J.I.A. **121**, III, 561–571

A COMPARISON BETWEEN THE MORTALITY OF NON-SMOKING AND SMOKING ASSURED LIVES IN THE UNITED KINGDOM

BY A. E. RENSHAW, B.Sc., Ph.D.
(of The City University, London)

ABSTRACT

A method of graduation by mathematical formula is extended to cover the simultaneous graduation of more than one comparable experience and is applied to the recent U.K. mortality experience of smoking and non-smoking assured lives.

KEYWORDS

Mortality; Assured Lives; Modelling

1. INTRODUCTION

IN a recent study conducted by the CMI Bureau, reported by Evans (1993) and in CMIR 13 (1993), the increased risk of premature death for smokers compared with non-smokers is demonstrated using well-tried actuarial methods. These comprise a comparison of the actual deaths reported in the various study groups with the expected deaths predicted on the basis of comparable standard actuarial life tables, together with the construction and comparison of the curve of deaths associated with the various study groups. The main purpose of this short paper is to present an alternative methodology for comparing the mortality between such well-defined groups. The approach, which does not involve recourse to standard tables, is based on an extension of the graduation methods outlined by Forfar, McCutcheon & Wilkie (1988) and has much wider application, see, for example, Renshaw, Haberman & Hatzopoulos (1994). In addition, the results stemming from the analysis are of considerable interest in their own right and are wholly supportive of the conclusions to be drawn from the CMI findings.

2. THE DATA

The data are denoted by:

$$(a_u, e_u)$$

comprising the actual numbers of deaths a_u, accruing from central exposures e_u, for a set of units $\{u\}$. The data currently available relate to the two-year calendar period 1988–89, and are based on policy rather than head counts.

Here, by way of illustration, we focus on assured lives (whole-life and endowment) for all policy durations combined. The data are categorised as follows:

gender (g) with	$i = 1$ female, $i = 2$ male
habit (h) with 3 levels	$j = 1$ non-smoker $j = 2$ smoker $j = 3$ undifferentiated
status (s) with two levels	$k = 1$ medical, $k = 2$ non-medical
age (x) arranged in 18 grouped levels	11–15, 16–20, . . . , 96–100

giving rise to $2\times3\times2\times18 = 216$ cross-classified cells or units, u. Write:

$$u \equiv (i,j,k,x).$$

The actual deaths a_u, are presented in Table 2.1. Ten of the cells, which are empty, are given zero weight in order to eliminate them from the ensuing analysis. A further 33 cells have zero deaths, 10 of which occur in the youngest age category 11–15 years, and a further 17 of which occur in the medical status category. All of these cells are retained in the ensuing analysis in which age x is modelled as an error-free variable and the remaining three covariates, gender, habit and status, modelled as categorical factors.

Table 2.1. Actual deaths, assured lives, 1988–89

Gender	Females						Males					
Status	Medical			Non-Medical			Medical			Non-Medical		
Habit	1	2	3	1	2	3	1	2	3	1	2	3
11–15	0	--	0	0	0	0	0	—	0	0	0	0
16–20	1	0	1	2	1	3	1	0	1	5	1	8
21–25	0	0	1	10	2	17	1	0	1	25	9	43
26–30	1	0	1	5	0	7	0	2	3	18	6	44
31–35	2	0	4	2	5	21	0	0	7	17	14	110
36–40	3	0	8	11	4	51	3	1	26	32	22	224
41–45	3	0	8	30	12	88	4	5	82	37	52	454
46–50	2	0	10	27	30	140	0	7	97	51	72	658
Age 51–55	4	1	18	33	19	171	8	9	164	103	113	1045
56–60	7	2	31	26	28	192	10	16	287	79	121	1339
61–65	5	1	24	22	19	110	6	20	366	56	87	1262
66–70	4	1	19	16	16	74	22	12	183	33	29	366
71–75	4	2	18	7	3	42	12	9	204	17	23	266
76–80	7	2	43	2	8	27	14	20	298	13	13	166
81–85	3	1	35	1	1	12	5	15	287	2	8	81
86–90	1	1	7	0	0	4	0	13	138	—	1	25
91–95	1	0	4	—	0	0	—	4	45	—	2	8
96–100	—	1	2	—	1	1	—	0	10	—	0	2

(habit: 1- non-smoker, 2- smoker, 3- undifferentiated)

3. The Model

We take as our starting point the Gompertz-Makeham graduation formula:

$$\mu_x = \mathrm{GM}_x(0, s) = \exp\left(\sum_{j=1}^{s-1} \beta_j x^j\right)$$

expressing the force of mortality μ_x, at age x, as an exponentiated polynomial of degree $(s-1)$ in x. The unknown parameters β_j lend flexibility to the formula. They are estimated in any graduation of data under the assumption that the actual numbers of deaths A_x, are distributed as independent Poisson responses with mean:

$$m_x = \mathrm{E}(A_x) = e_x \mu_x$$

where e_x denote the matching central exposures. Full details of this are available in Forfar *et al.* (1988). Scrutiny of Table 2.1 immediately reveals that it is not practical to apply such a technique to each separate column of Table 2.1 due to the low number of recorded deaths in many of the cells. We seek instead to broaden the technique so that it might still be applied to cross-classified mortality data of this type.

Thus motivated we target the force of mortality μ_u, a function of the units u described in Section 2, using the two-stage formula:

$$\log(\mu_u) = \sum_{v=o}^{p} z_{uv}\beta_v. \tag{3.1}$$

The (possible) explanatory variables—gender, habit, status and age—enter the right side through a variety of specified covariate structures (z_{uv}), while the unknown parameters β_v lend flexibility to the formula. For technical reasons we set $z_{uo} = 1$, and refer to β_0 as the general mean. The log function on the left side, which is monotonic and differentiable, maps the positive reals $\{\mu_u; \mu_u > 0\}$ onto the whole of the real line, and therefore has desirable technical properties. It is also consistent with the Gompertz-Makeham formula discussed previously. The formula also leads to sufficient statistics for the β_vs in the model-fitting process described next.

To fit such models to the data, it is necessary to estimate the β_vs and to assess the improvement in the goodness-of-fit of such formulae as more complex (nested) structures are contemplated for the right side of equation (3.1). To achieve this we require a further assumption. In keeping with modern actuarial graduation practice (see Forfar *et al.* 1988; Renshaw 1991, 1992), we model the actual number of deaths a_u, as independent over-dispersed Poisson response variables A_u, with mean and variance:

$$m_u = \mathrm{E}(A_u) = e_u \mu_u, \quad \mathrm{Var}(A_u) = \phi m_u. \tag{3.2}$$

The scale or dispersion parameter $\phi > 1$, is included, since the data are based on

policy rather than head counts. While there is evidence to suggest that ϕ varies with age, x, such variation has a very minor effect on the modelling process as a whole, and its effect is neglected for the purposes of this analysis. The corresponding expression for the log likelihood is then:

$$1 = \sum_{u=1}^{n} \{a_u \log(m_u) - m_u\} + \text{constant} \tag{3.3}$$

which is optimised to provide the maximum likelihood estimators for the β_vs. These enter expression (3.3) through the rearrangement:

$$m_u = e_u \exp\left\{\sum_{v=0}^{p} z_{uv}\beta_v\right\}$$

of the predictor-link relationship:

$$\log(m_u) = \eta_u = \log(e_u) + \sum_{v=0}^{p} z_{uv}\beta_v$$

comprising the log link function and linear predictor η_u. This formula is consistent with equation (3.1) on taking logs of the identity $m_u = e_u\mu_u$, taken from expressions (3.2). The term $\log(e_u)$, in the linear predictor does not involve a parameter and, as such, offsets the value of the general mean β_0. Denote the resulting maximum likelihood estimates for the current model c, by $\hat{\beta}_v$. These are computed by resorting to the interactive computer software package GLIM.

In addition to model fitting, we also require the means of model or formula selection. Define the deviance of the current model c, to be:

$$\mathrm{D}(c,f) = \sum_{u=1}^{n} d_u = \sum_{u=1}^{n} 2\left\{a_u \log\left(\frac{a_u}{\hat{m}_u}\right) - (a_u - \hat{m}_u)\right\} \tag{3.4}$$

where:

$$\hat{m}_u = e_u \exp\left\{\sum_{v=0}^{p} z_{uv}\hat{\beta}_v\right\}$$

denote the corresponding fitted values. The deviance is twice the difference between the log likelihood expression (3.3) evaluated when $m_u = a_u$ and when $m_u = \hat{m}_u$. While the latter are the fitted values under the current model c, we likewise interpret the former as the fitted values under the saturated or full model f. The model f, which is characterised by the fitted values $m_u = a_u$, implies a perfect fit for the data. Differences:

$$\{\mathrm{D}(c_1,f) - \mathrm{D}(c_2,f)\}/\phi$$

in the scaled deviances as we change from one (nested) predictor structure c_1, to a more complex predictor structure c_2, are used to assess the significance of the improvement in the model fit. These differences may be referred, as an

approximation, to the chi-square distribution subject to the appropriate degrees of freedom. At the same time we look at the significance of the individual parameter estimates $\hat{\beta}_v$, in order to safeguard against possible over-parameterisation. We also monitor the fit of the current model c, through a graphical analysis of the associated deviance residuals defined by:

$$r_u = \text{sign}(a_u - \hat{m}_u)\sqrt{d_u}$$

where d_u is the value of the uth component of the model deviance $D(c,f)$ defined by equation (3.4).

The scale parameter ϕ, is concerned with the second moment properties of the model and, as such, has an input into the construction of the standard errors of the parameter estimates, but not the estimates themselves, since it is assumed to be constant. The net effect is to increase the standard errors slightly to allow for the increased uncertainty induced by the presence of duplicate policies. This parameter is estimated by:

$$\hat{\phi} = \frac{D(c,f)}{\nu}$$

where the degrees of freedom ν, denote the number of observations minus the number of independent parameters in the predictor.

4. The Analysis

Applying the modelling technique outlined in Section 3 to the data described in Section 2, we begin with a search for a suitable model formula of the general type (3.1) which captures the underlying pattern of mortality in the data. The resultant deviance profile for just one of the many model building sequences possible is presented in Table 4.1. The table is to be interpreted by reading downwards. In this sequence, age effects are introduced first in the form of a polynomial predictor up to degree three and the coefficients then adjusted, in a variety of ways, to allow for the effects of the other three factors—gender, habit and status. It is possible to identify the specific nature of the various model structures by referring to the first column of Table 4.1. Thus reading down the table, the first nine graduation formulae fitted to the data in this sequence are:

$$\begin{aligned}
1 \quad &: \mu_{ijkx} = \exp\{\alpha\} && \text{(constant mortality throughout)}\\
+x \quad &: \mu_{ijkx} = \exp\{\alpha + \tau x\} && \text{(Gompertz Law } \forall\, i,j,k \text{ combined)}\\
+x^2 \quad &: \mu_{ijkx} = \exp\{\alpha + \tau x + \beta x^2\}\\
+x^3 \quad &: \mu_{ijkx} = \exp\{\alpha + \tau x + \beta x^2 + \gamma x^3\}\\
+g.x \quad &: \mu_{ijkx} = \exp\{\alpha + (\tau + \theta_i)x + \beta x^2 + \gamma x^3\}\\
+h.x \quad &: \mu_{ijkx} = \exp\{\alpha + (\tau + \theta_i + \psi_j)x + \beta x^2 + \gamma x^3\}\\
+g.h.x \quad &: \mu_{ijkx} = \exp\{\alpha + (\tau + \theta_i + \psi_j + (\theta\psi)_{ij})x + \beta x^2 + \gamma x^3\}\\
+s.x \quad &: \mu_{ijkx} = \exp\{\alpha + (\tau + \theta_i + \psi_j + \kappa_k + (\theta\psi)_{ij})x + \beta x^2 + \gamma x^3\}\\
+g \quad &: \mu_{ijkx} = \exp\{\alpha + \delta_i + (\tau + \theta_i + \psi_j + \kappa_k + (\theta\psi)_{ij})x + \beta x^2 + \gamma x^3\}
\end{aligned}$$

Table 4.1. Deviance profile, assured lives, 1988–89

Model	Deviation	First differences	
		Deviance	Degrees of freedom
1	18523		
		17499	1
$+x$	1024·0		
		35·4	1
$+x^2$	988·6		
		116·0	1
$+x^3$	872·6		
		353·6	1
$+g.x$	518·9		
		186·4	2
$+h.x$	332·5		
		2·4	2
$+g.h.x$	330·1		
		5·2	1
$+s.x$	324·9		
		0·6	1
$+g$	324·3		
		13·7	2
$+h$	310·6		
		3·9	2
$+g.h$	306·7		
		0·8	1
$+s$	305·9		
		5·6	1
$+g.x^2$	300·3		
		0·4	2
$+h.x^2$	299·9		
		0·5	2
$+g.h.x^2$	299·5		
		25·9	1
$+s.x^2$	273·6		

x—age, g—gender, h—smoking habit, s—medical status

and so on. Note how the predictor structures are built up sequentially, with the structure at any one stage in the process contained within the structure at any subsequent stage of the process, while both the (log) link function and the over-dispersed Poisson modelling distribution remain the same throughout. Such hierarchical predictor structures are said to be nested. An examination of the reduction in the (unscaled) deviance as more complex predictor structures are introduced, the third column of Table 4.1, together with an examination of the resulting parameters estimates and their standard errors, leads to the adoption of the model structure:

$$\log(\mu_{ijkx}) = \alpha + (\tau + \theta_i + \psi_j)x + \beta x^2 + \gamma x^3, \quad i = 1,2;\ j = 1,2,3. \qquad (4.1)$$

In particular, the entries in Column 3 of Table 4.1 may be referred, as an

Table 4.2. Parameter estimates, assured lives, 1988–89

$\hat{\alpha} = -7{\cdot}324\ (2{\cdot}686\times10^{-1})$

$\hat{\tau}$	$= -5{\cdot}031\times10^{-1}\ (8{\cdot}794\times10^{-2})$	$\hat{\theta}_2$	$= 4{\cdot}291\times10^{-2}\ (3{\cdot}477\times10^{-3})$
$\hat{\psi}_2$	$= 6{\cdot}456\times10^{-2}\ (6{\cdot}418\times10^{-3})$	$\hat{\psi}_3$	$= 4{\cdot}157\times10^{-2}\ (4{\cdot}991\times10^{-3})$
$\hat{\beta}$	$= 9{\cdot}222\times10^{-2}\ (9{\cdot}128\times10^{-3})$	$\hat{\gamma}$	$= -2{\cdot}819\times10^{-3}\ (3{\cdot}019\times10^{-4})$

Deviance is 332·52 on 199 degrees of freedom from 206 observations, with scale parameter $\hat{\phi} = 1{\cdot}671$

approximation, to the chi-square distribution subject to the degrees of freedom documented in the matching entry of Column 4 of Table 4.1. Since scale parameters of the order 2 are common for data sets of this type, chi-square critical values taken from standard tables are multiplied by a factor of 2 as a rough working guide. For technical reasons this predictor is subject to the two constraints $\theta_1 = \psi_1 = 0$, and therefore involves a total of seven parameters. The maximum likelihood estimates, their standard errors, together with the estimate for the over-dispersion parameter ϕ, are presented in Table 4.2. The ratio of any of the parameter estimates to its standard error, the so-called t-statistic, may be loosely interpreted as a standardised normal variate so that a value outside the range -2 to $+2$ may be interpreted as being statistically significant. This is the case for all seven parameter estimates in this model. For this model, both gender and habit impact additively on the coefficient of x, while the remaining coefficients of the cubic predictor are constant. Attempts to increase the complexity of the structure of the coefficients result in only small reductions in the model deviance, as is apparent from Table 4.1, and the simultaneous incorporation of non-significant parameter estimates, which is difficult to justify. In particular, the effect of status (non-medical or medical) is not statistically significant. Thus effectively the complete data set is modelled using a Gompertz-Makeham formula:

$$\mu_{ijkx} = \mathrm{GM}_x(0, 4)$$

defined by equation (4.1), in which one of the polynomial coefficients is additive in gender and status effects. Finally, as part of the residual analysis for this model, the histogram of deviance residuals together with the deviance residual plotted against age are reproduced in Figure 4.1. Both of these plots, together with other plots not reproduced here, are consistent with a satisfactory fit. This is regarded as a vital diagnostic checking procedure to ensure that the adopted model formula capsulates the underlying mortality patterns present in the data as a whole.

The force of mortality predicted by this model at five-yearly ages and cross-classified according to gender and habit (non-smoker, smoker, undifferentiated) is tabulated in Table 4.3. These rates form a useful complement to the CMI published analysis. The main features are as follows:

—In each category, the force of mortality initially decreases with age before starting on an upward trend in the mid to late twenties, consistent with known patterns of human mortality.

—Comparison within a given gender by habit at specific ages reveals that the force of mortality is consistently higher for smokers over non-smokers at all corresponding ages, with the force of mortality for the undifferentiated category intermediate, again for all corresponding ages. This is true of both females and males.

—While it is well established that mortality in females is lower than in males for

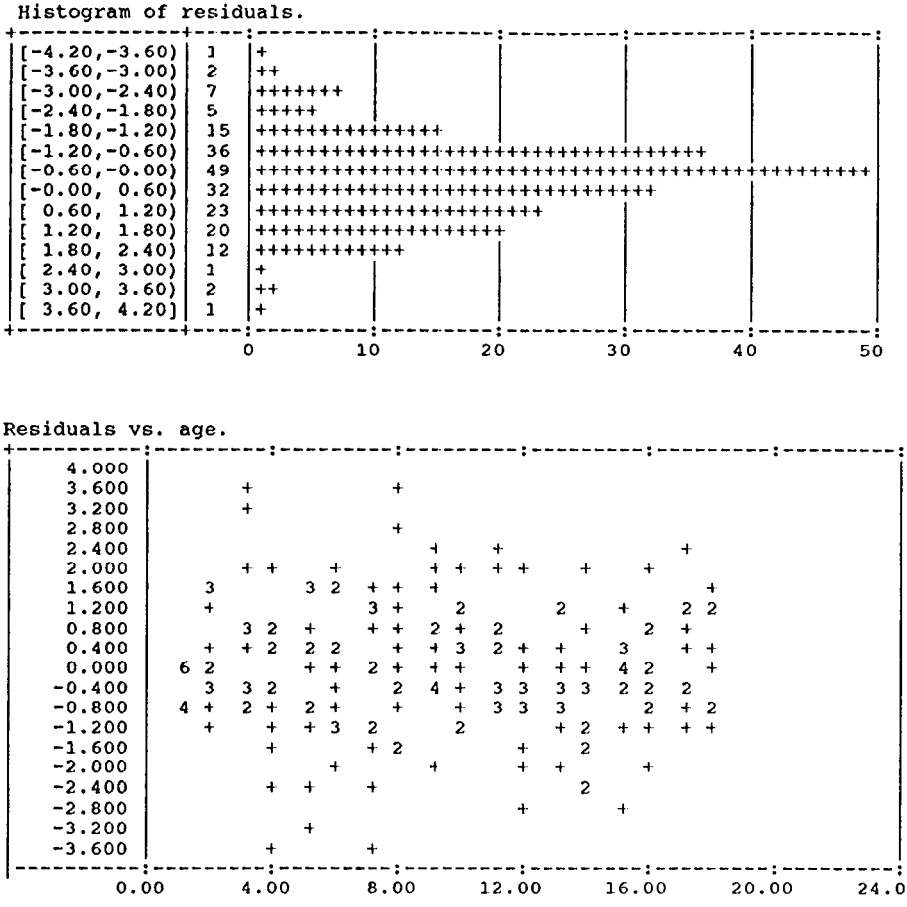


Figure 4.1. Deviance residual plots, assured lives, 1988–89

Table 4.3. Predicted force of mortality by age, gender and smoking habit, assured lives, 1988–89

Gender		Females			Males	
Habit	Non-Smoker	Smoker	Undiff'ed	Non-Smoker	Smoker	Undiff'ed
11–15	0·0004359	0·0004650	0·0004544	0·0004550	0·0004854	0·0004743
16–20	0·0003408	0·0003877	0·0003703	0·0003713	0·0004225	0·0004035
21–25	0·0003097	0·0003759	0·0003508	0·0003523	0·0004275	0·0003990
26–30	0·0003217	0·0004165	0·0003799	0·0003820	0·0004945	0·0004511
31–35	0·0003756	0·0005187	0·0004624	0·0004655	0·0006428	0·0005730
36–40	0·0004846	0·0007138	0·0006218	0·0006268	0·0009234	0·0008044
41–45	0·0006792	0·0010672	0·0009085	0·0009171	0·0014411	0·0012268
46–50	0·0010169	0·0017045	0·0014181	0·0014334	0·0024025	0·0019988
Age 51–55	0·0015992	0·0028593	0·0023248	0·0023530	0·0042068	0·0034204
56–60	0·0025973	0·0049533	0·0039358	0·0039890	0·0076074	0·0060447
61–65	0·0042831	0·0087132	0·0067659	0·0068665	0·0139684	0·0108467
66–70	0·0070516	0·0153016	0·0116120	0·0118003	0·0256061	0·0194317
71–75	0·0113960	0·0263779	0·0195623	0·0199063	0·0460763	0·0341711
76–80	0·0177749	0·0438866	0·0318073	0·0324100	0·0800211	0·0579962
81–85	0·0263092	0·0692899	0·0490771	0·0500741	0·1318791	0·0934082
86–90	0·0363334	0·1020720	0·0706528	0·0721847	0·2027898	0·1403682
91–95	0·0460316	0·1379413	0·0933106	0·0954617	0·2860669	0·1935104
96–100	0·0526031	0·1681462	0·1111572	0·1138723	0·3639939	0·2406272

comparable ages, all other things being equal, a comparison of the predicted force of mortality for female smokers with male non-smokers reveals that this situation is reversed for all ages. Here things are not all equal, certainly in at least one known respect.

To compute the predicted rates of mortality at intermediate ages the following transformed version of equation (4.1) is needed:

$$\log(\mu_{x+1/2}^{(ij)}) = \alpha + (\tau + \theta_i + \psi_j)\left(\frac{x-8}{5}\right) + \beta\left(\frac{x-8}{5}\right)^2 + \gamma\left(\frac{x-8}{5}\right)^3,$$

$$i = 1, 2; j = 1, 2, 3$$

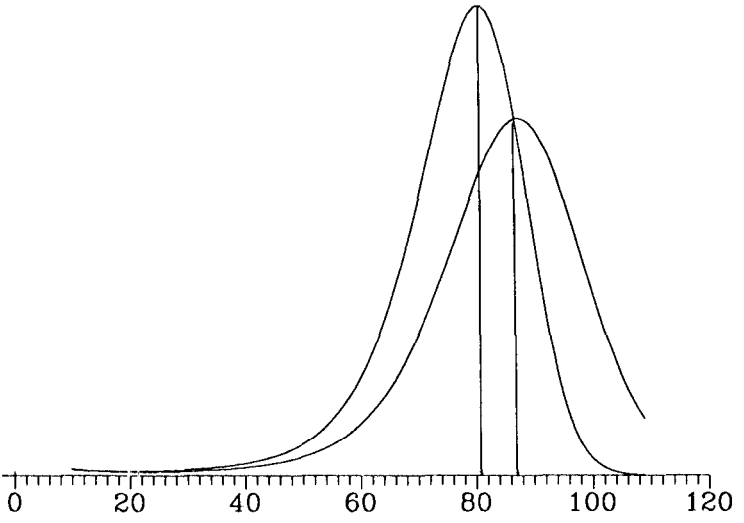
where x is now the age in years, and we have modified the notation identifying the unit associated with the specific value of μ in an obvious manner. Then, using the approximation:

$$q_x^{(ij)} = 1 - \exp(-\mu_{x+1/2}^{(ij)})$$

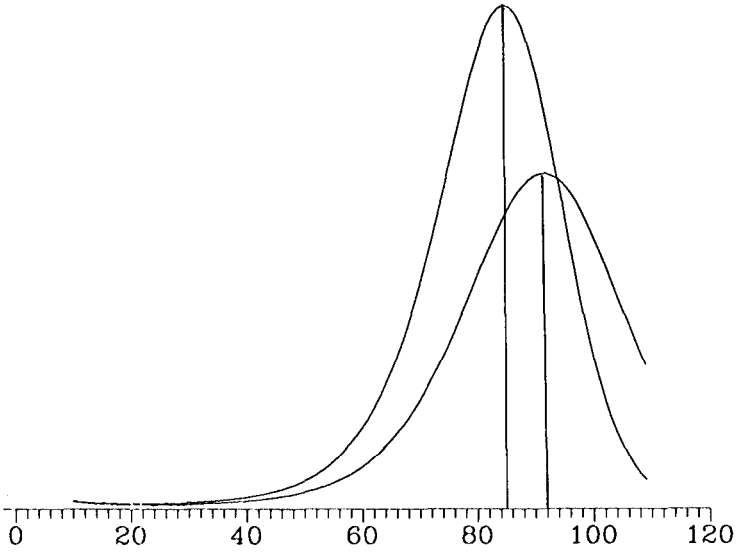
to compute the probability $q_x^{(ij)}$, that a life aged x dies before age $(x+1)$, the ordinates $l_x^{(ij)}\mu_x^{(ij)}$, of the curve of deaths are computed for individual ages x, making use of the identity:

$$l_{x+1}^{(ij)} = (1 - q_x^{(ij)})l_x^{(ij)}$$

and the arbitrary radix $l_{10}^{(ij)} = 100,000$. The resulting curves of death for non-



Curve of deaths, male non-smokers (right), smokers (left)



Curve of deaths, female non-smokers (right), smokers (left)

Figure 4.2.

smokers and for smokers are plotted on the same axes for females and for males separately in Figure 4.2. It is of interest to note that the maxima for these curves occur at the following ages:

	non-smoker	smoker
females	92	85
males	87	81

The findings of this analysis reinforce those of the CMI study group using more traditional actuarial methods.

5. Implementation

The re-analysis of the CMI smokers mortality data presented here was carried out using the GLIM software computer package. The interactive nature of the package means that it is possible to tailor a program to meet the specific needs of the problem. This package has much to offer the actuarial practitioner as a tool for analysing data in both life and non-life insurance, as is apparent from the ever-increasing number of actuarial applications of GLMs to appear in the literature.

Acknowledgement

The author wishes to thank both the CMI Bureau for permission to model the assured lives smokers study data set and Jillian Evans, secretary to the CMI Bureau, for helpful discussions about these data.

References

CMIB (1993). The mortality of smokers and non-smokers, 1988–89. *CMIR*, **13**, 109.

Evans, J. V. (1993). Smoker and non-smoker mortality. *The Actuary*, **3**, 5, 16–17.

Forfar, D. O., McCutcheon, J. J. & Wilkie, A. D. (1988). On graduation by mathematical formula. *J.I.A.* **115**, 1.

Renshaw, A. E. (1991). Actuarial graduation practice and generalised linear and non-linear models. *J.I.A.* **118**, 295.

Renshaw, A. E. (1992). Joint modelling for actuarial graduation and duplicate policies. *J.I.A.* **119**, 69.

Renshaw, A. E., Haberman, S. & Hatzopoulos, P. (1994). The modelling of recent mortality trends in United Kingdom male assured lives (submitted).