

Continuous Mortality Investigation

Critical Illness Committee

WORKING PAPER 14

**Methodology underlying the 1999-2002 CMI
Critical Illness experience investigation**

Continuous Mortality Investigation

Critical Illness Committee

WORKING PAPER 14

Methodology underlying the 1999-2002 CMI Critical Illness experience investigation

CONTENTS

1. INTRODUCTION	1
2. THE HISTORY OF THE INVESTIGATION	2
3. DATA REQUIREMENTS	3
4. DATA CHECKING	5
5. CLAIMS DATA AND ISSUES ARISING	5
5.1. <i>Dates relating to a claim</i>	5
5.2. <i>Delays between diagnosis and settlement of claims</i>	7
5.3. <i>Issues relating to the choice of “date of claim” for analysis purposes</i>	9
5.4. <i>Observed claim delay patterns</i>	10
5.5. <i>Underlying claim delay patterns</i>	10
6. ANALYSIS METHODOLOGY	13
6.1. <i>Cause of claim</i>	13
6.2. <i>Exposure calculation</i>	13
6.3. <i>Calculation of expected claims</i>	14
6.4. <i>CMI methodology for “date of claim” for the 1999-2002 quadrennium analyses</i>	15
6.5. <i>Estimation of missing dates of diagnosis</i>	16
6.6. <i>Allocation of claims to experience years for the 1999-2002 quadrennium analyses</i>	17
7. GROSSING-UP FACTORS	19
Appendix A: Impact of refining the date of diagnosis algorithm	22
Appendix B: Data requirements for each in force policy or claim record	24
Appendix C: Cause of claim groupings	25

Continuous Mortality Investigation

Critical Illness Committee

Methodology underlying the 1999-2002 CMI Critical Illness experience investigation

1. INTRODUCTION

- 1.1. This paper provides an update on the development of the CMI's analysis methodology for its Critical Illness investigation and supersedes its previous note on methodology that accompanied the 1998-1999 All Office analyses released in March 2003 and that was also used for the 2000 results released in December 2003.
- 1.2. The paper first sets out the brief history of the investigation (Section 2), then describes the CMI's data collection process (Sections 3 and 4). Key aspects of the claims data, and the consequent difficulties posed for the investigation, are covered (Section 5) before setting out the updated methodology that has been applied to the analysis of the 1999-2002 quadrennium experience (Section 6). The paper does not discuss the features of the data or the claims experience, except where they impact on the method of analysis.
- 1.3. It is important to note that the chosen methodology results in an understating of the true experience and the paper concludes by indicating the approximate grossing-up factors to correct for this in Section 7.
- 1.4. The main area of difficulty in developing a methodology for analysing claims experience has been the substantial delays observed in the claims data, firstly between the date of diagnosis of the underlying critical illness event and the date of notification to the insurer, and secondly by further significant delays between the date of notification and the dates the claim is finally admitted and settled. These delays indicate that Critical Illness business is subject to significant levels of "Incurred But Not Settled" (IBNS) claims, consisting of "Incurred But Not Reported" (IBNR) and "Reported But Not Settled" (RBNS) claims. For some claims, these delays can be measured in years rather than months or weeks.
- 1.5. These difficulties have been compounded by changes in the mix of business included in the data due to changes in the portfolios on which individual offices have submitted data as well as offices joining or leaving the investigation.
- 1.6. If claims are allocated to exposure years according to their date of diagnosis, it is clear that the final claims experience relating to an exposure year will not be known until a few years after the end of the exposure year. It is also clear that holding an exposure year open for an extended period, so that all claims with a date of diagnosis in that year that are eventually settled have been reported to the CMI, would be impractical
- 1.7. However, any investigation using a date other than diagnosis as the date of claim for analysis purposes will not reflect the true incidence of insurance liabilities. In particular, using the date of settlement will tend to underestimate overall experience where the

number of expected claims is growing (due to increasing volumes and/or ageing portfolios) and will overestimate the extent of any positive initial selection.

- 1.8. In summary, the method of analysis used by the CMI to address these issues is to:
 - allocate claims to exposure years according to their date of settlement;
 - within each exposure year, allocate claims to an age and duration according to their date of diagnosis or, where this is not provided, a date of diagnosis estimated from the other available dates of claim.
- 1.9. The CMI Critical Illness Committee believes that this provides a practical solution for a continuous investigation given the long delays that would occur in reporting results using alternative approaches. However, it should be noted that the method remains imperfect in several respects. In particular:
 - The mismatch between exposure and claims means that it will tend to underestimate the eventual experience for a particular exposure year when the insured portfolio is expanding and/or ageing, although estimates for the eventual experience for the year can be made.
 - the need to estimate many dates of diagnosis means that features of the results, especially any apparent initial selection effects, are subject to this additional estimation error risk.
 - there is a lack of precision regarding how the date of diagnosis is defined.
- 1.10. The paper discusses these issues and then concludes by indicating approximate grossing-up factors to broadly correct for the distortion to the results by exposure year. Such factors, given in Section 7, combine:
 - the removal of claims settled in the exposure year but diagnosed (or estimated to be diagnosed) in prior years; and
 - the addition of claims diagnosed in the exposure year but settled in later years or yet to be settled.

These factors have been derived using modified chain ladder techniques.

- 1.11. Feedback on this methodology is welcomed by the CMI. Please email any feedback, by 31/5/2005, to ci@cmib.org.uk.
- 1.12. After feedback has been received and evaluated, and any necessary revisions to the methodology have been made, the methodology will be formally documented in a CMI Report and the results for the quadrennium will be published to the profession. The methodology for future investigation periods will be subject to further review as data volumes increase.

2. THE HISTORY OF THE INVESTIGATION

- 2.1. The CMI initially launched a Critical Illness investigation with the aim of collecting data from 1995 on a continuous basis. Unfortunately, the volume of data submitted to that investigation did not make it feasible to produce any meaningful results. One key reason for the paucity of data was that much of the business was written by offices who were not then members of the CMI.

- 2.2. In view of the continuing importance of the product to the industry, the investigation was re-launched under the auspices of a dedicated Critical Illness Committee. The data requirements were reviewed and revised with a view to collecting data that were both more useful and allowed offices more flexibility in the method of provision. Potential contributors were contacted in 2000 and invited, initially, to contribute data in respect of 1998 and 1999. In order to maximise the volumes of data, non-member offices were also invited to contribute data in respect of these two years only. The 1998 and 1999 All Office analyses were released in March 2003 and accompanied by a note on the initial methodology used.
- 2.3. Following publication of the 1998 and 1999 analyses, non-members who contributed data were invited to join the CMI and to contribute data for 2000 and beyond and it is pleasing to note that most did so. All the offices were also asked to make every effort to provide dates of diagnosis for claims data submitted in the future in order to reduce the need for the CMI to estimate missing dates of diagnosis and hence reduce the impact of any estimation errors. The results for 2000 were issued in December 2003, using the same methodology as for previous years' analyses.
- 2.4. During the period of collecting data for the 2000-2002 experiences, further issues were identified by many offices in their own data. This led to several re-submissions of data by these offices, including re-submissions of data for the 1999 and 2000 experiences after the analyses for these years had already been published.
- 2.5. At the stage of analysing the experience for the 1999-2002 quadrennium, offices were again asked if they could provide dates of diagnosis for any claims where they had previously not done so. Some offices were able to provide dates of diagnosis for more claims and these have been used in the analyses for the quadrennium. Increased data volumes permitted additional investigation of the pattern of delays to claim settlement and this led to further development of the method of analysis, particularly in respect of estimating missing dates of diagnosis and estimating grossing-up factors to apply to the overall results. Experiences for 1999, 2000, 2001, 2002 and the 1999-2002 quadrennium are now being released to contributing offices using the revised data and methodology.

3. DATA REQUIREMENTS

- 3.1. Investigations are carried out on the following policy types:

Stand Alone - where the benefit is paid on diagnosis of critical illness
Full Acceleration - where the benefit is payable on the diagnosis of critical illness or death, whichever occurs first

- 3.2. The investigations include many types of Critical Illness cover including term, endowment and whole-of-life versions.
- 3.3. The investigations analyse the experience of these policy types by sex, smoker status, duration and broad age group. Two analyses are produced for Full Acceleration cases: one based on all claims (including deaths) and the other based only on critical illness (including TPD) claims – so akin to Stand Alone cover and excluding death claims. Each analysis is carried out on both a “lives” basis and an “amounts” basis. A summary of the analyses carried out is given in the results pack sent to contributing offices.

- 3.4. Data required for each investigation year (N) are:
- In force (i.e. on risk for benefits) on 1 January in year N.
 - Claims in year N.
 - In force on 1 January in year N + 1.
- 3.5. Data are collected on a “per policy” basis. Each in force policy, or claim, record should contain the information in Appendix B. In addition, the following requirements apply to data submitted:
- (i) Only directly written business should be included (i.e. no accepted reinsurance).
 - (ii) The investigation covers cases written on standard premium rates only. Rated cases should be excluded.
 - (iii) Multiple policies should be treated as a single policy where they arise from one underwriting process (e.g. clustered policies for tax purposes, automatic increments, etc.). If new underwriting is involved, a separate record should be submitted for the new policy element.
- 3.6. Offices are requested that claims submitted in respect of a year should be based, where possible, on date of settlement (i.e. data on a particular claim should be submitted in the year in which the claim is settled). If date of settlement is not known, date of admission or one of the other dates should be used to define whether a claim falls in the year. In any event, offices are requested to be consistent from one year to the next, such that no valid claim is either missed or double-counted.
- 3.7. The contributing offices are asked to provide four dates relating to each claim – diagnosis, notification, admission and settlement. In practice, most offices cannot supply all four of the dates requested in respect of a claim. In these circumstances, they are requested to supply the dates they can and otherwise leave fields blank. However, offices are asked to make every effort to provide the date of diagnosis as this is the key date used in the analyses.
- 3.8. Offices are asked to supply cause of claim as the critical illness event under which the claim was admitted. There is no specified coding for this field; the CMI converts the information provided to a standard categorisation as described in Section 6.1. Offices are additionally requested to specify the site of any cancer. Where cause is unknown, this should be stated.
- 3.9. Following publication of the 1998 and 1999 analyses, additional guidance was given to offices on commencement dates and the treatment of different benefit levels on flexible protection policies. Offices were also asked to exclude all claims within their free cover period from the data submitted to the CMI as these claims occur before commencement where there is no corresponding exposure.
- 3.10. Full details of the current data requirements and additional guidance are given in the Critical Illness Coding Guide version 3.2, available on request from the CMI.

4. DATA CHECKING

- 4.1. Submitted data is subject to a computerised checking process that generates error reports on three levels:
- (i) Illegal coding - field codings which are not recognised as a valid code and illogical dates of claim (such as settlement before diagnosis).
 - (ii) Suspect - field values which are considered unlikely to be valid (e.g. very high / very low sums insured, very extreme ages, etc.).
 - (iii) Warning - field values which may be invalid and should be checked (generally less extreme values than “suspects”).
- 4.2. Data sets are checked both in isolation and in comparison to other data sets (claims vs. in force). A report showing the results of the checking process is sent to the contributing office and may result in advice to the CMI on how certain records should be amended or in re-submission of the data.
- 4.3. While the checks help to clean up the data to some extent, they cannot be regarded as foolproof. For example, they cannot detect if claims records have been omitted from the data.
- 4.4. Furthermore, although offices were asked to eliminate multiple policies arising from one underwriting process, it does not appear that this has been done for some sections of the data. Indeed some offices have indicated that it is not possible for them to eliminate such policies. In such cases, CMI has asked the offices to ensure that the claims data are treated consistently with the in force data.

5. CLAIMS DATA AND ISSUES ARISING

5.1. Dates relating to a claim

- 5.1.1. The contributing offices are asked to provide four dates relating to each claim – diagnosis, notification, admission and settlement. If they are unable to provide all four, they are asked to provide as many as they can, but they must provide at least one. Although a minority of offices provide all four dates, the majority cannot and the 1999-2002 data contain various combinations of one, two and three dates.
- 5.1.2. Intuitively, one might expect the four dates to follow a set pattern – diagnosis, then notification, then admission and then settlement. This is not always the case and is thought to occur because of the lack of precision over the definition of date of diagnosis. With some claims, notification may be received before the claim criteria have been met. In such cases, notification may occur months before diagnosis, but then admission and settlement should follow relatively quickly after diagnosis. This phenomenon has been noted in particular for TPD claims where diagnosis may be defined by some offices as only occurring when the event is confirmed as both total and permanent.
- 5.1.3. On receipt of data, the CMI carry out validation checks on the dates of claim where offices provided more than one date for their claims. When the 2000-2002 data was validated, it became apparent that the data supplied by many offices was not of the required quality.

Although such checks had also been previously made to the 1998-1999 data before processing it, offices had made few corrections to their data at that time.

5.1.4. In particular, the validation indicated that some of the dates of claim were suspect as they implied that, either the claims were notified or settled on the same day as the date of diagnosis, or that claims were settled on the date of notification. As the dates of claim are central to the analyses carried out, the CMI again raised this issue with the offices. It then became clear that, for many offices, there were problems with the way the offices extracted data which also affected the data they had submitted for the 1998-1999 investigation years. When these data issues were resolved by the offices concerned, they re-submitted their data for 1999 as well as for later experience years.

5.1.5. Offices were also asked to confirm which of the dates of claim they had supplied was most reliable. Several offices agreed that some of the dates of claim provided were not reliable, and dates of claim indicated as unreliable were then deleted from the data analysed by the CMI. Offices also indicated that, where they had not been able to provide the date of diagnosis, the most reliable date provided was the date of settlement followed by the date of admission.

5.1.6. Following the data revisions, and given the issues regarding the reliability of the dates of claim, the CMI carried out the following process to maintain the integrity of the data being analysed:

- (i) Where the date of diagnosis submitted by an office was before the commencement date, the claim was excluded from the analyses.
- (ii) Where the date of diagnosis submitted by an office equalled the date of settlement, a revised date of diagnosis was estimated from the date of settlement (see Section 6.4), as offices had indicated that date of settlement was the more reliable.
- (iii) Where the date of diagnosis submitted by an office was close to the date of settlement, the claim was included in the analyses without adjustment.

5.1.7. The number and percentage of claim records in the 1999-2002 quadrennium containing each of the four dates, after the above process was carried out and dates indicated to be unreliable removed, are shown in Table 1 below.

Table 1. 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%

5.2. *Delays between diagnosis and settlement of claims*

5.2.1. The main area of difficulty in developing a methodology for analysing claims experience has been the substantial delays observed in the claims data, firstly between the date of diagnosis of the underlying critical illness event and the date of notification to the insurer, and secondly by further significant delays between the date of notification and the dates the claim is finally admitted and settled. For some claims, these delays can be measured in years rather than months or weeks.

5.2.2. The average delay between various pairs of dates, where we have them, and the volumes of data involved are shown in Table 2 below.

Table 2. Average delay between various dates of claim (in days). All 1999-2002 claims.

Events	Average number of days between events	Number of records	% of records containing both dates
Diagnosis – notification	104	5,356	45%
Diagnosis – admission	159	1,551	13%
Diagnosis – settlement	176	5,404	46%
Admission – settlement	5	3,686	31%

5.2.3. Although it may seem strange that the sum of the average times between diagnosis and admission and between admission and settlement is less than the average time between diagnosis and settlement, we must remember that we are dealing with different subsets of claims in each case and the average time between diagnosis and admission is based on a relatively small number of records. In practice, the directly observed average delay between admission and settlement is more reliable than that inferred from the other observations and is distributed within a small range.

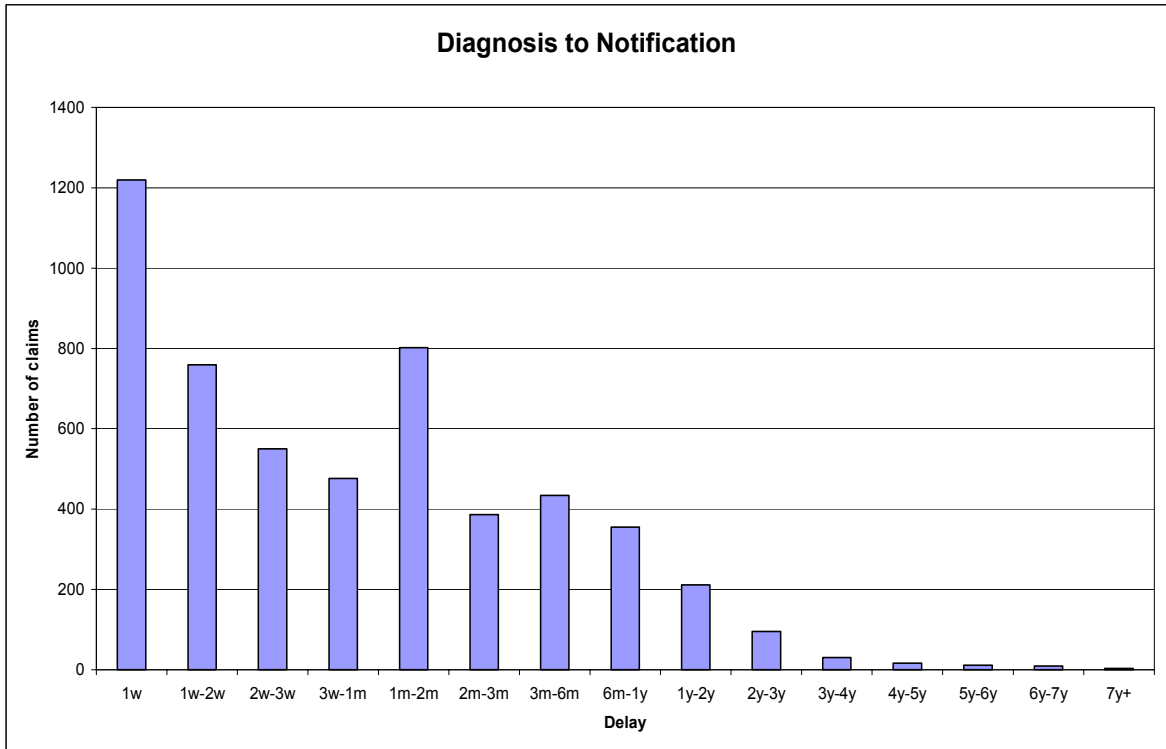
5.2.4. The distribution of the delays for 1999-2002 claims is also of interest. Figures 1 and 2 below show the observed distribution of delays from diagnosis to notification and from diagnosis to settlement. This confirms the existence of a number of cases where delays have run into several years but also shows relatively short delays in a significant proportion of claims.

5.2.5. These charts should be interpreted with care as they simply record the distribution of observed delays based on claims settled in the quadrennium. There is no attempt in these charts to adjust for changes in exposure over time. As the portfolio of Critical Illness business analysed has been growing rapidly over time, the longer observed delays will relate to claims diagnosed in earlier years when exposures were significantly smaller. The run-off of claims arising from a particular year's exposure, showing the underlying pattern of settlement delays, would be more skewed to the right (i.e. a higher proportion of longer delays than observed in figure 2 below).

5.2.6. In addition, the relative immaturity of the offices' portfolios means that if there are claims taking even longer periods to settle (extending beyond that seen in the claims data for the quadrennium), these will not yet be present in the data submitted to the CMI.

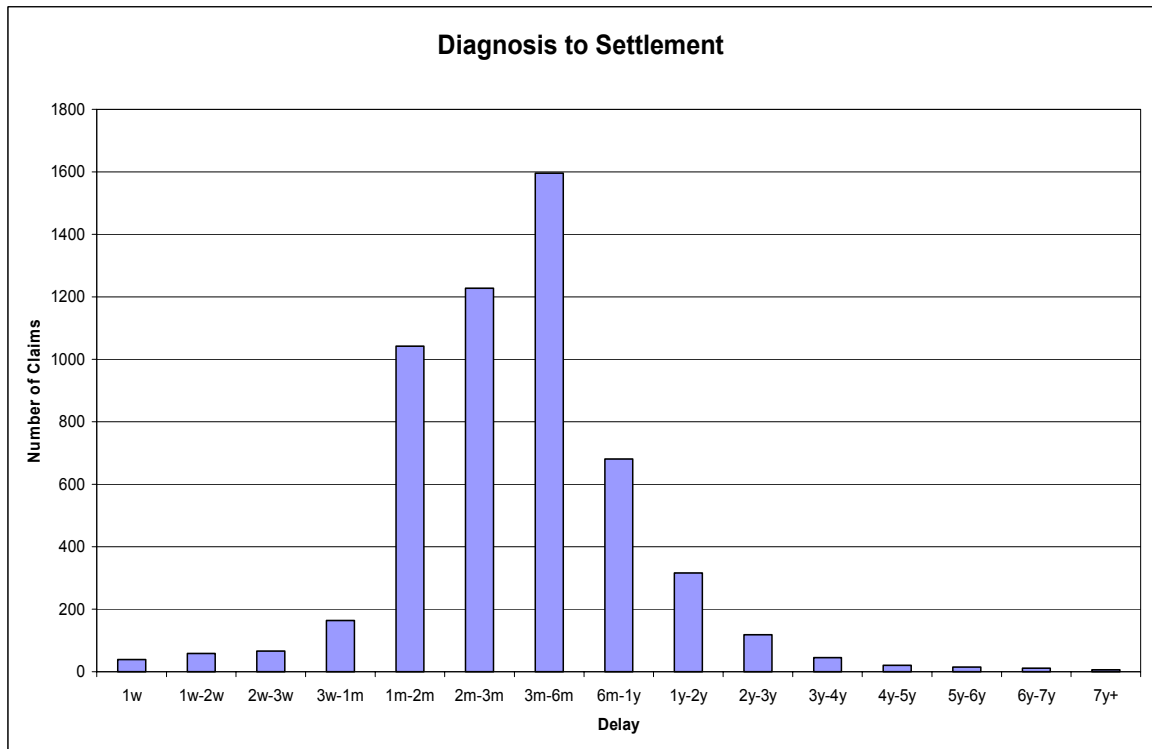
5.2.7. This underlying pattern of settlement delays is considered further in Section 5.5.

Figure 1. Observed distribution of delays between diagnosis and notification. All Critical Illness claims (including death claims). 1999-2002.



Note: This analysis is based on the 5,356 claims where both the dates of diagnosis and notification were submitted and after the process described in Section 5.1.6 had been carried out.

Figure 2. Observed distribution of delay between diagnosis and settlement. All Critical Illness claims (including death claims). 1999-2002.



Note: This analysis is based on the 5,404 claims where both the dates of diagnosis and settlement were submitted and after the process described in Section 5.1.6 had been carried out.

5.3. *Issues relating to the choice of “date of claim” for analysis purposes*

5.3.1. The Critical Illness Committee had to decide which of the four dates relating to a claim (diagnosis, notification, admission and settlement) should be used as the “date of claim” for analysis purposes. In fact, the date of claim is used for two distinct purposes:

- (a) To assign a claim to a particular year’s experience for the purpose of the analysis
- (b) To determine the age and policy duration at the time of claim.

5.3.2. Although using the date of settlement as the date of claim has the advantage of reflecting actual cashflow, this would not be appropriate for pricing or reserving as it would not reflect the actual underlying incidence of claims. The dates of notification and admission are even less appropriate as they neither reflect actual cashflow nor the actual underlying incidence of claims.

5.3.3. The date of diagnosis reflects the true cost to the insurer as measured by claims actually incurred in the period of insurance but not necessarily reported or settled. Setting the date of claim as the date of diagnosis is also consistent with the other CMI investigations where the date of the insured event, e.g. death, is used. However, using the date of diagnosis results in several practical problems:

- (a) Establishing date of diagnosis:
 - Only about half of the claims records for 1999-2002 have this information so missing dates of diagnosis need to be estimated.
 - The date of diagnosis is often not precisely defined and could be defined differently by different offices or even by different claims assessors within an office. A good example is TPD.
- (b) Using the date of diagnosis to assign claims to a particular year’s experience:
 - As claims are generally submitted to the CMI according to the year in which they are settled (this is the earliest opportunity for offices to report claims with certainty), this would require the reallocation of many claims settled in a single year to previous years. In particular, this would mean that any claims put back into a year before an office joined the investigation would be ignored as there would be no exposure information to match against it and in any event all claims pertaining to years prior to 1999 would be unusable. Whilst there are strong arguments for this approach which matches claims and exposure properly, they are somewhat undermined by the difficulties noted above in establishing a clear date of diagnosis.
 - Several years’ data would need to be collected before a year’s experience could be confidently closed off leading to problems with out-of-date information. Whilst estimates of outstanding claims could be made, these estimates would need to be updated as IBNS claims are replaced by actual claims.
 - The above problems are further complicated by offices joining and leaving the investigation and from new portfolios of data being added by individual offices from one year to the next.

The approach adopted to resolve these issues is described in Sections 6.4 to 6.6.

5.4. *Observed claim delay patterns*

5.4.1. Once all the offices had submitted all their data for the 1999-2002 quadrennium and the data had been validated, initial analysis showed that the average claim delays observed in the quadrennium data (Table 2) were longer than those observed in the 1998-1999 data. This was in line with the Committee's prior expectation that the observed claim delays would lengthen as more of the claims with long delays were actually settled and data in respect of these claims submitted to the CMI.

5.4.2. The delays inherent in the 1999-2002 are shown in Table 3 below. It can be noted that the longest observed delay between the Date of Diagnosis and the Date of Settlement is 9 years. The numbers also illustrate the paucity of data on claims with very long delays.

Table 3. Analysis of claims by year diagnosed and year settled for claim records with both dates after the process described in Section 5.1.6 has been carried out.

Year claim diagnosed	Year of claim settlement				Total
	1999	2000	2001	2002	
1993	0	0	0	1	1
1994	0	2	4	1	7
1995	0	1	8	5	14
1996	4	5	5	4	18
1997	11	23	11	6	51
1998	108	45	27	19	199
1999	356	319	55	22	752
2000	0	736	429	65	1,230
2001	0	0	1091	560	1,651
2002	0	0	0	1481	1,481
Total	479	1,131	1,630	2,164	5,404

5.4.3. As the claims data were now of greater volume, it became possible to carry out more detailed analysis on the claims data for the quadrennium. The analysis showed that:

- the average claim delays increased with the policy duration at the dates of notification, admission and settlement.
- the average delays varied by cause of claim – in particular, death claims were observed to be settled faster, on average, than critical illness claims.
- the average speed at which claims were observed to be settled differed significantly between the offices, particularly for claims settled within 2 years of diagnosis.

These features prompted the Committee to reconsider the methodology and the revised methodology is documented in Sections 6.4 to 6.6.

5.5. *Underlying claim delay patterns*

5.5.1. In order to further understand the impact of claim delays, the Committee attempted to derive an estimate of the true underlying pattern of delays in claim settlement. To do this, it was first necessary to consider the known distortions in the observed claim pattern, namely:

- the combination of large increases in expected claims (due to increasing business and/or ageing portfolios) for the offices contributing to the investigation and the

range of actual delays shown in Figure 2 which meant that observed delays are weighted towards shorter-delayed claims on recently-written business.

- the relative immaturity of the offices' portfolios means that if there are claims taking even longer periods to settle (extending beyond that seen in the claims data for the quadrennium), these will not yet be present in the data submitted to the CMI.
- further distortions due to offices either entering or leaving the investigation or adding new portfolios to the data submitted.

5.5.2. These factors mean that it is not possible to use a simple approach such as a chain ladder. For example, suppose that an office has contributed data in respect of a consistent portfolio of business throughout the quadrennium. A number of the claims settled in the period will relate to dates of diagnosis in 1998 and earlier. Necessarily, all claims that have taken 4 or more years to settle after diagnosis fall into this category and as we have no data on the exposure giving rise to these claims, our understanding of the tail is limited.

5.5.3. This situation is compounded when one considers that not all offices contributed data throughout the period and that in some cases new portfolios were introduced within an office (which in essence are akin to new offices joining the investigation).

5.5.4. Our analysis has therefore attempted to match up exposures correctly to delays of differing lengths in order to derive an estimate of the true underlying pattern of delays in claim settlement. To do this, we had to overcome what we have termed "fault lines" where a year-end is involved.

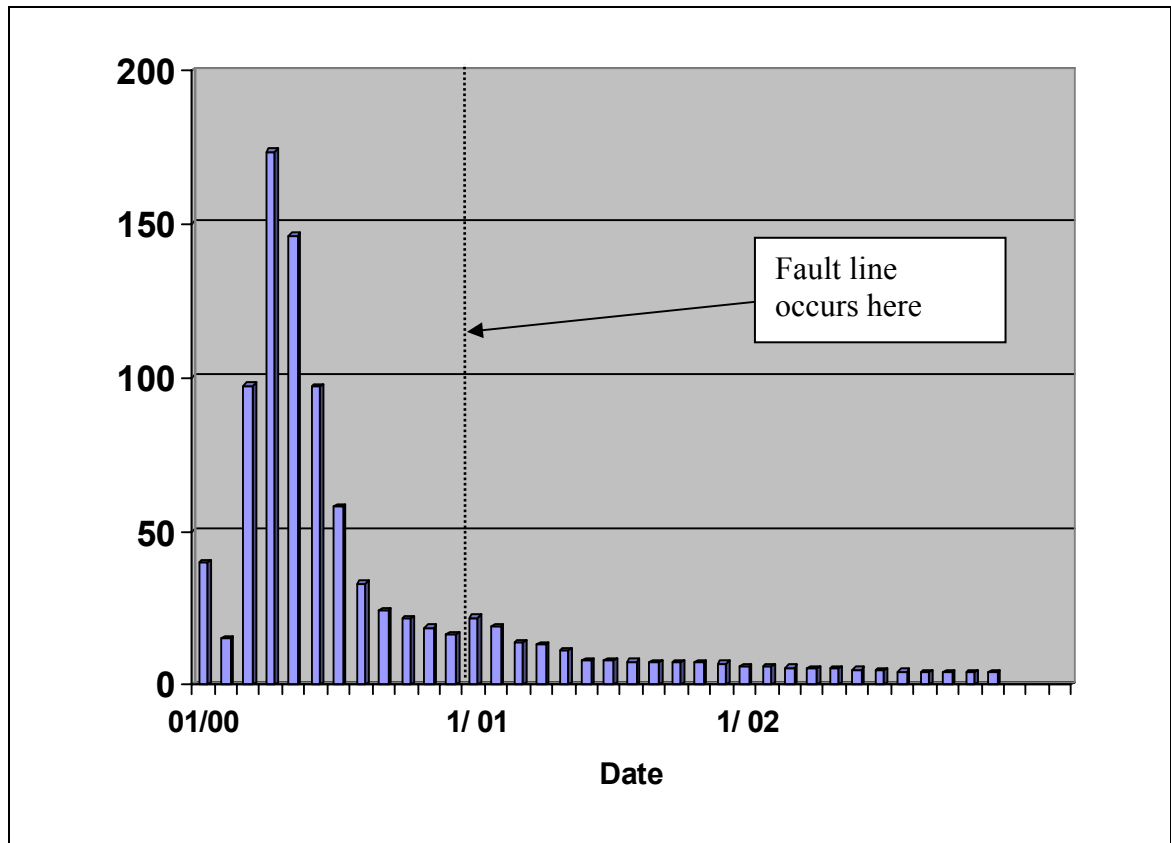
5.5.5. Our analysis looks at claims data tabulated in a two-way table by month of diagnosis and by month of settlement (for all those claims where both the dates were provided). It uses adjacent cells in the claims development process where the underlying exposure is identical, and hence the development from one period to the next can be considered reliable in order to derive the overall underlying claim delay pattern. Thus, for example, for claims diagnosed in January 2000, the numbers settled by month in each month from January to December 2000 can be taken as a reliable indicator of the claim delay pattern. Claims settled in January 2001 may not, however, as they may relate to an office (or portfolio) which was not included in the investigation for 2000 and so the claim development is inconsistent.

The effect of fault lines is illustrated in Figure 3 below which uses hypothetical (but realistic) numbers. The example assumes that new offices enter the investigation in 2001, resulting in a 50% increase in the underlying exposure. This results in a fault line between December 2000 and January 2001. Though not included in this example, further changes in exposure due to more offices entering the investigation in 2002 and so on would lead to additional fault lines.

The cohort of claims diagnosed in January 2000 therefore provides us with estimates of the delay pattern excluding possible fault lines between months 12 and 13, months 24 and 25 and so on. Claims diagnosed in February 2000 will provide us with an estimate of the delay pattern between months 12 and 13, etc but possibly not between months 11 and 12, months 23 and 24 and so on.

The claim delay pattern is then established using all cells not affected by fault lines.

Figure 3. Hypothetical distribution of delay between diagnosis and settlement to illustrate the impact of a “fault line” assuming 1000 claims with a date of diagnosis in January 2000.

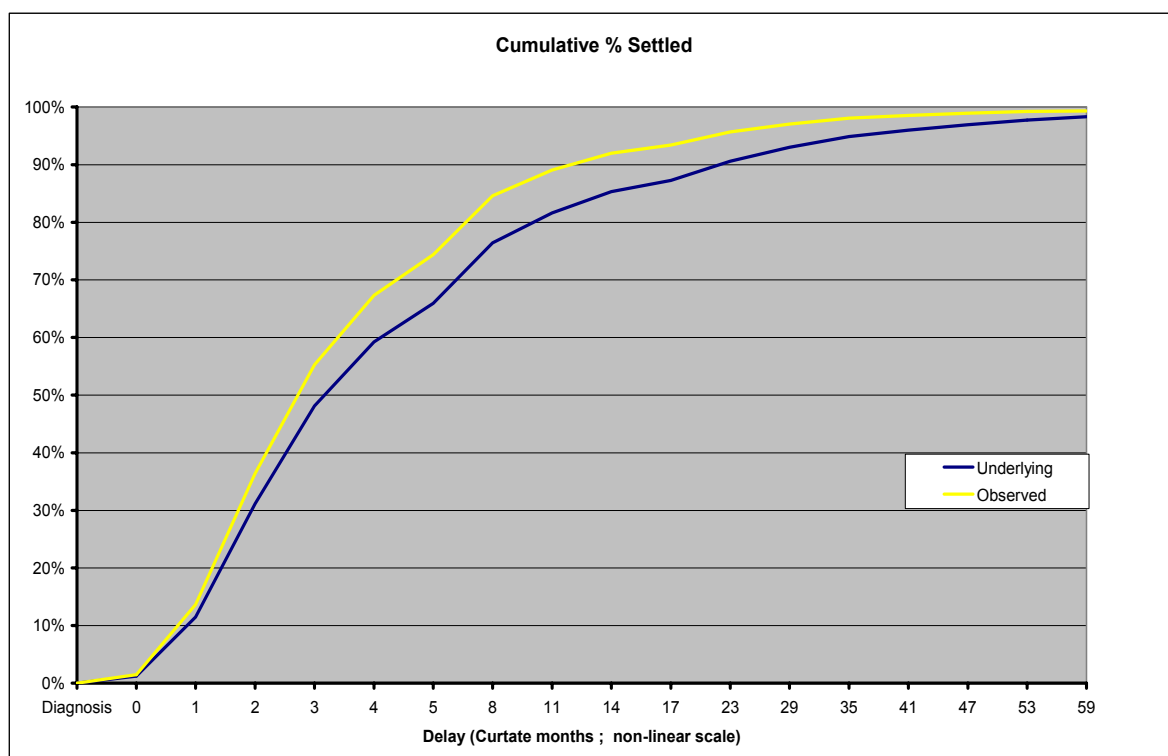


5.5.6. Some likely fault lines within the 1999-2002 data can be seen in Table 3, for example for year of diagnosis 1997 between 1999 and 2000. However because fault lines occur at company or portfolio level, others will be less visible because they are diluted by the volumes of other data.

5.5.7. The paucity of data with very long delays meant that the 1999-2002 data did not provide us with a reliable basis for estimating the underlying claim delay pattern in this area. Our approach was therefore to only use the actual data for claim delays up to 5 years. Beyond that point, we have used a simple extrapolation.

5.5.8. Figure 4 below compares the observed delay pattern with our estimate of the underlying delay pattern. Whereas the average observed delay between diagnosis and settlement is 176 days, the average for the derived underlying delay pattern is of the order of 260 days. The distribution is very skew and so the average delay increases markedly once correct weight is applied to the long tail.

Figure 4: Observed and underlying claim settlement delay patterns. Claims included in the 1999-2002 quadrennium analyses where offices provided both the date of diagnosis and the date of settlement.



5.5.9. It should be noted that in the approach described above no attempt has been made to differentiate claim delays by office or cause of claim. This is because the analysis methodology is data-hungry and current data volumes are insufficient to allow further subdivision.

6. ANALYSIS METHODOLOGY

6.1. Cause of claim

6.1.1. The cause of claim supplied by the offices is converted into one of 56 standard CMI subdivisions of cause of claim. About half of these refer to the specific site of a cancer, including “site not specified”. There is also a general “unknown” category where the office is unable to supply any cause, although the CMI encourages offices to supply the cause wherever possible. For the purposes of analyses to date, these detailed causes are merged into a smaller number of causes. In particular, currently all cancers are merged and the less frequent causes are merged into an “other” group. The detailed causes and the group to which they are currently allocated for the purpose of analysis are shown in Appendix C.

6.2. Exposure calculation

6.2.1. The in force data collected is essentially a list of in force policies at each year end. An exact method of exposure calculation is not possible as the CMI does not collect data on when policies move in and out of force during the year. A census method is therefore used

which assumes that all policies no longer in force at the end of the year exited half-way through the year and that new business entered mid-year.

6.2.2. For all in force data at 1/1/n, the age nearest and curtate duration is calculated as at 1/1/n. The data is then grouped for analysis purposes by age, duration, sex, smoker status, benefit type, and/or sales channel as required. Each census record will include both a “lives” and an “amounts” field.

6.2.3. The claims in calendar year n are grouped as for the in force data except that the age nearest and curtate duration calculation will be as at the “date of claim”, agreed by the CMI Critical Illness Committee to be the (actual or estimated) date of diagnosis (see Section 6.5). The claims are further sub-divided by cause (as allocated by CMI) for analysis and, in particular, deaths are separately identified for Full Acceleration business.

6.2.4. The exposure for age x and duration r in calendar year n ($E_{x,r,n}$) is then calculated as

$$E_{x,r,n} = \frac{1}{2} \times [P_{x,r,n} + P_{x,r,n+1} + \sum_{all\ c} {}^c\theta_{x,r,n}]$$

Where

$P_{x,r,n}$ = In force at 1/1/n for lives age x nearest and curtate duration r at 1/1/n

and

${}^c\theta_{x,r,n}$ = Claim of type c in calendar year n, age x nearest and curtate duration r at date of claim

6.2.5. Different exposure calculations are not required when analysing experience for a particular cause of claim as a multiple decrement table is adopted and used for the expected experience. However, any crude rates obtained will also be dependent q-type probabilities and the estimate of the probability of exiting by any one cause will depend on the probability of exiting by each of the other causes.

6.2.6. The critical illnesses covered by policies and definitions applied vary with offices and over time, including some retrospective changes by offices. However, the most serious of these critical illnesses, if not covered explicitly by the policy conditions would usually still be covered by the TPD condition. Given these variations and the data constraints, the Critical Illness Committee decided that it was not feasible to attempt to calculate a separate exposure for each of the critical illnesses covered, nor to reflect any variations in the definition of particular critical illness events. The situation will be kept under review in case there are major changes in the form of critical illness cover offered in the market.

6.3. *Calculation of expected claims*

6.3.1. The method of analysis involves comparing actual claims to expected claims, where the latter are based on a standard table:

- The main comparisons are carried out against the CIBT93 base table constructed by the Critical Illness Healthcare Study Group and published in the paper “A Critical Review” presented to the Staple Inn Actuarial Society on 14 March 2000. This table was developed from English population data in respect of 1993 and is not adjusted in any way to estimate rates for an insured experience (although the TPD element was derived from insured data).

- A less detailed comparison is carried out against the IC94 table developed in the paper “Reserving for Critical Illness Guarantees” presented to the Society of Actuaries in Ireland in November 1994. IC94 was largely derived from UK population data but adjusted for Irish experience as deemed appropriate. Further adjustments were then made to attempt to move from a population experience to an insured experience.

Both tables are Aggregate tables, (i.e. they contain no adjustment for smoker/non-smoker).

6.3.2. For each age and duration, the value of 100 A/E is calculated as:

$$100 \text{ A/E} = 100 \times {}_k\theta_{x,r,n} \div [E_{x,r,n} \times {}_kq_x]$$

where ${}_k\theta_{x,r,n}$ is the sum of ${}_c\theta_{x,r,n}$ for cause of claim (or group of causes) k and ${}_kq_x$ is the relevant rate of incidence from the standard table for cause of claim (or group of causes) k.

6.3.3. The exposure and claims used in the calculation are grouped by sex, smoker status, territory, benefit type and sales channel as required. They may also be grouped over a number of calendar years.

6.4. *CMI methodology for “date of claim” for the 1999-2002 quadrennium analyses*

6.4.1. In view of the issues set out in Section 5.3 above, the Committee decided against using a single date of claim both to determine the age and policy duration at the time of claim and to assign claims to a particular year’s experience. The Committee instead agreed the following methodology:

- Use the claims records submitted in respect of a particular year (the great majority of which relate to claims settled in that year) to compare against the expected claims derived from the exposure in that year. This is considered further in Section 6.6.
- Calculate the age and duration at the time of the claim based on the date of diagnosis, if supplied. If not supplied, the date of diagnosis is estimated from one of the other dates of claim supplied using the observed average delays between these dates as described in Section 6.5.

6.4.2. The Committee are conscious of a number of defects relating to this methodology:

- It does not correctly match the occurrence of the claim event (diagnosis) to exposure where the event happened in a prior year or if the claim is yet to be settled. One might therefore expect it to understate the eventual claims experience of a year when the insured portfolio is growing and/or ageing, as previous years’ delayed claims are matched to the investigation year’s exposure and the investigation year’s delayed claims are omitted.
- The estimation process for the date of diagnosis, when not supplied, is also crude as delays have been observed to vary by a number of factors including office and cause of claim which have not been taken into account.

6.4.3. However, the Critical Illness Committee took a pragmatic decision to use this methodology for the 1999-2002 quadrennium results as it believes that the methodology enables the key features of a year’s experience to be published relatively quickly. It also resists the tendency of showing a “false” pattern of initial selection due to delays between diagnosis and settlement if the date of settlement were used to determine the age and duration at the

time of claim. In any case, the Committee believes that the volumes of data currently available do not permit a more sophisticated approach.

6.5. *Estimation of missing dates of diagnosis*

6.5.1. As dates of diagnosis are only available for some 56% of 1999-2002 claims data, the missing dates of diagnosis have to be estimated from one of the other dates of claim. As information from offices indicated that the dates of settlement and admission (in that order) were more reliable than the date of notification, and the analyses showed that the average claim delay changed with the policy duration at these dates and the type of claim, the Committee decided on the following methodology to estimate these missing dates of diagnosis:

- (a) Use date of settlement, if available, less an adjustment based on the type of claim and the policy duration at the date of settlement from Table 5.
- (b) Where the date of settlement is not available, use date of admission, if available, less an adjustment based on the type of claim and the policy duration at the date of admission from Table 5.
- (c) Where neither the date of settlement nor the date of admission are available, use date of notification less an adjustment based on the type of claim and the policy duration at the date of notification from Table 5.
- (d) The estimated date of diagnosis is set to the policy commencement date if it would otherwise precede it. For the 1999-2002 quadrennium claims, this was not required as none of the estimated dates of diagnosis fell before the policy commencement date.

6.5.2. This methodology represents a refinement of that originally adopted for the 1998, 1999 and 2000 results. The previous methodology and the impact of the change is considered in Appendix A.

6.5.3. The source of the date of diagnosis used for the claims in the analyses of the 1999-2002 quadrennium is shown in Table 4 below.

Table 4. Source of date of diagnosis.

Source of date of diagnosis:		Number of claims	Percentage of claims
Actual date of diagnosis		6,649	56.3%
Estimated date of diagnosis based on:	Date of settlement	4,990	42.3%
	Date of admission	142	1.2%
	Date of notification	22	0.2%
Total		11,803	100.0%

6.5.4. The adjustments used to estimate the missing dates of diagnosis, based on the average observed delay within bands of policy duration at either the date of settlement, admission or notification, are given in Table 5 below.

6.5.5. Separate adjustments are used to estimate missing dates of diagnosis for critical illness claims and date of death for death claims as their observed and underlying delay patterns are very different. In general, death claims tend to get settled much faster. Though claim delay patterns can also be expected to vary by cause of critical illness, office and other

factors, the Committee decided against using adjustments that varied by these factors as there was insufficient data to estimate these adjustments credibly.

Table 5. Adjustments used to estimate missing dates of diagnosis from one of the other three dates of claim.

Duration at relevant date of claim in days	Average observed delay in days from date of diagnosis		
	Settlement	Admission	Notification
Critical Illness Claims			
<=91	53	48	19
92 – 183	82	77	29
184 – 365	104	99	50
366 – 730	125	120	68
731 – 1096	154	149	117
1097 – 1462	195	190	141
1463 – 1828	234	229	192
1829 – 2194	236	231	206
2195 – 2560	261	256	208
>=2561	298	293	253
Death Claims			
<=91	41	36	8
92 – 183	44	39	14
>=184	103	98	12

6.5.6. Observed delays, rather than underlying delays, are used to estimate missing dates of diagnosis because:

- Using observed delays implicitly reverses the distorting effects of increasing numbers of expected claims over the period. This is illustrated by the fact that for offices that did provide dates of diagnosis, it is the set of adjustments based on observed delays that would correctly estimate the dates of diagnosis from one of the other dates of claim.
- If the underlying delays were to be used to estimate missing dates of diagnosis, additional adjustments would be needed to remove the distorting effect of increasing numbers of expected claims.

6.5.7. Frequently, offices either provided the date of diagnosis for all their claims or not all. Therefore, the process for estimating missing dates of diagnosis implicitly assumes that the offices that did not provide dates of diagnosis have claim delay patterns that are similar, on average, to the offices that did.

6.6. *Allocation of claims to experience years for the 1999-2002 quadrennium analyses*

6.6.1. The Committee also considered whether it could improve the analysis methodology by using dates of diagnosis to allocate claims to particular years for the purpose of the analysis and to match it against that year's exposure. However the investigations indicated a number of difficulties in trying to use the date of diagnosis for this purpose.

- (a) Actual dates of diagnosis are only available for some 56% of claims and these mainly relate to specific offices – offices tend to either provide dates of diagnosis for all their claims data or not at all. Among the offices that do provide dates of diagnosis, there is

wide variation in the average delays observed due to both variations in the underlying settlement patterns as well as differences in business growth experienced in the years up to 2002.

Therefore, for offices that did not provide dates of diagnosis, the missing dates of diagnosis are estimated using average delays for offices that provided dates of diagnosis but may have very different settlement patterns and past growth rates. Also, estimated dates of diagnosis are point estimates and so do not allow for the whole range of possible underlying dates of diagnosis.

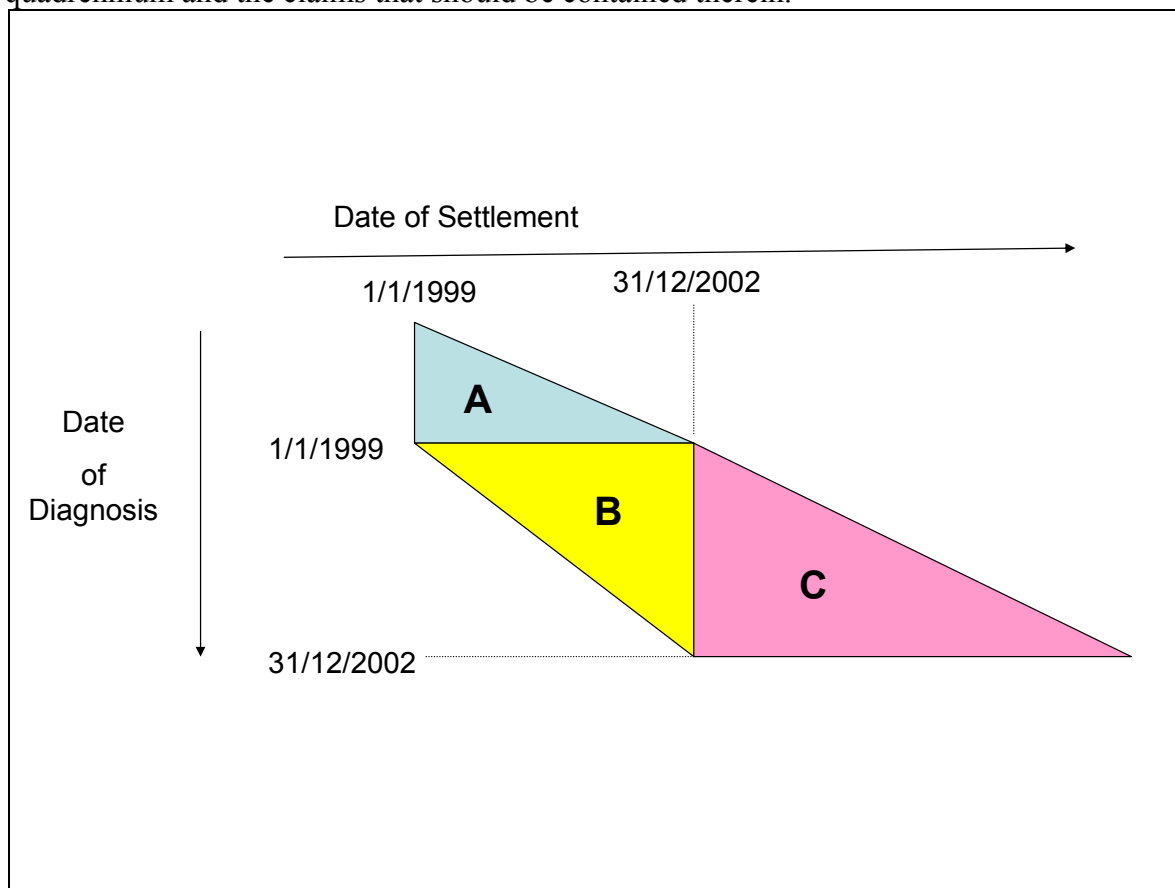
The estimated dates of diagnosis do not therefore provide a strong foundation for removing claims from the investigation.

- (b) Additionally, even if all offices had supplied dates of diagnosis for all their settled claims and these were used to allocate claims to experience years, chain ladder type adjustments for IBNS claims would still be required. However, these adjustments for IBNS claims would have to be office specific in order to allow for the difference between the offices in claims growth experienced (due to increasing business and/or ageing portfolios) over the years to 2002.

6.6.2. The Committee therefore considered that using estimated dates of diagnosis to allocate claims to experience years could lead to misleading results and provide a false sense of certainty. Instead, the Committee decided to leave unchanged the methodology used to assign a claim to a particular year's experience. The claims records submitted in respect of a particular year (the great majority of which related to claims settled in the year) would continue to be used to compare against the expected claims derived from the exposure in the year.

6.6.3. As a result the experience reports will understate the actual underlying experience as they do not allow for IBNS claims, though this understatement is reduced by the extent of claims settled in the year but diagnosed in previous years. The effective mis-statement for the whole quadrennium is illustrated schematically in Figure 5 below. The claims that are used in the investigation 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 but are included in the analysis by our approach. Claims in area C which also have dates of diagnosis within the investigation period have yet to be reported to the CMI. Understatement results if the claims in area C exceed the claims in area A as the Committee believes to be the case for the All Offices' experience given the growth in the number of claims over time.

Figure 5: Schematic diagram illustrating the claims that are contained in the 1999-2002 quadrennium and the claims that should be contained therein.



7. GROSSING-UP FACTORS

- 7.1. The Committee has therefore developed a model to estimate the grossing-up factors required to adjust for the mis-statement in reported results described in preceding sections. In principle, a factor should be applied to the results to allow for:
- the removal of claims settled in the period but diagnosed before the start of the period;
 - the addition of claims diagnosed in the period but settled after the period and reported to the CMI; and
 - the addition of claims diagnosed in the period and yet to be settled and reported to the CMI.
- 7.2. The model simulates a portfolio of critical illness business using quarter-year time intervals. The key assumptions to the model are:
- rates of business growth;
 - sets of claim rates and lapse rates; and
 - the underlying pattern of claim delays from date of diagnosis to date of settlement.
- For this, we have used the underlying pattern of claim delays as derived from our analyses as set out in Section 5.5.
- 7.3. From these inputs, exposures are calculated for each time interval in the projection and then claims diagnosed in these time intervals were derived using the assumed claim rates. The settlement of these claims can then be mapped using the claim delay patterns. The results from the model can be used to compare modelled experience as would be recorded

using the CMI methodology (allocating claims to each period using date of settlement) with the actual experience (allocating claims to each period using date of diagnosis). The required grossing-up factor is the ratio of the latter over former.

- 7.4. The model was tested by comparing the modelled “observed” pattern of claim delays with the actual pattern of observed claim delays within the claims data for 1999-2002. For this test, the growth and the claims rates were set to be consistent with the appropriate subset of the All Office experience over the 1999-2002 quadrennium (restricting to portfolios of business for which the claims data contained both date of diagnosis and date of settlement, and adjusting for changes in offices and portfolios of businesses reported to the CMI) and a close match was achieved.
- 7.5. As the grossing-up factor required depends on the business growth and claim rates experienced by the office, the Committee has decided to provide a table of sample adjustments applicable to a range of growth rates in the expected number of claims. Offices are therefore able to estimate the grossing-up factor most relevant to their own claims growth experience that would apply to the CMI analyses of their own office’s data. When considering results for an individual investigation year, the relevant rate of growth is approximated by the increase in expected claims for that year over the previous year’s expected claims, assuming a consistent portfolio of business.

Rate of growth in expected claims	Approximate grossing-up factor
0%	100%
10%	107%
20%	112%
30%	117%
50%	124%
75%	132%
100%	139%

The reported 100A/E in the CMI analyses should be multiplied by the appropriate grossing-up factor.

- 7.6. For the 1999-2002 quadrennium All Office experience, the average rate of growth was close to 25% per annum leading to an overall grossing-up factor of 115%. **Therefore the reported 100A/E for the 1999-2002 All Office experience should be multiplied by 1.15.**
- 7.7. When reviewing the results incorporating these grossing-up factors, offices need to bear in mind the following:
- The pattern of claim delays was derived from less than 50% of the total claims data largely relating to a sub-set of offices (which may not have a consistent definition of date of diagnosis) and is applied without adjustment to the remaining claims.
 - The pattern of claim delays may change over time.
 - The grossing-up factor will vary by investigation year so care should be taken in interpreting trends.
 - No allowance has been made in the claim delay patterns, and hence the grossing-up factor, for variation by office or cause of claim. Particular care should be taken in interpreting the reported results by distribution channel and cause of claim.

- The rate of growth in expected claims is likely to vary by duration leading to different grossing-up factors and therefore possible distortions in the results by duration.
- The rate of growth in expected claims is likely to vary by age leading to different grossing-up factors and therefore possible distortions in the results by age
- The rate of growth in expected claims may also vary by gender, smoker status and other factors leading to different grossing-up factors and therefore further distortions in the results.
- The grossing-up factors have been derived from lives experience and therefore may not be appropriate for amounts experience.

Appendix A: Impact of refining the date of diagnosis algorithm

Initially the CMI prepared results for 1998 and 1999 where the age and duration at claim were based on the date of settlement. As data volumes increased and the impact of claim delays became apparent, the Committee amended the approach to instead use date of diagnosis for the purpose of calculating the age and duration at claim. This methodology was documented in the CMI paper “Collection and analysis of data for the Critical Illness investigation” which was distributed with the 1998-1999 analyses.

In the 1998-2000 analyses initially published by the CMI, missing dates of diagnosis were estimated using the following algorithm:

Use date of diagnosis if available; otherwise approximate date of diagnosis by:

- a) Use date of settlement, if available, less 155 days.
- b) Otherwise, use date of notification, if available, less 80 days.
- c) Otherwise, use date of admission less 155 days.
- d) The estimated date of diagnosis is set to the policy commencement date if it would otherwise precede it.

As can be seen from the above, the duration of the policy at the date of claim was not used as a factor in estimating the missing dates of diagnosis. In addition, no differentiation was made between death and critical illness claims.

Following receipt of the of the 2001-2002 claim data, revisions to 1999-2000 claim data by many offices, further information from offices about the reliability of their dates of claim and further analysis of the 1999-2002 claim data, the algorithm used to estimate missing dates of diagnosis was refined as set out in Section 6.5. One improvement resulting from the amended methodology was that no claims were covered by (d) above.

The changes in methodology have negligible impact on the overall 100A/Es, as no claims are moved into or out of the analysis. However, they have a significant impact on the results by duration. In the following discussion the algorithm initially used to estimate missing dates of diagnosis for the 1998-2000 analyses is referred to as the “original date of diagnosis method”. The revised methodology used to produce the analyses for the 1999-2002 quadrennium is referred to as the “refined date of diagnosis method”.

Figure A1 compares the previously published 1999 results obtained using the “date of settlement method” (where the age and duration at claim are based on the date of settlement) and the “original date of diagnosis method”. Figure A2 compares the 1999-2002 quadrennium results obtained using the two diagnosis methodologies by select duration. The figures show 100 A/E by select duration in respect of all claims (including mortality) as compared against CIBT93 for Full Acceleration business for male lives using Aggregate (i.e. smoker/non-smoker combined) data.

Figure A1 shows that, compared to the use of the original date of diagnosis method, the date of settlement method appeared to produce a positive initial selection effect which is likely to be overstated for the reasons set out in Section 1.7. Figure A2 shows that using the refined date of diagnosis method, there appears to be a stronger positive selection effect, particularly in the first policy year. However, the observed positive initial selection should be treated with caution as the figures are presented before applying appropriate grossing-up factors.

Figure A1. Effect of revising the date of claim methodology from date of settlement method to original date of diagnosis method. Full Acceleration business, all claims (including mortality), Aggregate data. 100 A/E vs CIBT93, lives, 1999

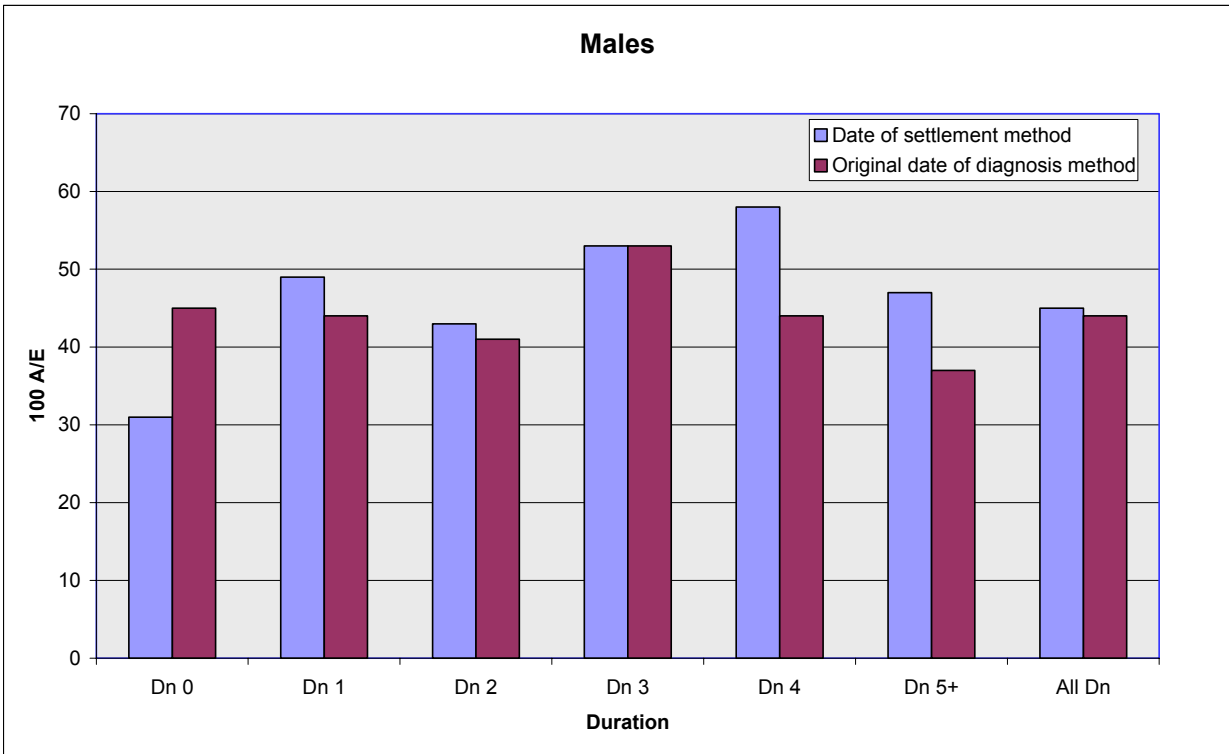
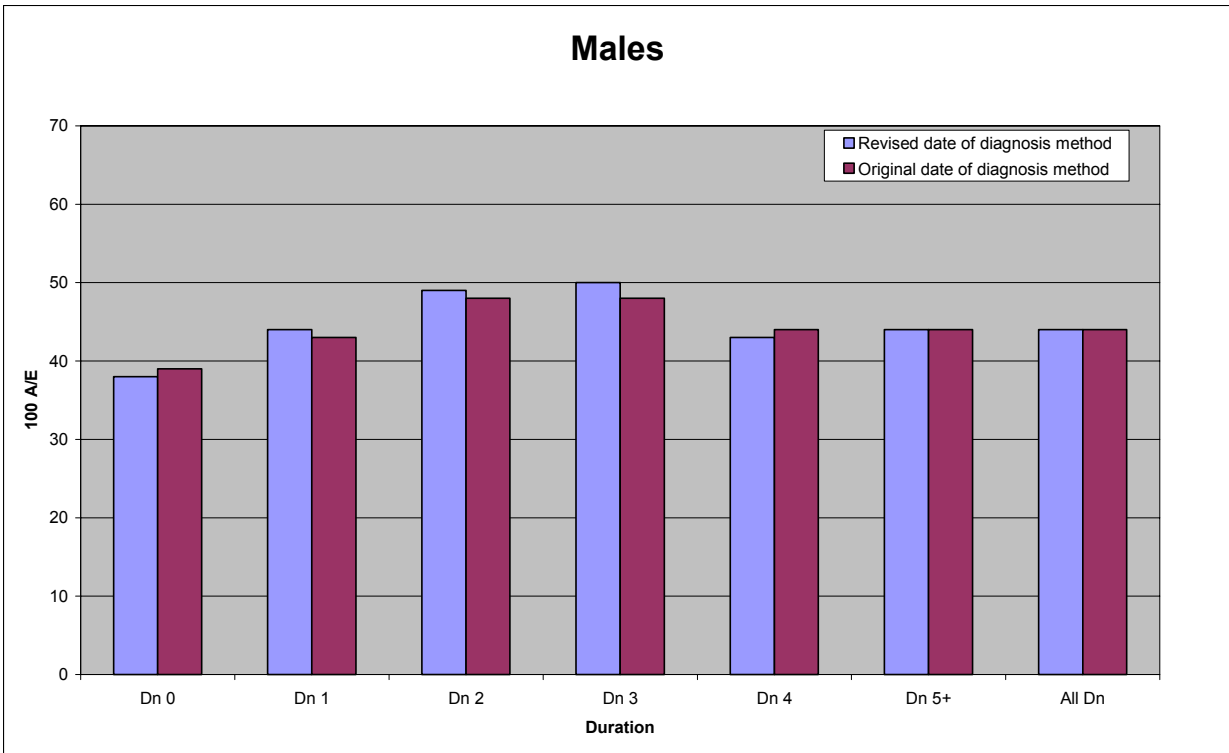


Figure A2. Effect of refining the date of claim methodology from original date of diagnosis method to refined date of diagnosis method. Full Acceleration business, all claims (including mortality), Aggregate data. 100 A/E vs CIBT93, lives, 1999 – 2002



Appendix B: Data requirements for each in force policy or claim record

Field Description (to appear in spreadsheet or database submissions)	Column Position (text submissions only)	Values or Format
Record type	1	I = in force C = claim
Office number	2-4	NNN
Record year	5-8	YYYY
Benefit type	9	S = stand alone F = full acceleration
Sex	10	M, F
Territory	11	1 = UK 2 = Eire
Smoker code	12	N = non smoker S = smoker U = undifferentiated
Date of birth	13-20	DDMMYYYY
Date of commencement	21-28	DDMMYYYY
Benefit Amount	29-36	NNNNNNNN
Identification Code	37-46	Any Alphanumeric
Product Code	47-56	Any Alphanumeric
Single or joint life	57	S = Single life J= Joint life (first event)
Sales channel	58	B = Bancassurer D = Direct Sales I = IFA O = Other U = Unknown
<i>The fields below should be completed only for claims records (C). They should be left blank for in force records.</i>		
Date of diagnosis/death	59-66	DDMMYYYY
Date of notification	67-74	DDMMYYYY
Data claim admitted	75-82	DDMMYYYY
Date of settlement	83-90	DDMMYYYY
Type of claim	91	D = Death C = Critical Illness
Cause of critical illness claim	92-140	Any Alphanumeric

Appendix C: Cause of claim groupings

	CMI Cause	Analysis Group
1	Deaths	Death
2	Terminal Illness	Death
3	Heart Attack	Heart Attack
4	Stroke	Stroke
5	Coronary Artery Bypass Graft (CABG)	CABG
6	Multiple Sclerosis	Multiple Sclerosis
7	Kidney Failure	Kidney Failure
8	Major Organ Transplant (MOT)	MOT
9	Total Permanent Disability (TPD)	TPD
10	Aorta Graft Surgery	Other
11	Benign Brain Tumour	Other
12	Blindness	Other
13	Deafness	Other
14	Heart Valve Replacement / Repair	Other
15	Loss of limbs	Other
16	Loss of speech	Other
17	Motor Neurone Disease	Other
18	Paralysis / Paraplegia	Other
19	Coma	Other
20	Parkinson's Disease	Other
21	Third Degree Burns	Other
22	Alzheimer's Disease	Other
23	Angioplasty	Other
24	Other	Other
25	Unknown	Unknown
26	Cancer - site not specified	Cancer
27	Malignant neoplasm of lip, oral cavity and pharynx	Cancer
28	Malignant neoplasm of digestive organs and peritoneum - unspecified	Cancer
29	Malignant neoplasm of oesophagus	Cancer
30	Malignant neoplasm of stomach	Cancer
31	Malignant neoplasm of small intestine including duodenum	Cancer
32	Malignant neoplasm of colon	Cancer
33	Malignant neoplasm of rectum, rectosigmoid junction and anus	Cancer
34	Malignant neoplasm of liver	Cancer
35	Malignant neoplasm of pancreas	Cancer
36	Malignant neoplasm of respiratory and intrathoracic organs - unspecified	Cancer
37	Malignant neoplasm of larynx	Cancer
38	Malignant neoplasm of trachea, bronchus and lung	Cancer
39	Malignant neoplasm of bone, connective tissue, skin and breast - unspecified	Cancer
40	Malignant neoplasm of bone and articular cartilage	Cancer
41	Malignant melanoma of skin	Cancer
42	Other malignant neoplasm of skin	Cancer
43	Malignant neoplasm of female breast	Cancer
44	Malignant neoplasm of genitourinary organs - unspecified	Cancer
45	Malignant neoplasm of ovary and uterine adnexa	Cancer
46	Malignant neoplasm of prostate	Cancer
47	Malignant neoplasm of testis	Cancer
48	Malignant neoplasm of bladder	Cancer
49	Malignant neoplasm of kidney and other urinary organs	Cancer
50	Malignant neoplasm of other sites	Cancer
51	Malignant neoplasm of brain	Cancer
52	Malignant neoplasm of Lymphatic and haematopoietic tissue	Cancer
53	Hodgkin's disease	Cancer
54	Leukaemia	Cancer
55	Myeloid Leukaemia	Cancer
56	Malignant neoplasm - multiple sites	Cancer