Continuous Mortality Investigation

Critical Illness Committee

WORKING PAPER 28

Progress towards an improved methodology for analysing CMI critical illness experience

July 2007

Continuous Mortality Investigation

Critical Illness Committee

WORKING PAPER 28

Progress towards an improved methodology for analysing CMI critical illness experience

CONTENTS

Continuous Mortality Investigation

Critical Illness Committee

Progress towards an improved methodology for analysing CMI critical illness experience

1. INTRODUCTION

1.1. This paper introduces a development to the methodology previously used by the CMI Critical Illness Committee to assist with the interpretation of critical illness insurance claims experience. It illustrates this development for experience in 1999-2002 and discusses the further work required to enable realistic claim rates for critical illness business to be produced.

1.2. The main area of difficulty in analysing claims experience within the CMI critical illness investigation has been the substantial delays from diagnosis of a claim to settlement. This and other issues with the data are discussed in Working Paper 14, which was published when the results for 1999-2002 were issued to member offices in May 2005.

1.3. Working Paper 14 also introduced the concept of a "grossing-up factor" that sought to adjust the reported experience from settled claims to diagnosed claims, that can be meaningfully compared with the exposure. The CMI CI Committee published Working Paper 18, later in 2005, to document the feedback received on Working Paper 14 which largely reinforced the Committee's desire to produce grossing-up factors for subsets of the data. These are required to understand the claims experience by age, gender, smoker status, duration, calendar year, cause of claim and other factors.

1.4. The method used to estimate grossing-up factors in Working Paper 14 required assumptions to be made about the growth in expected claims. The approach used to do this was relatively crude and it was then difficult to produce grossing-up factors for subsets of the data. We believe the development to the methodology outlined in this paper is a significant step forward in overcoming this difficulty.

1.5. Furthermore, the methodology used to calculate the underlying delays in claim settlement was data-intensive. This also inhibited estimation of grossing-up factors for subsets of the data. Further analysis of claim delays is still required before we are able to use our revised methodology to produce reliable results for subsets of the data and to develop realistic claim rates for critical illness business.

1.6. It is important to note that our previous methodology was developed to avoid the understatement of the experience that arises from comparing settled claims with the exposure in the corresponding year. The proposed development of the methodology seeks to make better use of the data fields available. We do not consider that this is

the only possible methodology which could be used but we believe it provides a practical solution and our analysis indicates that it produces credible results.

1.7. In particular, the methodology has been developed to cope with incomplete information on dates of claim within the dataset. The CMI CI investigation is now receiving date of diagnosis on an increasing proportion of claims and, in time, we hope that this proportion will be sufficiently high to allow us to use more conventional methods, such as deducting claims diagnosed before the investigation period from those settled during the period and applying IBNS adjustments to estimate the total diagnosed claims in the period. Whilst this proportion has increased again for the 2005 data collected to date, we think it unlikely that we will be able to adopt a more conventional approach yet. Even if we could, an alternative method is still required to analyse experience for the years to 2004.

1.8. The paper first sets out the background to this work in section 2, which includes a discussion on existing methodologies. We then introduce our proposed development to the methodology in section 3. Aspects of the revised methodology, together with the inputs and assumptions for an initial application, are discussed further in section 4. Section 5 sets out the results of this initial application of the revised methodology to the lives experience for full acceleration business in 1999-2002, whilst the sensitivity of these results to some of the assumptions is considered in section 6. Section 7 re-states our rationale for the development, compares its outputs with those from our previous methodology and considers the future use and development of the methodology. Section 8 sets out the further work required to complete an implementation of the revised methodology.

1.9. All feedback on this methodology will be warmly welcomed by the CMI CI Committee. Please see section 9 for details on providing feedback.

2. BACKGROUND

2.1. As noted above, the main area of difficulty in analysing claims experience within the CMI Critical Illness investigation has been the substantial delays from diagnosis of a claim to settlement. This and other issues are considered more fully in Working Paper 14, but are briefly summarised below for completeness. In this section we consider a conventional actuarial approach to addressing these issues and also our previous methodology, using grossing-up factors, as described in Working Paper 14.

2.2. The CMI collects data on critical illness business on a calendar year basis. Given that critical illness business is subject to significant delays between the date of diagnosis, when the critical illness claim is incurred, and the date of settlement, it would be impractical to wait for all the claims diagnosed in a particular year to be settled before collecting and analysing the data. The CMI therefore asks for claims to be submitted on the basis of claims settled during the year. This results in a mis-match between the exposure and claims. Given the substantial growth in business submitted to the investigation since it started, this mis-match is especially pronounced. The results that the CMI has released to date, providing ratios of actual settled claims to expected diagnosed claims, cannot therefore be considered a reliable guide to the true underlying experience.

2.3. We are therefore seeking a means to adjust the results to produce a reliable indicator of actual diagnosed claims to expected diagnosed claims.

Chain ladder technique

2.4. This is an established actuarial technique which might be considered appropriate to apply in the analysis of critical illness experience to enable the estimation of actual diagnosed claims.

2.5. The first step is to re-allocate the claims settled in each year to the appropriate years of diagnosis. The next step is to adjust the known diagnosed claims in each year by adding the "Incurred But Not Settled" (IBNS) claims that are expected to be settled in future years.

2.6. To do this one estimates the claim settlement delay distribution from the available data and assumes that the years that are not fully-developed will be subject to the same development pattern. In a simplistic application:

- Claims are documented by year of occurrence (*j*) and by curtate duration at settlement (*k*);
- The development ratio for any particular duration at settlement K is estimated as:

$$\Sigma\, C_{j,K+1} \,/\, \Sigma\, C_{j,K}$$

where $C_{j,k}$ is the cumulative claims paid by the end of year $(j+k)$ in respect of claims incurred in year j and the summation is over years of occurrence, j, for which settlements have progressed to duration $k+1$.

2.7. This approach therefore involves estimating the total actual diagnosed claims from those that are known using an assumption that the claim settlement delay distribution derived from previous periods can be used for the current period. These estimates would normally be refined over time, as more up-to-date information becomes available, until ultimately the estimates are entirely replaced by actual information.

2.8. Such an approach has been adopted for the Irish critical illness investigation, for example. This has not been used for the UK CMI investigation, principally because claim dates were not received for a significant proportion of data. In particular, the CMI only received a date of diagnosis for 56% of the claims settled during 1999-2002. As a result, we do not know which claims should be removed from the data because they relate to diagnoses before the investigation period. This led the Committee to consider that it could not reasonably adopt such an approach for the investigation periods reported to date and to seek alternative methodologies.

2.9. In addition, there were a number of issues that limited the confidence that the Committee felt could be placed on dates of diagnosis and on any claim settlement delay distribution derived from the data. These issues, which apply to all methodologies under consideration, are as follows:

- The incomplete information referred to above also means that we have less data from which to estimate the claim delay distribution.
- There was considerable growth in business volumes during 1999-2002 and in the preceding years. This means that claims with long delays to settlement are under-represented in the 1999-2002 settled claims, increasing the

uncertainty in any estimate of the claim delay distribution and the IBNS adjustment.

- There is a lack of precision and consistency between claims assessors regarding the definition of the date of diagnosis.

Grossing-up factors

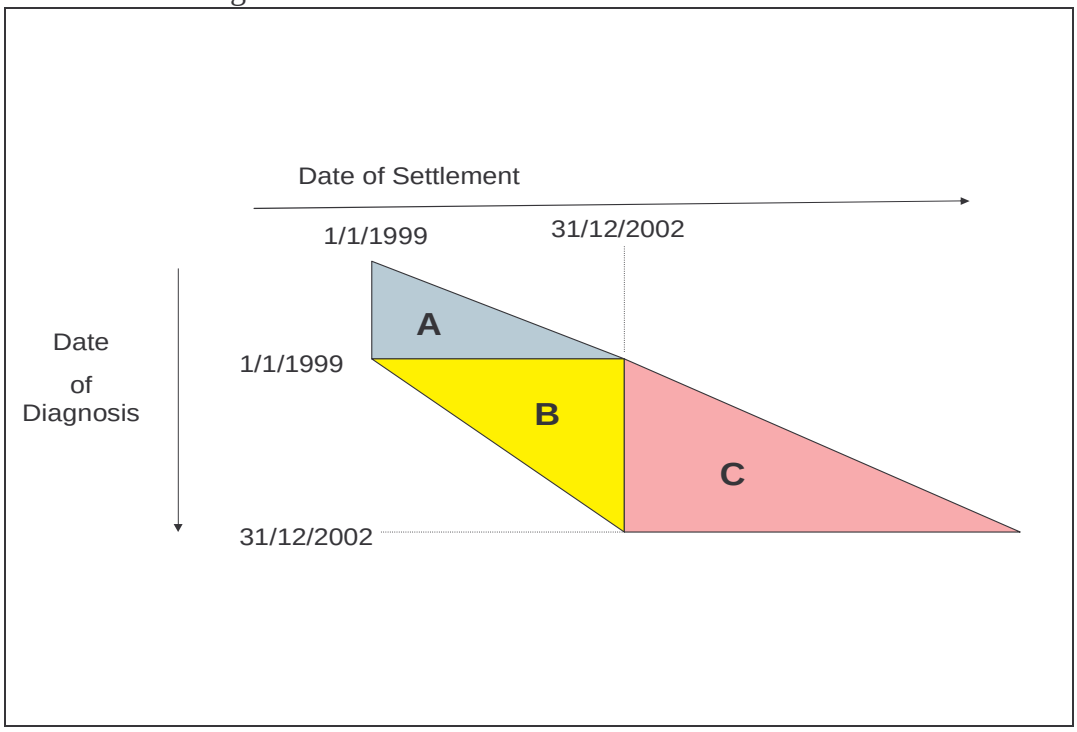
2.10. The critical illness results that have been issued to date by the CMI compare actual settled claims with expected diagnosed claims. In order to compensate for this mis-match, we have indicated the need to apply a "grossing-up factor", which seeks to convert actual settled claims to actual diagnosed claims. The method we used is described in Working Paper 14.

2.11. When the Committee released the 1999-2002 results, it indicated approximate grossing-up factors to correct for the distortion of the results. It is important to note that these grossing-up factors implicitly combined:

- the removal of claims settled in the investigation period but diagnosed (or estimated to be diagnosed) in prior periods; and
- the addition of claims diagnosed in the investigation period but settled in later periods or yet to be settled.

2.12. The schematic diagram reproduced below was used to illustrate the role of grossing-up factors. The actual settled claims reported to the CMI are shown as areas A and B. Claims in area B have dates of diagnosis falling within the investigation period. Claims in area A have dates of diagnosis preceding the investigation period. Note that the impact of unknown dates of diagnosis is that, in many cases, we do not know whether a particular claim falls into A or B.

2.13. Claims in area C also have dates of diagnosis within the investigation period but have yet to be reported to the CMI. Mis-statement in released results arises from the inequality between areas A and C. The 1999-2002 'All Offices' results released by the CMI under-stated true experience as the claims in area C exceeded the claims in area A due to the growth in the number of claims over time.



2.14. A key assumption involved in the estimation of grossing-up factors was the growth in expected claims, which drives the proportion of claims settled in the period that are estimated to have been diagnosed in prior periods, i.e. the split between A and B in the diagram above. The Committee expected this assumption to vary for subsets of the dataset – for example by duration – but thought it had insufficient information to estimate these accurately, as they relate to periods prior to the commencement of the investigation. Our proposed development to the methodology seeks to overcome this issue, recognising that we have important information available to us that was not fully utilised in the grossing-up factor approach.

2.15. As noted in 1.5, the method used to estimate the underlying claim delay distribution in Working Paper 14 was data-intensive, which also meant that it was difficult to apply to subsets of the data. This is an area where we intend to undertake further work, but it has not been resolved in our initial application of the revised methodology described in this Working Paper.

2.16. Finally it is worth noting that the CI Committee took the unprecedented step of making the 1999-2002 dataset available to member offices to allow them to undertake their own analyses and explore other methodologies. The dataset contained individual records for both in force and claims, with the minimum number of fields removed to preserve confidentiality of contributing offices' data.

3. OVERVIEW OF THE REVISED METHODOLOGY

3.1. In this section we aim to provide a succinct, generic description of the revised methodology. Further discussion of the inputs and assumptions required by the methodology and, in particular, those used in an initial application of the methodology, is contained in section 4. A pictorial representation of the methodology is provided in Appendix A.

3.2. We start with the known in force data. For the purpose of this description, we assume we have consistent data supplied throughout the period 1999-2002 (see section 4.2 for discussion of where this does not apply). The data therefore consists of five in force files at 1st January in each year from 1999 to 2003, inclusive. Each of these contains a list of in force records containing key risk factors such as gender and smoker status.

3.3. Note that we are assuming that we do not have access to in force data for the years before the investigation period, and the steps described below attempt to estimate this. It would of course be preferable to use actual data and this may be possible, for example, for an office seeking to use our approach to analyse its own experience.

3.4. This data can then be rolled back in time to previous year-ends: records in force at age X and duration T at 1/1/1999 are assumed to have been in force at age X-1 and duration T-1 at 1/1/1998 (unless of course, T=0, i.e. the policy was issued in 1998). Data is rolled back until the record with the longest duration at 1/1/1999 no longer exists, i.e. we are assuming that we have a record still in existence relating to the first year in which each office wrote business.

3.5. From this in force data, an estimate is made of the rate at which in force business goes off the books (the "off rate"). Note that business may go off for a number of reasons (maturity, expiry, surrender, lapse, claim) and the off rate is a composite variable encompassing all of these. It is likely that there were few maturities or expiries during the period with which we are concerned.

3.6. This off rate is assumed to apply in preceding years and is used to estimate the in force data at these previous year-ends that had gone off the books before 1/1/1999. We refer to this as "synthetic" in force data in the remainder of this paper. For example, if there were 100 policies in force at 1/1/1999 at age X and duration T and the assumed off rate is 9%, then (provided T is greater than zero) it will be assumed that there were 110 policies in force at 1/1/1998 at age X-1 and duration T-1; 100 known policies and 10 synthetic policies.

3.7. We can vary the off rate by age, duration, office or any other variable that is considered to be significant but we are nonetheless applying a rate estimated from a period where we have data to an earlier period. Whilst we recognise that generating synthetic data is a speculative assumption – and one that cannot be validated from the data supplied to the CMI – we believe that it has low potential to distort the results. This is discussed further in section 4.7.

3.8. With in force data for all these prior years available to us, we can now estimate the exposure for each of these years. This can be done using a census method, by averaging the start- and end-year in force figures. Alternatively, an exact method of calculating exposure can be employed using commencement dates and off dates, if these are known. To do this, assumptions regarding these dates would be required for the synthetic data.

3.9. Note that the exposure will retain whatever policy details are maintained in the rolling back of the in force data, e.g. age, gender, smoker status, sales channel, etc, provided that synthetic data is also generated using these fields.

3.10. Expected diagnosed claims are then calculated by multiplying the estimated exposed to risk by a claim rate. Note that these diagnosed claims relate not only to 1999-2002, but spread across all the years for which we have exposure.

3.11. Each diagnosed claim gives rise, in due course, to a settled claim. We therefore apply a claim delay distribution to each diagnosed claim to project it forward to settlement. We do not assign a single date of settlement to each claim; instead a proportion of each diagnosed claim is distributed into each category of age, duration and calendar year at settlement. For example, a claim diagnosed at age X and duration T in calendar year Y could be settled at (X,T,Y), (X+1,T,Y), (X,T+1,Y), (X,T,Y+1), (X+1,T+1,Y+1), etc and an appropriate proportion of each claim is assigned to each of these categories. These proportions are summed to give the total expected settled claims by age, duration and calendar year (and any other fields previously retained, such as smoker status).

3.12. We can now compare the expected settled claims during 1999-2002, estimated above, with the actual settled claims that have been provided to us by offices and are known.

3.13. If we stop at this point, this will tell us how the actual settled claims compare with the expected settled claims, based on the table of claim rates and the delay distribution that we have used. This is the basis on which results are presented in section 5 of this paper. Note that this does not achieve our original goal of comparing actual diagnosed claims to expected diagnosed claims. This is considered further in section 7.

3.14. It is however possible to take the methodology a stage further in order to estimate the claim rates applicable to diagnoses. This can be done by adjusting the initial set of assumed claim rates, perhaps using an iterative approach, so that the expected settled claims, by age and duration, compare with the actual settled claims to an acceptable level of accuracy. Note that these rates do not apply to a well-defined period, in that they relate to rates of claim diagnosis applicable to claims settled in the period 1999-2002. This is a slightly earlier period on average. We have not undertaken this additional step in our initial application.

3.15. We would expect that the process described in paragraphs 3.10 to 3.14 above would be carried out separately for male non-smokers, male smokers, female non-smokers and female smokers so that experience, and rates, are specific to each of these categories. Furthermore the process does not necessarily need to be carried out at an "all causes" level. Cause-specific claim rates and claim delay distributions could be used, in which case the actual settled claims should also relate to that specific cause of claim only.

4. AN INITIAL APPLICATION OF THE REVISED METHODOLOGY

4.1. In this section we provide more detail on the assumptions and inputs required by the methodology both generically and, more specifically, in our initial application to produce the results set out in section 5. This initial application uses lives experience for full acceleration business in 1999-2002

In force data

4.2. In Section 3, we considered applying the methodology to a consistent dataset over the period 1999-2002. However there were numerous changes in the composition of business included in the investigation during 1999-2002, due to changes in the portfolios on which individual offices submitted data and to offices joining or leaving the investigation. Overall these changes substantially increased the size and coverage of the investigation.

4.3. As a result, the CI Committee introduced the concept of "submission groups" within its work. These groups divide the data into subsets according to the years for which data is thought to have been submitted consistently. The data made available to member offices (see 2.16) included this field, to enable discontinuities within the data to be taken into account in further analysis, whilst not including any field as sensitive as an office number. The 1999-2002 data was therefore sub-divided into five submission groups, as follows:

Submission Group	Years of submission
1	Data submitted consistently throughout 1999-2002
2	Data submitted consistently throughout 1999-2001
3	Data submitted consistently throughout 2000-2002
4	Data submitted consistently for 1999 only
5	Data submitted consistently for 2002 only

4.4. Although not explicitly stated, the overview of the methodology in section 3 thus considered the application of the methodology to submission group 1. (This has also been used for the pictorial illustration in Appendix A.) The approach is easily adapted for the other submission groups. For example, for submission group 3, our starting data consists of four in force files at 1st January in each year from 2000 to 2003, inclusive. The roll back and inflating of exposure is therefore required for 1999 and prior years and only the expected settled claims during 2000-2002 will be compared with the actual settled claims in the final stage.

4.5. In our initial application, we have used in force data divided by gender, smoker status, age and duration. Furthermore we have considered only 'lives' experience, not 'amounts' at this stage – the methodology can obviously be applied to either.

4.6. We have grouped data by age nearest and curtate duration as at 1/1/Y within each in force file for our initial work. This has been done for pragmatic reasons (to avoid a computer-intensive routine, involving analysis at an individual record level) but restricts us to a census method for the calculation of exposure.

Roll-back of in force data

4.7. As explained in Section 3, there are two elements to the in force data estimated for prior years:

- The initial in force data is rolled back to the year each policy commenced; and
- The known in force data in these earlier years needs to be inflated for the policies that have gone off the books before the date of the first in force data file, in order to estimate the full exposure in the period. We termed this additional in force data, synthetic data.

4.8. There are two assumptions implicit within the roll-back of in force data to previous year-ends. The first is that records that are in force at age X, duration T at 1/1/Y are assumed to have been in force at age X-1 and duration T-1 at 1/1/(Y-1), unless, of course, T=0. The Committee believes that it will be exceptional for this not to be the case, for example even where a policy has lapsed and been reinstated, it may not have been in force at the previous year-end, but may well have been for the year-end preceding that. As a result, we consider this a reasonable assumption to make.

4.9. The second assumption is that the earliest year that the in force data is projected back to is derived from the first set of in force records submitted to the CMI (e.g. 1/1/1999 for submission group 1) and we are assuming that no older policies existed that have now all gone off the books. The Committee does not consider that this assumption has any material significance – in particular, exposure relating to the earliest years of critical illness business may not affect settled claims in 1999-2002, depending on the assumed length of the claim delay period.

4.10. Note that in rolling back the in force data, we can choose to retain whatever policy details we feel are likely to be significant to the claims experience, e.g. gender, smoker status, sales channel, etc, provided that synthetic data is also generated separately for these fields. In comparison to our previous methodology, which involved estimating the growth in expected claims, this approach uses the age and duration profile from the start in force to estimate the prior years' in force - this is the important information not previously being used that we referred to in section 2.14. Given that increasing volumes of critical illness business were sold during the 1990s, the impact of the rolling-back will be to move business towards younger ages and shorter durations, reflecting the business profile at that time.

Synthetic data and off rates

4.11. By definition we have to use some form of model or assumptions to generate the synthetic data. Our approach has been to estimate off rates from the period where we do have data and assume these can be applied in earlier years.

4.12. Intuitively this feels like an heroic assumption, since lapses and surrenders (which should form the majority of offs during this period) may be subject to a variety of factors, including economic conditions and changes in the price and design of products. However, as described in section 6, we have investigated the sensitivity of the results to this assumption and we feel confident it introduces little scope for distortion of the results. The following example illustrates the reasoning behind this:

- Suppose our initial in force data is 1/1/1999. The exposure in 1998 is calculated from the estimated in force at 1/1/1998 and the known in force at 1/1/1999.
- Only 10% of the 1/1/1998 in force data is synthetic (if the off rate is 9%), hence only around 5% of the 1998 exposure has been estimated. If volumes of new business have increased, then this proportion will be lower still.
- This exposure is used to calculate expected diagnosed claims in 1998, a proportion of which are then settled in 1998 so do not concern us.
- The error in the expected settled claims in 1999-2002 arising from having to estimate the 1/1/1998 in force is therefore small in comparison to the expected diagnosed claims in 1999-2002 that are settled during those years.
- Whilst the estimation error increases for the in force data in earlier years, the exposure in these years will be lower than in 1999-2002 due to the growth in business during the 1990s, plus a lower proportion of the expected diagnosed claims in the earlier years will have delays sufficiently long to bring them into the expected settled claims in 1999-2002.

4.13. Analysis of the progression of the in force data from 1/1/1999 to 1/1/2001 led us to use an off rate assumption of 9% p.a. The methodology would allow this to vary by office or duration, for example, but we have used a single assumption in our initial application since:

- We are applying the assumption to different calendar years, so any refinement may be spurious; and
- We believe the impact of the assumption to be low as discussed above and supported by the results of tests as set out in section 6 below.

We intend undertaking further analysis in this area before pursuing a more rigourous application of the methodology.

Exposure calculation and dates of events

4.14. In Section 3, we noted alternative methods for calculating the exposure.

4.15. If a census method is used, this implicitly assumes that policies enter and exit midway through the calendar year on average. This is the method that has been adopted for our initial application although we believe that a more accurate calculation is essential to produce accurate results. This is considered further in section 8 on Further Work.

4.16. Whilst we have made no explicit allowance in the exposure calculation for exposure after the date of diagnosis for claims, the exposure should not necessarily be regarded as central exposure. We suspect that policies will often remain premium-paying until (around) the date of settlement. Hence such policies will often exist in the in force data at year-end(s) following the date of diagnosis.

4.17. We therefore regard the definition of our exposure calculation as indeterminate. This potential mis-statement of exposure can only be addressed when we have more complete submission of dates of diagnosis.

4.18. Use of an exact method of exposure calculation would necessitate other assumptions:

- Within the data collected by the CMI we have date of commencement but we do not (currently) have date of exit, so this would need to be assumed for policies that leave the investigation;
- We would also need to estimate dates of commencement, dates of birth and dates of exit for the synthetic policies that we generate before the investigation period through the application of the off rate.

Claim rates

4.19. For our initial application we have calculated expected claims using the all causes rates from CIBT93, as these have been used as the main comparison basis within the results we have released to offices. CIBT93 does not differentiate between smokers and non-smokers or by duration, hence any differences in the experience will emerge in different values of Actual/Expected.

4.20. Note that when CIBT93 was published (in "A Critical Review"), the paper stated that the age definition of the CIBT93 table was age exact, which is the age definition of most standard actuarial tables. However, the paper also contained the results of an investigation into the claims experience in 1991-1997 and in producing the expected claims it appears that CIBT93 was used as if the rates were age nearest. After consulting with the authors, the CMI has used the table consistently with the latter interpretation, i.e. assuming the rates are age nearest, in the results it has released to date. Hence in this application we have again assumed that the CIBT93 rates are age nearest.

4.21. CIBT93 spans ages 20 to 80. We have a very small amount of exposure at younger ages and have assumed that the rates at these ages equal the rates at age 20. There is no exposure at ages above 80 in the 1999-2002 data.

4.22. We have used the CIBT93 rates without any adjustment but, as noted in 4.17 above, the definition of our exposure is unclear as to whether we should use central or initial rates.

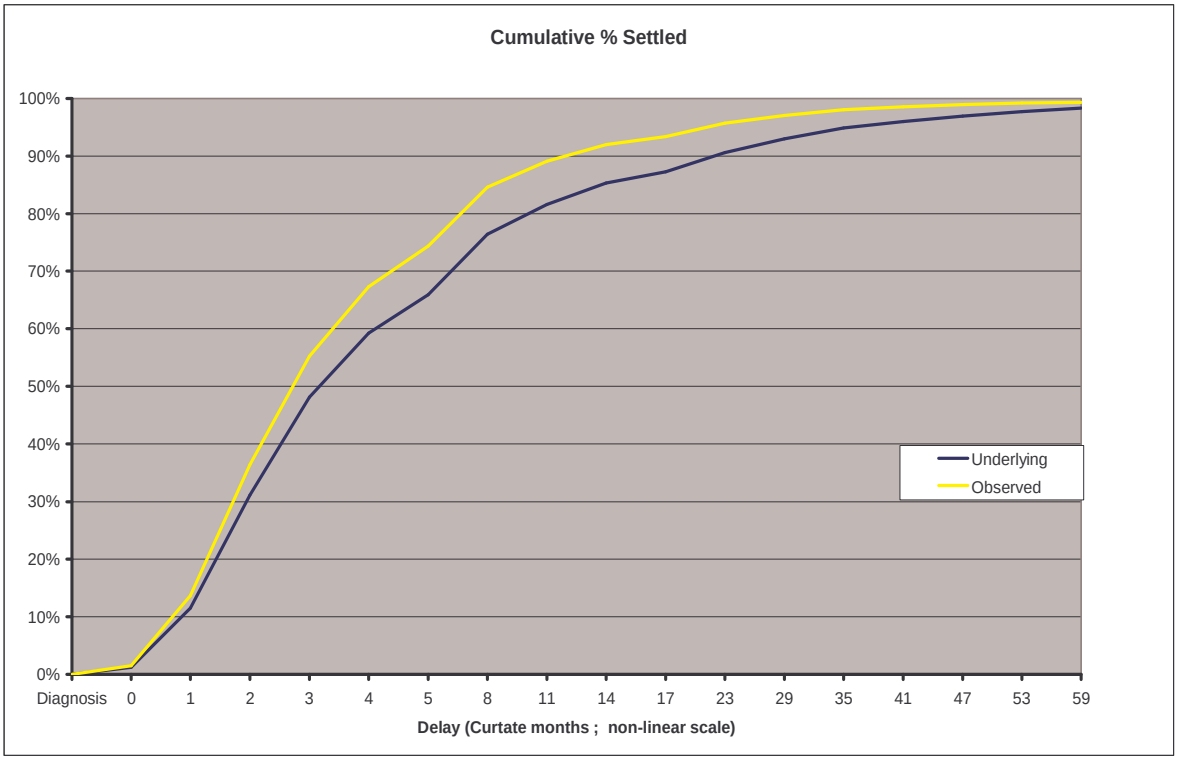
4.23. Note that we are potentially applying the claim rates across a relatively long period of exposure, perhaps from 1988 to 2002. The earliest years though have very little exposure and few (if any) of the claims diagnosed at that time will have delays to settlement sufficiently long to bring them into the expected settled claims in 1999-2002. In the initial application we are not rolling back in force data prior to 1/1/1993.

4.24. The methodology does not easily cater for varying the claim rates by calendar year, except in a very simplistic manner such as assuming they increase by *x%* each year. This is because the expected settled claims in 1999-2002 arise from diagnoses in a range of years. To illustrate this, suppose a given set of claim rates is a perfect fit to the overall experience during the investigation period, so that overall we have an A/E of 100%. Compare this to a situation where diagnosis rates increased during the period from, say 95% of the base table in the early part of the period to 105% in the latter part: in this situation, there will be fewer expected settled claims in 1999-2002 from the early years, but more from the latter years. This could still result in an overall A/E of 100%.

Claim delays

4.25. Applying claim rates to the exposure produces expected diagnosed claims at each age and duration for each calendar year within our extended investigation period. We then need to make an assumption about how long it takes for claims to be settled in order to convert these expected diagnosed claims into expected settled claims.

4.26. The estimation of the underlying delay distribution from the actual settled claims in 1999-2002 was detailed in section 5.5 of Working Paper 14 and illustrated in Figure 4 (reproduced below). The delay patterns shown reflect the cumulative proportion of claims settled by the end of each period from claims diagnosed during month 0 (rather than claims diagnosed at the start of month 0):



4.27. Note that both the observed and underlying claim delay curves were derived from the 1999-2002 settled claims where offices provided both the date of diagnosis and the date of settlement. In using either delay distribution we are therefore making an assumption that this distribution is also appropriate for use with the significant minority of claims where we do not have both dates. In particular, the useable data for this analysis comes from only a subset of offices in the full experience investigation, so we are assuming that the delay pattern exhibited by these offices applies equally to the other offices.

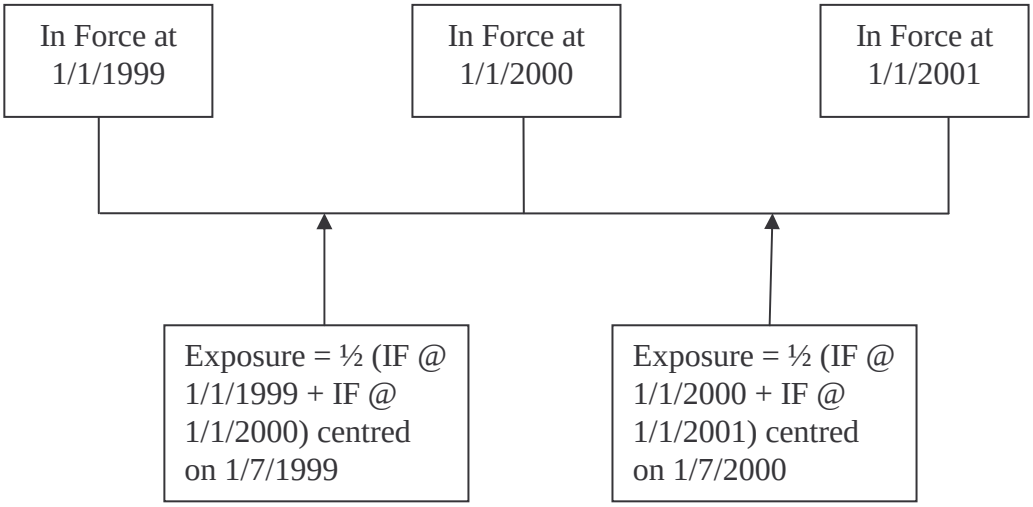
4.28. Note also that there was very little data with long delays to settlement, so the 1999-2002 data did not provide us with a reliable basis for estimating the underlying claim delay distribution in this area. Our approach was therefore to use the actual data for claim delays only up to 5 years. Beyond that point, we used a simple extrapolation in Working Paper 14. A slightly different approach has been used beyond 5 years in our initial application, as detailed in section 4.37 .

4.29. We have used the same underlying claim delay distribution in our current and previous work. In particular, in our initial application of the methodology, we have made the simplifying assumption that the same claim delay distribution can be applied to all subsets of the data, although we do not consider this to be realistic. This is the main area where further work is required.

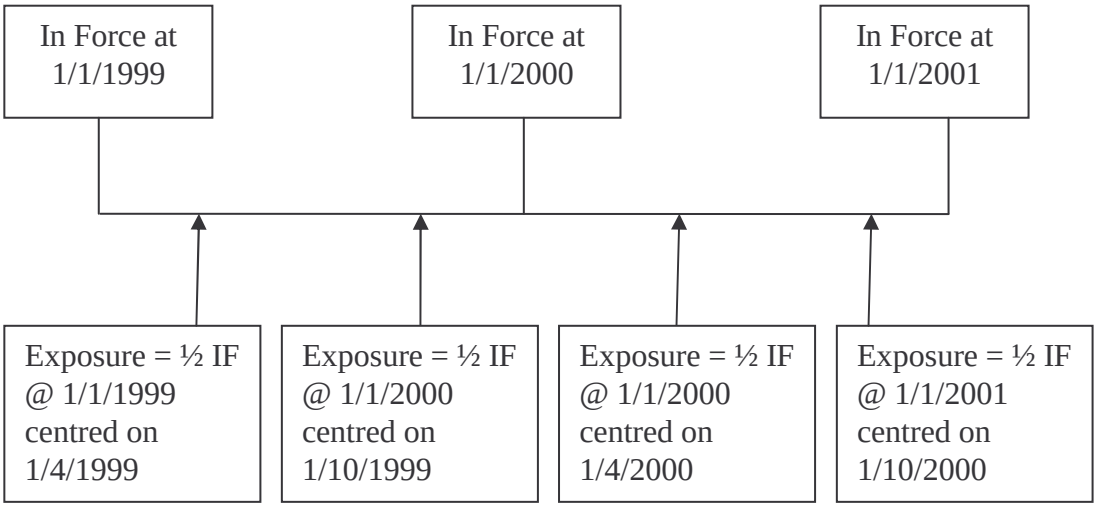
4.30. In applying claim delays to move from expected diagnosed claims to expected settled claims, we also need to give consideration to the age, duration and calendar year at the date of settlement. A sophisticated approach to this would be to consider the exposure on a daily basis, and thereby to derive expected diagnosed claims per day. These would then be projected forward to settlement on a daily basis, allowing for changes in age and duration based on the actual date of birth and date of commencement (or estimated dates, where the claims arise from synthetic exposure).

4.31. For the purposes of our initial application of the revised methodology, we have adopted a simpler approach. We have notionally sub-divided the exposure generated by the start in force and the end in force into two segments: consider a single policy in force at age X and duration T at 1/1/2000. If it remains in force throughout 2000, it will then feature in the in force data at 1/1/2001 at age X+1 and duration T+1. Ordinarily using a census method would generate exposure from that policy in 2000 of ½-year at (X,T) and ½-year at (X+1,T+1), both of which would be assumed to relate to mid-2000. For this policy, our calculation of the amount of exposure is unchanged, but the ½-year at (X,T) is assumed to relate to 1st April 2000 and the ½-year at (X+1,T+1) is assumed to relate to 1st October 2000. The difference between these approaches is illustrated below:

"Ordinary" treatment of exposure using a census method:



Treatment of exposure in our initial application:



4.32. We believe that this "split" approach is a better approximation to the true underlying situation for this initial application than the conventional census approach, where the exposure at a given age and duration is centred on the middle of the year, given that the growth in critical illness business during the (extended) period of exposure will have distorted the exposure by duration. This is probably best illustrated by means of a simplistic example:

- Suppose an office launched a critical illness product on 1st January 1997 and sold 1000 policies in the middle of every month thereafter and that there are no lapses.
- The in force data at 1/1/1998 will consist of 12,000 policies with curtate duration 0, sold during 1997.
- Ordinarily a census approach would attribute 6,000 life-years' exposure to this business in 1997 centred on 1st July, whereas our approach would attribute the same amount of exposure, but to 1st October.
- An exact approach would also attribute 6,000 life-years' exposure but to September.

The growth during the year is better reflected by our approach in this example, however for a rigourous application we clearly need to progress to an exact method of calculating exposure. Note that the effects of growth do not simply affect the first year, nor just duration 0 experience, but recur throughout the period, as the duration 1 exposure will be similarly weighted towards the end of 1998, and likewise the duration 2 exposure in 1999, etc.

4.33. This assumption is likely to have a material effect on the experience by duration and we do not consider that the results produced within our initial application by duration will necessarily be reliable and that further work is required to more accurately calculate exposure. This is considered further in section 8.

4.34. We now consider how the application of claim delays is affected by our unconventional treatment of exposure. Within the CMI Critical Illness investigation we have customarily grouped results by age nearest and curtate duration. In our initial application, exposure at age X nearest and curtate duration T in calendar year Y consists of:

½ $IF_{X,T,Y}$ centred on 1st April with an assumed age of X+¼ and an assumed duration of T+¾; and

½ $IF_{X,T,Y+1}$ centred on 1st October with an assumed age of X-¼ and an assumed duration of T+¼.

4.35. The impact of these assumptions is that of the expected diagnosed claims at age X nearest and curtate duration T in calendar year Y generated from the start in force:

- Claims settled within 3 months are assumed to be settled at ages between X+¼ and X+½, i.e. at age nearest X; at durations between T+¾ and T+1, i.e. at curtate duration T; and still within the current calendar year, Y. Hence they are estimated to be settled at X,T,Y.

Similarly:

- Claims settled from 3 to 9 months are estimated to be settled at X+1,T+1,Y;
- Claims settled from 9 to 15 months are estimated to be settled at X+1,T+1,Y+1;
- Claims settled from 15 to 21 months are estimated to be settled at X+2,T+2,Y+1;

etc

4.36. The expected diagnosed claims generated from the end in force are spread as follows:

- Claims settled within 3 months are assumed to be settled at X,T,Y
- Claims settled from 3 to 9 months are assumed to be settled at X,T,Y+1
- Claims settled from 9 to 15 months are assumed to be settled at X+1,T+1,Y+1
- Claims settled from 15 to 21 months are assumed to be settled at X+1,T+1,Y+2, etc

4.37. The figures used for the proportion of claims settled at the required number of months after diagnosis from the distribution illustrated in Working Paper 14 are:

Month	3	9	15	21	27	33	39	45	51	57	63
Cumulative Percentage Settled	39.4	71.2	83.5	86.8	89.8	92.4	94.4	95.7	96.7	97.5	97.8

The rates are then linearly interpolated to 100% at 69 months, after which we assume no further claims will be settled.

4.38. As noted in section 3.14 we have not undertaken the additional step of estimating claim rates at diagnosis in our initial application.

5. RESULTS OF THE INITIAL APPLICATION

5.1. In this section, we set out the results of using this methodology on the all causes experience on a lives basis for full acceleration business in 1999-2002. Results were sent to member offices in May 2005, at the same time as Working Paper 14 was published, but have not previously been made more widely available.

5.2. As previously noted, these critical illness results compared actual settled claims with expected diagnosed claims. Working Paper 14 provided a table of indicative grossing-up factors that varied by growth in expected claims, to attempt to correct the under-statement arising from the mis-match between exposure and claims. For the 1999-2002 'All Office' experience we estimated an overall grossing-up factor of 15%.

5.3. Results for our initial application of the revised methodology are shown in Appendix B for males and Appendix C for females, as described in 5.4 below. Comments on some of the features of the results are contained in the remaining paragraphs of this section.

5.4. The results in Appendices B and C are set out in 5 columns, as follows:

- Actual Settled Claims (ASC). These are the claims that offices advised to the CMI as settled during 1999-2002 (or the part of that period for which they contributed data). Although in total these are identical to the number of claims included in the results released in May 2005, their categorisation by age and duration differs. The released results used age and duration at date of diagnosis (or at our estimate of that date) whereas here they are categorised by age and duration at date of settlement (or at mid-year, if the CMI was only advised of year of settlement).
- Expected Diagnosed Claims (EDC). These are the claims that we expect to be diagnosed during 1999-2002 using CIBT93 to calculate the expected. Note that these numbers are not identical to those in the released results, as we have now used a slightly different measure of exposure, but without altering the claim rates from CIBT93.
- 100 x Actual Settled Claims / Expected Diagnosed Claims. This is effectively the comparison that we have provided in the released results, except for the revisions to age and duration and to exposure noted above.
- Expected Settled Claims (ESC). These are the claims that we expect to be settled during 1999-2002 using our revised methodology to calculate the historical exposure, CIBT93 to calculate the expected diagnosed claims, and the Working Paper 14 claim delay distribution to transform the claims from diagnosed to settled.
- 100 x Actual Settled Claims / Expected Settled Claims. This is the comparison generated by our revised methodology.

5.5. As shorthand, we refer to the values of 100 x Actual Settled Claims / Expected Settled Claims as "realistic" results in the following paragraphs, and to values of 100 x Actual Settled Claims / Expected Diagnosed Claims as "raw" results, noting that these differ from the results that were sent to member offices.

5.6. The first point to note is that the realistic results significantly exceed the raw results. This is to be expected. The overall level of increase is around 16%, which is very close to the grossing-up factor of 15% that we have previously indicated. This small difference is discussed further in section 7.7 but note that these two figures are not strictly comparable. We therefore believe that, overall, the revised methodology produces a realistic representation of true experience

5.7. The percentage differential between the realistic results and the raw results is very similar between males and females and between non-smokers and smokers. This is unsurprising given that:

- As noted in section 2.14, we believe that the differentials will largely reflect growth in expected claims and this will be closely correlated between the four categories; and
- The differentials will also depend on the claim delay distribution and, in our initial application of the methodology, we have not differentiated claim delays by gender or by smoker status. We believe that cause of claim is a significant determinant of claim delays and, of course, cause of claim differs by gender and, probably, by smoker status.

5.8. There is however considerable variation by duration. Indeed probably the most striking feature of these results is that the raw results appear to considerably under-state experience at duration 0. This is true to a much lesser extent at durations 1 and 2, and the differences between the raw results and the realistic results at durations 3, 4 and 5+ are relatively small.

5.9. It is important to note that this is again based on the assumption of a single delay distribution. This may not hold true and, in particular, one might intuitively expect that claims diagnosed at very short durations will be scrutinised with particular rigour, and hence delays to settlement will be longer. This would have the effect of reducing the realistic results at short durations (and increasing the results at other durations). Furthermore it is based on a census method of calculating exposure, as described in section 4.33, which will distort the results by duration and especially the results at duration 0 compared to an exact calculation.

5.10. It should also be noted that the apparent lack of selection (either positive or negative) in the realistic results says nothing directly about the shape of the select effect in the underlying claim rates. Claim diagnosis rates at duration 0 impact on expected settled claims at other durations too, and we have not yet undertaken the further step of deriving claim rates outlined in section 3.14.

5.11. The percentage differentials by age show considerably less variation.

6. SENSITIVITIES IN THE RESULTS

6.1. In this section we consider tests of the sensitivity of these results to some of the assumptions considered in sections 3 and 4. A limited amount of sensitivity testing has been undertaken to date and further testing is clearly required before we would adopt the revised methodology, however we did not wish to delay giving the Profession the opportunity to review and comment on the revised methodology.

6.2. A number of the assumptions covered in section 4 relate solely to our initial application of the revised methodology. These will be reviewed, and revised as appropriate, before we attempt to use the methodology to produce reliable results. Hence our focus in this section is on illustrating the sensitivity of the results to the key assumptions in section 3.

6.3. The results of the tests are shown in the form of graphs in Appendices D and E. In each case these graphs compare the expected settled claims in 1999-2002 using the revised assumption to the equivalent number using the original assumption, by age band and duration. These tests have been conducted on just the male non-smoker data.

Off rate

6.4. The first assumption we consider is the off rate which we introduced in sections 3.5 to 3.7. This is required in order to produce the synthetic in force data that we need to generate exposure in the years before the investigation period because we do not have access to actual data in those years.

6.5. We do not expect variations in the off rate to give rise to significant differences in the results, as explained in sections 4.11 to 4.13, but we recognise that this may not be immediately obvious.

6.6. As noted in section 4.13, our current best estimate for this assumption is 9% p.a. and we have not varied this for any subsets of the portfolio, although the methodology is not reliant on a single assumption. If further analysis of the data indicates that smokers have higher off rates than non-smokers, for example, then differential assumptions can be used for these categories of business as we are currently rolling back the known in force and estimating the synthetic in force separately for smokers and non-smokers. Similarly differences between males and females or by age and duration can easily be accommodated within the revised methodology.

6.7. Initially, we have tested the sensitivity of the results to different assumptions that are still applied across-the-board. The results of using assumptions of 5% p.a. and 20% p.a. are illustrated in Appendix D.

6.8. A lower off rate means less synthetic exposure, fewer expected diagnosed claims and hence fewer expected settled claims. Reducing the off rate from 9% to 5% reduces the expected settled claims in 1999-2002 by less than 0.5% for all ages and all durations combined. The reduction rises with duration and is greatest (around 1%) for durations 5+. There is no reduction at all for duration 0, because of our assumption within this initial application that the 1/1/1998 in force only generates claim diagnoses at 1/4/1998 with an exact duration of 0.75. Allowing for delays in settlement therefore means that any claim settled in 1999 (or later) is necessarily at curtate duration 1 (or higher) at the date of settlement. Using a more accurate method

of calculating exposure would mean that the duration 0 results would differ when the off rate is varied, but only to a marginal extent. We expect the general conclusion that the impact of a lower off rate has a greater effect the longer the duration will still hold true.

6.9. For all durations combined, the impact of reducing the off rate is lowest at younger ages and increases with age, however there is little obvious variation by age at specific durations.

6.10. Increasing the off rate from 9% to 20% increases the expected settled claims by 1.6% for all ages and all durations combined. The increase rises with duration and is greatest (around 3%) for durations 5+. Again, there is no change at all for duration 0, because of our other assumptions. Likewise, for all durations combined, the impact of increasing the off rate is lowest at younger ages and increases with age.

6.11. These tests demonstrate that the results are relatively insensitive to a change in the overall level of off rate assumed. However, it is also important to consider the impact of variations in the shape of underlying off rates. In this paper we consider just one variation, where the off rate varies by duration, as follows: 15% at duration 0, 12.5% at duration 1, 10% at duration 2, 7.5% at duration 3, 6% at duration 4 and 5% at durations 5 and over.

6.12. This is broadly consistent during 1999-2002 with the uniform rate of 9% used above. The impact of the amended structure of off rates is to imply higher overall off rates during the years before 1999, as increasing weight is given to the shorter durations in the early days of the critical illness product.

6.13. Overall, the expected settled claims increase by just 0.4% as a result of the substantial re-shaping of the off rate by duration. The impact varies by duration as follows:

- At duration 0, the expected settled claims remain unchanged due to the other assumptions in this initial application of the revised methodology (although even with a more accurate calculation, we expect little sensitivity here);
- At durations 5+, the expected settled claims reduce. This arises because for these claims, the reductions in the off rate at longer durations have greater impact than the increases at shorter durations; and
- For durations 1 to 4, the balance between the increase in off rates at short durations and the reduction at longer durations in this sensitivity test serves to increase the expected settled claims with the greatest effect at duration 3.

6.14. By age, the effect of re-shaping the off rate is to increase the expected settled claims most at younger ages, reflecting the weight of business to shorter durations.

6.15. We believe that the results of these tests demonstrate the lack of sensitivity in the expected settled claims to the off rate and that this sensitivity is greatly outweighed by the beneficial aspects of the methodology in projecting the shape and level of the exposure in prior years.

Claim delay distribution

6.16. The other sensitivity considered here is the variation in results arising from the use of different claim delay distributions. We consider this a key assumption within the

revised methodology and, unlike the sensitivity to the off rates discussed above, we expect the results of our revised methodology to show considerable variation to the assumed distribution. Indeed we anticipate that any methodology will require assumptions of the claim delay distribution, either explicitly (as with a chain ladder or this revised methodology) or implicitly (as in grossing-up factors).

6.17. Two alternative claim delay distributions are illustrated graphically in Appendix E, which also contains graphs illustrating the results of varying the distribution. The figures used for the proportion of claims settled at various months after diagnosis for these alternative distributions, analogous to those set out in section 4.37, are:

	Month	3	9	15	21	27	33	39	45	51	57	63
Cumulative Percentage Settled	Base	39.4	71.2	83.5	86.8	89.8	92.4	94.4	95.7	96.7	97.5	97.8
	Shorter	54.3	80.4	90.5	93.5	96.6	99.3	100.0				
	Longer	24.4	62.0	76.5	80.4	83.9	87.0	89.4	90.9	92.1	93.1	93.6

As noted previously, the rates for the base assumption are then linearly interpolated to 100% at 69 months, after which we assume no further claims will be settled. For simplicity, the rates for the "longer delays" scenarios have also been linearly interpolated to 100% at 69 months

6.18. As can be seen from Appendix E, the effect of varying the claim delay distribution has a substantial impact on the expected settled claims at duration 0. This is because of an asymmetry by duration:

- If claim delays are shorter, more of the expected diagnosed claims at duration 0 are expected to be settled with the curtate duration at settlement equal to zero.
- At other durations the effect is less significant. For example at duration 2, the shorter delay distribution means that fewer claims diagnosed at durations 0 and 1 are expected to be settled at duration 2, but this is more than offset by an increase in claims diagnosed at duration 2 not being delayed to durations 3+ at settlement.
- At durations 5+, the effect of the shorter claim delay distribution is to reduce the expected settled claims as this causes more claims to be settled at shorter durations. As we are not separating this category further in this analysis, no additional claims arise at durations 5+ from using the shorter delays.

6.19. The net effect of using the shorter claim delay distribution is to increase the overall expected settled claims by 2.6%, as the number of settled claims brought back, from 2003 or later, into the investigation period exceeds the number dropping out from the start of the investigation period.

6.20. Unsurprisingly, using the longer claim delay distribution has the opposite effect, with the expected settled claims at duration 0 reducing substantially and considerably less variation occurring at other durations. The net effect of using this longer claim delay distribution is to reduce the overall expected settled claims by around 9% as many more claims would be delayed in settlement until 2003 or later (particularly for durations 5+ at settlement).

7. USING THE REVISED METHODOLOGY

7.1. As noted in section 1.6, the methodology has been developed for pragmatic reasons and it is unlikely that it would have been considered from a theoretical perspective. Nor would we suggest the methodology is necessarily appropriate in other circumstances, for example where data is available to enable more direct calculation of actual rates of diagnosed claims. In this section we explain the rationale behind the development of this methodology, compare its outputs with those from our previous methodology and consider the future use and development of the methodology

7.2. We believe that the approach we propose has merit where dates of diagnosis are unknown for a significant proportion of claims, and hence there is uncertainty over which of the actual settled claims in a period are actual diagnosed claims in that same period and which relate to earlier periods. This is necessary information for a chain ladder or similar method to be used. The approach has also been developed with the following data characteristics in mind:

- Business volumes have increased, so that the synthetic data in earlier years has less weight compared to the known data in later years; and
- The claim delay distribution is such that most claims are settled within, say, 2 years so that the expected claims from the earlier years have substantially less weight than those from the later years.

It is these characteristics that mean that the results in 1999-2002 are relatively insensitive to the off rates assumed.

7.3. The key concern we had with regard to our previous methodology of grossing-up factors was the need to estimate growth in expected claims. Whilst some data exists on new business volumes, for example, that would allow us to estimate the growth in exposure, the growth in expected claims is also affected by the maturing of business, by age and duration.

7.4. Within the revised methodology, we make use of the fact that we know age and duration for the initial in force, and therefore can use this in rolling back the in force data. As illustrated in the diagrammatic representation in Appendix A, much of the prior year exposure is known, rather than synthetic, and this applies particularly to the exposure in the later prior years, which is most relevant to the settled claims in 1999-2002.

7.5. The CI Committee has previously focussed on the date of diagnosis of a claim, but at a very early stage took a pragmatic decision to collect claims according to the year of settlement. We still believe that the date of diagnosis is the most appropriate date to use when measuring experience. However as illustrated by the table below, reproduced from Working Paper 14, this focus is difficult to apply in relation to the 1999-2002 data, where we are missing dates of diagnosis for such a high proportion of claims:

Number and percentage of total claim records containing each date of claim. All 1999-2002 claims.

Date submitted by office	Number of claims	% of claims
Diagnosis	6,649	56%
Notification	9,755	83%
Admission	3,907	33%
Settlement	10,394	88%
Total	11,803	100%

7.6. In addition, there is a lack of clarity around the definition of the "date of diagnosis". We have sought to standardise the reporting of dates of claim with the Health Claims Forum, which published guidance during 2006 in this area. We are confident that this will increase consistency of practice in this area for claims settled from 2007, but it clearly has no impact on earlier years.

Further consideration of grossing-up factors

7.7. Whilst we do not anticipate making use of grossing-up factors in future, it is obviously important to compare the results of the revised methodology with those reported previously.

7.8. As noted in section 2.10, grossing-up factors, as described in Working Paper 14, are defined as the estimated ratio of actual diagnosed claims to actual settled claims. The results released to offices have been expressed as actual settled claims divided by expected diagnosed claims, thus multiplying the results by a grossing-up factor yields estimates of actual diagnosed claims divided by expected diagnosed claims in the period in question.

7.9. In the revised methodology we first estimate the exposure over an extended period, and then calculate expected diagnosed claims for that period before applying a claim delay distribution to arrive at the expected settled claims in 1999-2002. In section 5.6 we noted that the ratio of the realistic results to the raw results is not strictly comparable to the grossing-up factor we have previously used. This is because the comparisons in section 5 are based on experience of claims settled in 1999-2002 whereas, as noted above, grossing-up factors seek to adjust raw results to experience of claims diagnosed in 1999-2002. Hence the revised methodology and grossing-up factors relate to slightly different time periods, but we would nevertheless expect the difference between these to be small. (Indeed, there would be no difference if the underlying claim rates and delay distribution are unchanged throughout).

7.10. We therefore consider that the results of our initial application corroborate the overall grossing-up factor that was estimated at 15% for the 1999-2002 results. We do not however consider this grossing-up factor to be definitive, as:

- We intend to develop a more accurate calculation of exposure and the subsequent timing of expected diagnosed claims;
- We have used the claim delay distribution from Working Paper 14 to illustrate the revised methodology but anticipate this will change following further analysis.

7.11. In addition, the grossing-up factors were dependent on assumptions regarding the growth of business which have been replaced by what we consider to be a more reliable approach, involving rolling back of known in force data and estimating only the synthetic element. It is not surprising therefore that we arrive at a different answer.

Releasing annual results

7.12. As noted previously, the revised methodology provides results in the form of actual settled claims divided by expected settled claims for a given period. This differs from our original objective of actual diagnosed claims divided by expected diagnosed claims, but it brings a distinct advantage. Until such time as all claims arising in a particular year are settled, the figure for actual diagnosed claims will remain an estimate, incorporating an allowance for IBNS, even if all claims submitted to the CMI contain a date of diagnosis.

7.13. In contrast, best estimate results in the form of actual settled claims divided by expected settled claims can be calculated and released as soon as the exposure and the actual settled claims for that period are known, and may not need subsequent re-calculation to reflect settled claims in subsequent years.

7.14. As a result, using the revised methodology enables results from the CMI Critical Illness investigation to be released to offices in a more realistic manner than that in current use, without the need for delays due to further analysis. Such results would assume the underlying claim delay distribution has remained constant and this assumption clearly needs to be tested over time. However this, and other analysis, can be undertaken without holding up the release of annual results to offices.

Dates of diagnosis

7.15. Whilst we believe that our revised methodology makes good use of the data available, there is one aspect of the data which is not fully utilised, namely the (known) dates of diagnosis. These are used to estimate the claim delay distribution, but not otherwise.

7.16. We would prefer greater utilisation of known dates of diagnosis, however this is not straightforward given that we do not believe it is reasonable to assume the unknown dates are distributed consistently with the known dates, as different growth rates are likely to underlie the different categories of claims.

7.17. One possible approach to making greater use of the known dates of diagnosis is to reverse the part of the methodology described in section 3.11. Instead of calculating the expected diagnosed claims and using the claim delay distribution to decide whether or not the claim is expected to be settled during the investigation period, an alternative approach would be to:

- Allocate the claims with a known date of diagnosis to their correct year of diagnosis;
- For the claims where the date of diagnosis is unknown, use the underlying claim delay distribution, weighted by the estimated historical exposure, to allocate the claim to the various years of diagnosis;
- Ignore claims with an estimated or actual diagnosis date before the investigation period; and

- Apply a conventional IBNS adjustment to allow for the claims diagnosed during the investigation period that have yet to be settled.

7.18. We have yet to investigate this approach thoroughly. It would lead us to an estimate of actual diagnosed claims divided by expected diagnosed claims for the investigation period, as we desire. It would however sacrifice the definitive aspect of annual results noted above, in that the IBNS allowance would be refined as further information becomes available over time. We may therefore conclude that these approaches should be considered in parallel. We welcome comments on this alternative approach and suggestions for other adaptations to the methodology to utilise the known dates of diagnosis more fully.

8. FURTHER WORK

8.1. We would welcome feedback on the revised methodology described in this paper. Indeed we have deliberately produced this paper at this intermediate stage to encourage input which may influence the direction of the further work now planned.

8.2. We believe that the development outlined in this paper is a significant and positive step forward in separating the two key assumptions implicit in grossing-up factors and providing a more accurate way of estimating the growth in expected claims, especially for subsets of the data. It has however highlighted the need to undertake further analysis of claim delays to which the final results will prove sensitive. In the initial application described above we have simply used the rates previously derived in Working Paper 14 and applied these throughout. This further analysis is our top priority.

8.3. Initial work using GLM techniques has indicated that the key factors in determining claim delays are cause of claim and office, but we do not regard our initial analysis as conclusive in terms of demonstrating that duration or benefit amounts, for example, are not also significant factors. We therefore consider it appropriate to complete further analysis of claim delays before we seek to apply the methodology further or more accurately.

8.4. We have begun investigating ways to fit a parametric model of claim delays and we believe such a model would provide substantial benefits. In particular the use of a parametric model may enable us to investigate claim delays where data volumes are too low to be used reliably in the methodologies developed to date. In addition, we expect to have a similar number of settled claims in 2003-2005 as for 1999-2002, with a steadily increasing proportion of claims with both diagnosis date and settlement date recorded, so the volume of data available to analyse has increased significantly.

8.5. It is also worth noting that the need for reliable claim delay distributions is not peculiar to our methodology and was implicit within grossing-up factors. It would also be required by a chain ladder-type approach.

8.6. The second area of development required in relation to the methodology is to calculate exposure more accurately, using an exact method based on actual date of commencement, where known. Whilst this is unlikely to affect overall claims

experience significantly, we anticipate it will have a material impact on the exposure and hence the experience at duration 0. As a result we do not consider the results of our initial application to be reliable in this respect. This necessitates a more sophisticated implementation of the methodology than was used in the initial application, which was undertaken in spreadsheets based on grouped data.

8.7. Having completed these two areas of development there are a number of areas of further work to which we are keen to apply the methodology:

- Analysis of more recent data, from 2003 to 2005. These additional years contain substantial volumes of data as well as providing an insight on more recent experience.
- Amounts experience. The analysis above has considered only experience on a lives basis.
- Our initial application considers how experience varies by age, gender, smoker status and duration. Other factors which we may analyse include sales channel, product type, benefit amount, commencement year and office.
- Analysis by cause of claim. We are particularly keen to analyse experience for the main causes of claim, which may then start to give useful insights into variations in experience over time and into the effects of selection.

8.8. Finally, we intend to use the methodology to generate claim rates, as noted in section 3.14, which could be graduated to produce a published table.

9. REQUEST FOR FEEDBACK

9.1. Feedback on this methodology is welcomed by the CMI. Please e-mail any feedback, by 30 September 2007, to ci@cmib.org.uk.

9.2. After feedback has been received and evaluated, and any necessary revisions have been made, it is our intention to apply the further development of this methodology to the experience for 2003 and 2004 and, if all data has been received, 2005. Provided we do not encounter significant hurdles in the analysis of the claim delay distribution we hope to be able to complete and report on this work before the end of 2007.

9.3. Views on the relative priorities of the various analyses outlined in section 8 — and indeed ideas for other work — are also welcomed.

REFERENCES

CMI Working Paper 14 : Methodology underlying the 1999-2002 CMI Critical Illness experience investigation (May 2005)

CMI Working Paper 18 : 1999-2002 Critical Illness experience: Feedback on Working Paper 14 and Future Work (December 2005)

Irish Critical Illness experience 1995-2000, Report of the Critical Illness Working Party presented to the Society of Actuaries in Ireland on 3 November 2003

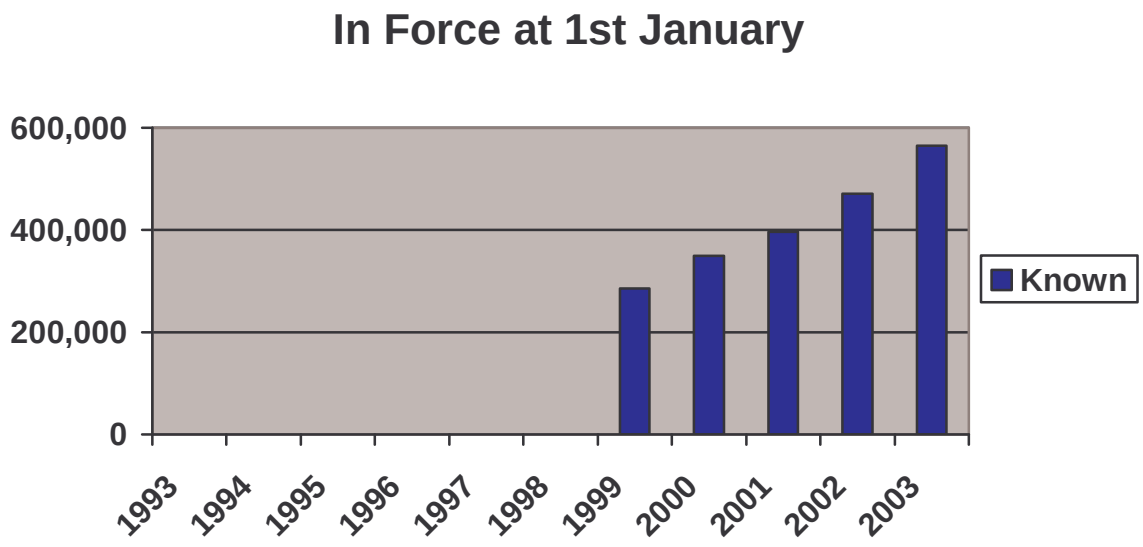
Irish Critical Illness experience 2001-2003, Report of the Critical Illness Working Party presented to the Society of Actuaries in Ireland on 29 May 2007

Critical Illness Healthcare Study Group : “A Critical Review” presented to the Staple Inn Actuarial Society on 14 March 2000.

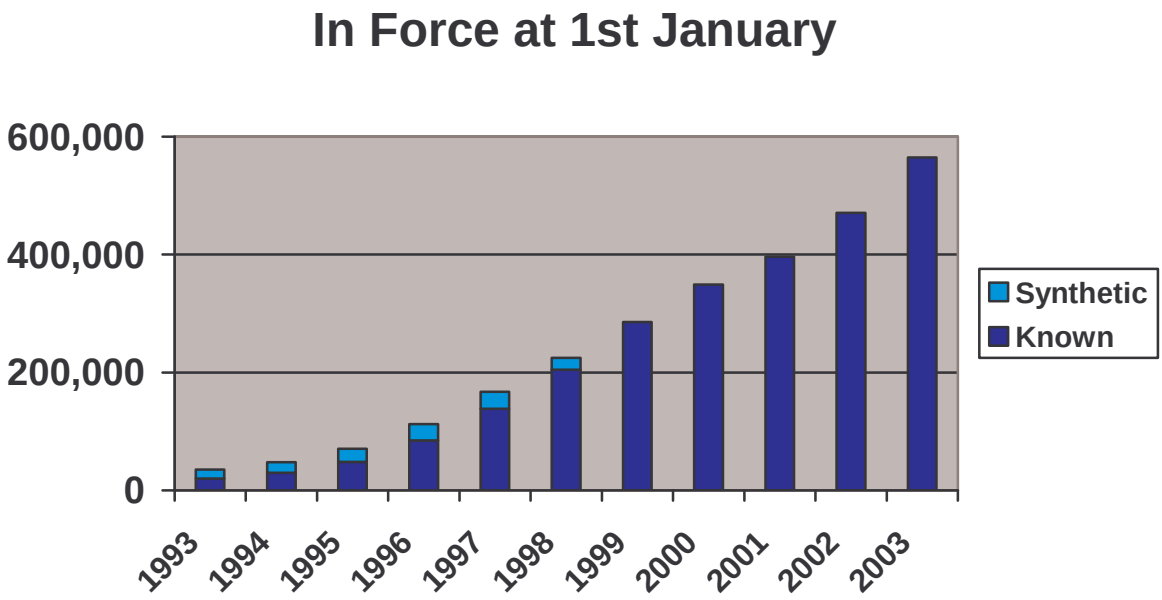
Appendix A: A pictorial representation of the revised methodology

Note: this appendix is included for illustrative purposes only. The numbers relate to male non-smokers in submission group 1

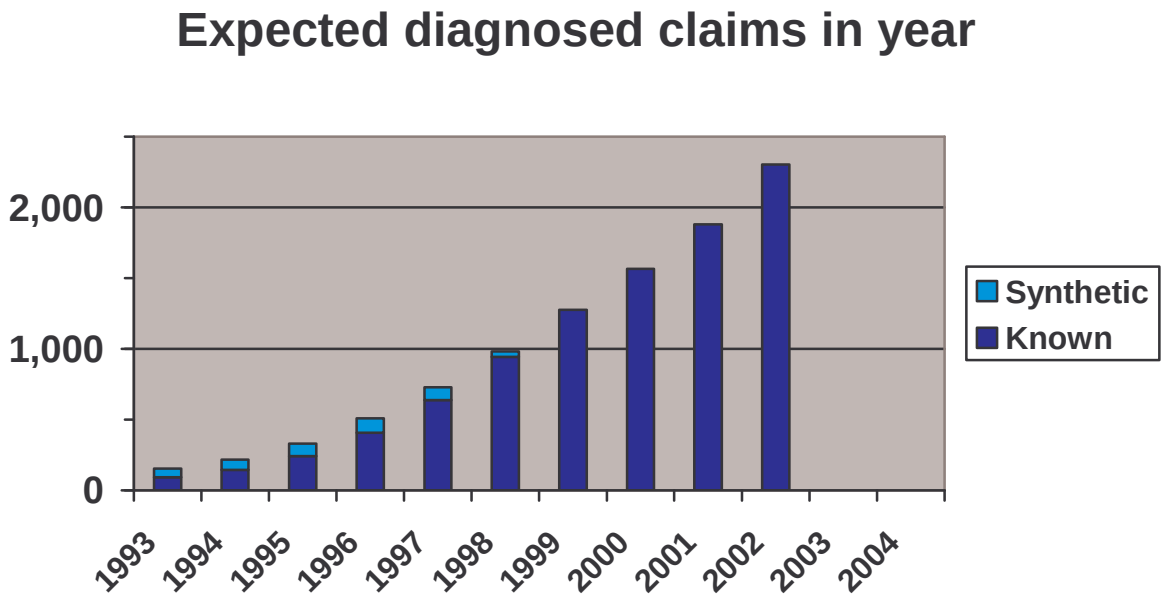
1. The approach starts with the known in force data (1/1/1999 to 1/1/2003):



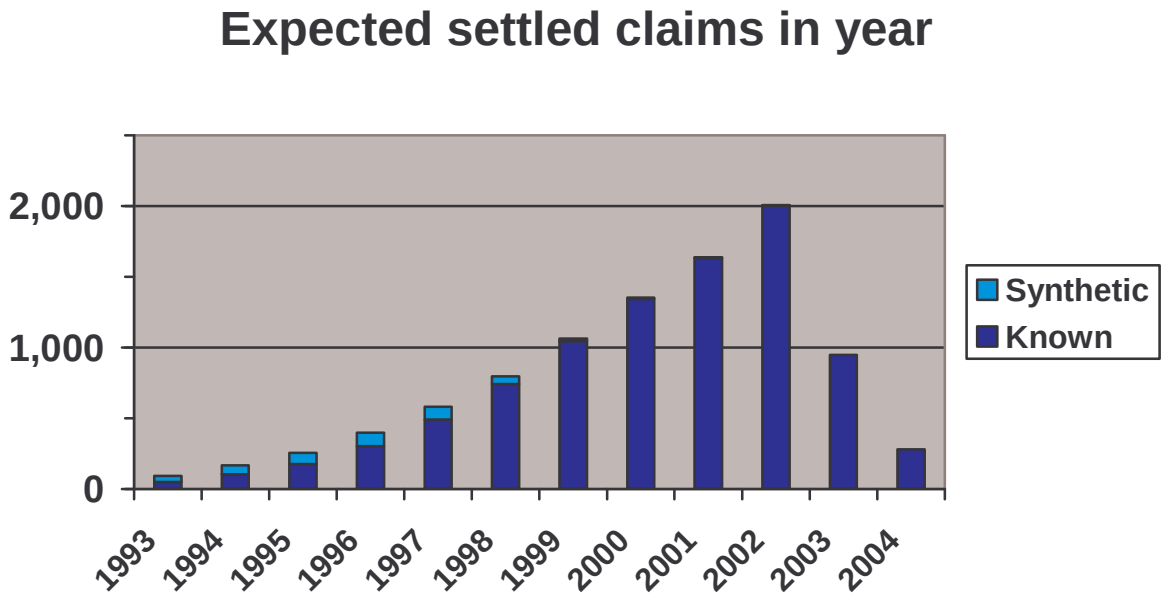
2. From the known in force, we estimate prior years' in force data:
   - part of this is a roll-back of known data (including adjusted age and duration); and
   - part is an estimate of the business that went off before data was submitted to the CMI (termed "synthetic" data)



3. From the estimated exposure in each year we calculate expected diagnosed claims in each year (at each age and duration) using CIBT93. The graph below indicates which claims are generated from "known" exposure and which from "synthetic" exposure.



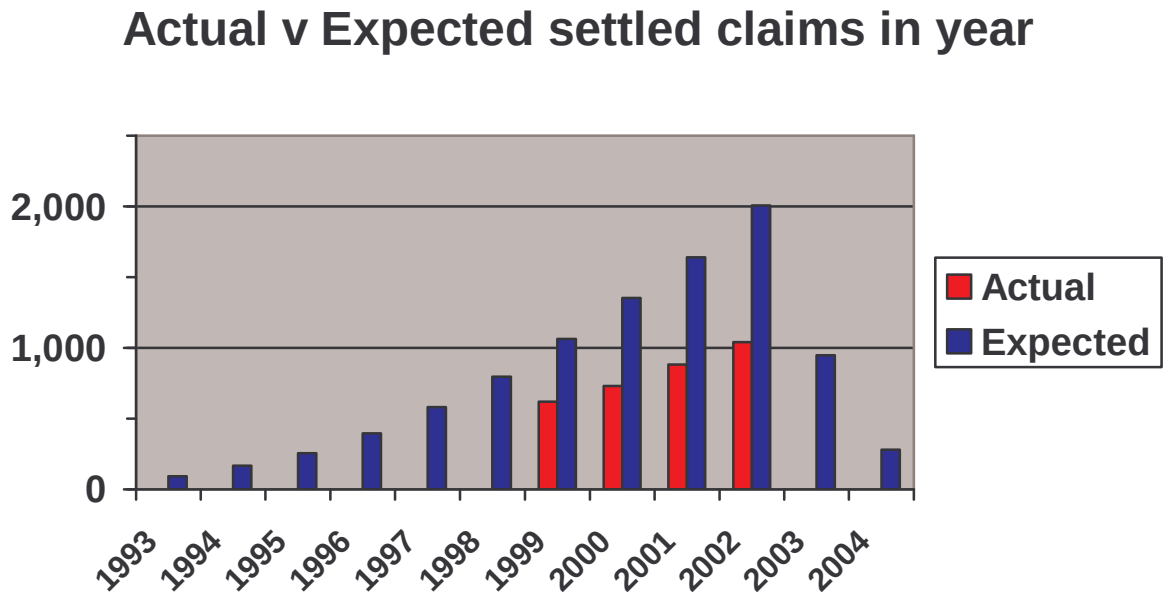
4. From the estimated diagnosed claims in each year we estimate settled claims in each year (at each adjusted age and duration) using an assumed claim delay distribution. The graph below indicates which of the settled claims are generated from known exposure and which from synthetic exposure.



Note that:

- The expected settled claims in the early years are under-stated, as we have not rolled back the in force to years before 1/1/1993, and hence do not have expected diagnosed claims for years prior to 1993.
- The expected settled claims in the years after 2002 are also under-stated, as there will be settled claims arising from diagnosed claims in 2003 and later that we have not estimated.
- Claims to be settled after 2004 are not shown in the graph.
- Claims settled in 1999-2002 that are attributable to synthetic exposure may not be clearly visible. For example of 1,063 expected settled claims in 1999, only 18 arise from synthetic exposure.

5. We then compare actual settled claims in 1999-2002 with expected settled claims in 1999-2002 to provide a measure of the experience relative to CIBT93, using the assumed claim delay distribution:



6. If desired, we can then amend our assumed claim rates in Step 3, above, iteratively in order to derive a set of diagnosed claim rates.

Appendix B: Male critical illness experience by age and duration

Full Acceleration business; Lives basis; All Causes (incl. mortality); 1999–2002; Expected based on CIBT93 and Working Paper 14 delay pattern (for Expected Settled Claims)

	Non-Smokers					Smokers				
	Actual Settled Claims (ASC)	**Expected Diagnosed Claims (EDC)**	**100 ASC/ EDC**	**Expected Settled Claims (ESC)**	**100 ASC/ ESC**	**Actual Settled Claims (ASC)**	**Expected Diagnosed Claims (EDC)**	**100 ASC/ EDC**	**Expected Settled Claims (ESC)**	**100 ASC/ ESC**
Duration 0										
< 30	96	320.8	30	173.3	55	46	105.6	44	56.9	81
31-40	171	720.0	24	385.5	44	86	226.7	38	121.1	71
41-50	132	785.0	17	418.5	32	113	217.8	52	116.3	97
51-60	96	568.2	17	304.5	32	72	132.0	55	71.0	101
61+	18	58.0	31	31.3	58	2	8.8	23	4.8	42
All	**513**	**2,452.0**	**21**	**1,313.0**	**39**	**319**	**690.9**	**46**	**370.1**	**86**
Duration 1										
< 30	109	218.0	50	190.1	57	30	67.0	45	59.6	50
31-40	228	559.0	41	477.3	48	120	166.1	72	144.3	83
41-50	203	614.5	33	524.4	39	129	165.8	78	143.9	90
51-60	164	485.5	34	415.6	39	87	112.3	77	97.7	89
61+	20	61.8	32	53.7	37	4	10.3	39	9.0	44
All	**724**	**1,938.8**	**37**	**1,661.1**	**44**	**370**	**521.5**	**71**	**454.5**	**81**
Duration 2										
< 30	65	149.0	44	142.5	46	20	43.1	46	42.0	48
31-40	195	448.0	44	415.8	47	81	125.4	65	118.8	68
41-50	179	490.5	36	455.7	39	103	127.5	81	120.7	85
51-60	163	393.4	41	369.7	44	77	90.8	85	87.1	88
61+	21	60.0	35	57.5	37	9	10.1	89	9.8	92
All	**623**	**1,541.0**	**40**	**1,441.2**	**43**	**290**	**396.9**	**73**	**378.5**	**77**
Duration 3										
< 30	68	99.4	68	100.9	67	20	27.7	72	28.7	70
31-40	143	370.2	39	360.6	40	70	98.6	71	98.2	71
41-50	182	414.6	44	403.9	45	100	104.1	96	103.1	97
51-60	146	331.4	44	327.6	45	65	76.0	85	76.5	85
61+	25	51.9	48	53.5	47	5	9.8	51	10.0	50
All	**564**	**1,267.5**	**44**	**1,246.6**	**45**	**260**	**316.2**	**82**	**316.4**	**82**
Duration 4										
< 30	39	62.6	62	65.0	60	11	16.6	66	17.6	62
31-40	137	301.1	45	297.1	46	68	76.7	89	77.3	88
41-50	122	359.2	34	352.4	35	60	85.2	70	85.4	70
51-60	125	289.7	43	288.5	43	67	64.2	104	65.5	102
61+	17	43.7	39	45.8	37	5	8.1	62	8.5	59
All	**440**	**1,056.3**	**42**	**1,048.8**	**42**	**211**	**250.8**	**84**	**254.4**	**83**
Duration 5+										
< 30	39	65.6	59	71.2	55	9	15.1	60	16.8	54
31-40	305	607.6	50	620.0	49	71	137.9	51	143.6	49
41-50	488	980.3	50	997.8	49	185	197.2	94	204.4	91
51-60	456	937.1	49	970.4	47	154	172.1	89	179.7	86
61+	109	176.8	62	196.4	55	23	23.3	99	26.5	87
All	**1,397**	**2,767.3**	**50**	**2,855.9**	**49**	**442**	**545.6**	**81**	**571.0**	**77**
All durations										
< 30	416	915.3	45	743.0	56	136	275.1	49	221.6	62
31-40	1,179	3,005.9	39	2,556.5	46	496	831.4	59	703.4	73
41-50	1,306	3,644.1	36	3,152.7	41	690	897.6	76	773.9	93
51-60	1,150	3,005.4	38	2,676.4	43	522	647.4	79	577.5	95
61+	210	452.3	46	438.2	48	48	70.3	65	68.6	75
All	**4,261**	**11,023.0**	**39**	**9,566.7**	**45**	**1,892**	**2,721.9**	**70**	**2,345.0**	**81**

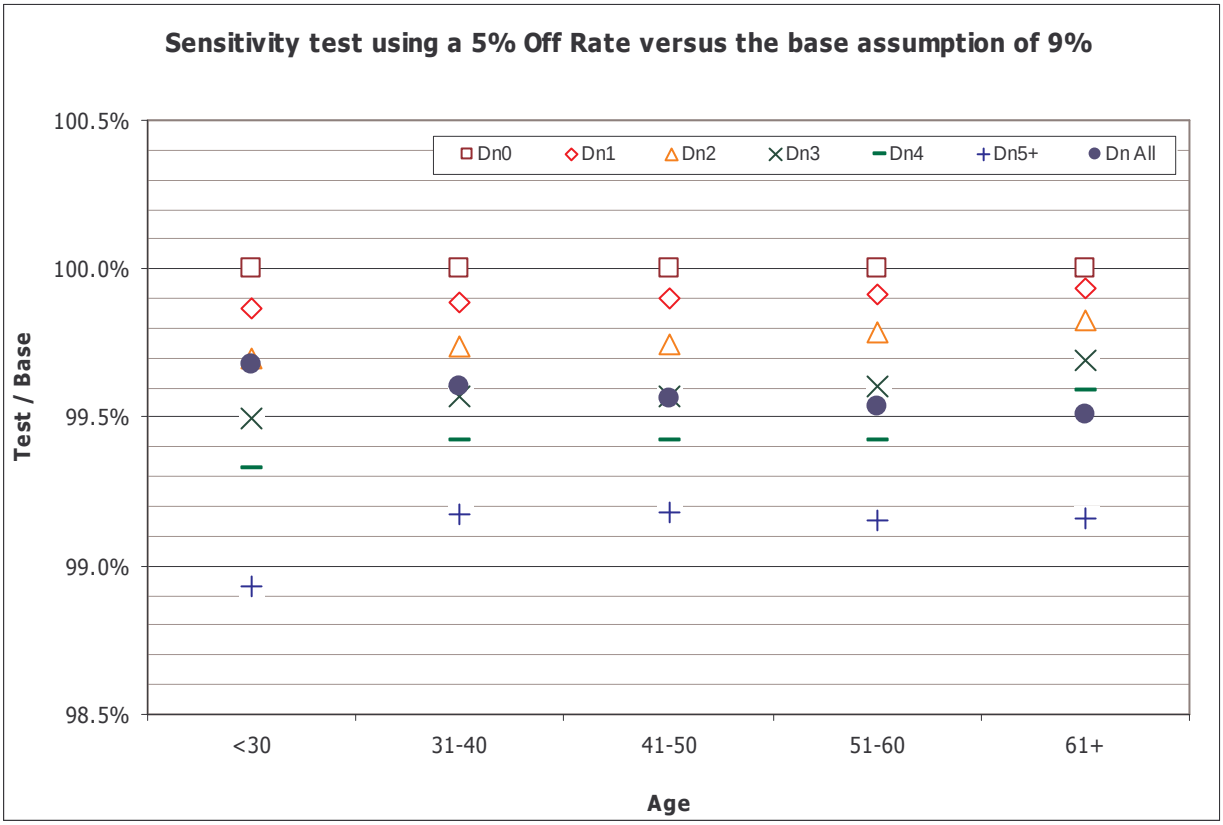
Appendix C: Female critical illness experience by age and duration

Full Acceleration business; Lives basis; All Causes (incl. mortality); 1999–2002; Expected based on CIBT93 and Working Paper 14 delay pattern (for Expected Settled Claims)

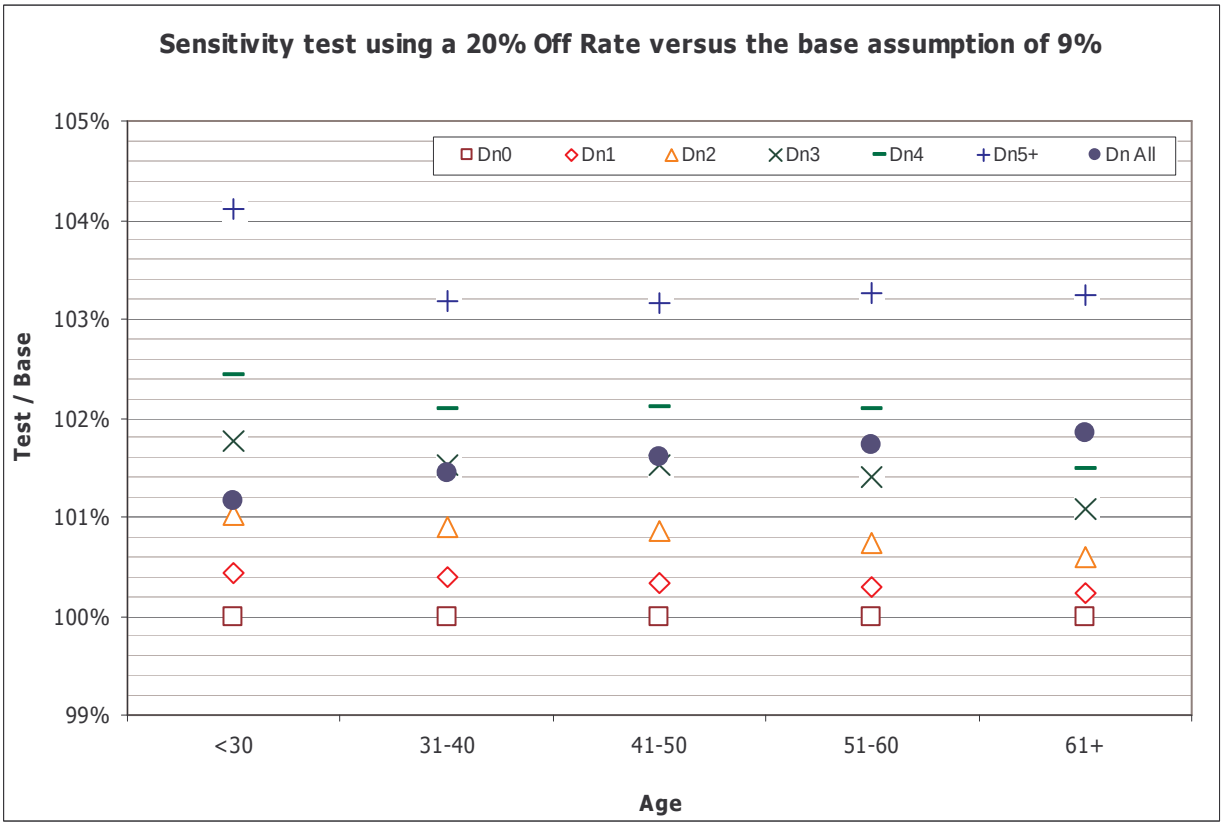
	Non-Smokers					Smokers				
	Actual Settled Claims (ASC)	**Expected Diagnosed Claims (EDC)**	**100 ASC/ EDC**	**Expected Settled Claims (ESC)**	**100 ASC/ ESC**	**Actual Settled Claims (ASC)**	**Expected Diagnosed Claims (EDC)**	**100 ASC/ EDC**	**Expected Settled Claims (ESC)**	**100 ASC/ ESC**
Duration 0										
< 30	60	273.2	22	147.4	41	20	67.7	30	36.4	55
31-40	158	568.2	28	303.6	52	41	135.7	30	72.5	57
41-50	123	503.0	24	268.2	46	33	123.6	27	66.1	50
51-60	72	261.4	28	139.5	52	23	57.7	40	31.1	74
61+	3	18.0	17	9.6	31	0	2.9	0	1.6	0
All	**416**	**1,623.9**	**26**	**868.4**	**48**	**117**	**387.6**	**30**	**207.6**	**56**
Duration 1										
< 30	81	191.6	42	164.8	49	20	45.0	44	39.3	51
31-40	205	443.1	46	378.2	54	59	100.5	59	87.0	68
41-50	182	398.3	46	340.8	53	53	95.9	55	83.0	64
51-60	97	225.1	43	194.0	50	31	50.4	62	43.9	71
61+	2	19.0	11	16.3	12	4	3.3	122	2.9	139
All	**567**	**1,277.1**	**44**	**1,094.1**	**52**	**167**	**295.0**	**57**	**256.1**	**65**
Duration 2										
< 30	71	134.5	53	126.3	56	21	30.5	69	29.1	72
31-40	192	357.8	54	331.8	58	39	76.6	51	72.3	54
41-50	180	322.3	56	299.9	60	52	73.7	71	69.6	75
51-60	95	187.4	51	176.9	54	28	41.6	67	39.9	70
61+	10	18.4	54	17.6	57	4	2.9	137	2.9	137
All	**548**	**1,020.4**	**54**	**952.6**	**58**	**144**	**225.3**	**64**	**213.8**	**67**
Duration 3										
< 30	39	93.3	42	92.5	42	9	20.7	44	20.9	43
31-40	147	299.2	49	291.1	50	31	61.0	51	60.4	51
41-50	130	275.5	47	269.0	48	39	60.2	65	59.7	65
51-60	72	159.3	45	158.7	45	29	35.7	81	35.8	81
61+	9	16.1	56	16.7	54	1	2.6	38	2.7	36
All	**397**	**843.3**	**47**	**828.0**	**48**	**109**	**180.2**	**60**	**179.5**	**61**
Duration 4										
< 30	31	61.7	50	62.3	50	8	13.1	61	13.5	59
31-40	142	247.2	57	243.5	58	21	48.9	43	49.0	43
41-50	134	238.2	56	235.1	57	42	49.6	85	49.7	84
51-60	52	138.9	37	139.6	37	16	30.5	52	31.1	51
61+	5	14.3	35	14.8	34	0	2.3	0	2.5	0
All	**364**	**700.3**	**52**	**695.3**	**52**	**87**	**144.5**	**60**	**145.8**	**60**
Duration 5+										
< 30	43	67.9	63	71.4	60	10	13.3	75	14.2	70
31-40	310	514.8	60	524.7	59	67	88.9	75	92.4	73
41-50	325	626.6	52	641.9	51	89	111.8	80	116.1	77
51-60	259	415.8	62	435.9	59	76	81.0	94	85.6	89
61+	28	58.8	48	65.0	43	13	7.7	170	8.9	147
All	**965**	**1,683.8**	**57**	**1,738.9**	**55**	**255**	**302.6**	**84**	**317.1**	**80**
All durations										
< 30	325	822.1	40	664.6	49	88	190.2	49	153.5	62
31-40	1,154	2,430.2	47	2,072.9	56	258	511.6	59	433.6	73
41-50	1,074	2,363.8	45	2,055.0	52	308	514.8	76	444.2	93
51-60	647	1,388.1	47	1,244.7	52	203	296.9	79	267.3	95
61+	57	144.6	39	140.0	41	22	21.7	65	21.4	75
All	**3,257**	**7,148.8**	**46**	**6,177.1**	**53**	**879**	**1,535.3**	**57**	**1,320.0**	**67**

Appendix D: Sensitivity tests on off rates

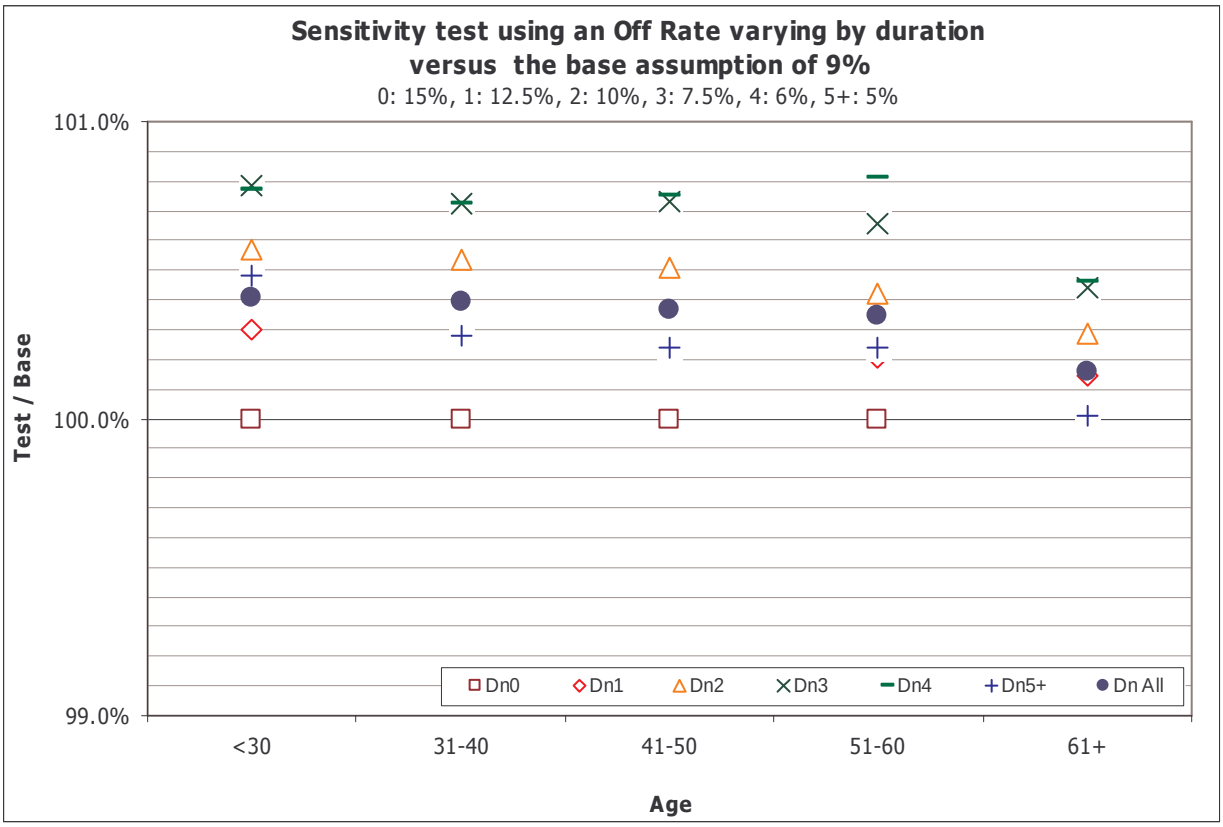
Using a uniform off rate of 5%



Using a uniform off rate of 20%



Using an off rate that varies by duration



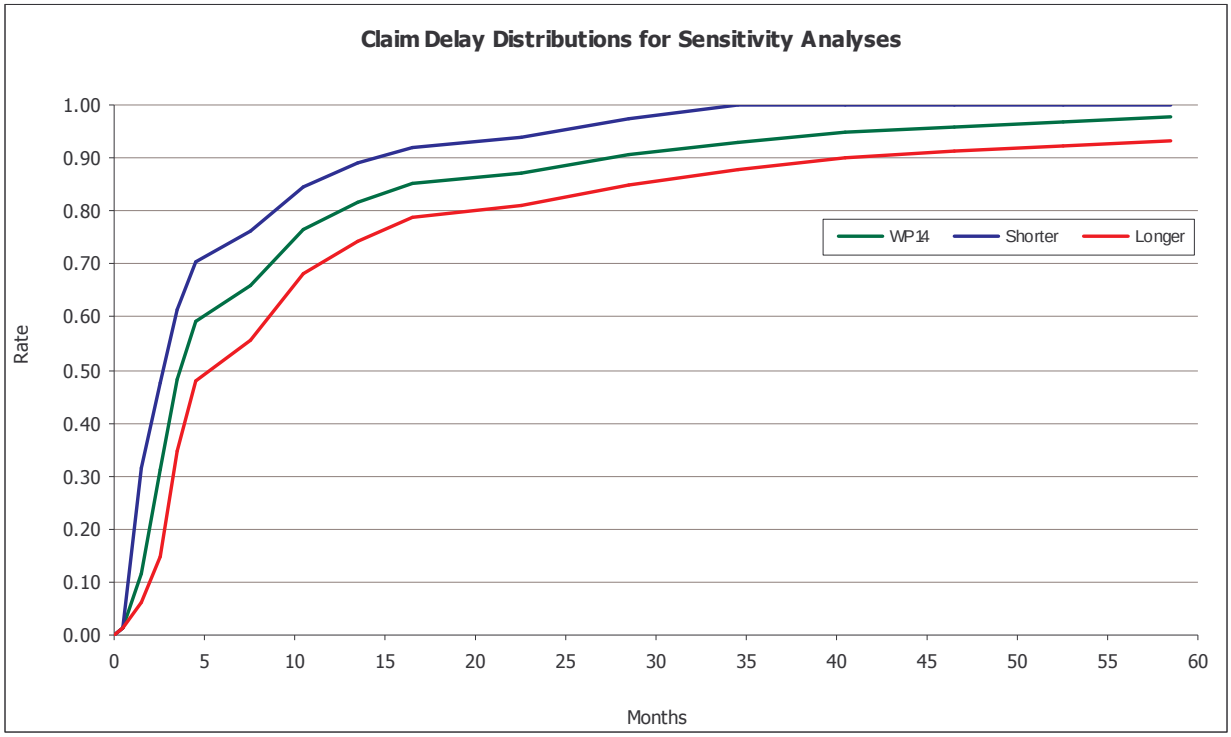
Notes:

In each case these graphs compare the expected settled claims in 1999-2002 using the revised assumption to the equivalent number using the original assumption, by age band and by duration.

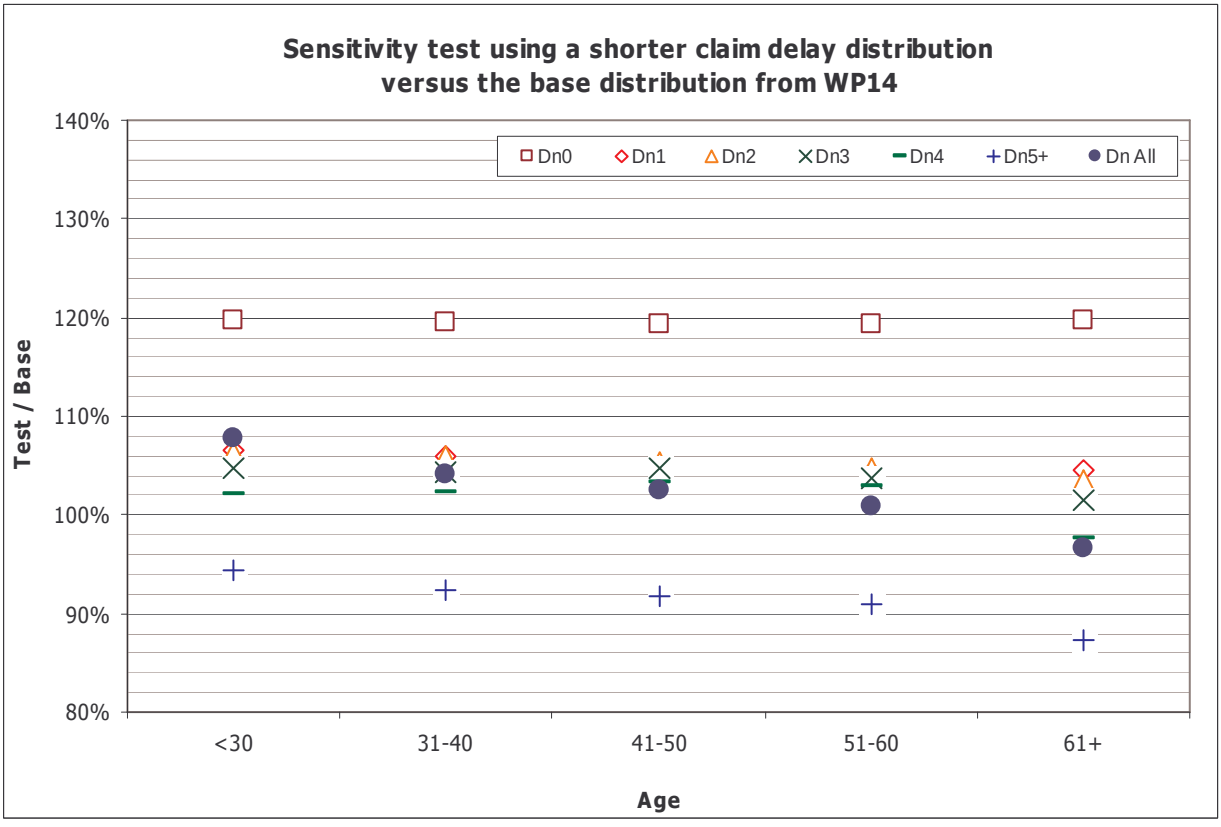
These tests have been conducted on just the male non-smoker dataset.

Appendix E: Sensitivity tests on the claim delay distribution

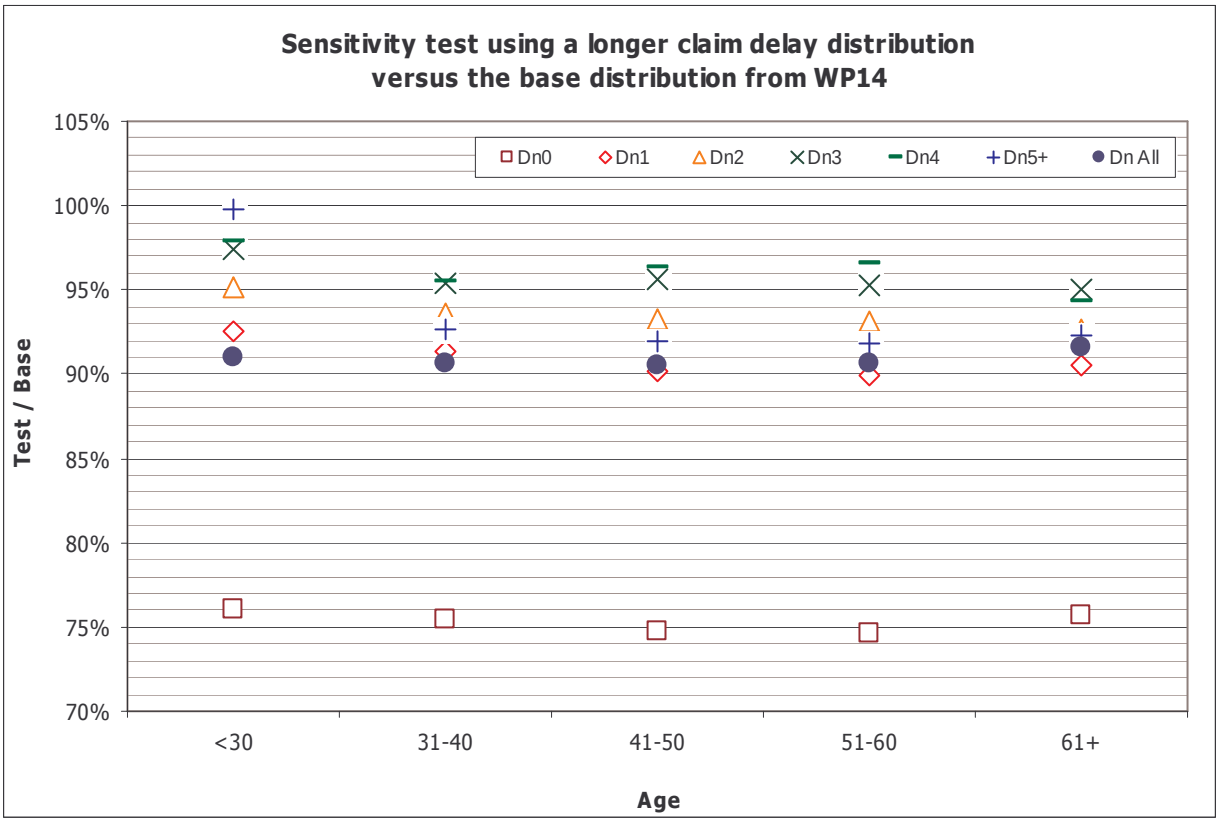
The claim delay distribution derived in Working Paper 14 has been used for the initial application of the revised methodology, except in section 6 of the paper where different distributions, illustrated in the graph below, have been used as sensitivity tests:



Using a shorter delay distribution.



Using a longer delay distribution.



Notes:

In each case these graphs compare the expected settled claims in 1999-2002 using the revised assumption to the equivalent number using the original assumption, by age band and by duration.

These tests have been conducted on just the male non-smoker dataset.