that the model with exposure affecting both promotion and initiation rates provided a better fit than one where exposure affected the initiation rate only. The TSCE model has been applied to other cancers; Meza *et al.* (2008) analysed lung cancer incidence using the TSCE model and found dose-dependencies in the promotion and malignant conversion rates in the development of lung cancer.

Moolgavkar *et al.* (2009) analysed pleural and peritoneal mesothelioma incidence in the Surveillance, Epidemiology, and End Results database (United States National Cancer Institute) over the period 1973 to 2005 using both the Armitage-Doll model and the TSCE model and although both described the data well, the latter provided a consistently better fit to the data.

Regardless of our choice of model, we are still presented with some challenges, one of which is the uncertainty surrounding more recent exposure levels. Although this uncertainty does not have a significant impact on the fit of historical data, it does significantly impact longer-term projections. In particular, the impact of the stock released parameter on the TSCE model projections is potentially large. We thus stress that a range spanning the lowest and highest confidence bands of all the optimal models should be considered the most representative of the true uncertainty in the projections.

7 CONCLUSIONS

Females

- There was a sharp increase in the implied exposure amongst females around the year 1948 with a rapid decline following; the implied exposure subsequently increased to a global peak around 1965, however there was greater uncertainty in the exposure levels after 1980.
- The background rate was estimated at 1.31 cases per million, corresponding to around 30 cases in 2007 amongst females.
- Although there was considerable uncertainty regarding current and recent exposure levels, a consistent finding was that the peak year of mortality is predicted to occur over a decade later in females than in males, with a lower number of peak deaths for females than males.

Males: Revised risk model

- The proportion of high exposed males who were exposed from ages 15, 25 and 35 were estimated at 15%, 49% and 36% respectively.
- The background rate was estimated at 0.94 cases per million, corresponding to around 19 cases in 2007 amongst males.
- The lowest deviance corresponded to a proportion of high exposed males of 0.05.
- Mesothelioma mortality amongst males is predicted to peak at around 1860 deaths in 2012, however the projections are sensitive to exposure assumptions.
- The number of cases is predicted to decline to under 200 cases in 2050.

Males: Two-stage clonal expansion model

- The best-fitting TSCE model was found to be one where both the initiation and promotion rates were dose-dependent.
- The proportion of high exposed males who were exposed from ages 15, 25 and 35 were estimated at 20%, 45% and 35% respectively; these estimates are similar to those obtained under the best-fitting revised risk model.
- The number of background cases was estimated to be around 9 cases in 2007 amongst males; this estimate is lower than that obtained under the best-fitting revised risk model.
- The lowest deviance corresponded to a proportion of high exposed males of 0.10.
- Mesothelioma mortality amongst males is predicted to peak at around 1780 deaths in 2010, however the projections are sensitive to exposure assumptions.
- The number of cases is predicted to decline to under 200 cases in 2050.

APPENDIX

A FEMALES

Prior distributions

The same prior distributions that were used in Tan and Warren (2009) for males for the parameters k, W, D, $Peakyear$ and $Rate$ have been used in the analysis of female data in Section 2. In addition, the prior distribution for k_b was chosen to be $U(0, 10)$ as it has been estimated at between 2 to 3 in exploratory analysis; the prior distribution for $Prop$ was chosen to be $U(0, 1)$ as it was assumed that the proportion of the peak exposure in the year 2000 was unlikely to exceed 1.

Proposal distributions

Each proposal distribution was chosen such that it was easy to sample from $q(\cdot|\theta_t)$, each step $\theta^* - \theta_{t-1}$ moves a reasonable distance in the parameter space, and the steps generated are not rejected too frequently. Apart from the proposal distributions for $Peakyear$, each distribution was chosen to be normal with a standard deviation such that the acceptance probability was approximately 0.2 to 0.5. The proposal distributions do not have an impact on the posterior parameter estimates, only on the convergence, mixing and autocorrelation of the chains generated by the Metropolis-Hastings algorithm. The prior and proposal distributions for the model parameters are shown in Table A1.

Table A1 Metropolis-Hastings Algorithm: Prior and proposal distributions for female model parameters

Parameter	Prior	Proposal
k	$U(0, 10)$	$N(k_{t-1}, 0.3^2)$
k_b	$U(0, 10)$	$N(k_{t-1}, 0.3^2)$
W_1	$U(0, 10)$	$N(w_{1,t-1}, 0.004^2)$
W_2	$U(0, 10)$	$N(w_{2,t-1}, 0.005^2)$
W_3	$U(0, 10)$	$N(w_{3,t-1}, 0.04^2)$
W_4	$U(0, 10)$	$N(w_{4,t-1}, 0.12^2)$
W_5	$U(0, 10)$	$N(w_{5,t-1}, 0.12^2)$
W_6	$U(0, 10)$	$N(w_{6,t-1}, 0.1^2)$
W_7	$U(0, 10)$	$N(w_{7,t-1}, 0.5^2)$
W_8	$U(0, 10)$	$N(w_{8,t-1}, 0.8^2)$
D_1, D_2, D_3	$U(-100, 200)$	$N(d_{1,t-1}, 2^2)$
D_4	$U(-100, 200)$	$N(d_{2,t-1}, 10^2)$
D_5	$U(-100, 200)$	$N(d_{3,t-1}, 6^2)$
D_6	$U(-100, 200)$	$N(d_{4,t-1}, 1.1^2)$
D_7	$U(-100, 200)$	$N(d_{5,t-1}, 1.5^2)$
D_8	$U(-100, 200)$	$N(d_{6,t-1}, 3.5^2)$
D_9	$U(-100, 200)$	$N(d_{7,t-1}, 8^2)$
$Peakyear$	$U(1950, 2000)$	$\mathrm{P}(Peakyear_t = Peakyear_{t-1} + 1) = 0.5$
		$\mathrm{P}(Peakyear_t = Peakyear_{t-1} - 1) = 0.5$
$Rate$	$U(0, 3)$	$N(Rate_{t-1}, 0.4^2)$
$Prop$	$U(0, 1)$	$N(Prop_{t-1}, 0.2^2)$

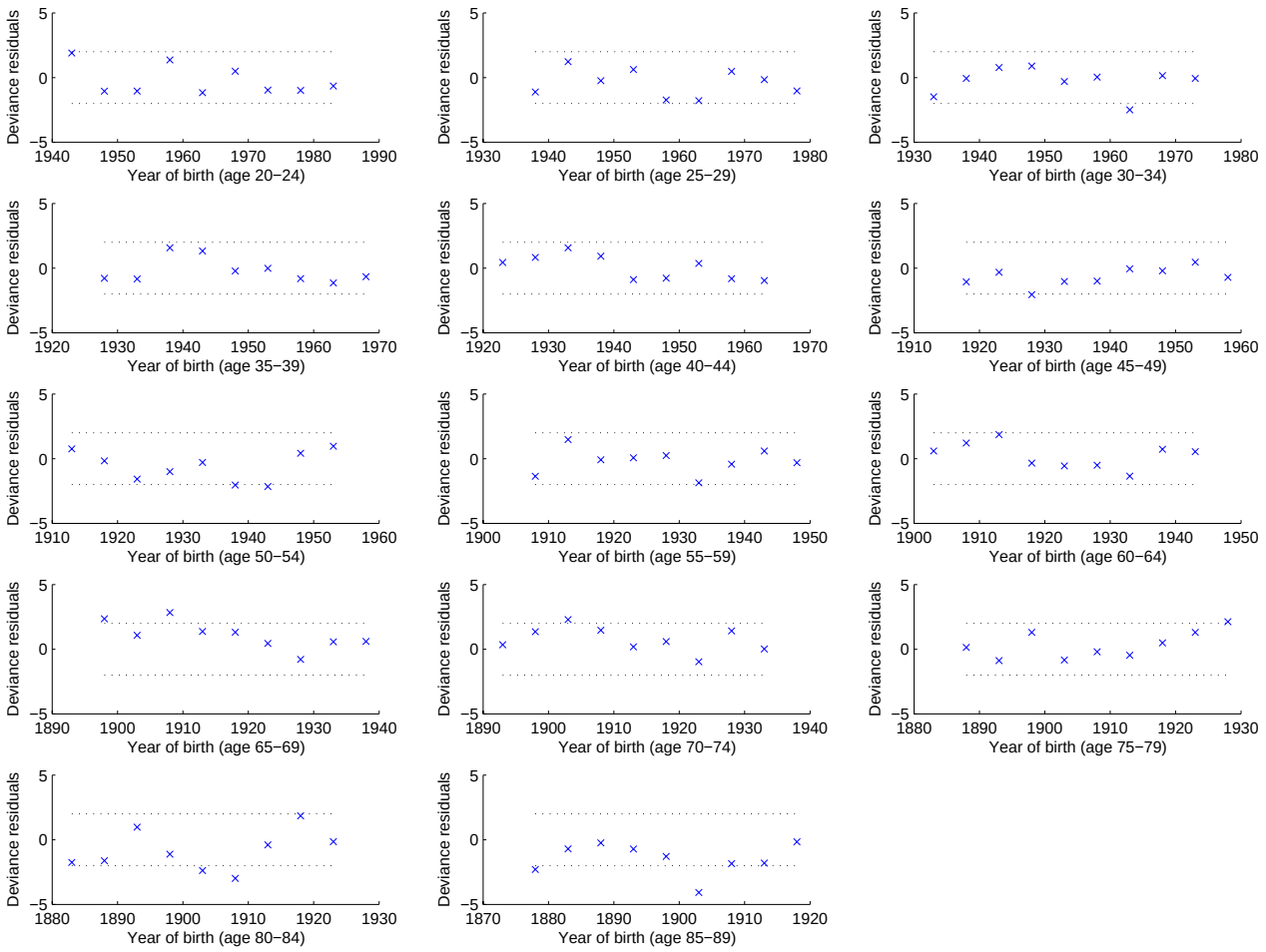

Figure A1 Deviance residuals by age group and birth cohort for F1

B REVISED RISK MODEL

Prior distributions

Non-informative prior distributions for each parameter were chosen by considering plausible ranges in addition to the estimates obtained using *fminsearch*.

The number of background cases in males (per million population) was estimated at around 1 in Tan and Warren (2009) and a uniform $U(0, 10)$ prior was chosen for $Rate$. Exploratory analysis using *fminsearch* suggested a range for β_2 between around 50 and 200, depending on the value of α chosen, hence uniform priors that incorporate that range have been used. The power of time associated with asbestos exposure has previously been estimated at between 2 and 3 (Hodgson *et al.*, 2005; Tan and Warren, 2009). It was unlikely that the risk decreased with time since exposure, hence the prior for k was chosen to be $U(0, 10)$. Similarly, the prior for k_b was chosen to be $U(0, 10)$. $High_{25}$ and $High_{35}$ were estimated at around 0.5 and 0.4 respectively in exploratory analysis. As the sum of $High_{15}$, $High_{25}$ and $High_{35}$ must equal 1 (and thus the sum of $High_{25}$ and $High_{35}$ can not be greater than 1), uniform $U(0, 0.6)$ and $U(0, 0.4)$ priors were chosen for $High_{25}$ and $High_{35}$ respectively, assuming that $High_{25}$ does not exceed 0.6 and $High_{35}$ does not exceed 0.4. $CrocRatio$ was estimated at around 0.1 in exploratory analysis and was considered unlikely to exceed 10. As it also must take a postive value, a uniform $U(0, 10)$ prior was chosen. $StockRel_{1980}$ can only take values between 0 and 1, so a uniform $U(0, 1)$ prior was chosen.

Proposal distributions

Each proposal distribution was chosen such that it was easy to sample from $q(\cdot|\theta_t)$, each step $\theta^* - \theta_{t-1}$ moves a reasonable distance in the parameter space, and the steps generated are not rejected too frequently. Each distribution was chosen to be normal with a standard deviation such that the acceptance probability was approximately 0.2 to 0.5. The proposal distributions do not have an impact on the posterior parameter estimates, only on the convergence, mixing and autocorrelation of the chains generated by the Metropolis-Hastings algorithm. Table A2 shows the prior and proposal distributions that have been used in fitting the model using the Metropolis-Hastings algorithm.

Table A2 Metropolis-Hastings Algorithm: Prior and proposal distributions for R1 model parameters

Parameter	Prior	Proposal
$Rate$	$U(0, 10)$	$N(Rate_{t-1}, 0.2^2)$
k	$U(1, 10)$	$N(k_{t-1}, 0.01^2)$
k_b	$U(0, 10)$	$N(k_{b,t-1}, 0.07^2)$
β_2	$U(50, 350)$	$N(\beta_{2,t-1}, 1.5^2)$
$High_{25}$	$U(0, 0.6)$	$N(High_{25,t-1}, 0.01^2)$
$High_{35}$	$U(0, 0.4)$	$N(High_{35,t-1}, 0.01^2)$
$CrocRatio$	$U(0, 10)$	$N(CrocRatio_{t-1}, 0.05^2)$
$StockRel_{1980}$	$U(0, 1)$	$N(StockRel_{1980,t-1}, 0.005^2)$

C TSCE MODEL

Prior distributions

Non-informative prior distributions for each parameter were chosen by considering plausible ranges in addition to the estimates obtained using *fminsearch* and, if common to both the revised risk and TSCE model, the estimates obtained under the revised risk model.

Proposal distributions

Each proposal distribution was chosen such that it was easy to sample from $q(\cdot|\theta_t)$, each step $\theta^* - \theta_{t-1}$ moves a reasonable distance in the parameter space, and the steps generated are not rejected too frequently. Each distribution was chosen to be normal with a standard deviation such that the acceptance probability was approximately 0.2 to 0.5. Table A3 shows the prior and proposal distributions that have been used in fitting the model using the Metropolis-Hastings algorithm.

Table A3 Metropolis-Hastings Algorithm: Prior and proposal distributions for T2 model parameters

Parameter	Prior	Proposal
α	$U(0, 5000)$	$N(\alpha_{t-1}, 15^2)$
γ_0	$U(0, 10)$	$N(\gamma_{0,t-1}, 0.0006^2)$
γ_1	$U(0, 10)$	$N(\gamma_{1,t-1}, 0.015^2)$
γ_2	$U(0, 10)$	$N(\gamma_{2,t-1}, 0.001^2)$
ν_0	$U(0, 10)$	$N(\nu_{0,t-1}, 0.000000003^2)$
ν_1	$U(0, 100)$	$N(\nu_{1,t-1}, 0.0003^2)$
$High_{25}$	$U(0, 0.6)$	$N(High_{25,t-1}, 0.02^2)$
$High_{35}$	$U(0, 0.4)$	$N(High_{35,t-1}, 0.01^2)$
$ChryRatio$	$U(0, 1)$	$N(CrocRatio_{t-1}, 0.001^2)$
$StockRel_{1980}$	$U(0, 0.05)$	$N(StockRel_{1980,t-1}, 0.003^2)$

Table A4 Cumulative hazard at age 89 for males in 1900 to 1980 birth cohorts with various exposure histories (figures in **bold** indicate the largest hazard for the associated exposure profile)

	Exposure history							
	Exposed from 15 for			Exposed from 25 for			Exposed from 35 for	
Birth cohort	5 years	10 years	40 years	5 years	10 years	30 years	5 years	10 years
1900	0.0029	0.0067	0.0858	0.0039	0.0115	0.0778	0.0134	0.0280
1910	0.0048	0.0144	0.2040	0.0184	0.0392	0.1884	0.0173	0.0566
1920	0.0219	0.0473	0.3446	0.0237	0.0792	0.3032	0.0657	**0.1391**
1930	**0.0282**	0.0950	**0.4185**	0.0858	**0.1839**	**0.3422**	**0.0755**	0.1354
1940	0.0993	**0.2134**	0.4158	**0.0978**	0.1777	0.2327	0.0386	0.0447
1950	0.1127	0.2056	0.2794	0.0507	0.0593	0.0767	0.0050	0.0091
1960	0.0590	0.0693	0.0985	0.0069	0.0130	0.0261	0.0040	0.0073
1970	0.0083	0.0160	0.0388	0.0055	0.0105	0.0208	0.0032	0.0058
1980	0.0067	0.0129	0.0311	0.0045	0.0084	0.0167	0.0026	0.0046

Table A5 Projections of male mesothelioma deaths under the TSCE model for various values of L and a_2

	Projection						
		L		a_2			
Year	10	12	15	0.00	0.006	0.010	0.020
2007	1743	1765	1775	1724	1751	1763	1791
2008	1766	1790	1803	1749	1777	1790	1830
2009	1772	1805	1822	1754	1786	1805	1849
2010	**1775**	**1811**	**1830**	**1762**	**1792**	**1815**	1866
2011	1764	1808	1830	1750	1787	1813	1869
2012	1754	1800	1824	1735	1779	1803	1875
2013	1736	1784	1811	1715	1763	1789	**1877**
2014	1712	1762	1791	1688	1738	1772	1865
2015	1677	1732	1763	1648	1711	1745	1848
2016	1636	1695	1728	1609	1674	1716	1837
2017	1596	1652	1686	1559	1634	1681	1812
2018	1543	1603	1638	1505	1587	1641	1787
2019	1482	1547	1582	1439	1534	1589	1757
2020	1421	1484	1520	1372	1473	1547	1726
2021	1356	1417	1451	1300	1411	1481	1686
2022	1280	1345	1378	1230	1344	1426	1645
2023	1211	1271	1301	1150	1278	1367	1599
2024	1132	1194	1222	1069	1208	1302	1552
2025	1055	1115	1139	991	1137	1237	1510
2026	981	1036	1055	911	1062	1173	1465
2027	907	956	970	827	993	1108	1420
2028	833	879	887	748	925	1047	1376
2029	761	804	806	677	855	983	1340
2030	691	733	729	604	799	934	1299
2031	634	668	658	541	738	878	1258
2032	571	606	589	478	681	829	1222
2033	518	545	521	420	633	778	1187
2034	467	488	457	366	578	732	1149
2035	417	435	398	313	532	687	1112
2036	372	387	344	268	488	642	1076
2037	328	338	289	224	448	605	1038
2038	290	297	243	185	408	569	1006
2039	260	263	205	152	380	539	972
2040	237	235	174	126	352	511	953
2041	212	212	148	106	331	490	925
2042	195	193	127	88	312	471	909
2043	180	177	110	73	297	455	882
2044	167	164	95	61	282	440	865
2045	159	152	83	52	271	424	843
2046	144	141	71	43	260	409	823
2047	135	131	61	35	246	395	800
2048	128	121	53	29	235	380	776
2049	122	113	45	23	222	366	749
2050	114	106	40	19	214	351	724

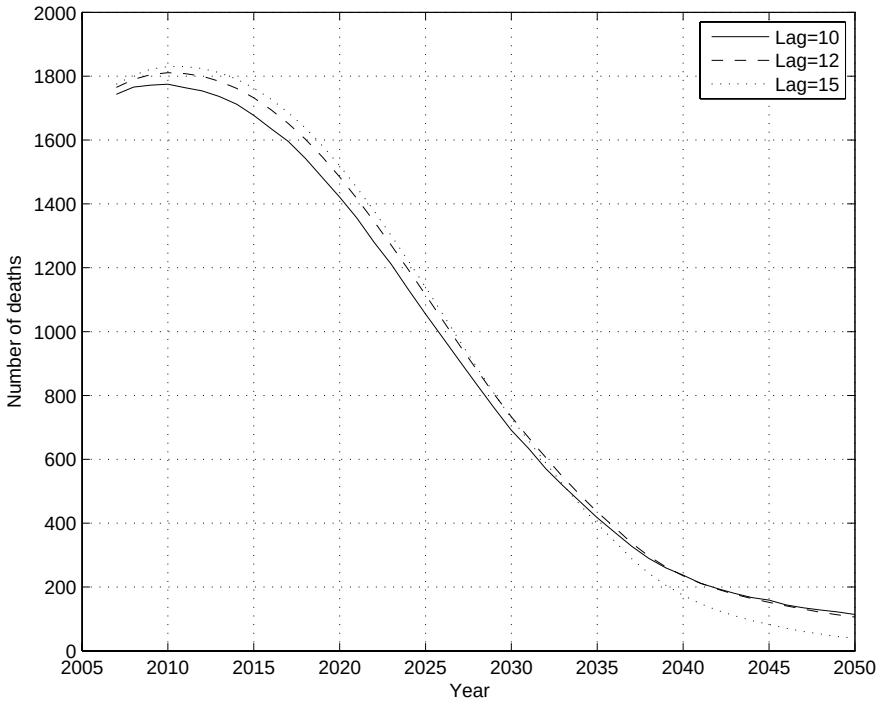

Figure A2 Comparison of the effects of lag on projections under the TSCE model

REFERENCES

Berry, G. (1999), 'Models for mesothelioma incidence following exposure to fibers in terms of timing and duration of exposure and the biopersistence of the fibers.', *Inhalation Toxicology* **11**, 111-130

Hastings, W. (1970), 'Monte Carlo Sampling Methods Using Markov Chains and Their Applications.', *Biometrika* **57**, 97-109

Hazelton, W., Clements, M. and Moolgavkar, S. (2005), 'Multistage Carcinogenesis and Lung Cancer Mortality in Three Cohorts.', **14(5)**, 1171-1181

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.*

Heidenreich, W., Luebeck, E. and Moolgavkar, S. (1997), 'Some Properties of the Hazard Function of the Two-Mutation Clonal Expansion Model.', *Risk Analysis* **17**, 391-399

Hodgson, J., McElvenny, D., Darnton, A., Price, M. and Peto, J. (2005), 'The expected burden of mesothelioma mortality in Great Britain from 2002 to 2050.', *British Journal of Cancer* **4**, 587-593

The MathWorks, Inc. (2009), 'MATLAB 7.9.0 (R2009b)'

Moolgavkar, S. and Knudson. (1981), 'Mutation and Cancer: A Model for Human Carcinogenesis.', *Journal of the National Cancer Institute* **66**, 1037-1052

Moolgavkar, S., Dewanji, A. and Venzon, D. (1988), 'A Stochastic Two-Stage Model for Cancer Risk Assessment. I. The Hazard Function and the Probability of Tumor', *Risk Analysis* **8**, 383-392

Moolgavkar, S., Turim, J., Brown, R. and Luebeck, E. (2001), 'Long Man-Made Fibres and Lung Cancer Risk.', *Regulatory Toxicology and Pharmacology* **33**, 138-146

Moolgavkar, S., Meza, R. and Turim, J. (2009), 'Pleural and peritoneal mesotheliomas in SEER: age effects and temporal trends, 1973-2005.', Cancer Causes Control **20**, 935-944

Peto, J., Matthews, F., Hodgson, J. and Jones, J. (1995), 'Continuing increase in mesothelioma mortality in Britain.', *Lancet* **345**, 535-539

Rake, C., Gilham, C., Hatch, J., Darnton, A., Hodgson, J. and Peto, J. (2009), 'Occupational, domestic and environmental mesothelioma risks in the British population: a case-control study.', *British Journal of Cancer* **100**, 1175-1183

Richardson, D. (2009), 'Lung cancer in chrysotile workers: analyses based on the two-stage clonal expansion model.', *Cancer Causes Control* **20**, 917-923

Tan, E. and Warren, N. (2009), 'Projection of mesothelioma mortality in Great Britain.', *HSE Research Report RR728*

Published by the Health and Safety Executive 07/11

HSE

Health and Safety Executive

Mesothelioma mortality in Great Britain

The revised risk and two-stage clonal expansion models

Asbestos is a known carcinogen that is the cause of the majority of mesothelioma cases worldwide. Various models have been used to describe the increase and likely future pattern of mesothelioma rates seen in many western countries – a legacy of past heavy industrial asbestos use. Following on from previous work (Tan and Warren, 2009), we analysed female mesothelioma mortality using the same risk model that was assumed for males. We also analysed mesothelioma mortality in males in Great Britain using two alternative risk models; the first is based on asbestos import data where the population is categorised into low and high exposure groups, with the calculation of risk based on the cumulative lung burden of the individual; the second is a two-stage clonal expansion model (TSCE), a biologically-based carcinogenesis model that assumes that the development of a malignant cell is the result of two critical and irreversible events, with asbestos lung burden as the measure of dose that enters the dose-response component of the TSCE model. We use Markov Chain Monte Carlo within a Bayesian framework to fit the models presented in this report.

Though considerably uncertain, peak mortality in females is predicted to occur over a decade later than in males, but with a substantially lower annual number of deaths. The updated models provide a reasonable basis for making relatively short-term projections of mesothelioma mortality in Britain. However, longer-term predictions comprise additional uncertainty not captured within the prediction intervals for the annual mortality rates. Taking this into account, 2100 deaths in 2016 represents our current best estimate of the upper limit for the male projections.

This report and the work it describes were funded by the Health and Safety Executive (HSE). Its contents, including any opinions and/or conclusions expressed, are those of the authors alone and do not necessarily reflect HSE policy.

RR876

www.hse.gov.uk