THE GENETICS OF BREAST AND OVARIAN CANCER III: A NEW MODEL OF FAMILY HISTORY WITH INSURANCE APPLICATIONS

By Eng Hock Gui, Baopeng Lu, Angus Macdonald, Howard Waters and Chessman Wekwete

ABSTRACT

Insurers' access to genetic test results is often restricted and the only genetic information that might be collected during underwriting in some countries is family history. Previous studies have included family history in a simple way but only for diseases which have no cause other than gene mutations, because then the event 'affected parent' contributes all possible information short of a genetic test result. We construct a model of breast cancer (BC) and ovarian cancer (OC) — common diseases with rare genetic variants — in which the development of a family history is represented explicitly as a transition between states, hence as part of the applicant's own life history. This allows the impact of a moratorium to be modelled. We then apply this family history model to life insurance in a semi-Markov framework and to critical illness (CI) insurance in a Markov framework to: (a) estimate premium ratings depending on genotype or family history; and (b) model the potential cost of adverse selection.

CONTACT ADDRESS

Angus Macdonald, Department of Actuarial Mathematics and Statistics, Heriot-Watt University, Edinburgh EH14 4AS, U.K. Tel: +44(0)131-451-3209; Fax: +44(0)131-451-3249; E-mail: A.S.Macdonald@ma.hw.ac.uk

1. Introduction

Breast cancer (BC) and ovarian cancer (OC) are relatively common, and both have inherited variants, accounting for a small proportion of all cases, that are now known to be caused primarily by mutations in the BRCA1 and BRCA2 genes. Genetic tests have been available in clinical practice since the mid 1990s, and inevitably this has led to a debate about the use of such test results in insurance underwriting.

Macdonald, Waters & Wekwete (2003a, 2003b) proposed a model of BC and OC leading to estimates of the probabilities that an applicant has a BRCA1/BRCA2 mutation given complete or incomplete knowledge of her family history, and applied the model to critical illness (CI) insurance underwriting. Here, we expand and update their work as follows:

(a) We propose a simplified model of 'family history' that greatly reduces the computational burden of the approach used before.

(b) We estimate survival rates after onset of BC and OC and extend the previous study to life insurance as well as CI insurance.

(c) We incorporate more recent estimates of rates of onset associated with these mutations, from epidemiological studies that better control the selection of subjects.

Thus, this paper represents Part III in the series launched by Macdonald, Waters & Wekwete (2003a, 2003b).

In Section 2 we review the current genetic epidemiology of BC and OC. In Section 3 we divide the general population into five sub-populations, according to the presence or absence of BRCA1 and BRCA2 mutations in the family and in the applicant for insurance, and estimate the proportion in each sub-population. In Section 4, we discuss the relationship between insurance and genetics and then propose a new way to include the development of a family history in an insurance model, with particular focus on underwriting, and estimate the relevant parameter (the 'rate of developing' a family history). In Section 5, we use a semi-Markov model to find life insurance premium ratings given either a known mutation or a family history. In Section 6, we investigate the potential effects of adverse selection under different moratoria on the use of genetic information. In Sections 7 and 8, we present similar calculations and estimate premium ratings and the effects of adverse selection in a CI insurance market. Our conclusions are given in Section 9.

2. The Epidemiology of Breast Cancer and Ovarian Cancer

2.1 *Mutation Frequencies*

Approximately 2% of BC is believed to occur in women who carry an inherited mutation in a gene called BRCA1 (Miki *et al.*, 1994; Futreal *et al.*, 1994); a smaller proportion of BC is due to inherited mutations in a gene called BRCA2 (Wooster *et al.*, 1995). Mutations in BRCA1 and BRCA2 also confer a predisposition to OC; about 5–13% of all OC results from the inheritance of gene mutations (Houlston *et al.*, 1991; Claus *et al.*, 1996). Mutations are rare, occurring in less than 0.2% of the population (Ford & Easton, 1995), but women with mutations are very likely to develop one or both of these cancers and develop them at relatively young ages (Easton *et al.*, 1997). The high incidence of BC and OC among mutation carriers suggests that the most highly predictive risk factor is a family history of BC and/or OC. In the case of BC, 4–5% of cases are believed to be due to inherited factors, and a family history of the disease has been the most reliable way to identify high-risk groups in the general population.

In early studies, most estimates of BC and OC risks were derived from high-risk populations, including families selected *because* they contained unusually large numbers of women who had had BC or OC. This probably leads to ascertainment bias; any families in which BRCA1 or BRCA2 mutations, by chance or for biological reasons, were not associated with unusually large numbers of cases, would not be included in the studies.

Table 1 shows some estimates of allele frequencies ('alleles' are different versions of the same gene, such variation being the source of genetic differences between individuals). Here, we assume that there are two alleles of each gene, one being the normal version and

Table 1: Estimates of the population frequency of mutations in BRCA1 and BRCA2 alleles.

Source	Frequency of mutations in	
	BRCA1 alleles	BRCA2 alleles
Claus *et al.* 1994	0.0033	Not in study
Parmigiani *et al.* 1998(a)	0.0006	0.00022
Parmigiani *et al.* 1998(b)	0.0008	0.0003
Parmigiani *et al.* 1998(c)	0.00045	0.000165
Antoniou *et al.* 2000	0.00128	0.00172
Antoniou *et al.* 2002	0.0005829	0.000676

Table 2: Published penetrance estimates of BRCA1 mutations, expressed as the cumulative probability of onset of BC and OC at ages 50 and 70, with 95% confidence intervals where available.

	Breast Cancer		Ovarian Cancer	
Source	Age 50	Age 70	Age 50	Age 70
Ford *et al.* (1994)	0.73 (0.49–0.87)	0.87 (0.72–0.95)	0.29 (0.16–0.40)	0.44 (0.28–0.56)
Easton *et al.* (1995)	0.51	0.85	0.29	0.44
Antoniou *et al.* (2000)	0.34 (0.17–0.60)	0.50 (0.26–0.82)	0.21 (0.8–0.47)	0.68 (0.36–0.94)
Anglian BCSG (2000)	0.32 (0.02–0.62)	0.47 (0.05–0.82)	0.11 (0.01–0.74)	0.36 (0.04–0.99)

the other a mutated version. Because mutations are rare and, in this case, dominantly inherited, we know that:

(a) the probability that a woman carries a mutation in one or other gene, is approximately $2f(1-f)$ where f is the relevant allele frequency; and

(b) the probability that a child, one of whose parents is a mutation carrier, inherits the mutation, is approximately $1/2$.

2.2 *Incidence Rates of BC and OC Associated with BRCA1 and BRCA2 Mutations*

Let $\mu_x^{BC,BRCA1}$ represent the rate of onset (alternatively force of onset, transition intensity) of BC at age x for a woman who carries a BRCA1 mutation, and define rates of onset relating to OC and to BRCA2 mutations analogously. This quantity may be reported in epidemiological studies. Alternatively, penetrances may be reported, which are cumulative probabilities of onset by age x assuming all other decrements to be absent. Let $q_x^{BC,BRCA1}$ be the penetrance of BRCA1 mutations in respect of BC by age x, and so on. Penetrances and onset rates are related by:

$$q_x^{BC,BRCA1} = 1 - \exp\left(-\int_0^x \mu_t^{BC,BRCA1} dt\right) \tag{1}$$

and so on. Tables 2 and 3 summarise published penetrance estimates of BRCA1 and BRCA2 mutations, respectively. The studies by Ford *et al.* (1994, 1998), Easton *et al.* (1995) and Antoniou *et al.* (2000) were based on multiple-case families while that by the

Table 3: Published penetrance estimates of BRCA2 mutations, expressed as the cumulative probability of onset of BC and OC at ages 50 and 70, with 95% confidence intervals where available.

	Breast Cancer		Ovarian Cancer	
Source	Age 50	Age 70	Age 50	Age 70
Ford *et al.* (1998)	0.28 (0.09–0.44)	0.84 (0.43–0.95)	0.004 (0.0–0.1)	0.27
Antoniou *et al.* (2000)	0.19	0.71	0.01	0.31
Anglian BCSG (2000)	0.18 (0.02–0.22)	0.56 (0.05–0.80)	0.03 (0.0–0.19)	0.10 (0.01–0.55)

Anglian Breast Cancer Study Group & Pharoah (2000) (Anglian BCSG) was population-based.

The recent meta-analysis by Antoniou *et al.* (2003) pooled pedigree data from 22 studies including 8,139 index cases unselected for family history. Of the 22 studies, 16 ascertained cases through population-based cancer registries, while the others were unselected, hospital-based studies. Age-specific cumulative risks were estimated by Kaplan-Meier methods, and approximate piecewise constant rates of onset were obtained as shown in Tables 4 and 5. These are ideal for modelling CI insurance, since the diagnosis of BC or OC will result in a valid claim under virtually all CI insurance policies. We fitted truncated Gamma functions to the piecewise constant onset rates, using unweighted least squares:

$$\mu_x^{BC,BRCA1} = \frac{1.39}{\Gamma(12.2)}\left(0.21^{12.2}e^{-0.21x}x^{11.2}\right) \tag{2}$$

$$\mu_x^{BC,BRCA2} = \frac{1.58}{\Gamma(8.37)}\left(0.10^{8.37}e^{-0.1x}x^{7.37}\right) \tag{3}$$

$$\mu_x^{OC,BRCA1} = \frac{11.19}{\Gamma(5.48)}\left(0.03^{5.48}e^{-0.03x}x^{4.48}\right) \tag{4}$$

$$\mu_x^{OC,BRCA2} = \frac{0.13}{\Gamma(57.95)}\left(1.00^{57.95}e^{-1.00x}x^{56.95}\right). \tag{5}$$

Figures 1–4 show the piecewise constant rates with 95% confidence intervals and our fitted onset rates. Note that in our applications we do not consider ages over 60.

2.3 Post-onset Mortality Rates

To model life insurance, we need to estimate post-onset mortality rates. According to Souhami & Tobias (1998), these depend on age, time since diagnosis and any factors that affect prognosis after treatment, including cancer staging at diagnosis (for example tumour size and the extent of tumour spread). However there is little evidence yet that BC or OC in mutation carriers is associated with very different mortality, and we assume it is not.

Coleman *et al.* (1999) give details of cancer registrations in England and Wales, including the exact age at diagnosis and the exact time to death after diagnosis. Details like tumour site and date of diagnosis, are also given. We consider BC cases recorded in
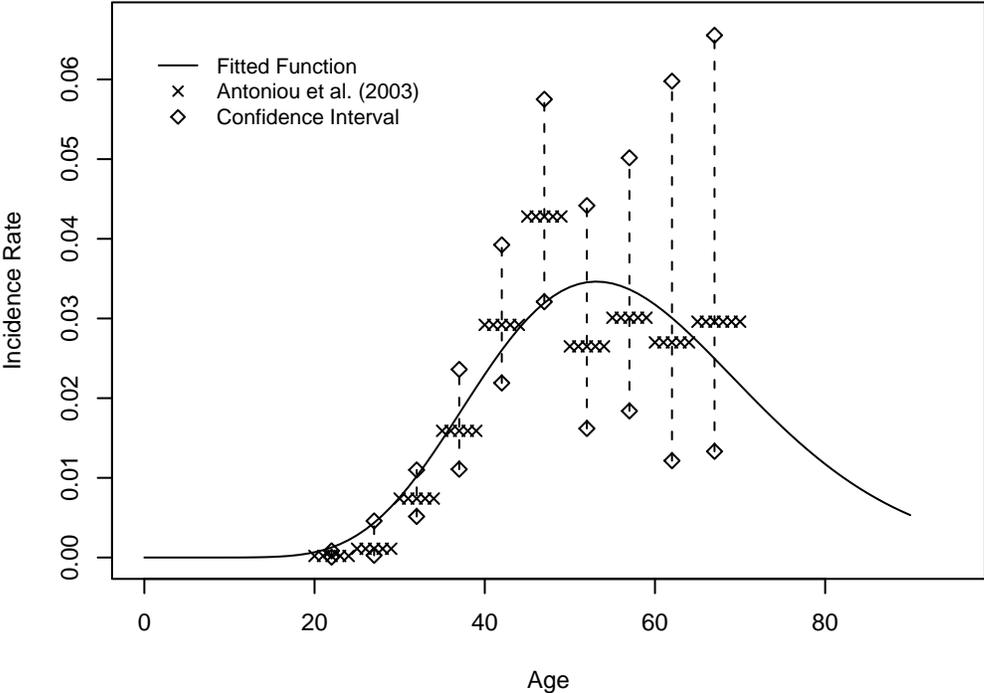
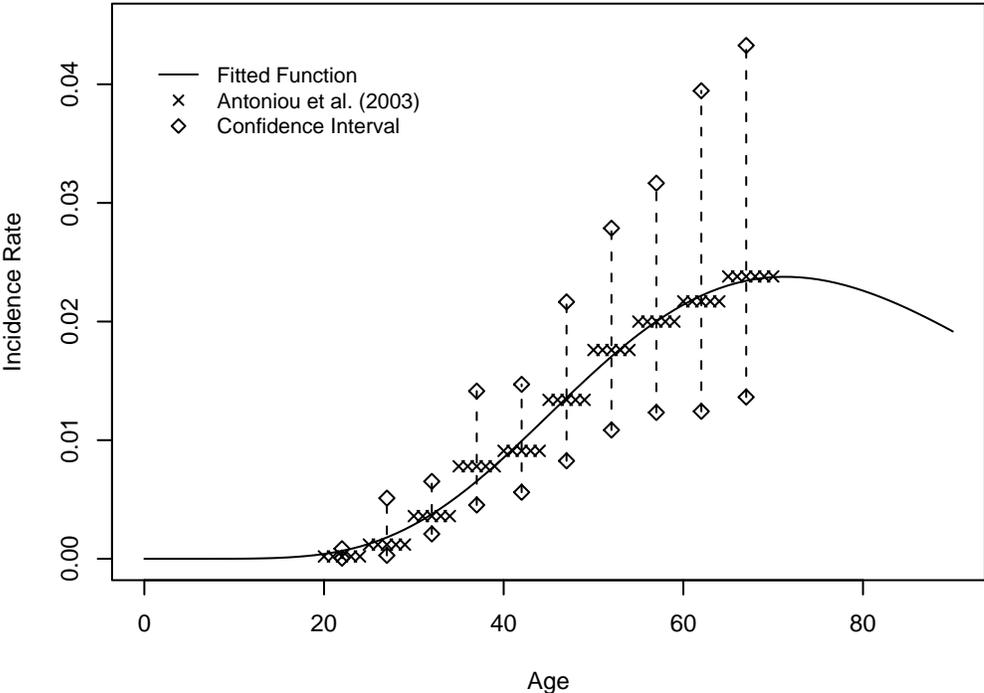Figure 1: Rate of onset of BC for BRCA1 mutation carriers.

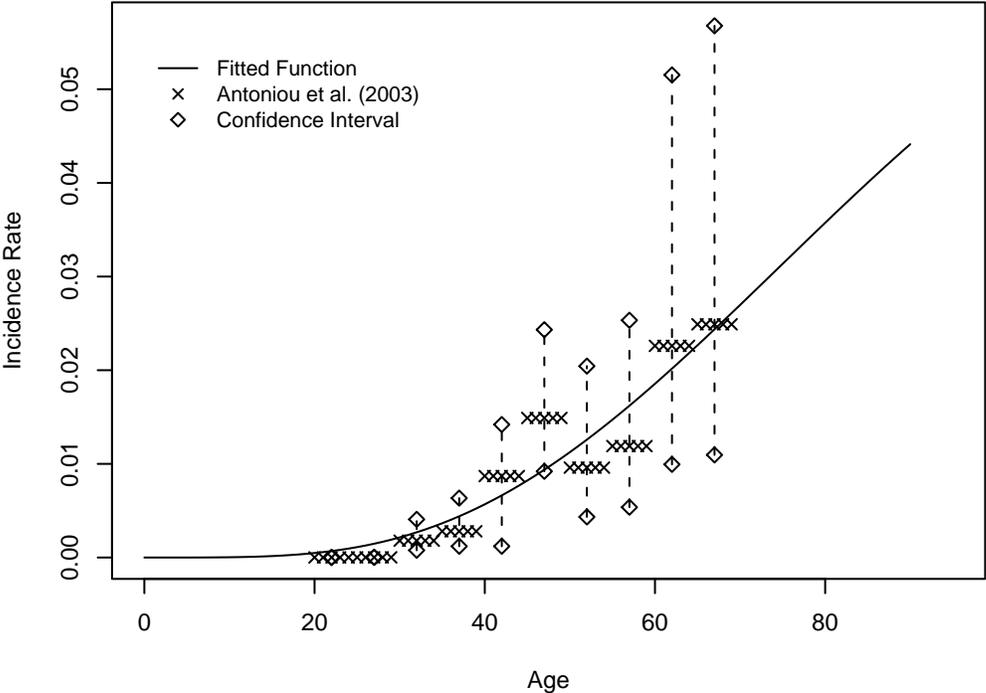Figure 2: Rate of onset of BC for BRCA2 mutation carriers.

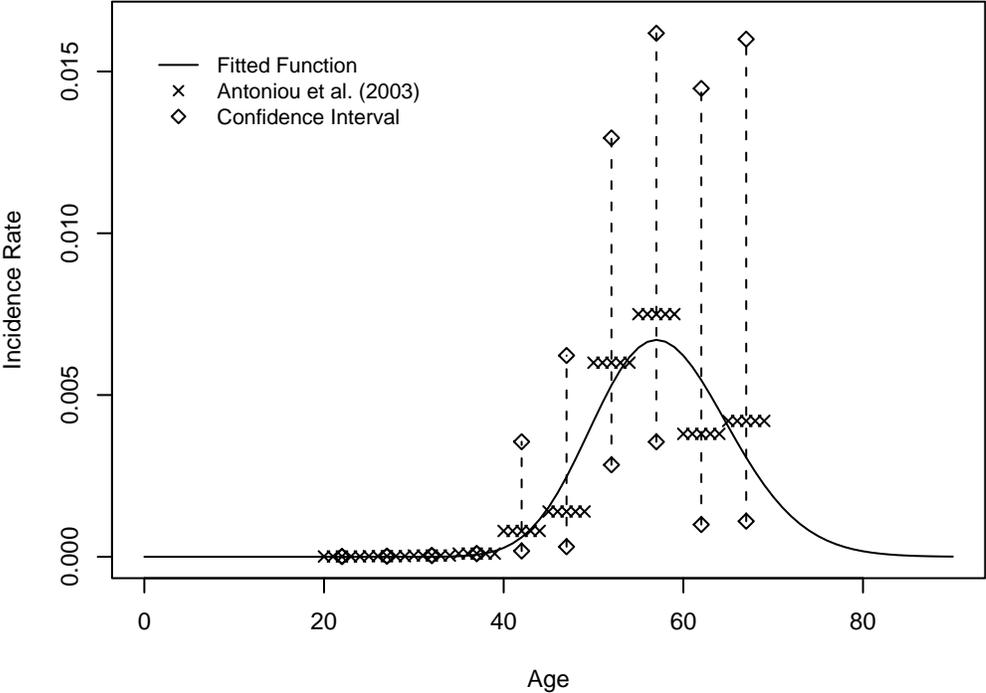Figure 3: Rate of onset of OC for BRCA1 mutation carriers.

Figure 4: Rate of onset of OC for BRCA2 mutation carriers.

Table 4: Estimated rates of onset of breast cancer (and 95% confidence intervals) in respect of BRCA1 and BRCA2 mutation carriers, based on the meta-analysis of Antoniou *et al.* (2003).

Age	BRCA1 mutation carriers		BRCA2 mutation carriers	
	Onset Rate	(CI)	Onset Rate	(CI)
20–24	0.0002	(0.000049–0.000835)	0.0002	(0.000047–0.000853)
25–29	0.0011	(0.000272–0.004594)	0.0012	(0.000284–0.005116)
30–34	0.0074	(0.00515–0.010988)	0.0036	(0.002093–0.006525)
35–39	0.0159	(0.01108–0.023609)	0.0078	(0.004534–0.014138)
40–44	0.0292	(0.02190–0.039238)	0.0091	(0.005607–0.014707)
45–49	0.0428	(0.03210–0.057513)	0.0134	(0.008257–0.021657)
50–54	0.0265	(0.01619–0.044167)	0.0176	(0.010853–0.027867)
55–59	0.0301	(0.01839–0.050167)	0.0200	(0.012333–0.031667)
60–64	0.0270	(0.01215–0.059786)	0.0217	(0.012428–0.039455)
65–69	0.0296	(0.01332–0.065543)	0.0238	(0.013631–0.043273)

Table 5: Estimated rates of onset of ovarian cancer (and 95% confidence intervals) in respect of BRCA1 and BRCA2 mutation carriers, based on the meta-analysis of Antoniou *et al.* (2003).

Age	BRCA1 mutation carriers		BRCA2 mutation carriers	
	Onset Rate	(CI)	Onset Rate	(CI)
20–24	0.00001	N/A	0.00001	N/A
25–29	0.00002	N/A	0.00002	N/A
30–34	0.0018	(0.000771–0.004078)	0.00004	N/A
35–39	0.0028	(0.001200–0.006343)	0.0001	N/A
40–44	0.0087	(0.005374–0.014201)	0.0008	(0.000178–0.003556)
45–49	0.0149	(0.009203–0.024322)	0.0014	(0.000311–0.006222)
50–54	0.0096	(0.004335–0.020439)	0.0060	(0.002842–0.012947)
55–59	0.0119	(0.005374–0.025335)	0.0075	(0.003553–0.016184)
60–64	0.0226	(0.009944–0.051528)	0.0038	(0.000995–0.014476)
65–69	0.0249	(0.010956–0.056772)	0.0042	(0.001100–0.016000)

1986–1990 inclusive and OC cases recorded in 1980–1990 inclusive, for women age 20–84 at diagnosis, but excluding:

(a) non-malignant tumours;

(b) persons whose vital status (dead or alive or censored) is unknown; and

(c) persons whose survival time after diagnosis is recorded as zero.

This leaves $110,697$ BC and $45,650$ OC cases. In Appendix A we fit a model of the mortality of women with BC or OC as a function of age and duration since diagnosis.

3. Risk Subpopulations

3.1 *Definition of Subpopulations*

Given our fitted onset and mortality rates (and corresponding morbidity and mortality rates for causes other than BC and OC) we can straightforwardly calculate premiums for known mutation carriers. However, this status is exactly what insurers in most jurisdictions are prevented from knowing or using, except perhaps for very large policies. Some countries have banned the use of family history as well as genetic test results, but in those that have not the actuarial problem reduces to estimating the probabilities:

$$\text{P[Applicant carries a causative mutation | Family history].} \qquad (6)$$

The interpretation of family history is simplest in respect of those disorders that have no causes other than mutations in a single gene, such as Huntington's disease (HD). Then the presence of the disease pinpoints a mutation carrier, and using Mendel's laws it is simple to calculate the genotype probabilities of blood relatives. If the disease is rare, dominantly inherited and has late onset, which most of those relevant to insurance are, the family history is usually an affected parent, and this supports the following simple model: the population can be divided into persons with no family history who are not at risk, and persons with a family history (taken to mean an affected parent), half of whom will be mutation carriers, the other half not. See Gutiérrez & Macdonald (2004) for an application of this model to HD.

BC, and to a lesser extent OC, is a common cancer of which a small proportion is attributable to mutations in single genes, so the above model cannot be used. Onset does not identify mutation carriers, there are no persons who are not at risk, and a history of BC and/or OC could arise by chance in any family. We are forced to take a slightly more elaborate approach. If we can model the probability of a family history developing, knowing whether or not a BRCA1 or BRCA2 mutation is present in the family, *and* if we know the prevalence of these mutations in the population, then Bayes' Theorem will allow us to estimate the probabilities (6). This suggests that we divide the population into subpopulations defined by the presence or absence of mutations in the family and in the applicant, defining five subpopulations that we label with index i:

(a) Family with no BRCA1 or the BRCA2 mutation carriers ($i=0$);

(b) BRCA1 mutation carrier family but applicant is not a carrier ($i=1$);

(c) BRCA1 mutation carrier family and applicant is a carrier ($i=2$);

(d) BRCA2 mutation carrier family but applicant is not a carrier ($i=3$);

(e) BRCA2 mutation carrier family and applicant is a carrier (i=4).

Here, we define a 'mutation carrier family' as a nuclear family (parents and siblings) in which one parent carries a mutation. Epidemiologists often study more extended pedigrees, but this definition is consistent with underwriting practice, in which it is uncommon for relatives other than parents and siblings to be taken into account. Since BRCA1 and BRCA2 mutations are rare, we ignore that possibility that the parents carry more than one mutated allele between them.

3.2 *The Distribution of Families in Subpopulations*

We take the allelic mutation frequencies estimated by Antoniou *et al.* (2002) as our starting point. Denote these $p_1 = 0.0005829$ for BRCA1 and $p_2 = 0.000676$ for BRCA2. From these we must calculate the proportion of families that fall into each subpopulation, namely the proportion of mother-father pairs in which a mutation is present. We estimate this by assuming that mutated and normal alleles are randomly distributed in the previous generation (that is, the grandparents of the applicant's generation) and that they are transmitted independently to offspring according to Mendel's laws. For simplicity, we ignore the possibility that any grandparent carries either two mutated copies of the same gene, or mutated copies of both genes.

Consider BRCA1 mutations. If we label a mutated BRCA1 allele A and a normal BRCA1 allele a, each grandparent has genotype Aa (carrier) or aa (non-carrier). Therefore the applicant's maternal grandparents may have included 0, 1 or 2 carriers, and the applicant's paternal grandparents likewise. Table 6 enumerates the possible combinations of genotypes of the grandparents, including the number of ways in which they can arise, each with equal probability. For example the genotype $(Aa, aa), (aa, aa)$ means that one grandparent is a mutation carrier, and this can arise in 4 ways because it could be any grandparent. The probability that any grandparent is a carrier is:

$$p_1^2 + 2p_1(1 - p_1) \approx 2p_1(1 - p_1) \tag{7}$$

where the omission of the first term may be justified on biological as well as numerical grounds because carriers of two mutations might not be viable. Table 6 sets out the calculation, including the probability the four grandparents will transmit a mutation to at least one of the applicant's parents. Summing the products of the terms in the last two columns, we find that the proportion of BRCA1 mutation carrier families in the population is 0.002328205. Similarly, we find the proportion of BRCA2 carrier families to be 0.002699435, therefore the proportion of non-carrier families to be 0.99497236. Remembering the assumption made at the end of Section 3.1, if the applicant is a member of a mutation carrying family she inherits the mutation with probability 1/2. Table 7 shows the resulting proportions of the five subpopulations in the whole population.

4. A Model of Family History

4.1 *Modelling Family History*

As mentioned in Section 3.1, family history is probably the most important item of information that may be used in underwriting. However, family histories are as diverse

Table 6: Calculation of the probability of at least one of the applicant's parents being a BRCA1 mutation carrier. $q_1 = 2p_1(1 - p_1)$ is approximately the probability that any given grandparent is a mutation carrier.

No. of Grandparent Carriers	Grandparents' Genotype	No. of Ways To Occur	Probability	Probability ≥ 1 Parent Is Carrier
0	$(aa, aa), (aa, aa)$	1	$(1 - q_1)^4$	0
1	$(Aa, aa), (aa, aa)$	4	$4q_1(1 - q_1)^3$	1/2
2	$(Aa, aa), (Aa, aa)$	4	$4q_1^2(1 - q_1)^2$	3/4
2	$(Aa, Aa), (aa, aa)$	2	$2q_1^2(1 - q_1)^2$	3/4
3	$(Aa, Aa), (Aa, aa)$	4	$4q_1^3(1 - q_1)$	7/8
4	$(Aa, Aa), (Aa, Aa)$	1	q_1^4	3/4

Table 7: Proportions of the five subpopulations in the whole population.

i	Subpopulation	Proportion
0	Non BRCA1/2 carrier family	0.994972360
1	BRCA1 carrier family, applicant non-BRCA1 carrier	0.001164102
2	BRCA1 carrier family, applicant BRCA1 carrier	0.001164102
3	BRCA2 carrier family, applicant non-BRCA2 carrier	0.001349718
4	BRCA2 carrier family, applicant BRCA2 carrier	0.001349718

as families themselves. Clinical geneticists may often obtain detailed pedigrees reaching back through several generations, in order to counsel patients as accurately as possible. Insurers, however, tend to seek simpler histories, not usually extending beyond parents and siblings. It may be argued that insurers should not base their decisions on less information than a clinician would regard as adequate, but insurers do not have the access and resources that would allow them to probe the personal details of an applicant's distant relatives. Given the difficulties of obtaining accurate medical information in the first place, concentrating on nuclear families may be reasonable. We will assume that an underwriter may learn how many of the applicant's parents and siblings suffered onset of BC or OC, and at what ages. We do *not* assume that the insurer learns about unaffected relatives; for example an applicant with two affected sisters may have ten other sisters or she may have none.

We model the insurers' use of this simple family history by supposing that there is a threshold beyond which the family history is a significant risk factor, and below which it is not. The definition of the threshold may include the number of affected relatives, the ages at which they were affected, and whether they are parents or siblings. Given a particular applicant and her family, whose members were born on known dates, we can pinpoint the exact time at which a 'family history' appeared; it is the time at which a relative was diagnosed with BC or OC and caused the threshold to be crossed. If the applicant is applying for insurance before that date she will not have a family history; if she is applying after that date she will. The key point is that we can define the 'onset' of
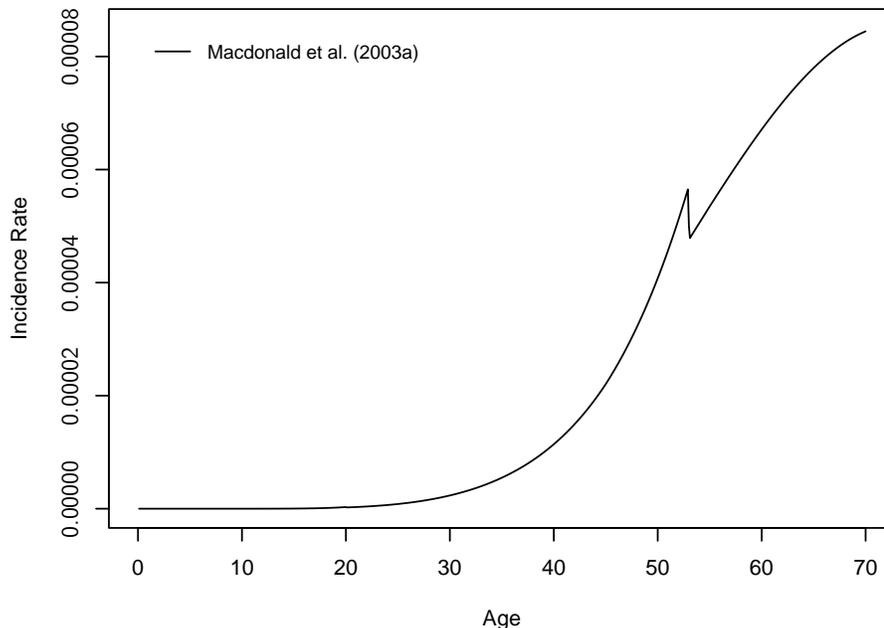
Figure 5: Rate of onset of family history for the non-mutation-carrying subpopulation, $i = 0$.

a family history as an event in the applicant's life history, so we can easily incorporate it into a multiple-state framework for modelling a life history. The only problem is to find the transition intensity that governs the 'onset' of family history.

The subpopulations defined in Section 3.1 give just the framework we need, both for finding this intensity and for insurance applications. Knowing whether or not the applicant has a parent who is a carrier allows us to find the probability that a family history will have appeared by the time the applicant reaches any age x. This leads to the required intensity. The only missing information is the details of the applicant's siblings and their ages; we can model this in a simple way and then average over the distribution of family sizes. And, knowledge of the applicant's own genotype determines her actual risk of BC and OC, which leads to the insurance applications.

4.2 *The Rate of Onset of a Family History*

We choose the following underwriting threshold for a family history: two first-degree relatives suffer onset of BC or OC before age 50. This is reasonably representative of actual practice. We also simplify the family structure by assuming that the applicant's siblings are all the same age as her, and their mother is 30 years older. Therefore the family structure is determined by the number of siblings. Let ps_j be the probability that the applicant has j sisters; Table 8 gives this distribution, based on Macdonald, Waters & Wekwete (2003a). We assume that no woman has more than six sisters.

Consider an applicant in subpopulation i. Let T_i be the random time at which she crosses the above threshold and develops a family history ($T_i = \infty$ if she never crosses it). We assume that the applicant's mother is free of BC and OC at age 30. Define $q_M^i(x)$ to

Table 8: Distribution of the number of the applicant's sisters. Source: Macdonald, Waters & Wekwete (2003a)

No. of sisters (j)	Probability (ps_j)
0	0.54759802
1	0.33055298
2	0.09749316
3	0.02111590
4	0.00285702
5	0.00035658
6	0.00002634

be the probability that she (the mother) has developed BC or OC by age $x + 30$ or age 50, whichever is smaller, conditional on being free of BC and OC at age 30. Define $q_S^i(x)$ similarly, to be the probability that any of the applicant's sisters has developed BC or OC by age x or age 50, whichever is smaller. Then:

$$
\begin{aligned}
\mathrm{P}[T_i \leq x] \; = \; & q_M^i(x) \times \sum_{j=1}^{j=6} ps_j(1 - (1 - q_S^i(x))^j) \\
& + (1 - q_M^i(x)) \times \sum_{j=2}^{j=6} ps_j(1 - (1 - q_S^i(x))^j - jq_S^i(x)(1 - q_S^i(x))^{j-1}). \quad (8)
\end{aligned}
$$

We split the calculation into three cases according to the sub-population. In the following, let $\mu_x^{BC,NC}$, $\mu_x^{OC,NC}$ be the incidence rates of BC and OC respectively for a non-mutation-carrier, taken to be the population onset rates, and let μ_s^{dead} be the force of mortality before onset of BC and OC. Then the probability that a non-carrier is healthy at age x, denoted p_x^{NC}, is given by:

$$
p_x^{NC} = \exp\left(-\int_0^x (\mu_s^{BC,NC} + \mu_s^{OC,NC} + \mu_s^{dead})ds\right) \quad (9)
$$

and we may define p_x^{BRCA1} and p_x^{BRCA2} similarly for BRCA1 and BRCA2 mutation carriers. (Note that these are not equivalent to the penetrances as defined in Equation (1) because the latter are single-decrement quantities).

(a) For subpopulation $i = 0$:

$$
q_M^0(x) = \int_0^{\min[x,20]} \frac{p_{x+30}^{NC}}{p_{30}^{NC}} (\mu_{30+r}^{BC,NC} + \mu_{30+r}^{OC,NC}) \, dr \quad (10)
$$

and:

$$
q_S^0(x) = \int_0^x p_r^{NC}(\mu_r^{BC,NC} + \mu_r^{OC,NC}) \, dr. \quad (11)
$$

Figure 6: Rate of onset of family history for the BRCA1 subpopulations, $i = 1$ and $i = 2$, based on the recent meta-analysis by Antoniou *et al.* (2003) with, for comparison, rate of onset of family history from Macdonald *et al.* (2003a), which was based on earlier studies.
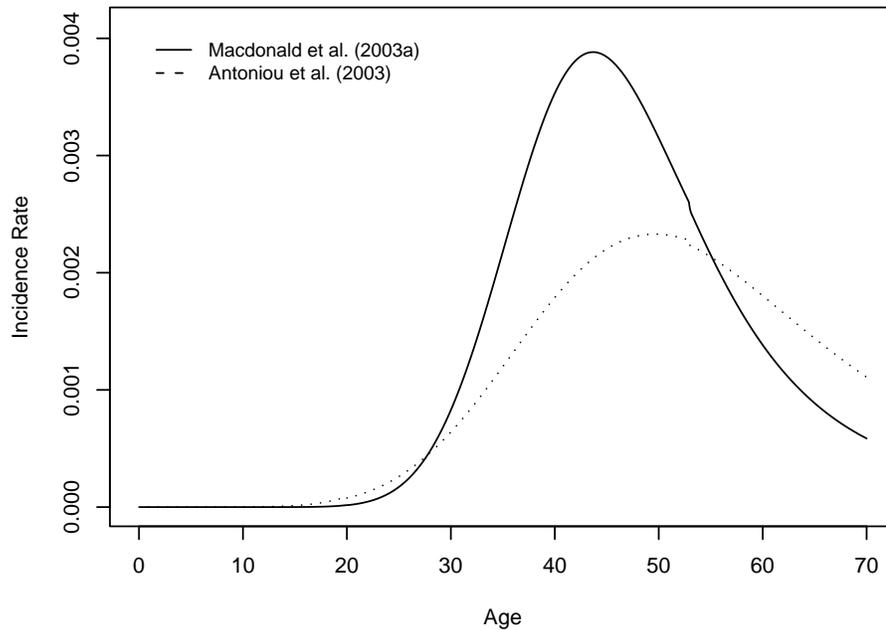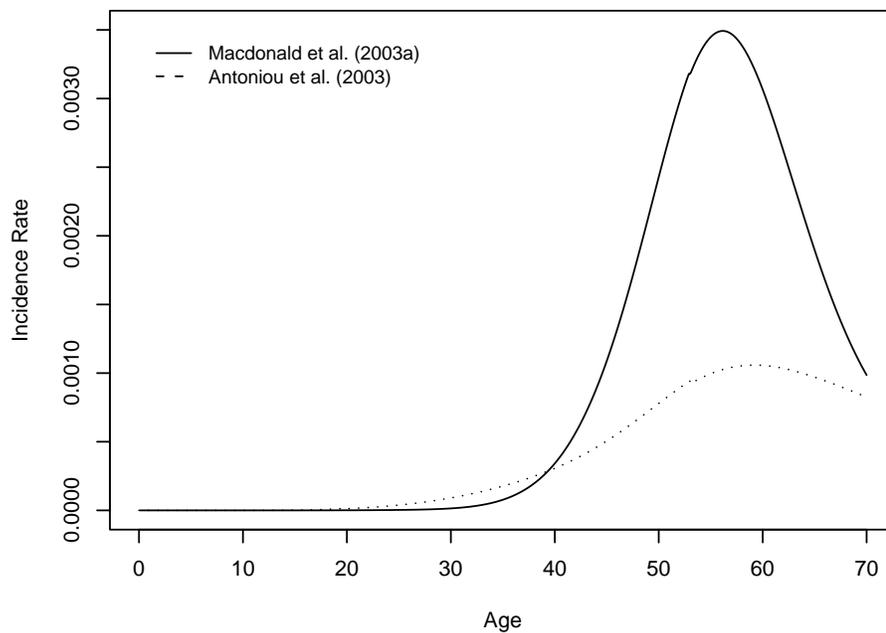

Figure 7: Rate of onset of family history for the BRCA2 subpopulations, $i = 3$ and $i = 4$, based on the recent meta-analysis by Antoniou *et al.* (2003) with, for comparison, rate of onset of family history from Macdonald *et al.* (2003a), which was based on earlier studies.

(b) For subpopulations $i = 1$ and $i = 2$, let c_M^i be the probability that the applicant's mother is a BRCA1 mutation carrier given she was healthy at age 30. We have $c_M^i = X/(X + Y)$ where:

$$X = p_{30}^{BRCA1} \times \text{P[Mother BRCA1 carrier]}$$
$$Y = p_{30}^{NC} \times \text{P[Mother not BRCA1 carrier]}$$

By the definition of a mutation carrier family, both carrier probabilities above are $1/2$. Then:

$$q_M^i(x) = (1 - c_M^i) \int_0^{\min[x,20]} \frac{p_{x+30}^{NC}}{p_{30}^{NC}} (\mu_{30+r}^{BC,NC} + \mu_{30+r}^{OC,NC}) \, dr \tag{12}$$

$$+ \, c_M^i \int_0^{\min[x,20]} \frac{p_{x+30}^{BRCA1}}{p_{30}^{BRCA1}} (\mu_{30+r}^{BC,BRCA1} + \mu_{30+r}^{OC,BRCA1}) \, dr \tag{13}$$

and:

$$q_S^i(x) = \frac{(1 - p_x^{BRCA1}) + (1 - p_x^{NC})}{2}. \tag{14}$$

(c) For subpopulations $i = 3$ and $i = 4$ we use the same method as in (b) above with the intensities relating to BRCA2 mutations in place of those relating to BRCA1 mutations.

We find the distributions of the T_i using the incidence rates from Macdonald, Waters & Wekwete (2003a) for non-mutation carriers, and those from Section 2.2 based on Antoniou *et al.* (2003) for mutation carriers, and then obtain the intensities of onset of family history by numerical differentiation. For comparison, we also tried the BC and OC onset rates used by Macdonald, Waters & Wekwete (2003a) based on Ford *et al..* (1998). Figures 5, 6 and 7 show the results.

Note that the discontinuity in these rates, which is most noticeable in Figure 5 but is also present in the other figures, is caused by the introduction of a BC screening program in the UK in 1988. Screening does not change a tumour's progression, but it may result in earlier diagnosis.

Clearly, the rates of onset of family histories are very much lower using the more recent, and supposedly less selective, study.
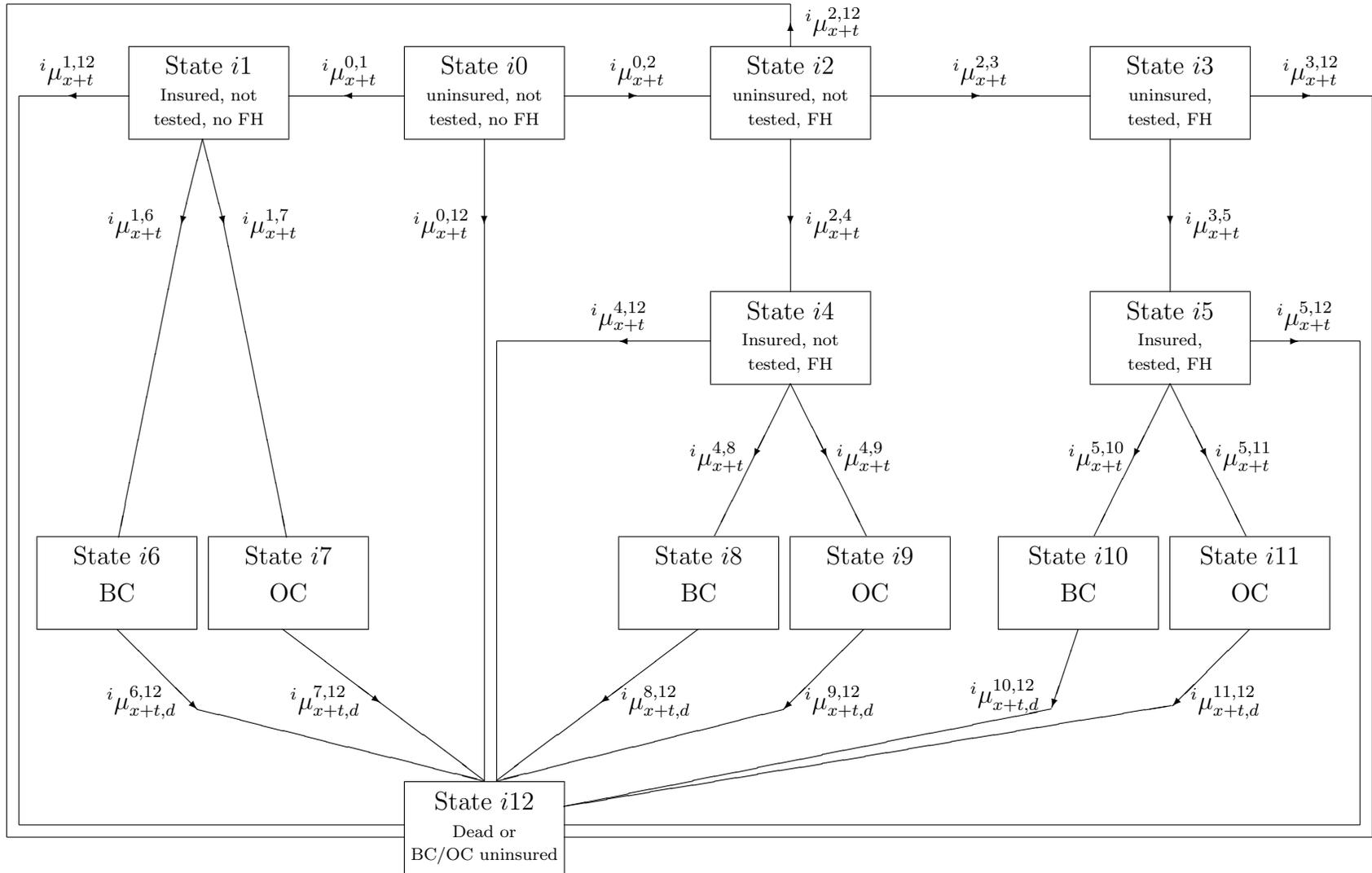
Figure 8: A semi-Markov model of family history, genetic testing, insurance purchase and life insurance events for a person in the ith subpopulation (FH = family history present). Intensities are functions of current age, $x+t$, and duration since onset of BC or OC, d.

5. Application of the Family History Model to Life Insurance

5.1 *A Life Insurance Model Including Family History*

The model in Figure 8 is a continuous-time semi-Markov model of a person's life history in a life insurance market. It represents both the insurance-buying behaviour and the claims experience of a person in the ith subpopulation. This model allows us to represent the following characteristics of a life insurance market:

(a) Mutation frequencies are represented by the proportions in each subpopulation;
(b) Underwriting classes are defined as sets of insured states, within each of which the same premium rate will be charged (which we assume to be based on the usual equivalence principle, ignoring all expenses and loadings). The effect of a moratorium may be to redefine the underwriting classes.
(c) The rate of insurance purchase by persons without a family history describes the market size.
(d) The rate of developing a family history is represented by the intensity $^i\mu_{x+t}^{0,2}$.
(e) Adverse selection is represented by: (i) the rate of insurance purchase after receiving an adverse test result; and, if family history may not be used, (ii) the rate of insurance purchase given a family history before testing; and (iii) any tendency to buy larger amounts of insurance given any adverse information.

The premium rate charged to persons in average good health is called the 'ordinary' rate, and this class is denoted OR. If the information disclosed to the insurer during underwriting indicates poor health or higher risk of poor health, then an extra premium may be charged, often expressed as a percentage of the OR rate. Often, for convenience, extra premiums increase in increments of 25% or 50% of the OR rate, rather than attempting any more precise rating. As reported in Macdonald (2003), about 95% of applications are accepted at OR, 4% are accepted with an extra premium, and 1% are declined.

5.2 *Numerical Procedures*

Suppose an individual in the ith subpopulation, insured at age x, is in state ij at age $x + t$. A premium will be paid continuously at rate b_{x+t}^{ij} per annum and a benefit of c_{x+t}^{ijk} will be received on transition into state ijk (these will be zero unless state ij is an insured state). If this were a Markov model we would define $_t^i p_x^{jk}$ to be the transition probability, that a person in state ij at age x is in state ik at age $x + t$, and $_t^i V_x^j$ to be the prospective reserve, namely the expected present value (EPV) at force of interest δ of the future loss in respect of a life in state ij, at age $x + t$. These values satisfy respectively Kolmogorov's Forward Equations:

$$\frac{\partial}{\partial t}\,_t^i p_x^{jk} = \sum_{l \neq k} {}_t^i p_x^{jl} \mu_{x+t}^{lk} - \sum_{l \neq k} {}_t^i p_x^{jk} \mu_{x+t}^{kl} \tag{15}$$

boundary condition $_0^i p_x^{jk} = \delta_{jk}$ (the Kronecker delta); and Thiele's Equations:

$$\frac{\partial}{\partial t}\,_t^i V_x^j = {}_t^i V_x^j \delta + b_{x+t}^{ij} - \sum_{k \neq j} {}^i\mu_{x+t}^{j,k}(c_{x+t}^{ijk} + {}_t^i V_x^k - {}_t^i V_x^j) \tag{16}$$

with the boundary condition $_T^i V_x^j = 0$, where T is the term of the policy. But the model is not Markov because rates of mortality after onset of BC or OC depend on duration since onset. To simplify calculation of first moments, as in Gutiérrez & Macdonald (2004), we modify the benefits and premiums so that on onset of BC or OC, premium payment ceases and a sum assured of $_{0,t}^{\;\;i} V_x^j$ is paid, where $_{d,t}^{\;\;i} V_x^j$ is the prospective reserve needed in respect of a life in state ij at age $x + t$ with duration d since entering the state. This is equivalent to the insurer immediately reinsuring the liability upon onset. These modified cashflows eliminate the duration-dependent intensities from the underlying multiple-state model, and are adapted to the resulting Markov model. We solved the resulting differential equations using a Runge-Kutta algorithm with step-size 0.0005 years and force of interest $\delta = 0.05$. The modified sum assured is found by numerical integration:

$$_{0,t}^{\;\;i} V_x^j = \int_0^{T-t} e^{-\delta s} \exp\left(- \int_0^s {}^i \mu_{x+t+r,r}^{j,12} \; dr \right) \left({}^i \mu_{x+t+s,s}^{j,12} \, c_{x+t+s}^{ij,12} - b_{x+t+s}^{ij} \right) \, ds \qquad (17)$$

based on the original cashflows.

5.3 *Sensitivity Analysis*

Macdonald *et al.* (2003b) allowed for ascertainment bias in an approximate fashion, by reducing by 50% or 75% the rate of onset for mutation carriers based on Ford *et al.* (1998), because the latter study used highly selected families. Antoniou *et al.* (2003), upon which this study is based, used less highly selected families but the possibility of ascertainment bias is still present, since selection was through an index (affected) patient. Therefore we allow approximately for possible bias by also considering onset rates 50% of those estimated directly from Antoniou *et al.* (2003).

5.4 *Premium Ratings for Mutation Carriers*

Table 9 shows premium rates, as a percentage of OR rates, for persons with a known BRCA1 or BRCA2 mutation. We observe the following:
(a) A BRCA1 mutation is much more severe than a BRCA2 mutation as we should expect from the rates of onset. The effect of reducing the penetrance of BRCA1 (by 50% of that observed) is substantial, especially where the additional risk is highest. All BRCA1 cases with reduced penetrance would be insurable given the typical limit for an offer of life cover in practice of 400–500% of OR rates.
(b) All cases of BRCA2 mutation carriers are insurable. The highest premium was about 260% of the OR premium, well within the the limit for insurability of 400–500% of OR rates.
(c) Using the onset rates for mutation carriers reduced to 50% of those fitted to the data, all premiums drop below the limit of 400–500% of OR rates and hence all cases become insurable.

5.5 *Premium Ratings for Persons With A Family History*

Subramanian *et al.* (1999) developed a double-decrement model of BC and OC to evaluate the increased mortality and term insurance costs for women who have a family history. In their model, the impact of a family history or a BRCA mutation was measured by the increase in the relative costs of term insurance. They concluded that family

history had a significant effect on premium rates. The ages at which the family members contributing to the history developed BC or OC was also significant. Our model, in a continuous-time multiple-state framework, allows us to formulate questions about insurance products in a more flexible way.

Given family history only, a level premium is calculated satisfying the equivalence principle, where the EPVs of the benefit and the premium are weighted averages of the EPVs in respect of each subpopulation, the weights being the probabilities of having a family history at the age when insurance was purchased. These are easily found by setting all the intensities in Figure 8 to zero, except $^i\mu_{x+t}^{0,2}$, $^i\mu_{x+t}^{0,12}$ and $^i\mu_{x+t}^{2,12}$. Table 10 shows the premiums for applicants with a family history of BC/OC as a percentage of those for standard risks. We make the following comments:

(a) Roughly, if the excess incidence rates were 50% of those observed, an additional premium of up to 50% might be charged; if the excess incidence rates were 100% of those observed, the extra premiums would about double. The highest premium is about 200% of the OR premium.

(b) For policies which expire after age 50, extra premiums decrease as the age at entry or the policy term increases. This is because surviving free of symptoms increases the probability that the applicant is not a mutation carrier. For example, the extra premium would not exceed 21% for a 10-year policy starting at age 50.

(c) All cases are within the limits of 400–500% of standard rates, and so would be regarded as insurable in the UK.

6. The Cost of Adverse Selection in a Life Insurance Market

Adverse selection may result if buyers of insurance have better information about their risk of claiming than has the seller. Ideally, insurance premiums should be set according to the mean risk of persons in the relevant underwriting class. In practice, underwriting classes are not truly homogeneous; when there is adverse selection, people who know they have an above-average risk of claiming may buy more insurance, while those who have a below-average risk may buy less. Then premiums based on the mean risk assuming uniform buying habits will be in deficit. Increasing the premium may skew peoples' buying habits yet further; perhaps an equilibrium will be reached, but it could be the extreme one in which only those at highest risk buy any insurance at all. Here, we do not explicitly model such a dynamic equilibrium, but we examine the premium increases that would result from a given pattern of insurance purchasing.

6.1 Parameterisation

We use the model in Figure 8 to evaluate the effect of adverse selection under various moratoria on the use of genetic information. Table 7 shows the proportions in each subpopulation. The incidence rates of BC and OC in BRCA1 and BRCA2 mutation carriers are shown in Table 4. Post-onset mortality rates are shown in Appendix B.

Table 9: Level net premiums for life insurance cover as a percentage of the premium for standard risks, for persons with a known BRCA1 or BRCA2 mutation. Excess BC/OC incidence rates are 100% and 50% of those observed.

Excess Incidence Rates	Age 20 at Entry				Age 30 at Entry			Age 40 at Entry		Age 50 at Entry
	10 Yrs	20 Yrs	30 Yrs	40 Yrs	10 Yrs	20 Yrs	30 Yrs	10 Yrs	20 Yrs	10 Yrs
	%	%	%	%	%	%	%	%	%	%
BRCA1:										
100%	395	600	587	461	645	604	462	464	386	297
50%	245	351	351	291	372	357	289	278	244	192
BRCA2:										
100%	211	258	256	233	247	250	228	204	206	185
50%	153	173	169	156	167	164	153	139	140	130

Table 10: Level net premiums for life insurance cover as a percentage of the premium for standard risks, for persons with a family history of BC and/or OC. Excess BC/OC incidence rates are 100% and 50% of those observed.

Excess Incidence Rates	Age 20 at Entry				Age 30 at Entry			Age 40 at Entry		Age 50 at Entry
	10 Yrs	20 Yrs	30 Yrs	40 Yrs	10 Yrs	20 Yrs	30 Yrs	10 Yrs	20 Yrs	10 Yrs
	%	%	%	%	%	%	%	%	%	%
100%	162	202	199	173	206	196	168	158	145	121
50%	130	152	151	139	153	150	136	128	123	110

The choice of the other parameters has been discussed elsewhere; except for the rate of genetic testing, which may be influenced by the features of the particular disorder, we follow the choices made by Gutiérrez & Macdonald (2004, Section 5.2) which we list below:

(a) *Insurance-Buying Behaviour:* We take $^i\mu_{x+t}^{0,1} = 0.05$ to represent a large insurance market and $^i\mu_{x+t}^{0,1} = 0.01$ to represent a smaller market.

(b) *The Rate of Insurance Purchase Given a Family History:* People with a family history, who are offered a higher than standard premium, may buy less insurance. We assume that in the large market, their rate of insurance purchase may be 0.05. 0.025 or 0.0 *per annum*; in the smaller market, they buy no insurance.

(c) *The Level of Adverse Selection:* Extreme adverse selection is represented by setting $^i\mu_{x+t}^{3,5} = 0.25$ in subpopulations 2 and 4. More moderate adverse selection is represented by a rate of purchase of $^i\mu_{x+t}^{3,5} = 2 \times {}^i\mu_{x+t}^{0,1}$.

The rate of genetic testing can be low in respect of severe disorders, especially if treatment is non-existent, traumatic or has a low rate of success. We assume $^i\mu_{x+t}^{2,3} = 0.02$, so that about 18% of people will have a test within 10 years of a family history emerging, and we will use higher rates in sensitivity analysis. Testing is generally only made available to women with a family history, and this often means an extensive pedigree collected by the clinician, so there is not an exact correspondence between this and our 'onset' of family history in the model, which is very much defined by insurance practice. Nevertheless, a rate of 0.02 *per annum* seems consistent with levels of genetic testing generally.

6.2 *Numerical Procedures*

We calculate premiums using the model in Figure 8 simply by conditioning on presence in the relevant insured state. Now we are also interested in insurance-buying behaviour, so we condition on presence, at outset (age 20) in the uninsured states. Note that a woman may have a family history by age 20, so she may be in one of the states $i0$ or one of the states $i2$, but we assume no genetic testing before age 20.

To keep all cashflows adapted to a Markov framework (as we did in Section 5.2 in respect of benefits) we need premium rates that depend only on the policyholder's current age, and not on age and duration since buying insurance. We achieve this by charging a rate of premium ρ_{x+t}^C, in the underwriting class consisting of the set of states C, equal to the expected rate of benefit outgo at age $x + t$. That is, we charge the weighted average of the intensities from the states in C into the various 'dead' states, the weights being the occupancy probabilities in the states in C each multiplied by the appropriate sum assured if this is not uniform. This method is similar to the current-cost charging used in unit-linked business. Let P_{20}^{ij} be the proportion of women in state ij at age 20 ($j = 0, 2$), and partition the underwriting class C into two subsets: C^α, those states with mortality independent of duration; and C^β, those states with mortality dependent on duration. Then:

$$\rho_{x+t}^C = \frac{\sum\limits_{k=0,2} \left(\sum\limits_{ij\in C^\alpha} {}^iP_{20}^k \; {}_tp_{20}^{kj} \; {}^i\mu_{x+t}^{j,12} + \sum\limits_{ij\in C^\beta} \int_0^t {}^iP_{20}^k \; {}_{s,t}p_{20}^{kj} \; {}^i\mu_{x+t,s}^{j,12} \; ds \right)}{\sum\limits_{k=0,2} \left(\sum\limits_{ij\in C^\alpha} {}^iP_{20}^k \; {}_tp_{20}^{kj} + \sum\limits_{ij\in C^\beta} \int_0^t {}^iP_{20}^k \; {}_{s,t}p_{20}^{kj} \; ds \right)} \tag{18}$$

Table 11: Percentage increases in premium rates for life insurance, under a moratorium on all genetic test results and adverse results respectively, for a market operating between ages 20 and 60.

Adverse Selection	Market Size	Rate of Purchase with A Family History	Moratorium on using	
			All test results %	Adverse results %
Severe	Large	Same as 'normal'	0.000219	0.000216
		Half as 'normal'	0.000331	0.000325
		Nil	0.000510	0.000500
	Small	Nil	0.002614	0.002575
Moderate	Large	Same as 'normal'	0.000093	0.000092
		Half as 'normal'	0.000186	0.000183
		Nil	0.000339	0.000332
	Small	Nil	0.000497	0.000490

where $_{s,t}{}^i p_x^{kj}$ is the probability that a person in state ik at age x will be in state ij at age $x + t$, with duration $s \leq t$ since entering it. We substitute rate ρ_{x+t+s}^C for rate b_{x+t+s}^{ij} in Equation (17).

If insurance is purchased exactly as assumed when calculating ρ_{x+t+s}^C, there is no adverse selection and the expected loss is zero. Under adverse selection, insurance is purchased at different rates than assumed but we continue to charge the premium rates ρ_{x+t+s}^C and the expected loss is positive. If we divide the latter by the EPV of all the premiums actually paid (that is, allowing for the rates of purchase with adverse selection occurring) this is exactly the proportionate increase in premiums that should be charged to everyone to meet the cost of the adverse selection.

6.3 *Moratoria on Genetic Test Results Alone*

First we assume that a moratorium allows family history to be used. Then there are two underwriting classes, the OR class and the 'family history' class, which we will denote FX. A strict moratorium will ban the use of all genetic test results, so the FX class will contain all insured states entered after developing a family history. A lenient moratorium (as in the UK) may allow someone who has a clear genetic test result to be offered the OR premium rate, in which case insured states entered after genetic testing in Subpopulations 0, 1 and 3 will be in the OR class.

Table 11 shows the percentage premium increases required to recoup the cost of adverse selection, under moratoria on all genetic test results (strict) and adverse test results (lenient) respectively. The values are very small, even negligible (we report several decimal places just so that differences can be seen) but there are some features to note:

(a) As we expect, the impact on the small market is larger than that on the large market.

(b) The costs are slightly higher under a moratorium on all genetic test results because under a moratorium on adverse test results, the cost is distributed over a larger OR class which includes those who can report negative test results.

Table 12: Percentage increases in premium rates for life insurance, using a moratorium on all genetic test results and adverse results respectively, under severe adverse selection, with annual rate of testing 0.04, for a market operating between ages 20 and 60.

Market Size	Rate of Purchase with A Family History	Moratorium on using	
		All test results %	Adverse results %
	Same as 'normal'	0.000393	0.000391
Large	Half as 'normal'	0.000590	0.000583
	Nil	0.000898	0.000879
Small	Nil	0.004655	0.004598

(c) It is natural that the more severe adverse selection will lead to higher costs; however, the effect is rather small, compared with the effect of market size.

We chose a rate of genetic testing of 0.02 *per annum*, corresponding to only modest demand. Table 12 shows the premium increases (for brevity, only under severe adverse selection) if the rate of testing is 0.04 *per annum*. Premiums increase a little, but the increases are still rather small.

6.4 *A Moratorium on Family History and Genetic Test Results*

Suppose insurers may not use any genetic information, including family history. All applicants will be charged the OR premium rate, which may increase for two reasons:

(a) Everyone can buy insurance on the same terms. Those who were previously in higher-risk underwriting classes and were offered a higher premium rate can now buy insurance at the same rate as those not at risk. Their insurance buying behaviour will change and, in the first instance, we suppose they become just as likely to buy insurance as anyone else, that is, at the normal rate. We do not call this adverse selection. In this case, the OR premium would rise just because of the creation of the new underwriting class.

(b) Those at risk might also buy insurance at a higher rate than normal, so there may be adverse selection as well, leading to a further increase in premiums. Here, people at risk include those with a family history who remain untested and those who have had a positive test result. We show both severe and moderate adverse selection for comparison.

The results are shown in Table 13. We see that the creation of the new underwriting class has a much larger effect than that of adverse selection. The effect of adverse selection is of a similar order of magnitude to that under the other moratoria, but the market size matters more; the premium increase in the small market is double that in the large market even under moderate adverse selection.

Table 13: Percentage increases in premium rates for life insurance, under a moratorium on genetic test results and family history, for a market operating between ages 20 and 60.

| | | Cost Arising From | |
Market Size	Change of Underwriting Classes %	Severe Adverse Selection %	Moderate Adverse Selection %
Large	0.477091	0.001677	0.000743
Small	0.431220	0.013602	0.001486

7. APPLICATION OF THE FAMILY HISTORY MODEL TO CI INSURANCE

7.1 *Critical Illness Insurance*

Critical illness insurance pays a lump sum on onset or diagnosis of any of a range of severe illnesses, such as cancer or heart attack, if they do not cause immediate death. The definition of 'immediate' is often within 28 days of diagnosis or onset. We refer to Dinani *et al.* (2000) for further details.

7.2 *CI Insurance Premium Ratings*

Figure 9 shows a model of a person's life history in a CI insurance market, in which:

(a) insurance purchase, development of family history and genetic tests are represented by states and intensities enclosed by the dashed box;
(b) CI claims could be triggered by: (i) breast cancer, (ii) ovarian cancer or (iii) other insured diseases given survival for at least 28 days;
(c) each state inside the dashed box has a transition into every state outside.

In previous sections, we estimated: (a) incidence rates of insurance purchase, developing a family history, and genetic testing in a life insurance market; and (b) onset rates of BC and OC, in the population and for BRCA1 or BRCA2 carriers. For other intensities we use the CI model proposed by Macdonald, Waters & Wekwete (2003b), described in Appendix B.

Since all intensities depend only on current age, the model is Markov and the numerical procedures are as in Section 5.2. Tables 14 and 15 give level net premiums for a CI policy as a percentage of the premium for standard risks, for mutation carriers and women with a family history, respectively. The premiums are calculated with estimated BC and OC intensities of 100% and 50% of those fitted to our data. In addition, Table 14 shows the premium rates for mutation carriers obtained by Macdonald, Waters & Wekwete (2003b). (Note that they gave premium rates as a percentage of an aggregate premium rate averaged over all genotypes. We have re-based their premiums using non-carriers for consistency with our figures.) We note that:
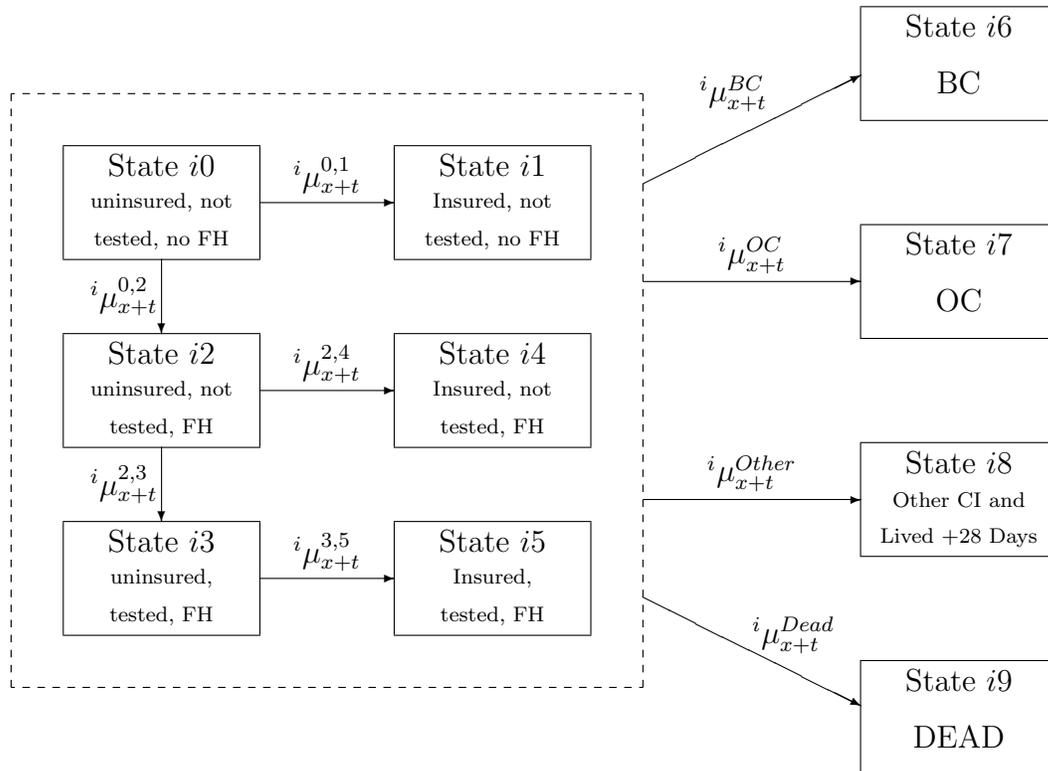
Figure 9: A Markov model of family history, genetic testing, insurance purchase and CI insurance events for a person in the ith risk subpopulation (FH = family history present).

(a) The relative premium increases are much higher than those for life insurance, especially for mutation carriers.

(b) Most mutation carriers are not insurable given that the typical limit for an offer of CI cover in practice is 300–350% of the standard rate.

(c) The effect of reducing the rates of onset (to allow for ascertainment bias) is considerable. All BRCA2 mutation carriers become insurable. Therefore, unless high intensities are well-founded based on studies free of ascertainment bias, cover could be extended to all BRCA2 mutation carriers.

(d) Policies could be issued to women with a family history for any age at entry and policy term, given the typical underwriting limit of 300–350% of the standard rate.

The premiums (relative to the standard rates) for mutation carriers obtained by Macdonald, Waters & Wekwete (2003b) are higher than ours, which is expected because since they used onset rates from a study believed to be highly susceptible to ascertainment bias. It is perhaps slightly surprising that the difference between our results and theirs is not uniformly greater.

We are unable to compare our premium rates for women with a family history with any from Macdonald, Waters & Wekwete (2003b), because they did not have the dynamic model of family history used here and had to use a different approach. (For an applicant at a fixed age, premiums were calculated for specimen family histories, involving second-degree relatives and family structures.) Indeed the fact that they could not properly reflect underwriting practice, especially when modelling adverse selection, motivated the dynamic model used here.

8. The Costs of Adverse Selection in a CI Insurance Market

We follow the same approach as in Section 6 to estimate the impact of adverse selection on the CI insurance market. The parameterisation of relevant parts of this model is the same as in Section 6.1.

8.1 *Moratoria on Genetic Test Results Alone*

Table 16 shows the percentage increases in premium rates for CI insurance under moratoria on genetic testing results but not family history. Compared with those for the life insurance model, see Table 11, the results are a little higher but otherwise similar.

8.2 *Moratorium on Family History and Genetic Test Results*

Table 17 shows the percentage increases in premium rates for CI insurance under moratoria on genetic testing results and family history. The results and comments are similar to those for life insurance. Note that without the dynamic model of family history Macdonald, Waters & Wekwete (2003b) could not estimate the effect of adverse selection under moratoria extending to family history.

Table 14: Level net premiums for CI cover as a percentage of the premium for standard risks, for persons with a known BRCA1 or BRCA2 mutation. For comparison, results from Macdonald, Waters & Wekwete (2003b) are reproduced in brackets. Excess BC/OC incidence rates are 100% and 50% of those observed.

Mutation	Incidence Rates	Age 20				Age 30			Age 40		Age 50
		10 Yrs	20 Yrs	30 Yrs	40 Yrs	10 Yrs	20 Yrs	30 Yrs	10 Yrs	20 Yrs	10 Yrs
		%	%	%	%	%	%	%	%	%	%
BRCA1	100%	977	1176	905	682	1347	967	725	842	654	532
						(1,843)	(1,554)	(1,121)	(1,610)	(1,123)	(747)
	50%	532	638	510	402	713	526	410	449	360	295
						(977)	(865)	(648)	(861)	(614)	(423)
BRCA2	100%	366	416	361	317	449	369	323	338	308	296
						(211)	(486)	(584)	(689)	(776)	(1,067)
	50%	225	246	215	195	259	215	195	195	184	177
						(155)	(300)	(384)	(400)	(476)	(601)

Table 15: Level net premiums for CI cover as a percentage of the premium for standard risks, for persons with a family history of BC/OC. Excess BC/OC incidence rates are 100% and 50% of those observed.

Rates	Age 20				Age 30			Age 40		Age 50
	10 Yrs	20 Yrs	30 Yrs	40 Yrs	10 Yrs	20 Yrs	30 Yrs	10 Yrs	20 Yrs	10 Yrs
	%	%	%	%	%	%	%	%	%	%
100%	278	312	254	209	330	251	204	208	174	142
50%	189	208	181	159	217	179	156	154	139	120

Table 16: Percentage increases in premium rates under a moratorium on all genetic testing results and adverse results respectively, for a CI insurance market operating between ages 20 and 60.

Adverse Selection	Market Size	Rate of Purchase with A Family History	Moratorium on using	
			All test results %	Adverse results %
Severe	Large	Same as 'normal'	0.000358	0.000353
		Half as 'normal'	0.000549	0.000541
		Nil	0.000861	0.000847
	Small	Nil	0.004503	0.004444
Moderate	Large	Same as 'normal'	0.000156	0.000154
		Half as 'normal'	0.000314	0.000310
		Nil	0.000582	0.000572
	Small	Nil	0.000883	0.000872

9. CONCLUSIONS

9.1 *The Family History Model*

Macdonald (2003) describes two modelling approaches in the actuarial literature on genetics and insurance. One is called 'top-down' (Macdonald, 1999) in which a simple model of all genetic diseases is used, with no attempt to model specific mutations. These models might be useful if they can show that, even under extremely adverse assumptions, the cost of the adverse selection would be negligible. But when adverse selection might not be negligible, the 'top-down' approach is not suitable, and a 'bottom-up' approach based on particular diseases is needed. Macdonald, Waters, & Wekwete (2003b) proposed such a model for CI insurance in respect of BC and OC. However, they defined family history in a static way, by way of specific histories presented by a woman applying for insurance. Also, they considered affected aunts as well as first-degree relatives, which is close to clinical practice but less usual in underwriting practice. The chief drawbacks were:

(a) Family histories could not be incorporated into their model of adverse selection, in respect of either: (i) adverse selection arising from genetic tests when family history can be used in underwriting; or (ii) a moratorium extending to family history. Yet subsequent developments in several countries, including Sweden and the UK, have meant that the use, or not, of family histories is the key question.

(b) Later work showed that it is possible to represent family history quite easily in models of diseases with exclusively genetic causes (Gutiérrez & Macdonald, 2004) but not in models of common diseases with rare genetic variants, such as BC and OC.

In view of the above, we have specified here a model of family history that is simpler, because it is based on underwriting practice and uses first-degree relatives only, but also

Table 17: Percentage increases in premium rates under a moratorium on genetic testing results and family history, for a CI insurance market operating between ages 20 and 60.

Market Size	Cost From Change of Underwriting Classes %	Cost From Severe Selection %	Cost From Moderate Selection %
Large	0.733737	0.003149	0.001434
Small	0.678819	0.030161	0.003408

dynamic, because the appearance of a family history is an event in the applicant's life history.

9.2 *Premium Ratings*

We reach different conclusions about premium ratings for life insurance and CI insurance:

(a) The (relative) premium increases for CI insurance are much bigger than those for life insurance, and in most cases known mutation carriers would not be insurable.

(b) In respect of life insurance, most mutation carriers are insurable given the limit of 400–500% of the standard rate in practice, with a few exceptions (BRCA1 mutation carriers given the highest onset rates). In contrast, for CI insurance most mutation carriers are not insurable given the limit for CI cover of 300–350% of the standard rate.

(c) All cases for women with a family history for life (or CI) insurance are insurable. Especially for those at ages over 50, extra premiums of less than 50% could be offered.

9.3 *Reduced Onset Rates to Allow for Ascertainment Bias*

We fitted onset rates of BC or OC for mutation carriers based on Antoniou *et al.* (2003), which used data from 22 studies that had tested patients with BC or OC who were unselected for family history of BRCA1 or BRCA2 mutation. The benefit (compared with earlier studies) was that these estimates should be less vulnerable to ascertainment bias and hence closer to the true onset rates among randomly selected mutation carriers. However, the prevalence of mutations in unselected series of cases is so low that very large numbers of cases are needed. Also, we should note that the underlying families were still selected on the basis of one affected index patient; these were not truly prospective studies, so ascertainment bias may still exist but, we hope, to a smaller extent. Therefore, it is reasonable to reduce the estimates to make allowance for the possible bias. We reduce the rates of onset of BC and OC for mutation carriers fitted to the rates in Antoniou *et al.* (2003) by 50%. The effect of these lower onset rates is substantial:

(a) Extra premiums could decrease by a half or more in most cases, especially where the genetic risks are highest.

(b) All cases for life insurance become insurable and most cases for CI insurance become insurable with some exceptions for BRCA1 mutation carriers.

9.4 *The Impact of Adverse Selection*

The costs (premium increases) to recoup the impact even of severe adverse selection are very small, in most circumstances less than 0.01%. However, we cannot say that it can be entirely neglected, because BRCA1 and BRCA2 mutations are rare. The largest premium increase arises when severe adverse selection happens in a small market, but compared with the effect of imposing a moratorium on the use of family history, it is still small.

Acknowledgements

References

Anglian Breast Cancer Study Group and Pharoah, P. (2000). Prevalence and penetrance of BRCA1 and BRCA2 in a population based series of breast cancer cases. *British Journal of Cancer*, **83**, 1301–1308.

Antoniou, A.C., Gayther, S.A., Stratton, J.F., Ponder, B.J. & Easton, D.F. (2000). Risk models for familial breast and ovarian cancer. *Genetic Epidemiology*, **18**, 173–190.

Antoniou, A.C., Pharoah, P.P.D., Mcmullan, G., Day, N.E., Stratton, M.R., Peto, J., Ponder, B.J. & Easton, D.F. (2002). A comprehensive model for familial breast cancer incorporating BRCA1, BRCA2 and other genes. *British Journal of Cancer*, **86**, 76–83.

Antoniou, A.C., Pharoah, P.P.D., Narod, S., Risch, H. A., Eyfjord, J. E., Hopper, J. L., Loman, N., Olsson, H., Johannsson, O., Borg, A., Pasini, B., Radice, P., Manoukian, S., Eccles, D. M., Tang, N., Olah, E., Anton-Culver H., Warner, E., Lubinski, J., Gronwald, J., Gorski, B., Tulinius, H., Thorlacius, S., Eerola, H., Nevanlinna, H., Syrjäkoski, K., Kallioniemi, O.-P., Thompson, D., Evans, C., Peto, J., Lalloo, F., Evans, D. G., & Easton, D.F. (2003). Average risks of breast and ovarian cancer associated with mutations in BRCA1 or BRCA2 detected in case series unselected for family history: A combined analysis of 22 studies. *American Journal of Human Genetics*, **72**, 1117–1130.

Claus, E.B., Risch, N.J. & Thompson, W.D. (1994). Autosomal dominant inheritance of early-onset breast cancer. Implications for risk prediction. *Cancer*, **73**, 643–651.

Claus, E.B., Schildkraut, J.M., Thompson, W.D. & Risch, N.J. (1996). The genetic attributable risk of breast and ovarian cancer. *Cancer*, **77**, 2318–2324.

Coleman, M., Babb, P., Damiecki, P., Grosclaude, P.C., Honjo, S., Jones, J., Knerer, G., Pitard, A., Quinn, M.J., Sloggett, A. & De Stavola, B.L. (1999). Cancer survival trends for England and Wales, 1971-1995: Deprivation and NHS Region. *Office of National Statistics, London*.

Dinani, A., Grimshaw, D., Robjohns, N., Somerville, A.S. & Staffurth, J. (2000). *A critical review: report of the critical illness healthcare study group*. Presented to the Staple Inn Actuarial Society, London, on 14 March 2000.

EASTON, D.F., FORD, D. & BISHOP, D.T. (1995). Breast and ovarian cancer incidence in BRCA1-mutation carriers. *American Journal of Human Genetics*, **56**, 265–271.

EASTON, D.F., STEELE, L., FIELDS, P., ORMISTON, W., AVERILL, D., DALY, P.A., MC-MANUS, R., NEUHAUSEN, S.L., FORD, D., WOOSTER, R., CANNON-ALBRIGHT, L.A., STRATTON, M.R. & GOLDGAR, D.E. (1997). Cancer risks in two large breast cancer families linked to BRCA2 on chromosome 13q12-13. *American Journal of Human Genetics*, **61**, 120–128.

FORD, D., EASTON, D.F., BISHOP, D.T., NAROD, S.A. & GOLDGAR, D.E. (1994). Risks of cancer in BRCA1 mutation carriers. *The Lancet*, **343**, 692–695.

FORD, D. & EASTON, D.F. (1995). The genetics of breast and ovarian cancer. *British Journal of Cancer*, **72**, 805–812.

FORD, D., EASTON, D.F., STRATTON, M. NAROD, S., GOLDGAR, D., DEVILEE, P., BISHOP, D.T., WEBER, B., LENOIR, G., CHANG-CLAUDE, J., SOBOL, H., TEARE, M.D., STRUEWING, J., ARASON, A., SCHERNECK, S., PETO, J., REBBECK, T.R., TONIN, P., NEUHAUSEN, S., BARKARDOTTIR, R., EYFJORD, J., LYNCH, H., PONDER, B.A., GAYTHER, S.A., ZELADA-HEDMAN, M. AND THE BREAST CANCER LINKAGE CONSORTIUM (1998). Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families. *American Journal of Human Genetics*, **62**, 676–89.

FUTREAL, P.A., LIU, Q., SHATTUCK-ELDENS, D., COCHRAN, C., HARSHMAN, K., TAVTIGIAN, S., BENNETT, L.M., HAUGEN-STRANO, A., SWENSEN, J. & MIKI, Y. (1994). BRCA1 mutations in primary breast and ovarian carcinomas. *Science*, **266**, 120–122.

GUTIÉRREZ, M.C. & MACDONALD, A.S. (2004). Huntington's disease, critical illness insurance and life insurance. *Scandinavian Actuarial Journal*, **2004**, 279–313.

HOULSTON, R.S., COLLINS, A., SLACK, J., CAMPBELL, S., COLLINS, W.P., WHITEHEAD, M.I. & MORTON, N.E. (1991). Genetic epidemiology of ovarian cancer: segregation analysis. *Annals of Human Genetics*, **55**, 291–9.

MACDONALD, A.S. (1999). Modelling the impact of genetics on insurance. *North American Actuarial Journal*, **3(1)**, 83-101.

MACDONALD, A.S. (2003). Genetics and insurance: what we have learned so far?. *Scandinavian Actuarial Journal*, **2003**, 324–348.

MACDONALD, A.S., WATERS, H.R. & WEKWETE, C.T. (2003a). The genetics of breast and ovarian cancer I: A model of family history. *Scandinavian Actuarial Journal*, **2003**, 1–27.

MACDONALD, A.S., WATERS, H.R. & WEKWETE, C.T. (2003b). The genetics of breast and ovarian cancer II: A model of critical illness insurance. *Scandinavian Actuarial Journal*, **2003**, 28–50.

MIKI, Y., SWENSEN, J., SHATTUCK-EIDENS, D., FUTREAL, P. A., HARSHMAN, K., TAVTIGIAN, S., LIU, Q., COCHRAN, C., BENNETT, L. M., DING, W., BELL, R., ROSENTHAL, J., HUSSEY, C., TRAN, T., MCCLURE, M., FRYE, C., HATTIER, T., PHELPS, R., HAUGENSTRANO, A., KATCHER, H., YAKOMO, K., GHOLAMI, Z., SHAFFER, D., STONE, S., BAYER, S., WRAY, C., BOGDEN, R., DAYANANTH, P., WARD, J., TONIN, P., NAROD, S., BRISTOW, P.K., NORRIS, F.H., HELVERING, L., MORRISON, P., ROSTECK, P., LAI, M., BARRETT, J.C., LEWIS, C., NEUHAUSEN, S., CANNONALBRIGHT, L., GOLDGAR, D., WISEMAN, R., KAMB, A. & SKOLNICK, M.H. (1994). A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. *Science*, **266**, 61–71.

PARMIGIANI, G., BERRY, D. & AGUILAR, O. (1998). Determining carrier probabilities for breast cancer susceptibility genes BRCA1 and BRCA2. *American Journal of Human Genetics*, **62**, 145–158.

SOUHAMI, T. & TOBIAS, J. (1998). *Cancer and its management.* Blackwell Science.3rd edition

SUBRAMANIAN, K., LEMAIRE, J., HERSHEY, J.C., PAULY, M.V., ARMSTRONG, K. & ASCH, D.A. (1999). Estimating adverse selection costs from genetic testing for breast and ovarian cancer: The case of life insurance. *Journal of Risk and Insurance*, **66**, 531–550.

WOOSTER, R., BIGNELL, G., LANCASTER, J., SWIFT, S., SEAL, S., MANGION, J., COLLINS, N., GREGORY, S., GUMBS, C. & MICKLEM, G. (1995). Identification of the breast cancer susceptibility gene BRCA2. *Nature*, **378**, 789–792.

APPENDIX A

POST-ONSET MORTALITY RATES

From Coleman *et al.* (1999) we have 110,697 BC cases in 1986–1990 and 45,650 OC cases in 1980–1990. Defining 'age' as age nearest birthday at diagnosis, we follow up each case until the earliest of death, 31 December 1995 or exit from observation for any other reason. We have, for each year after diagnosis:

(a) $E_{x,d}$, the exact time spent under observation between integer durations d and $d+1$ (curtate duration d) by lives aged x nearest birthday; and

(b) $A_{x,d}$, the number of lives aged x nearest birthday who die at curtate duration d

from which we compute crude forces of mortality: $\hat{\mu}_{x,d} = A_{x,d}/E_{x,d}$. Assuming Poisson errors, we fit the following GLMs to each duration separately up to 6 years, and then to 6 years and over.

$0 \leq \text{Duration} \leq 1$

$$\mu_{x,d}^{BC} = 2.00266 - 0.1507811x + 0.004264272x^2 - 5.27552 \times 10^{-5}x^3 + 2.456224 \times 10^{-7}x^4$$

$$\mu_{x,d}^{OC} = exp(-2.71394 + 0.023657x + 0.1960156 \times 10^{-3}x^2)$$

$1 \leq \text{Duration} \leq 2$

$$\mu_{x,d}^{BC} = exp(0.6037712037 - 0.015001751x + 0.1111315 \times 10^{-3}x^2)$$
$$\times (0.0474102 + 0.307835 \times 10^{-3}x + exp(3.06993 - 0.284105x + 0.00266558x^2))$$

$$\mu_{x,d}^{OC} = exp(-9.6097147 + 0.3634146x - 0.005149204x^2) + 0.2471276 \times 10^{-4}x^3$$

$2 \leq \text{Duration} \leq 3$

$$
\begin{aligned}
\mu_{x,d}^{BC} &= exp(0.4971436820 - 0.015001751x + 0.1111315 \times 10^{-3}x^2) \\
&\quad \times (0.0474102 + 0.307835 \times 10^{-3}x + exp(3.06993 - 0.284105x + 0.00266558x^2)) \\
\mu_{x,d}^{OC} &= exp(-10.1965014 + 0.3634146x - 0.005149204x^2) + 0.2471276 \times 10^{-4}x^3
\end{aligned}
$$

$3 \leq \text{Duration} \leq 4$

$$
\begin{aligned}
\mu_{x,d}^{BC} &= exp(0.3905161603 - 0.015001751x + 0.1111315 \times 10^{-3}x^2) \\
&\quad \times (0.0474102 + 0.307835 \times 10^{-3}x + exp(3.06993 - 0.284105x + 0.00266558x^2)) \\
\mu_{x,d}^{OC} &= exp(-13.4719011 + 0.52647732x - 0.008227498x^2) + 0.4354431 \times 10^{-4}x^3
\end{aligned}
$$

$4 \leq \text{Duration} \leq 5$

$$
\begin{aligned}
\mu_{x,d}^{BC} &= exp(0.352482834 - 0.003144911x) \\
&\quad \times (0.02902753 + exp(-0.1624326x + 0.00164027x^2)) \\
\mu_{x,d}^{OC} &= exp(-14.1632748 + 0.52647732x - 0.008227498x^2) + 0.4354431 \times 10^{-4}x^3
\end{aligned}
$$

$5 \leq \text{Duration} \leq 6$

$$
\begin{aligned}
\mu_{x,d}^{BC} &= exp(0.082950962 - 0.000880887x) \\
&\quad \times (0.02902753 + exp(-0.1624326x + 0.00164027x^2)) \\
\mu_{x,d}^{OC} &= exp(-14.8546485 + 0.5384382x - 0.008227498x^2) + 0.4354431 \times 10^{-4}x^3
\end{aligned}
$$

$6 \leq \text{Duration}$

$$
\begin{aligned}
\mu_{x,d}^{BC} &= exp(-0.18658091 + 0.001383137x) \\
&\quad \times (0.02902753 + exp(-0.1624326x + 0.00164027x^2)) \\
\mu_{x,d}^{OC} &= exp(-15.5460222 + 0.54441864x - 0.008227498x^2) + 0.4354431 \times 10^{-4}x^3
\end{aligned}
$$

APPENDIX B

THE CRITICAL ILLNESS INSURANCE MODEL

Macdonald, Waters & Wekwete (2003b) proposed the following CI insurance model for females, including as major illnesses: (a) cancers except BC, OC OC and non-malignant skin cancer; (b) heart attack; and (c) stroke, allowing for survival for at least 28 days. In the following, x denotes age.

(a) *Cancers Other Than BC and OC:*

$$
\begin{aligned}
\mu_x^{cancer} &= exp(-10.3995 + 0.08235x) \qquad \text{for } x < 40 \\
\mu_x^{cancer} &= 0.00808 - 0.00019x \\
&\quad + 0.000016(x - 35)^2 - 0.000000144(x - 35)^3 \qquad \text{for } x > 64
\end{aligned}
$$

with linear interpolation between ages 40 and 64.

(b) *Stroke:*

$$\mu_x^{stroke} = \exp(-11.45 + 0.085x).$$

(c) *Heart Attack:*

$$\mu_x^{heart} = 0.58 \left(\frac{0.16^{16.34} \exp(-0.16x)x^{15.34}}{\Gamma(16.34)} \right).$$

(d) *The 28-day Survival Probabilities:* The 28-day survival probability is irrelevant for cancers, but is given below for strokes and heart attacks . The intensities of strokes and heart attacks as CI claim events should be reduced and mortality rates should be increased accordingly.

$$
\begin{aligned}
p_x^{heart} &= 0.8983095 - 0.00235911x - 0.00001359781x^2 \\
p_x^{stroke} &= 0.8718412 + 0.001566578x - 0.00003711161x^2.
\end{aligned}
$$

(e) *Total Onset Rate of CIs other than BC and OC*: Other minor causes account for about 15% of CI claims, therefore:

$$\mu_x^{total} = 1.15(\mu_x^{cancer} + p_x^{stroke}\mu_x^{stroke} + p_x^{heart}\mu_x^{heart}) + 0.15(\mu_x^{BC} + \mu_x^{OC}).$$

(f) *Mortality Rate Excluding CI Events*: The mortality rate is that of English Life Tables No. 15 (Females) (ELT15F) modified to remove deaths caused by conditions that would lead to a CI claim. The ratio θ_x of the number of deaths from these causes to the total number of deaths is:

$$
\begin{aligned}
\theta_x &= -0.02612913 + 0.1046405x \\
&\quad -0.01181445x^2 + 0.0004671351x^3 - 0.000005790098x^4 \quad \text{for } x < 35 \\
\theta_x &= -1.345136 + 0.08972161x - 0.001199781x^2 + 0.000004867845x^3 \quad \text{for } x \geq 35.
\end{aligned}
$$