MODELLING ADVERSE SELECTION IN THE PRESENCE OF A COMMON GENETIC DISORDER: THE BREAST CANCER POLYGENE

BY

ANGUS S. MACDONALD AND KENNETH R. MCIVOR

ABSTRACT

The cost of adverse selection in the life and critical illness (CI) insurance markets, brought about by restrictions on insurers' use of genetic test information, has been studied for a variety of rare single-gene disorders. Only now do we have a study of a common disorder (breast cancer) that accounts for the risk associated with multiple genes. Such a collection of genes is called a polygene. We take two approaches to modelling the severity of adverse selection which may result from insurers being unable to take account of tests for polygenes as well as major genes. First, we look at several genetic testing scenarios, with a corresponding range of possible insurance-buying behaviours, in a market model for CI insurance. Because a relatively large proportion of the population is exposed to adverse polygenic risk, the costs of adverse selection are potentially much greater than have been associated with rare single genes. Second, we use utility models to map out when adverse selection will appear, and which risk groups will cause it. Levels of risk aversion consistent with some empirical studies do not lead to significant adverse selection in our model, but lower levels of risk aversion could effectively eliminate the market.

KEYWORDS

Adverse Selection; BRCA1; BRCA2; Breast Cancer; Critical Illness Insurance; Polygene.

1. INTRODUCTION

1.1. Major Genes and Polygenes

The link between high risks of breast cancer (BC) and ovarian cancer (OC), and rare mutations in either of the BRCA1 and BRCA2 genes, is well-established. Several actuarial studies (Lemaire *et al.*, 1999; Subramanian *et al.*, 2000; Macdonald, Waters & Wekwete, 2003a, 2003b; Gui *et al.*, 2006) have considered the implications for the markets in life insurance and critical illness (CI) insurance

Astin Bulletin 39(2), 373-402. doi: 10.2143/AST.39.2.2044640

(known as dread disease insurance in continental Europe). Single genes associated with greatly increased disease risk are called 'major genes'. On the whole, these studies have found that insurance premiums for mutation carriers may be greatly elevated, but that banning the use of genetic test results would not lead to significant costs arising from adverse selection, because mutations are so rare. Moreover, insurers may often still have regard to family history.

More recently a different type of genetic risk called 'polygenic' has been associated with BC. What we will call a 'polygene' is a collection of genes, each with several variants, called 'alleles', not necessarily rare. Adverse configurations of a polygene might confer susceptibility to a particular disease, or beneficial configurations might protect against it. These outcomes could also be strongly influenced by the environment. It is likely that we all carry some 'good' and some 'bad' polygene configurations, but this is quite speculative at this stage.

Genetic science is moving on, from understanding single-gene disorders to beginning to understand polygenic disorders. We highlight some key differences between the two, that could affect questions of insurance.

(a) Mutations in major genes are rare, but polygenes will be present (in all their varieties of configurations) in every person. So, instead of tiny numbers of people with very high risks, larger numbers with modestly increased risks may contribute to the cost of adverse selection.
(b) Genetic testing for major gene mutations is, in many countries, controlled by public health services and is only offered if family history suggests a mutation is present. Polygenic risk, however, need not be associated with any clear family history, so genetic testing for risky polygenes will be initiated in some other way, such as screening programmes or selling direct to the public.

That said, research is at a very early stage and genetic testing for risky polygenes is not yet feasible.

1.2. The Polygenic Breast Cancer Model of Antoniou *et al.* (2002)

Antoniou *et al.* (2002) estimated rates of onset of BC and OC in a model assuming the presence of: (a) rare mutations in the major genes BRCA1 and BRCA2; and (b) a polygene affecting BC risk only. The polygene was modelled as a collection of three genes, each with a protective and a deleterious allele. Since everyone possesses two copies of every nuclear gene (ignoring sex chromosomes) each person carries six of these alleles. The quantitative effect of the polygene was governed by a number called the 'polygenotype', denoted P, which was the sum of the contributions of each allele: an adverse allele contributed +1/2, and a protective allele contributed –1/2. Thus P was an integer between –3 (beneficial) and +3 (adverse). P was assumed to act multiplicatively on the onset rate of BC as follows:

$$\text{Onset Rate for Polygenotype } P = \text{Baseline Onset Rate} \times \exp(cP) \qquad (1)$$

https://doi.org/10.2143/AST.39.2.2044640 Published online by Cambridge University Press

where the baseline onset rate is that for $P = 0$, and the constant c is just a scale factor. Assuming each allele to be equally common, and inherited independently of the others, the distribution in the population of the quantity $(P + 3)$ is Binomial(6,1/2).

Macdonald & McIvor (2006) used this polygenic model to estimate critical illness (CI) insurance prices. They considered the influence of the polygene on the development of family histories of BC/OC, and their use in underwriting.

1.3. Multi-State Market Models for Adverse Selection

Our aim is to study the impact of the polygenic model on adverse selection in the CI insurance market. Adverse selection arises when an individual is able to obtain insurance without having to disclose all the information that the insurer requires to price the contract accurately, so the applicant obtains cover below its full cost. An obvious first approach is to adapt the models used to address this question in respect of single-gene disorders. These are Markov multi-state models of the life history of a person who starts healthy and uninsured, and who may pass through a series of states representing relevant events, such as acquiring a family history, taking a genetic test, buying CI insurance, and suffering onset of BC/OC. By allowing some transition intensities to depend on the person's current knowledge, we can model how that influences various decisions. By basing premium rates on the insurer's lack of knowledge, we can model their exposure to adverse selection. The outcomes translate directly into the uniform increase in premiums that would be needed to recoup the cost of the adverse selection, a simple and relevant measure. We do this in Section 2, where a new consideration is the different routes to genetic testing that may come to be associated with polygenes.

1.4. An Economic Framework for Adverse Selection

An acknowledged weakness of the multi-state market model described above is its lack of economic rationale. That is, it posits certain behaviour (adverse selection) on the part of market participants. This behaviour changes the price of CI policies. There it ends, in our model, but an economist would iterate the process, price and behaviour changing each other, until (or if) equilibrium was reached.

It is hard to incorporate equilibrium prices into the Markov market model directly. Other approaches exist, considering the insurance-buying decision at a fixed time, setting the applicant's knowledge (and utility function) against the insurer's ignorance (and pooling instincts), and seeing whether insurance would be purchased, by whom, and perhaps also how much (Hoy & Witt, 2007; Macdonald & Tapadar, 2008). Using the polygenic BC model, we implement this approach in Section 3. We discuss the different insights that the two approaches bring, and draw conclusions, in Section 4.

https://doi.org/10.2143/AST.39.2.2044640 Published online by Cambridge University Press

2. Multi-State Market Models for Adverse Selection

2.1. The Basic Market Model

We begin with the model in Figure 1. Each genotype is represented by a version of this model, with different rates of onset of BC and OC. The model represents the life history of a person, as yet uninsured, who may buy insurance before or after having a genetic test. Premiums are payable while in either of the insured states, and the benefits are payable on transition from either of these states into a 'critical illness' state (which represents the onset of BC, OC, or another critical illness).

Some of the assumptions we make are as follows:

(a) Large and small markets are represented by 'normal' rates of insurance purchase of 0.05 or 0.01 *per annum*, respectively.
(b) In both markets, low risk polygenotype carriers may buy less insurance than the 'normal' rate. These carriers purchase at the normal rate, half of the normal rate, or not at all.
(c) Genetic testing occurs at three possible rates *per annum*: 0.02972 (low), 0.04458 (medium), or 0.08916 (high), based on an uptake proportion of 59% (Ropka *et al.*, 2006) over a period of 30, 20, or 10 years of testing respectively. Also, testing may only occur between ages 20 and 40 (when it has high priority).
(d) 'Severe' adverse selection means that high-risk polygenotype carriers will purchase insurance at rate 0.25 *per annum*.

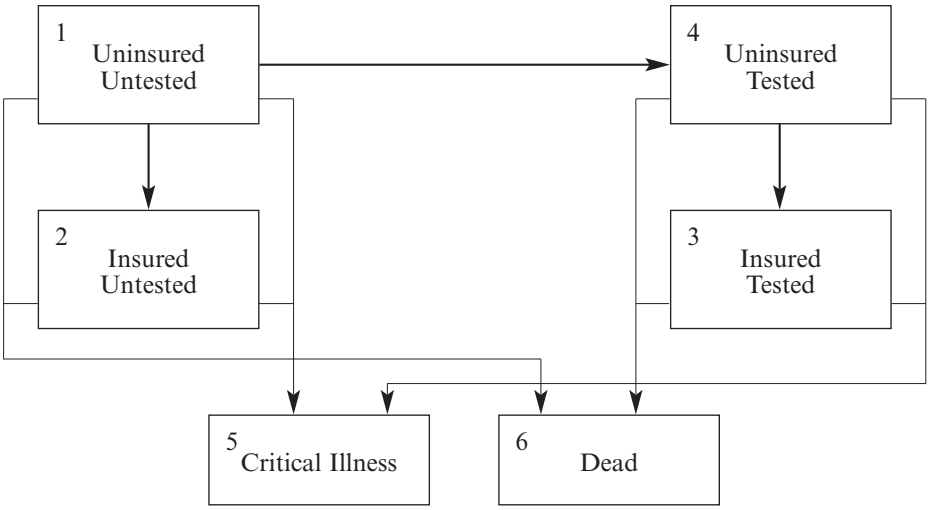

Figure 1: A model of the behaviour of a genetic subpopulation with respect to the purchasing of CI insurance. Genetic testing is available at an equal rate to all subpopulations.

https://doi.org/10.2143/AST.39.2.2044640 Published online by Cambridge University Press

All other intensities, governing transitions into the 'Dead' and 'Critical Illness' states, are as in Macdonald & McIvor (2006) and are omitted for brevity. NPEPVs of benefits and premiums are found by solving Thiele's differential equations backwards numerically with force of interest $\delta = 0.05$. Occupancy probabilities are found by solving Kolmogorov's Forward Equations. For both, we use a step-size of 0.0005 years.

2.2. A Genetic Screening Programme for the Polygene Only

For simplicity, the first possibility we consider is that a genetic screening programme exists for the polygenotype only, not extending to the BRCA1/2 genotypes. There are seven polygenotypes, therefore a 42-state model. We assume that the distribution of new-born persons in the seven sub-populations is Binomial(6,1/2), and that mortality and morbidity before age 20 does not depend on genotype (so that the expected proportions in each starting state at age 20 have not changed). Since the rate of BC onset is negligible before about age 30, this assumption seems reasonable.

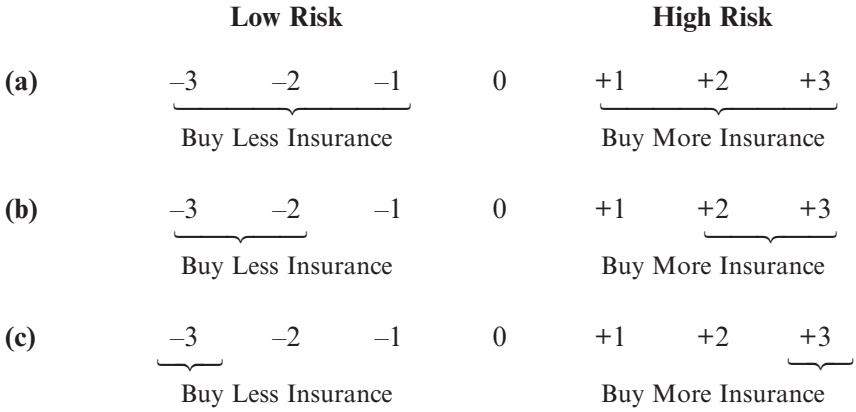

FIGURE 2: Three possible behaviours of tested polygenotype carriers in the adverse selection model, labelled (a), (b) and (c).

Tested carriers may alter their insurance-buying habits, in one of two ways: carriers of deleterious polygenotypes may buy more insurance, or carriers of protective polygenotypes may buy less insurance. This latter behaviour is uncommon in adverse selection studies; it is usually assumed that individuals who receive negative test results for rare, severe, mutations will purchase insurance at the normal market rate. One study (Subramanian *et al.*, 1999) performed a sensitivity analysis where tested non-carriers could reduce their coverage. It is plausible that this makes more sense from an economic point of view. Figure 2 shows three scenarios of differing severity.

https://doi.org/10.2143/AST.39.2.2044640 Published online by Cambridge University Press

The percentage by which all premiums must be raised in order to negate the adverse selection costs is:

$$100 \times \left(\frac{\text{(EPV[Loss|Adverse Selection] – EPV[Loss|No Adverse Selection]}}{\text{EPV[Premium Income|Adverse Selection]}} \right). \quad (2)$$

Table 1 shows the premium increases needed to absorb the costs of the adverse selection under each scenario in Figure 2. Compared with previous results based on major genes only (Gui *et al.*, 2006; Gutierrez & Macdonald, 2007) these are very high. This is because deleterious polygenotypes are more common than major gene mutations, and also because these authors did not have to consider the possibility that carriers of beneficial genotypes would buy less insurance. Note the large fall in costs between Scenarios (b) and (c). In Scenario (c), adverse selection is confined to the tails of the Binomial distribution of polygenotypes.

Curiously, in a small market the cost of adverse selection is always higher in Scenario (b) than in (a). This is because premium increases are relative to a baseline 'ordinary' rate (OR rate), which is higher in Scenario (a) than in (b).

TABLE 1

COSTS OF ADVERSE SELECTION RESULTING FROM HIGH RISK POLYGENOTYPE CARRIERS BUYING MORE INSURANCE THAN LOW RISK POLYGENOTYPE CARRIERS IN A CRITICAL ILLNESS INSURANCE MARKET OPEN TO FEMALES BETWEEN AGES 20-60. SCREENING AVAILABLE FOR THE POLYGENE ONLY.

Genetic Testing	Market Size	Insurance Purchasing of Low Risk Polygenotypes	Premium Increase in Scenario (a) %	(b) %	(c) %
Low	Large	Normal	1.05975	0.90051	0.26447
		Half	1.69947	1.42748	0.30825
		Nil	2.86994	2.36206	0.38349
	Small	Normal	6.81382	7.03421	1.03701
		Half	7.72964	7.90872	1.10420
		Nil	8.86328	8.94001	1.17995
Medium	Large	Normal	1.39994	1.20315	0.29909
		Half	2.27895	1.93093	0.35897
		Nil	3.95151	3.25381	0.46294
	Small	Normal	8.25781	9.04200	1.32143
		Half	9.49144	10.27268	1.41322
		Nil	11.10982	11.75999	1.51728
High	Large	Normal	2.01615	1.77037	0.36453
		Half	3.40261	2.92799	0.45793
		Nil	6.32401	5.15661	0.62418
	Small	Normal	10.31240	12.36857	1.83292
		Half	12.19959	14.37086	1.97494
		Nil	15.09918	16.92565	2.13791

https://doi.org/10.2143/AST.39.2.2044640 Published online by Cambridge University Press

We repeated the experiment for the case where high-risk polygenotype carriers purchase insurance at the normal rate, hence they do not contribute to adverse selection. However low-risk polygenotype carriers may still modify their purchasing behaviour by purchasing at half the normal rate or not purchasing at all. Table 2 shows the effect of such behaviour on the cost of adverse selection. (When the low-risk polygenotype carriers purchase at the normal rate, there is no adverse selection.)

2.3. A Genetic Screening Programme for the Polygene and Major Genes

Now we consider the possibility that screening is available for the major BRCA1 and BRCA2 genotypes, as well as for the polygenotype. We have $3 \times 7 = 21$ distinct genotypes, and 126 states in the model. We assume population frequencies of BRCA1 and BRCA2 mutations of 0.0010181 and 0.0013577 respectively (Antoniou *et al.*, 2002).

We consider the same adverse selection scenarios as in Figure 2, but additionally those who carry adverse BRCA1/2 mutations exercise selection (that

TABLE 2

COSTS OF ADVERSE SELECTION RESULTING FROM LOW RISK POLYGENOTYPE CARRIERS BUYING LESS INSURANCE THAN NORMAL IN A CRITICAL ILLNESS INSURANCE MARKET OPEN TO FEMALES BETWEEN AGES 20-60. HIGH RISK POLYGENOTYPE CARRIERS BUY INSURANCE AT NORMAL RATE. SCREENING AVAILABLE FOR THE POLYGENE ONLY.

Genetic Testing	Market Size	Insurance Purchasing of Low Risk Polygenotypes	Premium Increase in Scenario (a) %	(b) %	(c) %
Low	Large	Normal	0.00000	0.00000	0.00000
		Half	0.64468	0.51917	0.16535
		Nil	1.82361	1.44036	0.44408
	Small	Normal	0.00000	0.00000	0.00000
		Half	0.98952	0.78697	0.24367
		Nil	2.19323	1.71326	0.51771
Medium	Large	Normal	0.00000	0.00000	0.00000
		Half	0.88799	0.71343	0.22642
		Nil	2.57479	2.01092	0.61215
	Small	Normal	0.00000	0.00000	0.00000
		Half	1.36056	1.07732	0.33161
		Nil	3.08502	2.37443	0.70689
High	Large	Normal	0.00000	0.00000	0.00000
		Half	1.40648	1.12406	0.35433
		Nil	4.35040	3.28932	0.97370
	Small	Normal	0.00000	0.00000	0.00000
		Half	2.13881	1.67679	0.51027
		Nil	5.14799	3.79710	1.09563

https://doi.org/10.2143/AST.39.2.2044640 Published online by Cambridge University Press

TABLE 3

COSTS OF ADVERSE SELECTION RESULTING FROM HIGH RISK POLYGENOTYPE CARRIERS BUYING MORE INSURANCE THAN LOW RISK POLYGENOTYPE CARRIERS IN A CRITICAL ILLNESS INSURANCE MARKET OPEN TO FEMALES BETWEEN AGES 20-60. SCREENING AVAILABLE FOR MAJOR GENES AND THE POLYGENE.

Genetic Testing	Market Size	Insurance Purchasing of Low Risk Polygenotypes	Premium Increase in Scenario (a) %	(b) %	(c) %
Low	Large	Normal	1.08112	0.93445	0.34798
		Half	1.73147	1.47837	0.53487
		Nil	2.92044	2.44180	0.84843
	Small	Normal	6.92408	7.25457	2.86768
		Half	7.85526	8.15781	3.15438
		Nil	9.00878	9.22405	3.47696
Medium	Large	Normal	1.42838	1.24882	0.46798
		Half	2.32238	2.00052	0.72474
		Nil	4.02274	3.36579	1.16023
	Small	Normal	8.39054	9.32273	3.81134
		Half	9.64599	10.59581	4.20875
		Nil	11.29504	12.13682	4.65893
High	Large	Normal	2.05799	1.83889	0.69839
		Half	3.46956	3.03642	1.10290
		Nil	6.44544	5.34242	1.80734
	Small	Normal	10.47855	12.75016	5.53492
		Half	12.40261	14.82749	6.16727
		Nil	15.36631	17.48591	6.89426

is to say, they decide to insure, without sharing their knowledge of their elevated risk) regardless of polygenotype. The resulting premium increases are shown in Table 3. They are not much larger than those in Table 1, the greatest increase being in Scenario (c). Compared with screening for the polygene alone, the adverse selection costs if screening is extended to BRCA1/2 mutations are not high.

We also considered the possibility that some BRCA1/2 mutation carriers do not buy more insurance, if they carry a protective polygenotype which 'voids' the BRCA1/2 risk. It was shown in Macdonald & McIvor (2006) that BRCA1/2 mutation carriers with $P = -3$ could plausibly obtain CI insurance at ordinary rates. The effect was very small, and we omit the results.

2.4. More Limited Genetic Testing for the Polygene and Major Genes

Only about 25% of BC cases are hereditary and only about 20% of these are caused by identifiable (major) genes, so mass screening programmes for these would be ineffective. Much more likely is that testing will continue to be offered

https://doi.org/10.2143/AST.39.2.2044640 Published online by Cambridge University Press

only to women who present a relevant family history of BC/OC, who are more likely to carry deleterious genes. Therefore we adjust our original model as shown in Figure 3, to include the development of a relevant family history, which is now taken to be a prerequisite for genetic testing.

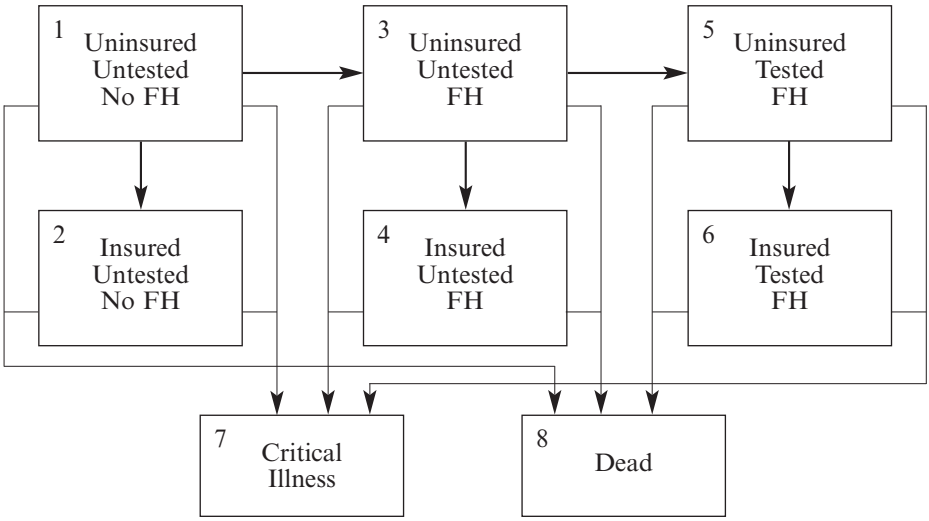

FIGURE 3: A model of the behaviour of a genetic subpopulation with respect to the purchasing of CI insurance. Genetic testing is available only after the appearance of a relevant family history (FH) of BC/OC.

We define a 'relevant' family history to mean that a healthy woman has two or more female first-degree relatives (FDRs) who contracted BC/OC before age 50. (FDR normally means parents and siblings but in what follows we limit it to mother and sisters.) The incidence of a relevant family history was calculated using the formula common in epidemiology:

$$\text{Incidence Rate} = \frac{\text{Number of new cases arising in specified time period}}{\text{Number of individuals at risk during the time period}}. \qquad (3)$$

In order to be categorised as 'at risk' of a relevant family history developing a woman must be healthy, with either: (a) no FDRs affected before age 50 and at least two unaffected FDRs under age 50; or (b) one FDR affected before age 50 and at least one unaffected FDR under age 50. If a relevant family history develops, each healthy daughter contributes as a 'new case' in the incidence at that time. These rates were estimated by simulation in Macdonald & McIvor (2006). The incidence rate is shown in Figure 4 for the subpopulations without BRCA mutations in the family, and in Figure 5 for the BRCA1 and BRCA2 carrier families. Since we assume all siblings to be the same age, no relevant family history can develop after age 50.

https://doi.org/10.2143/AST.39.2.2044640 Published online by Cambridge University Press

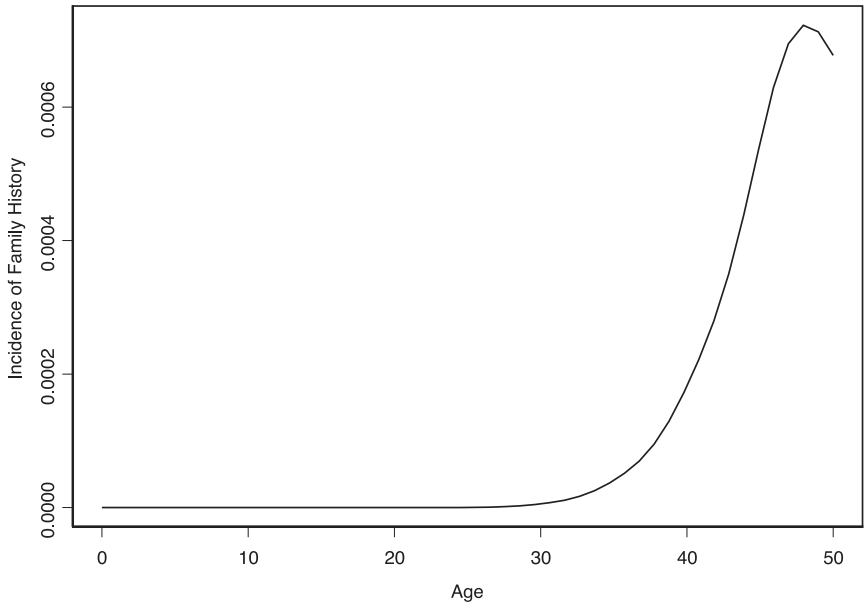

FIGURE 4: The incidence of a relevant family history for the subpopulations without BRCA mutations.

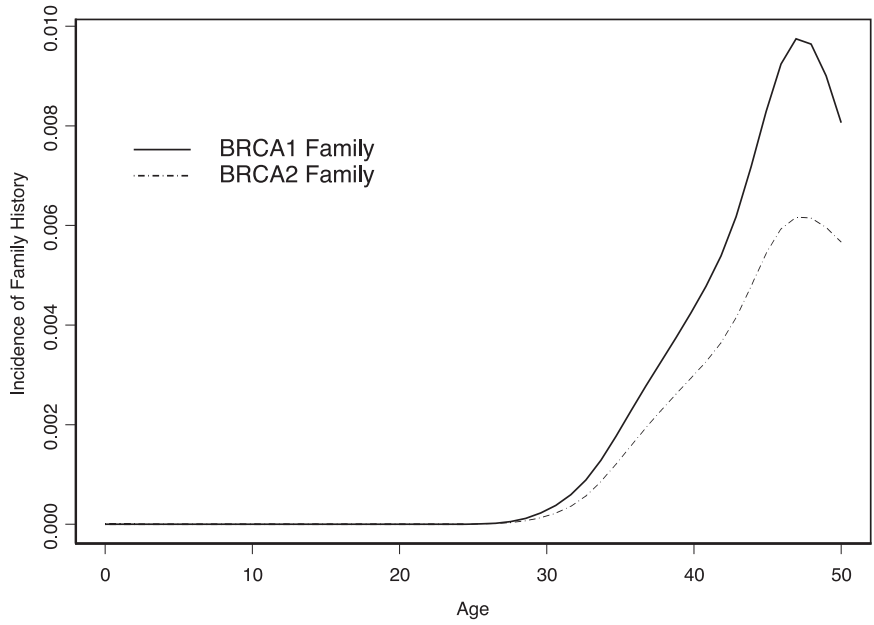

FIGURE 5: The incidence of a relevant family history for the subpopulations with BRCA1/2 mutations in the family.

https://doi.org/10.2143/AST.39.2.2044640 Published online by Cambridge University Press

TABLE 4

COSTS OF ADVERSE SELECTION RESULTING FROM HIGH RISK POLYGENOTYPE CARRIERS BUYING MORE INSURANCE THAN LOW RISK POLYGENOTYPE CARRIERS IN A CRITICAL ILLNESS INSURANCE MARKET OPEN TO FEMALES BETWEEN AGES 20-60. TESTING AVAILABLE FOR MAJOR GENES AND THE POLYGENE AFTER THE ONSET OF A RELEVANT FAMILY HISTORY.

Genetic Testing	Market Size	Insurance Purchasing of Low Risk Polygenotypes	Premium Increase in Scenario (a) %	(b) %	(c) %
Low	Large	Normal	0.00026	0.00025	0.00020
		Half	0.00034	0.00031	0.00025
		Nil	0.00044	0.00041	0.00031
	Small	Normal	0.00180	0.00164	0.00118
		Half	0.00192	0.00175	0.00125
		Nil	0.00206	0.00187	0.00134
Medium	Large	Normal	0.00034	0.00032	0.00025
		Half	0.00044	0.00041	0.00032
		Nil	0.00059	0.00055	0.00041
	Small	Normal	0.00255	0.00232	0.00165
		Half	0.00273	0.00248	0.00177
		Nil	0.00292	0.00266	0.00189
High	Large	Normal	0.00053	0.00049	0.00038
		Half	0.00071	0.00066	0.00049
		Nil	0.00098	0.00089	0.00065
	Small	Normal	0.00445	0.00404	0.00287
		Half	0.00477	0.00433	0.00306
		Nil	0.00511	0.00463	0.00327

The rate of genetic testing *per annum* among women who have developed a relevant family history we take to be 0.04012 (low), 0.06020 (medium) or 0.12040 (high), based on a proportion of 70% (Ropka *et al.*, 2006) being tested over a 30, 20 or 10 year period, respectively.

The results are in Table 4. The costs of adverse selection are greatly reduced when a relevant family history is a prerequisite for a genetic test. Once again a small insurance market suffers higher relative costs.

2.5. Separate Testing for the Polygene and Major Genes

Perhaps a more realistic situation is that testing for major genes is conducted through a public health service, once a relevant family history has signalled the risk, and testing for the polygene may be sought privately (by asymptomatic individuals). Therefore, we have two different testing events: one for the BRCA1/2 genes and one for the polygene, leading to the model shown in Figure 6. Both the family-history related and the non-family-history related testing rates may be at the low, medium or high levels.

https://doi.org/10.2143/AST.39.2.2044640 Published online by Cambridge University Press

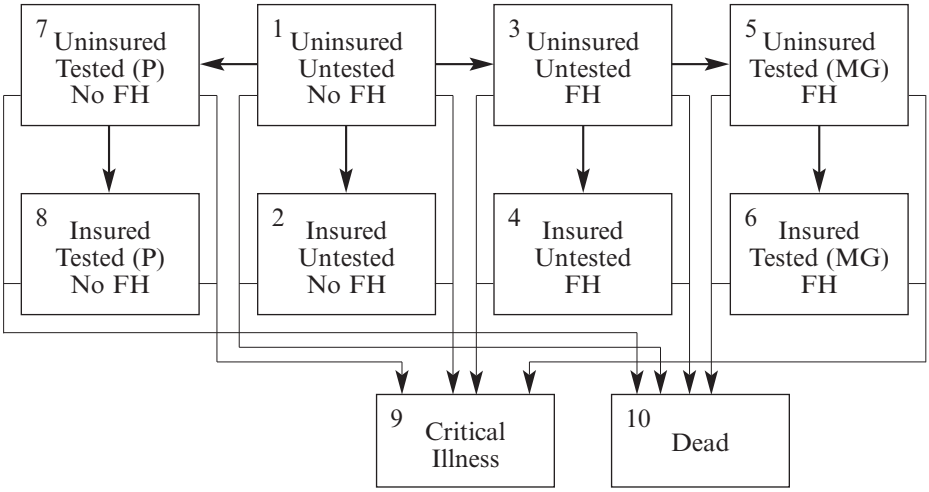

FIGURE 6: A model of the behaviour of a genetic subpopulation with respect to the purchasing of CI insurance. Genetic testing for major genes (MG) is available only after the appearance of a relevant family history (FH) of BC/OC. Testing for the polygene (P) is available before a relevant family history has appeared.

TABLE 5

COSTS OF ADVERSE SELECTION RESULTING FROM HIGH RISK POLYGENOTYPE CARRIERS BUYING MORE INSURANCE THAN LOW RISK POLYGENOTYPE CARRIERS IN A CRITICAL ILLNESS INSURANCE MARKET OPEN TO FEMALES BETWEEN AGES 20-60. SEPARATE TESTING FOR POLYGENE AND MAJOR GENES.

Genetic Testing	Market Size	Insurance Purchasing of Low Risk Polygenotypes	Premium Increase in Scenario (a) %	(b) %	(c) %
Low	Large	Normal	1.04241	0.89201	0.30311
		Half	1.67193	1.41431	0.46840
		Nil	2.82281	2.34075	0.74706
	Small	Normal	6.71995	6.97657	2.53289
		Half	7.61851	7.84459	2.78778
		Nil	8.72852	8.86823	3.07451
Medium	Large	Normal	1.37917	1.19288	0.40759
		Half	2.24492	1.91474	0.63453
		Nil	3.88979	3.22682	1.02114
	Small	Normal	8.15742	8.97684	3.36875
		Half	9.36749	10.19844	3.72104
		Nil	10.95022	11.67450	4.11975
High	Large	Normal	1.99302	1.75879	0.60791
		Half	3.36192	2.90895	0.96494
		Nil	6.23846	5.12262	1.58888
	Small	Normal	10.22137	12.30461	4.89643
		Half	12.07653	14.29422	5.45403
		Nil	14.91496	16.83171	6.09361

https://doi.org/10.2143/AST.39.2.2044640 Published online by Cambridge University Press

TABLE 6

COSTS OF ADVERSE SELECTION RESULTING FROM HIGH RISK POLYGENOTYPE CARRIERS BUYING MORE INSURANCE THAN LOW RISK POLYGENOTYPE CARRIERS IN A CRITICAL ILLNESS INSURANCE MARKET OPEN TO FEMALES BETWEEN AGES 20-60. SEPARATE TESTING FOR POLYGENE AND MAJOR GENES. MODEST ADVERSE SELECTION.

Genetic Testing	Market Size	Insurance Purchasing of Low Risk Polygenotypes	Premium Increase in Scenario (a) %	(b) %	(c) %
Low	Large	Normal	0.60275	0.50524	0.16798
		Half	1.23472	1.02422	0.33281
		Nil	2.39027	1.94512	0.61072
	Small	Normal	5.59863	5.54929	1.94805
		Half	6.51345	6.40019	2.20002
		Nil	7.64109	7.40364	2.48350
Medium	Large	Normal	0.80890	0.68244	0.22773
		Half	1.67911	1.39829	0.45385
		Nil	3.33224	2.70007	0.83911
	Small	Normal	6.92804	7.23268	2.61005
		Half	8.16462	8.42626	2.95730
		Nil	9.77516	9.86818	3.35034
High	Large	Normal	1.20769	1.03200	0.34723
		Half	2.58640	2.16860	0.70242
		Nil	5.47923	4.35764	1.32324
	Small	Normal	8.96365	10.18049	3.86596
		Half	10.86549	12.11832	4.41319
		Nil	13.74869	14.58806	5.04095

Our results, in Table 5, show somewhat smaller costs than the 'combined testing' model (Table 3). In fact, the costs are close to those of the polygene-only screening programme (Table 1).

We have so far assumed that 'severe' adverse selection takes place, defined as a rate of insurance purchase of 0.25 *per annum* by carriers of high-risk polygenotypes and BRCA1/2 mutations. We assume a less severe rate of adverse selection to be 0.1 *per annum*. Table 6 gives the costs of adverse selection when we apply this more modest rate of adverse selection. Note that the costs are still very high in the small market and in general the reduction in costs is small.

3. AN ECONOMIC FRAMEWORK FOR ADVERSE SELECTION

3.1. Utility Functions

A weakness of the models just described, shared by the models used by previous authors, is the lack of any economic rationale for the insurance-buying decisions. In particular, adverse selection is assumed to cause premium rates

https://doi.org/10.2143/AST.39.2.2044640 Published online by Cambridge University Press

TABLE 7

THE FOUR UTILITY FUNCTIONS PARAMETERISED BY MACDONALD & TAPADAR (2009).

Family	Utility Function	Parameter	Model
Iso-Elastic	$U(w) = \begin{cases} (w^\lambda - 1)/\lambda & \lambda < 1 \text{ and } \lambda \neq 0 \\ \log(w) & \lambda = 0 \end{cases}$	$\lambda = 0.5$ $\lambda = 0$ $\lambda = -8$	1 2 3
Negative Exponential	$U(w) = -\exp(-Aw)$	$A = 9 \times 10^{-5}$	4

to change, which ought in turn to affect insurance-buying decisions — the start of the classic 'adverse selection spiral'. Without pursuing this sequence of price and behavioural changes, we cannot be sure that premiums will eventually reach an equilibrium close to the changes suggested above. The usual approach to studying market equilibria starts with utility functions. It is hard to introduce these fully in the Markov models just described, but we can, nevertheless, use them to describe limits on the behaviour of market participants.

The utility function $U(w)$, with $U'(w) > 0$ and $U''(w) < 0$, can be interpreted as an increasing concave relation that describes the relative satisfaction gained from holding wealth w. Macdonald & Tapadar (2009) parameterised four utility functions, three from the Iso-Elastic family and one from the Negative Exponential family. We use the same utility functions and will refer to them as Models 1, 2, 3 and 4, as shown in Table 7. Models 1 and 2 have low risk-aversion. Models 3 and 4 were parameterised using data from a 1995 Italian thought-experiment (Eisenhauer & Ventura, 2003), adjusted for the sterling/lira exchange rate and UK price inflation up to 2006, and have higher risk-aversion.

We assume that a risk-averse individual facing uncertainty will seek to maximise his or her expected utility. For example, suppose an individual with total wealth W faces a loss L with probability q. The *actuarial value* (or *fair value*) of insurance against this random event is qL, but the individual would be prepared to pay premium $\Pi L \geq qL$ if:

$$U(W - \Pi L) > qU(W - L) + (1 - q)U(W). \tag{4}$$

The premium per unit of loss Π^* at which an individual would no longer purchase insurance is found by converting the inequality in Equation (4) to an equality and solving it.

Now suppose the population is stratified into separate subpopulations, within each of which the loss probability is different. Suppose that individuals are able to discover which stratum they are in, for example, by genetic testing. If the insurer charges everyone the same rate of premium, persons in each stratum will decide whether or not to insure, using their own level of risk. If the premium rate is high enough, low-risk individuals will leave the market. This is the boundary at which adverse selection begins to affect the market.

Macdonald & Tapadar (2009) explored this aspect of adverse selection using a hypothetical genetic model with two genotypes interacting with two levels of an environmental factor. We can now extend their study, to the more realistic setting of BC/OC and a real parameterised model.

3.2. Critical Illness Insurance Premiums

We will use the CI model in Figure 7 to calculate single premiums for stand-alone CI policies. We do not consider CI policies sold as riders to life insurance (so-called 'accelerated' benefits) because of the need to model survival after onset; this would be useful future work.

In the CI model a unit sum assured is payable on transition from the Healthy state to any CI state (BC, OC or Other Critical Illness). For simplicity, and consistency with previous studies of insurance and utility (Hoy & Witt, 2007; Macdonald & Tapadar, 2009), let the force of interest be $\delta = 0$. This means that EPVs are equivalent to the probabilities of the CI event occurring.

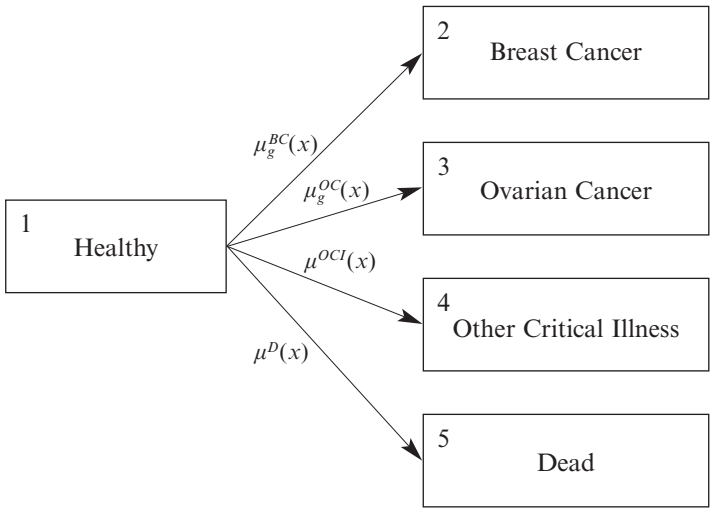

FIGURE 7: A model of the life history of a critical illness insurance policyholder, beginning in the Healthy state. Transition to the non-Healthy state d at age x is governed by an intensity $\mu^d(x)$ depending on age x or, in the case of BC and OC, $\mu_g^d(x)$ depending on genotype g as well.

For convenience, introduce genotype BRCA0 to represent non-carriers of BRCA1/2 mutations. Let (P, M) denote the genotype of a woman with polygenotype P and major genotype BRCAM, and let $\Pi(P, M)$ represent the CI single premium per unit benefit she would be charged, for a given entry age and policy term.

https://doi.org/10.2143/AST.39.2.2044640 Published online by Cambridge University Press

3.3. Threshold Premiums

For a given entry age and policy term, let $\bar{\Pi}$ be the single premium per unit benefit offered to all women when the insurer has no information regarding genotype. Thus $\bar{\Pi}$ is the weighted average premium:

$$\bar{\Pi} = \sum_{P,M} \omega(P,M)\,\Pi(P,M) \tag{5}$$

where $\omega(P,M)$ is the proportion of the population with genotype (P,M). These weighted average premiums are given in Table 8. Insurance will be bought by women with genotype (P,M) if $\bar{\Pi} \le \Pi^*(P,M)$ where $\Pi^*(P,M)$ is the solution of:

$$U(W - \Pi^*(P,M)L) = \Pi(P,M)\,U(W-L) + (1 - \Pi(P,M))\,U(W). \tag{6}$$

We call $\Pi^*(P,M)$ the *threshold premium* for the onset of adverse selection in respect of genotype (P,M). Clearly, adverse selection will appear first when $\bar{\Pi} > \Pi^*(-3,0)$. Tables 9 and 10 show the values of $\Pi^*(-3,0)$ for a selection of CI policies and a range of losses L when initial wealth W = £100,000. We can see that as the ratio L/W of loss to wealth increases, the threshold premium increases, implying greater propensity to insure more serious losses.

TABLE 8

THE WEIGHTED AVERAGE SINGLE PREMIUM FOR VARIOUS CI POLICIES (SEE EQUATION (5)).

Age	Term	Rate of Premium $\bar{\Pi}$
20	10	0.00670
	20	0.03081
	30	0.09549
	40	0.20315
30	10	0.02436
	20	0.08969
	30	0.19840
40	10	0.06782
	20	0.18009
50	10	0.12331

From Tables 9 and 10 we can roughly deduce what level of loss ratio L/W will initiate adverse selection. These ratios are about 0.85 for Model 1 and 0.55 for Model 2. However for Models 3 and 4, $\bar{\Pi} < \Pi^*(-3,0)$ for almost all levels of loss L we have tabulated. By replacing $\Pi^*(P,M)$ with $\bar{\Pi}$ in Equation (6) and solving for L with genotype $(-3,0)$, we can find the levels of loss that would initiate adverse selection. These are given in Table 11. We can see that adverse selection could occur under Models 3 and 4, but only for very low levels of insured loss relative to wealth (for which CI insurance is certainly unnecessary).

https://doi.org/10.2143/AST.39.2.2044640 Published online by Cambridge University Press

TABLE 9

THRESHOLD PREMIUM RATES $\Pi^*(-3,0)$ AT WHICH ADVERSE SELECTION WILL APPEAR, FOR A VARIETY OF CI POLICIES AND INITIAL WEALTH $W = £100{,}000$.

	Age	Term	Loss to Wealth Ratio								
			0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
Model 1	20	10	0.00598	0.00615	0.00634	0.00656	0.00682	0.00713	0.00752	0.00804	0.00884
		20	0.02222	0.02284	0.02355	0.02436	0.02530	0.02644	0.02786	0.02976	0.03267
		30	0.06491	0.06665	0.06862	0.07088	0.07352	0.07670	0.08068	0.08601	0.09416
		40	0.15116	0.15487	0.15905	0.16385	0.16946	0.17621	0.18466	0.19597	0.21329
	30	10	0.01639	0.01685	0.01737	0.01797	0.01867	0.01951	0.02056	0.02197	0.02414
		20	0.05944	0.06105	0.06286	0.06495	0.06738	0.07031	0.07397	0.07888	0.08639
		30	0.14646	0.15006	0.15414	0.15881	0.16428	0.17085	0.17908	0.19010	0.20697
	40	10	0.04394	0.04515	0.04651	0.04807	0.04990	0.05210	0.05485	0.05853	0.06417
		20	0.13274	0.13605	0.13980	0.14411	0.14914	0.15519	0.16276	0.17289	0.18843
	50	10	0.09358	0.09602	0.09878	0.10194	0.10564	0.11009	0.11566	0.12311	0.13453
Model 2	20	10	0.00614	0.00650	0.00692	0.00743	0.00806	0.00887	0.00999	0.01167	0.01481
		20	0.02280	0.02411	0.02566	0.02751	0.02981	0.03276	0.03678	0.04283	0.05407
		30	0.06652	0.07018	0.07447	0.07960	0.08591	0.09398	0.10490	0.12115	0.15079
		40	0.15456	0.16226	0.17122	0.18185	0.19480	0.21115	0.23294	0.26469	0.32059
	30	10	0.01681	0.01779	0.01893	0.02031	0.02202	0.02421	0.02721	0.03172	0.04012
		20	0.06093	0.06430	0.06826	0.07299	0.07882	0.08627	0.09636	0.11141	0.13892
		30	0.14977	0.15727	0.16601	0.17638	0.18901	0.20498	0.22627	0.25734	0.31214
	40	10	0.04506	0.04759	0.05057	0.05414	0.05855	0.06419	0.07185	0.08332	0.10442
		20	0.13579	0.14270	0.15076	0.16034	0.17203	0.18684	0.20662	0.23560	0.28698
	50	10	0.09583	0.10094	0.10691	0.11403	0.12277	0.13389	0.14885	0.17098	0.21086

TABLE 10

THRESHOLD PREMIUM RATES $\Pi^*(-3,0)$ AT WHICH ADVERSE SELECTION WILL APPEAR, FOR A VARIETY OF CI POLICIES AND INITIAL WEALTH $W = £100{,}000$.

	Age	Term	Loss to Wealth Ratio								
			0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Model 3	20	10	0.00959	0.01778	0.03769	0.09005	0.21518	0.41504	0.61436	0.77441	0.89972
		20	0.03526	0.06339	0.12395	0.24326	0.41797	0.59426	0.73691	0.84639	0.93172
		30	0.10009	0.16789	0.28336	0.44015	0.59803	0.72657	0.82357	0.89704	0.95424
		40	0.22084	0.33224	0.47458	0.61700	0.73361	0.82046	0.88435	0.93252	0.97001
	30	10	0.02611	0.04743	0.09527	0.19800	0.36742	0.55406	0.71014	0.83071	0.92476
		20	0.09200	0.15558	0.26661	0.42239	0.58347	0.71628	0.81688	0.89313	0.95250
		30	0.21457	0.32447	0.46658	0.61031	0.72871	0.81711	0.88218	0.93126	0.96945
	40	10	0.06871	0.11905	0.21407	0.36307	0.53293	0.68014	0.79334	0.87938	0.94639
		20	0.19610	0.30114	0.44206	0.58947	0.71337	0.80661	0.87540	0.92730	0.96769
	50	10	0.14167	0.22829	0.35984	0.51576	0.65780	0.76832	0.85064	0.91285	0.96127
Model 4	20	10	0.00941	0.01611	0.02880	0.05236	0.09286	0.15297	0.22655	0.30203	0.37102
		20	0.03459	0.05767	0.09737	0.15879	0.23877	0.32473	0.40437	0.47237	0.52858
		30	0.09825	0.15418	0.23365	0.32782	0.42065	0.50101	0.56629	0.61830	0.65991
		40	0.21711	0.30980	0.41326	0.50963	0.58899	0.65076	0.69824	0.73510	0.76422
	30	10	0.02561	0.04309	0.07416	0.12504	0.19655	0.27936	0.36036	0.43167	0.49155
		20	0.09030	0.14272	0.21875	0.31112	0.40414	0.48585	0.55273	0.60623	0.64910
		30	0.21093	0.30234	0.40540	0.50224	0.58244	0.64506	0.69327	0.73071	0.76031
	40	10	0.06743	0.10885	0.17290	0.25738	0.34930	0.43451	0.50642	0.56483	0.61198
		20	0.19272	0.27997	0.38148	0.47950	0.56218	0.62737	0.67781	0.71708	0.74815
	50	10	0.13913	0.21081	0.30338	0.40215	0.49159	0.56506	0.62310	0.66872	0.70499

TABLE 11

MINIMUM LOSS AT WHICH ADVERSE SELECTION OCCURS WITH WEALTH W = £100,000.

Age	Term	Model			
		1 £	2 £	3 £	4 £
20	10	45,600	25,100	3,100	3,100
	20	84,300	53,800	7,500	7,700
	30	100,000	61,600	9,100	9,400
	40	84,800	55,500	8,000	8,300
30	10	100,000	60,600	8,800	9,000
	20	100,000	63,800	9,500	9,900
	30	85,600	56,200	8,200	8,500
40	10	100,000	65,200	9,800	10,200
	20	85,300	55,800	8,100	8,300
50	10	80,300	50,600	7,000	7,200

3.4. Parameterisation of the Polygenic Model

The heterogeneity of risk introduced by the polygene, therefore the potential for adverse selection, is directly related to the standard deviation of the exponent cP in Equation (1). Denote this standard deviation σ_R, which Antoniou *et al.* (2002) estimated to be 1.291. Here we ask, what would be the effect of a larger or smaller σ_R?

For larger σ_R, persons with the (–3,0) genotype have lower relative risk and hence value insurance less, so we expect adverse selection to appear more readily. On the other hand, if $\sigma_R = 0$ there would be no adverse selection on account of the polygene, so as $\sigma_R \to 0$ there may sometimes be a non-zero value of σ_R at which adverse selection disappears. The major genes still play a role of course, but here they constitute a fixed background.

We will call the levels of σ_R at which adverse selection appears or disappears *threshold standard deviations* and denote them σ_R^*. They will depend on the entry age, policy term, utility model, wealth W and loss L. Table 12 shows values of σ_R^*. Boldface indicates values that are less than the actual estimate $\hat{\sigma}_R = 1.291$, meaning that we should expect adverse selection to appear, given the model of Antoniou *et al.* (2002). In agreement with Section 3.3, higher loss ratios L/W mean individuals with genotype (–3,0) would be less motivated to insure, so σ_R^* increases with L.

The missing figures in Table 12 are cases where the propensity to insure is sufficiently strong, that persons with genotype (–3,0) will insure unless their relative risk is very small, hence σ_R^* is very large. In fact we could not compute them because of numerical overflow, which is why they are missing, but for all practical purposes these are cases where adverse selection will never appear. This

https://doi.org/10.2143/AST.39.2.2044640 Published online by Cambridge University Press

TABLE 12

THRESHOLD STANDARD DEVIATIONS σ_R^* AT WHICH ADVERSE SELECTION APPEARS, FOR WEALTH $W = £100{,}000$. FIGURES IN BOLD INDICATE VALUES OF σ_R^* LOWER THAN THE ESTIMATE $\hat{\sigma}_R$ OF ANTONIOU ET AL. (2002). MISSING VALUES INDICATE VALUES OF σ_R^* TOO LARGE TO COMPUTE, IN PRACTICAL TERMS MEANING THAT ADVERSE SELECTION WILL NEVER APPEAR.

Model	Age	Term	Loss to Wealth Ratio L/W								
			0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	20	10	**0.193**	**0.485**	**0.844**	**1.152**	1.393	1.592	1.770	1.942	2.130
		20	**0.050**	**0.131**	**0.232**	**0.363**	**0.530**	**0.733**	**0.957**	**1.189**	1.441
		30	**0.039**	**0.101**	**0.175**	**0.268**	**0.387**	**0.542**	**0.738**	**0.974**	**1.253**
		40	**0.048**	**0.119**	**0.207**	**0.319**	**0.467**	**0.658**	**0.889**	**1.151**	1.467
	30	10	**0.038**	**0.103**	**0.183**	**0.284**	**0.412**	**0.576**	**0.778**	**1.008**	**1.270**
		20	**0.036**	**0.094**	**0.163**	**0.249**	**0.359**	**0.501**	**0.686**	**0.914**	**1.192**
		30	**0.047**	**0.116**	**0.201**	**0.311**	**0.455**	**0.640**	**0.868**	**1.127**	1.441
	40	10	**0.036**	**0.092**	**0.159**	**0.242**	**0.347**	**0.483**	**0.659**	**0.880**	**1.153**
		20	**0.049**	**0.119**	**0.207**	**0.319**	**0.466**	**0.655**	**0.883**	**1.140**	1.445
	50	10	**0.060**	**0.146**	**0.256**	**0.399**	**0.583**	**0.803**	**1.040**	**1.286**	1.565
2	20	10	**0.462**	**1.073**	1.465	1.734	1.949	2.138	2.320	2.515	2.761
		20	**0.125**	**0.322**	**0.594**	**0.907**	**1.194**	1.447	1.685	1.934	2.259
		30	**0.096**	**0.238**	**0.432**	**0.685**	**0.970**	**1.248**	1.523	1.832	2.295
		40	**0.113**	**0.280**	**0.516**	**0.815**	**1.126**	1.432	1.767	2.210	3.136
	30	10	**0.098**	**0.253**	**0.463**	**0.732**	**1.014**	**1.277**	1.524	1.777	2.089
		20	**0.089**	**0.222**	**0.400**	**0.636**	**0.911**	**1.188**	1.463	1.766	2.209
		30	**0.110**	**0.273**	**0.502**	**0.795**	**1.104**	1.408	1.738	2.171	3.044
	40	10	**0.087**	**0.216**	**0.387**	**0.613**	**0.881**	**1.153**	1.421	1.708	2.101
		20	**0.113**	**0.281**	**0.516**	**0.813**	**1.120**	1.418	1.736	2.149	2.915
	50	10	**0.139**	**0.351**	**0.647**	**0.976**	**1.276**	1.551	1.837	2.196	2.767
3	20	10	2.263	2.930	3.705	4.867					
		20	1.605	2.475	3.532						
		30	1.412	2.536	4.606						
		40	1.587	3.377							
	30	10	1.444	2.292	3.245						
		20	1.355	2.449	4.421						
		30	1.563	3.301							
	40	10	1.322	2.337	3.979						
		20	1.574	3.204							
	50	10	1.708	3.039							
4	20	10	2.234	2.840	3.381	4.155	4.915	7.458			
		20	1.566	2.346	3.127	4.244					
		30	1.366	2.345	3.673						
		40	1.529	2.920							
	30	10	1.404	2.177	2.878	3.699					
		20	1.309	2.266	3.452						
		30	1.506	2.855	4.733						
	40	10	**1.277**	2.175	3.109	4.810					
		20	1.519	2.792	4.637						
	50	10	1.659	2.768	4.548						

https://doi.org/10.2143/AST.39.2.2044640 Published online by Cambridge University Press

observation has an exact analogue in Macdonald & Tapadar (2009). Note that most missing values appear under utility Models 3 and 4, which were the two models parameterised from real economic data; Models 1 and 2, in contrast, were chosen simply to illustrate a lower level of risk-aversion.

3.5. Adverse Selection by Multiple Subpopulations

Previously we considered the case where adverse selection is triggered when the lowest-risk subpopulation refuses to purchase insurance. However, given a population with such a broad range of risks, it may also be of interest to consider the prospect that more than one low-risk subpopulation will no longer purchase insurance. (In fact, our discrete subpopulations defined by a polygene with three bi-allelic loci, meaning that there are two variants for each of the three genes, is itself the result of discretising an underlying Normal distribution for the polygenic risk in Equation (1)). Here we suppose that adverse selection extends to both the $(-3,0)$ and $(-2,0)$ genotypes. Adverse selection first occurs within the $(-3,0)$ genotype who, as a result, are removed from the risk pool. Our attention is then directed on the point where persons with the $(-2,0)$ genotype stop purchasing insurance, given a new actuarially fair premium rate $\bar{\Pi}_2 > \bar{\Pi}$ which no longer includes the $(-3,0)$ genotype.

The threshold premium is $\Pi^*(-2,0)$, and values of this are given in Tables 13 and 14. Table 15 shows the threshold standard deviations, as in Section 3.4. These latter values are higher than the values in Table 12, since the polygenic risk must be greater to trigger adverse selection within a second subpopulation.

3.6. The Polygenotype as a Continuous Random Variable

We mentioned in Section 3.5 that the Binomial model of the polygenotype is in fact a discretised version of a model in which the numerical value of the polygenotype has a continuous distribution, usually assumed to be Normal (Strachan & Read, 2004). One reason to discretise the polygenotype is in order to model the transmission of the polygenotype from parents to children, because the basis of inheritance is the transmission of discrete genes (Lange, 1997).

Since P is the sum of six independent random variables, each taking values $-1/2$ and $+1/2$ with equal probability, we see it has mean 0 and variance $3/2$. If we revert to the continuous Normal model of P, we equate moments and assume that $P \sim \mathcal{N}(0,3/2)$. Denote its density $f_P(p)$, let the (discrete) distribution of the major genotype M be $f_M(m)$, and let the (mixed) distribution of the combined genotype be $f_{(P,M)}(p,m) = f_P(p)\, f_M(m)$. Then the insurer's fair value premium, denoted $\bar{\Pi}^c$ is:

$$\bar{\Pi}^c = \sum_m \int_{-\infty}^{\infty} \Pi(p,m)\, f_{(P,M)}(p,m)\, dp. \tag{7}$$

https://doi.org/10.2143/AST.39.2.2044640 Published online by Cambridge University Press

TABLE 13

THRESHOLD PREMIUM RATES $\Pi^*(-2,0)$ AT WHICH ADVERSE SELECTION BY BOTH THE $P=-3$ AND $P=-2$ POLYGENOTYPE SUBPOPULATIONS WILL TAKE PLACE, FOR A VARIETY OF CI POLICIES AND INITIAL WEALTH $W=$£100,000.

	Age	Term	Loss to Wealth Ratio								
			0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Model 1	20	10	0.00601	0.00618	0.00638	0.00660	0.00686	0.00717	0.00756	0.00808	0.00888
		20	0.02256	0.02319	0.02390	0.02472	0.02568	0.02683	0.02828	0.03021	0.03316
		30	0.06615	0.06791	0.06993	0.07223	0.07492	0.07816	0.08221	0.08763	0.09594
		40	0.15343	0.15718	0.16142	0.16628	0.17196	0.17879	0.18735	0.19880	0.21634
	30	10	0.01669	0.01716	0.01769	0.01830	0.01902	0.01987	0.02095	0.02238	0.02458
		20	0.06067	0.06231	0.06416	0.06628	0.06876	0.07174	0.07548	0.08048	0.08814
		30	0.14872	0.15237	0.15650	0.16123	0.16677	0.17343	0.18177	0.19292	0.21002
	40	10	0.04490	0.04613	0.04752	0.04911	0.05098	0.05322	0.05603	0.05979	0.06555
		20	0.13478	0.13814	0.14194	0.14629	0.15139	0.15752	0.16519	0.17546	0.19119
	50	10	0.09482	0.09729	0.10008	0.10328	0.10703	0.11153	0.11717	0.12471	0.13626
Model 2	20	10	0.00617	0.00653	0.00696	0.00747	0.00811	0.00892	0.01004	0.01173	0.01489
		20	0.02314	0.02448	0.02604	0.02793	0.03025	0.03325	0.03733	0.04347	0.05486
		30	0.06779	0.07152	0.07588	0.08111	0.08753	0.09574	0.10684	0.12337	0.15348
		40	0.15688	0.16467	0.17373	0.18449	0.19759	0.21413	0.23615	0.26823	0.32464
	30	10	0.01713	0.01812	0.01929	0.02069	0.02243	0.02466	0.02771	0.03231	0.04086
		20	0.06218	0.06562	0.06965	0.07448	0.08041	0.08801	0.09829	0.11361	0.14160
		30	0.15208	0.15967	0.16852	0.17901	0.19180	0.20796	0.22949	0.26089	0.31622
	40	10	0.04604	0.04862	0.05166	0.05531	0.05980	0.06556	0.07337	0.08507	0.10658
		20	0.13787	0.14487	0.15303	0.16273	0.17457	0.18955	0.20957	0.23886	0.29077
	50	10	0.09710	0.10227	0.10831	0.11551	0.12435	0.13559	0.15073	0.17309	0.21338

TABLE 14

THRESHOLD PREMIUM RATES $\Pi^*(-2,0)$ AT WHICH ADVERSE SELECTION BY BOTH THE $P=-3$ AND $P=-2$ POLYGENOTYPE SUBPOPULATIONS WILL TAKE PLACE, FOR A VARIETY OF CI POLICIES AND INITIAL WEALTH $W=$£100,000.

	Age	Term	Loss to Wealth Ratio								
			0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Model 3	20	10	0.00965	0.01787	0.03789	0.09047	0.21591	0.41581	0.61491	0.77473	0.89987
		20	0.03579	0.06429	0.12553	0.24563	0.42049	0.59621	0.73820	0.84714	0.93206
		30	0.10193	0.17066	0.28707	0.44402	0.60117	0.72878	0.82501	0.89788	0.95461
		40	0.22385	0.33595	0.47837	0.62016	0.73592	0.82203	0.88537	0.93312	0.97027
	30	10	0.02659	0.04828	0.09683	0.20058	0.37046	0.55653	0.71179	0.83168	0.92519
		20	0.09382	0.15837	0.27044	0.42650	0.58685	0.71868	0.81844	0.89405	0.95291
		30	0.21760	0.32822	0.47045	0.61356	0.73109	0.81873	0.88324	0.93187	0.96972
	40	10	0.07017	0.12137	0.21755	0.36718	0.53654	0.68275	0.79504	0.88038	0.94683
		20	0.19887	0.30467	0.44583	0.59270	0.71576	0.80825	0.87646	0.92792	0.96796
	50	10	0.14344	0.23076	0.36279	0.51852	0.65992	0.76980	0.85160	0.91341	0.96151
Model 4	20	10	0.00946	0.01620	0.02896	0.05262	0.09327	0.15354	0.22721	0.30270	0.37165
		20	0.03510	0.05850	0.09866	0.16061	0.24097	0.32702	0.40657	0.47438	0.53040
		30	0.10006	0.15676	0.23696	0.33150	0.42425	0.50431	0.56923	0.62092	0.66225
		40	0.22008	0.31338	0.41700	0.51313	0.59209	0.65345	0.70059	0.73717	0.76606
	30	10	0.02608	0.04387	0.07541	0.12692	0.19898	0.28204	0.36299	0.43413	0.49380
		20	0.09209	0.14531	0.22214	0.31496	0.40796	0.48936	0.55587	0.60903	0.65161
		30	0.21391	0.30594	0.40920	0.50582	0.58562	0.64782	0.69568	0.73284	0.76220
	40	10	0.06885	0.11099	0.17589	0.26100	0.35309	0.43811	0.50969	0.56777	0.61461
		20	0.19545	0.28336	0.38514	0.48300	0.56531	0.63011	0.68020	0.71919	0.75003
	50	10	0.14087	0.21314	0.30612	0.40496	0.49421	0.56740	0.62515	0.67054	0.70662

TABLE 15

THRESHOLD STANDARD DEVIATIONS σ_R^* AT WHICH ADVERSE SELECTION APPEARS FOR THE (–2,0) GENOTYPE, FOR WEALTH $W = £100{,}000$. FIGURES IN BOLD INDICATE VALUES OF σ_R^* LOWER THAN THE ESTIMATE $\hat{\sigma}_R$ OF ANTONIOU ET AL. (2002). MISSING VALUES INDICATE VALUES OF σ_R^* TOO LARGE TO COMPUTE, IN PRACTICAL TERMS MEANING THAT ADVERSE SELECTION WILL NEVER APPEAR.

Model	Age	Term	Loss to Wealth Ratio								
			0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	20	10	**0.266**	**0.592**	**0.914**	**1.184**	1.405	1.595	1.768	1.938	2.124
		20	**0.072**	**0.184**	**0.315**	**0.467**	**0.639**	**0.824**	**1.019**	**1.225**	1.457
		30	**0.057**	**0.143**	**0.243**	**0.360**	**0.497**	**0.655**	**0.835**	**1.041**	**1.290**
		40	**0.069**	**0.168**	**0.284**	**0.421**	**0.581**	**0.764**	**0.968**	**1.199**	1.489
	30	10	**0.054**	**0.146**	**0.253**	**0.378**	**0.522**	**0.685**	**0.866**	**1.067**	1.300
		20	**0.052**	**0.134**	**0.228**	**0.338**	**0.466**	**0.616**	**0.789**	**0.990**	**1.236**
		30	**0.068**	**0.164**	**0.277**	**0.411**	**0.568**	**0.747**	**0.949**	**1.178**	1.465
	40	10	**0.052**	**0.131**	**0.222**	**0.328**	**0.452**	**0.597**	**0.765**	**0.961**	**1.200**
		20	**0.070**	**0.168**	**0.284**	**0.420**	**0.579**	**0.760**	**0.962**	**1.188**	1.468
	50	10	**0.086**	**0.205**	**0.344**	**0.507**	**0.690**	**0.888**	**1.094**	1.315	1.576
2	20	10	**0.568**	**1.113**	1.473	1.733	1.944	2.132	2.313	2.507	2.752
		20	**0.176**	**0.422**	**0.700**	**0.976**	**1.229**	1.463	1.689	1.932	2.251
		30	**0.136**	**0.323**	**0.544**	**0.788**	**1.037**	**1.285**	1.539	1.835	2.285
		40	**0.159**	**0.375**	**0.629**	**0.902**	**1.176**	1.455	1.772	2.199	3.086
	30	10	**0.139**	**0.340**	**0.575**	**0.826**	**1.072**	1.307	1.537	1.780	2.086
		20	**0.127**	**0.303**	**0.511**	**0.744**	**0.987**	**1.232**	1.484	1.772	2.201
		30	**0.156**	**0.367**	**0.616**	**0.885**	**1.157**	1.433	1.745	2.162	2.997
	40	10	**0.125**	**0.295**	**0.497**	**0.723**	**0.961**	**1.200**	1.444	1.716	2.097
		20	**0.160**	**0.375**	**0.629**	**0.900**	**1.170**	1.441	1.743	2.140	2.876
	50	10	**0.195**	**0.455**	**0.749**	**1.038**	1.306	1.563	1.837	2.186	2.745
3	20	10	2.256	2.921	3.687	4.845					
		20	1.612	2.465	3.509						
		30	1.435	2.521	4.551						
		40	1.600	3.324							
	30	10	1.461	2.285	3.227						
		20	1.383	2.436	4.366						
		30	1.578	3.248							
	40	10	1.352	2.328	3.917						
		20	1.587	3.153							
	50	10	1.712	3.006							
4	20	10	2.227	2.831	3.368	4.137	4.892	7.375			
		20	1.575	2.337	3.109	4.178					
		30	1.392	2.334	3.615						
		40	1.545	2.878							
	30	10	1.424	2.171	2.863	3.673					
		20	1.340	2.257	3.400						
		30	1.524	2.816	4.658						
	40	10	1.312	2.168	3.082	4.756					
		20	1.535	2.759	4.571						
	50	10	1.665	2.746	4.493						

https://doi.org/10.2143/AST.39.2.2044640 Published online by Cambridge University Press

However, as in the discrete polygenotype model, we assume that the insurer adjusts the premium to reflect the composition of the risk pool. If everyone with genotype $(p,0)$ with $p < p^*$ has declined to insure and has left the risk pool, the actuarial premium for those who remain is:

$$\bar{\Pi}^c(p^*) = \frac{\int_{p^*}^{\infty} \Pi(p,0) f_{(P,M)}(p,0)\,dp + \int_{-\infty}^{\infty} \Pi(p,1) f_{(P,M)}(p,1)\,dp + \int_{-\infty}^{\infty} \Pi(p,2) f_{(P,M)}(p,2)\,dp}{1 - \int_{-\infty}^{p^*} f_{(P,M)}(p,0)\,dp}. \tag{8}$$

where the notation makes explicit the dependence of the actuarial premium on p^*. Following the same general idea as before, we find the threshold values of the polygenotype p^*, such that those with genotypes $(p,0)$ (for $p < p^*$) do not insure at premium $\bar{\Pi}^c(p^*)$, and everyone else does. These thresholds are given in Table 16 for utility Model 1 and in Table 17 for utility Model 2. They are markedly higher than those found previously. Indeed there are instances where nearly the entire BRCA0 subpopulation would decline to buy insurance. This could easily 'spill over' into the BRCA1 and BRCA2 subpopulations, although for the purposes of demonstration we have assumed it does not. Such behaviour would be disastrous for CI business.

However, we find that adverse selection does not usually occur under utility Models 3 and 4. It is so infrequent that instead of showing tables we list the exceptions here (all of which occur when $L/W = 0.1$).

(a) In Model 3 a policy with entry age 40 and term 10 years has $p^* = -3.74$ representing 0.1% of the population.
(b) In Model 4 a policy with entry age 30 and term 20 years has $p^* = -2.86$ representing 1% of the population, and a policy with entry age 40 and term 10 years has $p^* = 2.74$ representing 98.5% of the population.

4. Conclusions

4.1. Background

It is clear that a major development of genetics in future will concern the interactions of multiple genes contributing to common disorders. One of the very few epidemiological studies so far to estimate onset rates associated with a complex disorder is that of Antoniou *et al.* (2002). They modelled a polygene influencing BC risk, in addition to the known BRCA1/2 major genes. Risky polygenes are common enough that adverse selection becomes an option for a larger proportion of the population. Furthermore, the relative risks attributed to the polygene range from 0.04 to 23.62, many times more extreme than the assumptions in Macdonald, Pritchard & Tapadar (2006) for example. Our aim has been to apply this model to the study of adverse selection in a CI insurance market, both along the lines of previous studies and from more of an economic viewpoint.

https://doi.org/10.2143/AST.39.2.2044640 Published online by Cambridge University Press

TABLE 16

THE POLYGENOTYPE p^* AT WHICH ADVERSE SELECTION OCCURS UNDER THE DYNAMIC INSURER PRICING METHOD FOR A VARIETY OF POLICY ENTRY AGES AND TERMS, WITH $\sigma_R = 1.291$, $W = £100{,}000$ AND MODEL 1 UTILITY. THE FIGURES IN PARENTHESES SHOW THE PROPORTION OF THE POPULATION WHO WILL NOT PURCHASE INSURANCE.

Age	Term	Loss to Wealth Ratio								
		0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
20	10	3.25 (99.4%)	3.21 (99.3%)	3.15 (99.3%)	3.09 (99.2%)	$-\infty$ (0.0%)	$-\infty$ (0.0%)	$-\infty$ (0.0%)	$-\infty$ (0.0%)	$-\infty$ (0.0%)
	20	3.63 (99.6%)	3.60 (99.6%)	3.57 (99.6%)	3.53 (99.6%)	3.49 (99.5%)	3.43 (99.5%)	3.34 (99.4%)	3.22 (99.3%)	$-\infty$ (0.0%)
	30	3.40 (99.5%)	3.38 (99.5%)	3.35 (99.5%)	3.32 (99.4%)	3.27 (99.4%)	3.22 (99.3%)	3.14 (99.2%)	3.00 (99.0%)	2.68 (98.3%)
	40	3.25 (99.4%)	3.23 (99.3%)	3.20 (99.3%)	3.17 (99.3%)	3.13 (99.2%)	3.08 (99.2%)	3.00 (99.0%)	2.86 (98.8%)	$-\infty$ (0.0%)
30	10	3.68 (99.6%)	3.65 (99.6%)	3.62 (99.6%)	3.58 (99.6%)	3.54 (99.6%)	3.48 (99.5%)	3.40 (99.5%)	3.29 (99.4%)	3.06 (99.1%)
	20	3.41 (99.5%)	3.39 (99.5%)	3.36 (99.5%)	3.33 (99.4%)	3.28 (99.4%)	3.23 (99.3%)	3.15 (99.3%)	3.02 (99.1%)	2.72 (98.4%)
	30	3.25 (99.4%)	3.23 (99.3%)	3.21 (99.3%)	3.18 (99.3%)	3.14 (99.2%)	3.08 (99.2%)	3.01 (99.1%)	2.87 (98.8%)	$-\infty$ (0.0%)
40	10	3.42 (99.5%)	3.40 (99.5%)	3.37 (99.5%)	3.33 (99.4%)	3.29 (99.4%)	3.23 (99.3%)	3.15 (99.3%)	3.02 (99.1%)	2.73 (98.5%)
	20	3.26 (99.4%)	3.24 (99.4%)	3.21 (99.3%)	3.18 (99.3%)	3.14 (99.2%)	3.09 (99.2%)	3.01 (99.1%)	2.86 (98.8%)	$-\infty$ (0.0%)
50	10	3.32 (99.4%)	3.30 (99.4%)	3.27 (99.4%)	3.23 (99.3%)	3.18 (99.3%)	3.12 (99.2%)	3.03 (99.1%)	2.87 (98.8%)	$-\infty$ (0.0%)

TABLE 17

THE POLYGENOTYPE p^* AT WHICH ADVERSE SELECTION OCCURS UNDER THE DYNAMIC INSURER PRICING METHOD FOR A VARIETY OF POLICY ENTRY AGES AND TERMS, WITH $\sigma_R = 1.291$, $W = £100{,}000$ AND MODEL 2 UTILITY. THE FIGURES IN PARENTHESES SHOW THE PROPORTION OF THE POPULATION WHO WILL NOT PURCHASE INSURANCE.

Age	Term	Loss to Wealth Ratio								
		0.1	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
20	10	3.21 (99.3%)	3.11 (99.2%)	−∞ (0.0%)	−∞ (0.0%)	−∞ (0.0%)	−∞ (0.0%)	−∞ (0.0%)	−∞ (0.0%)	−∞ (0.0%)
	20	3.61 (99.6%)	3.55 (99.6%)	3.47 (99.5%)	3.37 (99.5%)	3.23 (99.3%)	−∞ (0.0%)	−∞ (0.0%)	−∞ (0.0%)	−∞ (0.0%)
	30	3.38 (99.5%)	3.33 (99.4%)	3.27 (99.4%)	3.18 (99.3%)	3.05 (99.1%)	2.81 (98.7%)	−∞ (0.0%)	−∞ (0.0%)	−∞ (0.0%)
	40	3.23 (99.3%)	3.18 (99.3%)	3.13 (99.2%)	3.05 (99.1%)	2.93 (98.9%)	−∞ (0.0%)	−∞ (0.0%)	−∞ (0.0%)	−∞ (0.0%)
30	10	3.66 (99.6%)	3.60 (99.6%)	3.52 (99.6%)	3.43 (99.5%)	3.30 (99.4%)	3.08 (99.2%)	−∞ (0.0%)	−∞ (0.0%)	−∞ (0.0%)
	20	3.39 (99.5%)	3.34 (99.4%)	3.28 (99.4%)	3.19 (99.3%)	3.06 (99.1%)	2.83 (98.7%)	−∞ (0.0%)	−∞ (0.0%)	−∞ (0.0%)
	30	3.23 (99.3%)	3.19 (99.3%)	3.13 (99.2%)	3.06 (99.1%)	2.94 (98.9%)	−∞ (0.0%)	−∞ (0.0%)	−∞ (0.0%)	−∞ (0.0%)
40	10	3.40 (99.5%)	3.35 (99.5%)	3.28 (99.4%)	3.19 (99.3%)	3.05 (99.1%)	2.81 (98.7%)	−∞ (0.0%)	−∞ (0.0%)	−∞ (0.0%)
	20	3.24 (99.4%)	3.20 (99.3%)	3.14 (99.2%)	3.06 (99.1%)	2.94 (98.9%)	−∞ (0.0%)	−∞ (0.0%)	−∞ (0.0%)	−∞ (0.0%)
50	10	3.30 (99.4%)	3.24 (99.4%)	3.17 (99.3%)	3.08 (99.2%)	2.92 (98.9%)	−∞ (0.0%)	−∞ (0.0%)	−∞ (0.0%)	−∞ (0.0%)

4.2. Multi-State Market Models

The most striking conclusion is the magnitude of the possible adverse selection costs associated with the polygene. In the worst case (Scenario (a) small insurance market, high level of genetic testing) the necessary premium increase approaches 17%. We must bear in mind that this cost arises solely from the contribution of BC/OC, and that there are several other genetic disorders which must have a polygenic component when major genes have been allowed for. This contrasts with the findings of most studies of major genes alone, where the rarity of mutations keeps adverse selection costs very small.

The inflated costs in the small market are perhaps a cause for concern. Should polygene testing become available, youthful insurance markets may be exposed to high costs from a moratorium on genetic testing (even at modest levels of testing).

The contrast between Tables 4 and 6 is perhaps the most revealing. In Table 4 we assumed all genetic tests to follow the development of a relevant family history. The premium increases, even under severe adverse selection, were of the order of 0.001%. In Table 6 we assumed that testing for polygenes was freed from this regime. Even under more moderate adverse selection, the premium increases were of the order of 1%, a thousand-fold increase.

Nevertheless, these are worst case outcomes. They assume the existence of accurate, freely available genetic tests, no treatments to modify the known risks, active and well-informed market participants (who can all afford insurance) and so on. And, women may not be all that likely to curtail their cover because of a beneficial test result, because they would lose cover against all other illnesses as well as BC. Reality is likely to fall well short of the worst outcome even if conditions would let it exist.

4.3. An Economic Framework for Adverse Selection

Introducing the utility framework, with informed purchasers *versus* risk-pooling insurers, characterises adverse selection in another way, by mapping out when it ought or ought not to appear. This approach too has many assumptions and simplifications, including single premiums, perfectly rational agents, universal knowledge (on the part of individuals) of genotype, fixed insurance coverage and a known utility function. We suspect, nevertheless, that blurring all these in more realistic ways is liable to reduce rather than to increase adverse selection.

Our chief observation is the strong dependence on the utility function assumed. It is comforting that those that were parameterised using empirical data (Models 3 and 4) showed adverse selection to be very limited. In most cases, it would not appear at all under the assumed conditions. However, those utilities chosen just to illustrate lower risk aversion (Models 1 and 2) led to adverse selection in most cases, including sometimes for polygenic risks less extreme than those of the empirical BC model.

Unlike the multi-state market model, the economic model does not let us estimate what the costs of adverse selection would be. However, under what might be the most realistic assumptions, of Section 3.6, we could estimate the proportion of the population who would leave the CI market under equilibrium. Under Models 1 and 2 this proportion was generally very high, in many cases the market effectively disappearing. Under Models 3 and 4 this proportion was nearly always zero, adverse selection being absent. Clearly, relatively modest shifts in preferences as expressed by utilities can move the market from one extreme to the other.

Acknowledgements

This work was carried out at the Genetics and Insurance Research Centre at Heriot-Watt University. We would like to thank the sponsors for funding, and members of the Steering Committee for helpful comments at various stages. KM was funded by the Engineering and Physical Sciences Research Council.

References

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

Eisenhauer, J.G. and Ventura, L. (2003) Survey measures of risk aversion and prudence. *Applied Economics*, **35(13)**, 1477-1484.

Gui, E.H., Lu, B., Macdonald, A.S., Waters, H.R. and Wekwete, C.T. (2006) The genetics of breast and ovarian cancer III: A new model of family history with insurance applications. *Scandinavian Actuarial Journal*, **2006**, 338-367.

Gutierrez, M.C. and Macdonald A.S. (2007) Adult polycystic kidney disease and insurance: A case study in genetic heterogeneity. *North American Actuarial Journal*, **11(1)**, 90-118.

Hoy, M. and Witt, J. (2007) Welfare effects of banning genetic information in the life insurance market: The case of BRCA1/2 genes. *Journal of Risk and Insurance*, **74(3)**, 523-546.

Lange, K. (1997) An approximate model of polygenic inheritance. *Genetics*, **147**, 1423-1430.

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

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

Macdonald, A.S. and McIvor, K.R. (2006) Application of a polygenic model of breast and ovarian cancer to critical illness insurance. *Annals of Actuarial Science*, **1**, 319-343.

Macdonald, A.S., Pritchard, D.J. and Tapadar, P. (2006) The impact of multifactorial genetic disorders on critical illness insurance: A simulation study based on UK Biobank. *ASTIN Bulletin*, **36**, 311-346.

Macdonald, A.S. and Tapadar, P. (2009) Multifactorial genetic disorders and adverse selection: Epidemiology meets economics. *To appear in Journal of Risk and Insurance*.

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

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

Ropka, M.E., Wenzel, J., Phillips, E.K., Siadaty, M. and Philbrick, J.T. (2006) Uptake rates for breast cancer genetic testing: A systematic review. *Cancer Epidemiology Biomarkers and Prevention*, **15(5)**, 840-855.

STRACHAN, T. and Read, A.P. (2004) *Human Molecular Genetics 3*. Garland Publishing.
SUBRAMANIAN, K., LEMAIRE, J., HERSHEY, J.C., PAULY, M.V., ARMSTRONG, K. and ASCH, D.A. (1999) Estimating adverse selection costs from genetic testing for breast and ovarian cancer: The case of life insurance. *Journal of Risk and Insurance*, **66**, 531-550.

ANGUS MACDONALD
Department of Actuarial Mathematics and Statistics
and the Maxwell Institute for Mathematical Sciences
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

https://doi.org/10.2143/AST.39.2.2044640 Published online by Cambridge University Press