A DYNAMIC FAMILY HISTORY MODEL OF HEREDITARY NONPOLYPOSIS COLORECTAL CANCER AND CRITICAL ILLNESS INSURANCE

By Li Lu, Angus Macdonald, Howard Waters and Fei Yu

ABSTRACT

Hereditary Nonpolyposis Colorectal Cancer (HNPCC) is characterised by the familial aggregation of cancer of the colon and rectum (CRC). It may be caused by any of five mutations in DNA mismatch repair (MMR) genes or by non-genetic factors, such as life style. However, it accounts for only about 2% of CRC, which is a very common cancer. Previous actuarial models, of diseases with only genetic causes, assumed that a family history of the disease shows mutations to be present, but this is not true of HNPCC. This is a significant limitation, since the best information available to an underwriter (especially if the use of genetic test results is banned) is likely to be knowledge of a family history of CRC. We present a Markov model of CRC and HNPCC, which includes the presence of a family history of CRC as a state, and estimate its intensities allowing for MMR genotype. Using this we find the MMR mutation probabilities for an insurance applicant with a family history of CRC. Our model greatly simplifies the intensive computational burden of finding such probabilities by integrating over complex models of hidden family structure. We estimate the costs of critical illness insurance given the applicant's genotype or the presence of a family history. We then consider what the cost of adverse selection might be, if insurers are unable to use genetic tests or family history information. We also consider the effect of using alternative definitions of a family history in underwriting.

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

1.1 Genetic Testing, Family History and Insurance

Before genetic testing was developed, the only basis for suspecting that a disease might be genetic in origin was a tendency of cases to cluster within families. Epidemiologists and insurance underwriters were both interested in such 'family histories', but neither had any more evidence. Now, genetic testing has identified many of the genes in which errors, called mutations, may cause disease. The simplest cases arise when a mutation in a single gene can cause disease, called the 'single-gene disorders'. We may divide the single-gene disorders into two groups, which are significantly different for actuarial purposes:

(a) Some are purely genetic in origin; the disease has no known cause except defects in particular genes. Huntington's disease (HD) and adult polycystic kidney disease (APKD) are examples.
(b) Others are common diseases, of which some but not all cases are linked to mutations in single genes. Most important are breast cancer (BC) and ovarian cancer (OC), where 5–10% of cases are caused by mutations in the BRCA1 and BRCA2 genes, and colorectal cancer (CRC), where an inherited form called hereditary nonpolyposis colorectal cancer (HNPCC) is responsible for up to 5% of cases (see Section 2.3).

The difference matters because regulators or governments in several jurisdictions have limited the use that insurers may make of genetic information. These moratoria (we will call them such, even if the legal form they take is an outright ban) nearly all forbid the use of adverse genetic test results, sometimes with exceptions for very large policies. Therefore the question: "what is the appropriate premium rate for a known carrier of a given mutation?" is only applicable, in practice, in these exceptional cases. Moratoria in some countries (for example Sweden) also ban the use of family histories, but in others (for example the UK) insurers are free to use family histories. Thus, the most important practical question is: "what is the appropriate premium rate for a person with a given family history of a genetic disorder?". A necessary step in answering this is to estimate the probabilities:

$$P[\text{ Applicant carries a causative mutation} \mid \text{Family history}]. \tag{1}$$

This is simplest if the disease is purely genetic in origin, for all the family members who have the disease *must* carry the mutation, and Mendel's laws let us find the above probabilities, usually quite easily. We may reasonably assume that 'family history' means an affected parent (this covers most cases), and partition the population into three groups:
(a) Persons with an affected parent, who have themselves inherited the mutation.
(b) Persons with an affected parent, who have themselves *not* inherited the mutation.
(c) Persons with no family history, who we assume are not at risk of the disease.

The combination of rarity and dominant inheritance means that if the mutation affects a proportion p of the population, the proportions in these three groups at birth are p, p and $1 - 2p$, respectively, from which initial condition we can find the proportions when the applicant is any age x. Thus we can weight the expected present values (EPVs) of insurance cashflows according to the probabilities (1). This model has been used to study HD (Gutiérrez and Macdonald, 2004), APKD (Gutiérrez and Macdonald, 2003) and early-onset Alzheimer's disease (Gui and Macdonald, 2002).

But if a small subset of a common disease originates in single-gene defects, affected family members need not carry a causative mutation, and a family history could arise wholly or partly by chance. (Cases in the absence of defective genes are called 'phenocopies' or 'sporadic cases'.) Then the population may not be partitioned as above, and the probabilities (1) are not simple to find. Moreover, the following complications arise.
(a) Mutations may not cause disease in all who carry them (called 'incomplete penetrance').

(b) The same disease may be caused by defects in any of several genes. For example, breast cancer may be caused by defects in the BRCA1 or BRCA2 genes.
(c) Defects in a gene may cause more than one disease. For example, mutations in BRCA1 can cause breast or ovarian cancer.
(d) Defects at different points in the same gene may present different risks of the disease.

Previous researchers have modelled BC/OC (Lemaire *et al.*, 2000; Subramanian *et al.*, 1999). In particular Macdonald *et al.* (2003a,b) calculated the probability that a BRCA1 or BRCA2 mutation was present in a woman applying for critical illness (CI) insurance, based on a specific family history of BC/OC at the time of application. The family histories modelled included affected mother, aunts and sisters. This model was limited in some respects however:
(a) The family histories were more typical of those collected by clinicians than by underwriters. An insurance proposal form usually asks only about first-degree relatives (FDRs); that is, parents and siblings. Obtaining accurate information is not trivial.
(b) The family histories were static, not dynamic, in nature. For example, suppose a woman age x applied for insurance having two affected sisters. The development of the family history over time was ignored, so it was not known if, say, only one sister had been affected when the applicant was age $x - 1$, and it was the recent onset in respect of her second sister that had influenced her decision to buy insurance. Thus the model was limited in its approach to measuring adverse selection.
(c) The approach was computationally very intensive, as all unknown factors about the family had to be integrated out in a high-dimensional application of Bayes' theorem.

This paper has two main aims. First, to parameterise a model of the onset of CRC in general, and HNPCC in particular. This is the first actuarial study of HNPCC, one of the most common single-gene disorders. Second, to introduce a new dynamic model of family history which is not limited to purely genetic disorders, and will allow family history underwriting and adverse selection to be modelled in a simple way, similar to that described above in respect of purely genetic disorders. In another paper we will apply the same model to BC and OC.

1.2 *Colorectal Cancer and HNPCC*

In the UK, CRC is the second commonest cancer affecting women (after BC) and the third most common cancer in men (after prostate and lung cancer); more than 35,000 people are diagnosed with CRC every year. It can be sporadic or hereditary, and the hereditary forms, of which HNPCC is one, together account for about 10% of cases.

The 'nonpolyposis' in HNPCC means that there are few if any polyps (small precancerous growths) in the colon. This is also true of most sporadic CRCs, so HNPCC is not easily distinguished from sporadic CRC by clinical signs alone. It is caused by mutations in any one of the DNA mismatch repair (MMR) genes MLH1, MSH2 MSH6, PMS1 and PMS2. These mutations may also cause cancer of the endometrium and other organs.

In Section 2, we describe the genetic basis of HNPCC and the rôle of the MMR genes. A model of HNPCC is presented in Section 3 and transition intensities and CI premium rates (for known mutation carriers) are estimated in Sections 4 and 5, respectively.

1.3 *A Dynamic Family History Model*

Suppose a person X has a father with HD. Then no other information, except a genetic test result, will change the probability that that X carries a HD mutation. In effect, we make the simplification:

$$\text{Family history} \; = \; \text{Affected parent.}$$

However, if X's father has had CRC, it might be hereditary or sporadic. Then other information does matter; for example, whether X has a sister who has had CRC, or has ten older brothers who have not had CRC. We cannot make the simplification above. But we note that it amounts to turning the presence of a family history into a question with a 'yes' or 'no' answer: is either parent affected or not? Insurance underwriters to some extent do the same, by setting thresholds below which a family history is deemed unimportant, and above which it matters. In the example before, a woman had one affected sister at age $x - 1$, but by age x she had two. If the relevant threshold for underwriting was 'two or more affected FDRs', then she crossed that threshold between age $x - 1$ and age x. This suggests that, given an explicit definition of family history, we can model its appearance as an event, or a transition between states, along with all the other transitions in the model. Then, we partition the population in the following way:
(a) Persons in a family in which the mutation is present, who do themselves carry it.
(b) Persons in a family in which the mutation is present, who do not themselves carry it.
(c) Persons in a family in which the mutation is not present, who cannot carry it.

This is not the same as the partition mentioned in Section 1.1. There, the presence or absence of a family history defined the partition. Here, a family history may appear — the underwriting threshold may be crossed — in any of the three groups, it is the presence or absence of mutations, in the family and in the applicant, that determines the partition. Each person's actual risk depends solely on whether or not they carry the mutation, therefore on the group they are in, as before. But now their risk of developing a family history also depends on which group they are in, in a dynamic way.

To implement the model, we need to estimate the probability that a family history has appeared by any age x, in each group; in other words, the distribution of the random time until onset of family history. This requires: a model of onset of HNPCC; a model of family structure; and an underwriting criterion for a family history to have developed. From the distribution, we obtain a transition intensity between states 'No Family History' and 'Family History' in a Markov or semi-Markov model of a person's life history. This is described in Sections 6 and 7. The chief problem is that the insurance underwriter usually does not know about unaffected family members, so as in Macdonald *et al.* (2003a,b) we use the model of family structure to integrate that information out. The computational burden of this approach was mentioned above, but here it matters less because it need only be done once in order to find the intensity of onset of a family history.

With the paremterised model, we find CI insurance premiums based on a family history of CRC in Section 8, and estimate the cost of adverse selection in a market with moratoria in Section 9. Our conclusions are in Section 10.

2. The Genetic Basis of HNPCC

2.1 *Defining HNPCC Clinically*

Before the genetic basis of HNPCC was identified, the disease was recognised by familial aggregation of CRCs that had an early age of onset, an excess of primary tumours (often multiple) located in the proximal colon, and an excess occurrence of cancers in certain other organs, in particular the endometrium. See Warthin (1913); Lynch and Krush (1971); Boland and Troncale (1984) for a brief history of the diagnosis of HNPCC.

Criteria to provide uniform diagnoses were developed by the International Collaborative Group on HNPCC in Amsterdam in 1991 (Vasen *et al.*, 1991). The Amsterdam criteria have three major requirements:

(a) histologically verified CRC in three or more relatives, one of whom is a first-degree relative of the other two;

(b) CRC in at least two generations; and

(c) one or more CRC cases diagnosed before age 50.

Recent developments suggest that the Amsterdam criteria are too restrictive, and underestimate the prevalence of HNPCC. They ignore extracolonic cancers, and small families rarely meet them. Improved criteria have been proposed, such as the modified Amsterdam criteria (Bellacosa *et al.*, 1996) and Bethesda criteria (Rodriguez-Bigas *et al.*, 1997), see also Mitchell *et al.* (2002). Note, however, that even the original Amsterdam criteria demand more information than is usually obtained by an insurance underwriter.

The average age at diagnosis of first CRC in HNPCC patients is 44 years. Lynch and Smyrk (1996) suggested that a patient younger than 45 years should prompt a careful perusal of family history and probably genetic testing.

2.2 *HNPCC Genes*

Here we summarise briefly the detailed account given in Lu (2006). By 1993, linkage studies had identified HNPCC genes on chromosomes 2 and 3, which led Ionov *et al.* (1993) and Thibodeau *et al.* (1993) to observe that some tumours have microsatellite instability, a form of error in DNA replication not present in normal tissue. Peltomaki *et al.* (1993), Aaltonen *et al.* (1993) and later Liu *et al.* (1996) found that most CRCs from HNPCC patients showed microsatellite instability. This suggested that HNPCC might be caused by mutations in DNA mismatch repair (MMR) genes responsible for correcting faults in DNA replication.

Much evidence proved that more than one gene contributes to DNA mismatch repair and the HNPCC syndrome could be associated with mutations of any one of them. Mutations of five genes, namely MSH2, MLH1, PMS1, PMS2, MSH6, have so far been identified to be responsible for most cases of HNPCC.

Microsatellite instability is seen in sporadic CRC as well but only accounts for 10%–15% of cases (Lynch and Smyrk, 1996; Marra and Boland, 1995). Risinger *et al.* (1993) found microsatellite instability in 17% of sporadic endometrial cancers and endometrial cancers from HNPCC kindreds.

2.3 *The Frequency of HNPCC and of MMR Mutations*

We can estimate the frequency of HNPCC by combining estimates of the prevalence of CRC and the proportion of CRC due to HNPCC. Lu (2006) gives details of many studies into the frequency of HNPCC, which we omit for brevity. We have chosen to follow Aaltonen *et al.* (1998), who found that 2% of all CRC patients in Finland have HNPCC mutations. The Finnish Cancer Registry also found the risk of CRC by age 85 is 5% in the general population. This suggests that the population frequency of HNPCC is 1/1000, which we take as our assumption.

MLH1 and MSH2 mutations account for 90% of all HNPCC cases (Lynch and de la Chapelle, 2003), and we assume: (a) that they have roughly equal frequencies (consistent with the results of Vasen *et al.* (2001) that we use later); and (b) that the other minor MMR mutations have similar prognoses to those of MLH1 and MSH2 mutations. The penetrance of CRC in respect of MLH1 and MSH2 mutation carriers is estimated at about 65% and 70% by age 75 years, therefore we model all HNPCC mutations approximately by assuming that the population frequencies of these two mutations are $0.0005/0.65 = 0.000769$ and $0.0005/0.7 = 0.000714$ respectively.

3. A Model for HNPCC and Critical Illness Insurance

We represent the life history of a critical illness (CI) policyholder by the Markov model in Figure 1. The onset of an HNPCC associated cancer (CRC, endometrial cancer (EC, females only), or other extracolonic cancers) will trigger payment of the benefit. The policyholder is also at risk of other critical illnesses which include heart attack, stroke and other cancers not associated with HNPCC. We use g to denote genotype, and the transition intensity μ_x^{gjk} represents the transition rate between states gj and gk at age x, in respect of a person with genotype g.

Let $_tp_x^{gjk}$ denote the probability of being in state gk at age $x+t$ conditional on being in state gj at age x (we call this a 'transition probability'), and let OP_x^{gj} be the probability of being in state gj at age x (we call this an 'occupancy probability'). We obtain these probabilities by solving numerically the Kolmogorov forward equations.

Also, let $_tV_x^{gj}$ be the expected present value (EPV) of the future discounted loss at age $x+t$, conditional on being in state gj at age x (that is, the prospective policy value). We obtain these by solving numerically Thiele's differential equations.

See Hoem (1988) for details of these differential equations. To solve them, we used a Runge-Kutta algorithm with step-size 0.0005 years, and (in Thiele's equations) a force of interest 0.05 per year.

4. Estimating Onset Rates of HNPCC Associated Cancers

4.1 *The Data*

In two related papers, Vasen *et al.* (1996, 2001) studied the onset of HNPCC based on data collected from the Netherlands HNPCC registry. From January to July 2000, 193 families were registered, of which 116 met the original or modified Amsterdam criteria while 77 were suspected of HNPCC. Cancer diagnosis was verified by medical or pathologic reports. Mutation analysis was performed in respect of the members of all these families.
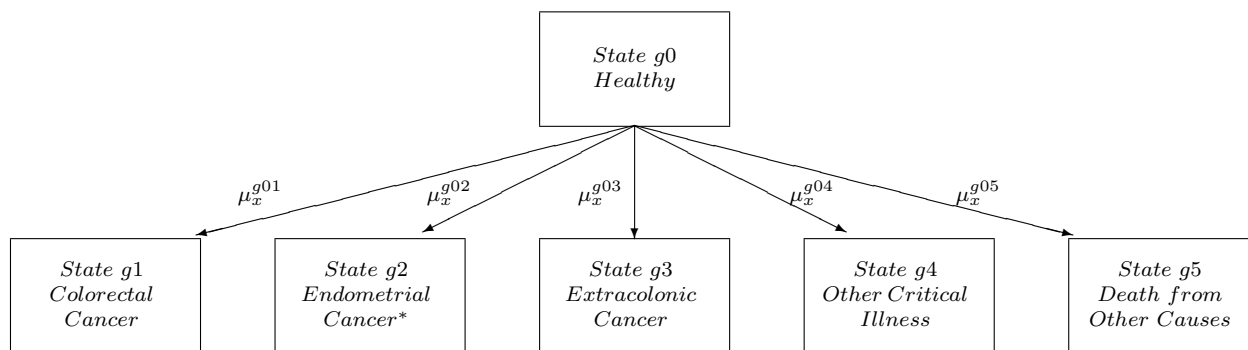
Figure 1: Model 1: A model of HNPCC and critical illness insurance, given exact genotype labelled g. (*Females only.)

Besides these Dutch families, 58 Norwegian families, also suspected of HNPCC, were investigated by mutation analysis.

Based upon the mutation analysis, the families with various MMR gene mutation were distinguished from families with no such mutation, but which met the original or modified Amsterdam criteria. Three groups of individuals were regarded as mutation carriers:

(a) putative carriers, who had CRC or EC, but excluding proven non-carriers;
(b) relatives identified as carrying a mutation; and
(c) obligate carriers, based on their position in the pedigree.

Table 1 shows the cumulative risks of CRC and EC, calculated using Kaplan-Meier methods, extracted from the graphs in Vasen *et al.* (2001), and Table 2 shows the cumulative risks of other cancers cited by Vasen *et al.* (2001).

4.2 *Onset of Colorectal Cancer*

The age-related cumulative risk (penetrance), denoted $F(x)$, can be regarded as the probability of developing a cancer by age x, if we assume there were no other causes of decrement. Given estimates of these probabilities, we fit smooth functions to them, or their logit transformation, by unweighted least squares. The rate of onset (hazard rate), $\mu(x)$, is then given by:

$$\mu(x) = \frac{F'(x)}{1 - F(x)} \tag{2}$$

and can be computed by analytical or numerical differentiation.

We fitted a polynomial function to the logit transforms of the non-zero values in each column in Table 1. The fitted functions were extrapolated to the ages where values were zero or unavailable. Equations (3) and (4) give the fitted penetrances of CRC, for male and female MLH1 mutation carriers respectively and Equations (5) and (6) for male and female MSH2 mutation carriers. Figure 2 illustrates the observed and fitted probabilities and the hazard rates derived from them.

Table 1: Cumulative risks (%) of developing colorectal cancer (CRC) or endometrial cancer (EC, females only) by age for carriers of MLH1 or MSH2 mutations, given sex and type of mutation. (Source: Vasen *et al.* (2001).)

	MLH1 Mutation			MSH2 Mutation		
Sex	M	F	F	M	F	F
Cancer	CRC	CRC	EC	CRC	CRC	EC
Age	%	%	%	%	%	%
20	0.0	0.0	0.0	0.0	0.0	0.0
25	0.0	0.0	0.0	2.4	0.7	0.7
30	6.9	5.7	0.0	6.0	1.9	1.2
35	14.9	11.2	0.0	13.6	2.6	2.7
40	20.0	16.9	2.4	28.6	11.7	5.0
45	29.2	26.3	5.4	40.5	12.6	8.9
50	43.1	30.4	11.6	56.1	34.1	20.0
55	51.0	40.0	19.8	63.3	47.1	24.5
60	55.9	45.1	24.6	70.5	49.1	35.0
65	55.9	51.2	24.6	73.2	51.5	37.3
70	64.9	54.1	24.6	73.2	54.1	37.3

Table 2: Cumulative risks (%) of developing other extracolonic cancers by age (years) in carriers of MLH1 or MSH2 mutations. A 95% confidence interval for the life time risk is also shown. Source: Vasen *et al.* (2001).

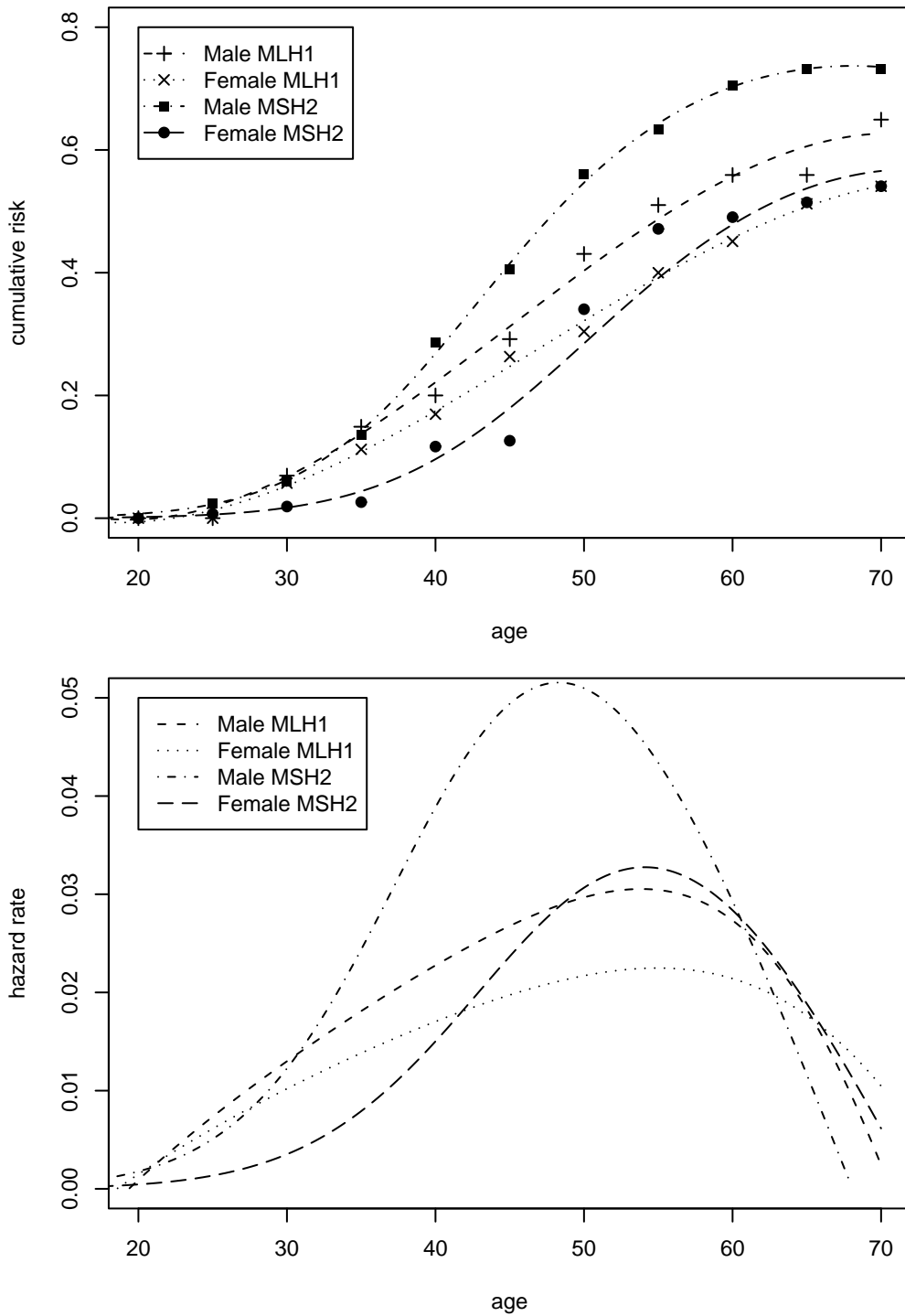|	MLH1 mutation					MSH2 mutation				
Age	40	50	60	70	95%CI	40	50	60	70	95%CI
	%	%	%	%	%	%	%	%	%	%
Stomach	0.0	0.4	2.1	2.1	0.0–4.7	0.7	1.2	4.3	4.3	0.5–8.1
Urinary Tract	0.0	0.0	0.0	1.3	0.0–3.9	0.4	1.5	5.4	12.0	3.5–20.0
Small Bowel	0.0	0.7	4.0	7.2	1.5–12.9	0.8	1.2	4.5	4.5	0.5–8.5
Ovary	0.6	2.3	3.4	3.4	0.0–6.8	1.7	6.4	10.4	10.4	3.2–17.6
Brain	0.0	0.0	0.0	0.0	—	1.2	1.2	1.2	1.2	0.0–2.6

Figure 2: The observed and fitted probability (top) and rate of onset (bottom) of colorectal cancer for male and female MLH1 or MSH2 mutation carriers. Data: Vasen *et al.* (2001).

$$F(x)_{MLH1.m}^{CRC} = 0.3406 - 0.039040x + 0.001294x^2 - 0.000009611x^3 \quad (22 \le x \le 70). \quad (3)$$

$$F(x)_{MLH1.f}^{CRC} = 0.2193 - 0.02570x + 0.0008584x^2 - 0.000006075x^3 \quad (23 \le x \le 70). \quad (4)$$

$$F(x)_{MSH2.m}^{CRC} = \frac{\exp(-10.92 + 0.3512x - 0.002579x^2)}{1 + \exp(-10.92 + 0.3512x - 0.002579x^2)} \quad (20 \le x \le 68). \quad (5)$$

$$F(x)_{MSH2.f}^{CRC} = \frac{\exp(-12.36 + 0.3498x - 0.002421x^2)}{1 + \exp(-12.36 + 0.3498x - 0.002421x^2)} \quad (20 \le x \le 70). \quad (6)$$

We assume that $F(x)_{MLH1.m}^{CRC} = 0$ below age 22 and $F(x)_{MLH1.f}^{CRC} = 0$ below age 23.

4.3 Onset of Endometrial Cancer and Extracolonic Cancers

The risk of endometrial cancer (EC) among female mutation carriers is much less than the risk of CRC (the life-time risk is less than 40%). We fitted the penetrance functions shown in Equations (7) and (8) for MLH1 and MSH2 mutation carriers, respectively. The observed and fitted penetrances and the rates of onset are illustrated in Figure 3.

$$F(x)_{MLH1}^{EC} = \frac{\exp(-17.78 + 0.4975x - 0.003655x^2)}{1 + \exp(-17.78 + 0.4975x - 0.003655x^2)} \quad (20 \le x \le 65). \quad (7)$$

$$F(x)_{MSH2}^{EC} = \frac{\exp(-4.307 - 0.1973x + 0.008763x^2 - 7.432 \times 10^{-5}x^3)}{1 + \exp(-4.307 - 0.1973x + 0.008763x^2 - 7.432 \times 10^{-5}x^3)} \quad (20 \le x \le 65). \quad (8)$$

MMR mutation carriers may be at higher risk than non-carriers of developing extracolonic cancers. Table 2 presents these risks in respect of type of cancer and type of mutation. The penetrances are relatively small, compared with those of CRC and EC.

Equations (9) and (10) show fitted penetrances of extracolonic cancers for male and female MLH1 mutation carriers respectively and Equations (11) and (12) for male and female MSH2 mutation carriers. Figure 4 illustrates the observed and fitted penetrances and the onset rates.

$$F_{MLH1.m}^{OECC}(x) = \frac{\exp(-30.6539 + 0.8128x - 0.0058x^2)}{1 + \exp(-30.6359 + 0.8128x - 0.0058x^2)} \quad 20 \le x \le 70. \quad (9)$$

$$F_{MLH1.f}^{OECC}(x) = \frac{\exp(-18.76 + 0.4747x - 0.003334x^2)}{1 + \exp(-18.76 + 0.4747x - 0.003334x^2)} \quad 20 \le x \le 70. \quad (10)$$
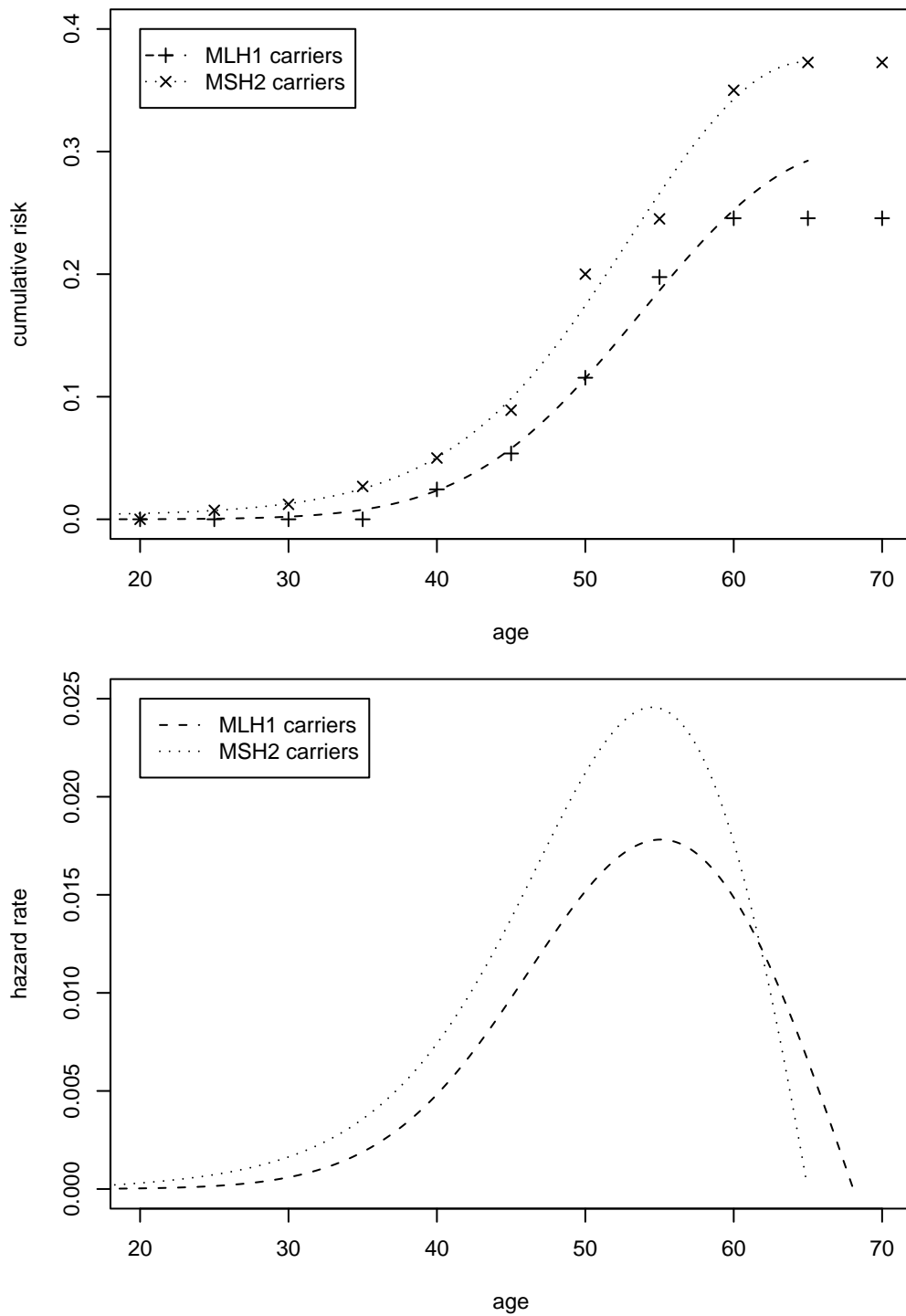
Figure 3: The observed and fitted probability (top) and rate of onset (bottom) of endometrial cancer (EC) for MLH1 and MSH2 mutation carriers. Data: Vasen *et al.* (2001).

Table 3: Cumulative risks (%) of developing any one of the five extrocolonic cancers for MLH1 or MSH2 mutation carriers. The cancers include cancer of: stomach, urinary tract, small bowel, ovary (female only) and brain.

Age	MLH1 Mutation				MSH2 Mutation			
	40	50	60	70	40	50	60	70
Male Risk	0.0%	1.10%	6.02%	10.33%	3.07%	5.00%	14.58%	20.54%
Female Risk	0.6%	3.37%	9.21%	13.38%	4.71%	11.08%	23.46%	28.8%

$$F_{MSH2.m}^{OECC}(x) = \frac{\exp(29.80 - 2.014x + 0.039x^2 - 0.00023x^3)}{1 + \exp(29.80 - 2.014x + 0.039x^2 - 0.00023x^3)} \qquad 42 \le x \le 65. \quad (11)$$

$$F_{MSH2.f}^{OECC}(x) = \frac{\exp(-10.45 + 0.2501x - 0.001618x^2)}{1 + \exp(-10.45 + 0.2501x - 0.001618x^2)} \qquad 20 \le x \le 70. \quad (12)$$

Below age 42, the intensity corresponding to $F_{MSH2.m}^{OECC}(x)$ is extrapolated linearly to zero at birth.

4.4 *Population Risk of CRC and EC*

Point estimates of rates of onset of CRC and EC in the population were given by:

$$\frac{\text{No. of cases of onset at age } x}{\text{Person-years exposed to risk at age } x}$$

based on the National Cancer Statistics in England and Wales in 1988–1992, and the corresponding population totals. We fitted the following beta functions to the onset rates of CRC for both males and females, on the age range 0–89 years, using weighted least squares, with the exposures as convenient weights.

$$\mu_{pop.m}^{CRC}(x) = 0.001400 \frac{\Gamma(10.196)}{\Gamma(8.196)\Gamma(2)} \left(\frac{80x}{89^2}\right)^{7.196} \left(1 - \frac{80x}{89^2}\right) \qquad (13)$$

$$\mu_{pop.f}^{CRC}(x) = 0.001092 \frac{\Gamma(8.207)}{\Gamma(6.742)\Gamma(1.465)} \left(\frac{80x}{89^2}\right)^{5.742} \left(1 - \frac{80x}{89^2}\right)^{0.465}. \qquad (14)$$

In the case of EC, we fitted an exponential function up to age 54 (but assumed to be zero before age 20) and a polynomial beyond age 57, blended linearly between these ages.

$$\mu_{pop}^{EC}(x) = \begin{cases} \exp(-17.32 + 0.09261x + 0.004273x^2 - 5.200 \times 10^{-5}x^3) & 20 \le x \le 54 \\ -4.665 \times 10^{-4} + 2.446 \times 10^{-5}x - 1.555 \times 10^{-7}x^2 & 57 \le x \le 89 \end{cases} \qquad (15)$$

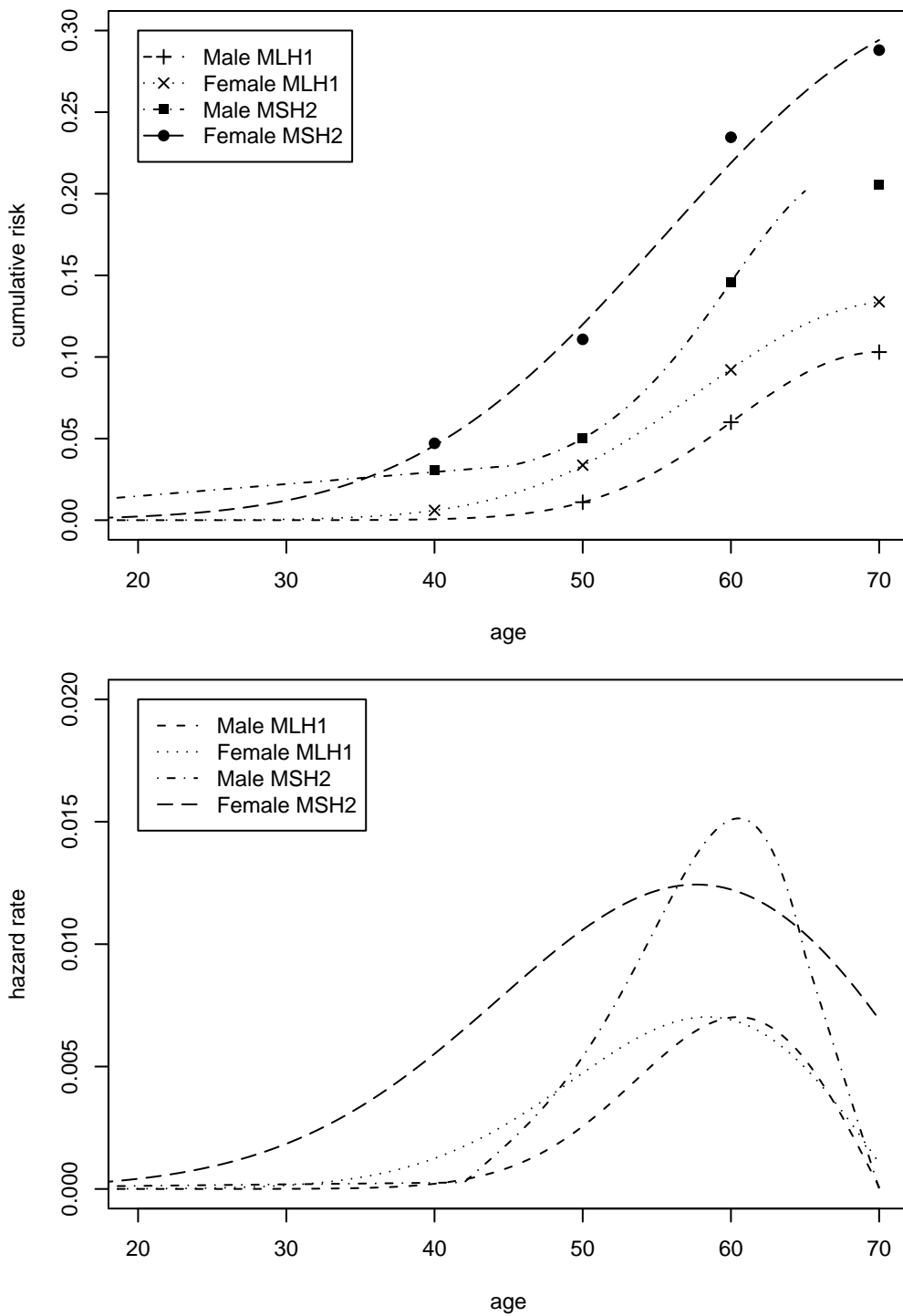These rates are illustrated in Figure 5.

Figure 4: The observed and fitted probability (top) and rate of onset (bottom) of extracolonic cancers (OECC) for MLH1 and MSH2 mutation carriers. Data: Vasen *et al.* 2001 Vasen *et al.* (2001).
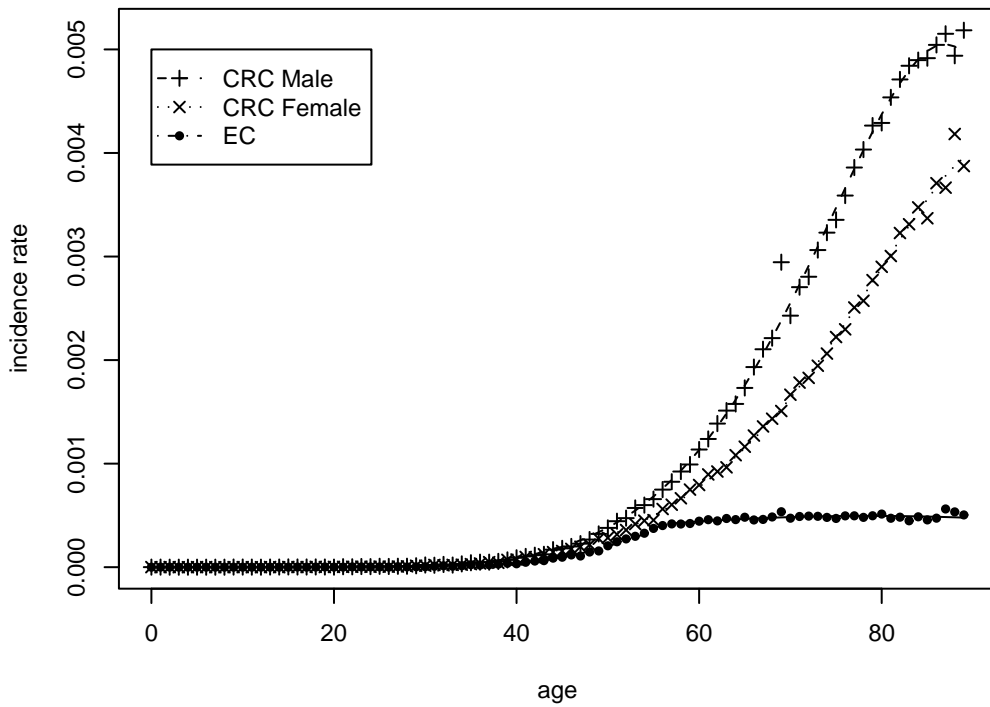
Figure 5: Point and smoothed estimates of the rates of onset of CRC and EC for the general population. Source: National Cancer Statistics in England and Wales, 1988–1992.

Table 4: Level net premiums for a unit sum assured for level CI cover for persons with a known gentotype. The premium rates for known mutation carriers are presented as a percentage of the premium for non-carriers (standard risks).

Age at Entry	Policy Term	Non-carrier		MLH1 carrier		MSH2 carrier	
		Males	Females	Males	Females	Males	Females
				%	%	%	%
20	10	0.00031	0.00054	2,040	992	1,996	674
	20	0.00062	0.00089	1,727	1,068	2,095	905
	30	0.00122	0.00145	1,101	903	1,457	922
	40	0.00212	0.00214	716	709	918	746
30	10	0.00112	0.00148	1,662	1,169	2,289	1,085
	20	0.00218	0.00241	1,033	950	1,613	1,058
	30	0.00370	0.00353	680	743	972	848
40	10	0.00399	0.00398	774	875	1,342	1,167
	20	0.00653	0.00576	530	696	837	927
50	10	0.01113	0.00892	409	604	549	842

5. Critical Illness Insurance Premiums by Genotype

Using the intensities obtained in Section 4, we calculate the level premium for a £1 sum assured under a CI insurance contract for an identified MMR mutation carrier, for various terms and entry ages.

The rates of developing other critical illnesses is based on the model described in Gutiérrez and Macdonald (2003) and we refer readers there for details. The rate of mortality is based on the English Life Tables No. 15 adjusted to exclude deaths which follow a CI insurance claim.

Table 4 shows examples of the premium rates for a unit sum assured for a healthy person whose genotype is known. The standard premium rate is assumed to be that for non-carriers, and it is shown while the premium rates for mutation carriers are presented as a percentage of those standard premium rates. The extra premiums are 4 to 23 times the basic premium. Note that most insurers would decline the application should the extra premium be 200% to 250% higher than the basic premium.

We can find other features in Table 4:
(a) The premium rates are more sensitive to entry age than to policy term.
(b) In the case of male MLH1 carriers, a lower entry age results in a larger extra premium than does a higher entry age. The extra premium for a 20-year policy on entry at age 20 is 17 times the basic premium, compared with 5 times at entry age 40.
(c) Female MSH2 mutation carriers would be charged much lower premiums than male carriers at low entry ages with short terms, because the rate of onset of CRC for males is about two or more times that for females before age 40. After age 40, the onset of EC results in roughly the same overall risk of cancer for both sexes.

Table 5: Subpopulations defined by an insurance applicant's own genotype and any mutation present in their family.

Subpopulation	Mutation Present in Family	Mutation Carried by Applicant	Applicant's Genotype
$i = 1$	None	None	$g_i = 1$
$i = 2$	MLH1	MLH1	$g_i = 2$
$i = 3$	MLH1	None	$g_i = 1$
$i = 4$	MSH2	MSH2	$g_i = 3$
$i = 5$	MSH2	None	$g_i = 1$

6. A MODEL OF FAMILY HISTORY OF HNPCC

6.1 *Family History Does Not Always Reveal the Presence of a Mutation*

In practice it is unlikely that insurers would learn an applicant's genotype, and possibly illegal to use that information anyway. The most they might learn is that there is a history of disease among the (healthy) applicant's first-degree and (less likely) second-degree relatives. If the disease has no cause except mutations in a single gene, Mendel's laws then govern the applicant's chances of carrying the mutation. But non-carriers of MMR mutations can also develop CRC, EC and other extracolonic cancers, so the applicant's MMR genotype probabilities, *given* a family history of these cancers, are not determined by Mendel's laws alone. However, these genotype probabilities are clearly the correct basis for underwriting. In Section 7 we will present a method for estimating these probabilities. The model on which this method is based is set out in the next section.

6.2 *Modelling Onset of Family History*

Based on our genotype model in Figure 1, we partition the population into five groups or subpopulations, labelled by the letter i, according to the combination of the person's own genotype *and*, in groups 2–5, the presence of a specific mutation in their family; see Table 5. The genotype of an applicant in Subpopulation i is denoted g_i, and these values are also shown in Table 5. We make the following assumptions.

(a) Persons in all five subpopulations are at risk of CRC and (females only) EC.
(b) A person's risk of CRC or EC depends only on their own genotype; the rates of onset are as estimated in Section 4. Given the genotype, the family history is irrelevant.
(c) The proportion of the population in Subpopulations 2 and 4 are given by the population mutation frequencies, which we took to be 0.000769 and 0.000714, respectively.
(d) A child of a mutation carrier has probability 1/2, at birth, of being a carrier. Hence the proportions in Subpopulations 3 and 5 are the same as those in Subpopulations 2 and 4, respectively.

The key innovation in the model is that, provided the definition of family history is sufficiently precise, then the 'onset' of a family history is an event that occurs at a precise time, so it can be modelled as a transition between states, governed by a transition
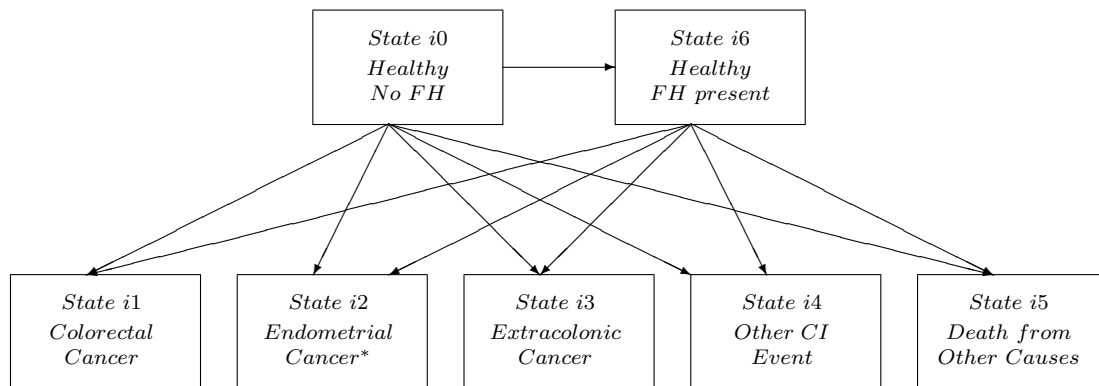
Figure 6: Model 2: A model of HNPCC, development of a family history, and critical illness insurance, in subpopulation i. (*Females only.)

intensity capable of being computed from the other model assumptions. The events that precipitate the 'onset' of a family history are, of course, events in the life histories of the applicant's relatives.

Figure 6 shows how we represent the development of a family history for Subpopulation i as a transition between the states labelled $i0$ and $i6$. Because of assumption (b) above, we have $\mu_x^{i0j} = \mu_x^{i6j}$ for $j = 1, \ldots, 5$. Note that a family could exist at birth, so the applicant starts in state $i6$, depending on the definition of family history being used.

The probability that a healthy person age x is a mutation carrier, given only the presence or absence of a family history, is proportional to the occupancy probabilities in the relevant states as follows:

$$P[\text{ In Subpopulation } i \mid \text{No Family History by Age } x\ ] \quad = \quad \frac{\text{OP}_x^{i0}}{\sum_{j=1}^{5} \text{OP}_x^{j0}} \qquad (16)$$

$$P[\text{ In Subpopulation } i \mid \text{Family History by Age } x\ ] \quad = \quad \frac{\text{OP}_x^{i6}}{\sum_{j=1}^{5} \text{OP}_x^{j6}}. \qquad (17)$$

7. ESTIMATING THE RATE OF 'ONSET' OF FAMILY HISTORY

7.1 *Defining Family History*

Underwriting practice can differ considerably from clinical practice, usually taking account only of affected FDRs. For example, Table 6 shows the underwriting guidelines for applicants with a family history of CRC used by one large reinsurer. We consider three scenarios that might represent underwriting criteria:
(a) Scenario 1: at least one FDR develops CRC before age 60;
(b) Scenario 2: at least two FDRs develop CRC before age 60;
(c) Scenario 3: at least one FDR develops CRC before age 60 and at least one FDR develops EC before age 60.

Table 6: Example underwriting guidelines for applicants with a family history of CRC.

Applicant	Family History	Premium Rate
below 50	one FDR with CRC before age 60	+50/Excl
below 50	two or more FDR(s) with CRC before age 60	CMO/Decline/Excl
over 50	exclude those with one or more FDR with CRC	Standard
over 50	one FDR with CRC before age 60	+25
over 50	two or more FDRs with CRC before age 60	+75/CMO/Excl

$\big($FDR =First Degree relative, CMO = Refer Chief Medical Officer, Excl = Exclude$\big)$

Table 7: The distribution of the number of siblings. Source: Macdonald *et al.* (2003a).

No. of siblings k	0	1	2	3	4	5	≥ 6
Probability P[k]	0.2311	0.5006	0.1973	0.0541	0.0124	0.0034	0.0011

7.2 *A Model of Family Structure*

Given a person in Subpopulation i, we can compute: (a) the probability that they already have a family history when they are born; and (b) the probability that they have developed a family history by age x, which we denote ${}^{i}F_x^{fh}$. By numerical differentiation, we then obtain the intensity μ_x^{i06} in Figure 6. Clearly, the family histories that it is possible for a person to have depend on the number and configuration of their FDRs, and since this is usually not known by the insurer, we need a simple model of family structure over which we can average to obtain ${}^{i}F_x^{fh}$. We make the following assumptions:

(a) The parents are 30 years older than the applicant and any siblings are all the same age as the applicant.

(b) The distribution of the number of siblings is as given by Macdonald *et al.* (2003a) and is shown in Table 7. We assume the applicant has no more than 6 siblings.

(c) Siblings could be male or female with equal probability.

(d) The applicant's siblings in a mutation-carrying family (Subpopulations 2–5) are each carriers with probability 1/2, independently of each other.

(e) Exactly one of the parents in a mutation-carrying family is a carrier. We neglect the rare possibility that both parents are carriers.

7.3 *The Probability of a Family History at Birth*

A person may be born with a family history if one or both parents develops CRC before age 30 (note that EC affects fertility). Both parents doing so is such a rare event that we neglect it, so we need only consider Scenario 1, and the possibility that one parent develops CRC before age 30.

If we assume that mortality is the same before and after onset of CRC (perhaps a reasonable approximation to make before age 30) and if we ignore all decrements except death and onset of CRC, then we can compute, for a parent with sex s and genotype g:

$$\omega_s^g = \text{P[Healthy at Age 30]} = \frac{\text{P[Alive, No CRC at Age 30]}}{\text{P[Alive, No CRC at Age 30]} + \text{P[Alive, CRC at Age 30]}}.$$

Then the probability that neither parent has had CRC by age 30 in Subpopulation i is approximately:

$$\frac{\omega_m^1 \, \omega_f^{g_i} + \omega_m^{g_i} \, \omega_f^1}{2}. \tag{18}$$

Thus we find that under Scenario 1, the probability of a family history at birth is 0.06 in Subpopulations 2 and 3, 0.04 in Subpopulations 4 and 5, and less than 0.001 in Subpopulation 1.

7.4 *The Probability of Developing a Family History after Birth*

Most people will not be born with a family history, but will be at risk of developing one thereafter, depending on the life histories of their parents and siblings. We illustrate the calculation for Scenario 1; the details for the other scenarios are in the Appendix.

Consider a healthy person age x in Subpopulation i. Let $^iP_x^{0P}$ denote the probability that neither of their parents has developed CRC by age $x + 30$, and let $^iP_x^{0S}$ denote the probability that none of their siblings has developed CRC by age x. So $^iF_x^{fh}$ can be written as two terms, depending on whether or not a parent develops CRC:

$$^iF_x^{fh} = (1 - {}^iP_x^{0P}) + {}^iP_x^{0P} \times (1 - {}^iP_x^{0S}). \tag{19}$$

Referring to the model in Figure 1, let $_sP_{x,t}^{g_ijk}$ denote that probability that a person of sex s and age x in Subpopulation i, with genotype g_i, who is currently in state g_ij, is in state g_ik at age $x+t$. In particular, $_sP_{x,t}^{g_i01}$ is the probability that a healthy person has contracted CRC. Then the first term in Equation (19) is evaluated in the same way as Equation (18):

$$^iP_x^{0P} = \frac{1}{2}(1 - {}_mP_{30,x\wedge30}^{101})(1 - {}_fP_{30,x\wedge30}^{g_i01}) + \frac{1}{2}(1 - {}_mP_{30,x\wedge30}^{g_i01})(1 - {}_fP_{30,x\wedge30}^{101}) \tag{20}$$

where $x \wedge 30$ indicates that onset of CRC after age 60 does not contribute to a family history. Given the assumptions in Section 7.2, the unconditional probability that any sibling has had CRC by age x or by age 60 (whichever is less), denoted $^iP_{x\wedge60}$, is $({}_mP_{0,x\wedge60}^{101} + {}_mP_{0,x\wedge60}^{g_i01} + {}_fP_{0,x\wedge60}^{101} + {}_fP_{0,x\wedge60}^{g_i01})/4$, and:

$$^iP_x^{0S} = \sum_{k=0}^{6} \text{P}[k] \, (1 - {}^iP_{x\wedge60})^k \tag{21}$$

where $\text{P}[k]$ is given in Table 7.

7.5 *Illustration*

Figure 7 illustrates the rate of developing a family history under each scenario. Note that the scales are not all the same. We see the following features:
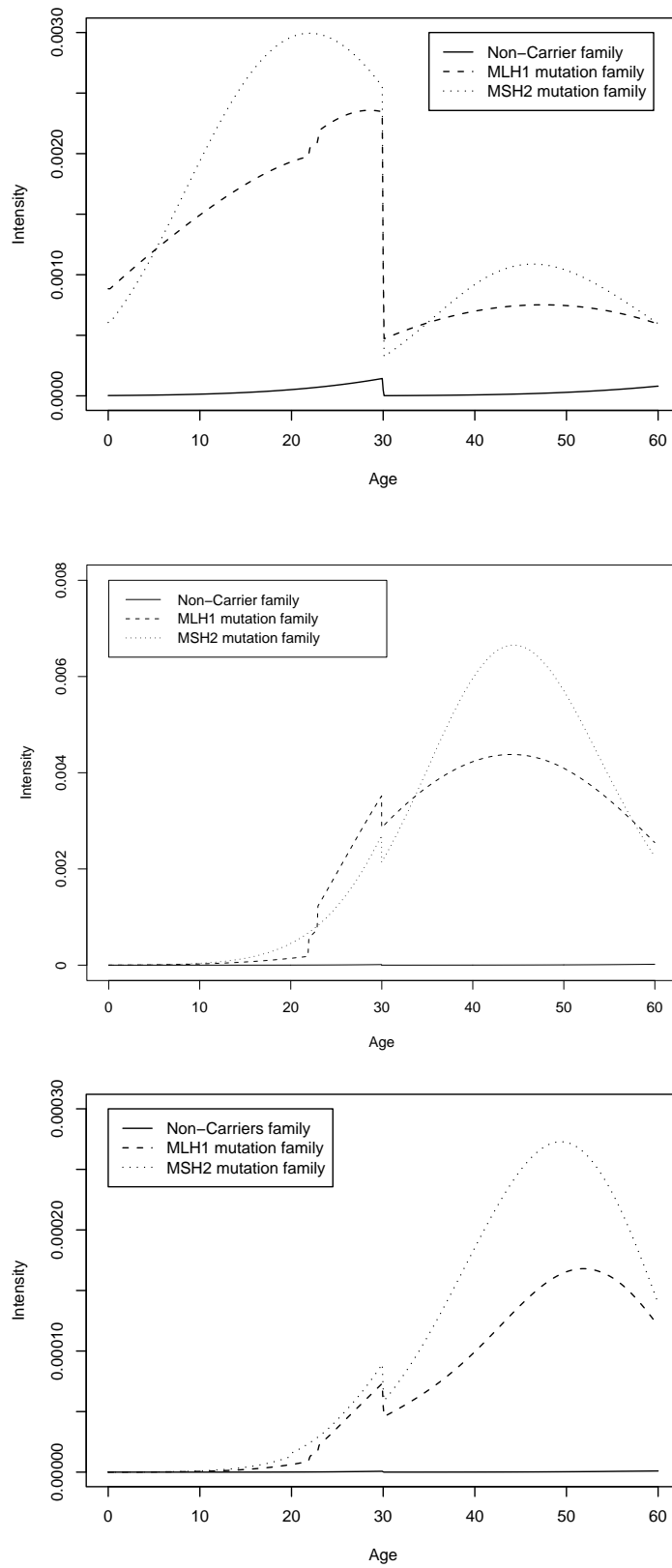
Figure 7: The rate of developing family history for applicants in non-MMR-mutation-carrying families, MLH1 mutation-carrying families and MSH2 mutation-carrying families under family history Scenarios 1 (top), 2 (middle) and 3 (bottom).

(a) The rates of 'onset' are considerably higher under Scenario 1, which uses the weakest definition of family history.
(b) There is a downward jump at age 30 in all cases. This is because the parents then reach age 60 and can no longer contribute to a family history.
(c) The rates of 'onset' for mutation carriers increase after age 30, peak at about age 45 and then decrease. This pattern is consistent with the onset rate of CRC in Figure 2.
(d) The rate of 'onset' for MLH1 mutation carriers jumps at age 20 under all scenarios. This is because the rate of onset of CRC is assumed to be zero before age 20.
(e) The risk of developing a family history after age 60 is zero, by definition.

8. Premium Rates Allowing for a Family History of CRC

Here, we consider what premium rates would be charged to a person age x applying for CI insurance, who has a family history of CRC, assuming this information may be used in underwriting. Let $\mathcal{C}$ denote the set of states included in this underwriting class. The EPV of any insurance cashflows in respect of an arbitrary member of this underwriting class is just a weighted average of the EPVs in respect of each genotype represented in that underwriting class, the weights being the proportions OP_x^{ij} present therein at age x. Let B_x^{ij} and C_x^{ij} be the EPVs of a unit benefit and a unit premium, respectively, in respect of an applicant age x in state ij of the model in Figure 6. The level premium, denoted $\mathrm{LP}_x^{\mathcal{C}}$, is:

$$\mathrm{LP}_x^{\mathcal{C}} = \frac{\sum_{ij \in \mathcal{C}} \mathrm{OP}_x^{ij} \, B_x^{ij}}{\sum_{ij \in \mathcal{C}} \mathrm{OP}_x^{ij} \, C_x^{ij}}.$$

Table 8 gives the level premium for level CI insurance for a person with a family history of CRC, expressed as a percentage of the OR premium. We observe the following:
(a) The extra premium depends on the definition of family history. Only the intensity μ^{i02} in Figure 6 is altered; the other intensities are unchanged.
(b) The extra premium varies from 30% to 374% for different entry ages and terms. This means that practically all applicants for CI cover with a family history of CRC or other HNPCC-related cancers are insurable, under current underwriting practice. This is significantly more lenient than, for example, Table 6 would suggest.

9. The Extended Critical Illness Insurance Model

9.1 *An Insurance Market Model*
So far we have considered a transparent insurance market in which there is no moratorium and no adverse selection.
(a) Model 1 represented the life history of a person who had already bought CI insurance.
(b) Model 2 extended Model 1 by allowing for the development of a family history.
(c) To model adverse selection and moratoria, we need to extend Model 2 to make the purchase of insurance an event in the life history. We will call this Model 3.

Table 8: Level premiums for a term insurance with £1 sum assured for level CI cover for persons with family history, as a percentage of the premium for a standard risk, using three different definitions of family history.

Age at Entry	Policy Term	Scenario 1		Scenario 2		Scenario 3	
		Male %	Female %	Male %	Female %	Male %	Female %
20	10	395	217	474	235	398	210
	20	365	236	456	267	380	234
	30	265	219	326	253	277	222
	40	197	188	232	215	203	191
30	10	259	191	473	315	343	235
	20	192	174	319	276	242	212
	30	153	153	226	226	181	180
40	10	177	176	373	368	302	295
	20	142	151	254	286	212	233
50	10	130	147	219	287	195	250

   Model 3 is shown (partially) in Figure 8. It is a model of the entire CI insurance market operating up to age 60. A person may buy CI insurance at any time, before or after developing a family history, or taking a genetic test. Persons in States $i0$ to $i5$ are at risk of suffering a critical illness or death, and we refer to these states as 'at-risk' states. For simplicity we combine all states representing a claim (States $i1$ to $i5$ in Figures 1 and 6) into a single 'CI Event' state (State $i6$ in Figure 8). We assume the following:

(a) A person is born healthy with or without a family history of CRC, depending on the criterion of family history being used. Hence he/she is in state $i0$ or state $i2$ at birth.

(b) A person may buy CI insurance after age 20, if they are in State $i0$ or State $i2$.

(c) After the onset of a family history, an individual may be advised to have a genetic test. We ignore the unlikely event of genetic testing in the absence of a family history. This assumption would change if a screening program for MMR mutations were introduced, but at present there are no such programs.

(d) People with a family history or an adverse genetic test result may be charged higher than standard premium rates. Hence they may buy insurance at a different rate from persons charged the standard rate. This is modelled by varying the intensities μ_x^{i01}, μ_x^{i23} and μ_x^{i45}.

(e) Once insurance has been purchased, the subsequent onset of a family history or taking of a genetic test are irrelevent.

(f) A payment of $b_x^{ij6} = 1$ is made on entering state $i6$ from an 'at-risk' state ij at age x.

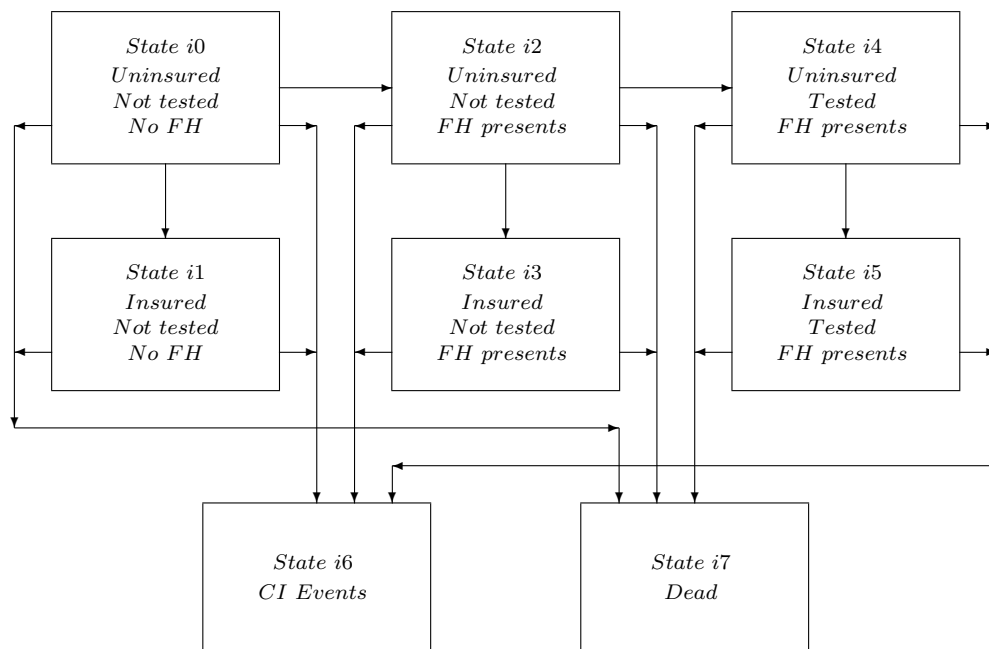(g) A premium at rate ψ_x^{ij} is paid continuously by a person age x in an insured state ij.

Figure 8: Model 3: A model of the critical illness insurance market allowing for family history of CRC and genetic testing, for a person in Subpopulation i. All states are at risk of death and critical illness (the same events as shown in Figure 1).

9.2 *Definition of Underwriting Classes*

The insurer may divide the whole population into several underwriting classes, according to the risk factors used. Everyone in an underwriting class is charged the same rate of premium, which we assume satisfies the equivalence principle within that class. Table 9 lists possible components of each underwriting class under various restrictions on the information that can be used in underwriting.

Consider line No. 1 in Table 9 as an example. There are three underwriting classes, the 'ordinary rates' class (OR), the 'family history' class (FX) and the 'positive genetic test result' class. The OR class includes persons with no family history and those who have a family history but have been verified as non-carriers by genetic testing. The FX class consists of those who have a family history except those who have genetic test results. The 'positive genetic test result' class consists those who are verified carriers. The FX underwriting class is heterogeneous.

Given an underwriting class denoted $\mathcal{C}$, EPVs of insurance cashflows in respect of its members will be weighted averages of the EPVs in respect of the various genotypes represented in $\mathcal{C}$. The weights, which we denote $W_x^{ij}(\mathcal{C})$, are just the occupancy probabilities in Model 3, which will depend on movements among the various states, including the purchase of insurance, the onset of family histories and the taking of genetic tests. (We include $\mathcal{C}$ in the notation explicitly because we will consider alternative definitions of underwriting classes shortly. Clearly $W_x^{ij}(\mathcal{C}) = 0$ if $ij \notin \mathcal{C}$, and note that the $W_x^{ij}(\mathcal{C})$ are not the same as the occupancy probabilities OP_x^{ij} defined earlier in respect of Model 2.) We need to make some reasonable assumptions about the intensities representing all these events. We believe the following assumptions are reasonable:
(a) The size of the insurance market is defined by the rate of purchase of insurance

Table 9: Possible underwriting classes with five sub-populations labelled by i. (F) denotes persons who have a family history and (U) denotes those who do not. $\oplus$ denotes a positive test result and $\ominus$ denotes a negative test result.

| | Factors Allowed in Underwriting | | | | Composition of Underwriting Classes | | |
No.	Genetic Testing Exists?	Family History	-ve Test Result	+ve Test Result	Ordinary Rates Class	Rated for Family History	Rated for Genetic Test
1	Yes	Yes	Yes	Yes	$i = 1\text{–}5$ (U) $i = 1,3,5$ ($\ominus$)	$i = 1\text{–}5$ (F)	$i = 2,4$ ($\oplus$)
2	Yes	Yes	Yes	No	$i = 1\text{–}5$ (U) $i = 1,3,5$ ($\ominus$)	$i = 1\text{–}5$ (F) $i = 2,4$ ($\oplus$)	
3	Yes	Yes	No	No	$i = 1\text{–}5$ (U)	$i = 1\text{–}5$ (F, $\oplus$ and $\ominus$)	
4	Yes	No	No	No	$i = 1\text{–}5$		
5	No	Yes	No	No	$i = 1\text{–}5$ (U)	$i = 1\text{–}5$ (F)	
6	No	No	No	No	$i = 1\text{–}5$		

      in Subpopulation 1. An annual rate $\mu_x^{101} = 0.05$ represents a large market while $\mu_x^{101} = 0.01$ represents a smaller market. If these apply between ages 20 and 60, 86% of people will buy insurance in the large market and 33% in the smaller market.

(b) Persons in subpopulations 2–5 with no family history purchase insurance at the same rate as persons in Subpopulation 1.

(c) Rates of onset of family history μ_x^{i02} were obtained in Section 4.

(d) The rate of genetic testing is $\mu_x^{i24} = 0.02$ per annum. This implies that 18% of people would have a genetic test within 10 years of the onset of a family history.

(e) A person in the FX class would usually be charged a higher premium than the OR class. This would discourage them from buying insurance at the normal rate. In the large market we assume that they buy insurance at the normal rate, or at half that rate, or they do not buy insurance at all. We assume that they do not buy insurance in the smaller market.

(f) We make the extreme assumption that 'adverse selectors' buy insurance at rate $\mu_x^{i45} = 0.25$ per annum. So, if a moratorium allows adverse selection to happen, 70% of people will buy insurance within 5 years of receiving an adverse genetic test result.

9.3 *Modelling Adverse Selection*

      We use Model 3 (Figure 8) to measure losses caused by adverse selection using the method described by Macdonald (2003b). The decision to buy insurance is an event whose timing is uncertain, therefore the level premium that would be charged for cover until age 60 depends on the age at purchase. This cashflow is not adapted to a Markov framework, which would require all cashflows to depend solely on the current age. The simplest solution is to charge a rate of premium, within each underwriting class, equal to the expected claims in the next time interval dt. That is, the rate of premium $\psi_{x+t}^{\mathcal{C}}$ in

Table 10: Percentage increases in premium rates arising from severe adverse selection. Moratoria on the use of genetic test results, family history underwriting still allowed. CI market operating between ages 20 and 60. The rate of genetic testing is 0.02 per annum.

Size of Market	Insurance Purchasing of People with Family History	Family History Scenario	Moratorium on Using			
			Adverse Test Results		All Test Results	
			Males	Females	Males	Females
			%	%	%	%
Large	Normal	1	0.009	0.009	0.009	0.009
		2	0.000	0.000	0.000	0.000
		3	0.000	0.000	0.000	0.000
	Half	1	0.015	0.016	0.015	0.016
		2	0.000	0.000	0.000	0.000
		3	0.000	0.000	0.000	0.000
	Nil	1	0.026	0.029	0.027	0.030
		2	0.000	0.000	0.000	0.000
		3	0.000	0.000	0.000	0.000
Small	Nil	1	0.083	0.095	0.084	0.096
		2	0.001	0.002	0.001	0.002
		3	0.000	0.001	0.000	0.001

underwriting class $\mathcal{C}$ is:

$$\psi_{x+t}^{\mathcal{C}} = \frac{\sum_{ij \in \mathcal{C}} W_{x+t}^{ij}(\mathcal{C}) \mu_{x+t}^{ij6}}{\sum_{ij \in \mathcal{C}} W_{x+t}^{ij}(\mathcal{C})}. \tag{22}$$

Using these rates of premium, we can compute the EPV of benefits payable in each underwriting class, $B^{\mathcal{C}}$, and the expected value of the insurance loss, $L^{\mathcal{C}}$. The EPV of the premiums payable is then $B^{\mathcal{C}} - L^{\mathcal{C}}$. If there is no adverse selection and people buy insurance at the rate assumed when calculating the premium rate $\psi_{x+t}^{\mathcal{C}}$, then $L^{\mathcal{C}} = 0$.

Suppose that, before a moratorium is imposed or agreed, there are just two underwriting classes, OR and FX, because genetic testing does not exist. After a moratorium is agreed, there are still just two underwriting classes, because genetic test results may not be used, but their composition may have changed because of the extra information available to insured persons. We will denote these changed underwriting classes OR* and FX*. Then adverse selection may give rise to a cost because:

(a) we continue to calculate premium rates $\psi_{x+t}^{\mathcal{C}}$ using the weights $W_{x+t}^{ij}(\text{OR})$ and $W_{x+t}^{ij}(\text{FX})$; but

(b) EPVs are based on premium rates $\psi_{x+t}^{\mathcal{C}}$ and the new weights $W_{x+t}^{ij}(\text{OR}^*)$ and $W_{x+t}^{ij}(\text{FX}^*)$.

Table 11: Percentage increases in premium rates arising from severe adverse selection. Moratoria on the use of genetic test results, family history underwriting still allowed. CI market operating between ages 20 and 60. The rate of genetic testing is 0.10 per annum.

| Size of Market | Insurance Purchasing of At-risk Individuals | Family History Definition | Moratorium on Using | | | |
| | | | Adverse Test Results | | All Test Results | |
			Males %	Females %	Males %	Females %
Large	Normal	1	0.024	0.024	0.026	0.026
		2	0.000	0.001	0.000	0.001
		3	0.000	0.000	0.000	0.000
	Half	1	0.040	0.041	0.043	0.044
		2	0.001	0.001	0.001	0.001
		3	0.000	0.000	0.000	0.000
	Nil	1	0.067	0.071	0.074	0.079
		2	0.001	0.001	0.001	0.001
		3	0.000	0.000	0.000	0.000
Small	Nil	1	0.214	0.233	0.233	0.255
		2	0.004	0.005	0.005	0.006
		3	0.001	0.002	0.002	0.002

9.4 *A Moratorium on Genetic Test Results*

Define the following EPVs:

$$
\begin{aligned}
B = B^{\text{OR}} + B^{\text{FX}} \quad &= \quad \text{EPV of benefits with no adverse selection} \\
L = L^{\text{OR}} + L^{\text{FX}} \quad &= \quad \text{EPV of loss with no adverse selection} \\
B^* = B^{\text{OR}^*} + B^{\text{FX}^*} \quad &= \quad \text{EPV of benefits with adverse selection} \\
L^* = L^{\text{OR}^*} + L^{\text{FX}^*} \quad &= \quad \text{EPV of loss with adverse selection.}
\end{aligned}
\tag{23}
$$

Hence the insurer who sets premium rates assuming no adverse selection would suffer loss L^* and would raise all premium rates by a proportion $(L^* - L)/(B^* - L^*)$ to compensate.

Table 10 shows the percentage increases in premium rate arising from severe adverse selection, under a moratorium on genetic test results, the use of family history still being permitted. We show the effects of two versions of this moratorium (lines 2 and 3 in Table 9), depending on whether or not a person with a clear test result may be offered the OR rate of premium or not. The difference is tiny, and indeed the costs of adverse selection are barely observable. All figures indicate premium increases of 0.1% or less.

As noted by Macdonald (2003a,b) and Gutiérrez and Macdonald (2003), we expect the cost of adverse selection to increase if the moratorium bans the use of all test results. The reason is that a small group of verified non-carriers would be included in the FX^* class, and the rate of premium in that class would be reduced.

We found that the insurance loss was nearly proportional to the sum assured taken out by 'adverse selectors', so as this increases above the average (the second possible form of adverse selection) multiples of the premium increases in Table 10 give a sufficient

indication of the cost of adverse selection.

Our choice of a modest rate of genetic testing (0.02 per annum) assumes that doctors act as gatekeepers, controlling access to genetic tests according to clinical need. This is true in the UK, where genetic tests are not easily available outside the National Health Service. Of our three scenarios, Scenario 2 is perhaps closest to the clinical criteria for diagnosing HNPCC. However, this means that our rate of genetic testing is probably not all that extreme, therefore the costs in Table 10 are not worst-case in this respect. In Table 11 we assume genetic testing to be more widespread, with a rate of testing of 0.1 per year, so that about 85% of people are tested within 20 years of developing a family history. The costs are considerably higher than those in Table 10, although their absolute values are still small (the largest premium increase is less than 0.3%).

9.5 *A Moratorium on Genetic Tests and Family History*

Extending a moratorium to family history information might affect an insurance market in two ways, both leading to an increase in premiums, as described by Macdonald (2003b). First, it will create a single underwriting class, that we denote $OR^{\dagger}$. People with a family history, who before would tend to buy less insurance than normal, could now buy insurance at the normal rate (ignoring the small effect of the consequent premium increase on the demand for insurance). This cannot be called adverse selection, if its motivation is simply to fulfil the normal needs for insurance rather than to exploit hidden information. Let $B^{\dagger}$ and $L^{\dagger}$ be the EPVs of benefits and losses in this market. Let $\tilde{B}$ be the EPV of benefits assuming that family histories were never observed, either by individuals or by insurers. Then every healthy person would buy insurance at the normal rate. So the first effect of the moratorium would be to increase the OR rate of premium by the proportion $B^{\dagger}/\tilde{B} - 1$.

An additional premium increase would arise if there was adverse selection in addition to the above change in behaviour. Let $L^{\ddagger}$ and $B^{\ddagger}$ be the EPVs of losses and benefits respectively given severe adverse selection (rate of insurance purchase 0.25). Hence premiums rise further, in the proportion $(L^{\ddagger} - L^{\dagger})/(B^{\ddagger} - L^{\ddagger})$. This time, we assume this behaviour extends to those who have a family history but have not had a genetic test, as well as those who have an adverse test result.

Table 12 illustrates each of these increases. We make the following points:
(a) Only the market size matters, because the insurance-buying behaviour of people charged higher premiums is no longer relevant.
(b) The premium increases caused by extending the OR class are all fairly small.
(c) The cost of adverse selection is also small; the largest premium increase, in the smaller market, is about 0.3%.
(d) The premium increases are largest under family history Scenario 1.
(e) As before, the insurance losses when there is adverse selection are practically proportional to the sums assured taken out by 'adverse selectors'.

10. SUMMARY AND CONCLUSIONS

Several common cancers, including CRC, have a small subset of inherited versions, not easily distinguished from sporadic cases by clinical means. These are some of the most

Table 12: Percentage increases in OR premium rates arising from new underwriting classes, and in all premiums arising from severe adverse selection, following a moratorium on the use of all genetic test results and family history. CI market operating between ages 20 and 60.

Size of Market	Family History Definition	OR premium Increases Arising from New Underwriting Classes		Premium Increases Arising from Severe Adverse Selection	
		Males %	Females %	Males %	Females %
Large	1	0.164	0.170	0.048	0.044
	2	0.002	0.002	0.001	0.001
	3	0.001	0.001	0.000	0.000
Small	1	0.172	0.180	0.304	0.307
	2	0.002	0.013	0.010	0.013
	3	0.001	0.001	0.003	0.004

important single-gene disorders, and HNPCC is a typical example. Given a purely genetic disorder (such as HD and APKD) a family history of the cancer reveals with certainty the presence of a causative mutation, and this was a key assumption in previous actuarial models of underwriting and adverse selection (Macdonald (2003a,b)). A family history of one of these common cancers does *not* reveal the presence of a causative mutation, so these models fail, or require excessive computation.

We focussed on HNPCC, and used the device of modelling the appearance of a family history as an event in a person's life history, represented by transition into a new state in a multiple-state model. Thus, given a clear enough definition of 'family history', the problem reduces to computing the age-related rate of 'onset' of a family history. This was obtained by applying relevant epidemiological data to a simple model of a nuclear family. Our conclusions in respect of CI insurance were as follows:

(a) The extra premiums for known mutation carriers are very high (Table 4). In all cases, the extra premiums exceed 250%, which is about the limit of insurability. A carrier of a MMR gene mutation is therefore very likely to be declined, if the mutation status is disclosed to the insurer.

(b) In an insurance market with neither a moratorium nor adverse selection, the extra premiums for persons with a family history of CRC vary between 30% and 374% for different entry ages and terms and definitions of family history (Table 8). This suggests that people with a family history are, in most cases, within the normal limits of insurability.

(c) The cost of adverse selection, caused by a moratorium on genetic test results or family history, is very small. In all cases we examined, premium increases to offset the cost of adverse selection would be less than 1%. This conclusion holds even in some sensitivity tests making even stronger assumptions about the severity of adverse selection.

(d) The definitions of family history we used in this paper are not identical to those used in insurance underwriting, nor to those used in a clinical setting, although they simplify and represent both reasonably well. The results do depend quite strongly on the definition of a family history; under family history Scenarios 2 and 3 the costs of adverse selection are less than those under Scenario 1.

(e) This is a simple and effective approach to modelling inherited disorders which are not purely genetic. It eliminates much of the computational burden involved in previous approaches (Macdonald *et al.*, 2003a,b). It can be extended to other disorders which have both genetic and sporadic causes, of which breast cancer is probably the most important.

We recognise that the models we have used have many elements, each of which requires assumptions that are hard to verify. Many of the model's parameters are either based, indirectly, on quite sparse data, or are reasonable guesses in the absence of any data. Some specific limitations of the model are the following.

(a) The actuary wants to price policies with a wide range of entry ages and policy terms. Thus models are parameterised by age-dependent intensities (or probabilities). This is very demanding; most medical studies of modest size publish simpler statistics whose statistical properties are robust given modest sample sizes. Thus extra assumptions (for example, the form of the onset rates) are needed to parameterise the model using published studies.

(b) Other choices of model might have advantages, for example a Cox or other regression model might make best use of the data, with genotype and perhaps sex as covariates, and would be directly applicable to our actuarial problems. However, we did not have access to the data.

(c) Without access to the data, it is impossible to estimate any statistical properties such as standard errors of premium rates, yet these are all derived from quite a small study.

(d) The model of adverse selection assumes perfectly inelastic demand for insurance. An economist's model would seek a new market equbribium once adverse selection had increased prices. This is perhaps a small defect when mutations are very rare.

Finally, we note that other studies have estimated the cost of adverse selection for other genetic disorders, including HD, APKD, breast cancer and ovarian cancer. In all these cases, the cost, in terms of premium increases, have been less than or are around 1%. But we are not yet in a position to say conclusively that the aggregate effects of all these genetic disorders are still acceptably minor.

References

Aaltonen, L., Peltomaki, P., Leach, F., Sistonen, P., Pylkkanen, L. and Mecklin, L. (1993). Clues to the pathologenesis of familial colorectal cancer. *Science*, **260**, 812–816.

Aaltonen, L., Salovaara, R., Kristo, P., Canzian, F., Hemminki, A., Peltomaki, P., Chadwick, R., Kaarianen, H., Eskelinen, M., Jarvinen, H., Mecklin, J. and Chapelle, A. (1998). Incidence of hereditary nonpolyposis colorectal cancer and the feasibility of molecular screening for the disease. *The New England Journal of Medicine*, **338**, 1481–1487.

Bellacosa, A., Genuari, M. and Anti, M. (1996). Hereditary nonpolyposis colorectal cancer: review of clinical, molecular genetics, and counseling aspects. *American Journal of Medical Genetics*, **62**, 353–364.

Boland, C. and Troncale, F. (1984). Familial colonic cancer without antecedent polyposis. *Annual of Internal Medicine*, **100**, 700–701.

Gui, E. and Macdonald, A. (2002). A Nelson-Aalen estimate of the incidence rates of Early-Onset Alzheimer's disease associated with the Presenilin-1 gene. *ASTIN Bulletin*, **32**(1), 1–42.

Gutiérrez, C. and Macdonald, A. (2003). Adult polycystic kidney disease and critical illness insurance. *North American Actuarial Journal*, **7**(2), 93–115.

Gutiérrez, C. and Macdonald, A. (2004). Huntington's disease, critical illness insurance and life insurance. *Scandinavian Actuarial Journal*, pages 279–313.

Hoem, J. (1988). The versatility of the markov chain as a tool in the mathematics of life insurance. *Transactions of the 23rd International Congress of Actuaries, Helsinki*, **S**, 171–202.

Ionov, Y., Peinado, A., Malkhosyan, S., Shibata, D. and Perucho, M. (1993). Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogeneisis. *Nature*, **363**, 558–561.

Lemaire, J., Subramanian, K., Armstrong, K. and Asch, D. (2000). Pricing term insurance in the presence of a family history of breast and ovarian cancer. *North American Actuarial Journal*, **4**, 75–87.

Liu, B., Parsons, R., Papadopoulos, N., Nicolaides., N., Lynch, H. and Watson, P. (1996). Analysis of mismatch repair genes in hereditary nonpolyposis colorectal cancer patients. *Natural Medicine*, **2**, 169–174.

Lu, L. (2006). *Some Actuarial and Statistical Investigations into Topics on Genetics and Insurance*. Ph.D. Thesis, Heriot-Watt University, Edinburgh.

Lynch, H. and de la Chapelle, A. (2003). Hereditary colorectal cancer. *The New England Journal of Medicine*, **348**, 919–932.

Lynch, H. and Krush, A. (1971). Cancer family "G" revisited: 1895–1970. *Cancer*, **27**, 1505–1511.

Lynch, H. and Smyrk, T. (1996). Hereditary nonpolyposis colorectal cancer (Lynch syndrome). *Cancer*, **78**, 1149–1167.

Macdonald, A. (2003a). Genetics and insurance: What we have learned so far? *Scandinavian Actuarial Journal*, **2003**, 324–328.

Macdonald, A. (2003b). Moratoria on the use of genetic tests and family history for mortgage-related life insurance. *British Actuarial Journal*, **9**, 279–311.

Macdonald, A., Waters, H. and Wekwete, C. (2003a). The genetics of breast and ovarian cancer I: A model of family history. *Scandinavian Actuarial Journal*, **2003**, 1–27.

Macdonald, A., Waters, H. and Wekwete, C. (2003b). The genetics of breast and ovarian cancer II: A model of critical illness insurance. *Scandinavian Actuarial Journal*, **2003**, 28–50.

Marra, G. and Boland, C. (1995). Hereditary nonpolyposis colorectal cancer: the syndrome, the genes and historical perspectives. *Journal of the National Cancer Institute*, **87**, 1114–1125.

Mitchell, R., Farrington, S., Dunlop, M. and Campbell, H. (2002). Mismatch repair genes $hMLH1$ and $hMSH2$ and colorectal cancer: A HuGE review. *American Journal of Epidemiology*, **156**, 885–902.

Peltomaki, P., Lothe, R., Aaltonen, L., Pylkkanen, L., Nystromlahti, L. M. and Seruca, R. (1993). Microsatellite instability is associated with tumors that characterize the hereditary nonpolyposis colorectal carcinoma syndrome. *Cancer Research*, **53**, 5853–5855.

Risinger, J., Berchuck, A., Kohler, M., Watson, P., Lynch, H. and Boyd, J. (1993). Genetic instability of microsatellites in endometrial carcinoma. *Cancer Research*, **53**, 5100–5103.

Rodriguez-Bigas, M., Boland, C. and Hamilton, S. (1997). A National Cancer Institute workshop on hereditary non-polyposis colorectal cancer syndrome: Meeting highlights and Bethesda guidelines. *Journal of the National Cancer Institute*, **89**, 1758–1762.

Subramanian, K., Lemaire, J., Hershey, J., Pauly, M., Armstrong, K. and Asch, D. (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.

Thibodeau, S., Bren, G. and Schaid, D. (1993). Microsatellite instability in cancer of the proximal colon. *Science*, **260**, 816–819.

Vasen, A., Meckin, J. and Khan, P. (1991). The international collaborative group on hereditary non-polyposis colorectal cancer. *Dis Colon Rectum*, **34**, 424–425.

Vasen, H., Stormorken, A., Menko, F., Nagengast, F., Kleibeuker, J., Griffioen, G. and Taal, B. (2001). MSH2 mutation carriers are at higher risk of cancer than MLH1 mutation carriers: A study of hereditary nonpolyposis colorectal cancer families. *Jounal of Clinical Oncology*, **19**, 4074–4080.

VASEN, H., WIJNEN, J., MENKO, F., KLEIBEUKER, J., TAAL, B., GRIFFIOEN, G., NAGENGAST, F., MEIJERS-HEIJBOER, E., BERTARIO, L., VARESCO, L., BISGAARD, M.-L., MOHR, J., FODDE, R. and KHAN, P. (1996). Cancer risk in families with hereditary nonpolyposis colorectal cancer diagnosed by mutation analysis. *Gastroenterology*, **110**, 1020–1027.

WARTHIN, A. (1913). Heredity with reference to carcinoma. *Archives of Internal Medicine*, **12**, 546–555.

APPENDIX

THE PROBABILITY OF ONSET OF FAMILY HISTORY AFTER BIRTH

Here we develop the age-related probability distributions of onset of a family history, under Scenarios 2 and 3 of Section 7.1, and the simplifying assumptions of Section 7.2.

(a) Family History Scenario 2:

We begin with notation. Recall that in Section 7.3 we defined ω_s^g to be the probability that a person of gender s and genotype g is alive and healthy at age 30. Similarly, we define ν_s^g to be the probability that such a person is alive and has had CRC by age 30. Next, consider a person age x in subpopulation i. Let $^iP_x^{0P}$, $^iP_x^{1P}$ and $^iP_x^{2P}$ denote the probabilities that neither of their parents, one of their parent or both of their parents, respectively, suffer onset of CRC by age $x + 30$. Let $^iP_x^{0S}$, $^iP_x^{1S}$ and $^iP_x^{2S}$ denote the probabilities that they have zero, one or two siblings, respectively, who have suffered onset of CRC by age x.

The probability that they have a family history as defined under Scenario 2 (at least two FDRs develop CRC before age 60) is denoted $^iF_x^{fh}$ and is:

$$^iF_x^{fh} = {}^iP_x^{0P} \times (1 - {}^iP_x^{0S} - {}^iP_x^{1S}) + {}^iP_x^{1P} \times (1 - {}^iP_x^{0S}) + {}^iP_x^{2P}.$$

We find the terms in $^iF_x^{fh}$ as follows. Note that we are working with Model 1 (Figure 1) and that $_sP_{x,t}^{gjk}$ denotes the probability that a person of sex s and genotype g who is in state gj at age x will be in state gk at age $x + t$. For brevity in the following, define $_sQ_{x,t}^{gjk} = 1 - {}_sP_{x,t}^{gjk}$.

$$
\begin{aligned}
{}^iP_x^{0P} \;=\; & \text{P[Neither parent has CRC at age } x + 30 \mid \text{Both alive at age 30]} \\
\;=\; & \frac{1}{2} \left\{ \omega_m^{g_i} \; {}_mQ_{30,x\wedge30}^{g_i01} \; \omega_f^1 \; {}_fQ_{30,x\wedge30}^{101} + \omega_m^1 \; {}_mQ_{30,x\wedge30}^{101} \; \omega_f^{g_i} \; {}_fQ_{30,x\wedge30}^{g_i01} \right\}.
\end{aligned}
$$

$$
\begin{aligned}
{}^iP_x^{1P} \;=\; & \text{P[One parent has CRC at age } x + 30 \mid \text{Both alive at age 30]} \\
\;=\; & \frac{1}{2} \left\{ (\nu_m^{g_i} + \omega_m^{g_i} \; {}_mP_{30,x\wedge30}^{g_i01}) \; \omega_f^1 \; {}_fQ_{30,x\wedge30}^{101} + \omega_m^{g_i} \; {}_mQ_{30,x\wedge30}^{g_i01} \; (\nu_f^1 + \omega_f^1 \; {}_fP_{30,x\wedge30}^{101}) \right. \\
& \left. + (\nu_m^1 + \omega_m^1 \; {}_mP_{30,x\wedge30}^{101}) \omega_f^{g_i} \; {}_fQ_{30,x\wedge30}^{g_i01} + \omega_m^1 \; {}_mQ_{30,x\wedge30}^{101} (\nu_f^{g_i} + \omega_f^{g_i} \; {}_fP_{30,x\wedge30}^{g_i01}) \right\}.
\end{aligned}
$$

$$
\begin{aligned}
{}^{i}P_x^{2P} &= \text{P[Both parents have CRC at age } x+30 \mid \text{ Both alive at age 30]} \\
&= \frac{1}{2} \left\{ (\nu_m^{g_i} + \omega_{mm}^{g_i} P_{30,x\wedge30}^{g_i01})(\nu_f^1 + \omega_{ff}^1 P_{30,x\wedge30}^{101}) \right. \\
&\quad \left. + (\nu_m^1 + \omega_{mm}^1 P_{30,x\wedge30}^{101})(\nu_f^{g_i} + \omega_{ff}^{g_i} P_{30,x\wedge30}^{g_i01}) \right\}.
\end{aligned}
$$

${}^{i}P_x^{0S}$ is the same as in Equation (21). With ${}^{i}P_{x\wedge60}$ defined as it was just before Equation (21), we also have:

$$
{}^{i}P_x^{1S} = \sum_{k=0}^{k=6} \text{P}[k] \; k \; {}^{i}P_{x\wedge60} \; (1 - {}^{i}P_{x\wedge60})^{k-1}.
$$

(b) Family History Scenario 3:

We must introduce endometrial cancer (EC). Under Scenario 3, at least one FDR must have EC and at least one must have CRC. Let ζ_f^g be the probability that a woman of genotype g is alive at age 30 but has had EC. Clearly $\zeta_m^g = 0$. Also, let ${}^{i}P_x^{0PC}$ and ${}^{i}P_x^{0PE}$ be the probabilities that neither parent has had EC or CRC, respectively, and modify the notation for similar probabilities similarly. Then:

$$
{}^{i}F_x^{fh} = {}^{i}P_x^{0P} \times {}^{i}P^{2SCE} + {}^{i}P_x^{1PC} \times (1 - {}^{i}P_x^{0SE}) + {}^{i}P_x^{1PE} \times (1 - {}^{i}P_x^{0SC}) + {}^{i}P_x^{2PCE}
$$

Again we use the notation of Model 1 (Figure 1). Ignoring other diseases, we have $\omega_s^g + \nu_s^g + \zeta_s^g = 1$ and approximately:

$$
\omega_s^{g_i} = \frac{{}_sP_{0,30}^{g_i00}}{{}_sP_{0,30}^{g_i00} + {}_sP_{0,30}^{g_i01} + {}_sP_{0,30}^{g_i02}}
$$

and similarly for ν_s^g and ζ_s^g. Then:

$$
\begin{aligned}
{}^{i}P_x^{0P} &= \text{P[Neither parent has CRC or EC at age } x+30 \mid \text{ Both alive at age 30]} \\
&= \frac{1}{2} \left\{ \omega_m^{g_i} \, {}_mQ_{30,x\wedge30}^{g_i01}) \, \omega_f^1 \, ({}_fQ_{30,x\wedge30}^{101} - {}_fP_{30,x\wedge30}^{102}) \right. \\
&\quad \left. + \omega_m^1 \, {}_mQ_{30,x\wedge30}^{101} \, \omega_f^{g_i} \, ({}_fQ_{30,x\wedge30}^{g_i01} - {}_fP_{30,x\wedge30}^{g_i02}) \right\}.
\end{aligned}
$$

$$
\begin{aligned}
{}^{i}P_x^{1PC} &= \text{P[One parent has CRC, neither has EC at age } x+30 \mid \text{ Both alive at age 30]} \\
&= \frac{1}{2} \left\{ (\nu_m^{g_i} + \omega_{mm}^{g_i} P_{30,x\wedge30}^{g_i01}) \, \omega_f^1 \, ({}_fQ_{30,x\wedge30}^{101} - {}_fP_{30,x\wedge30}^{102}) \right. \\
&\quad + \omega_m^{g_i} \, {}_mQ_{30,x\wedge30}^{g_i01} \, (\nu_f^1 + \omega_f^1 \, {}_fP_{30,x\wedge30}^{101}) \\
&\quad + (\nu_m^1 + \omega_m^1 \, {}_mP_{30,x\wedge30}^{101}) \, \omega_f^{g_i} \, ({}_fQ_{30,x\wedge30}^{g_i01} - {}_fP_{30,x\wedge30}^{g_i02}) \\
&\quad \left. + \omega_m^1 \, {}_mQ_{30,x\wedge30}^{101} \, (\nu_f^{g_i} + \omega_{ff}^{g_i} P_{30,x\wedge30}^{g_i01}) \right\}.
\end{aligned}
$$

$$
\begin{aligned}
{}^{i}P_{x}^{1PE} &= \text{P[One parent has EC, neither has CRC at age } x+30 \mid \text{Both alive at age 30]} \\
&= \frac{1}{2}\left\{ \omega_{m}^{g_{i}}\, {}_{m}Q_{30,x\wedge 30}^{g_{i}01}\,(\zeta_{f}^{1}+\omega_{ff}^{1}P_{30,x\wedge 30}^{102}) + \omega_{m}^{1}\, {}_{m}Q_{30,x\wedge 30}^{101}\,(\zeta_{f}^{g_{i}}+\omega_{f}^{g_{i}}\, {}_{f}P_{30,x\wedge 30}^{g_{i}02}) \right\}.
\end{aligned}
$$

$$
\begin{aligned}
{}^{i}P_{x}^{2PCE} &= \text{P[One parent has CRC, one EC at age } x+30 \mid \text{Both alive at age 30]} \\
&= \frac{1}{2}\left\{ (\nu_{m}^{g_{i}}+\omega_{mm}^{g_{i}}P_{30,x\wedge 30}^{g_{i}01})(\zeta_{f}^{1}+\omega_{ff}^{1}P_{30,x\wedge 30}^{102}) \right. \\
&\qquad \left. + (\nu_{m}^{1}+\omega_{mm}^{1}P_{30,x\wedge 30}^{101})\,(\zeta_{f}^{g_{i}}+\omega_{f}^{g_{i}}\, {}_{f}P_{30,x\wedge 30}^{g_{i}02}) \right\}.
\end{aligned}
$$

$$
\begin{aligned}
{}^{i}P_{x}^{0SC} &= \text{P[No sibling has CRC at age } x\text{]} \\
&= \sum_{k=0}^{k=6}\frac{\text{P}[k]}{2^{k}}\sum_{u=0}^{u=k}\binom{k}{u}\left(\frac{{}_{m}Q_{0,x\wedge 60}^{101}+{}_{m}Q_{0,x\wedge 60}^{g_{i}01}}{2}\right)^{u}\left(\frac{{}_{f}Q_{0,x\wedge 60}^{101}+{}_{f}Q_{0,x\wedge 60}^{g_{i}01}}{2}\right)^{k-u}.
\end{aligned}
$$

$$
\begin{aligned}
{}^{i}P_{x}^{0SE} &= \text{P[No sibling has EC at age } x\text{]} \\
&= \sum_{k=0}^{k=6}\frac{\text{P}[k]}{2^{k}}\sum_{u=0}^{u=k}\binom{k}{u}\left(\frac{{}_{f}Q_{0,x\wedge 60}^{102}+{}_{f}Q_{0,x\wedge 60}^{g_{i}02}}{2}\right)^{k-u}.
\end{aligned}
$$

$$
\begin{aligned}
{}^{i}P_{x}^{0SCE} &= \text{P[No sibling has CRC or EC at age } x\text{]} \\
&= \sum_{k=0}^{k=6}\frac{\text{P}[k]}{2^{k}}\sum_{u=0}^{u=k}\binom{k}{u}\left(\frac{{}_{m}Q_{0,x\wedge 60}^{102}+{}_{m}Q_{0,x\wedge 60}^{g_{i}02}}{2}\right)^{u} \\
&\qquad \times \left(1-\frac{{}_{f}P_{0,x\wedge 60}^{101}+{}_{f}P_{0,x\wedge 60}^{g_{i}01}+{}_{f}P_{0,x\wedge 60}^{102}+{}_{f}P_{0,x\wedge 60}^{g_{i}02}}{2}\right)^{k-u}.
\end{aligned}
$$

$$
\begin{aligned}
{}^{i}P_{x}^{2SCE} &= \text{P[At least one sibling has CRC and one has EC at age } x\text{]} \\
&= 1 - {}^{i}P_{x}^{0SC} - {}^{i}P_{x}^{0SE} + {}^{i}P_{x}^{0SCE}.
\end{aligned}
$$