A MATHEMATICAL MODEL OF ALZHEIMER'S DISEASE AND THE APOE GENE

By

ANGUS MACDONALD AND DELME PRITCHARD

Heriot-Watt University, Edinburgh

ABSTRACT

Alzheimer's disease (AD) accounts for a significant proportion of long-term care costs. The recent discovery that the $\varepsilon4$ allele of the ApoE gene indicates a predisposition to earlier onset of AD raises questions about the potential for adverse selection in long-term care insurance, about long-term care costs in general, and about the potential effects on costs of gene therapy, or better targetted treatments for AD. This paper describes a simple Markov model for AD, and the estimation of the transition intensities from the medical and epidemiological literature.

KEYWORDS

Alzheimer's Disease, Apolipoprotein E Gene, Genetics, Markov Models.

1. INTRODUCTION

Molecular genetics has far-reaching implications for all aspects of health economics, including the effectiveness, or even practicability, of insurance-based funding of all forms of care. In particular, if genetic test results are known by the individual but not by the insurer, there may be a tendency for those most at risk to buy more insurance, known as adverse selection. This is a natural subject for quantitative modelling.

Such discussion as has taken place (in the United Kingdom) has tended to be confined to life insurance, in which context it has been suggested that the overall costs of adverse selection might be limited (Macdonald, 1997, 1999; Pritchard, 1997). However, these papers acknowledge:
(a) that different conclusions might hold in respect of other forms of insurance; and
(b) the lack of sound epidemiological data in respect of any but a few genetic conditions.

Thus, the model proposed by Macdonald (1997, 1999) was based on parameter estimates intended to be extreme, and to suggest an upper bound on the cost of adverse selection, rather than on data-based statistical estimates. Tan (1997) applied a similar model to annuity business, with results that suggested higher costs.

The U.K. Government has so far avoided legislating on the very sensitive issue of the use of genetic test information by insurers, but instead has set up the Genetics and Insurance Committee (GAIC), charged with the assessment of the likely relevance and reliability of genetic test information as it relates to different kinds of insurance. Such considerations require actuarial models of the insurance process, allowing for the effects of specific genes on mortality and morbidity. To date, the only such study is Lemaire et al. (1999) and Subramanian & Lemaire (1999), treating breast and ovarian cancer and the BRCA1 and BRCA2 genes.

In this paper we propose a simple Markov model for Alzheimer's disease (AD) and estimate its transition intensities from medical and epidemiological studies. Genetic variation arises because the $\varepsilon 4$ allele of the Apolipoprotein E (ApoE) gene is known to indicate a predisposition to earlier onset of AD.

In Section 2 we give a brief outline of genetics terminology, then describe AD and the evidence for a genetic component of AD. In Section 3 we specify the model, and in Sections 4 and 5 we estimate the transition intensities. We show some results in Section 6, and our conclusions are in Section 7.

Applications of the model include the study of long-term care insurance, (Macdonald & Pritchard, 1999) and the study of long-term care costs (Warren et al., 1999). The processes leading to LTC insurance claims are complex, when compared with other forms of insurance, and there are no insurance data to speak of; therefore we (and insurance companies themselves) have to rely on data collected and published for a variety of reasons, mostly in the medical literature. As a result, our model is far from definitive, but we found the process of extracting information from the medical literature and putting it to actuarial use very instructive, and we suggest that any shortcomings of our model shed useful light on the problems that might be faced in the future.

2. THE GENETICS OF ALZHEIMER'S DISEASE

For a recent survey of the genetic epidemiology of AD, see Slooter & van Duijn (1997); Breteler et al. (1992), reviews the position before much was known about the genetic component of AD.

2.1. A Brief Primer on Human Genetics

The following discussion gives only the briefest definitions of genetic terms, and omits all complications. For a proper account, see a text such as Strachan & Read (1999).

(a) Each cell of the human body (except sperm and egg cells) has a nucleus containing 23 pairs of chromosomes. Each cell carries an identical set of chromosomes at birth, unless some have been damaged.

(b) A chromosome is a very long double strand (the famous double helix) or linear molecule, made up of a sequence of base pairs of DNA.

(c) A gene is a sequence of base pairs at a fixed location on a given chromosome that acts as a functional unit, coding for the production of a protein (usually). The term 'gene' strictly refers to the locus on the chromosome.

(d) The gene at a given locus might have several variants (rather as chemical elements might have several isotopes); these are known as alleles. Different alleles are simply different sequences of base pairs. Differences between alleles are the result of mutations, namely alterations to a gene caused by (for example) errors in replication of the DNA when cells divide. If the mutation occurs in an adult (somatic) cell, it can only spread through cell division (which is what happens in cancers). If the mutation occurs in a gene carried by a sperm or egg cell, it can be passed to offspring.

(e) Because the chromosomes are paired (except the X and Y chromosomes that determine sex) each cell has two alleles of every gene, one inherited from each parent. They may be the same (homozygosity) or different (heterozygosity).

(f) Mutant genes will encode variants of the protein that the gene produces. These may have a beneficial, neutral or adverse effect on the cell or organism. If the altered protein is radically different, the cell will probably die, or the organism will fail to develop, or it will not survive to reproductive age. Less radical variations may manifest themselves as susceptibility to disease, or may be harmless but non-functional.

(g) The great range of genetic diseases arises from the range of effects of the protein products of different alleles, and the simple combinatorics of inheritance.

   (1) If an allele encodes a harmless but non-functional protein product, the disease will appear only in homozygous individuals (autosomal recessive inheritance, such as cystic fibrosis).

   (2) Heterozygous cells will produce a mixture of variants of the protein product; if just one of these is lethal it will cause disease (autosomal dominant inheritance, such as Huntington's disease).

   (3) In between these extremes, alleles encode protein products that are more or less dangerous or fully-functional, the effect often depending on the environment. Then different levels of susceptibility to disease will appear, and homozygous individuals will often be more susceptible than heterozygous individuals.

(h) Lifestyle and environment can affect the potency of susceptibility genes. For example, the activity of the protein produced by a dangerous allele may be enhanced in certain environments, or a protective allele may be put at greater risk of being knocked out by a mutation.

(i) Although the outline above mentions only single genes, most processes in the human body are complicated and involve very many proteins, each encoded by its own gene with its various alleles. Most genetic disease will result from combinations of several alleles, lifestyle and environment; the term for this is multifactorial disorder.

2.2. Cognitive Impairment

The term 'cognitive impairment' covers AD, which accounts for by far the greater number of cases, and other forms of mental deterioration, chiefly vascular in origin (for example, arising from strokes). Assessment is liable to be imprecise, making it difficult to decide on an exact date of inception of cognitive impairment, if such a thing exists. Moreover, although AD is the commonest form of cognitive impairment, it is hard to diagnose with certainty except by post-mortem examination. These factors introduce considerable uncertainty into epidemiological studies of AD. Breteler *et al.* (1992) noted that:

(a) AD itself can have a significant vascular component;
(b) some of the neuropathological symptoms of AD (see Section 2.4) can also be symptoms of vascular dementia; and
(c) studies by Tierney *et al.* (1988) found that post-mortem examination confirmed only 64-86% of diagnoses of AD.

2.3. The Pathology of Alzheimer's Disease

The pathology of AD includes:
(a) senile plaques (deposits on the outside of neurones (brain cells), consisting largely of the protein β-amyloid);
(b) neurofibrillary tangles (connections between neurones);
(c) amyloid angiopathy (deposits of amyloid protein in the arteries of the brain);
(d) loss of neurones; and
(e) decreased activity of choline acetyltransferase (an enzyme).

Therefore, any gene whose expression leads to the production, or over-production, of substances associated with these changes is potentially a genetic marker for AD.

2.4. General Evidence of a Genetic Contribution to Alzheimer's Disease

AD is a disease of old age; it is rare below ages 60-70. These rare cases are called 'early-onset' AD, which should not be confused with early onset of AD within the usual age range. We are concerned only with the latter.

Families with a history of AD are sometimes observed, but AD also occurs sporadically (that is, in the absence of a family history of AD) and it

is always possible that a case of AD in an affected family is, in fact, sporadic. The differences between familial and sporadic AD are not clear, although the former may be marked by earlier onset and more rapid progression.

A very few families have several cases of early-onset AD in several generations, consistent with autosomal dominant transmission (Levy-Lahad & Bird, 1996), and three genes have been found. First was the gene encoding for amyloid precursor protein (APP), involved in the production of β-amyloid. It resides on chromosome 21, which is the chromosome affected in cases of Down syndrome, sufferers of which often develop AD in middle age. Several mutations have been found, but they are rare. Later, mutations in two genes labelled presenilin-1 (PS-1) and presenilin-2 (PS-2) were identified, which appeared to be associated with AD, though the mechanisms remain unclear.

Familial AD is not restricted to early-onset cases and family history remains an important risk factor for late-onset AD (Jarvik et al., 1996). Susceptibility genes have been identified, of which the most studied is that which codes for apolipoprotein E.

2.5. The Apolipoprotein E Gene

This summary is not comprehensive, but we hope that it is detailed enough to give an actuarial reader an impression of the progress made in understanding genetic factors of AD, as well as some of the problems and (perhaps most important) the great speed at which human genetics is advancing.

Apolipoprotein E (ApoE) is found in senile plaques and neurofibrillary tangles in AD patients. It has also been studied because of its rôle in lipid metabolism. The gene that encodes it is on chromosome 19, which was linked to families with late-onset AD by Pericak-Vance et al. (1991), making it a clear candidate gene for familial AD. Strittmatter et al. (1993) confirmed this hypothesis, which was rapidly supported by many other studies. The basic facts are as follows:

(a) The ApoE gene has three common alleles – $\varepsilon 2$, $\varepsilon 3$ and $\varepsilon 4$ – whose frequencies are roughly 0.09, 0.77 and 0.14 respectively.

(b) Since each offspring receives one allele from each parent, there are six possible genotypes ($\varepsilon 2/\varepsilon 2$, $\varepsilon 2/\varepsilon 3$, $\varepsilon 2/\varepsilon 4$, $\varepsilon 3/\varepsilon 3$, $\varepsilon 3/\varepsilon 4$ and $\varepsilon 4/\varepsilon 4$). Offspring with two copies of the same allele are called homozygotes, while those with two different alleles are called heterozygotes.

(c) The ApoE $\varepsilon 4$ allele increases the risk of AD in a dose related fashion, such that $\varepsilon 4$ homozygotes ($\varepsilon 4/\varepsilon 4$) are at greater risk than $\varepsilon 4$ heterozygotes ($\varepsilon 2/\varepsilon 4$, $\varepsilon 3/\varepsilon 4$), who in turn are at greater risk than those without the $\varepsilon 4$ allele (Bickeböller et al., 1997; Corder et al., 1994; van Duijn et al., 1995; Farrer et al., 1997; Jarvik et al., 1996; Kuusisto et al., 1994; Lehtovirta et al., 1995; Mayeux et al., 1993; Myers et al., 1996; Poirier et al., 1993; Tsai et al., 1994). See Section 5.2

for risk estimates. The risk depends on age, being highest at ages 60-70, tapering off at older ages (Bickeböller *et al.*, 1997; Corder *et al.*, 1994; Farrer *et al.*, 1997).

(d) It is also possible that the ε4 allele is associated with earlier onset of AD (not to be confused with early-onset AD). The effect may be dose dependent (Farrer *et al.*, 1997; Frisoni *et al.*, 1995; Gomez-Isla *et al.*, 1996); or not (Lehtovirta *et al.*, 1995; Stern *et al.*, 1997; Corder *et al.*, 1995); or it may not exist at all (Liddell *et al.*, 1994; Masullo *et al.*, 1998; Norrman *et al.*, 1995).

(e) Investigations into the rate of mental decline of AD patients by genotype found no evidence for any difference (Basun *et al.*, 1995; Gomez-Isla *et al.*, 1996; Masullo *et al.*, 1998; Norrman *et al.*, 1995). There is conflicting evidence about mortality. It is possible that younger age at onset should imply longer survival times, because of the usual age-related mortality differentials, and therefore that the ε4 allele should be associated with longer life after onset of AD. While some studies support this (Corder *et al.*, 1995; Gomez-Isla *et al.*, 1996; Norrman *et al.*, 1995), others have found no difference (Basun *et al.*, 1995; Stern *et al.*, 1997). If ε4 is associated with lighter mortality in AD patients then risk estimates from cross-sectional studies (the vast majority to date) should be interpreted with caution. An incidence study (Evans *et al.*, 1997) confirmed ε4 to be a significant risk factor, but the estimated increased risk of onset was at the lower end of the reported range.

(f) In contrast, the ε2 allele has been found to have a protective effect against late-onset AD (Corder *et al.*, 1994; Farrer *et al.*, 1997; Gomez-Isla *et al.*, 1996; Jarvik *et al.*, 1996; Lambert *et al.*, 1998; Masullo *et al.*, 1998). However, a study of early-onset AD patients (van Duijn *et al.*, 1995), found a higher frequency of the ε2 allele, and an association of ε2 with a more aggressive form of AD, suggesting different rôles of ApoE in early-onset and late-onset AD. Findings relating to the ε2 allele are based on the ε2/ε3 genotype, as ε2 homozygotes are rare. The risk attached to the ε2/ε4 genotype is not clear, possibly because ε2 and ε4 have opposite effects (Jarvik *et al.*, 1996; Levy-Lahad *et al.*, 1996).

ApoE ε4 is the most important genetic risk factor for AD identified yet. Though it is neither necessary nor sufficient to cause AD it does increase susceptibility. Approximately 26% of Caucasians carry at least one ε4 allele and it has been estimated that between 42% and 79% of AD cases are attributable to the associated excess risk (Nalbantoglu *et al.*, 1994).

2.6. Other Genetic Factors of Alzheimer's Disease

In 1997, a gene for the K-variant of butyrylcholinesterase (BCHE K), not a risk factor by itself, was found to act in synergy with ApoE ε4, such that carriers of both (an estimated 6% of Caucasians) were at over 30 times the risk of AD as a person with neither (Lehmann *et al.*, 1997). Subsequent

studies (Brindle *et al.*, 1998; Singleton *et al.*, 1998) failed to reproduce the result. Although some explanations have been advanced, caution is advisable in using BCHE K as a risk factor for AD.

Payami *et al.* (1997) reported an association between AD and the A2 allele of the human leukocyte antigen (HLA); the HLA-A2 phenotype and ApoE $\varepsilon4/\varepsilon4$ genotype had similar and additive effects on reducing age at onset of AD, at ages below 60 and above 75. Further studies would be needed to confirm these findings.

Poduslo *et al.* (1998) found the apolipoprotein CI (apo CI) gene to be a risk factor for early-onset and late-onset AD, whether sporadic or familial. Apo CI A homozygotes had 4 to 5 times the odds of developing AD, heterozygotes about twice the risk. This was not unexpected, since Apo CI is closely linked to ApoE and in linkage disequilibrium with ApoE and AD. Linkage disequilibrium is the non-random assortment, in a population, of two genes on the same chromosome (the strength of the linkage is inversely proportional to the distance between them). It was thought that the association of AD with ApoE may be more significant.

3. A Model for Alzheimer's Disease

3.1. Model Specification

We use a continuous-time multiple state model. For general comments on these models, see Macdonald (1996a) or Waters (1984). Lives with each ApoE genotype are assumed to form a homogeneous population, suffering the different risks of AD discussed in Sections 2.5 and 5.2.

An important reason for using these models is that they allow the most complete representation of the underlying process. It is then necessary to estimate a large number of transition intensities, for which adequate data do not always exist, but it is preferable to obtain a clear picture of the data needed than to sweep the issue under the carpet by working with some less adequate model in the first place. In particular:

(a) if some simpler model is eventually recommended for use, because of missing data or for computational convenience, it is important to be able to assess its soundness in practice; and

(b) if missing data become available later, for example, as the insured lives experience develops, it is a hindrance if too much has been invested in a model that cannot incorporate it.

Modern computing power is such that the computational demands of multiple state models (numerical integration of differential equations) can quite reasonably be met, for arbitrarily complex Markov models (Norberg, 1995) and for many semi-Markov models (Waters & Wilkie, 1987; Waters, 1991). The techniques can all be found in standard texts on numerical analysis, and no actuary should be prevented from choosing an adequate model by the need to use them.

Figure 1 shows a simple model of AD. Each genotype is represented by such a model; the transition intensities in each model will differ, representing the different genetic risks. x denotes the age at outset (for example, when insurance is purchased) and t the elapsed duration. The choice of states is dictated entirely by the events that have been studied in the medical and epidemiological literature. For certain purposes, it would be desirable to model other events, such as the start of a long-term insurance claim. No data about that event are available; however a major event that has been studied is institutionalisation. Although becoming institutionalised need not coincide with the start of an insurance claim, it is the best available proxy.

Macdonald (1999) considered frailty models as an alternative to Markov models, for genetics and insurance applications. They offer the advantage of a simple model of the genetic variability, if that is justified by the circumstances. They may be especially useful for modelling multifactorial disorders, or genes with very many alleles or mutations, but for a single gene with just a few alleles it seems reasonable to model each separately. Other possible models (such as Cox-type models) might be useful for modelling individual transitions but do not lend themselves to the inclusion of payments contingent upon complicated life histories.
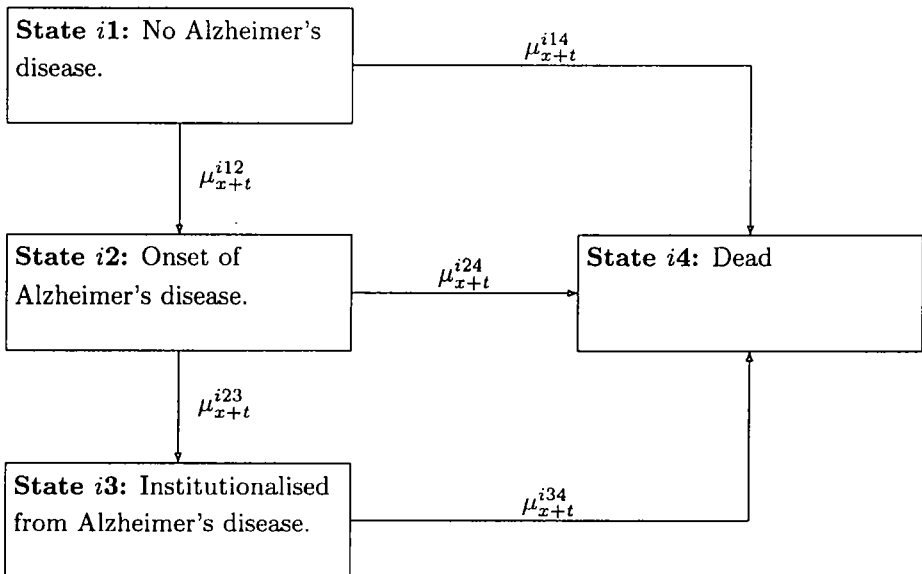

FIGURE 1: A simple model of Alzheimer's disease in the ith of M subgroups, each representing a different ApoE genotype. x is the age at outset, and t the elapsed duration.

3.2. An Expanded Long-Term Care Model

AD alone does not account for all long-term care costs. Broadly speaking, the need for care arises because of cognitive disorder (including AD) or loss of ability to perform Activities of Daily Living (ADLs) such as dressing, washing and feeding. A comprehensive model of long-term care costs can be specified in terms of these causes, with AD included as a component, and the impact of the ApoE gene on overall care costs can thereby be studied. However, incorporating AD explicitly in an expanded model will require data that describe, at the level of individual lives, the progress of AD and the loss of ADLs. Until such data are available, we can only estimate overall care costs by crude approximations: see Macdonald & Pritchard (1999).

4. ESTIMATION OF TRANSITION INTENSITIES NOT DEPENDING ON APOE GENOTYPE

In this section we estimate the transition intensities for the events: onset of AD; institutionalisation; and death. All of these must be 'estimated' from results reported in the medical and epidemiological literature. Ideally, we would work with the original data, but these are almost never available. Reported results are not always ideal for the extraction of parameters for an actuarial model; often the age groups used are very wide, and not the same in different surveys; sometimes only graphs (such as Kaplan-Meier survival curves) are given.

4.1. Prospective and Case-Based Studies

A most important distinction must be drawn when estimating transition intensities from epidemiological studies (see, for example, Kahn & Sempos (1989), Clayton & Hills (1993), Lilienfeld & Hills (1993), Selvin (1996)).
(a) *Prospective* studies, based on samples of the general population, ought to yield the most reliable estimates of population risk, but are expensive and time-consuming. Moreover, they are rarely even begun until substantial evidence of an effect has been accumulated from other studies.
(b) *Case-based* studies, based on affected persons (and controls) often yield relative risks greatly in excess of the true population risks, precisely because the subjects are affected or at risk. However, early studies into any medical condition are almost inevitably of this type.

Our current knowledge of most genetic disorders is derived from case-based studies; this is certainly true of ApoE and AD (see Section 2.5). It is very likely that estimates of risk conferred by ApoE genotype will fall as more prospective studies are carried out (see the comment on Evans *et al.* (1997) at the end of Section 2.5(e)), but this will take time.

The approach we adopt is as follows:
(a) in Section 4.2, we state assumptions about the general level of mortality;
(b) in Section 4.3, we estimate the aggregate incidence of AD, denoted μ_{x+t}^{AD}, which has been investigated extensively;
(c) in Section 4.4, we estimate the intensity of institutionalisation, following the onset of AD (that is, μ_{x+t}^{i23}) and the force of mortality following the onset of AD (that is, μ_{x+t}^{i24});
(d) in Section 4.5, we estimate the force of mortality for lives institutionalised with AD (that is, μ_{x+t}^{i34});
(e) in Section 5.1, we estimate the population frequencies of the ApoE alleles; and
(f) in Section 5.2, we estimate the incidence of AD for each genotype using odds ratios from the genetic studies: this gives estimates of μ_{x+t}^{i12}.

4.2. Baseline Mortality Tables

For convenience, we choose parametric approximations to the AM80 and AF80 Ultimate mortality tables as bases for mortality assumptions; for use in the model they are adjusted in a variety of ways. Gompertz curves were fitted to μ_{x+t} at ages 65-120, using log-linear least squares (see equation (1)):

$$^{AM\,80}\mu_{x+t} = 0.0000941\,16e^{0.084554(x+t)} \tag{1}$$

$$^{AF80}\mu_{x+t} = 0.000025934e^{0.093605(x+t)}.$$

Experiments with the AM80 and AF80 tables themselves showed that the Gompertz approximations had a negligible effect in long-term care applications; we use them because they are sometimes useful in numerical work. For insurance use, some allowance must be made for future improvements in mortality. No experience is available to help, but following discussion with some actuaries experienced in pricing long-term care insurance, we have chosen 65% of these baseline tables as the aggregate mortality assumptions.

4.3. The Onset of Alzheimer's Disease in the Population

AD has been the subject of some large-scale epidemiological studies, many of them pre-dating the discovery of the ApoE gene. Some of these report incidence rates, or 'occurrence/exposure' rates, which are exactly the estimates we need for transition intensities.

There is general agreement, in this literature, on the shape of the intensity μ_{x+t}^{AD} in the age range 60-85 years; it is very low at ages 60-64 (about 0 to 0.002) and increases rapidly with age, approximately doubling every 5 years. Sayetta (1986) and Hebert et al. (1995) found that a Gompertz curve gave the best fit, despite trying a number of more complex models.

A number of studies report the incidence of AD (that is, the intensity μ_{x+t}^{AD}) by age but not by genotype, including Copeland *et al.* (1992), Hagnell *et al.* (1992), Kokmen *et al.* (1993), Letenneur *et al.* (1994), Nilsson (1984), Ott *et al.* (1998), Rocca *et al.* (1998) and Rorsman *et al.* (1986). Of particular interest, however, is the recent meta-analysis of the incidence of AD by Jorm & Jolley (1998):

(a) it draws on 23 studies world-wide, including 13 European studies;

(b) the analysis is carried out separately for Europe, the U.S.A. and East Asia;

(c) the incidence of AD is estimated by severity, categorised as Mild + and Moderate + AD, where Mild + includes all cases classified as mild or worse; and

(d) point estimates of μ_{x+t}^{AD} were obtained for 5-year age groups from 65 to 95, and no *a priori* shape of μ_{x+t}^{AD} was assumed.

We estimated μ_{x+t}^{AD} from Jorm & Jolley (1998) using the figures from the European studies and for Mild + AD. The estimates, 95% confidence limits and the log-linear least squares Gompertz fit:

$$\mu_{x+t}^{AD} = 1.31275 \times 10^{-7} e^{0.145961(x+t)} \tag{2}$$

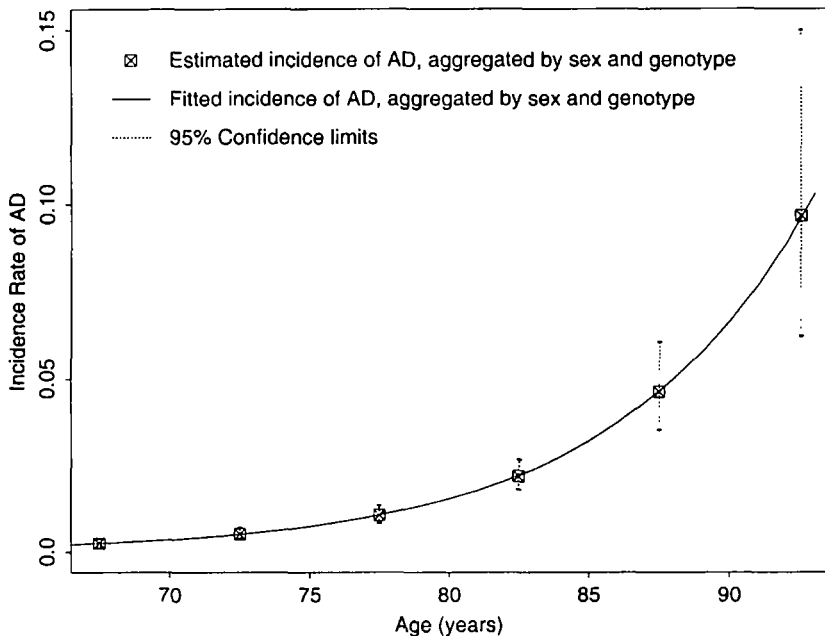are shown in Figure 2. It is clear that a Gompertz curve is a very good fit.

FIGURE 2: Aggregate incidence of Alzheimer's disease: point estimates and 95% confidence intervals.
Source: Jorm & Jolley (1998).

Data on the incidence of AD among the very elderly (> 90 years) are sparse, so estimates at these ages have wide confidence intervals and the trend is uncertain. The meta-analysis by Gao *et al.* (1998) found that the rate of increase in μ^{AD}_{x+t} slowed down with age, but other studies found no evidence of this (Hebert *et al.*, 1995; Jorm & Jolley, 1998; Letenneur *et al.*, 1994). We have simply extrapolated the Gompertz formula above to all ages; the effect of this assumption will depend on the particular application, or type of insurance, and this should be investigated when the model is used (see, for example, Macdonald & Pritchard (1999)).

Many studies have found men and women to be at the same risk of AD (Kokmen *et al.*, 1993; Nilsson, 1984; Ott *et al.*, 1998; Rocca *et al.*, 1998) and, when differences have been reported (Gao *et al.*, 1998; Jorm & Jolley, 1998; Letenneur *et al.*, 1994), women were found to be at greater risk only at very old ages. Some experiments (not described here) in applying the model to AD-related long-term care insurance costs using different rates of AD for men and women showed that it made little difference, and here we have used the aggregate rate (equation (2)).

4.4. Time from Onset of Alzheimer's Disease to Institutionalisation or Death

The available data do not allow us to analyse μ^{i23}_{x+t}, μ^{i24}_{x+t} or μ^{i34}_{x+t} by genotype.

Table 1 summarises the literature on time to the first of institutionalisation or death ('first event') for AD patients. Some studies give times from entry to the study rather than from onset, which is usually not observed directly. A striking feature is that few lives die before becoming institutionalised. This may seem surprising as AD patients have generally been reported to suffer higher mortality than healthy lives (see Section 4.5). However, AD's debilitating effects are not sudden, and we may expect patients to be in receipt of informal care between onset and institutionalisation, which might lead μ^{i24}_{x+t} to be relatively light.

We used the data from the study by Jost & Grossberg (1995). Although it is not the largest study, it does have advantages:
(a) it is a brain bank study, so all AD cases were confirmed by autopsy (the only reliable method of diagnosis);
(b) there were no censored cases; and
(c) the time from onset to institutionalisation is estimated.

TABLE 1

MEAN AND MEDIAN TIMES TO INSTITUTIONALISATION (INST'N) OR FIRST EVENT FOR AD PATIENTS

Reference	Age at		Time (years) to		% for which 1st event is death
	Onset	Entry	Inst'n	1st Event	
Berg et al. (1988)		71.4 (1)			7.1%
Heyman et al. (1997)		72.0 (2)		3.1 (2)	13.1%
Jost et al. (1995)	75.1 (3)		4.3 (1)		15.0%
Severson et al. (1994)		79.4 (1)	2.5 (2) (4)		10.0%

(1) Mean.

(2) Median.

(3) Mean age at onset of AD, if institutionalised, estimated as (mean age at institutionalisation − mean time to institutionalisation).

(4) Median time from onset estimated as 5.6 years.

Since we cannot distinguish genotypes here, we will just write μ_{x+t}^{23} and μ_{x+t}^{24} instead of μ_{x+t}^{i23} and μ_{x+t}^{i24}, respectively. Guided by these data we derive moment estimates of μ_{x+t}^{23} (the force of institutionalisation) and μ_{x+t}^{24} (the force of mortality of an AD patient prior to institutionalisation). We define below the usual indicator functions (I_j) and sample path functions (N_{jk}) in respect of a single life (see Macdonald (1996b)):

$$I_j(t) = \begin{cases} 1 & \text{if life is in state } j \text{ at time } t \\ 0 & \text{otherwise} \end{cases}$$

$$dN_{jk}(t) = \begin{cases} 1 & \text{if life transfers from state } j \text{ to state } k \text{ at time } t \\ 0 & \text{otherwise} \end{cases}$$

$$N_{jk}(T) = \int_0^T dN_{jk}(t) = \text{No. of transfers from state } j \text{ to state } k$$

Also let $P_{x,y}^{jj}$ be the probability that a life in state j at age x is in state j at age y. Then equation (3) below is the mean age at onset of AD, given that the life was eventually institutionalised with AD (as in Jost & Grossberg (1995)):

$$E\left[x + \int_x^\omega I_1(t)dt \,\Big|\, N_{23}(\omega - x) = 1 \text{ and } I_1(x) = 1\right] =$$

$$x + \frac{\int_x^\omega (t-x)\mu_t^{12} P_{xt}^{11}\{\int_t^\omega \mu_s^{23} P_{ts}^{22} ds\} dt}{\int_x^\omega \mu_t^{12} P_{xt}^{11}\{\int_t^\omega \mu_s^{23} P_{ts}^{22} ds\} dt} \;; \tag{3}$$

equation (4) is the mean time from onset of AD to institutionalisation:

$$E\left[\int_x^\omega I_2(t)dt \,\Big|\, N_{23}(\omega - x) = 1 \text{ and } I_1(x) = 1\right] =$$

$$\frac{\int_x^\omega \mu_t^{12} P_{xt}^{11}\left\{\int_t^\omega (s-t)\mu_s^{23} P_{ts}^{22} ds\right\}dt}{\int_x^\omega \mu_t^{12} P_{xt}^{11}\left\{\int_t^\omega \mu_s^{23} P_{ts}^{22} ds\right\}dt};$$

and equation (5) is the probability that an AD patient dies before becoming institutionalised. The upper age bound, denoted ω, is taken to be 120 years:

$$P\left[N_{24}(\omega - x) = 1 \,\Big|\, N_{12}(\omega - x) = 1 \text{ and } I_1(x) = 1\right] =$$

$$\frac{\int_x^\omega \mu_t^{12} P_{xt}^{11}\left\{\int_t^\omega \mu_s^{24} P_{ts}^{22} ds\right\}dt}{\int_x^\omega \mu_t^{12} P_{xt}^{11} dt} \tag{5}$$

Setting equations (3), (4) and (5) equal to their estimated values from Table 1, we obtain 3 equations, which can be solved for at most 3 unknown parameters. The parametric forms we chose were as follows:

(a) $\mu_{x+t}^{12} = A + \mu_{x+t}^{AD}$, where μ_{x+t}^{AD} is given by equation (2). This Makeham term adjusts the incidence of AD to a level that gives the same mean age at onset (for AD patients who become institutionalised).

(b) $\mu_{x+t}^{23} = D$. We felt that the data did not support anything more elaborate than a constant intensity.

(c) $\mu_{x+t}^{24} = P\mu_{x+t}^{14}$. That is, the mortality of an AD patient before becoming institutionalised is a proportion of baseline mortality.

(d) μ_{x+t}^{14}, baseline mortality, was taken as AM80 mortality, using the Gompertz approximation given by equation (1). Although it is appropriate to allow for future improvements in mortality in applications, it is not appropriate to do so in estimation based on past data. The values of D and P do not depend strongly on the baseline mortality.

Solving these equations numerically yields the solutions:

$$A = 0.02025038 \quad D = 0.18895779 \quad P = 0.33502488.$$

The transition intensities μ_{x+t}^{23} and μ_{x+t}^{24} are summarised in Table 3.

The Makeham term, A, is a nuisance parameter used to adjust the incidence of AD so that the mean age at onset in the model is the same as that in the data. Its only purpose here is to improve the estimation of the other terms, as the survival of a cohort of AD patients is strongly related to their mean age at onset. It does not furnish an estimate of the incidence of AD in the whole population, which was described in Section 4.3.

4.5. Mortality of Lives with Alzheimer's Disease

AD patients have been found to suffer higher mortality than the general population (Barclay et al., 1985(b); Bonaiuto et al., 1995; Bracco et al., 1994;

Burns *et al.*, 1991; van Dijk *et al.*, 1991; Evans *et al.*, 1991; Heyman *et al.*, 1996; Mölsä *et al.*, 1986; Treves *et al.*, 1986). However, there is little agreement on the magnitude of the increase, or its dependence on age at onset, duration since onset, sex, race, level of education, marital status, level of cognitive impairment, familial/non-familial AD and level of behavioural impairment. The main factors we need to consider are:

(a) *The magnitude of the increase in mortality for AD lives.* The mortality of lives with AD has been investigated using different methodologies. For example, Evans *et al.* (1991), estimated the relative risk of death for AD patients as 1.44 (95% confidence interval 1.05-1.96) times that of the unaffected. Others have suggested that AD has only a small impact on mortality: Barclay *et al.* (1985a) claimed that well-tended individuals may have life expectancy close to normal, and Sayetta *et al.* (1986) found that survival did not depend on disease acquisition.

(b) *The effect of age at onset on relative mortality.* The mortality of patients with AD increases with age (Bonaiuto *et al.* 1995; Burns *et al.*, 1991). Most studies into survival times have found no relation between age at entry into the study and relative survival (Barclay *et al.*, 1985b; Bracco *et al.*, 1994; Heyman *et al.*, 1996; Stern *et al.*, 1995; Mölsä *et al.*, 1986), except that Barclay *et al.* (1985b) found that younger lives had shorter relative survival times. Diesfeldt *et al.* (1986), investigating survival from onset of AD, found that AD patients with onset before age 76 had reduced survival times, but not those with later onset. Comparing the two methods of investigation, Walsh *et al.* (1990) found that older age at onset affected survival adversely, whereas older age at entry into the study did not; a possible explanation was that older patients have symptoms for a shorter time before presentation. Although no definitive relationship between age at onset and relative survival emerges, it is clear that:
   (1) survival with AD depends on age; and
   (2) if age at onset affects relative mortality, the relationship is only weak, but possibly stronger at younger ages.
   In terms of the model in Figure 1, this suggests that mortality in state $i3$ (institutionalised from AD) could be modelled by the addition of a Makeham term to the normal force of mortality; the latter is age dependent, and the Makeham term will be less significant at older ages.

(c) *The effect of the duration of AD on relative survival.* Perhaps surprisingly, the duration of AD has not been found to be associated with increased mortality (Barclay *et al.*, 1985a; Bracco *et al.*, 1994; Burns *et al.*, 1991; Diesfeldt *et al.*, 1986; Heyman *et al.*, 1996; Sayetta *et al.*, 1986; Walsh *et al.*, 1990). That is, AD patients with long duration of symptoms do not suffer higher mortality than patients, of the same age, with short duration of symptoms. In terms of the model, this means that the mortality of lives in states 2 and 3 (onset of AD and institutionalised from AD) does not depend on the time spent in these states. This is especially convenient, as it allows us to work in a Markov framework.

(d) *The effect of gender on relative survival with AD.* Many researchers have found that the differences in survival between men and women with AD can be explained by the usual mortality differential between men and women (Beard *et al.*, 1994; Bonaiuto *et al.*, 1995; Bracco *et al.*, 1994; Burns *et al.*, 1991; Heyman *et al.*, 1996; Walsh *et al.*, 1990), though Barclay *et al.* (1985a), did find greater differences. In terms of modelling, allowing for the normal differences in mortality between genders should be sufficient.

Table 2 summarises the literature on survival with AD. Since we cannot distinguish genotypes here, we will just write μ^{34}_{x+t} instead of μ^{i34}_{x+t}. As in the previous section, we can write down the mean age at onset (see equation (6)) and the mean survival time (see equation (7)) in the model of Figure 1:

$$E\left[x + \int_x^\omega I_1(t)dt \mid N_{12}(\omega - x) = 1 \text{ and } I_1(x) = 1\right] =$$

$$x + \frac{\int_x^\omega (t - x)\mu_t^{12} P_{xt}^{11} dt}{\int_x^\omega \mu_t^{12} P_{xt}^{11} dt} \tag{6}$$

$$E\left[\int_x^\omega I_2(t) + I_3(t)dt \mid N_{12}(\omega - x) = 1 \text{ and } I_1(x) = 1\right] =$$

$$\frac{\int_x^\omega \mu_t^{12} P_{xt}^{11} \left\{\int_t^\omega (s - t)(\mu_s^{23} + \mu_s^{24}) P_{ts}^{22} ds + \int_t^\omega \mu_s^{23} P_{ts}^{23} \int_s^\omega (r - s)\mu_r^{34} P_{sr}^{33} dr\, ds\right\} dt}{\int_x^\omega \mu_t^{12} P_{xt}^{11} dt} \tag{7}$$

TABLE 2

SUMMARY STATISTICS ON SURVIVAL TIMES OF AD PATIENTS

Reference	Mean (Median) Age at Onset	Mean (Median) Survival Time	Addition to μ^{34}_{x+t}
Barclay *et al.* (1985a)	(73.3) yrs	(8.1) yrs	0.15829
Bracco *et al.* (1994)	(72.4) yrs	7.3 yrs	0.25259
Diesfeldt *et al.* (1986)	75.6 yrs	7.2 yrs	0.21056
Heyman *et al.* (1996)	(69.2) yrs	(9.7) yrs	0.10993
Jost *et al.* (1995)	75.1 yrs	8.11 yrs	0.13345
Kokmen *et al.* (1988)	80.4 yrs	6.2 yrs	0.26420
Treves *et al.* (1986)	73.9 yrs	(9.3) yrs	0.08135
Average			0.17291

Setting equations (6) and (7) equal to their estimated values in Table 2, and noting that we have estimates of μ_{x+t}^{23} and μ_{x+t}^{24} from the previous section, we obtain 2 equations, which can be solved for at most 2 unknown parameters. The parametric forms we used are as follows:

(a) $\mu_{x+t}^{12} = A + \mu_{x+t}^{AD}$, where μ_{x+t}^{AD} is given by equation (2). This is just the addition of a Makeham term to the force of incidence of AD, shifting the latter to a level that gives the estimated age at onset.

(b) $\mu_{x+t}^{34} = K + {}^{AM80}\mu_{x+t}$. This is a Makeham term as discussed in (d) above.

The estimated values of K for each of the references cited are given in the last column of Table 2. They range from about 0.08 to 0.27, with an average of 0.173. The Makeham term A is, again, only included to improve the estimation of the other terms (see the end of the previous section).

4.6. Summary of the Transition Intensities for the AD Model

For clarity, we summarise the transition intensities estimated here. They all have the form:

$$\mu_{x+t}^{jk} = A + DBe^{C(x+t)} \tag{8}$$

and the calculated values are given in Table 3. Three values are given for μ_{x+t}^{34}, an upper bound, mean value and lower bound to enable a check of how sensitive the results are, in any particular investigation, to this term.

TABLE 3

SUMMARY OF TRANSITION INTENSITIES FOR THE AD MODEL WITH BASELINE MORTALITY 100% (65%) OF AM80 AND AF80

Transition Intensity	A	D	B ($\times 10^{-5}$)		C ($\times 10^{-2}$)	
			Male	Female	Male	Female
μ_{x+t}^{24}	0	0.33502	9.4116	2.5934	8.4554	9.3605
		(0.21776)				
μ_{x+t}^{23}	0.18896	0.00				
μ_{x+t}^{34} (Lower bound)	0.08	1.00 (0.65)	9.4116	2.5934	8.4554	9.3605
μ_{x+t}^{34} (Mean)	0.17291	1.00 (0.65)	9.4116	2.5934	8.4554	9.3605
μ_{x+t}^{34} (Upper bound)	0.27	1.00 (0.65)	9.4116	2.5934	8.4554	9.3605

5. ESTIMATION OF TRANSITION INTENSITIES DEPENDING ON ApoE GENOTYPE

5.1. Population Frequencies of the ApoE Genotypes

Table 4 shows the population frequencies of the ApoE genotypes estimated in several studies. Others are Corder *et al.* (1994, 1995), Gomez-Isla *et al.* (1996), Lehtovirta *et al.* (1996), Liddell *et al.* (1994), Lopez *et al.* (1998), Nalbantoglu *et al.* (1994), Poirier *et al.* (1993), Roses (1995) and Tsai *et al.* (1994).

Some features are clear: the $\varepsilon3/\varepsilon4$ genotype is not uncommon (about 21%) while the $\varepsilon2/\varepsilon4$ and $\varepsilon4/\varepsilon4$ genotypes are quite uncommon (about 3% and 1% respectively). We might expect to find lower proportions of 'dangerous' genotypes at older ages, because these lives suffer a higher rate of AD onset, but the two age-related studies (Bickeböller *et al.* (1997) and Corder *et al.* (1995)) gave conflicting results. However, there is reasonable agreement on the gene frequencies at around ages 60-70, which is what we need for our modelling.

Farrer *et al.* (1997) is a meta-analysis, combining the results of 40 other studies, including 6,264 Caucasian subjects. As it is the largest study, and differentiates by ethnic group, gender and ascertainment methods, and as the ApoE $\varepsilon4$ allele was found with the same frequency in respect of AD diagnosed at autopsy and clinically diagnosed probable AD, we use its estimated gene frequencies, namely: $\varepsilon2/\varepsilon2$ 0.008; $\varepsilon2/\varepsilon3$ 0.127; $\varepsilon2/\varepsilon4$ 0.026; $\varepsilon3/\varepsilon3$ 0.609; $\varepsilon3/\varepsilon4$ 0.213; $\varepsilon4/\varepsilon4$ 0.018. The sharp-eyed will notice that these sum to 1.001, because of roundings used in Farrer *et al.* (1997), but we have left this small discrepancy unadjusted.

5.2. Genetic Risk of Alzheimer's Disease

When we have a heterogeneous population, it is often convenient to think of a given intensity in each sub-population as a multiple (not necessarily constant) of a 'baseline' intensity, either in one of the sub-populations or in an aggregated 'average' population. Similarly, if p_1 and p_2 are the probabilities of an event in populations 1 and 2 respectively, the relative risk in population 2 (with respect to population 1) is p_2/p_1. A related quantity is the odds ratio: the odds in populations 1 and 2, respectively, are $p_1/(1 - p_1)$ and $p_2/(1 - p_2)$, and the odds ratio is:

$$\frac{p_2(1 - p_1)}{p_1(1 - p_2)} . \tag{9}$$

When intensities are small, the odds ratio is a good approximation to the relative risk. In many studies, the published results are either relative risks or odds ratios.

TABLE 4

ESTIMATED POPULATION FREQUENCY OF APOE GENOTYPES

Reference	Country / Ethnicity	No. of Lives	Sex	Age Group	Allele Frequency			Genotype Frequency					
					$\varepsilon2$	$\varepsilon3$	$\varepsilon4$	$\varepsilon2/\varepsilon2$	$\varepsilon2/\varepsilon3$	$\varepsilon2/\varepsilon4$	$\varepsilon3/\varepsilon3$	$\varepsilon3/\varepsilon4$	$\varepsilon4/\varepsilon4$
Farrer *et al.* (1997)	Caucasian	6,262	M & F	All	0.084	0.779	0.137	0.01	0.13	0.03	0.61	0.21	0.02
	Afr-Amer	240	M & F	All	0.083	0.727	0.190	0.01	0.13	0.02	0.50	0.32	0.02
	Hispanic	267	M & F	All	0.067	0.823	0.110	0.00	0.12	0.01	0.67	0.18	0.02
	Japanese	1,977	M & F	All	0.042	0.869	0.089	0.00	0.07	0.01	0.76	0.16	0.01
Bickeböller *et al.* (1997)	France	1,030	M & F	All	0.085	0.770	0.145	0.01	0.12	0.03	0.59	0.24	0.01
		316	M	All	0.070	0.815	0.105	0.01	0.10	0.02	0.67	0.19	0.00
		40	M	< 60	0.050	0.800	0.150	0.00	0.08	0.02	0.62	0.28	0.00
		93	M	60-69	0.070	0.845	0.085	0.01	0.12	0.00	0.70	0.17	0.00
		80	M	70-79	0.060	0.840	0.090	0.00	0.10	0.02	0.71	0.16	0.00
		103	M	≥ 80	0.081	0.795	0.125	0.01	0.11	0.03	0.64	0.20	0.01
		714	F	All	0.090	0.750	0.170	0.01	0.13	0.03	0.55	0.27	0.02
		47	F	< 60	0.075	0.735	0.190	0.00	0.15	0.00	0.51	0.30	0.04
		75	F	60-69	0.075	0.730	0.195	0.01	0.11	0.02	0.52	0.31	0.03
		143	F	70-79	0.065	0.765	0.160	0.00	0.12	0.01	0.56	0.29	0.01
		449	F	≥ 80	0.095	0.750	0.165	0.01	0.13	0.04	0.56	0.25	0.02
van Duijn *et al.* (1995)	Netherlands	532	M & F	< 65	0.103	0.731	0.165	0.01	0.17	0.02	0.51	0.27	0.02
		228	M	< 65	0.105	0.705	0.190	0.01	0.16	0.03	0.48	0.29	0.03
		304	F	< 65	0.100	0.745	0.155	0.01	0.17	0.01	0.53	0.26	0.02
Evans *et al.* (1997)	E. Boston	490	M & F	≥ 65	0.062	0.854	0.084	0.01	0.09	0.01	0.74	0.14	0.01
Jarvik *et al.* (1996)	not given	310	M & F	48-98	0.098	0.750	0.132	0.01	0.12	0.05	0.59	0.20	0.01
		117	M	48-98	0.068	0.791	0.124	0.02	0.09	0.02	0.64	0.21	0.01
		193	F	48-98	0.117	0.725	0.137	0.01	0.15	0.07	0.55	0.20	0.01
Lambert *et al.* (1998)	not given	308	M & F	not given	0.081	0.805	0.114	0.00	0.13	0.02	0.65	0.18	0.01
Levy-Lahad *et al.* (1996)	not given	304	M & F	not given	0.100	0.765	0.135	0.01	0.13	0.05	0.60	0.21	0.01
Lucotte *et al.* (1997)	France	248	M & F	≥ 65	0.069	0.804	0.127	0.02	0.09	0.01	0.65	0.22	0.01

Few studies report prospectively the incidence of AD by genotype. Two that do are Evans *et al.*, (1997) and Slooter *et al.*, (1998). Both have quite small study populations, and neither provides age specific estimates of AD risk, so they are not appropriate for our purposes.

Table 5 gives the Odds Ratios (ORs) of AD and 95% confidence intervals from a number of genetic studies. The 'reference' populations (also shown in the table) were either the $\varepsilon3/\varepsilon3$ genotype or the three non-$\varepsilon4$ genotypes combined. The estimated ORs vary considerably across studies. For example, estimates of the OR for the $\varepsilon3/\varepsilon4$ genotype (relative to the $\varepsilon3/\varepsilon3$ genotype) range from 1.8% to 3.7%, and for the $\varepsilon4/\varepsilon4$ genotype, from 6.2% to 30.7%. Some of the variation may be explained by the differences between the studies themselves. In particular, differences may arise from: the method of ascertainment of patients, the countries of study, the method of diagnosis of AD, the age structure of the samples, the reference/risk genotypes, and whether they are cross-sectional or prospective studies. We make the following observations:
(a) The study by Lopez *et al.* (1998) suggests that the risk of AD associated with the ApoE $\varepsilon4$ allele may be different in different countries.
(b) In support of the above, Mayeux *et al.* (1998) found that the association between ApoE and AD may depend on ethnic group and, in particular, may not be present in black populations.
(c) Despite the differences between studies: the presence of one or two $\varepsilon4$ alleles is consistently reported to be significantly associated with AD; and homozygotes are generally reported to have higher risk of onset of AD than heterozygotes.
(d) The weakest associations between ApoE and AD were reported in the two prospective studies, those by Evans *et al.* (1997) and Slooter *et al.* (1998). This is as expected for the reasons given in Section 4.1.

For our purposes, the genetic risk of AD at different ages is important. Few studies have considered this; the odds ratios from two that have are given in Table 6. Bickeböller *et al.* (1997) is based on hospital admissions, and Corder *et al.* (1994) on autopsy cases; both use $\varepsilon3/\varepsilon3$ as the reference population. Although some ORs are missing, because of small sample sizes, the trends are fairly clear:
(a) The odds of AD among the higher risk genotypes ($\varepsilon3/\varepsilon4$ and $\varepsilon4/\varepsilon4$) fall with age. This may be expected as higher risk genotypes will succumb to AD more rapidly, reducing the proportion of such genotypes within the population.
(b) Conversely, the protection conferred by the lower risk genotype ($\varepsilon2/\varepsilon3$) seems to weaken with age, possibly as this genotype becomes more common in the remaining population.

TABLE 5

AGGREGATED ODDS RATIOS OF AD FOR THE APOE ε4 ALLELE

Reference	Ascertainment Scheme (1)	Reference Genotype (2)	Genotype (2)	Odds Ratio	
				Mean	95% CI
Evans et al. (1997)	P	ε3/ε3	ε3/ε4 & ε4/ε4	2.27	1.1-4.9
Frisoni et al. (1995)	C	-/-	ε4/-	6.6	2.2.-19.5
			ε4/ε4	17.9	4.5-70.5
Jarvik et al. (1996)	C	ε3/ε3	ε2/ε3	0.4	0.2-0.96
			ε2/ε4	1.4	0.6-3
			ε3/ε4	3.1	2-4.7
			ε4/ε4	30.7	7-131
Kuusisto et al. (1994)	P	-/-	ε4/-	2.7	1.4-5.2
			ε4/ε4	9.1	3.5-23.4
Lambert et al. (1998)	C	-/-	ε4/- & ε4/ε4	4.66	3.14-6.93
Lehtovirta et al. (1995)	C	-/-	ε4/-	5.1	2.4-11.1
			ε4/ε4	21.4	2.8-166.3
Liddell et al. (1994)	C	-/-	ε4/-	2.2	1.1-4.7
			ε4/ε4	10.7	2.3-48.8
Lopez et al. (1998)	C	-/-	ε4/- & ε4/ε4	2.34 (3)	1.03-5.55
			ε4/- & ε4/ε4	3.64 (4)	1.78-7.69
Mayeux et al. (1993)	P	-/-	ε4/-	4.2 (5)	1.8-9.5
			ε4/ε4	17.9 (5)	4.6-69.8
			ε4/- & ε4/ε4	15.3 (6)	3.0-78.1
			ε4/- & ε4/ε4	0.7 (7)	0.1-6.4
			ε4/- & ε4/ε4	4.5 (8)	0.7-27.7
Myers et al. (1996)	P	ε3/ε3	ε3/ε4	3.7	1.9-7.5
			ε4/ε4	30.1	10.7-84.4
Nalbantoglu et al. (1994)	A	-/-	ε4/- & ε4/ε4	15.5	6.2-38.5
Slooter et al. (1998)	P	ε3/ε3	ε2/ε3	0.4	0.1-1.0
			ε2/ε4	1.3	0.2-8.5
			ε3/ε4	1.8	1.0-3.1
			ε4/ε4	6.2	1.4-28.2
Tsai et al. (1994)	C	-/-	ε4/- & ε4/ε4	4.6	1.9-12.3
			ε4/-	3.6	1.5-9.8

(1) Ascertainment Scheme: C indicates clinic/hospital; P, population/community; and A, autopsy/brainbank.

(2) Dash (-) represents ε2 or ε3 alleles.

(3) Study population – Gerona, Spain.

(4) Study population – Pittsburgh, USA.

(5) Mixture of White, Black and Hispanic ethnic groups.

(6) White ethnic group only.

(7) Black ethnic group only.

(8) Hispanic ethnic group only.

TABLE 6

ODDS RATIOS OF AD BY GENOTYPE AND AGE

Bickeböller et al. (1997)				Corder et al. (1994)			
Age Group	Genotype	Odds Ratio		Age Group	Genotype	Odds Ratio	
		Mean	95% CI			Mean	95% CI
60-69	ε2/ε3	0.3	0.0-2.3	60-66	ε3/ε3	0.1	−
	ε2/ε4	−	−		ε2/ε4	1.2	−
	ε3/ε4	3.1	1.4-6.9		ε3/ε4	11.1	−
	ε4/ε4	29.1	3.6-239.5		ε4/ε4	123.8	−
70-79	ε2/ε3	0.4	0.1-2.3	67-74	ε2/ε3	0.3	−
	ε2/ε4	−	−		ε2/ε4	1.1	−
	ε3/ε4	3.2	1.5-6.6		ε3/ε4	4.6	−
	ε4/ε4	−	−		ε4/ε4	20.8	−
80+	ε2/ε3	0.3	0.0-2.6	75-92	ε2/ε3	0.5	−
	ε2/ε4	−	−		ε2/ε4	1.6	−
	ε3/ε4	1.3	0.5-3.4		ε3/ε4	3.2	−
	ε4/ε4	−	−		ε4/ε4	10.0	−
60+	ε2/ε3	0.4	0.1-0.9	60+	ε2/ε3	0.3	−
	ε2/ε4	1.6	0.5-5.5		ε2/ε4	1.1	−
	ε3/ε4	2.2	1.5-3.5		ε3/ε4	4.4	−
	ε4/ε4	11.2	4.0-31.6		ε4/ε4	19.3	−

These trends are supported by the meta-analysis by Farrer et al. (1997), and as it is from this study that we take our estimates of the ApoE genotype risks, we cite some relevant details:

(a) The aggregate relative odds from Farrer et al. (1997) (relative to the ε3/ε3 genotype) are shown in Figure 3.

(b) The genotype risks of AD were not significantly different in respect of Caucasian males and females, except in the case of the ε3/ε4 genotype, for which women had a significantly higher risk of AD. The relative odds of AD by ApoE for Caucasian men and women are shown in Figures 4 and 5. (The authors kindly provided us with the numerical values of the odds ratios; confidence intervals were not available.)

(c) The genotypes ε2/ε2 and ε2/ε3 were combined as there were very few ε2/ε2 genotypes, and the risks associated with the two genotypes appeared to be similar.

(d) Note that the risks associated with the ApoE ε4 allele were considerably higher than those found in the two population-based studies by Evans et al., (1997) and Slooter et al., (1998).

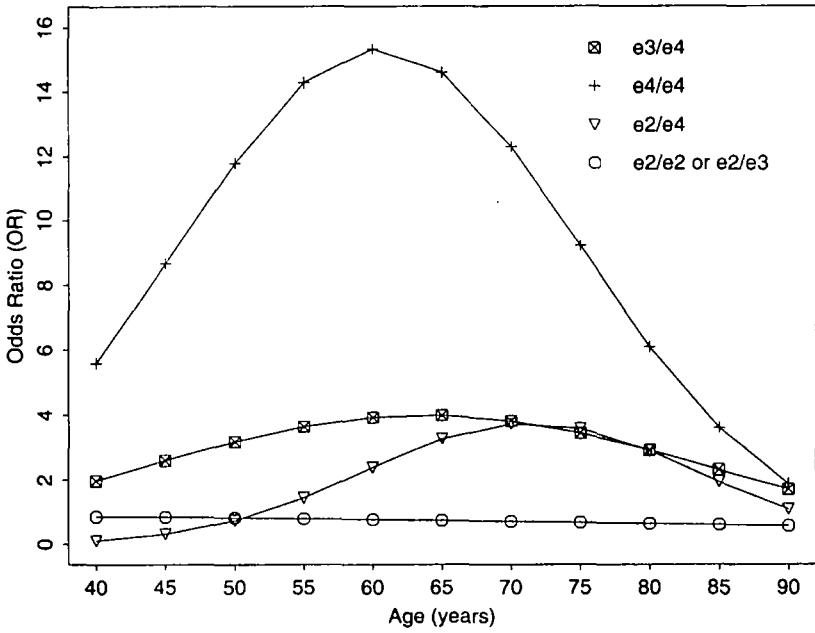FIGURE 3: Odds ratios (ORs) of AD relative to $\varepsilon3/\varepsilon3$ genotype for males and females combined. Source: Farrer *et al.* (1997).
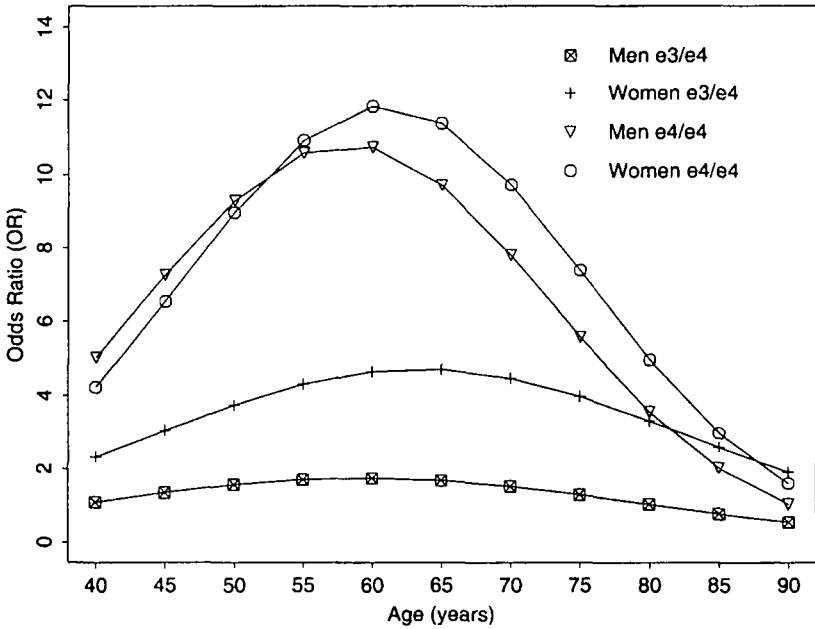
FIGURE 4: Odds ratios (ORs) of AD relative to $\varepsilon3/\varepsilon3$ genotype for $\varepsilon3/\varepsilon4$ and $\varepsilon4/\varepsilon4$ genotypes. Source: Farrer *et al.* (1997).
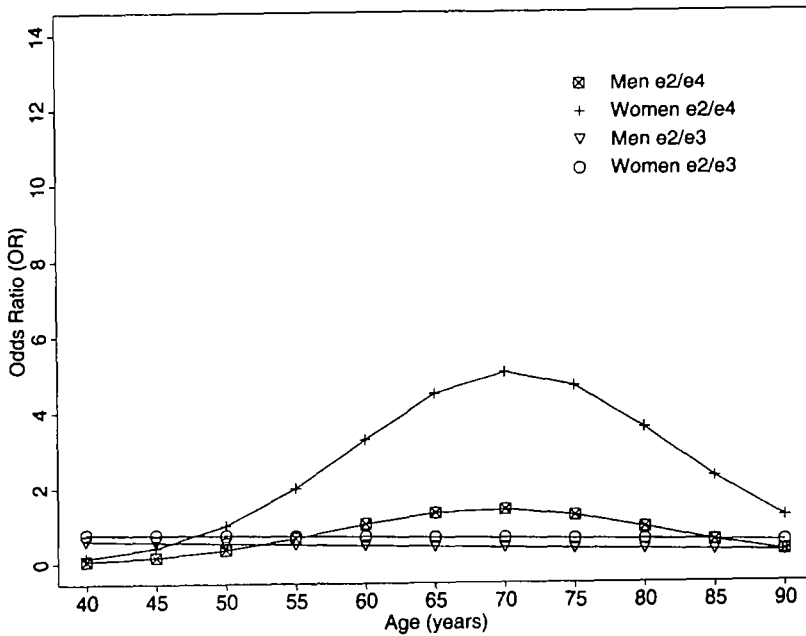
FIGURE 5: Odds ratios (ORs) of AD relative to $\varepsilon3/\varepsilon3$ genotype for $\varepsilon2/\varepsilon2$ or $\varepsilon2/\varepsilon3$ and $\varepsilon2/\varepsilon4$ genotypes.
Source: Farrer *et al.* (1997).

For use in our model, these odds ratios have to be converted into relative risks. More precisely, we have to find a plausible set of age- and genotype-dependent transition intensities that are consistent with the odds ratios and together are consistent with the aggregate incidence of AD. There is no unique solution to this problem. The method we used was to model the incidence of AD for the *i*th genotype as:

$$\mu_{x+t}^{i12} = r_1 f_{x+t}^{i} \mu_{x+t}^{AD} \tag{10}$$

where:
(a) μ_{x+t}^{AD} is the aggregate incidence rate of AD (from Section 4.3);
(b) f_{x+t}^{i} is a parametric function representing the risk relative to the aggregate incidence rate, where we take $f_{x+t}^{i} = 1$ in the case of the $\varepsilon3/\varepsilon3$ genotype; and
(c) r_1 is a constant chosen so that the aggregate incidence of AD based on the modelled intensities is consistent with the aggregate incidence μ_{x+t}^{AD}.

We did this for males and females separately and combined. We confined our attention to ages 60 and over, in order to get a better fit in the age range of interest in applications. The form of the ORs, either rising to a peak and then falling, or gently declining, suggested a similar pattern of relative risks,

and we found the following family of functions satisfactory (note that constant relative risks, or proportional hazards, result in odds ratios with exponential growth):

$$f^i_{x+t} = Ee^{-F((x+t)-k_1)^2 - G((x+t)-k_2)} + H. \tag{11}$$

Actuaries will recognize this as a GM(1,3) function, familiar in the graduation of life tables (Forfar, McCutcheon & Wilkie, 1988), although as described below either F or G is set to zero. We found this flexible enough to give a good approximation to the ORs, and also suitable for extrapolating beyond age 90. The fitting procedure was as follows:
(a) by considering the form of the OR, we set either $F = 0$ (giving an exponential function) or $G = 0$ (giving a bell-curve function), and set H equal to 0 or 1;
(b) the best value of k_1 or k_2 was found, to the nearest integer, by inspection;
(c) the resulting ORs were calculated from the model in Figure 1 using the intensities from previous sections; and
(d) the remaining coefficients (either E and F, or E and G) were fitted by least squares.

For the calculations in (c) above we used the following parameters:
(a) $\mu^{14}_{x+t} = 0.65 \times {}^{AM80}\mu_{x+t}$ for males and $\mu^{14}_{x+t} = 0.65 \times {}^{AF80}\mu_{x+t}$ for females and in aggregate, where ${}^{AM80}\mu_{x+t}$ and ${}^{AF80}\mu_{x+t}$ are given by equation 1.

(b) $\mu^{23}_{x+t} = 0.189$, calculated in Section 4.4.

(c) $\mu^{24}_{x+t} = 0.335 \times \mu^{14}_{x+t}$ calculated in Section 4.4.

(d) $\mu^{34}_{x+t} = 0.173 + \mu^{14}_{x+t}$, the mean value calculated in Section 4.5.

The fitted parameters are given in Table 7. Figure 6 shows that the modelled ORs closely reproduce the estimates from Farrer *et al.* (1997) (see Figure 3; only the aggregate ORs are shown, and the modelled ORs start at age 61 because we start with unaffected lives at age 60).

To determine the parameter r_1, we calculated the aggregate incidence of AD in the whole model, and fitted this to μ^{AD}_{x+t} by least squares. If ${}^{i}P^{11}_{xt}$ is the probability that a life with genotype i, healthy at age x, is unaffected by AD at age $x + t$, and if p^i_x is the population frequency of the ith genotype at age x then the aggregate incidence of AD is:

$$\text{Aggregate incidence of AD at age } (x + t) = r_1 \left\{ \sum_{i=1}^{5} p^i_x \, {}^{i}P^{11}_{xt} f^i_{x+t} \right\} \mu^{AD}_{x+t} \tag{12}$$

TABLE 7

PARAMETERS FOR THE RELATIVE RISK OF AD FOR MALES, FEMALES AND IN AGGREGATE, BY GENOTYPE

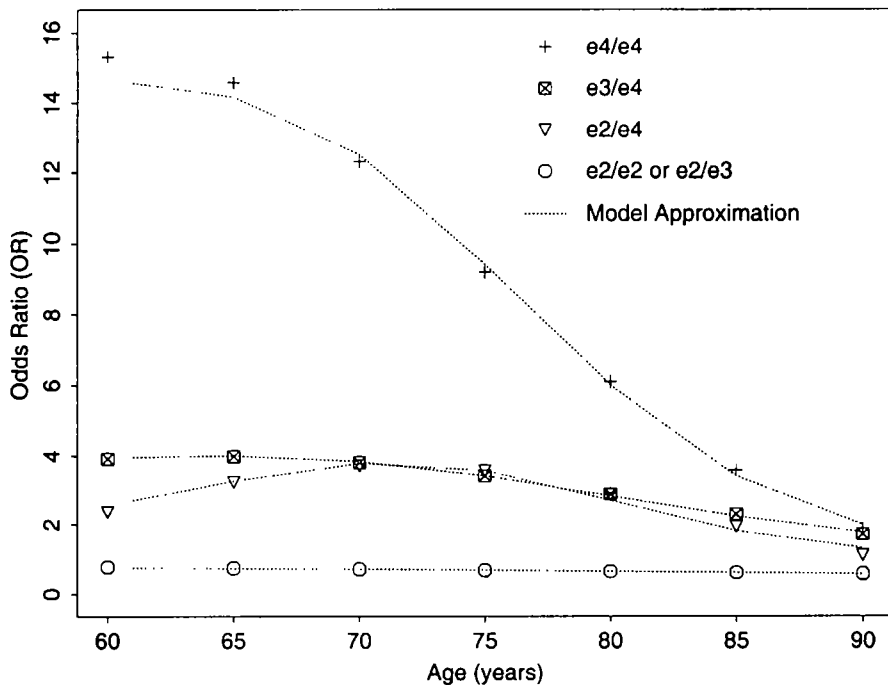| Gender	Genotype	Parameter Values						
		E	F	G	H	k_1	k_2	r_1
Both	$\varepsilon4/\varepsilon4$	13.5	0.00529	0	1	60	–	0.93
	$\varepsilon3/\varepsilon4$	2.98	0.00312	0	1	62	–	
	$\varepsilon2/\varepsilon4$	2.87	0.00938	0	1	68	–	
	$\varepsilon2/\varepsilon2$ & $\varepsilon2/\varepsilon3$	0.754	0	0.00859	0	–	60	
Female	$\varepsilon4/\varepsilon4$	10.4	0.00504	0	1	60	–	0.88
	$\varepsilon3/\varepsilon4$	3.68	0.00319	0	1	62	–	
	$\varepsilon2/\varepsilon4$	4.21	0.01020	0	1	68	–	
	$\varepsilon2/\varepsilon2$ & $\varepsilon2/\varepsilon3$	0.675	0	0.00692	0	–	60	
Male	$\varepsilon4/\varepsilon4$	8.94	0.00656	0	1	60	–	1.27
	$\varepsilon3/\varepsilon4$	1.92	0.00103	0	0	51	–	
	$\varepsilon2/\varepsilon4$	1.42	0.00506	0	0	67	–	
	$\varepsilon2/\varepsilon2$ & $\varepsilon2/\varepsilon3$	0.434	0	0.01600	0	–	60	

FIGURE 6: Odds ratios (ORs) of AD relative to $\varepsilon3/\varepsilon3$ genotype from Farrer et al. (1997), compared with ORs computed using modelled relative risk functions.

We took $x = 60$, and for the p_x^i we used the gene frequencies of the Caucasian control populations in Farrer *et al.* (1997), which were given in Section 5.1. The incidence of AD, μ_{x+t}^{AD}, was taken as that estimated in equation 2 and the occupancy probabilities, $^iP_{xt}^{11}$, were calculated by solving Kolmogorov's forward equations numerically using a Runge-Kutta algorithm with step size 0.0005 years (Conte & De Boor, 1972).

The values of r_1 are given in Table 7. The adjustment to the overall level only had a marginal effect on the modelled ORs for the individual genotypes.

The relative risk functions for males and females are given in Figures 7 and 8. For females, the $\varepsilon4/\varepsilon4$, $\varepsilon3/\varepsilon4$ and $\varepsilon2/\varepsilon4$ genotypes are unambiguously high-risk; the relative risks exceed 1.0 at all ages. For males, only the $\varepsilon4/\varepsilon4$ genotype confers higher risks at all ages. The $\varepsilon2$ allele appears to be protective, so the $\varepsilon2/\varepsilon2$ and $\varepsilon2/\varepsilon3$ genotypes are low-risk, while the $\varepsilon3/\varepsilon4$ and $\varepsilon2/\varepsilon4$ genotypes are initially at higher risk but are at lower risk from about age 75. The protection apparently given by the $\varepsilon2$ allele in males means that the $\varepsilon3/\varepsilon3$ genotype confers slightly higher than average risk; this is why, in Table 7, $r_1 > 1$ for males. It must be remembered that data in respect of males are relatively sparse, and data in respect of very old males even sparser, so these effects should be treated with caution; we can have more confidence in the relative risks in respect of females.
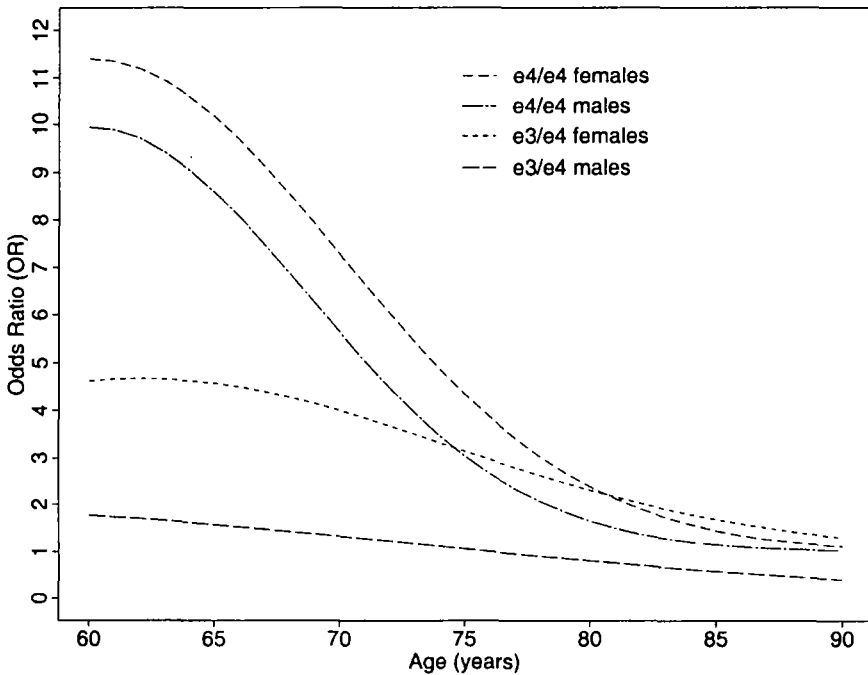

FIGURE 7: Modelled risk of AD, relative to the $\varepsilon3/\varepsilon3$ genotype, for $\varepsilon4/\varepsilon4$ and $\varepsilon3/\varepsilon4$ genotypes. Based on odds ratios from Farrer *et al.* (1997).
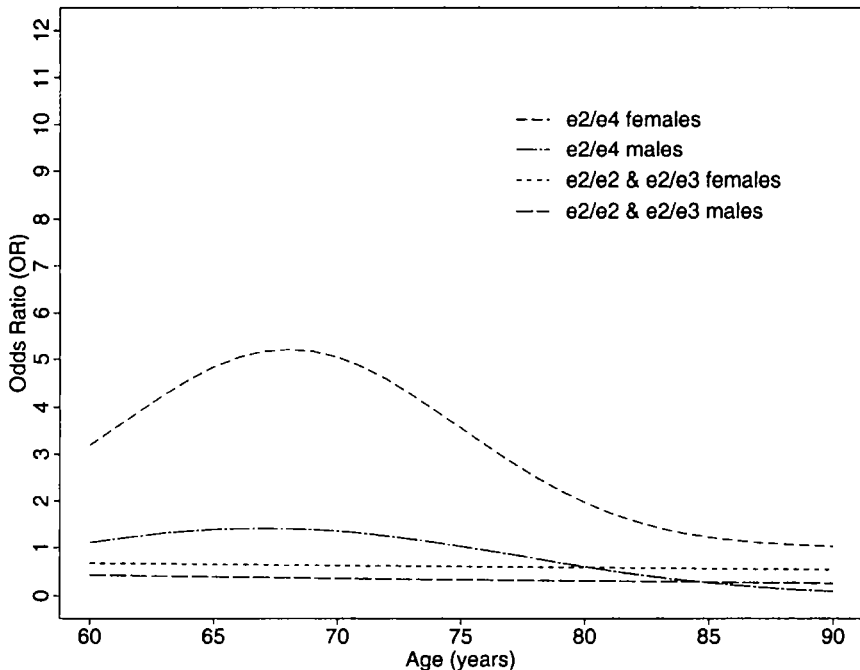
FIGURE 8: Modelled risk of AD, relative to the $\varepsilon3/\varepsilon3$ genotype, for $\varepsilon2/\varepsilon4$ and $\varepsilon2/\varepsilon2$ & $\varepsilon2/\varepsilon3$ genotypes. Based on odds ratios from Farrer *et al.* (1997).

These risk estimates probably overstate the true population risks, perhaps quite substantially, as they are from clinic- and autopsy-based studies, which investigate precisely the subjects that are affected or already known to be at risk. To allow for this possibility we will also consider models assuming that the true relative risks are a proportion $m < 1$ of those estimated above. We do this by adjusting equation (10) so that for genotype *i*:

$$\mu_{x+t}^{i12} = r_m\{(f_{x+t}^i - 1)m + 1\}\mu_{x+t}^{AD} \tag{13}$$

where f_{x+t}^i is as above, and r_m is chosen as above so that the aggregate incidence of AD in the model is consistent with μ_{x+t}^{AD}. Values of $r_{0.5}$ and $r_{0.25}$ are shown in Table 8.

TABLE 8

r_m FOR $m = 1$, 0.5 AND 0.25

Gender	r_1	$r_{0.5}$	$r_{0.25}$
Both	0.93	0.96	0.97
Female	0.88	0.94	0.97
Male	1.27	1.11	1.05

Figure 9 shows that the aggregate incidence of AD in the genetic model for both sexes combined is quite close to μ^{AD}_{x+t}. It also shows that increasing the level of relative risk tends to overstate the incidence of AD at younger ages, and to understate it at older ages; the reason is that higher relative risks deplete the high-risk groups more quickly, leaving a relatively healthier population at older ages.
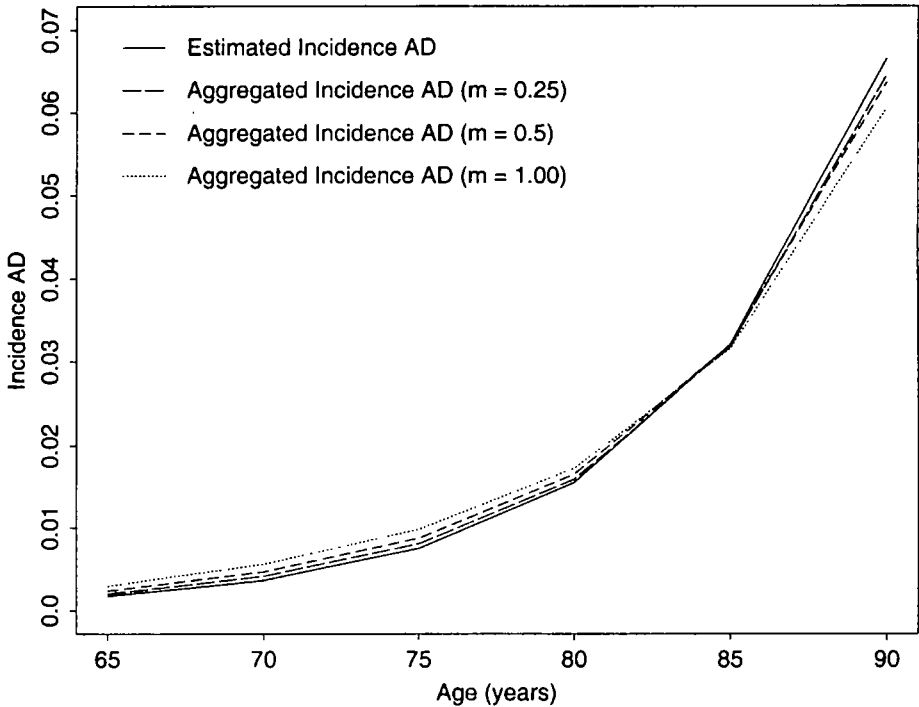

FIGURE 9: Comparison of estimated population incidence of AD μ^{AD}_{x+t} with the aggregated incidence of AD for different levels of relative risk, males and females combined.

Decreasing the level of relative risks for high-risk genotypes means increasing the relative risks for low-risk genotypes. Using a lower value of r_m will diminish any effects of the (possibly anomalous) feature, noted above, that the $\varepsilon3/\varepsilon4$ genotype is low-risk for males.

5.3. Comment on Model Selection

We chose a simple model for the relative risks (equations (10) and (11)). We did consider alternatives, in particular cubic polynomials and Gamma functions, but these gave poorer fits, and were less suitable for extrapolation (cubics to older ages and Gamma functions to younger ages). Also, it is easily seen that if an OR is specified as a function of time, and the transition

intensity in the reference population is given, the transition intensity in the second population is determined (as the solution to an ODE); Figure 6, therefore, gives good support for our choice of model. Further refinement seemed somewhat spurious, given the data we were using, and in view of the major sensitivity analysis needed in respect of the dominant parameter m.

6. RESULTS

6.1. Occupancy Probabilities

Figures 10 and 11 show occupancy probabilities up to age 90 for females healthy at age 60, with high $(m = 1)$ and low $(m = 0.25)$ relative risks, respectively. Each shows:

(a) Occupancy probabilities in respect of each genotype (with $\varepsilon2/\varepsilon2$ and $\varepsilon2/\varepsilon3$ combined).

(b) Occupancy probabilities calculated by aggregating all the genotypes in the model. In the notation of equation (12) the probability of being in state j $(j = 1, 2, 3, 4)$ at age $60 + t$ is $\sum_{i=1}^{i=5} p_{60}^i \, {}^i P_{60,t}^{1j}$, where the sum is over all genotypes. These are labelled 'Aggregated Genotypes'.

(c) Occupancy probabilities based on the aggregate incidence of AD, μ_{x+t}^{AD}. These are labelled 'Aggregate Model'.

With high relative risks $(m = 1)$, the effect of the $\varepsilon4/\varepsilon4$ allele is clear; AD cases rise to a peak in the early 70s, by which time over 10% of the original cohort are in one of the AD states, and then fall away. A similar but smaller effect can be seen for the $\varepsilon3/\varepsilon4$ genotype. With low relative risks $(m = 0.25)$ these features are all diminished; in particular the peaks in the early 70s disappear.

We omit the corresponding figures for males; the differences are as we would expect, given the modelled relative risks.

For males and females with low relative risks (Figure 11) the aggregated results from the genetic model are very close to the results from the aggregate (population) model. For females with high relative risks, the rate of onset of AD seems to be too low at younger ages and too high at older ages.

6.2. Prevalence Rates

Also of interest are prevalence rates, namely the proportion of those alive at every age who are in each of the three live states. Figures 12 and 13 show these, for females, including, for convenience, the two AD states combined. We omit the corresponding figures for males.

The most striking feature is the prevalence of AD in respect of the $\varepsilon4/\varepsilon4$ genotype under high relative risks (Figure 12); it increases almost linearly. Again, for males and females the aggregated results from the genetic model are quite close to those from the aggregate model. Moreover, they fall within the range of prevalence rates actually observed. Breteler *et al.* (1992)

cite the following rates: 47.2% at ages 85 and over (Evans *et al.*, (1989)); 31.7% at ages 85 and over (Pfeffer *et al.*, 1987); and 28.0% at ages 90 and over (O'Connor *et al.*, 1989); some other studies gave lower figures.

6.3. Gene Frequencies at Higher Ages

We have assumed that the gene frequencies given by Farrer *et al.* (1997) are appropriate for age 60. They will change with age, as higher-risk genotypes die more quickly. We must estimate these if we wish to consider entrants (to a study, or into insurance) at ages over 60. Table 9 shows estimates of the gene frequencies in respect of lives unaffected by AD at ages 65, 70 and 75. Using the notation of equation (12), these are given by:

$$p_{60+t}^i = \frac{p_{60}^i \, {}^iP_{60,t}^{11}}{\sum_{j=1}^{j=5} p_{60}^j \, {}^jP_{60,t}^{11}} \; . \tag{14}$$

These are not the gene frequencies in respect of the whole population; lives alive but who have AD are omitted (as is appropriate for insurance applications). Nor are they the gene frequencies in respect of the healthy population; lives with disabilities other than AD are included.

Gene frequencies in the whole population at older ages can also be estimated, as:

$$\frac{p_{60}^i \left({}^iP_{60,t}^{11} + {}^iP_{60,t}^{12} + {}^iP_{60,t}^{13} \right)}{\sum_{j=1}^{j=5} p_{60}^j \left({}^jP_{60,t}^{11} + {}^jP_{60,t}^{12} + {}^jP_{60,t}^{13} \right)} \tag{15}$$

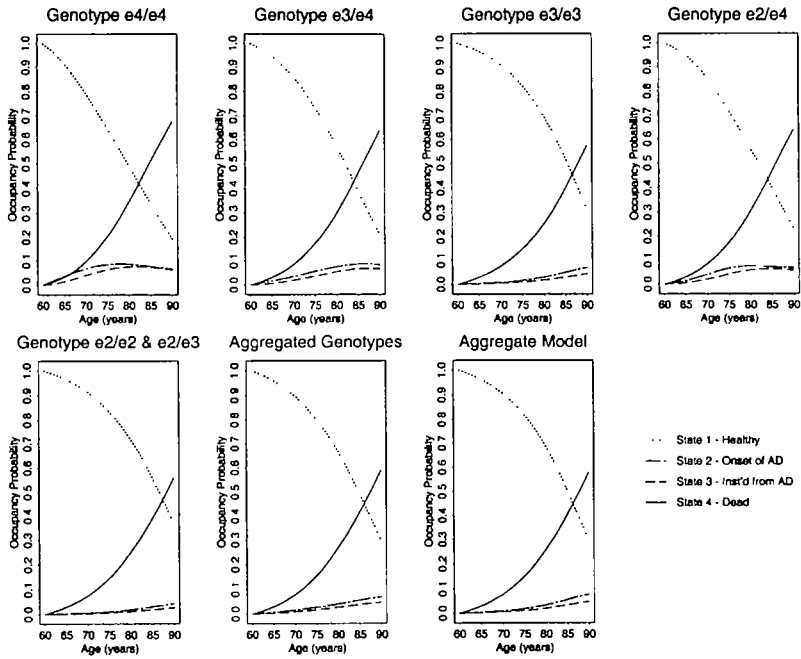but these are not so relevant for insurance applications.

FIGURE 10: Occupancy probabilities for females, healthy at age 60, high relative risks ($m = 1$).
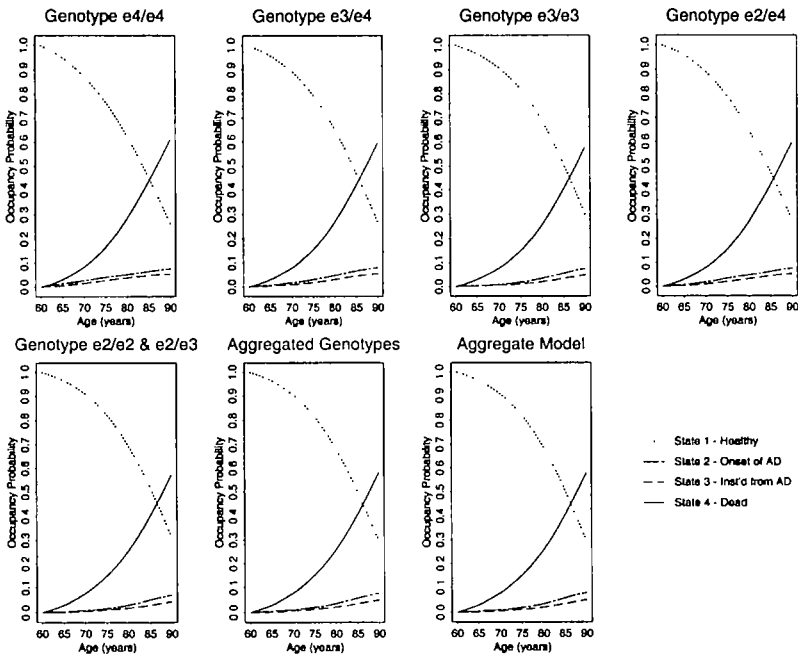
FIGURE 11: Occupancy probabilities for females, healthy at age 60, low relative risks ($m = 0.25$).
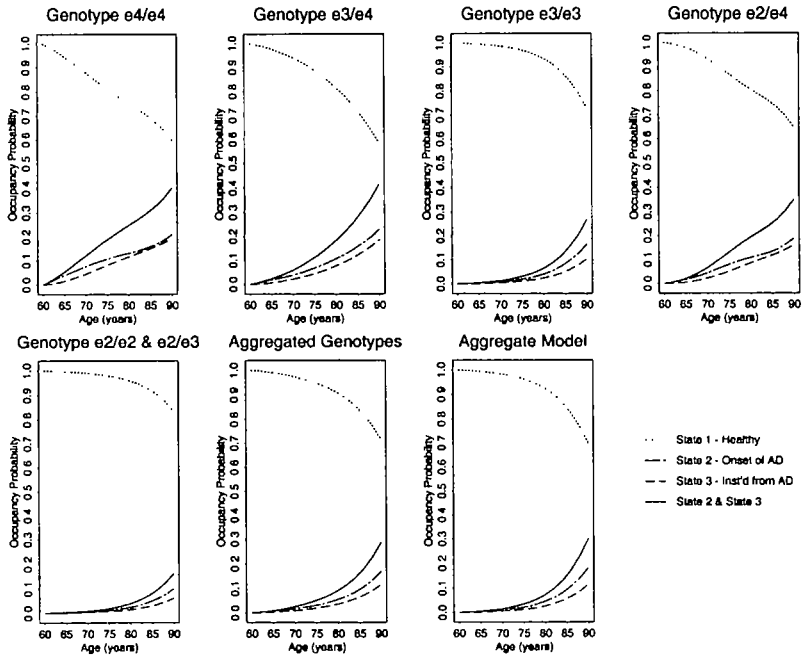
FIGURE 12: Prevalence rate of Alzheimer's disease for females healthy at age 60, high relative risks ($m = 1$).
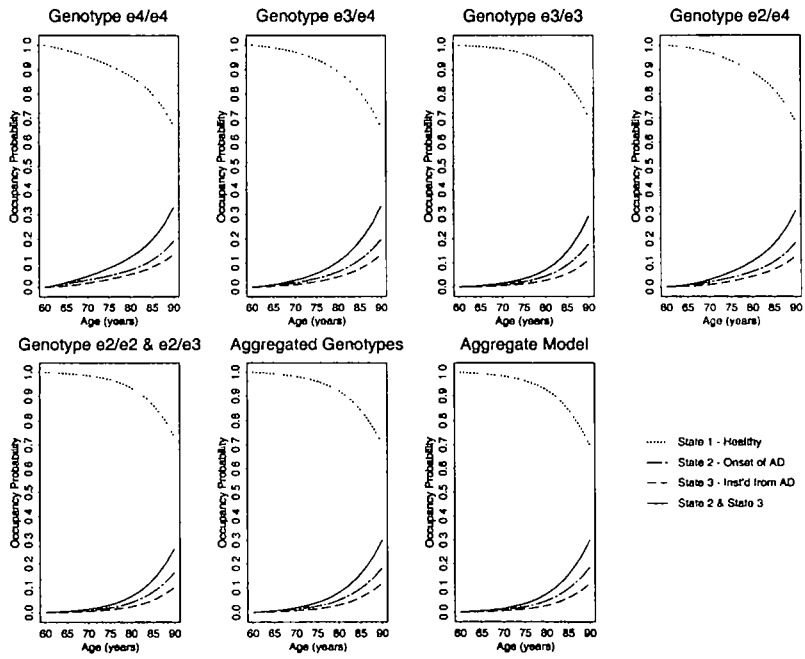

FIGURE 13: Prevalence rate of Alzheimer's disease for females healthy at age 60, low relative risks ($m = 0.25$).

ANGUS MACDONALD AND DELME PRITCHARD

TABLE 9

FREQUENCIES OF APOE GENOTYPES AMONG LIVES FREE OF ALZHEIMER'S DISEASE AT AGES 65, 70 AND 75,
ESTIMATED BY SOLVING THE KOLMOGOROV EQUATIONS FORWARD FROM AGE 60

Gender	Proportion of relative risk, m	Age	Gene Frequencies in AD-free population				
			$\varepsilon4/\varepsilon4$	$\varepsilon3/\varepsilon4$	$\varepsilon3/\varepsilon3$	$\varepsilon2/\varepsilon4$	$\varepsilon2/\varepsilon2$ & $\varepsilon2/\varepsilon3$
M & F	1.00	65	0.0168	0.2103	0.6114	0.0258	0.1357
		70	0.0151	0.2055	0.6168	0.0251	0.1374
		75	0.0133	0.1978	0.6246	0.0241	0.1403
	0.50	65	0.0174	0.2115	0.6100	0.0259	0.1353
		70	0.0164	0.2090	0.6128	0.0255	0.1362
		75	0.0154	0.2050	0.6170	0.0250	0.1377
	0.25	65	0.0177	0.2121	0.6092	0.0259	0.1351
		70	0.0172	0.2109	0.6106	0.0258	0.1355
		75	0.0166	0.2088	0.6128	0.0255	0.1363
F	1.00	65	0.0171	0.2097	0.6116	0.0257	0.1358
		70	0.0159	0.2041	0.6176	0.0247	0.1377
		75	0.0145	0.1951	0.6263	0.0232	0.1409
	0.50	65	0.0175	0.2112	0.6101	0.0258	0.1354
		70	0.0168	0.2081	0.6133	0.0253	0.1364
		75	0.0160	0.2033	0.6181	0.0245	0.1381
	0.25	65	0.0177	0.2119	0.6093	0.0259	0.1351
		70	0.0174	0.2104	0.6110	0.0256	0.1357
		75	0.0169	0.2078	0.6135	0.0252	0.1365
M	1.00	65	0.0169	0.2120	0.6094	0.0260	0.1357
		70	0.0153	0.2110	0.6105	0.0258	0.1373
		75	0.0138	0.2097	0.6105	0.0257	0.1403
	0.50	65	0.0175	0.2125	0.6088	0.0260	0.1352
		70	0.0168	0.2120	0.6094	0.0259	0.1359
		75	0.0160	0.2115	0.6094	0.0258	0.1372
	0.25	65	0.0177	0.2126	0.6086	0.0260	0.1350
		70	0.0174	0.2124	0.6089	0.0259	0.1354
		75	0.0170	0.2122	0.6089	0.0259	0.1360

7. Conclusions

7.1. The Model

We have specified and calibrated a simple continuous-time Markov model of AD allowing for the variability of the ApoE gene, suitable for use in insurance and other applications, which will be the subject of further studies (for example, Macdonald & Pritchard (1999)). Much uncertainty remains:

(a) No single study yet exists that would allow all the intensities in the model to be estimated simultaneously. The estimation is based on a number of different studies, some quite small, of different populations, with different research protocols and methods of analysis, and very likely different definitions of 'onset of AD' and 'instititutionalisation'.

(b) The relative risks of the ApoE genotypes are based on case-based studies, not prospective population studies, and the risks associated with the $\varepsilon 4$ allele are almost certain to be lower than those estimated to date. We have been unable to do more than to show what effect this might have.

Nevertheless, certain features of our model ought to be robust. Whatever reduction in relative risks we have used, we have adjusted the genotype-specific incidence rates of AD so that the aggregated (population) incidence rates are close to those actually observed. The latter is one of the few reasonably reliable benchmarks available. Further, our model produces prevalence rates of AD that fall within the range of those observed in many studies.

As well as the intensities, we have estimated the ApoE gene frequencies at ages up to 75, in respect of lives unaffected by AD at these ages. These are needed in (for example) insurance applications.

7.2. Discussion

(a) The model specification is dictated entirely by the events studied in the medical and epidemiological literature, and not by the events that might be of interest in any particular application. If it is the case that actuarial models might, in future, need to incorporate more medical detail, it would be very useful to try to collaborate with medical and other researchers.

(b) The published conclusions of medical papers are usually in the form of summary statistics (means, medians, odds ratios and confidence intervals) and if age-related outcomes are shown they are usually in the form of graphs. These are not ideal for actuarial use. AD is a major condition, much studied, but we have had to make some crude assumptions in order to calibrate the model from published data only. There must be many medical data sets that could furnish age-related

estimates of incidence rates, if only they could be re-analysed. Again, closer collaboration between actuaries and other researchers would be valuable.

(c) Another common type of medical statistic is prevalence rates. The difference between prevalence rates and incidence rates is exactly the difference between the Manchester Unity approach to modelling Permanent Health Insurance, and the multiple-state model approach. Prevalence rates are often easier to estimate, as they can be based on census-type surveys, but it would be helpful if the greater versatility of incidence rates (transition intensities) was more widely appreciated.

(d) As a consequence of fitting the intensities using published summary statistics, it is impossible to estimate even crude confidence intervals for them. In any application, therefore, sensitivity analysis is especially important.

(e) Several epidemiological authors have suggested the use of individual patient data rather than summary or published data, partly to avoid publication bias in meta-analyses. Useful references are Piantadosi (1997), Green, Benedetti & Crowley (1997) and Friedman, Furberg & DeMets (1998).

(f) It is now about six years since the rôle of the ApoE gene in AD was confirmed. Since then, the gene has been intensively studied, to the point that meta-analyses including thousands of lives have been published. Even so, little is known about population risk, and data are very scarce in places, so that:

(1) we have had to reduce relative risks rather arbitrarily to allow for the selectiveness of case-based studies; and

(2) the relative risks for males are suspect, with the $\varepsilon2$ allele conferring such strong protection that carriers of the $\varepsilon4$ allele are not necessarily at higher risk overall.

If this is typical, it seems likely that the time between the identification of a gene disorder, and the assessment of its impact on insurance, will be of the order of ten years.

(g) Despite the fact that AD is a much-studied condition, many studies reach conflicting conclusions. When setting up an actuarial model, it is necessary to consider the body of medical research in its entirety (hence the large number of references). Inevitably, some studies must be chosen as the basis of the model, but to confine one's attention only to these risks overlooking important features or sources of variation, and could impair the credibility of the results in the eyes of medical experts.

(h) ApoE is a relatively simple gene to consider, since it has only three relevant alleles, hence six genotypes. Other genes are more complex; for example, the BRCA1 gene (predisposing to breast and ovarian cancer) has several hundred known mutations, some of which have only been observed in a single family.

(i) We cannot stress too strongly the speed at which human genetics is developing. This work started in late 1997, and since then the volume of

papers on AD and the ApoE gene has increased greatly, as the references show. It is very likely that assessments of the impact of specific genetic tests on insurance will have to be revisited quite frequently, if they are to remain credible.

Points (a) to (c) above suggest ways in which medical data might be made more useful for actuarial models, but there is no *a priori* reason why medical studies should be planned with that in mind. However, actuarial models derived from insurance practice, capable of dealing with fairly general payments while in different states or on transition between different states, could make a useful contribution to health economics, and it would be helpful to pursue collaborations from that point of view.

ACKNOWLEDGEMENTS

REFERENCES

BARCLAY, L.L., ZEMCOV, A., BLASS, J.P. and McDOWELL F.H. (1985a) Factors associated with duration of survival in Alzheimer's disease. *Biological Psychiatry* **20**, 86-93.
BARCLAY, L.L., ZEMCOV, A., BLASS, J.P. and SANSONE, J. (1985b) Survival in Alzheimer's disease and vascular dementias. *Neurology* **35**, 834-840.
BASUN, H., GRUT, M., WINBLAD, B. and LANNFELDT, L. (1995) Apolipoprotein ε4 allele and disease progression in patients with late onset Alzheimer's disease. *Neuroscience Letters* **183**, 32-34.
BEARD, C.M., KOKMEN, E., O'BRIEN, P.C. and KURLAND, L.T. (1994) Are patients with Alzheimer's disease surviving longer in recent years? *Neurology* **44**, 1869-1871.
BERG, L., MILLER, J.P., STORANDT, M., DUCHEK, J., MORRIS, J.C., RUBIN, E.H., BURKE, W.J. and COBEN, L.A. (1988) Mild senile dementia of the Alzheimer type: 2. Longitudinal assessment. *Annals of Neurology* **23**, 477-484.
BICKEBÖLLER, H., CAMPION, D., BRICE, A., AMOUYEL, P., HANNEQUIN, D., DIDIERJEAN, O., PENET, C., MARTIN, C., PÉREZ-TUR, J., MICHON, A., DUBOIS, B., LEDOZE, F., THOMAS-ANTERION, C., PASQUIER, F., PUEL, M., DEMONET, J-F., MOREAUD, O., BABRON, M-C., MEULIEN, D., GUEZ, D., CHARTIER-HARLIN, M-C., FREBOURG, T., AGID, Y., MARTINEZ, M. and CLERGET-DARPOUX, F. (1997) Apolipoprotein E and Alzheimer's disease: genotype-specific risks by age and sex. *American Journal of Human Genetics* **60**, 439-446.
BONAIUTO, S., MELE, M., GALLUZO, L. and GIANNANDREA, E. (1995) Survival and dementia: A 7-year follow up of an Italian population. *Archives of Gerontology and Geriatrics* **20**, 105-113.
BRACCO, L., GALLATO, R., GRIGOLETTO, F., LIPPI, A., LEPORE, V., BINO, G., LAZZARO, M.P., CARELLA, F., PICCOLO, T., POZZILLI, C., GIOMETTO, B. and AMADUCCI, L. (1994) Factors affecting course and survival in Alzheimer's disease. *Archives of Neurology* **51**, 1213-1219.
BRETELER, M.B.B., CLAUS, J.J., VAN DUIJN, C.M., LAUNER, L.J. and HOFMAN, A. (1992) Epidemiology of Alzheimer's disease. *Epidemiologic Reviews* **14**, 59-82.

BRINDLE, N., SONG, Y., ROGAEVA, E., PREMKUMAR, S., LEVESQUE, G., YU, G., IKEDA, M., NISHIMURA, M., PATERSON, A., SORBI, S., DUARA, R., FARRER, L. and ST GEORGE-HYSLOP, P. (1998) Analysis of the butyrycholinesterase gene and nearby chromosome 3 markers in Alzheimer's disease. *Human Molecular Genetics* 7, 933-935.

BURNS, A., LEWIS, G., JACOBY, R. and LEVY, R. (1991) Factors affecting survival in Alzheimer's disease. *Psychological Medicine* 21, 363-370.

CLAYTON, D. and HILLS, M. (1993) *Statistical methods in epidemiology.* Oxford University Press.

CONTE, S.D. and DE BOOR, C. (1972) *Elementary numerical analysis.* McGraw-Hill, New York.

COPELAND, J.R.M., DAVIDSON, I.A., DEWEY, M.E., GILMORE, C., LARKIN, B.A., MCWILLIAM, C., SAUNDERS, P.A., SCOTT, A., SHARMA, V. and SULLIVAN, C. (1992) Alzheimer's disease, other dementias, depression and pseudo-dementia: prevalence, incidence and three-year outcome in Liverpool. *British Journal of Psychiatry* 161, 230-239.

CORDER, E.H., SAUNDERS, A.M., STRITTMATTER, W.J., SCHMECHEL, D.E., GASKELL, P.C., RIMMLER, J.B., LOCKE, P.A., CONNEALLY, P.M., SCHMADER, K.E., SMALL, G.W., ROSES, A.D., HAINES, J.L. and PERICAK-VANCE, M.A. (1994) Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer's disease. *Nature Genetics* 7, 180-184.

CORDER, E.H., SAUNDERS, A.M., STRITTMATTER, W.J., SCHMECHEL, D.E., GASKELL, P.C., RIMMLER, J.B., LOCKE, P.A., CONNEALLY, P.M., SCHMADER, K.E., TANZI, R.E., GUSELLA, J.F., SMALL, G.W., ROSES, A.D., PERICAK-VANCE, M.A. and HAINES, J.L. (1995) Apolipoprotein E, survival in Alzheimer's disease and the competing risks of death and Alzheimer's disease. *Neurology* 45, 1323-1328.

DIESFELDT, H.F.A, VAN HOUTE, L.R. and MOERKENS, R.M. (1986) Duration of survival in senile dementia. *Acta Psychiatr Scandinavica* 73, 366-371.

VAN DIJK, P.T.M., DIPPEL, D.W.J. and HABBEMA, J. DIK F. (1991) Survival of patients with dementia. *Journal of the American Geriatric Society* 39, 603-610.

VAN DUIJN, C.M., DE KNIFF, P., WEHNERT, A., DE VOECHT, J., BRONZOVA, J.B., HAVEKES, L.M., HOFMAN, A. and VAN BROECKHOVEN, C. (1995) The apolipoprotein E ε4 allele is associated with an increased risk of early-onset Alzheimer's disease and a reduced survival. *Annals of Neurology* 37, 605-610.

EVANS, D.A., FUNKENSTEIN, H., ALBERT, M.S., SCHERR, P.A., COOK, N.R., CHOWN, M.J., HEBERT, L.E., HENNEKENS, C.H. and TAYLOR, J.O. (1989) Prevalence of Alzheimer's disease in a community population of older persons: Higher than previously reported. *Journal of the American Medical Association* 262, 2551-2556.

EVANS, D.A., SMITH, L.A., SCHERR, P.A., ALBERT, M.S., FUNKENSTEIN, H.H. and HEBERT, L.E. (1991) Risk of death from Alzheimer's disease in a community population of older persons. *American Journal of Epidemiology* 134, 403-412.

EVANS, D.A., BECKETT, L.A., FIELD, T.S., FENG, L., ALBERT, M.S., BENNETT, D.A., TYCKO, B. and MAYEUX R. (1997) Apolipoprotein E ε4 and incidence of Alzheimer's disease in a community population of elder persons. *Journal of the American Medical Association* 277, 822-824.

FARRER, L.A., CUPPLES, L.A., HAINES, J.L., HYMAN, B., KUKULL, W.A., MAYEUX, R., MYERS, R.H., PERICAK-VANCE, M.A., RISCH, N., VAN DUIJN, C.M. and APOE and ALZHEIMER DISEASE META ANALYSIS CONSORTIUM (1997) Effects of age, gender and ethnicity on the association between apolipoprotein E genotype and Alzheimer's disease. *Journal of the American Medical Association* 278, 1349-1356.

FORFAR, D.O., MCCUTCHEON, J.J. and WILKIE, A.D. (1988) On graduation by mathematical formula. *J.I.A.* 115, 1-149 and (with discussion) *T.F.A.* 41, 97-269.

FRIEDMAN, L.M., FURBERG, C.D. and DEMETS, D.L. (1998) *Fundamentals of clinical trials.* Springer-Verlag, New York.

FRISONI, G.B., GOVONI, S., GEROLDI, C., BIANCHETTI, A., CALABRES, L., FRANCESCHINI, G. and TRABUCCHI, M. (1995) Gene dose of the ε4 allele of apolipoprotein E and disease progression in sporadic late onset Alzheimer's disease. *Annals of Neurology* 37, 596-604.

GAO, S., HENDRIE, H.C., HALL, K.S. and HUI, S. (1998) The relationship between age, sex and the incidence of dementia and Alzheimer's disease. *Archives of General Psychiatry* 55, 809-815.

GOMEZ-ISLA, T., WEST, H.L., REBECK, G.W., HARR, S.D., GOWDON, J.H., LOCASCIO, J.J., PERLS, T.T., LEPSITZ, L.A. and HYMAN, B.T. (1996) Clinical and pathological correlates of apolipoprotein E ε4 in Alzheimer's disease. *Annals of Neurology* 39, 62-70.

GREEN, S., BENEDETTI, J. and CROWLEY, J. (1997) *Clinical trials in oncology*. Chapman and Hall, London.

HAGNELL, O., OJESJO, L. and RORSMAN, B. (1992) Incidence of dementia in the Lundby study. *Neuroepidemiology* 11, 61-66.

HEBERT, L.E., SCHERR, P.A., BECKETT, L.A., ALBERT, M.S., PILGRIM, D.M., CHOWN, M.J., FUNKENSTEIN, H.H. and EVANS, D.A. (1995) Age specific incidence of Alzheimer's disease in a community population. *Journal of the American Medical Association* 273, 1354-1359.

HEYMAN, A., PETERSON, B., FILLENBAUM, G. and PIEPER, C. (1996) The consortium to establish a registry for Alzheimer's disease (CERAD) Part XIV: Demographic and clinical predictors of survival in patients with Alzheimer's disease. *Neurology* 46, 656-660.

HEYMAN, A., PETERSON, B., FILLENBAUM, G. and PIEPER, C. (1997) Predictors of time to institutionalization of patients with Alzheimer's disease: the CERAD experience part XVII. *Neurology* 48, 1304-1309.

JARVIK, G.P., LARSON, E.B., GODDARD, K., KUKULL, W.A., SCHELLENBERG, G.D. and WIJSMAN, E.M. (1996) Influence of apolipoprotein E genotype on the transmission of Alzheimer's disease in a community-based sample. *American Journal of Human Genetics* 58, 191-200.

JORM, A.F. and JOLLEY, D. (1998) The incidence of dementia – a meta analysis. *Neurology* 51, 728-733.

JOST, B.C. and GROSSBERG, G.T. (1995) The natural history of Alzheimer's disease: a brain bank study. *Journal of the American Geriatrics Society* 43, 1248-1255.

KAHN, H. and SEMPOS, C.T. (1989) *Statistical methods in epidemiology*. Oxford University Press.

KOKMEN, E., CHANDRA, V. and SCHOENBERG, B. (1988) Trends in incidence of dementing illness in Rochester, Minnesota, in three quinquennial periods, 1960-1974. *Neurology* 38, 975-980.

KOKMEN, E., BEARD, R.N., O'BRIEN, P.C., OFFORD, M.S. and KURLAND, L.T. (1993) Is the incidence of dementing illness changing? *Neurology* 43, 1887-1892.

KUUSISTO, J., KOIVISTO, K., MYKKÄNEN, L., HELKALA, E-L., VANHANEN, M., HÄNNINEN, T., PYÖRÄLÄ, K., KESÄNIEMI, Y.A., RIEKKINON, P. and LAAKSO, K. (1994) Association of apolipoprotein E phenotypes with late onset Alzheimer's disease: population based study. *British Medical Journal* 309, 636-638.

LAMBERT, J-C., PASQUIER, F., COTTEL, D., FRIGARD, B., AMOUYEL, P. and CHARTIER-HARLIN, M-C. (1998) A new polymorphism in the APOE promoter associated with risk of developing Alzheimer's disease, *Human Molecular Genetics* 7, 533-540.

LEHMANN, D.J., JOHNSTON, C. and SMITH, A.D. (1997) Synergy between the genes for butyrylcholinesterase K variant and apolipoprotein E4 in late-onset confirmed Alzheimer's disease. *Human Molecular Genetics* 6, 1933-1936.

LEHTOVIRTA, M., HELISALMI, S., MANNERMAA, A., SOININEN, H., KOIVISTO, K., RYYNÄNEN, M. and RIEKKINEN SR. P. (1995) Apolipoprotein E polymorphism and Alzheimer's disease in Eastern Finland. *Neuroscience Letters* 185, 13-15.

LEMAIRE, J., SUBRAMANIAN, K., ARMSTRONG, K. and ASCH, D.A. (2000) Pricing term insurance in the presence of a family history of breast or ovarian cancer. *To appear in North American Actuarial Journal*.

LETENNEUR, L., COMMENGES, D., DARTIGUES, J.F. and BARBERGERGATEAU, P. (1994) Incidence of dementia and Alzheimer's disease in elderly community residents of south-western France. *International Journal of Epidemiology* 23, 1256-1261.

LEVY-LAHAD, E. and BIRD, T.D. (1996) Genetic factors in Alzheimer's disease: a review of recent advances. *Annals of Neurology* 40, 829-840.

LIDDELL, M., WILLIAMS, J., BAYER, A., KAISER, F. and OWEN, M. (1994) Confirmation of association between the ε4 allele of apolipoprotein E and Alzheimer's disease. *Journal of Medical Genetics* 31, 197-200.

LILIENFELD, D.E. and STOLLEY, P.D. (1994) *Foundations of epidemiology*. Oxford University Press.

LOPEZ, O.L., LOPEZ-POUSA, S., KAMBOH, M.I., ADROER, R., OLIVA, R., LOZANO-GALLEGO, M., BECKER, T.B. and DEKOSKY, S.T. (1998) Apolipoprotein E polymorphism in Alzheimer's disease: a comparative study of two research populations from Spain and the United States. *European Neurology* **39**, 229-233.

LUCOTTE, G., DAVID, F., LEGRAND, C., TURPIN, J-C. and CLAVEL, C. (1997) Predictive value of apolipoprotein E ε4 allele genotyping in Alzheimer's disease. *Alzheimer's Research* **3**, 7-9.

MACDONALD, A.S. (1996a) An actuarial survey of statistical models for decrement and transition data I: Multiple state, Poisson and binomial models. *British Actuarial Journal* **2**, 129-155.

MACDONALD, A.S. (1996b) An actuarial survey of statistical models for decrement and transition data III: Counting process models. *British Actuarial Journal* **2**, 703-726.

MACDONALD, A.S. (1997) How will improved forecasts of individual lifetimes affect underwriting?. *Philosophical Transactions of the Royal Society* B, **352**, 1067-1075, and (with discussion) *British Actuarial Journal* **3**, 1009-1025 and 1044-1058.

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

MACDONALD, A.S. and PRITCHARD, D.J. (1999) Genetics, Alzheimer's disease and long-term care insurance. *Genetics and Insurance Research Report No. 99/2, Department of Actuarial Mathematics and Statistics, Heriot-Watt University, Edinburgh.*

MASSULLO, C., DANIELE, A., SERIPA, D., FILIPPINI, V., GRAVINA, C., CARBONE, G., GAINOTTI, G. and FAZIO, V.M. (1998) Apolipoprotein E genotype in sporadic early- and late-onset Alzheimer's disease. *Dementia and Geriatric Cognitive Disorders* **9**, 121-125.

MAYEUX, R., STERN, Y., OTTMAN, R., TATEMICHI, T.K., TANG, M-X., MAESTRE, G., NGAI, C., TYCKO, B. and GINSBERG, H. (1993) The apolipoprotein ε4 allele in patients with Alzheimer's disease. *Annals of Neurology* **34**, 752-754.

MÖLSÄ, P.K., MARTTILA, R.J. and RINNES, U.K. (1982) Epidemiology of dementia in a Finnish population. *Acta Neurologica Scandinavica* **65**, 541-552.

MÖLSÄ, P.K., MARTTILA, R.J. and RINNES, U.K. (1986) Survival and cause of death in Alzheimer's disease and multi-infarct dementia. *Acta Neurologica Scandinavica* **74**, 103-107.

MYERS, R.H., SCHAEFFER, E.J., WILSON, P.W.F., D'AGOSTINO, R., ORDOVAS, J.M., ESPINO, A., WHITE, R.F., KNOEFEL, J.E., COBB, J.L., MCNULTY, K.A., BEISER, A. and WOLF, P.A. (1996) Apolipoprotein E ε4 association with dementia in a population based study: the Framingham study. *Neurology* **46**, 673-677.

NALBANTOGLU, J., GILFIX, B.M., BERTRAND, P., ROBITAILLE, Y., GAUTHIER, S., ROSENBLATT, D.S. and POIRIER, J. (1994) Predictive value of apolipoprotein E genotyping in Alzheimer's disease: results of an autopsy series and an analysis of several combined studies. *Annals of Neurology* **36**, 889-895.

NILSSON, L.V. (1984) Incidence of severe dementia in an urban sample followed from 70 to 79 years of age. *Acta Psychiatrica Scandinavica* **70**, 478-486.

NORBERG, R. (1995) Differential equations for moments of present values in life insurance. *Insurance: Mathematics and Economics* **17**, 171-180.

NORRMAN, J., BROOKES, A.J., YATES, C. and CLAIR, D.S. (1995) Apolipoprotein E genotype and its effect on duration and severity of early and late onset Alzheimer's disease. *British Journal of Psychiatry* **167**, 533-536.

O'CONNOR, D.W., POLLITT, P.A., HYDE, J.B. *et al.* (1989) The prevalence of dementia as measured by the Cambridge Mental Disorders of the Elderly Examination. *Acta Psychiatrica Scandinavica* **79**, 190-198.

OTT, A., BRETELER, M.M.B., VAN HARSKAMP, F., STIJNEN T. and HOFMAN, A. (1998) Incidence and risk of dementia. *American Journal of Epidemiology* **147**, 574-580.

PAYAMI, H., SCHELLENBERG, G.D., ZAREPARSI, S., KAYE, J., SEXTON, G.J., HEAD, M.A., MATSUYAMA, S.S., JARVIK, L.F., MILLER, B., MCMANUS, D.Q., BIRD, T.D., KATZMAN, R., HESTON, L., NORMAN, D. and SMALL, G.W. (1997) Evidence for association of HLA-A2 allele with onset age of Alzheimer's disease. *Neurology* **48**, 512-518.

PERICAK-VANCE, M.A., BEBOUT, J.L., GASKELL, P.C., YAMAOKA, L.H., HUNG, W.Y., ALBERTS, M.J., WALKER, A.P., BARTLETT, R.J., HAYNES, C.A., WELSH, K.A., EARL, N.L., HEYMAN, A., CLARK, C.M. and ROSES, A.D. (1991) Linkage studies in familial Alzheimer's disease: evidence for chromosome 19 linkage. *American Journal of Human Genetics* **48**, 1034-1050.

PFEFFER, R.I., AFIFI, A.A. and CHANCE J.M. (1987) Prevalence of Alzheimer's disease in a retirement community. *American Journal of Epidemiology* **125**, 420-436.

PIANTADOSI, S. (1997) *Clinical trials: A methodologic perspective.* John Wiley and Sons, New York.

PODUSLO, S.E., NEAL, M., HERRING, K. and SHELLY, J. (1998) The apolipoprotein CI A allele as a risk factor for Alzheimer's disease. *Neurochemical Research* **23**, 361-367.

POIRIER, J., DAVIGNON, J., BOUTHILLIER, D., KOGAN, S., BERTRAND, P. and GAUTHIER, S. (1993) Apolipoprotein E polymorphism and Alzheimer's disease. *Lancet* **342**, 697-699.

PRITCHARD, D.J. (1997) *Life assurance: financial implications of a change in insuring behaviour resulting from individuals' increased knowledge of their genetic predispositions.* M.Sc. dissertation, Heriot-Watt University, Edinburgh.

ROCCA, W.A., CHA, R.H., WARING, S.C. and KOKMEN, E. (1998) Incidence of dementia and Alzheimer's disease: a reanalysis of data from Rochester, Minnesota, 1975-1984. *American Journal of Epidemiology* **148**, 51-62.

RORSMAN, B., HAGNELL, O. and LANKE, J. (1986) Prevalence and incidence of senile and multi-infarct dementia in the Lundby study: a comparison between the time periods 1947-1957 and 1957-1972. *Neuropsychobiology* **15**, 122-129.

ROSES, A.D. (1995) Apolipoprotein E genotyping in the differential diagnosis, not prediction, of Alzheimer's disease. *Annals of Neurology* **38**, 6-14.

SÁYETTA, R.B. (1986) Rates of senile dementia – Alzheimer's type in the Baltimore longitudinal study. *Journal of Chronic Disabilities* **39**, 271-286.

SELVIN, S. (1996) *Statistical analysis of epidemiologic data.* Oxford University Press.

SEVERSON, M.A., SMITH, G.E., TANGALOS, E.G., PETERSON, R.C., KOKMEN, E., IVNIK, R.J., ATKINSON, E.J. and KURLAND, L.T. (1994) Patterns and predictors of institutionalization in community-based dementia patients. *Journal of the American Geriatrics Society* **42**, 181-185.

SINGLETON, A., SMITH, G., GIBSON, A.M., WOODWARD, R., PERRY, R.H., INCE, P.G., EDWARDSON, J.A. and MORRIS, C.M. (1998) No association between the K variant of the butyrylcholinesterase gene and pathologically confirmed Alzheimer's disease. *Human Molecular Genetics* **7**, 937-939.

SLOOTER, A.J.C. and VAN DUIJN, C.M. (1997) Genetic epidemiology of Alzheimer disease. *Epidemiologic Reviews* **19**, 107-119.

SLOOTER, A.J.C., CRUTS, M., KALMIJN, S., HOFMAN, A., BRETELER, M.M.B., VAN BROECKHOVEN, C. and VAN DUIJN, C.M. (1998) Risk estimates of dementia by apolipoprotein E genotype from a population-based incidence study: the Rotterdam study. *Archives of Neurology* **55**, 964-968.

STERN, G., TANG, M.X., DENARO, J. and MAYEUX, R. (1995) Increased risk of mortality in Alzheimer's disease patients with more advanced educational and occupational attainment. *Annals of Neurology* **37**, 590-595.

STERN, Y., BRANDT, J., ALBERT, M., JACOBS, D.M., LIU, X., BELL, K., MARDER, K., SANO, M., ALBERT, S., CASTENADA, C.D-C., BYLSMA, F., TYCKO, B. and MAYEUX, R. (1997) The absence of an apolipoprotein $\varepsilon 4$ allele is associated with a more aggressive form of Alzheimer's disease. *Annals of Neurology* **41**, 615-620.

STRACHAN, T. and READ, A.P. (1999) *Human molecular genetics (second edition).* BIOS Scientific Publishers, Oxford.

STRITTMATTER, W.J., SAUNDERS, A.M., SCHMECHEL, D., PERICAK-VANCE, M., ENGHILD, J., SALVESEN, G.S. and ROSES, A.D. (1993) Apolipoprotein E: High-avidity binding to β-amyloid and increased frequency of type-4 allele in late-onset Alzheimer's disease. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 1977-1981.

SUBRAMANIAN, K. and LEMAIRE, J. (2000) Estimating adverse selection costs in a market with genetic testing for breast and ovarian cancer: the case of life insurance. *To appear in Journal of Risk and Insurance.*

TAN, K.W. (1997) *The financial impact of genetic testing on annuities.* M.Sc. dissertation, Heriot-Watt University, Edinburgh.

TIERNEY, M.C., FISHER, R.H., LEWIS, A.J., ZORZITTO, M.L., SNOW, W.G., REID, D.W. and NIEUWSTRATEN, P. (1988) The NINCDS-ADRDA work group criteria for the clinical diagnosis of probable Alzheimer's disease: a clinicopathologic study of 57 cases. *Neurology* **38**, 359-364.

TREVES, T., KORCZYN, A.D., ZILBER, N., KAHANA, E., LEIBOWITZ, Y., ALTER, M. and SCHOENBERG, B.S. (1986) Presenile dementia in Israel. *Archives of Neurology* **43**, 26-29.

TSAI, M-S., TANGALOS, E.G., PETERSEN, R.C., SMITH, G.E., SCHAID, D.J., KOKMEN, E., IVNIK, R.J. and THIBODEAU, S.N. (1994) Apolipoprotein E: risk factor for Alzheimer's disease. *American Journal of Human Genetics* **54**, 643-649.

WALSH, J.S., WELCH, H.G. and LARSON, E.B. (1990) Survival of outpatients with Alzheimer-type dementia. *Annals of Internal Medicine* **113**, 429-434.

WARREN, V., BRETT, P., MACDONALD, A.S., PLUMB, R.H. and READ, A.P. (1999) *Genetic tests and future need for long-term care in the UK: report of a Work Group of the Continuing Care Conference Genetic Tests and Long-term Care Study Group.* Continuing Care Conference, London.

WATERS, H.R. (1984) An approach to the study of multiple state models. *Journal of the Institute of Actuaries* **111**, 363-374.

WATERS, H.R. and WILKIE, A.D. (1987) A short note on the construction of life tables and multiple decrement tables. *Journal of the Institute of Actuaries* **114**, 569-580.

WATERS, H.R. (1991) Computational procedures for the model. *Continuous Mortality Investigation Bureau Report* **12**, 79-96.

ANGUS MACDONALD AND DELME PRITCHARD
Department of Actuarial Mathematics and Statistics
Heriot-Watt University
Edinburgh EH14 4AS
United Kingdom