HUNTINGTON'S DISEASE AND INSURANCE II: CRITICAL ILLNESS AND LIFE INSURANCE

By Cristina Gutiérrez and Angus Macdonald

ABSTRACT

In Part I we proposed a model of Huntington's disease (HD), a highly penetrant, dominantly inherited, fatal neurological disorder. Although it is a single-gene disorder, mutations are variable in their effects, depending on the number of times that the CAG trinucleotide is repeated in a certain region of the HD gene. The model covered: (a) rates of onset, depending on CAG repeat length as well as age; (b) post-onset rates of mortality; and (c) the distribution of CAG repeat lengths in the population. Using these, we study the critical illness and life insurance markets. We calculate premiums based on genetic test results that disclose the CAG repeat length, or more simply on a family history of HD. These vary widely with age and policy term; some are exceptionally high, but in a large number of cases cover could be offered within normal underwriting limits. We then consider the possible costs of adverse selection, in terms of increased premiums, under various possible moratoria on the use of genetic information, including family history. These are uniformly very small, because of the rarity of HD, but do show that the costs would be much larger in relative terms if family history could not be used in underwriting. We point out some difficulties involved in applying a moratorium that recognises simply a dichotomy between 'carriers' and 'non-carriers' of any mutation in a gene when these mutations are, in fact, very variable in their effects. These complexities suggest that restrictions on the disclosure, rather than on the use, of genetic information, if it became established as a principle, could deprive insurers of information needed for risk management even if not used in underwriting.

KEYWORDS

Critical Illness Insurance; Family History; Genetic Tests; Huntington's Disease; Life Insurance; Moratorium; 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

Huntington's disease (HD) is a dominantly inherited, fatal brain disease, caused by mutations in the HD gene. It has an extensive epidemiological literature, because it has been studied on the basis of family history for a long time before genetic tests were developed. What modern molecular genetics has revealed, however, is that the causative mutations are variable in their structure and in their effects. A certain region of the gene contains, in sequence, a variable number of CAG trinucleotides, each encoding the amino acid glutamine. This sequence is unstable and liable to expand in successive generations. The vast majority of people have fewer than 35 CAG repeats, and are not at risk of HD.

If there are 40 or more CAG repeats, onset of HD is practically certain. An 'intermediate allele' with 36–39 CAG repeats presents some risk of HD, but onset is not certain.

In Part I, we surveyed the literature of HD and proposed models for the features of HD most relevant for insurance applications:
(a) the rate of onset as a function of age and CAG repeat length;
(b) post-onset survival rates; and
(c) the distribution of CAG repeat lengths in the population.

In this Part, we apply these models to critical illness (CI) and life insurance. To avoid repetition, we refer the reader to Part I for all details of the genetical background and terminology, and of the HD model itself.

CI insurance (also known as dread disease or trauma insurance) is, in principle, the easiest contract to model, because payment is related to onset, and rates of onset are age-dependent, so a Markov model can be used to calculate premiums and reserves. We describe such a model in Section 2. However, onset of HD does not necessarily trigger a CI claim, as the criteria for disability might be reached only after the disease has progressed to a later stage. We allow for this with an accelerated lifetime model based on the post-onset mortality. Then, in Section 3 we obtain CI insurance premiums allowing for a genetic test that reveals the number of CAG repeats, or a family history of HD; and in Section 4.1, we model the possible costs of adverse selection in the CI insurance market, under various moratoria on the use of genetic test results or family history.

In Section 5 we propose a semi-Markov model for the life insurance market, and we consider premium ratings and adverse selection costs in Sections 6 and 7, respectively. Our conclusions are in Section 8.

2. A Critical Illness Insurance Model

2.1 *Model Specification*

We wish to address two questions:
(a) If insurers do have access to genetic information relating to HD, whether that is family history or a test result, how would premiums be affected? This is the question addressed by Smith (1998) in respect of life insurance.
(b) If insurers do not have access to such information, because of a moratorium on its use, what is the potential cost, to insurers or to other insured persons, of adverse selection?

The model in Figure 1 lets us address both questions. It is a continuous-time, discrete-state Markov model representing both the CI insurance-buying behaviour and the claims experience of a person with a given genotype denoted g_i.
(a) Figure 1 is a model of a person's life history in an insurance *market*. They start uninsured in state $i0$, and may buy a CI policy before or after having had a genetic test. If they are more likely to buy insurance after having had an adverse test result, adverse selection will appear and its cost can be measured.
(b) Premiums depending on genotype can be found simply by assuming that a person starts in one of the insured states ($i1$ or $i3$) on the policy inception date.

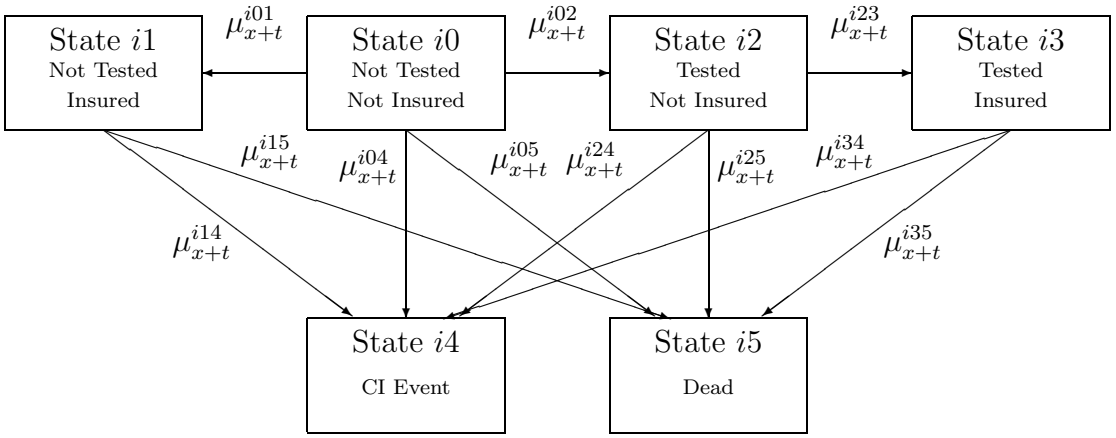

Figure 1: A Markov model of the insurance purchase and CI insurance events for a person with genotype g_i.

Figure 2 shows how the model is extended to the entire population allowing for genetic heterogeneity. For simplicity, it shows an aggregate model in which CAG repeat length is not considered, and a person at risk either is or is not a carrier of a HD mutation. For our later work, we have a separate sub-population for each of CAG repeat lengths 36 to 50, hence 102 states in 17 sub-populations. The proportions starting in the states labelled $i0$ are determined by the mutation frequencies modelled in Part I, Section 4.5, while intensities into the CI claiming states ($i4$) will depend on genotype.

As in Macdonald (2001) or Gutiérrez & Macdonald (2001) this model can represent many features of the problem:

(a) Mutation frequencies are the proportions in the starting states in each sub-population.
(b) The rate of insurance purchase in the 'not at risk' sub-population will determine the market size.
(c) The rate of genetic testing is explicit.
(d) Modified insurance-buying behaviour (both rate of purchase and amount purchased):
    (1) upon being heavily rated-up because of family history; or
    (2) when in possession of information that need not be disclosed
    is represented by the rates of purchase in the appropriate at-risk sub-populations.
(e) Underwriting classes are represented by sets of insured states within each of which the same premium rate will be charged.

2.2 *Rates of Onset of Non-Genetic Critical Illnesses*

There is no standard industry model for CI insurance in the U.K. or elsewhere. We use the model from Gutiérrez & Macdonald (2001) which is described briefly in the Appendix. This provides intensities of CI claims, and a force of mortality adjusted for CI claims, based on population data for males and females. Therefore, it is not suitable for direct application to insurance populations, but we are interested only in relative costs when

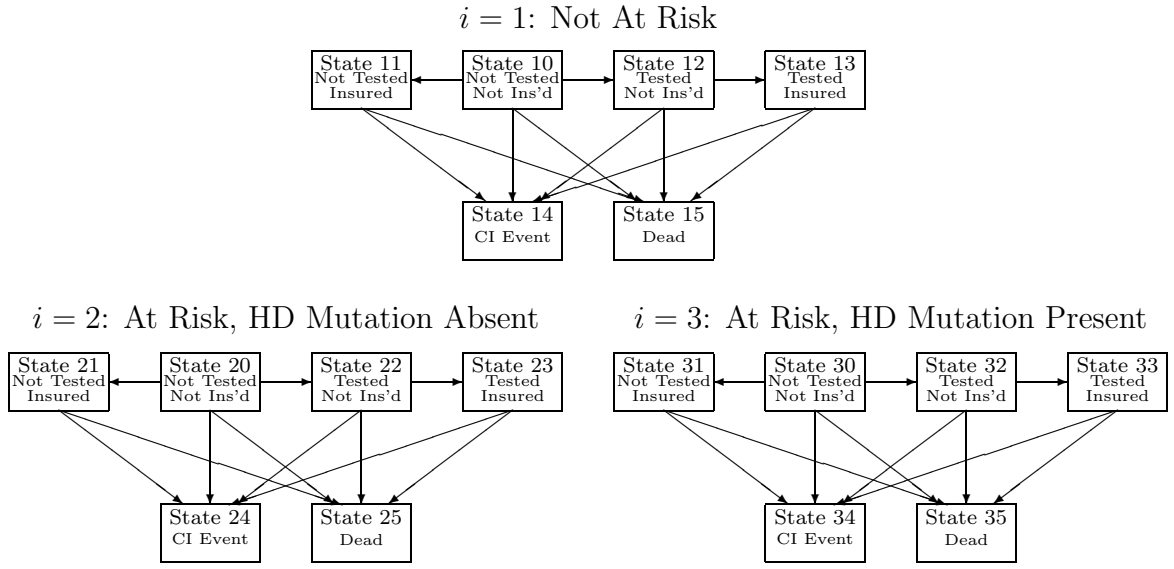

Figure 2: A Markov model of Critical Illness insurance allowing for family history of HD and genetic testing.

the CI claim rates are augmented by onset of HD. We made no attempt to remove deaths related to HD from the population mortality, as their impact there is negligible. Other CI models have been proposed by Dinani *et al.* (2000) and Macdonald, Waters & Wekwete (2003a, 2003b).

2.3 *The Timing of Critical Illness Insurance Claims*

Recall that Brinkman *et al.* (1997) defined age at onset as "... the first time a patient has either neurological or psychiatric symptoms that represented a permanent change from the normal state." This falls short of any criteria for a CI claim to be made, so any claim is likely to be some time after onset in the model. There are no studies of progression of HD that would allow us to specify meaningful states of health, one of which might represent a CI claim, and we are guided by two qualitative observations:

(a) Harper (1996) described three stages of HD (see Part I, Table 1) of which the second might and the third almost certainly would lead to a successful CI claim, which would therefore be roughly 5 or 10 years after onset.
(b) The ABI's genetics adviser suggested that a CI claim might follow about 10 years after onset, in the context of preparing an application to the Genetics and Insurance Committee in the U.K. (Professor J. A. Raeburn, personal communication).

It would be possible to translate these suggestions into a simple deterministic adjustment to the model. For example, we might assume that the claim will be delayed by a certain period of y years after onset, if that epoch falls within the policy term. Thus instead of paying a sum assured of £1 at onset at age $x + t$, we pay a sum of $\exp(-\delta y)$ if age $x + t + y$ falls within the policy term, or zero otherwise, where δ is the force of interest (and premiums would continue to be paid accordingly). This would entail a number of

minor assumptions, such as a person with HD being removed from the risk of other CI events between onset and claim payment, but there is a more serious objection here. The ability to foresee the future, from the time of onset, makes it impossible to calculate a meaningful premium under the equivalence principle when we consider the possible costs of adverse selection (Section 4.1).

We therefore take the suggestions in (a) and (b) above to be deterministic expressions of a probabilistic desire, and the accelerated lifetime model suggests itself as an obvious model. Given the distribution $F_X(x)$ of a random variable X representing a lifetime (here, the duration-dependent distribution of the lifetime after onset) we multiply the timescale by a constant $\phi \geq 1$ to obtain a new random variable Y such that:

$$F_Y(x) = \mathrm{P}[Y \leq x] = \mathrm{P}[X \leq \phi x] = F_X(\phi x). \tag{1}$$

See Collett (1994) for details. Clearly, the median of Y will be $1/\phi$ times the median of X, and this is at least consistent with a possible interpretation of the stages in Part I, Table 1. The corresponding relation between the intensities associated with X and Y is:

$$\mu_Y(x) = \phi \mu_X(\phi x). \tag{2}$$

We see that $\phi = 1.5$ and $\phi = 3$ correspond to claims being paid after 2/3 or 1/3, respectively, of the survival time after onset (on average) and this gives a simple probabilistic interpretation of Harper's three stages of HD. We will show premiums based on both of these assumptions, because it is quite possible that different insurers would apply different criteria in assessing a claim, influenced by a variety of factors including, perhaps, public image.

The models shown in Figures 1 and 2 correspond to $\phi = \infty$. They must be modified by the addition of a separate state representing onset of HD, with subsequent transition into the 'CI Event' state being possible, if $\phi < \infty$, but it is obvious how to do this. We assume that persons in this state (after onset but before claim payment) are still at risk of other CI events or death.

Note that the duration-dependent survival rates after onset mean that the model is semi-Markov if $\phi < \infty$. However, we can bring the computations back within a Markov framework, following Gui & Macdonald (2002). On transition from an insured, healthy state into an insured, HD state at age $x + t$, the insurer must set up the appropriate reserve, which we denote ${}_{t,0}V_x^{HD}$ (by definition, the duration at the moment of transition is zero). All the policy values in other states remain the same if the insurer 'pays out' the amount ${}_{t,0}V_x^{HD}$ as a 'sum assured' at age $x + t$, rather than setting up the reserve, collecting further premiums and paying out subsequent CI claims. These amounts depend on age only, so are adapted to a Markov framework. Of course, this only works for first moments.

2.4 *Numerical Methods*

Once the intensities in the model have all been fixed or estimated, we proceed by solving Kolmogorov's forward equations for occupancy probabilities, or Thiele's equations for expected present values (EPVs) of insurance cash-flows. With μ_x^{jk} the transition intensity between distinct states j and k and ${}_tp_x^{jk}$ the probability that a person in state

j at age x will be in state k at age $x+t$ (the occupancy probability), Kolmogorov's equations are:

$$\frac{\partial}{\partial t}{}_tp_x^{jk} = \sum_{l\neq k} {}_tp_x^{jl}\mu_{x+t}^{lk} - \sum_{l\neq k} {}_tp_x^{jk}\mu_{x+t}^{kl}. \tag{3}$$

(Note that we omit the i denoting genotype g_i for brevity here.) We can add insurance cash-flows to the model, with the convention that positive cash-flows are received by the insurer. If a continuous payment is made at rate b_x^j per annum while in state j at age x, or a lump sum of b_x^{jk} is made on transition from state j to state k at age x, Thiele's equations for the statewise prospective reserves ${}_tV_x^j$, at force of interest δ, at age $x+t$ are:

$$\frac{\partial}{\partial t}{}_tV_x^j = \delta_t V_x^j + b_{x+t}^j - \sum_{k\neq j} \mu_{x+t}^{jk}\left(b_{x+t}^{jk} + {}_tV_x^k - {}_tV_x^j\right). \tag{4}$$

These must be solved numerically. We used a fourth-order Runge-Kutta algorithm with step-size 0.0005 years. In all the calculations for this paper we used a force of interest of 0.05 per annum.

3. Critical Illness Insurance Underwriting

3.1 *Underwriting With Known CAG Repeat Length: 40–50 CAG Repeats*

The only individuals who might undergo a presymptomatic genetic test for HD are blood relatives of HD sufferers. Because of its rarity and its strongly Mendelian pattern of inheritance, there are no grounds for testing someone without symptoms who does not come from an affected family. In the absence of a genetic test, someone at risk carries the mutation with a probability that depends on their relatives. For example:

(a) a person with an unaffected parent and no affected siblings, but an affected grandparent, carries the mutation with a probability that diminishes as their parent grows older and remains unaffected; or

(b) a person with an affected parent, and who has children themselves, carries the mutation with a probability that diminishes as long as all their own children remain unaffected.

This is similar to the kind of information that might be used in life insurance underwriting, although the pedigrees used by clinical geneticists would usually be more detailed and more thoroughly checked.

Using the rates of onset from Part I, Section 4, we obtain level premium rates, payable continuously, for a CI insurance policy with a level sum assured, for various terms and entry ages. For 40–50 CAG repeats these are shown in Tables 1 and 2, expressed as a percentage of the standard premium rate. In Table 1 we assume that a claim is paid on reaching Stage 2 of Harper's progression, represented by $\phi = 3$ in the accelerated lifetime model. In Table 2 we assume that a claim is paid on reaching Stage 3 of Harper's progression, represented by $\phi = 1.5$ in the accelerated lifetime model.

(a) The stage at which a CI claim would be admitted is clearly very important. If we follow the assumption made by the ABI's genetics adviser, Table 2 might be more realistic, so our following remarks refer to it.

(b) Given that most CI insurers will decline risks rated above about 300% to 350% of the standard premium rate, we see that a substantial number of cases are within these limits, especially for shorter terms and older ages. For example, a person age 50 could always be offered terms, even with 50 CAG repeats. On the other hand, some premiums are extremely high, over 1,000% of the standard premium. The premiums are very dependent on the age and policy term. However, the premiums for older persons with a large number of CAG repeats are probably hypothetical as the chances of receiving such an application are small.

(c) The range of results shows that the CAG repeat length, if known, would be a most important risk factor. For example for a man age 20 seeking a 20-year policy, the premiums range from 103% to 2,674% of the standard rate. This has significant implications for policy on disclosure of genetic test results. For example, the man above could almost certainly get cover at standard rates if he disclosed a genetic test result, but since this would be an *adverse* result the insurer would not be allowed to use it.

(d) The differences between males and females are because of the different standard premium rates, our HD model is for males and females combined. However, these differences are relatively trivial.

(e) This kind of information does not become irrelevant if genetic test results may not be used in underwriting, as it still contributes to an understanding of the risk pool.

(f) We remarked in Section 2.3 that the delay between onset and payment of a CI claim might most simply be represented by a deterministic period. If we assumed a delay of 10 years, consistent with payment at Stage 3 of Harper's progression (Part I, Table 1), then no 10-year policies would have a premium higher than standard. Table 2 shows that this would be very misleading; with a high number of CAG repeats, even 10-year policies can attract a very high rating. This is a striking example of the errors that can be made if key stochastic features of a model are replaced by deterministic 'equivalents', even if they appear to be chosen sensibly.

3.2 *Underwriting With Known CAG Repeat Length: 36–39 CAG Repeats*

Tables 3 and 4 show the level CI premiums, as a percentage of the standard premiums, with intermediate alleles of 36–39 CAG repeats. In Table 3 we assume that a claim is paid on reaching Stage 2 of Harper's progression, represented by $\phi = 3$ in the accelerated lifetime model. In Table 4 we assume that a claim is paid on reaching Stage 3 of Harper's progression, represented by $\phi = 1.5$ in the accelerated lifetime model.

The results are quite striking. Only with 38 or more CAG repeats (in Table 3) or 39 CAG repeats (in Table 4) would there be any real question of a premium higher than standard, and then only for longer terms. Even if our extrapolated function is only right in its general features, we can conclude that access to CI insurance should be possible for anyone with a presymptomatic test result in the intermediate range, and mostly at standard rates. This is much more encouraging than might previously have been supposed.

Table 1: Level net premium for level CI cover for persons with a known HD mutation, as a percentage of the premium for standard risks. Claims arising at Stage 2 of Harper's progression (Part I, Table 1).

Sex of Applicant	Age at Entry (Years)	Policy Term (Years)	Premium as Percentage of Standard Number of CAG Repeats 40 %	41 %	42 %	43 %	44 %	45 %	46 %	47 %	48 %	49 %	50 %
Female	20	10	100	101	107	125	176	294	523	911	1,487	2,249	3,168
		20	111	141	221	389	682	1,118	1,690	2,373	3,125	3,907	4,680
		30	168	270	442	683	978	1,311	1,667	2,040	2,424	2,816	3,208
		40	246	361	502	658	828	1,013	1,215	1,437	1,676	1,928	2,185
	30	10	115	154	249	429	705	1,066	1,480	1,913	2,336	2,732	3,091
		20	183	308	515	800	1,137	1,492	1,837	2,154	2,434	2,676	2,882
		30	268	406	578	771	975	1,181	1,380	1,565	1,732	1,880	2,007
	40	10	180	274	405	559	718	871	1,011	1,136	1,246	1,341	1,424
		20	298	443	604	760	899	1,018	1,117	1,199	1,268	1,324	1,372
	50	10	247	330	414	494	567	630	685	733	774	809	840
Male	20	10	100	102	112	143	231	433	828	1,495	2,485	3,796	5,376
		20	115	159	275	519	942	1,574	2,403	3,392	4,482	5,614	6,736
		30	180	300	503	787	1,136	1,529	1,951	2,392	2,848	3,312	3,776
		40	242	356	495	651	821	1,007	1,211	1,435	1,677	1,931	2,190
	30	10	120	171	297	534	900	1,376	1,923	2,494	3,054	3,577	4,051
		20	192	328	556	869	1,239	1,629	2,009	2,358	2,666	2,932	3,159
		30	256	385	546	728	921	1,117	1,306	1,483	1,642	1,782	1,904
	40	10	179	273	403	556	714	866	1,005	1,129	1,238	1,333	1,415
		20	271	396	535	671	792	896	982	1,054	1,114	1,164	1,205
	50	10	216	281	348	411	468	518	562	599	632	660	684

Table 2: Level net premium for level CI cover for persons with a known HD mutation, as a percentage of the premium for standard risks. Claims arising at Stage 3 of Harper's progression (Part I, Table 1).

Sex of Applicant	Age at Entry (Years)	Policy Term (Years)	Premium as Percentage of Standard Number of CAG Repeats 40 %	41 %	42 %	43 %	44 %	45 %	46 %	47 %	48 %	49 %	50 %
Female	20	10	100	100	101	104	113	134	178	255	374	537	738
		20	102	109	130	176	266	412	621	888	1,200	1,536	1,877
		30	119	153	219	323	467	643	842	1,054	1,270	1,479	1,678
		40	157	214	291	386	492	607	728	853	979	1,104	1,223
	30	10	103	110	130	167	227	307	401	500	598	690	775
		20	123	163	236	346	482	632	779	915	1,034	1,136	1,221
		30	165	230	317	420	529	635	733	818	891	951	1,002
	40	10	115	134	160	192	225	257	287	314	338	360	378
		20	169	225	289	351	407	455	494	526	552	573	591
	50	10	133	153	173	192	209	224	238	249	260	269	276
Male	20	10	100	100	102	107	122	159	235	367	571	851	1,198
		20	103	113	143	211	340	552	855	1,242	1,693	2,181	2,674
		30	123	163	240	363	532	740	976	1,227	1,481	1,729	1,964
		40	156	211	288	381	487	601	723	848	975	1,100	1,220
	30	10	104	114	139	189	268	374	497	628	757	880	992
		20	125	169	250	370	520	684	846	995	1,126	1,238	1,331
		30	160	220	302	398	501	601	692	773	842	900	947
	40	10	115	134	160	191	224	256	286	313	337	358	376
		20	159	207	262	316	365	406	440	467	490	509	524
	50	10	126	141	157	172	186	198	208	218	226	233	239

Table 3: Level net premium for level CI cover for persons with a known 'intermediate allele' HD mutation (36–39 CAG repeats), as a percentage of the premium for standard risks. Claims arising at Stage 2 of Harper's progression (Part I, Table 1).

		Premium as Percentage of Standard							
		Females				Males			
Age at	Policy	No. of CAG Repeats				No. of CAG Repeats			
Entry	Term	36	37	38	39	36	37	38	39
(Years)	(Years)	%	%	%	%	%	%	%	%
20	10	100	100	100	100	100	100	100	100
	20	100	100	100	102	100	100	100	103
	30	100	101	104	120	100	101	105	124
	40	101	105	121	165	101	104	121	163
30	10	100	100	100	103	100	100	101	104
	20	100	101	105	125	100	101	106	128
	30	101	105	124	174	101	105	122	169
40	10	100	101	106	127	100	101	106	127
	20	101	107	130	191	101	106	126	178
50	10	101	107	130	177	101	106	123	160

We stress that our model for intermediate alleles is based on extrapolating the function fitted to 40–50 CAG repeats, and there were no data for 36–38 CAG repeats in Brinkman *et al.* (1997). However, the premiums there are sufficiently low that we think our conclusions are robust.

3.3 *Underwriting Based on Family History Only*

If an applicant of a given age has a family history of HD, but no genetic test result is known, a level premium is computed using the equivalence principle, where the expected present values (EPVs) of a unit benefit and a unit premium are weighted averages of the EPVs in respect of each possible genotype (including non-carriers), the weights being the probabilities of being alive and healthy at the given age. These are obtained as the occupancy probabilities in a model in which, at age 20, half of all persons at risk are non-carriers, and other half have CAG repeat lengths distributed according to the estimates in Part I, Section 4.5. From Part 1, Section 2.1, we assume that 18.75 per 100,000 persons are mutation carriers.

Table 5 shows the level CI premiums as a percentage of standard premiums. Again, the stage at which a claim would be made has the greatest bearing on the results. If it is at Stage 3 of Harper's progression (Part I, Table 1), then all but young persons seeking very long-term cover could be offered terms, and older persons could be offered rates close to standard. We would expect premiums to fall with age, because survival free of symptoms increases the chance of not being a carrier. However, this pattern of premiums may give rise to a lapse and re-entry risk. We have assumed that level premiums will be payable throughout the term of a policy, but for longer policy terms it might be advisable

Table 4: Level net premium for level CI cover for persons with a known 'intermediate allele' HD mutation (36–39 CAG repeats), as a percentage of the premium for standard risks. Claims arising at Stage 3 of Harper's progression (Part I, Table 1).

		Premium as Percentage of Standard							
		Females				Males			
Age at	Policy	No. of CAG Repeats				No. of CAG Repeats			
Entry	Term	36	37	38	39	36	37	38	39
(Years)	(Years)	%	%	%	%	%	%	%	%
20	10	100	100	100	100	100	100	100	100
	20	100	100	100	100	100	100	100	101
	30	100	100	101	105	100	100	101	106
	40	100	101	107	123	100	101	106	122
30	10	100	100	100	101	100	100	100	101
	20	100	100	101	106	100	100	101	107
	30	100	101	108	126	100	101	107	124
40	10	100	100	101	105	100	100	101	105
	20	100	102	109	130	100	102	108	125
50	10	100	102	106	117	100	101	105	113

Table 5: Level net premiums for CI cover as a percentage of the premium for standard risks, for persons with a family history of HD (affected parent or sibling).

		Claims arising			
		At Stage 2		At Stage 3	
Age at	Policy	of Part I, Table 1		of Part I, Table 1	
Entry	Term	Females	Males	Females	Males
(Years)	(Years)	%	%	%	%
20	10	263	380	132	156
	20	503	684	246	311
	30	480	549	289	323
	40	388	387	268	266
30	10	266	320	137	148
	20	335	358	195	204
	30	296	284	203	197
40	10	172	171	115	115
	20	202	188	142	136
50	10	128	122	107	105

in practice to charge a higher extra premium for a shorter term (as recommended by Brackenridge & Elder (1998) for life insurance).

It is often supposed that, even if a moratorium is imposed on the use of genetic test results, persons who disclose a 'clear' result will be underwritten as normal. However the fact that premiums given a family history in some cases (younger persons, up to 44 CAG repeats) lie within the range of premiums defined by individual CAG repeat lengths raises awkward questions about how far this 'common sense' approach might extend. We discuss this in detail in Section 8.3.

3.4 *Ascertainment Bias and Underwriting*

Ascertainment is the process by which persons or families come to the attention of researchers. It is usually presumed that ascertainment is incomplete, because persons with milder symptoms or later onset, or families with few affected members, will be more easily overlooked. The effect of under-ascertainment is to overstate estimates of rates of onset. It is regarded as a central problem in genetic epidemiology.

In Part I, Section 4.3, we referred to Falush *et al.* (2000), who found under-ascertainment with 40 CAG repeats, and extreme under-ascertainment with 36–38 CAG repeats. We could conclude from this that the figures in Tables 3 and 4 should be even lower than they are. However, an applicant for insurance with a *known* number of CAG repeats in the intermediate range is presumably a member of the 'ascertained' group, just because they have been tested, so it is more reasonable to say that the rates of onset, overstated because of low ascertainment, are in fact appropriate for the group of tested individuals.

In general, it is possible that under-ascertainment is less relevant in actuarial studies than it is in epidemiology, at least as far as genetic testing is concerned. It results in estimated rates of onset for a selected group rather than for the population, but it is just the same selected group who would ever approach an insurer with knowledge of a test result (whether or not they had to disclose it). This does assume that genetic tests for severe disorders are only ever taken for a reason, which is likely to be true in the U.K. but perhaps not everywhere else.

4. Critical Illness Insurance and Adverse Selection

4.1 *Adverse Selection and Moratoria on Genetic Information in Underwriting*

Macdonald (2001) and Gutiérrez & Macdonald (2001) used the model in Figure 1 to illustrate the potential costs arising from adverse selection if there were a moratorium on the use of genetic information in underwriting. The current position in the U.K. is that genetic test results may not be used for life insurance of up to £500,000 or other forms of insurance of up to £300,000, but family history can still be used. In some other countries (Sweden for example) family history also may not be used.

The methodology has been extensively discussed in the two papers cited above, so we will just summarise it here:

(a) An underwriting class is defined as a set of (insured) states in the model, within each of which the same rate of premium will be charged. In the absence of a moratorium the insurer would presumably partition the states according to homogeneity of risks, but a moratorium may force very different risks into the same underwriting class.

(b) The rate of premium within each underwriting class is calculated using the equivalence principle. However, level premiums cannot be used because they depend on age at purchase of insurance, so that a person in one of the insured states at age 50 (say) would pay a different rate of premium if they had entered it at age 30 than if they had entered it at age 40. This is not compatible with Thiele's equations (Equation (4)). The solution is to charge a rate of premium equal to the weighted average of the intensities from the insured states in the underwriting class to the CI claim state. If $\mathcal{C}$ is a set of states representing an underwriting class, the rate of premium is:

$$\rho^{\mathcal{C}}_{x+t} = \frac{\sum\limits_{ij\in\mathcal{C}} p_i \ {}_tp_x^{i0j} \ \mu_{x+t}^{ij4}}{\sum\limits_{ij\in\mathcal{C}} p_i \ {}_tp_x^{i0j}} \tag{5}$$

where p_i is the proportion who start in state $i0$ at age x. This device means that we must first solve Kolmogorov's equations to obtain the occupancy probabilities, then solve Thiele's equations using these rates of premium.

(c) The result of solving Thiele's equations is the EPV of the insurance loss conditional on being in any state. The weighted average of these EPVs in the starting states $(10, 20$ and so on), the weights being the occupancy probabilities at outset based on the mutation frequencies, is the EPV of the insurance loss in respect of the entire market.

(d) The insurer calculates the rates of premium assuming no adverse selection takes place. If this is borne out, the EPV of the loss is zero, because the equivalence principle has been correctly applied. If there is adverse selection, however, the EPV of the loss will be non-zero. This is the 'cost' of adverse selection. To recoup it, the insurer would have to increase premiums by:

$$\frac{\text{EPV of loss with adverse selection}}{\text{EPV of premiums payable with adverse selection}} \tag{6}$$

and this is the quantity that we take as our measure of the cost of adverse selection.

4.2 *Parameterisation*

We must choose intensities to represent 'normal' insurance purchase in each underwriting class, adverse selection, and genetic testing. To a large extent these are speculative:

(a) the CI insurance market is small but growing in the U.K., or hardly established in most other countries;

(b) little is known about how peoples' insurance-buying behaviour is changed by knowledge of genetic risks; and

(c) genetic testing is in its infancy.

We represent larger and smaller markets by constant rates of insurance purchase of 0.05 or 0.01 per annum, respectively, over the age range 20–60. That is, we assume that the market operates between these ages, and all CI policies have cover expiring at age 60. Clearly this could be refined if age-related rates of purchase were available. However,

persons offered a much higher premium because of a family history of HD might not be so likely to buy insurance; to cover the range of possibilities we suppose that in the larger market they buy insurance at rate 0.05, 0.025 or 0 per annum (the latter could also represent declinature on the part of the insurer) and in the smaller market we suppose that they do not buy insurance at all. This too could be refined, if we had a good model of the elasticity of demand for CI insurance. A moderate level of adverse selection is represented by intensities of 0.1 in the larger market, and 0.02 in the smaller market (twice the 'normal' rates). A severe level of adverse selection is represented by an intensity of 0.25 per annum; this is so high that most people will have bought insurance within a few years.

The prevalence of genetic testing for HD was discussed in Part I, Section 3.1; with testing having been widely available since about 1994, about 10–20% of at-risk persons have been tested. We assume that most testing takes place at relatively young ages, in the model at ages 20–40. A rate of 0.014 per annum over these ages means that about 10% would be tested after 8 years (and 24% after 20 years) thus representing a modest rate of testing. We take this as our baseline, and compare it with:

(a) the same rate of testing of 0.014 per annum, but extending over ages 20–60; and
(b) a rate of testing of 0.035 per annum, over ages 20–40.

On balance, we believe that the baseline rate of testing, low and in line with observed prevalence, is most appropriate. Perhaps the most plausible reason for higher rates of testing in future might be the development of effective treatments for HD, which would offset any increase in costs, although in ways impossible to predict. It is the combination of high rates of testing and no treatment at all that seems least plausible.

4.3 *Moratoria on Genetic Test Results*

A moratorium may forbid the use of all genetic test results, or (perhaps more likely) allow the use of negative test results that would allow someone to be offered the standard premium rate. In either case there will be two underwriting classes: one including everyone charged the standard premium, and another whose members will be offered a premium based on family history.

Tables 6 and 7 show the percentage premium increases needed to recoup the cost of moderate adverse selection in respect of undisclosed genetic tests for HD mutations, assumes claims arise at Stage 2 or Stage 3, respectively, of Harper's progression (Part I, Table 1). They are all small (less than 0.02%) but there is a pattern:

(a) They increase substantially, the less people at risk take up the offer of insurance at premiums increased because of family history. It is probably most realistic to suppose that such persons do not buy insurance, because they often might not be offered it.
(b) The costs are much more substantial in the smaller market.
(c) The costs are very slightly higher under the moratorium on adverse test results only. As in Gutiérrez & Macdonald (2001), this is because the premium rate in the 'family history' underwriting class is now more heavily weighted by mutation carriers.
(d) Extending the period during which genetic testing is assumed to take place has little effect. Mainly this is because of the high penetrance of HD mutations; relatively few mutation carriers will ever undergo a presymptomatic test after age 40. Increasing

Table 6: Percentage increases in CI insurance premium rates arising from moderate adverse selection. Moratoria on the use of genetic test results, family history underwriting still allowed. CI market operating between ages 20 and 60. Claims arising at Stage 2 of Harper's progression (Part I, Table 1).

Size of Market	Insurance Purchasing of At-Risk Individuals	Rate of Testing	Age Range of Testing	Moratorium on Using All test results		Moratorium on Using Adverse test results	
				Females	Males	Females	Males
				%	%	%	%
Large	Normal	0.014	20–40	0.002	0.002	0.002	0.002
	Half	0.014	20–40	0.005	0.004	0.004	0.004
	Nil	0.014	20–40	0.009	0.008	0.008	0.008
Small	Nil	0.014	20–40	0.008	0.008	0.008	0.007
Large	Normal	0.014	20–60	0.002	0.002	0.002	0.002
	Half	0.014	20–60	0.005	0.004	0.004	0.004
	Nil	0.014	20–60	0.009	0.009	0.009	0.008
Small	Nil	0.014	20–60	0.009	0.008	0.008	0.007
Large	Normal	0.035	20–40	0.005	0.005	0.004	0.004
	Half	0.035	20–40	0.010	0.010	0.009	0.009
	Nil	0.035	20–40	0.020	0.019	0.017	0.016
Small	Nil	0.035	20–40	0.019	0.017	0.015	0.014

Table 7: Percentage increases in CI insurance premium rates arising from moderate adverse selection. Moratoria on the use of genetic test results, family history underwriting still allowed. CI market operating between ages 20 and 60. Claims arising at Stage 3 of Harper's progression (Part I, Table 1).

Size of Market	Insurance Purchasing of At-Risk Individuals	Rate of Testing	Age Range of Testing	Moratorium on Using All test results		Moratorium on Using Adverse test results	
				Females	Males	Females	Males
				%	%	%	%
Large	Normal	0.014	20–40	0.001	0.001	0.001	0.001
	Half	0.014	20–40	0.002	0.002	0.002	0.002
	Nil	0.014	20–40	0.005	0.004	0.004	0.004
Small	Nil	0.014	20–40	0.004	0.003	0.003	0.003
Large	Normal	0.014	20–60	0.001	0.001	0.001	0.001
	Half	0.014	20–60	0.002	0.002	0.002	0.002
	Nil	0.014	20–60	0.004	0.004	0.004	0.004
Small	Nil	0.014	20–60	0.004	0.003	0.003	0.003
Large	Normal	0.035	20–40	0.003	0.003	0.002	0.002
	Half	0.035	20–40	0.006	0.005	0.005	0.004
	Nil	0.035	20–40	0.010	0.010	0.008	0.008
Small	Nil	0.035	20–40	0.009	0.008	0.007	0.006

Table 8: Percentage increases in standard premium rates for CI insurance arising from new underwriting classes, and in all premiums arising from moderate or 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. Claims arising at Stage 2 of Harper's progression (Part I, Table 1).

	OR Premium Increases Arising From New		Premium Increases Arising From		Premium Increases Arising From	
Size of	Underwriting Classes		Moderate Adverse Selection		Severe Adverse Selection	
Market	Females	Males	Females	Males		
	%	%	%	%	%	%
Large	0.069	0.064	0.034	0.032	0.066	0.062
Small	0.057	0.051	0.055	0.051	0.350	0.308

the rate at which testing takes place, however, increases the costs substantially.

The small magnitude of these costs, even assuming very severe adverse selection, cannot by itself be taken to mean that all genetic testing is irrelevent for CI insurance. HD is one, quite rare, member of the universe of genetic disorders, and we would have to complete a program of modelling the others before we could reach any such conclusion.

4.4 *Moratoria on Family History and Genetic Test Results*

A moratorium on family history as well as genetic tests has two results:

(a) Those who were previously in higher-risk underwriting classes can now buy insurance in the normal way; that is, at the same rate as persons not at risk. This will increase premiums, but it is not adverse selection. We call this the cost of defining new underwriting classes. Note that it does not depend on the rate at which at-risk persons previously bought insurance.

(b) However, these same peoples' knowledge of their genetic risk might lead them to buy insurance at a rate higher than normal, so there might be further premium increases for that reason. Moreover, this group now includes those who have a family history but who have not been tested, which means that the rate of genetic testing is relevant only to the extent that testing removes non-carriers from the at-risk group.

Tables 8 and 9 show these two levels of premium increases separately, assuming claims arise at Stage 2 or Stage 3, respectively, of Harper's progression (Part I, Table 1). Both moderate and severe levels of adverse selection are shown (see Section 4.6). The rate of genetic testing was 0.014 per annum with moderate adverse selection, and 0.035 per annum with severe adverse selection, between ages 20 and 40; other assumptions made almost no difference and we omit them. In the large market, the cost of the new underwriting class is high compared with the previous costs of a moratorium on adverse selection, approaching 0.07%, in Table 8, and the cost of further adverse selection is about half as much again. In the smaller market, the cost of additional (severe) adverse selection is higher, exceeding 0.05%. But in absolute terms these increases are negligible.

Table 9: Percentage increases in standard premium rates for CI insurance arising from new underwriting classes, and in all premiums arising from moderate or 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. Claims arising at Stage 3 of Harper's progression (Part I, Table 1).

	OR Premium Increases Arising From New		Premium Increases Arising From		Premium Increases Arising From	
Size of	Underwriting Classes		Moderate Adverse Selection		Severe Adverse Selection	
Market	Females	Males	Females	Males		
	%	%	%	%	%	%
Large	0.038	0.035	0.021	0.019	0.041	0.039
Small	0.030	0.026	0.031	0.028	0.225	0.207

Table 10: Percentage increases in CI insurance 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. Claims arising at Stage 2 of Harper's progression (Part I, Table 1).

	Insurance Purchasing		Age	Moratorium on Using			
Size of	of At-Risk	Rate of	Range of	All test results		Adverse test results	
Market	Individuals	Testing	Testing	Females	Males	Females	Males
				%	%	%	%
Large	Normal	0.035	20-40	0.011	0.011	0.010	0.010
	Half	0.035	20-40	0.018	0.017	0.016	0.015
	Nil	0.035	20-40	0.029	0.027	0.025	0.024
Small	Nil	0.035	20-40	0.090	0.082	0.078	0.071

4.5 *Higher Sums Assured*

All the results shown here suppose that 'adverse selectors' buy the same amount of insurance as normal; they do not insure themselves for above-average amounts. This possibility is the second component of adverse selection. It is easy to see that in this case, the premium increases are proportionate to the multiple of the average sum assured taken out by 'adverse selectors', so for brevity we omit the tables.

4.6 *More Extreme Adverse Selection*

Our 'severe' rate of adverse selection, 0.25 per annum, is deliberately extreme. It implies that nearly all people in a position to exploit non-disclosure will do so within a few years. Arguably this is unlikely, but to show the worst that might be expected we combine severe adverse selection with a higher rate of genetic testing (0.035 per annum up to age 40). Tables 8 and 9 included the results for a moratorium on family history, and Tables 10 and 11 show the costs of moratoria on genetic test results, assuming CI

Table 11: Percentage increases in CI insurance 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. Claims arising at Stage 3 of Harper's progression (Part I, Table 1).

Size of Market	Insurance Purchasing of At-Risk Individuals	Rate of Testing	Age Range of Testing	Moratorium on Using All test results		Moratorium on Using Adverse test results	
				Females	Males	Females	Males
				%	%	%	%
Large	Normal	0.035	20-40	0.007	0.006	0.006	0.006
	Half	0.035	20-40	0.010	0.010	0.009	0.009
	Nil	0.035	20-40	0.016	0.015	0.014	0.013
Small	Nil	0.035	20-40	0.049	0.045	0.042	0.039

claims to be paid at Stage 2 or Stage 3, respectively, of Part I, Table 1:

(a) The worst case under a moratorium on genetic test results, in the smaller market, would be premium increases of about 0.05% or 0.1%, depending on the stage at which a claim would be payable. Of course this is much larger than before, but as an extreme upper limit it is very modest.

(b) With a moratorium on family history, premiums in the smaller market could increase by as much as 0.35% because of adverse selection. This is because persons at risk because of a family history can buy insurance at the new standard rates and are assumed to do so at rate 0.25 per annum, which is so high that not many are tested before buying insurance. When family history underwriting is allowed, an adverse test result is a prerequisite for adverse selection to occur.

Any of these numbers could be increased if adverse selection extended to sums assured higher than average, in proportion to the excess. However we believe that they show that only in the most extreme circumstances could adverse selection in respect of HD alone have any impact on a CI insurance market of any reasonable size.

5. A Life Insurance Model

Figure 3 shows a semi-Markov model of a life insurance market, similar to that in Gui & Macdonald (2002). The intensities $\mu^{i46}_{x+t,d}$ and $\mu^{i56}_{x+t,d}$, representing post-onset mortality, depend on both age and duration. In fact the mortality rates found in Part I, Section 4.4 depended on duration alone, but Wilkie (2000) pointed out that a duration-dependent post-onset rate of mortality may be lower than the usual age-related rate of mortality, especially at high ages. We therefore take these intensities to be the greater of those based on Equations (3) to (5) of Part I (as appropriate) or those of English Life Tables No. 15.

State $i6$ is labelled 'Dead, or HD and Not Insured'. If an uninsured person develops symptoms of HD, they become uninsurable, so for simplicity we represent this by transition into the only absorbing state in the model.

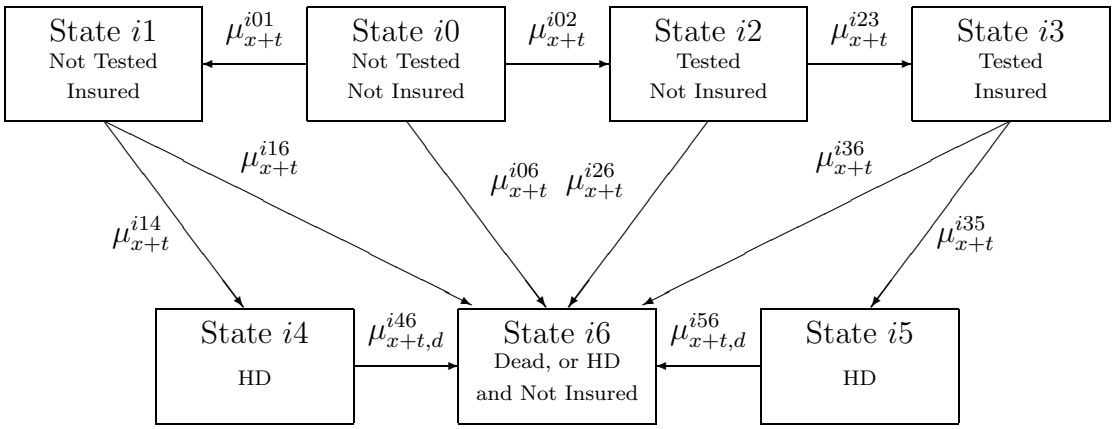

Figure 3: A semi-Markov model of insurance purchase and life insurance events for a person with genotype g_i.

It is necessary to have separate states representing onset of HD from tested and untested insured states, because:
(a) they could be in different underwriting classes, therefore contributing to the calculation of different premium rates when we model adverse selection (Section 7); and
(b) part of adverse selection after receiving an adverse test result could be to take out a larger sum assured.

Although the model is semi-Markov, EPVs can be calculated in a Markov framework, hence Thiele's equations can be used, in the same way as described in Section 2.3.

6. Life Insurance Underwriting

6.1 *Underwriting With Known CAG Repeat Length: 40–50 CAG Repeats*

Using the model of Figure 1 in Part I, we can easily write down the EPVs of a unit sum assured and a unit annual premium payable continuously while alive, between ages x and $x+n$, for a person with genotype g_i:

$$\text{EPV[Benefit]} = \int_0^n e^{-\delta t} \int_0^t {}_sp_x^{i00}\mu_{x+s}^{i01}\ {}_{t-s}p_{x+s,0}^{i11}\mu_{x+t,t-s}^{i12}\ ds\ dt \tag{7}$$

$$\text{EPV[Premium]} = \int_0^n e^{-\delta t}\left({}_tp_x^{i00} + \int_0^t {}_sp_x^{i00}\mu_{x+s}^{i01}\ {}_{t-s}p_{x+s,0}^{i11}\ ds\right)\ dt. \tag{8}$$

Alternatively, we can obtain these EPVs from the solutions of Thiele's equations in respect of the insured states in the model shown in Figure 3 (extended appropriately to different genotypes as in Figure 2).

Table 12: Level net premium for level life insurance cover for persons with a known HD mutation, with 40–50 CAG repeats, as a percentage of the premium for standard risks.

Sex of Applicant	Age at Entry (Years)	Policy Term (Years)	Premium as Percentage of Standard Number of CAG Repeats 40 %	41 %	42 %	43 %	44 %	45 %	46 %	47 %	48 %	49 %	50 %
Female	20	10	100	100	100	102	105	114	132	166	219	293	387
		20	101	105	117	147	209	315	475	690	951	1,242	1,545
		30	112	138	192	288	432	624	853	1,107	1,371	1,631	1,877
		40	141	192	272	381	513	664	825	990	1,154	1,310	1,456
	30	10	101	106	117	139	175	225	285	349	414	477	535
		20	116	146	208	307	438	588	741	885	1,014	1,125	1,220
		30	147	206	294	408	535	662	780	884	972	1,044	1,104
	40	10	106	114	126	141	158	174	190	205	219	231	242
		20	142	181	229	279	326	366	401	430	454	474	491
	50	10	108	114	120	126	132	137	142	147	151	155	158
Male	20	10	100	100	100	101	102	105	111	123	142	169	203
		20	101	102	108	121	148	196	269	367	487	621	760
		30	106	118	146	195	270	369	490	624	764	902	1,032
		40	119	144	186	244	316	399	488	581	672	760	842
	30	10	101	103	108	120	139	165	196	230	264	298	329
		20	109	126	161	219	295	384	475	561	638	705	762
		30	124	155	205	270	344	419	490	552	604	648	684
	40	10	103	107	113	121	130	138	147	155	163	170	176
		20	120	140	165	192	218	241	261	278	292	304	314
	50	10	102	104	106	108	109	111	113	114	116	117	119

Table 13: Level net premium for level life insurance cover for persons with a known 'intermediate allele' HD mutation, as a percentage of the premium for standard risks.

		Premium as Percentage of Standard							
		Females				Males			
Age at	Policy	No. of CAG Repeats				No. of CAG Repeats			
Entry	Term	36	37	38	39	36	37	38	39
(Years)	(Years)	%	%	%	%	%	%	%	%
20	10	100	100	100	100	100	100	100	100
	20	100	100	100	100	100	100	100	100
	30	100	100	101	103	100	100	100	101
	40	100	101	104	114	100	100	101	106
30	10	100	100	100	100	100	100	100	100
	20	100	100	101	104	100	100	100	102
	30	100	101	104	117	100	100	102	108
40	10	100	100	100	102	100	100	100	101
	20	100	101	104	116	100	100	102	108
50	10	100	100	101	104	100	100	100	101

Table 12 shows level premiums for a level amount of life insurance, expressed as a percentage of the premium for standard risks.

6.2 *Underwriting With Known CAG Repeat Length: 36–39 CAG Repeats*

Table 13 shows level premiums for a level amount of life insurance, expressed as a percentage of the premium for standard risks, given 36–39 CAG repeats. The highest is 117%, for a female age 30 seeking 30 years' cover. This table suggests that people with intermediate alleles might be offered life insurance at standard rates in all cases.

6.3 *Underwriting Based on Family History Only*

Brackenridge & Elder (1998) provide guidelines for the selection of risks and premium ratings given *a priori* genetic risk of 50%. A summary of the guidelines on the basis of family risk but not genetic testing is presented in Table 14. They can be modified if the last forbear known to be heterozygous was a grandparent, implying a genetic risk of 25%, or a great-grandparent, giving a genetic risk of 12.5%. We have included in the table, for comparison, the resulting premiums as a percentage of standard rates, for males.

The guidelines in Brackenridge & Elder (1998) in the case that a genetic test result is available are based on the earlier tests for markers, not direct analysis of the HD gene; their references are Gusella *et al.* (1983) and Brock *et al.* (1989). They are therefore based on an assessment of the risk of heterozygosity, given an adverse test result, that is now out of date. Brackenridge & Elder (1998) is without doubt the authority on medical underwriting, so this is an interesting sign of the great speed at which medical genetics is developing.

Smith (1998), whose model we described in Part I, gave sample extra premiums for term and endowment assurances, for mutation carriers and (by a simple Bayesian

Table 14: Guidelines for rating the risk of Huntington's disease in life insurance, on the basis of family risk but not genetic testing. Source: Brackenridge & Elder (1998) .

Information in Proposal	Suggested Rating	Premium (male) as Percentage of Standard Rates %
Symptoms present	Uninsurable	n/a
Age under 21	Decline	n/a
Age 21–35	+7 *per mille* for 10 years	901 (age 21)
Age 36–45	+5 *per mille* for 10 years	378 (age 36)
Age 46–55	+5 *per mille* until age 55	202 (age 46)
Age 56 and over	Standard	

Table 15: Examples of percentage extra premiums for term assurance, for a male with an affected father. Source: Smith (1998).

	Mutation Carrier		Asymptomatic At-Risk	
Age	Term $\leq$ 10 Yrs	Term $>$ 10 Yrs	Term $\leq$ 10 Yrs	Term $>$ 10 Yrs
< 35	+325%	Decline	+125%	+200%
35–45	+175%	+275%	+50%	+75%
44–55	+50%	+100%	Standard	Standard
> 55	Standard	+25%	Standard	Standard

argument) for asymptomatic at-risk individuals. The latter were insurable at all ages and terms, sometimes with only a modest extra premium; the former were not always insurable, assuming a 400% extra premium limit. An example in respect of term assurance is shown in Table 15. Under endowment assurances, coverage could always be offered, at an extra premium not exceeding about 40%. Overall, Smith's conclusions were that life insurance could be provided to people at risk of HD, or even sometimes to mutation carriers, at lower cost than had often previously been assumed (as exemplified by five different underwriting manuals).

Table 16 shows level premiums for a level amount of life insurance, expressed as a percentage of the premium for standard risks, in respect of an applicant with an affected parent or sibling. These were calculated in the same way as the CI insurance premiums in Section 3.3.

(a) They are broadly consistent with the results in Table 15 from Smith (1998) (note that these were expressed as percentage extra premiums).
(b) As with CI insurance premiums, they are sometimes (younger lives, up to 45 CAG repeats) higher than the premiums based on a genetic test result.
(c) They are very much lower than the premiums implied by the suggested ratings in Brackenridge & Elder (1998) (Table 14).

Table 16: Level net premiums for level life insurance cover as percentage of the level premium for standard risks, for persons with a family history of HD (affected parent or sibling).

Age at Entry (Years)	Policy Term (Years)	Females %	Males %
20	10	114	105
	20	211	150
	30	297	202
	40	293	203
30	10	122	112
	20	187	151
	30	208	160
40	10	107	103
	20	130	115
50	10	102	101

7. Life Insurance and Adverse Selection

7.1 *A Model of the Life Insurance Market*

Figure 3 shows a model of a life insurance market, for a person with genotype g_i. We would like to use this to illustrate the potential costs of adverse selection, as we did for CI insurance in Section 4.1. However, the computations are complicated by the duration-dependence of mortality rates after onset of HD. Note the following:

(a) There are separate states representing onset of HD for tested and untested insured persons, because the benefit is payable on exit from either of these states and could be higher for tested persons.

(b) We assume that people who have not bought insurance before HD appears cannot buy it afterwards, so the intensities μ^{i06}_{x+t} and μ^{i26}_{x+t} include death and onset of HD.

In Section 4.1, the premium rate payable by an insured person at age $x+t$ was given by Equation (5). Here we must allow for the duration-dependent survival rates after onset. For simplicity, suppose that genotype g_i by itself is an underwriting class. The rate of premium we want is that which would be charged in the absence of adverse selection, so assume there is no genetic testing ($\mu^{i02}_{x+t} = 0$). Then the weighted average intensity into the dead state from the two insured states $i1$ and $i4$, denoted $\mu^{\mathcal{C}}_{x+t}$, is:

$$\mu^{\mathcal{C}}_{x+t} = \frac{{}_tp^{i01}_x \mu^{i16}_{x+t} + \int_0^t {}_{t,z}p^{i04}_x \mu^{i46}_{x+t,z}\, dz}{{}_tp^{i01}_x + \int_0^t {}_{t,z}p^{i04}_x\, dz}. \tag{9}$$

The superscript '$\mathcal{C}$' in $\mu^{\mathcal{C}}_{x+t}$ indicates that this is a rate in respect of an underwriting class; if genotypes are combined into underwriting classes, we would extend the weighted average accordingly as in Equation (5). These are the rates of premium we use when

Table 17: Percentage increases in premium rates for life insurance arising from moderate adverse selection. Moratoria on the use of genetic test results, family history underwriting still allowed. Life insurance market operating between ages 20 and 60.

Size of Market	Insurance Purchasing of At-Risk Individuals	Rate of Testing	Age Range of Testing	Moratorium on Using All test results Females %	All test results Males %	Adverse test results Females %	Adverse test results Males %
Large	Normal	0.014	20–40	0.001	0.001	0.001	0.001
	Half	0.014	20–40	0.002	0.001	0.002	0.001
	Nil	0.014	20–40	0.004	0.002	0.004	0.002
Small	Nil	0.014	20–40	0.003	0.002	0.003	0.002
Large	Normal	0.014	20–60	0.001	0.001	0.001	0.001
	Half	0.014	20–60	0.002	0.001	0.002	0.001
	Nil	0.014	20–60	0.004	0.002	0.004	0.002
Small	Nil	0.014	20–60	0.003	0.002	0.002	0.001
Large	Normal	0.035	20–40	0.003	0.002	0.003	0.001
	Half	0.035	20–40	0.006	0.003	0.005	0.003
	Nil	0.035	20–40	0.010	0.005	0.008	0.004
Small	Nil	0.035	20–40	0.008	0.004	0.005	0.003

adverse selection is introduced, because they are the rates an insurer would use in a 'normal' market.

We use the same device as in Section 2.3 to bring the calculations back within a Markov framework, by 'paying' a 'sum assured' equal to the policy value on entering an HD state from an insured state. The statewise policy value ${}_{t,0}V_x^{i4}$ (and likewise ${}_{t,0}V_x^{i5}$), for policy term n years and sum assured £1, is:

$$ {}_{t,0}V_x^{i4} = \int_0^{n-t} e^{-\delta s} {}_sp_{x+t,0}^{i44}(\mu_{x+t+s,s}^{i46} - \mu_{x+t+s}^{\mathcal{C}})\, ds. \tag{10} $$

7.2 *Parameterisation*

The parameterisation of the model, in respect of insurance purchase and genetic testing, is exactly the same as in Section 4.2. Recall that 'moderate' adverse selection means insurance is purchased at twice the 'normal' rate, and 'severe' adverse selection means that the rate of purchase is 0.25 per annum. Clearly, we might regard the larger market as more relevant for life insurance.

7.3 *Moratoria on Genetic Test Results*

Table 17 shows the percentage increases in premium rates arising from moderate adverse selection, under moratoria covering genetic test results but not family history. They can truly be described as negligible; the insurer who can calculate a premium to

Table 18: Percentage increases in standard premium rates for life insurance arising from new underwriting classes, and in all premiums arising from moderate or severe adverse selection, following a moratorium on the use of all genetic test results and family history. Life insurance market operating between ages 20 and 60.

	OR Premium Increases Arising From New		Premium Increases Arising From		Premium Increases Arising From	
Size of	Underwriting Classes		Moderate Adverse Selection		Severe Adverse Selection	
Market	Females	Males	Females	Males		
	%	%	%	%	%	%
Large	0.042	0.022	0.024	0.014	0.049	0.029
Small	0.032	0.015	0.032	0.018	0.256	0.148

an accuracy of 0.01%, which is the largest cost in the table, does not exist. Nevertheless we should hesitate to say that adverse selection in respect of genetic disorders *in toto* is negligible, as HD is just one (admittedly severe) disorder among several.

7.4 *Moratoria on Family History and Genetic Test Results*

Table 18 shows the percentage increases in standard premium rates arising from the creation of a single underwriting class, and in all premiums arising from moderate or severe adverse selection (parameterised as in Section 4.4), following a moratorium on the use of all genetic test results and family history. The premium increases are now larger but hardly more significant, up to about 0.04%, even if there were no adverse selection, and over 0.06% if there was (for females).

7.5 *More Extreme Adverse Selection*

As for CI insurance, we show in Table 19 the effect of severe adverse selection and the higher rate of genetic testing on the costs under moratoria on genetic test results. The costs under a moratorium on family history were included in Table 18. These support our previous conclusion: only in the most extreme (and unlikely) circumstances, and in the smaller market, would these costs even be discernable.

8. Conclusions

8.1 *Critical Illness Insurance*

The definition of onset in Brinkman *et al.* (1997) is based on the earliest indications of HD, whereas a CI claim is likely to arise some time later. Harper (1996) provides a detailed description of three stages of HD, each corresponding roughly to 5 years within an overall 15-year survival. This suggested an accelerated lifetime model, applied to the post-onset survival rates, to represent claim payments on entering Stage 2 or Stage 3 of Harper's progression. These were used in the CI insurance model from Gutiérrez & Macdonald (2001). We found the following:

Table 19: Percentage increases in premium rates for life insurance arising from severe adverse selection. Moratoria on the use of genetic test results, family history underwriting still allowed. Life insurance market operating between ages 20 and 60.

Size of Market	Insurance Purchasing of At-Risk Individuals	Rate of Testing	Age Range of Testing	Moratorium on Using All test results Females %	All test results Males %	Adverse test results Females %	Adverse test results Males %
Large	Normal	0.035	20-40	0.008	0.004	0.007	0.004
	Half	0.035	20-40	0.011	0.006	0.010	0.006
	Nil	0.035	20-40	0.016	0.009	0.014	0.008
Small	Nil	0.035	20-40	0.050	0.028	0.043	0.024

(a) CI premium rates, as a proportion of standard rates, varied greatly with age and policy term, as well as CAG repeat length. There were quite large differences between males and females. In many cases the premiums fell within the limits currently regarded as insurable by UK companies.

(b) In all cases persons with an 'intermediate' allele (36–39 CAG repeats) could be offered CI cover, mostly at standard rates. In practice such test results might be ignored.

(c) Premiums based on family history alone (affected parent or sibling) fell naturally in the range of those for individual CAG repeat lengths, with the following features:
   (1) If a CI claim corresponded to Stage 3 in Harper's progression, CI cover could be offered in all cases, though younger persons would be near the limit of current UK underwriting practice.
   (2) Level premiums dropped sharply for older persons (over 40), leading to a possible lapse and re-entry risk.
   (3) A younger person, carrying a mutation with a low CAG repeat number, could often be offered a premium lower than that based on family history, but under most moratoria this would require the insurer to underwrite on the basis of a disclosed adverse test result, which would be disallowed.
   (4) For high CAG repeat numbers, premiums for 10-year policies were very much higher than standard. This showed that it could be quite misleading to treat the delay between onset and claiming in a deterministic way, for example by assuming that claims were paid 10 years after onset.

(d) The potential costs of adverse selection were very small, because of the rarity of HD mutations. In absolute terms they appeared negligible, but of course this is only one genetic disorder among several. If family history underwriting was still allowed, premium increases did not exceed 0.1% even assuming a rather high rate of genetic testing and extreme adverse selection. Only in a much smaller market than we assumed, or if adverse selection included gross over-insurance, would premiums be affected noticeably. A moratorium on the use of family history would increase premiums by much more even in the absence of adverse selection, but the absolute

increases are still very small.

8.2 *Life Insurance*

Our conclusions in respect of life insurance are very much along the same lines as those in respect of CI insurance, and we shall avoid repetition. We found premiums to be affordable in many cases, either with a relatively low number of CAG repeats or just with a family history. The results in the latter case were broadly consistent with those of Smith (1998) and much lower than the ratings in Brackenridge & Elder (1998). Premiums given any intermediate allele, even 39 CAG repeats, were close to standard rates, and such test results could perhaps be ignored for life insurance. The potential costs of adverse selection arising from HD alone were negligible, but until we can place this in the context of all single-gene disorders we cannot draw firm conclusions.

It is interesting to compare the life insurance premiums given a family history of HD with the premiums given a family history of another neurological genetic disorder, early-onset Alzheimer's disease (EOAD), associated with Presenilin-1 gene mutations, given in Gui & Macdonald (1992). In many cases the premiums in respect of HD are lower than those in respect of EOAD. The reason is that mortality after onset of HD is considerably lighter than mortality after onset of EOAD associated with Presenilin-1 mutations (or more accurately, the available estimates suggest that this is so). HD might spring to mind as the prototype of a severe, single-gene disorder, but age at onset is only half of the story.

8.3 *Variable Single-Gene Disorders: A Problem for a Moratorium?*

This is the first study to consider insurance pricing in the presence of a variable disease-causing mutation. Previous studies, including Smith (1998) but also all studies into other genetic disorders, have assumed that all mutations in each gene involved have the same penetrance.

Classifying people as either mutation carriers or non-carriers, hence the crisp dichotomy of 'adverse' and 'clear' genetic test results, is perhaps consistent with the simplified view of genetic disorders as homogeneous, but is not consistent with genuine variability or heterogeneity. DNA-based genetic testing (if reliable) is more accurate than family history information, in the sense that it resolves the uncertainly in the latter, but that does not mean that any such ranking may be assumed when considering the medical or financial consequences. The examples given here of premium ratings given an 'adverse' test result that are much lower than those based on the family history show this clearly.

This ambiguity could affect the definition of an adverse test result in any moratorium that allows family history to be used to underwrite. There seems to be broad agreement in the UK industry that a clear test result will, in practice, be taken into account however a moratorium is worded. We included this possibility in our models, finding that the premiums charged to those who remain with a family history rise, with the removal of some non-carriers from this particular risk pool. What might happen were we to extend this, seemingly reasonable, process to a variable genetic disorder such as HD? First, those with test results showing a small number of CAG repeats could, in some cases, disclose them to get lower premiums than those based on their family history, so the latter would

have to rise, perhaps uncovering another group of mutation carriers who could get lower premiums by disclosing test results, and so on until the limits of insurability were reached. Thus common sense and the best of intentions could lead to the use of adverse genetic tests in underwriting.

One response would be to interpret a moratorium more strictly, and impose a family history rating regardless of the disclosure of a clear or 'mildly' adverse test result. How then should we respond to the applicant who has no mutation, or a mutation with 'only' 40 CAG repeats? To ignore the fact would be to tell him or her that as a member of a family at risk of HD they have a duty to pool their risks with those of all members of all other such families, regardless of their personal circumstances. The burden is theirs, not to be shared with those at risk of other single-gene disorders (who will presumably bear their own burdens), or of mutifactorial disorders, or of anything else. In passing, we note that our suggestion that most carriers of 'intermediate' HD alleles (36–39 CAG repeats) might be offered standard premium rates, while it seems sensible, is in fact a step in the direction of using adverse test results to underwrite. Once the principle is allowed for clear test results, where should we stop? It will be interesting to see how the implications of a moratorium are worked out as more is learned about heterogeneity and variability of single-gene disorders.

8.4 *Disclosure for Pricing* versus *Disclosure for Risk Management*

This paper shows that complex questions of insurance management can emerge even from consideration of rare, single-gene disorders. Premium rating is only one among several such questions. Fears about discrimination have led insurers to accept, or regulators to impose, moratoria on the use of genetic information for premium rating. It seems that most moratoria in practice have gone further than this, and ban the disclosure of the relevant information. In the UK, this stems partly from the principles underlying data protection: since the insurer will not use the information for pricing, they should not ask for it. Of course this provides strong reassurance that it will not be used in pricing, but it also stops the insurer from using it in other aspects of risk management that would in no way disadvantage those at risk, and might strengthen the overall position of the insurance pool. For example if a group of insurers entered into a pooling arrangement to share the costs of subsidising premiums for at-risk applicants, they would need to know how to share the costs and how to reserve for the liability. Indeed the regulator might require this of them.

Public trust in the insurance industry is lacking, sometimes with good reason. Understandably, therefore, it is disclosure, rather than particular uses, of genetic information that has been restricted at first. And although actuarial research into genetic disorders is far from complete, studies such as this one indicate that the risks in absolute terms are so small that this might not matter. This should not obscure the important point of principle, however: if government or society wishes to guarantee access to insurance, there will come a point when it is unreasonable to achieve this simply by depriving insurance managers of information they might need for the sound conduct of the business.

Acknowledgements

This work was carried out at the Genetics and Insurance Research Centre at Heriot-Watt University, which receives funding from a consortium of insurance companies through the Association of British Insurers. We thank the sponsors for funding, and members of the Steering Committee for helpful comments at various stages. We are grateful to Professor Peter Harper for his help.

References

Brackenridge, R. & Elder, J. (1998). *Medical selection of life risks. Fourth Edition.* Macmillan.

Brinkman, R., Mezei, M., Theilmann, J., Almqvist, E. & Hayden, M. (1997). The likelihood of being sffected with Huntington disease by a particular age, for a specific CAG size. *American Journal of Human Genetics*, **60**, 1202–1210.

Brock, D., Curtis, A., Barron, L., Dinwoodie, D., Crosbie, A., Mennie, M., Millan, F., Raeburn, J., Holloway, S. & Wright, A. (1989). Predictive testing for Huntington's disease with linked DNA markers. *The Lancet*, **2(8661)**, 463–466.

Collett, D. (1994). *Modelling survival data in medical research.* Chapman & Hall, London.

Dinani, A., Grimshaw D., Robjohns N., Somerville S., Spry A., Staffurth J. (2000). *A critical review. Report of the Critical Illness Healthcare Study Group.* Unpublished manuscript, The Staple Inn Actuarial Society.

Falush, D., Almqvist, E.W., Brinkmann, R.R., Iwasa, Y. & Hayden, M.R. (2000). Measurement of mutational flow implies both a high new-mutation rate for Huntington disease and a substantial underascertainment of late-onset cases. *American Journal of Human Genetics*, **68**, 373–385.

Gui, E.H. & Macdonald, A.S. (2002). Early-onset Alzheimer's disease, critical illness insurance and life insurance. *Research Report No. 02/2, Genetics and Insurance Research Centre, Heriot-Watt University, Edinburgh.*

Gusella, J., Wexler, N., Conneally, M., Naylor, S., Anderson, M., Tanzi, R., Watkins, P., Ottina, K., Wallace, M., Sakaguchi, A., Young, A., Shoulson, I., Bonilla, E. & Martin, J. (1983). A polymorphic DNA marker genetically linked to Huntington's disease. *Nature*, **306**, 234–238.

Gutiérrez, M.C. & Macdonald, A.S. (2001). Adult polycystic kidney disease and critical illness insurance. *Research Report No. 01/4, Department of Actuarial Mathematics and Statistics, Heriot-Watt University.*

Harper, P.S. (1996). *Huntington's disease.* W.B. Saunders.

Macdonald, A.S. (2001). Moratoria on the use of genetic tests and family history for mortgage-related life insurance. *Genetics and Insurance Research Centre, Heriot-Watt University: Research Report No. 01/3.*

Macdonald, A.S., Waters, H.R. & Wekwete, C.T. (2003a). The genetics of breast and ovarian cancer I: A model of family history. *To appear in Scandinavian Actuarial Journal.*

Macdonald, A.S., Waters, H.R. & Wekwete, C.T. (2003b). The genetics of breast and ovarian cancer II: A model of critical illness insurance. *To appear in Scandinavian Actuarial Journal.*

Smith, C. (1998). *Huntington's chorea: A mathematical model for life insurance*. Unpublished manuscript, Swiss Re, Zurich.

Wilkie, A.D. (2000). *Report by the Independent Actuary on the Application by the Association of British Insurers to the Genetics and Insurance Committee for Approval to use Genetic Test Results for Insurance Risk Assessment — Use of Huntington's Disease Test in Life Insurance*. Report available by request to the Genetics and Insurance Committee.

APPENDIX

THE CRITICAL ILLNESS INSURANCE MODEL

Gutiérrez & Macdonald (2001) obtained the following model for CI insurance based on medical studies and population data. Full references can be found in that paper.

(a) Rates of onset were found for:

(1) *Cancer (excluding non-malignant skin cancers)*: For males:

$$\begin{aligned}\mu_x^c &= \exp(-11.25 + 0.105x) \quad (x < 51)\\ \mu_x^c &= \exp(0.2591585 - 0.01247354x + 0.0001916916x^2 - 8.952933 \times 10^{-7}x^3) \quad (x \geq 60)\end{aligned}$$

with linear interpolation between ages 51 and 60, and for females:

$$\begin{aligned}\mu_x^c &= \exp(-10.78 + 0.123x - 0.00033x^2) \quad (x < 53)\\ \mu_x^c &= -0.01545632 + 0.0003805097x \quad (x \geq 53).\end{aligned}$$

(2) *Heart Attack*: For males:

$$\begin{aligned}\mu_x^h &= \exp(-13.2238 + 0.152568x) \quad (x < 44)\\ \mu_x^h &= (-0.01245109 + 0.000315605x) \quad (x > 49)\end{aligned}$$

with linear interpolation between ages 44 and 49, and for females:

$$\mu_x^h = \left(0.598694\left(\frac{0.15317^{15.6412}\exp(-0.15317x)x^{14.6412}}{\Gamma(15.6412)}\right)\right).$$

(3) *Stroke*: For males:

$$\mu_x^s = \exp(-16.9524 + 0.294973x - 0.001904x^2 + 0.00000159449x^3)$$

and for females:

$$\mu_x^s = \exp(-11.1477 + 0.081076x).$$

(b) 28-day survival factors for heart attack and stroke victims were taken from Dinani *et al.* (2000) (this relates to the common contractual condition, that payment depends on surviving for 28 days). Let p_x^h and p_x^s be the 28-day survival probabilities after the first-ever heart attack or stroke, respectively, and $q_x^h = 1 - p_x^h$, $q_x^s = 1 - p_x^s$ the corresponding mortality rates. From Dinani *et al.* (2000), $q_x^h = 0.21$ at ages 20–80 for females, and q_x^h for males is given in Table 20. From the same source, $p_x^s = (0.9 - 0.002x)/0.9$ for both males and females.

Table 20: 28-Day mortality rates ($q_x^h = 1 - p_x^h$) following heart attack. Based on Dinani *et al.* (2000).

age	q_x^h	age	q_x^h	age	q_x^h	age	q_x^h
20–39	0.15	47–52	0.18	58–59	0.21	65–74	0.24
40–42	0.16	53–56	0.19	60–61	0.22	75–79	0.25
43–46	0.17	57	0.20	62–64	0.23	80+	0.26

(c) Other minor causes of CI insurance claims amount to about 15% of those arising from cancer, heart attack and stroke. Therefore the aggregate rate of CI claims is:

$$\mu_x^{CI} = 1.15(\mu_x^c + p_x^h \times \mu_x^h + p_x^s \times \mu_x^s).$$

(d) Population mortality rates (English Life Tables No. 15) were adjusted to exclude deaths which would have followed a CI insurance claim.