

# **Continuous Mortality Investigation**

## **Critical Illness Committee**

### **WORKING PAPER 33**

#### **A new methodology for analysing CMI critical illness experience**

*The views expressed in this working paper are those of the members of the CMI Critical Illness Committee and do not necessarily represent the views of the CMI, the Actuarial Profession or the employers of Committee members. This paper is a discussion document and the Committee fully acknowledges that there may be other valid approaches to analysing the CMI's critical illness data. This paper does not provide any form of guidance and should not be relied on as such. The CMI and the Committee members accept no liability whatsoever for any direct or consequential loss arising from any use of this working paper or its contents.*

July 2008

© 2008 Institute of Actuaries and Faculty of Actuaries

The text in this document may be reproduced free of charge in any format or medium providing that it is reproduced accurately and not used in a misleading context. The material must be acknowledged as Institute of Actuaries and Faculty of Actuaries copyright and the title of the document specified.

## EXECUTIVE SUMMARY

The background to CMI Working Paper 33 – “A new methodology for analysing CMI critical illness experience” is that:

- The CMI collects data on critical illness business on a calendar year basis. Given the often significant time-intervals between the date of diagnosis, when a critical illness claim is incurred, and the date of settlement, it would be impractical to wait for all the claims diagnosed in a particular calendar 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 volumes, 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. 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.
- Working Paper 14 also introduced the concept of a ‘grossing-up factor’ which sought to provide an overall adjustment to the reported experience.
- The method used to estimate grossing-up factors required estimates to be made of the growth in expected claims, which were necessarily relatively crude, and made it difficult to produce grossing-up factors for subsets of the data (e.g. by age and duration).

In July 2007, the Committee published Working Paper 28 which outlined revisions to the previous methodology to make better use of the data fields available and to reduce the uncertainty inherent in the previous approach. Our work since its publication is described in this paper. In summary:

- Working Paper 28 outlined the new methodology and described a fairly crude application to demonstrate the concept. We have now progressed to a more sophisticated implementation which is described in this paper.
- The claims data and the time-interval between diagnosis and settlement of a claim are discussed in this paper, and a refined model of the claim development distribution is introduced.
- Working Paper 28 introduced the concept of ‘off rates’, which are a necessary feature of the revised methodology. Further analysis of off rates is described in this paper.

We have used this methodology and our best estimates of the claim development distribution and off rates to produce results for accelerated critical illness experience on a lives basis for the years 1999-2004. These are the first results that we have calculated that properly match claims to exposure, but they do so in terms of settled claims, not diagnosed claims, and so need careful interpretation. Results for 1999-2002 are contained in this paper; results for 2003 and 2004 are being sent to member offices. The interpretation of these results is discussed in this paper and we also illustrate the sensitivity of the results to some of the key assumptions.

The Committee now intends to use the revised methodology to produce results for other subsets of critical illness business and to generate realistic claim rates.

This paper highlights a number of specific areas for feedback but comments on all aspects would be welcomed by the Committee.

# **Continuous Mortality Investigation**

## **Critical Illness Committee**

### **WORKING PAPER 33**

#### **A new methodology for analysing CMI critical illness experience**

#### **CONTENTS**

<b>EXECUTIVE SUMMARY</b>	<b>1</b>
<b>CONTENTS</b>	<b>2</b>
<b>1. INTRODUCTION</b>	<b>3</b>
<b>2. FEEDBACK ON WORKING PAPER 28</b>	<b>6</b>
<b>3. CLAIM DATES AND THE PROGRESSION TO SETTLED CLAIMS</b>	<b>9</b>
<b>4. THE INTERVAL FROM DIAGNOSIS TO SETTLEMENT: DATA</b>	<b>16</b>
<b>5. THE INTERVAL FROM DIAGNOSIS TO SETTLEMENT: MODELLING</b>	<b>21</b>
<b>6. THE INTERVAL FROM DIAGNOSIS TO SETTLEMENT: SENSITIVITIES</b>	<b>33</b>
<b>7. DEVELOPMENT OF THE METHODOLOGY</b>	<b>40</b>
<b>8. ANALYSIS OF OFF RATES</b>	<b>50</b>
<b>9. 1999-2002 RESULTS</b>	<b>54</b>
<b>10. ASSUMPTIONS AND SENSITIVITIES IN THE RESULTS</b>	<b>59</b>
<b>11. SUMMARY AND FURTHER WORK</b>	<b>65</b>
<b>REFERENCES</b>	<b>68</b>
<b>Appendix A: Defining the date of diagnosis</b>	<b>69</b>
<b>Appendix B: Reconciliation of results</b>	<b>72</b>
<b>Appendix C: Accelerated critical illness experience 1999-2002</b>	<b>76</b>

# **Continuous Mortality Investigation**

## **Critical Illness Committee**

### **A new methodology for analysing CMI critical illness experience**

#### **1. INTRODUCTION**

- 1.1. This paper presents the results of recent work by the CMI Critical Illness Committee, in particular with regard to modelling the development, between the date of diagnosis and the date of settlement, of critical illness insurance claims.
- 1.2. The work described in this paper builds on the methodology described in Working Paper 28. Having derived best estimates of the claim development distribution and off rates, we have produced results for accelerated critical illness experience on a lives basis for the years 1999-2004. These are the first results that we have calculated that properly match claims to exposure, but they do so in terms of settled claims, not diagnosed claims, and so need careful interpretation. Our intention is to progress to using the methodology to derive realistic claim diagnosis rates for UK critical illness business.
- 1.3. The CMI collects data on critical illness business on a calendar year basis. Given the often significant time-intervals between the date of diagnosis, when a 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 a calendar year but this results in a mis-match between the exposure and claims in that year. Given the substantial growth in business volumes, this mis-match is especially pronounced. The results that the CMI has released to member offices to date, providing ratios of actual settled claims to expected diagnosed claims, cannot therefore be considered a reliable guide to the true underlying experience. 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.
- 1.4. Working Paper 14 also introduced the concept of a 'grossing-up factor' which sought to avoid the understatement of the experience arising from comparing actual settled claims with expected diagnosed claims based on the exposure in the corresponding year. The CMI Critical Illness Committee published Working Paper 18, later in 2005, to document the feedback received on Working Paper 14; this 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.
- 1.5. The method used to derive grossing-up factors in Working Paper 14 required estimates to be made of 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. Furthermore, the methodology used to calculate the underlying claim development distribution was data-intensive, which also inhibited estimation of grossing-up factors for smaller subsets of the data.

- 1.6. In July 2007, the Committee published Working Paper 28 which outlined revisions to the previous methodology to make better use of the data fields available and to reduce the uncertainty inherent in the previous approach.
- 1.7. The Committee received a limited amount of feedback on Working Paper 28. This is summarised in section 2 of this paper, together with our responses.
- 1.8. The main area of further work indicated in Working Paper 28 was further analysis of the claim development distribution. This facilitated the use of the revised methodology to produce results that properly match claims to exposure for subsets of the data. The claims data and the time-interval between diagnosis and settlement of a claim are discussed in section 3. In section 4 we derive the set of claims data on which our subsequent modelling work is based. The modelling itself is then described in sections 5 and 6.
- 1.9. As noted in Working Paper 28, the initial application of the revised methodology described therein did not produce accurately calculated exposure. This required a more sophisticated implementation of the methodology, using actual dates of commencement where known. A more detailed implementation is described in section 7 of this paper.
- 1.10. Working Paper 28 introduced the concept of ‘off rates’, which are a necessary feature of the revised methodology. We have undertaken some further analysis of off rates since Working Paper 28 was published, in order to provide a better estimate of the off rates we should assume and to inform the choice of sensitivity test scenarios we have considered. This analysis is described in section 8.
- 1.11. The further work described in this paper has enabled the Committee to produce results for accelerated critical illness experience on a lives basis for the years 1999-2004. These results properly match claims to exposure, but they do so in terms of settled claims, not diagnosed claims, and so need careful interpretation. Summary results for 1999-2002 are contained in section 9 and in Appendix C. In addition, corresponding results for 2003 and 2004 are being sent to member offices.
- 1.12. Section 10 summarises the assumptions underlying these results and also includes results based on alternative estimates for the claim development distribution and off rates, to illustrate the sensitivity of the results to these assumptions.
- 1.13. The Committee intends to use the revised methodology to produce results for other subsets of critical illness business and to generate realistic claim diagnosis rates. Section 11 summarises the content of this paper then outlines the further work and investigations which the Committee now plans to undertake.
- 1.14. Finally, we draw readers’ attention to Appendix A which summarises our efforts to develop a better definition of the date of diagnosis of a critical illness claim. We invite

feedback on whether we should also seek to standardise the definition of date of notification for critical illness claims.

- 1.15. All feedback on this paper will be warmly welcomed by the CMI Critical Illness Committee. Details on providing feedback are also contained in section 11.

## 2. FEEDBACK ON WORKING PAPER 28

- 2.1. The Committee invited feedback on Working Paper 28 which outlined proposed revisions to the methodology we use for analysing critical illness experience. We received a limited amount of feedback which is summarised in this section, together with our responses.
- 2.2. Much of the feedback was received from others involved in the CMI's work and related primarily to technical aspects of our work. This was unsurprising given that Working Paper 28 focussed on the methodology we are pursuing, and that the revisions to the methodology had not then reached the stage of generating realistic results. Indeed some of the feedback recognised that the Committee was already seeking to address many of the points raised. Feedback has been collated under four headings:
  - Statistical modelling
  - Data issues
  - Bases for comparisons of Actual and Expected
  - Estimating 'off rates'

### *Statistical modelling*

- 2.3. Comments were received that the Committee has yet to set out a clear stochastic model for the critical illness claims process and that doing so would help to identify what could or should be estimated from the data. The Committee accepts this comment. Although we are seeking to move in this direction, we are constrained by the data that we have and the issues inherent therein, for example incomplete data on dates of diagnosis. As a result we continue to seek pragmatic means to make best use of the data we have and to overcome these practical issues. A clear stochastic model is therefore a goal for the Committee, but not our highest priority at the current time.
- 2.4. In particular it was suggested that one possible consequence of this lack of a clear model is reduced clarity over the measure of exposure referred to in paragraphs 4.16 and 4.22 of Working Paper 28. The response suggested that practice between offices as to when a policy is moved out of the in force may vary. Whilst we accept this is likely to be the case between admission and settlement of a claim, we suspect that few if any offices remove a policy from the in force on notification, due to the significant proportion of notified claims that are not settled. Hence we believe that, for the majority of claims, exposure will continue to accrue after the date of diagnosis. What is less clear is whether this is "inappropriate exposure", as the respondent suggested, if life offices continue to collect premiums during this period (and if they subsequently retain them, when the claim payment is made). Once we have progressed to producing realistic claim rates, this is an area we may consider further, perhaps surveying offices as the respondent suggested.
- 2.5. Comments were also received on the desirability of applying standard statistical methodology to the claim development distribution. This was our intention, and progress to date is documented later in this paper. We recognise that further work will also be desirable in this area.

*Data issues*

- 2.6. One response suggested that our analyses could be greatly improved if contributing offices were able to submit more complete data, in terms of on-risk dates and transition dates. This is undoubtedly true, and is part of the impetus behind the CMI's 'Per Policy' data initiative, however in the interim the Committee is seeking to produce the most meaningful analyses it can from the data available. The efforts of offices to supply data are greatly appreciated by the Committee, and we strongly encourage the adoption of 'Per Policy' data to allow more accurate and more meaningful future analyses.

*Bases for comparisons of Actual and Expected*

- 2.7. One response welcomed the move away from comparing Actual Settled Claims with Expected Diagnosed Claims. It was suggested that our previous approach, adjusting by means of grossing-up factors, was inappropriate as it sought to move towards Assumed Diagnosed Claims / Expected Diagnosed Claims and that the correct comparison is Actual Settled Claims with Expected Settled Claims. We adopted Actual Settled Claims / Expected Settled Claims for the results in Working Paper 28, and intend to continue to include these in future results. Actual Settled Claims are known soon after the event (subject of course to the usual time-lags in offices submitting data) and Expected Settled Claims can be estimated using our methodology from the known and synthetic in force data, which again is determined once we have the end year in force. We hope that a comparison of Actual Settled Claims with Expected Settled Claims will therefore provide a convenient means of issuing results to offices soon after data have been collected and validated, whilst recognising that such results will be vulnerable to any changes in the claim settlement process. Particular care may be needed in interpreting individual office results, where such changes may be more material.
- 2.8. However diagnosis (or undergoing of a qualifying surgical procedure) is the insurable event triggering claims under critical illness policies and is therefore very much the "actual" event we want to measure for experience investigation purposes. Hence we also see merit in seeking to analyse Actual Diagnosed Claims / Expected Diagnosed Claims. Furthermore, whichever way we analyse critical illness claims experience, our goal is to produce rates of claim diagnosis, not rates of claim settlement.
- 2.9. In time it may be possible to count Actual Diagnosed Claims accurately if we receive dates of diagnosis on (nearly) all claims and we can then analyse Actual Diagnosed Claims / Expected Diagnosed Claims, which we believe to be the most appropriate measure. This will necessitate estimating diagnosis dates on the small proportion of cases without them and allowing explicitly for an Incurred But Not Settled (IBNS) claims adjustment. Even then we see merit in continuing to also analyse Actual Settled Claims / Expected Settled Claims, as such results will be definitive, whereas Actual Diagnosed Claims will initially contain a significant IBNS adjustment and will be subject to subsequent revision.
- 2.10. Until we receive dates of diagnosis on (nearly) all claims we do see value in seeking to produce the best possible estimate of Actual Diagnosed Claims / Expected Diagnosed Claims and, for the foreseeable future, will seek to analyse both Actual Settled Claims / Expected Settled Claims and Actual Diagnosed Claims / Expected Diagnosed Claims.



- 2.11. We received one enquiry as to whether the new approach would mean delays before annual actual/expected results could be released to offices. This is certainly not our intention, and we anticipate we might release actual settled claims/expected settled claims results for year Y based on a stated table of claim rates and a stated claim development distribution derived from data from years preceding Y. This is little different from current CMI practice where, for example, Life Office Mortality results for years up to 2004 were released using “92” Series tables as a comparison; 2005 with comparisons against both “92” Series and “00” Series tables, and 2006 is expected to be compared with “00” Series tables only.
- 2.12. The added complication is that apparent divergences in observed critical illness experience may arise from either the number of settled claims or from a change in the claim development distribution. If, after further analysis, it becomes apparent that the claim development distribution has altered significantly from that used in preparing an initial set of results, then this can be re-visited.

*Estimating ‘off rates’*

- 2.13. The Committee’s use of off rates was welcomed in the responses we received. However there were a number of supplementary comments. These are considered below.
- 2.14. It was suggested that we had not clearly specified how these off rates varied, for example by age, duration and office. Working Paper 28 did state (in 4.13) that “The methodology would allow this to vary by office or duration, for example, but we have used a single assumption in our initial application...” We went on to indicate our reasons for this, namely that:
- “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” [of Working Paper 28].
- 2.15. We also indicated that we intended to undertake further analysis in this area. We have subsequently completed some further analysis of off rates and this is described in section 8 of this paper.
- 2.16. However we repeat the comment in Working Paper 28 that even if we find significant differences between off rates during the investigation period, we will then be applying these to earlier calendar years, so it is questionable whether the accuracy of the end results would be increased.

### 3. CLAIM DATES AND THE PROGRESSION TO SETTLED CLAIMS

- 3.1. Before we proceed to considering the claim development distribution, we first provide an overview of the data collected by the CMI on settled claims. Note that we are considering only claims settled in the period 1999-2004 ignoring both the small volume of earlier data, and that from 2005, which was incomplete when we started this analysis. In addition, we focus on the number of claims, not claim amounts. A review of the 1999-2002 data can be found in section 5 of Working Paper 14.
- 3.2. A total of 21,365 claims settled in 1999-2004 have been submitted to the CMI and included in the All Office results for those years.

#### *Claim dates*

- 3.3. The CMI requests 4 dates of claim for each settled claim submitted – date of diagnosis, date of notification, date of admission and date of settlement. These dates are not always received and in certain instances appear inconsistent.
- 3.4. Table 3.1 shows the number and percentage of claim records containing each of the four dates for the 1999-2002 quadrennium and separately for 2003 and 2004:

Table 3.1. Number and percentage of total claim records containing each date of claim. All 1999-2004 settled claims.

<b>Dates submitted by office</b>	<b>1999-2002</b>		<b>2003</b>		<b>2004</b>	
	<b>Number of claims</b>	<b>% of claims</b>	<b>Number of claims</b>	<b>% of claims</b>	<b>Number of claims</b>	<b>% of claims</b>
<b>Diagnosis</b>	6,649	56%	3,145	70%	3,789	75%
<b>Notification</b>	10,137	86%	4,052	90%	4,524	90%
<b>Admission</b>	3,928	33%	2,823	62%	3,405	68%
<b>Settlement</b>	10,394	88%	4,354	96%	4,853	96%
<b>Total</b>	<b>11,803</b>	<b>100%</b>	<b>4,519</b>	<b>100%</b>	<b>5,043</b>	<b>100%</b>

- 3.5. Note that the values for date of notification and date of admission for 1999-2002 in Table 3.1 differ from those in Table 1 in Working Paper 14. The earlier values were compiled after the date of notification was removed if it was equal to the date of diagnosis, and similarly for the date of admission. No adjustments have been applied to dates of claim in compiling the values above, other than the data checks undertaken at the time individual office data is processed.
- 3.6. It is clear that the CMI is receiving each date on an increasing proportion of the claims data. What is not clear from Table 3.1 is that this trend was also present in the 1999-2002 data: for example, the percentage of claims submitted with date of diagnosis increased from 37% in 1999 to 52%, 61% and 66% in 2000 to 2002 respectively.
- 3.7. For many offices, this percentage is not changing significantly; some offices have not submitted any dates of diagnosis and others submit them on all (or virtually all) of their

claims. Table 3.2 illustrates the position for the six largest offices and it will be observed that significant increases in the total (for these six offices) arise from office F beginning to submit dates of diagnosis on an increased percentage of claims in 2000 and office B joining the investigation in 2003. The increase from 2003 to 2004 arises from a further increase in the percentage of claims submitted with dates of diagnosis for office F and a greater proportion of the claims from these six offices arising from those submitting dates of diagnosis on a high proportion of claims.

Table 3.2. Percentage of total claim records containing date of diagnosis for the six largest offices (in terms of settled claims in 2004). All 1999-2004 claims.

<b>Office</b>	<b>1999</b>	<b>2000</b>	<b>2001</b>	<b>2002</b>	<b>2003</b>	<b>2004</b>
<b>A</b>	17%	15%	17%	15%	15%	17%
<b>B</b>	n/a	n/a	n/a	n/a	67%	75%
<b>C</b>	100%	100%	100%	100%	100%	100%
<b>D</b>	0%	0%	0%	0%	0%	0%
<b>E</b>	99%	98%	100%	99%	100%	100%
<b>F</b>	0%	78%	73%	70%	72%	87%
<b>Total</b>	<b>38%</b>	<b>55%</b>	<b>55%</b>	<b>58%</b>	<b>61%</b>	<b>69%</b>

- 3.8. In time, the Committee hopes that the proportion of claims with a reliable date of diagnosis will be sufficiently high to allow the use of 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. However in the interim, our methodology has been developed to make allowance for incomplete data on dates of claim.

*Progression to settled claims*

- 3.9. As noted previously, we are seeking to base our analysis of claims experience on dates of diagnosis, when claims are incurred. However, we collect data on the basis of settled claims so the time-interval between these two dates is of great significance to our analysis.
- 3.10. Whilst our modelling work, described in subsequent sections, focuses on the interval between the date of diagnosis and the date of settlement, we first consider the intervals between dates more generally. Note that within this section we illustrate the intervals using crude average time-intervals (although our modelling work seeks to reflect the complexities in the data, discussed in section 5).
- 3.11. In the analysis summarised below, in addition to the overall interval between diagnosis and settlement, we consider the separate intervals between diagnosis and notification, between notification and admission, and between admission and settlement. Whilst the latter stages of this progression can be assumed to be well-ordered – i.e. we presume a claim will not be admitted until it has been both diagnosed and notified, nor settled before it has been admitted – the date of diagnosis will not necessarily occur before the date of notification. This is due to ambiguities in the definitions of both the date of diagnosis and the date of notification.
- 3.12. The date of diagnosis could relate to the original diagnosis of the underlying disease, or to a later date when the definition specified within the critical illness policy is met.

Recognising this issue, the Committee launched an initiative in conjunction with the Health Claims Forum to draw up guidelines for claims assessors on what constitutes the 'date of diagnosis' for Critical Illness claims. This is described more fully in Appendix A and, in essence, we now regard the date of diagnosis to be defined as "**the date at which the critical illness definition was fulfilled**". More detail on the interpretation is included within Appendix A, but since companies were asked to adopt this guidance from 1 January 2007, no common standard was in place during the investigation period we are currently considering.

- 3.13. A similar ambiguity applies to date of notification, which could represent the date at which the insurer is originally notified of a (possible) claim or a later date at which the insurer is notified of the evidence required to admit the claim. The guidance developed with the Health Claims Forum did not extend to date of notification and, with hindsight, it would have been preferable to have clarified the interpretation of both dates at the same time. Again this is discussed further in Appendix A. Given that date of notification is not as important to our analysis as date of diagnosis, the Committee does not see standardisation of its interpretation as a key priority, but we welcome views on this.

*Policyholder or insurer?*

- 3.14. It is helpful to recognise that the interval between diagnosis and settlement can be considered to be comprised of two distinct periods: the first up to the date the policyholder/claimant notifies the insurer of a claim, and the second from then to the date when the insurer settles the claim. This makes intuitive interpretation of the total interval between diagnosis and settlement difficult. For example, claimants might be quicker to notify insurers for policies with higher sums insured, yet the insurer may scrutinise such claims more closely. Similarly, claimants might be quicker to notify insurers for policies that have recently been taken out, as they are more familiar with the cover, but again the insurer may scrutinise more closely claims diagnosed soon after a policy is issued.

- 3.15. We have not attempted any rigorous analysis of the overall interval in this way at this stage; but we have observed an interesting feature of those claims with particularly long intervals from diagnosis to settlement. Of the 100 claims where the interval from diagnosis to settlement exceeds 1500 days and where we also have a date of notification:

- The crude average interval from diagnosis to settlement is 2,078 days;
- The crude average interval from diagnosis to notification is 1,849 days; and
- The crude average interval from notification to settlement is 229 days.

I.e. for this subset of claims, the claim dates supplied to the CMI indicate that nearly 90% of the total interval appears to arise prior to notification by the policyholder/claimant. Indeed in only 7 of these 100 claims is the interval principally between notification and settlement.

- 3.16. In contrast, for the majority of claims (i.e. excluding these 100), the crude average interval between diagnosis and notification and the crude average interval between notification and settlement are remarkably similar, at 84 days and 82 days respectively.

*Intervals between dates of claim*

3.17. Table 3.3 shows the average observed intervals between various dates of claim, where we have them, and the volumes of data involved for the 1999-2002 quadrennium and for 2003 and 2004 combined.

3.18. Note that the values for 1999-2002 in Table 3.3 differ from those in Table 2 in Working Paper 14. As noted in paragraph 3.5, this arises due to elimination of some of the dates of claim in our earlier work.

Table 3.3. Crude average interval between various dates of claim (in days). All 1999-2004 settled claims.

<b>Pairs of Events</b>	<b>1999-2002</b>			<b>2003-2004</b>		
	<b>Average number of days between events</b>	<b>Number of records</b>	<b>% of records containing both dates</b>	<b>Average number of days between events</b>	<b>Number of records</b>	<b>% of records containing both dates</b>
<b>Diagnosis to notification</b>	97	5,738	49%	92	6,275	66%
<b>Notification to admission</b>	93	3,620	31%	79	5,898	62%
<b>Admission to settlement</b>	6	3,706	31%	15	6,220	65%
<b>Diagnosis to settlement</b>	176	5,404	46%	180	6,585	69%

3.19. Claims with diagnosis or notification dates on or after date of settlement have been excluded from further analysis in this section, but all other claims with relevant dates have been included. Slightly different criteria for inclusion were used in the analysis of the 1999-2002 data in Working Paper 14, leading to some minor differences in the tables below from those that appeared in that paper.

3.20. Two of the four time-intervals shown above demonstrate a reduction in length. Intervals between diagnosis and notification are five days shorter on average for the 2003-2004 data compared with that from 1999-2002 and the average interval between notification and admission has reduced by 14 days, but the interval between admission and settlement has increased by 9 days. The average for the overall interval between diagnosis and settlement, of primary importance to us, has increased by four days.

3.21. These changes are clearly not self-consistent; this is because the intervals are being measured for different subsets of claims, where we have the relevant dates. The relationships between the dates appear more consistent for 2003-4 than for the 1999-2002 data, probably due to the increased proportion of claims contributing to more than one of each of the intervals.

3.22. Table 3.4 shows these intervals for each calendar year. This suggests that the features noted above are not symptomatic of any underlying trend except, perhaps, for admission to settlement. These intervals are considered further below.

Table 3.4. Crude average interval between various dates of claim (in days) by calendar year. All settled claims.

	1999	2000	2001	2002	2003	2004	1999-2004
<b>Diagnosis to notification</b>	71	100	102	100	97	87	94
<b>Notification to admission</b>	98	99	83	101	77	81	85
<b>Admission to settlement</b>	1	10	4	7	12	17	12
<b>Diagnosis to settlement</b>	143	177	182	178	177	184	179

*Diagnosis to notification*

3.23. The interval between diagnosis and notification appears stable throughout 2000-2003.

The shorter average for 2004 is heavily influenced by two large offices:

- One office had an average interval between these two dates of 110 days in 2002, 111 days in 2003 and just 86 days in 2004. This is also the only office where date of notification occurs before date of diagnosis for some claims. This is feasible if date of notification is based on the original notification by the policyholder and date of diagnosis is interpreted as confirmation of a valid claim perhaps requiring permanence to be established.
- Another office started contributing data in 2003 and has shorter than average intervals between these two dates (an average of 62 days over 2003-4).

*Notification to admission*

3.24. The interval between notification and admission appears to have fallen over the period, with the interval of similar length in 1999, 2000 and 2002 and in 2001, 2003 and 2004. As with the interval between diagnosis and notification, this does not appear to be indicative of any overall trend in the claims data, but is heavily influenced by three large offices:

- One office that first started providing dates of admission in 2001 had the shortest average interval between notification and admission (an average of just 45 days across all four years, with little fluctuation).
- One office with a shorter than average interval length (69 days), first started providing dates of admission in 2003.
- Prior to 2003, many of the claims with both relevant dates had been provided by a single office. This office has a longer than average interval length throughout the period, although it did drop in 2001, which contributed significantly to the fall for all offices apparent in Table 3.4.

*Admission to settlement*

3.25. This is the shortest of the intervals being considered, at an average of just 12 days over 1999-2004 as a whole. There is however a marked increase in the interval in 2003 and again in 2004. Again this appears to be heavily influenced by two large offices:

- The office noted earlier that started contributing data in 2003 has a longer than average interval, of just over 40 days.
- Several offices show a lengthening of a few days between 2003 and 2004; the average for one particular office increased from 12 days in 2003 to 24 days in 2004.

3.26. In general there is great variation between the lengths of the admission to settlement interval. Whilst most are less than a week, there are many of over 50 days, and even

some of more than 100 days, which clearly inflate the average. Thus, although the magnitude of the average interval is small, quite large variations from year to year are not necessarily indicative of any trend. Death claims in particular show large variation in the length of this interval, perhaps whilst probate is established. Indeed the variations appear greater for single life policies than joint life - the average interval for death claims on single life policies is 35 days compared to 7 days for joint life policies.

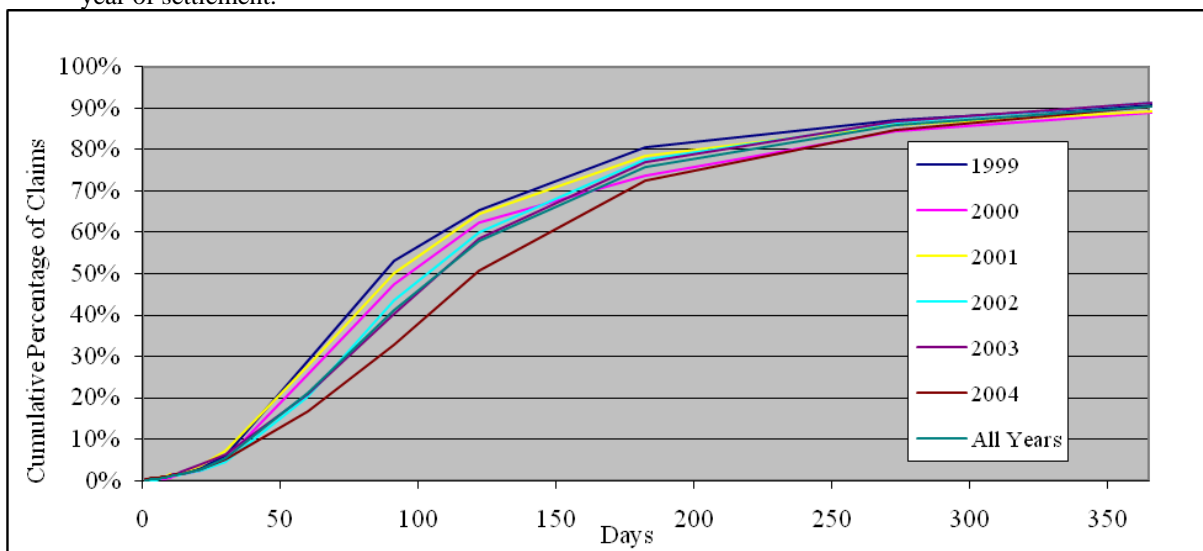
*Diagnosis to settlement*

3.27. The interval between date of diagnosis and date of settlement is the most important for our analyses. Both dates are available for 69% of the 2003-2004 claims, a significant increase from the 46% available from the 1999-2002 claims. The proportion of claims with both dates increased most rapidly from 1999 to 2001, though there has still been a steady increase since.

3.28. Table 3.4 shows that the average interval between diagnosis and settlement remained fairly constant throughout 2000-2004. However, further analysis suggests that considering only the average interval may mask a subtle trend, which is considered below.

3.29. Figure 3.1 illustrates the crude claim development distribution during the first year from date of diagnosis, by calendar year of settlement. There appears to have been a decrease in the rate of claim settlement between one month and six months over the period. However, after six months this reverses so that the proportion of the total claims settled within a year is very similar in each year of this period.

Figure 3.1. Crude claim development distribution during the first year from date of diagnosis, by calendar year of settlement.



3.30. The difference in the first year is offset by a gradual reduction over the period of claims with longer intervals, leaving the average interval from diagnosis to settlement relatively stable. The table below shows claims with an interval over two years as a percentage of all settled claims:

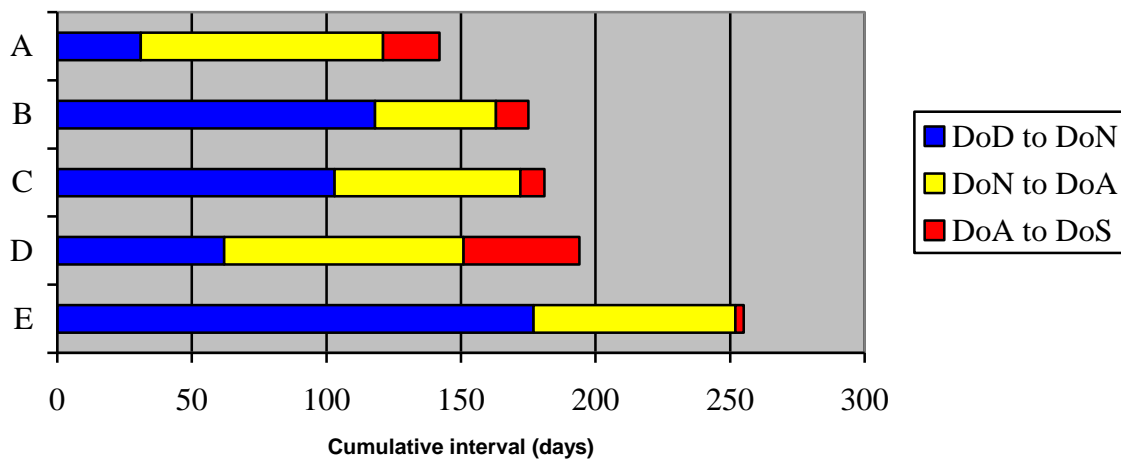
1999	2000	2001	2002	2003	2004	1999-2004
1.9%	4.4%	4.8%	3.6%	3.6%	2.8%	3.5%

3.31. Differences in the crude claim development distribution over the period are difficult to interpret as they may result from differences in the growth of claims and changes in the mix of business and offices.

*Variation between offices*

3.32. It will be apparent from the explanatory comments in previous paragraphs that there are significant variations between offices. Noting our earlier comments in paragraphs 3.14 to 3.16, these variations might arise from differences in policyholder behaviour or in processes and practices within offices, including interpretation of date of diagnosis and date of notification. They might also reflect differences in the underlying products, or the maturity of the portfolio. Figure 3.2 illustrates the differences in intervals for five large offices that supply dates of claim for a significant percentage of the claims they submit to the CMI.

Figure 3.2. Crude average cumulative interval to settlement (in days) for five selected offices.



Note that office A in Figure 3.2 is not necessarily office A in Table 3.2. The labels used for anonymised offices in subsequent sections of this paper are also not necessarily consistent



#### **4. THE INTERVAL FROM DIAGNOSIS TO SETTLEMENT: DATA**

- 4.1. In Working Paper 28, we highlighted the need to undertake further analysis of the time period between diagnosis and settlement. This need is not peculiar to the methodology developed in Working Paper 14 and Working Paper 28 but would also be required by a chain ladder-type approach, for example.
- 4.2. In section 5 we describe our initial attempts at modelling the development of claims in the interval between diagnosis and settlement, since that is the interval for which we need to estimate the distribution to apply within our methodology. This is referred to as the 'claim development distribution' in this paper.
- 4.3. First though we derive the set of claims data on which we undertake this modelling. We have chosen to use the entire set of claims data available to us, i.e. claims settled in 1999-2004, even though in subsequent sections of this paper we only show results for 1999-2002. The additional years contain more claims data than the original four and consequently we anticipated the six-year period would provide a more reliable basis for modelling the distribution. We consider the effect of only using claims settled in 1999-2002 data later in this paper.
- 4.4. For the purpose of analysing the claim development distribution we have focussed only on those claims where we have date of diagnosis and date of settlement. Removing the 9,376 records where we do not have both dates leaves 11,989 claims.
- 4.5. Note that in the analysis described in Working Paper 14, we also included claims where we had both the date of diagnosis and the date of admission (but no date of settlement), using date of admission + 5 days as a proxy for date of settlement. A similar approach now would add 82 further claims (prior to the adjustments described in the following paragraphs). We have not included such claims in this latest analysis as the additional claims are not significant (due to the addition of 2003 and 2004 data). In addition, the interval between admission and settlement appears to have increased for a number of offices from 1999-2002 to 2003 and then to 2004, as noted in section 3.25. This has increased the estimate of the adjustment required from admission to settlement, but also introduced a concern around the apparent variability in this interval which we had expected to be fairly stable.
- 4.6. The current application of our proposed methodology is limited to accelerated business. Since the claim development distribution that we derive will only be applied to expected diagnosed claims on accelerated business, and since the distribution may differ between accelerated and stand-alone business, we have restricted our attention to claims on accelerated business. This removes 1,565 claims on stand-alone critical illness policies, which we do not use further in this paper. This leaves us with 10,424 claims with both the date of diagnosis and the date of settlement on accelerated business.

##### *Claims with very short intervals between diagnosis and settlement*

- 4.7. One of the routine checks on critical illness data undertaken by the CMI when individual office data is processed is that the date of settlement is not before the date of diagnosis. Where a claim record fails this test, the dates are queried with the office

concerned. This sometimes results in one or other date being revised but often the office advises that the date of diagnosis is unreliable for that claim and should be deleted.

- 4.8. This particular check does not identify claims with very short intervals between diagnosis and settlement, and there are a number of such claims, including 17 where the two dates are equal, as shown in Table 4.1.

Table 4.1. Number of claims with intervals between date of diagnosis and date of settlement up to and including 14 days. All accelerated claims settled in 1999-2004 where both dates are submitted.

<b>Interval (days)</b>	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
<b>Number of claims</b>	17	10	9	4	3	13	18	10	15	20	10	13	16	17	18

- 4.9. The issue of the definition of date of diagnosis is considered further in Appendix A. Where the date of diagnosis has been interpreted as the date that permanence has been established, short periods from diagnosis to settlement may exist, but a period as short as 0 or 1 day seems highly unlikely. We have removed those claims with an interval from diagnosis to settlement of 0 days from our subsequent analysis, but retained all claims with intervals of 1 day or greater. This removes 17 claims, leaving 10,407 claims for analysis. The 17 claims are effectively treated in an identical manner to those where we have received a date of settlement but no date of diagnosis. We acknowledge that the removal of claims with an interval of 0 days, but retention of those with an interval of 1 day, is entirely arbitrary, however it is consistent with the approach taken for Working Paper 14.

*“Duplicates”*

- 4.10. The CMI has not received sufficient information to enable us to identify all multiple claims on one individual. Indeed if these claims arise on policies with different offices then it would not necessarily be appropriate to remove such “duplicates” anyway for the purposes of modelling the claim settlement process, as the dates of diagnosis and settlement, the date of commencement and even the cause of claim may all differ.
- 4.11. Nevertheless where we have clear examples of duplicates – which we define as an exact match on office, gender, date of birth, date of diagnosis, date of settlement and cause of claim – we have amalgamated these claims. In amalgamating these duplicate claims, our approach has been to add together the sums insured but otherwise to use the policy details (commencement date, smoker status, etc) from the latest record. (Within this paper, the way in which we have amalgamated these duplicates only affects the investigation of the sensitivity of the claim development distribution in section 6. Furthermore we have only adjusted for duplicates here and not in other areas such as exposure and claims in the results.)
- 4.12. Using this approach reduces the data available for analysis to 9,778 claims.

*Increasing proportion of claims with dates of claim supplied*

- 4.13. It has been noted in section 3.6 that the CMI has been receiving an increasing proportion of claims with the various dates of claim supplied. As illustrated there are some apparent step-changes in the proportion of claims for which we received both required

dates, sub-divided by office and calendar year. The treatment of these is considered later (paragraph 5.14) but, in addition, there are some subsets (by office and calendar year) where the proportion of claims with both dates was very low.

4.14. Our approach has simply been to exclude such claims from our subsequent work on modelling the interval between diagnosis and settlement, since there may be a material bias in the claims with both dates, where these are a low proportion of the total claims settled by that office in that calendar year. Such a bias might arise if the claims submitted only relate to one, or a small number of, causes of claim.

4.15. This approach is illustrated for two specimen offices in Table 4.2.

- There is a clear step-change for office A between 1999 and 2000. Given that there was a relatively small number of claims in 1999, those claims have been excluded (as explained in paragraph 4.14).
- There is also a clear step-change for office B between 2002 and 2003. Given that there were substantial numbers of claims in 1999-2002, those claims have not been excluded, but allowance has been made for the increase in proportion (as is explained later in paragraph 5.14).

Note that office B also illustrates the arbitrary nature of this approach, since we have not taken specific account of the increase from 63% of claims in 1999 to 76% in 2000, but would have done so had the proportion in 1999 been much lower.

Table 4.2. Percentage of accelerated claims with both date of diagnosis and date of settlement as a percentage of all accelerated claims by calendar year for two offices.

Office	1999	2000	2001	2002	2003	2004
A	31%	98%	100%	100%	100%	100%
B	63%	76%	83%	79%	100%	100%

4.16. This has led to the removal of a further 160 claims; including these might have introduced a bias, whereas their exclusion only reduces the claims available for modelling by around 1.6%.

### Summary

4.17. The derivation of the set of claims on which our modelling is based is summarised in Table 4.3.

Table 4.3. Summary of the derivation of the claims data used in modelling work in section 5.

<b>Total settled claims in 1999-2004</b>	<b>21,365</b>
Minus claims without date of diagnosis or date of settlement	9,376
Minus claims on stand-alone critical illness	1,565
Minus claims with date of diagnosis date equal to date of settlement	17
Minus "duplicates"	629
Minus groups of claims where a very low proportion have both dates	160
<b>Total claims on accelerated business settled in 1999-2004 used in subsequent modelling</b>	<b>9,618</b>

4.18. For completeness, we note that these claims include 41 where the smoker status was not advised to the CMI. In some other sections of this paper, the small amount of undifferentiated exposure and claims has not been included.

- 4.19. It is important to note that the subsequent modelling is only based on around 45% of the total claims, and around 52% of the claims on full acceleration business, submitted to the CMI. One of the key assumptions underlying the results presented later in this paper is that the claim development distribution derived from this subset of claims can be applied to the full dataset.
- 4.20. In some respects, the question of whether this subset is unbiased cannot be tested – the most obvious example being office. We have noted in section 3 some differences in the observed intervals between offices and we believe that such differences exist, perhaps because of differences in processes but, more significantly, because of the interpretation of date of diagnosis for some claims, such as those depending on permanence. However we can only observe such differences where we have the relevant dates of claim (for a significant proportion of claims). Where an office has not submitted any dates of diagnosis, say, then we have no information on claim development for that office and can do no better than assume an average distribution.
- 4.21. One bias within the subset that we have observed is with regard to cause of claim. Table 4.4 shows the percentage of settled claims on accelerated business in 1999-2004 for both the full dataset and the subset used for modelling for selected causes of claim.

Table 4.4. Percentage of settled claims on accelerated business in 1999-2004 by cause of claim in the full dataset and in the claims data used in modelling work in section 5.

Cause of claim	% of settled claims on accelerated business in 1999-2004	
	Full dataset	Modelling subset
<b>Cancer</b>	46.2%	52.9%
<b>Heart Attack</b>	11.3%	13.7%
<b>Stroke</b>	5.0%	6.0%
<b>CABG</b>	2.2%	2.4%
<b>MS</b>	4.5%	4.4%
<b>TPD</b>	3.4%	2.8%
<b>Death</b>	21.0%	11.3%
<b>Other (incl. unknown)</b>	6.4%	6.5%
<b>Total</b>	<b>100.0%</b>	<b>100.0%</b>

- 4.22. It will be observed that death, MS and TPD claims are all under-represented in the subset. As shown in Table 4.5, MS and TPD claims each has a lower than average proportion of dates of diagnosis submitted and we suggest that this is probably due to the uncertainty over what was meant by date of diagnosis for such claims.

Table 4.5. Percentage of all settled claims on accelerated business in 1999-2004 with date of diagnosis and with date of settlement, by cause of claim.

Cause of claim	% of all settled claims in 1999-2004 with:	
	Date of diagnosis	Date of settlement
<b>Cancer</b>	65%	99%
<b>Heart Attack</b>	69%	99%
<b>Stroke</b>	67%	99%
<b>CABG</b>	61%	99%
<b>MS</b>	55%	99%
<b>TPD</b>	46%	98%
<b>Death</b>	65%	63%
<b>Other (incl. unknown)</b>	62%	93%
<b>Total</b>	<b>64%</b>	<b>91%</b>

- 4.23. In contrast for death claims, the bias seems to be caused by a lower than average proportion of claims having a date of settlement. (Note that, as claims are submitted by year of settlement, we always know the year, even if we do not know the date, of settlement.)
- 4.24. The consequence of this under-representation of death, MS and TPD claims is that other causes, such as cancer, feature more prominently in the subset used for modelling than they should. The effect of this is considered later, in section 6.

## 5. THE INTERVAL FROM DIAGNOSIS TO SETTLEMENT: MODELLING

5.1. This section first sets out the features of the claims data that have influenced our approach to modelling. It then considers how these might be addressed by two approaches:

- direct calculation of the cumulative distribution; or
- fitting a parametric model.

5.2. We illustrate the results of these two alternative approaches by applying them to the set of claims settled in 1999-2004 arising from full acceleration business, from any cause, and on a “lives” basis – that is giving equal weight to each claim record irrespective of the claim amount. For this purpose, the claims records are restricted as outlined in section 4, principally due to the requirement for records to have both a date of diagnosis and a date of settlement, giving 9,618 usable claim records in total for this analysis.

5.3. As an initial working assumption, we disregard possible variations in the claim development distribution over time, and between subsets of the data, and we derive an aggregate distribution from the full available dataset. We consider this assumption again in section 6.

### *Adopted Notation*

5.4. Suppose we group reported claim dates into a series of uniform calendar time periods and let  $C[i,j]$  be the number of reported claims diagnosed in calendar period  $i$  and settled in calendar period  $j$  (where  $j \geq i$ ). Hence:

- The total reported diagnosed claims in a period I is

$$\sum_{i \in I} \sum_{j=i}^{\infty} C[i,j]$$

- The total reported settled claims in a period J is

$$\sum_{j \in J} \sum_{i=0}^j C[i,j]$$

5.5. Also let:

- $T$  be the number of time periods between diagnosis and settlement;
- $F(T)$  be the cumulative proportion of claims settled within  $T$  time periods after diagnosis; and
- $f(T)$  be the probability of a claim being settled exactly  $T$  time periods after diagnosis.

Note that although it may often be the case for distributions that  $F(0)=0$ , this is not the case in paragraphs 5.6 to 5.26 where the chosen time period is a calendar month so that some claims will be settled in the same time period in which they are diagnosed.

5.6. Further suppose, for an exposure period I and an unbiased subset of business, we have complete claims data – that is all claim events with diagnosis dates in period I have been reported and settled. We could then derive the cumulative distribution of the interval from diagnosis to settlement directly as:

$$F(T) = \frac{\sum_{i \in I} \sum_{j=i}^{j=i+T} C[i,j]}{\sum_{i \in I} \sum_{j=i}^{\infty} C[i,j]}$$

5.7. We could approximate the probability density function in a similar way:

$$f(T) = \frac{\sum_{i \in I} C[i, i+T]}{\sum_{i \in I} \sum_{j=i}^{\infty} C[i, j]}$$

or we could fit a parametric model to the data to obtain smooth density and distribution functions.

*Crude Estimation of the Claim Development Distribution*

5.8. We have previously referred to the results of a version of this calculation, simplistically applied to the recorded data without regard to the completeness requirement, as the “crude” or “observed” claim development distribution.

5.9. As the claims data are characterised by having settlement dates in the observation period J, this crude calculation actually delivers:

$$F_{crude}(T) = \frac{\sum_{j \in J} \sum_{i=j-T}^{i=j} C[i, j]}{\sum_{j \in J} \sum_{i=0}^{i=j} C[i, j]}$$

and

$$f_{crude}(T) = \frac{\sum_{j \in J} C[j-T, j]}{\sum_{j \in J} \sum_{i=0}^{i=j} C[i, j]}$$

summed over the recorded claims.

*Modelling problems inherent in the form of the claims data*

5.10. There are, however, a number of issues inherent in the CMI critical illness claims dataset which frustrate this methodology in addressing the true underlying claim development distribution.

5.11. The first three of these, described in paragraphs 5.12 to 5.14, are closely related and lead us to introduce the concept of the ‘Effective Observation Period’ for each office or portfolio of business:

5.12. We do not, in practice, have a complete set of claims data for any exposure period I because claims are only reported to the investigation on the basis of settlement dates falling within the observation period J (1999 to 2004, for the purposes of this analysis):

- For claim events with dates of diagnosis before the observation period J, some may have been settled also before period J and therefore not included in the reported claims dataset.
- Some claim events with dates of diagnosis before or during the observation period J may not have been settled by the end of period J and therefore such “open” claims also will not be included in the reported claims dataset.

Thus we cannot calculate the denominator, nor sometimes even the numerator, for the equation in 5.6. Even where we do know the number of claims settled in the period T, we cannot directly calculate the rate of settlement because we do not know the number of “open” claims “exposed to settlement” at that point. We must therefore modify our methodology to work only with the subset of claims “truncated” on both sides by the restriction to settlement dates within the observation period.

5.13. A further complication arises from changes in the data contributed within the overall observation period,  $J$ .

- Where an office ceases to submit data after time period  $j_{end}$  in  $J$ , then for that office we will have settled claims reported in cells  $[i, j_{end}]$  but not in  $[i, j_{end}+1]$ . The Effective Observation Period for the office is not the overall period  $J$ , but is limited by the earlier cut-off at  $j_{end}$ .
- Where an office starts to submit data in time period  $j_{start}$  in  $J$ , then for that office we will have settled claims reported in cells  $[i, j_{start}]$  but not in  $[i, j_{start}-1]$ . The Effective Observation Period for the office is limited by the later cut-in at  $j_{start}$ .

Thus the Effective Observation Period varies from office to office and we must also allow for this, reflecting the “truncation” of the reported claims data at the boundaries of actual observation for each office.

Such changes in data submission were referred to as ‘fault lines’ in Working Paper 14. In many cases, data changes occur not because an office starts or ceases to submit data, but because it changes (usually increases) the portfolios covered. Suppose an office previously submitted data for portfolio A but then starts to submit for portfolios A and B:

- If the CMI can identify the claims in subsequent submissions by portfolio, then portfolio A could be used in the claim development distribution analysis throughout and portfolio B could be treated as if it were a new office.
- If the CMI cannot segregate the claims by portfolio in this way, then portfolio A has to be regarded as having left the investigation and portfolio (A+B) treated as a new office for the claim development distribution analysis.

In practice, additional portfolios have only been separated out if the CMI is able to distinguish the original and additional portfolios in both the claims data and the in force data; in all other cases a change in business covered has been treated as if the original office left the investigation and was replaced by a new office.

5.14. The final complication arising from the form of the claims data is that the subset of claims with both date of diagnosis and date of settlement is an increasing proportion of the total claims over the period, as noted in section 3. Thus, similar to the issues posed by changes in data submitted, we may not have a consistent underlying dataset to allow valid comparison of cells  $[i, j+1]$  and  $[i, j]$ .

In theory we might allow for this by replacing  $C[i, j]$  with  $C[i, j] \div fn(j)$ , where  $fn(j)$  is the proportion of claims settled in period  $j$  for which we receive both required dates. However, inspection of the data suggests  $fn(j)$  would be difficult to work with as it is not a smooth or continuous function, would vary by office, and may also be dependent on the diagnosis period,  $i$ . In addition, where  $fn(j)$  is low there is likely to be material bias, for example by cause of claim, in the mix of claims in the subset with both required dates compared to the full set of claims.



We have therefore adopted a simpler approach. Considering the proportion of claims settled for which we received both required dates, sub-divided by office and calendar year, some step changes were apparent. These can be treated identically to fault lines in 5.13 above - that is, we treat claims settled before and after the step change as separate portfolios of business for the purpose of deriving the claims development distribution, restricting the Effective Observation Period as appropriate to each consistent subset of reported claims. In addition, the few subsets with a very low proportion of claims reported with both dates were simply excluded completely from the analysis because of the likelihood of material bias (see 4.13).

- 5.15. A separate issue that we also have to contend with is the low numbers of claims with long intervals between diagnosis and settlement, and cannot know the ultimate length of the distribution. This inevitably leads to considerable uncertainty and estimation error in modelling the tail of the distribution. This issue is exacerbated when we seek to estimate the claim development distribution for subsets of the full claims dataset, which necessarily reduces the numbers of claims available to us.

*The Working Paper 14 approach*

- 5.16. Our original approach to deriving the underlying claim development distribution was outlined in Working Paper 14 and is described here using the notation developed above.

- 5.17. Suppose, as in 5.7 above, we have complete claims data for an exposure period I and an unbiased subset of business. We could then directly calculate the probability of a claim being settled exactly T periods after diagnosis as:

$$f(T) = \frac{\sum_{i \in I} C[i, i+T]}{\sum_{i \in I} \sum_{j=i}^{\infty} C[i, j]}$$

- 5.18. Unfortunately, as noted in 5.12 the form of the data does not allow us to calculate the denominator of this equation. However, if the claims data are complete and consistent within the observation period J – that is all claim events with diagnosis dates in some period  $i$  and settlement dates in period J have been reported – then we can calculate the numerator provided  $i+T$  falls within the observation period J. We can therefore also calculate  $R(T)$ , the ratio of the probabilities of a claim being settled T+1 periods and T periods after diagnosis:

$$R(T) = \frac{f(T+1)}{f(T)} = \frac{\sum_i C[i, i+T+1]}{\sum_i C[i, i+T]}$$

where the sum is over all periods,  $i$ , for which  $i+T$  and  $i+T+1$  both fall within the observation period, J.

- 5.19. We can then combine these ratios to build an estimate of the cumulative distribution for the time-interval, T, from diagnosis to settlement using:

Define  $S(0) = 1$

and

$$S(T) = \prod_{t=0}^{T-1} R(t) \quad \text{for } T > 0$$

then

$$f(T) = \frac{s(T)}{\sum_{t=0}^T s(t)}$$

and

$$F(T) = \sum_{t=0}^T f(t)$$

5.20. The essential requirements for this approach are that:

- The reported claims data are consistent and complete within the observation period J and diagnosis periods used; and
- There is sufficient data so that  $\sum_i C[i, i+T] > 0$ , and therefore  $R(T) > 0$ , for all intervals T between diagnosis and settlement.

We therefore need to arrange or modify the calculations to address the problems noted in paragraphs 5.12 to 5.14 such that these conditions are met.

5.21. We must restrict the calculations to respect the boundaries of the Effective Observation Period for which we have consistent and complete claims data. This is best illustrated by an example. Suppose we have collected reported claims data in respect of settled claims in time periods t to t+3. The claims data could be tabulated as shown below by time period of diagnosis and interval to settlement:

Period of Diagnosis	Number of time periods from diagnosis to settlement				
	0	1	2	3	4
t-4					C[t-4, t ]
t-3				C[t-3, t ]	C[t-3, t+1]
t-2			C[t-2, t ]	C[t-2, t+1]	C[t-2, t+2]
t-1		C[t-1, t ]	C[t-1, t+1]	C[t-1, t+2]	C[t-1, t+3]
t	C[t , t ]	C[t , t+1]	C[t , t+2]	C[t , t+3]	
t+1	C[t+1, t+1]	C[t+1, t+2]	C[t+1, t+3]		
t+2	C[t+2, t+2]	C[t+2, t+3]			
t+3	C[t+3, t+3]				
t+4					

For the set of claim events with date of diagnosis in a particular time period, we can see a portion of their claim development distribution by looking along the row for the relevant time period. However, the observed claim counts cut a diagonal path across the table and, in combining the data to calculate  $R(T)$ , we must ensure we make only valid comparisons by staying within the area of the observations. For example:

$$R(0) = \frac{C[t,t+1]+C[t+1,t+2]+C[t+2,t+3]}{C[t,t]+C[t+1,t+1]+C[t+2,t+2]}$$

and

$$R(1) = \frac{C[t-1,t+1]+C[t,t+2]+C[t+1,t+3]}{C[t-1,t]+C[t,t+1]+C[t+1,t+2]}$$

Note  $C[t-1,t]$  cannot be used in calculating  $R(0)$  because its left-hand partner  $C[t-1,t-1]$  lies outside of the Effective Observation Period; it can however validly be used in calculating  $R(1)$ .

Similarly,  $C[t+2,t+3]$  is validly used in calculating  $R(0)$ , but cannot be used in calculating  $R(1)$  because its right-hand partner  $C[t+2,t+4]$  lies outside of the Effective Observation Period.

This approach, grouping the available claims data by diagnosis period and time to settlement and then selecting appropriate matched pairs of cells for the calculation of  $R(T)$  resolves the problem noted in 5.12

- 5.22. We can deal in a similar way with the fault line issues noted in 5.13 and 5.14. By restricting the Effective Observation Period for each office or portfolio of business, such that the reported claims reflect an unbiased subset of claims arising from consistent exposure across the observation period, we can ensure none of the pairs of cells used in calculating the  $R(T)$  are made invalid by crossing any fault line or inconsistency in the data.
- 5.23. In practice, we chose to group the claims data by calendar month of diagnosis date and by number of calendar months interval from diagnosis to settlement. Grouping by calendar periods is convenient because the edges of the observation period are marked by calendar year-end boundaries. Any fault lines, through offices or portfolios of business entering or leaving the investigation, or through step changes in the proportion of claims reported with both the required dates, also lie along calendar year-end boundaries.
- 5.24. The issue noted in 5.15 of rapidly falling data volumes, as the interval between diagnosis and settlement lengthens, leads to empty cells and the problem of some  $R(T) = 0$ . We addressed this by grouping period to settlement into quarter-years or even half-years (measured from calendar month of diagnosis date) to calculate the  $R(T)$  ratios for intervals greater than 6 months.
- 5.25. Finally, the issue of unknown ultimate length of the claim development distribution was resolved pragmatically. The approach outlined above was used for intervals up to 5 years only and then run on for a further 5 years by simple extrapolation; the distribution was assumed complete 10 years after diagnosis date. The extrapolation assumed an exponential run down of  $f(T)$  over years 6 to 10, fitted to the pattern derived from the data for years 3 to 5.
- 5.26. Sample points from the claim development distribution derived with this methodology, applied to the set of claims settled in 1999-2004 and arising from full acceleration business, from any cause, and on a lives basis, are shown in table 5.1. Also shown for comparison are the equivalent values from Working Paper 14 when this methodology was applied to data for claims settled in 1999-2002 (see chart in paragraph 5.5.8 in Working Paper 14).

Table 5.1. Cumulative percentage of claims settled by selected time-intervals (in months) from month of diagnosis to month of settlement derived using the approach set out in Working Paper 14 and applied to claims settled in 1999-2004 and claims settled in 1999-2002.

<b>Claim Development Distribution: <math>F(T)</math>=cumulative percentage settled at time-interval <math>T</math>, where <math>T</math> = interval (in months) from diagnosis month to settlement month</b>										
<b>T</b>	<b>0</b>	<b>1</b>	<b>2</b>	<b>5</b>	<b>8</b>	<b>11</b>	<b>17</b>	<b>23</b>	<b>35</b>	<b>59</b>
<b>1999-2004</b>	2.0	12.5	29.6	67.0	79.6	85.3	91.1	93.9	97.0	99.3
<b>1999-2002</b>	1.2	11.5	31.2	65.9	76.4	81.6	87.3	90.4	94.9	98.3

Notes:

(1) For Working Paper 14 the claims dataset contained records from both full acceleration and stand alone critical illness business, from any claim cause. Calculations were also on a “lives” basis. The methodology applied now is a refinement of that used for Working Paper 14, particularly in respect of smoothing and extrapolating the  $R(T)$  ratios for  $T > 3$  years.

(2) The chart from Working Paper 14 was repeated in paragraph 4.26 of Working Paper 28; these values are consistent with that chart. However, they do not reconcile with the table in paragraph 4.37 of Working Paper 28, which shows erroneous interpolated figures for intervals from 9 months. These erroneous values were then used to produce the indicative results shown in Working Paper 28.

*The rationale for seeking an alternative, parametric model approach*

5.27. The Working Paper 14 approach provided us with a realistic model of the underlying claim development distribution for all claims combined. However we were concerned the sparsity of data when claims were sub-divided by risk factors might have prevented us from using this approach to derive claim development distributions for subsets of the portfolio, which was expected to be a necessary step for us to be able to produce realistic results, and would certainly be required to produce realistic claim rates by cause of claim.

Furthermore, this approach only yielded estimates of the cumulative distribution function,  $F(T)$ , at particular intervals determined by the way in which we grouped the intervals from diagnosis to settlement: i.e. calendar months for the first half-year, then by quarter-year or half-year; all measured from the calendar month (not exact date) of diagnosis. We could “join the dots” by linear interpolation or more realistic curve fitting, but an alternative methodology which delivered an easier to use estimate of the complete claim development distribution would be desirable.

5.28. As indicated in Working Paper 28, we were therefore keen to explore alternative approaches to deriving underlying claim development distributions.

5.29. There are a number of features of parametric modelling and the associated estimation procedures, that made this approach potentially very attractive, if a suitable model could be found and fitted:

- Most models will provide a smooth, well-behaved distribution, fitted to the claims data.
- Once the model has been fitted, it will provide a unique probability of settlement associated with every interval from the exact date of diagnosis.

- The entire distribution can be easily summarised, in terms of the parameter values.
- Most models would also provide statistical properties such as a mean and variance for the fitted parameters, which allow us to generate probability distributions for the model.
- The goodness of fit of the model to the data can be measured by conventional statistical tests.
- We should be able to fit the model even where we have limited data (although of course the confidence intervals around the fit will be wider). This is potentially particularly useful for smaller subsets of the data.
- We may be able to interpret the parameters so that, for example, increasing the value of a particular parameter might imply a slower rate of settlement in the early months.
- If the parameters alter smoothly over time, then the model may have some predictive power, so that we can estimate the claim development distribution for future periods.

5.30. Note that the Working Paper 14 approach could be extended to provide us with some of these features, but not all; hence the potential attraction of a parametric model.

#### *Fitting a parametric model*

5.31. We assume the distribution of the time-interval,  $T$ , between diagnosis and settlement can be adequately described by a relatively simple model specified by a formula with a small number of parameters. We have used the maximum likelihood approach to derive parameter values which optimise the fit of the modelled probability density function,  $f(T)$ , to the observed “probability of settlement at time  $T$ ”, where  $T$  is measured, in whole days, from the date of diagnosis.

5.32. The numerator for the observed probability of settlement at a time-interval  $T$  days after diagnosis, is simply the number of claims (within a particular subset) that are reported settled at that interval. The calculation of the denominator (the number of claims exposed to settlement) is however complicated by the problems noted in 5.12:

- An adjustment for left censoring is required because some of the dates of diagnosis fall before the observation period and hence these claims have been “exposed to settlement” during that prior period (see 5.12).
- An adjustment for right censoring is required because some claims with dates of diagnosis before or during the observation period are not settled in the period and hence are “exposed to settlement” in subsequent years (see 5.12).
- Further adjustments for left and right censoring are required because of changes in the data submitted at office level, restricting the Effective Observation Period for each office or portfolio of business (see 5.13).
- Likewise, further adjustments for left and right censoring are required because of step changes in the proportion of claims reported with both date of diagnosis and date of settlement (see 5.14).

5.33. We take account of data being left and right censored by adjusting the likelihood function. The appropriate adjustment follows from noting that we wish to estimate  $f(t_k)$ , the probability that  $T = t_k$ , but we are observing the conditional probability  $f'(t_k)$  given

that, for the Effective Observation Period, we cannot have observed the event before  $a_k$  or after  $b_k$  (as defined below). We therefore need to divide the unconditional probability  $f(t_k)$  by the probability of the event falling within this interval. Hence we adjust the normal likelihood formula (L) as follows:

$$L'(\alpha, \beta, \dots) = \prod_{k=1}^n f'(t_k; \alpha, \beta, \dots) = \prod_{k=1}^n \left[ \frac{f(t_k; \alpha, \beta, \dots)}{F(b_k) - F(a_k)} \right]$$

Where:

- $f(t)$  is the probability density function for the model we are fitting i.e. the time-interval between diagnosis and settlement;
- $F(t)$  is the corresponding cumulative distribution function;
- $t_k$  is the observed interval from diagnosis to settlement for each claim  $k$ ;
- $a_k$  is the interval, for each claim  $k$ , from its date of diagnosis to the point at which we first have the potential to observe settlement of the claim;
- $b_k$  is the interval, for each claim  $k$ , from its date of diagnosis to the point after which we no longer have the potential to observe settlement of the claim;
- $\alpha, \beta$ , etc are the parameters of the model; and
- The product is over all settled claims reported in the (summed) Effective Observation Periods.

5.34. A left censoring adjustment is applied to set the values of  $a_k$  and counts the number of days between the date of diagnosis and the start of the Effective Observation Period (the later of the start of the investigation period and when that office or portfolio started submitting data on a consistent basis). No left censoring adjustment is required if the date of diagnosis is after the start of the Effective Observation Period, so for many claims  $a_k = 0$ . For each claim  $k$ ,

$$a_k = \max\{\text{Start of Effective Observation Period} - \text{date of diagnosis}, 0\}$$

5.35. A right censoring adjustment is applied to set the values of  $b_k$  and counts the days from the date of diagnosis to the end of the Effective Observation Period (the earlier of the end of the investigation period and when the office or portfolio stopped submitting data on a consistent basis). For each claim  $k$ ,

$$b_k = \{\text{End of Effective Observation Period} - \text{date of diagnosis}\}$$

5.36. These censoring adjustments reduce the denominator used for each claim record to the proportion of claims we would expect to be settled within the Effective Observation Period for the business underlying the claim. Equivalently, we can express this as weighting each observation  $f(t_k)$  to allow for other claims expected to have been excluded from the analysis because their date of settlement falls outside of the Effective Observation Period.

5.37. A clear correspondence can be drawn with the Working Paper 14 approach. The censoring boundaries  $a_k$  and  $b_k$  for each claim are where the “fault lines” (including observation period boundaries), applying to the office or portfolio of business from which the claim arises, cut across the time-interval from the claim’s date of diagnosis.

5.38. To facilitate the mathematical process of solving for the parameters, we make a common transformation and seek to maximise the log-likelihood function:

$$\ln L'(\alpha, \beta, \dots) = \sum_{k=1}^n \ln f'(t_k; \alpha, \beta, \dots)$$

5.39. Note, it is unclear how effectively this approach may handle the issue flagged in 5.15. Whilst a well-chosen model may give us some intuitive confidence over the shape of the tail of the distribution, the model's fit to the sparse data at long intervals after diagnosis may be questionable. In particular, note that the adjustment of dividing by  $F(b_k) - F(a_k)$  becomes large at the tail of the distribution and may make the fitting process there somewhat circular.

*The Burr distribution*

5.40. The Committee has investigated a limited number of models to apply to the claim development distribution and has chosen to work with the Burr distribution for this Working Paper. Whilst the Burr model appears to provide a reasonable fit to this aggregate data, we do not claim that it is necessarily the best model. Neither do we claim that it achieves all the desirable features of a parametric model we set out in paragraph 5.29 and the Committee welcomes suggestions of other possible models for consideration.

5.41. The Burr model, as we have used it, is a 3-parameter model with the following probability density function:

$$f(t) = \frac{\alpha\gamma\lambda^\alpha t^{\gamma-1}}{(\lambda + t^\gamma)^{\alpha+1}}$$

and cumulative distribution function:

$$F(t) = 1 - \left(\frac{\lambda}{\lambda + t^\gamma}\right)^\alpha$$

5.42. In simple terms, these three parameters give flexibility over:

- The peak rate of settlement ( $\alpha$ ),
- When this peak rate occurs ( $\lambda$ ), and
- The thickness of the tail ( $\gamma$ ).

5.43. Using this methodology, applied to the set of claims settled in 1999-2004 and arising from full acceleration business from any cause and on a lives basis, as derived in section 4, results in a Burr model of the claim development distribution with the parameters shown in Table 5.2. Note that in this and subsequent applications of the Burr model in this paper we have rounded the parameters to 4 decimal places ( $\alpha$  and  $\gamma$ ) or the nearest integer ( $\lambda$ ).

5.44. Table 5.2 also shows 95% confidence intervals for the parameters. The distributions derived using the upper and lower bounds for alpha and gamma are inverted, i.e. the lower parameter values, calculated using  $(\hat{e} - 1.96\sigma)$ , produce a distribution with longer intervals to settlement than the higher parameter values. This is because the higher parameter values produce a larger number of claim settlements for the same input, in terms of number of days.

Table 5.2. Parameter values for the Burr model of the claim development distribution for the set of claims settled in 1999-2004 that was derived in section 4. Parameter values are shown for the best fit and for 95% confidence intervals around the parameter values.

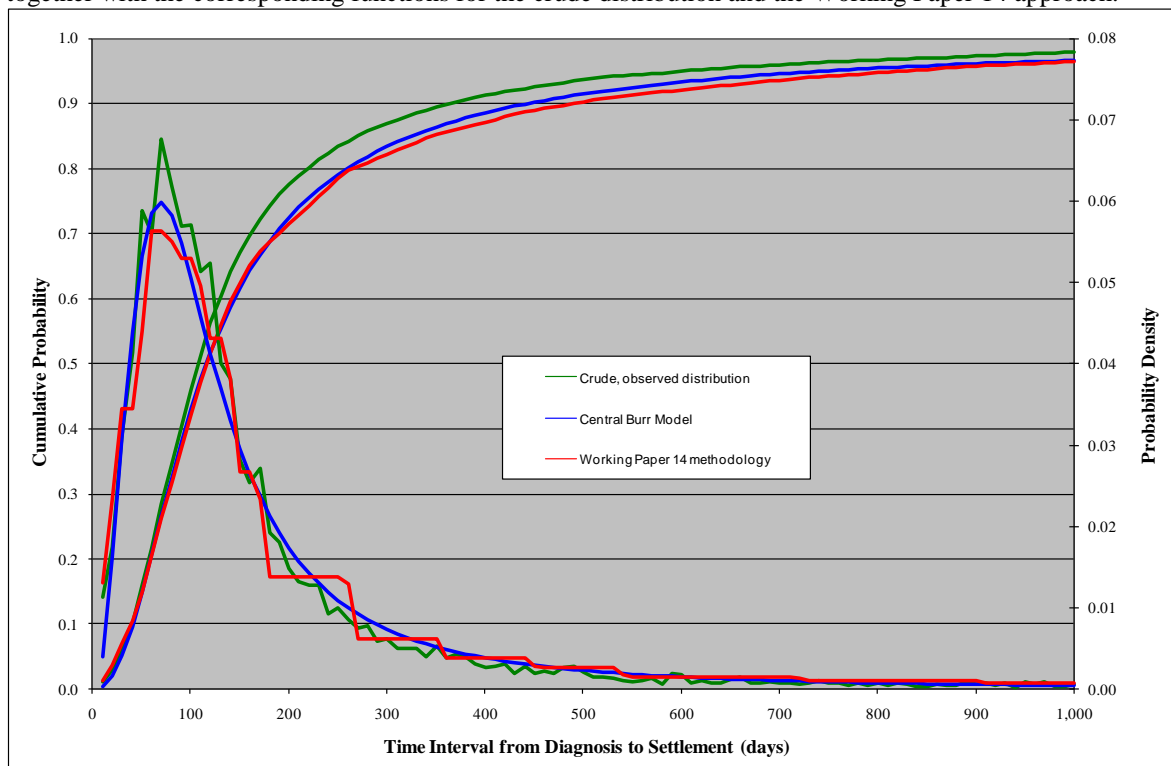
<b>Burr model of claim development distribution</b>			
	$\hat{\epsilon} - 1.96\sigma$	$\hat{\epsilon}$ (‘central’)	$\hat{\epsilon} + 1.96\sigma$
$\alpha$	0.5026	<b>0.5574</b>	0.6121
$\lambda$	21,904	<b>33,856</b>	45,807
$\gamma$	2.2832	<b>2.3852</b>	2.4872

5.45. In subsequent use of these distributions in this paper we refer to the distribution using the best fit assumptions as the ‘central’ distribution, to the distribution using the higher parameter values for alpha and gamma and the lower parameter value for lambda as the ‘short’ distribution, and to that using the lower values for alpha and gamma and the higher value for lambda as the ‘long’ distribution. Note that the short and long distributions are derived using 95% confidence intervals for the parameter values, but carry no exact probabilistic interpretation in relation to the claim development distribution itself.

*Comparison of the results*

5.46. Figure 5.1 shows a comparison of the estimates of the claim development distribution as derived through the two alternative approaches described in this section, along with the crude, observed distribution (see 5.7 paragraph 5.8). Both the cumulative distribution and probability density functions are shown.

Figure 5.1. Probability density function and cumulative distribution function for the central Burr model of the claim development distribution for the set of claims settled in 1999-2004 that was derived in section 4, together with the corresponding functions for the crude distribution and the Working Paper 14 approach.





Note that:

- For the distribution derived from Working Paper 14 methodology, linear interpolation has been assumed between the estimated points of the cumulative distribution function. This is a simple and convenient assumption, but not a necessary or realistic one.
- The probability density is evaluated over time-intervals of 10 days.

5.47. A simpler, practical way of comparing these distributions is to consider how they would combine with an assumed underlying rate of growth in claims to affect the reported experience. Using Working Paper 14 terminology, we calculate and compare the grossing-up factors which would result from these claim development distributions if applied to business with 25% p.a. underlying growth in claims.

Table 5.3. Implied grossing-up factors for three models of the claim development distribution assuming 25% p.a. growth in underlying claims for the set of claims settled in 1999-2004 that was derived in section 4.

<b>Model</b>	<b>Implied grossing-up factor</b>
Working Paper 14 methodology	12.97%
Central Burr Model	12.75%
Crude, observed distribution	11.02%

5.48. We have chosen this 25% p.a. underlying growth in claims to illustrate the implied grossing-up factors, as that was our estimate of the approximate growth rate for the 1999-2002 quadrennium of CMI critical illness data. The overall implied grossing-up factors, using claims data for 1999-2004, are still close to the 15% we estimated in Working Paper 14.

5.49. It is clear from Figure 5.1 and Table 5.3 that the two alternative approaches give very similar estimates for the overall claim development distribution. The difference in the grossing-up factors that each implies is not significant in the context of interpreting overall critical illness claims experience. The difference between these two estimates of the underlying distribution and the crude distribution is rather more significant.

#### *Truncation of the Burr distribution*

5.50. As a practical measure to achieve a finite distribution and thereby ease the calculation burden in using the Burr model within our revised methodology, we have chosen to truncate the distribution. This avoids the need to apply a very long tail (beyond that indicated by the data) and consequently a long back-projection period for the exposure. We have modified the modelled distribution by applying linear interpolation to the fitted cumulative distribution over the period from the end of year 3 to the end of year 7 from diagnosis. This is a shorter period than used in Working Paper 14; Table 5.4 shows that this practical measure does not cause any significant distortion overall.

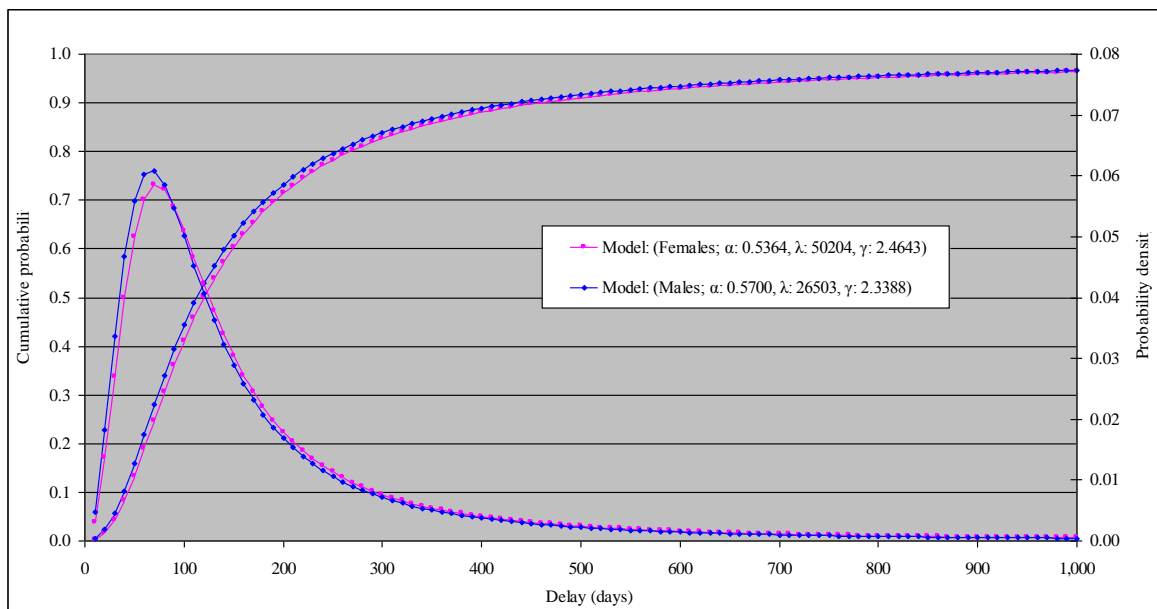
Table 5.4. Implied grossing-up factors, assuming 25% p.a. growth in expected claims for the Burr model without truncation and for selected truncation patterns.

<b>Truncation pattern</b>	<b>Implied grossing-up factor</b>
Burr Model without truncation	12.75%
Truncation; linear years 3 to 7	12.83%
Truncation; linear years 4 to 6	12.69%
Truncation at end year 5	12.67%
Truncation; linear years 3 to 5	12.54%

## 6. THE INTERVAL FROM DIAGNOSIS TO SETTLEMENT: SENSITIVITIES

- 6.1. In this section we describe the results of fitting the parametric model to subsets of the data. In doing so, we are seeking to demonstrate that it is reasonable at this stage of our work to use a single claim development distribution and that this is unlikely to substantially distort the results. We do not however claim (or seek to demonstrate) that this is an optimal approach, nor that there may not be subtle differences that emerge from subsequent work. In particular we believe that such differences introduce no greater uncertainty into our results than other assumptions that we are required to make, due to incomplete data.
- 6.2. In section 10 we consider the sensitivity of the 1999-2002 results to alternative claim development distributions. These do not relate to particular subsets of the data, derived in the following paragraphs, but we do compare the distributions for the subsets with those later used in the sensitivity tests towards the end of this section.
- 6.3. The first separation of the data was by gender. It was not thought that any significant difference in fit would occur between male data and female data, unless there was some other underlying reason, e.g. a difference in the claim development distribution between causes of claim that have greater prevalence for males or females.
- 6.4. Figure 6.1 appears to indicate that there is little variation between the two fitted curves.

Figure 6.1. Probability density function and cumulative distribution function for Burr models of the claim development distribution for males and females in the set of claims settled in 1999-2004 that was derived in section 4.



- 6.5. Comparison of the parameter values in Table 6.1 also supports the hypothesis, as most of the parameters for each of the fits lie inside the confidence intervals of the other. In addition all but one of the parameter values sit within the confidence intervals of the central distribution shown in Table 5.2.

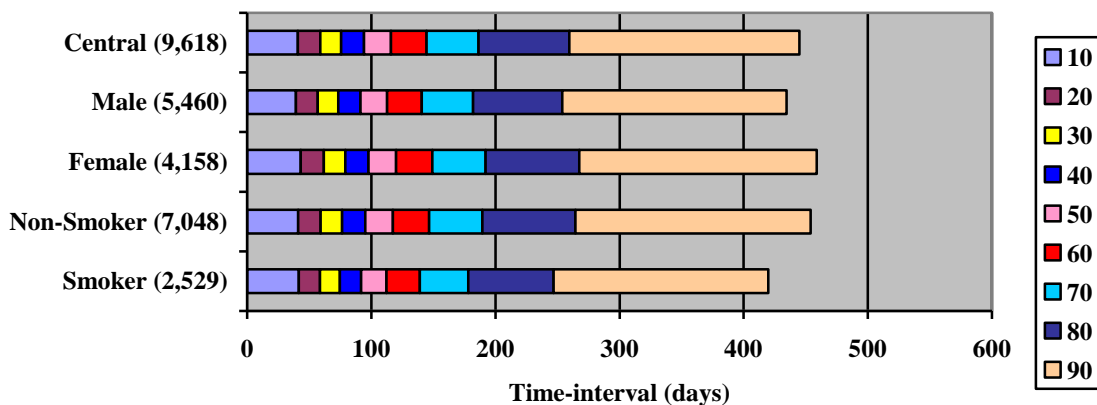
Table 6.1. Parameter values for the Burr model of the claim development distribution for the male and female subsets of the claims settled in 1999-2004 that was derived in section 4. Parameter values are shown for the best fit and for 95% confidence intervals.

	Male			Female		
	$\hat{e} - 1.96\sigma$	$\hat{e}$	$\hat{e} + 1.96\sigma$	$\hat{e} - 1.96\sigma$	$\hat{e}$	$\hat{e} + 1.96\sigma$
$\alpha$	0.4955	0.5700	0.6446	0.4559	0.5364	0.6168
$\lambda$	14,867	26,503	38,140	20,067	50,204	80,340
$\gamma$	2.2097	2.3388	2.4680	2.2949	2.4643	2.6336

6.6. Figure 6.2 compares the Burr models for males and females with the central distribution for all claims specified in Table 5.2. The chart shows the cumulative time-intervals from diagnosis to settlement for each decile of claims to be settled using each of the Burr models of the claim development distribution.

6.7. Note that in Figure 6.2 (and similar figures later in this section) only deciles up to 90% are shown, since the cumulative distribution function tends to, but never actually reaches, 100%; i.e. all the comparisons are of the unadjusted fitted distributions, with no truncation of the tail, as discussed in paragraph 5.50. In order to provide an indication of the credibility of the Burr models applied to the subsets of the data, Figure 6.2 (and subsequent diagrams) shows the number of claims on which each model is based in brackets. These can be compared to the total number (of 9,618 claims) derived in Table 4.3.

Figure 6.2. Cumulative time-intervals from diagnosis to settlement for each decile of claims to be settled using Burr models of the claim development distribution for all claims and for males and females and for non-smokers and smokers in the set of claims settled in 1999-2004 that was derived in section 4. Figures in brackets are the numbers of claims on which each model is based.



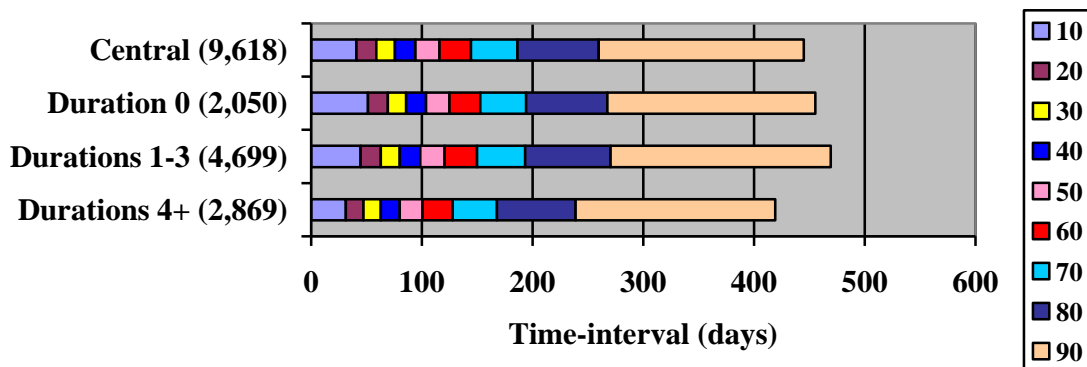
6.8. Figure 6.2 also compares the Burr models for non-smokers and smokers and, as with males and females, there is little variation between the distributions. In addition, all four of these variations sit comfortably between the short and long distributions at most points, the only exception being that the female distribution exhibits a slightly slower settlement rate than the long distribution until about 60 days.

6.9. We have considered various other subsets of the data. Some such as age (separated between ages 40 or under, 41 to 50 and 51+, where age is defined as age nearest at date of diagnosis) and single or joint life also show little variation and are not included here.

6.10. Other subsets that are considered below are duration, calendar year and benefit amount.

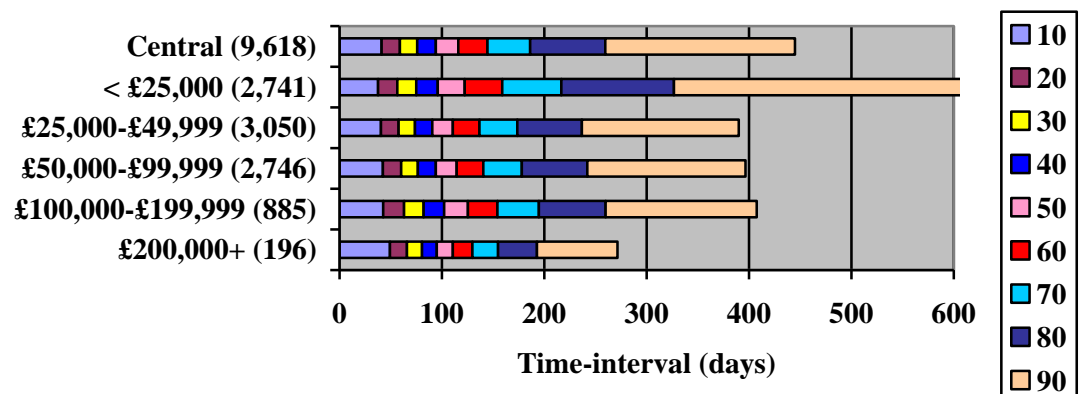
6.11. The deciles for the Burr models fitted by duration are shown in Figure 6.3. Duration has been separated between duration 0, durations 1 to 3 combined and durations 4+, based on observed differences in the crude distribution (with duration defined as curtate duration at date of diagnosis). Although the overall difference in the distributions is small, the first decile appears to reduce with increasing duration, as one might expect if offices scrutinise short duration claims more closely.

Figure 6.3. Cumulative time-intervals from diagnosis to settlement for each decile of claims to be settled using Burr models of the claim development distribution for all claims and by duration in the set of claims settled in 1999-2004 that was derived in section 4. Figures in brackets are the numbers of claims on which each model is based.



6.12. We also anticipated possible variation by benefit amount, reflecting both claimants' desire for the sum insured and insurers' desire to investigate larger claims more fully.

Figure 6.4. Cumulative time-intervals from diagnosis to settlement for each decile of claims to be settled using Burr models of the claim development distribution for all claims and by benefit amount in the set of claims settled in 1999-2004 that was derived in section 4. Figures in brackets are the numbers of claims on which each model is based.

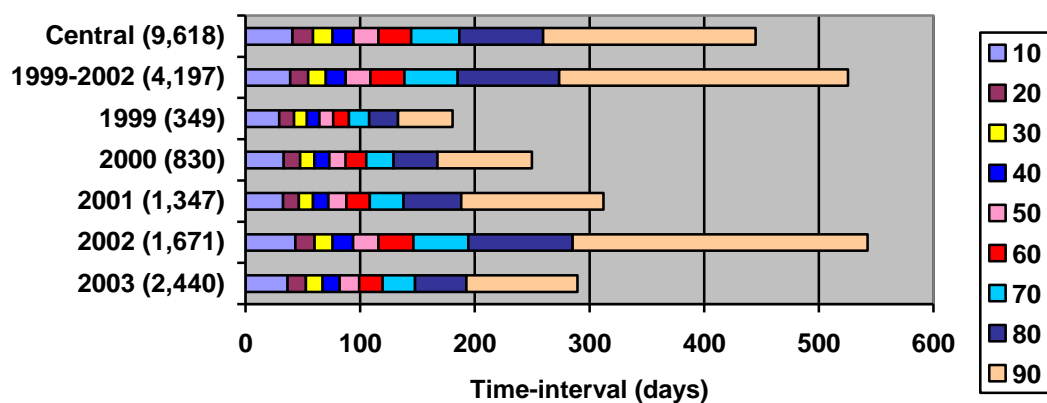


6.13. The deciles for the Burr models fitted by benefit amount are shown in Figure 6.4, with amounts separated into five bands, as shown. There appears to be little difference in settling around the first 50% of claims, but thereafter the smallest claims appear to take longest to settle, and the largest claims the shortest.

6.14. Figure 6.5 shows the corresponding values for the Burr models fitted by calendar year of settlement. In addition to the individual years, values for the distribution based on 1999-2002 combined are also shown.

6.15. There appears to be a gradual lengthening of the claim development distribution from 1999 to 2001, followed by a substantial lengthening in 2002. The distribution for 2003 is then much closer to that for 2001. Notwithstanding the increasing numbers of claims underlying these distributions by year, the distribution for 1999-2002 does not appear, from Figure 6.5, to be an average of the separate years.

Figure 6.5. Cumulative time-intervals from diagnosis to settlement for each decile of claims to be settled using Burr models of the claim development distribution for all claims and by calendar year in the set of claims settled in 1999-2004 that was derived in section 4. Figures in brackets are the numbers of claims on which each model is based.



6.16. Note also that these fits include all the relevant claims settled in each particular year, and no adjustment has been made in respect of offices joining or leaving the investigation.

6.17. These results appear spurious, as they arise in the fits obtained from the Burr model but not in the crude distributions (see Figure 3.1) nor in values produced using the Working Paper 14 approach. We believe this may indicate the unsuitability of a 3-parameter model for some subsets of the data. In particular, within some subsets, such as by calendar year, we have little data at long intervals to settlement and these may be exerting more influence on the fit of the Burr model at shorter intervals than is desirable.

6.18. Furthermore note that no values are included for 2004, as the software used failed to find a fit for the 2004 data. There was no obvious reason for this, but we have not yet investigated this further. These issues may necessitate consideration of different techniques for the tail from the main distribution in future work.

6.19. Despite the differences noted above, we believe that it is reasonable for us to use a single claim development distribution at this stage of our work, where the distribution is assumed to apply across all of the investigation period. In making this assertion it is important to note:

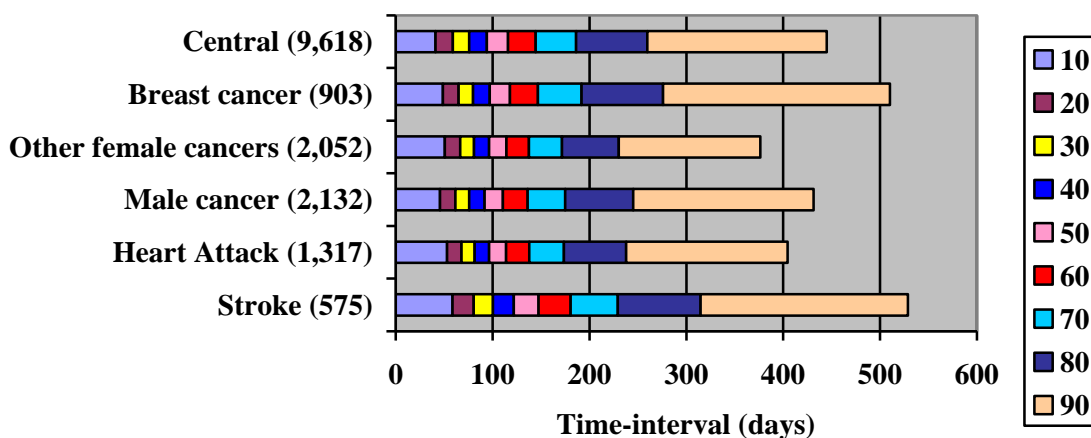
- The context in which the claims development distribution is being used. This was described in Working Paper 28, and is explained again from paragraph 7.47; and
- The sensitivity of the results to different fits of the claim development distribution, considered in section 10.

6.20. All the comparisons above have been undertaken at an “all causes” level and for all offices combined; we consider variations by these factors in the following paragraphs.

6.21. We have previously noted significant differences in the crude claim development distribution between the major causes of critical illness claims and therefore anticipate differences in the underlying distribution too. Whilst we consider that the Burr model provides a reasonable fit to the claim development distribution at an all causes level, it does not necessarily provide a good model for each of the causes individually. In particular, the crude distributions for death and TPD have quite different shapes and we have struggled to fit the Burr model to these subsets of the data. It is likely that we will need to investigate other models to pursue work at an individual cause level.

6.22. Figure 6.6 shows the cumulative time-intervals from diagnosis to settlement for the Burr model fitted to selected causes of claim, where it does provide a reasonable fit. Heart attack and male cancers show a very similar fit (to each other and to all causes) throughout; other female cancers show a similar pattern. Breast cancer claims show a similar pattern until around 50% are settled – subsequent claims are settled more slowly – whilst stroke claims appear to take longer to be settled throughout, presumably because of the need to establish permanence of neurological deficit.

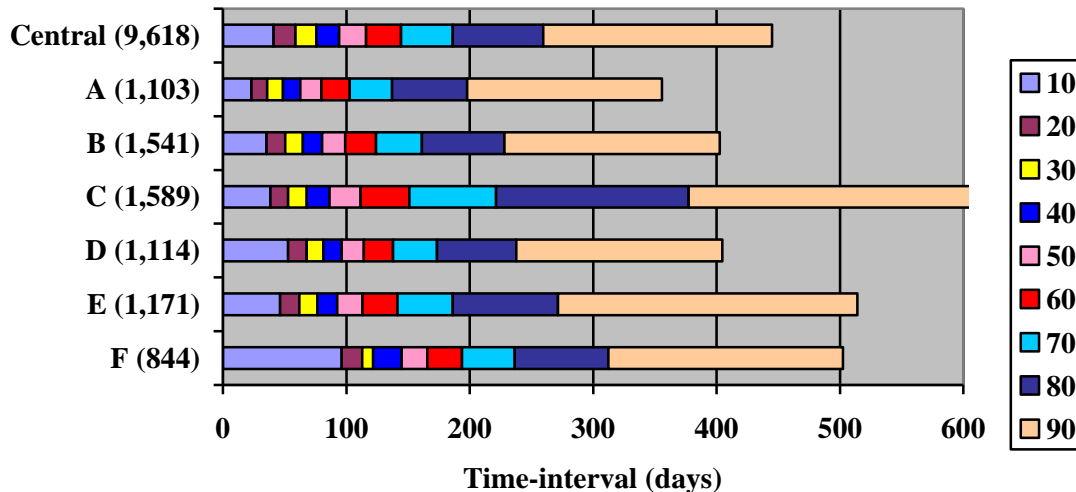
Figure 6.6. Cumulative time-intervals from diagnosis to settlement for each decile of claims to be settled using Burr models of the claim development distribution for selected causes of claim in the set of claims settled in 1999-2004 that was derived in section 4. Figures in brackets are the numbers of claims on which each model is based.



6.23. We noted variations between offices in the crude time-intervals between diagnosis and settlement in paragraph 3.32 and therefore anticipated significant differences in the claim development distributions also. Figure 6.7 shows cumulative time-intervals for six of the larger offices within the 1999-2004 dataset. Considerable variation emerges, not only in the overall pace of settlement, but in different areas of the claims. For example,

based on the fitted Burr models, office C has a very similar settlement pattern to that for “all claims” until around 60% of claims are settled, but then has a much slower rate of settlement, so that 80% of claims are only settled after 938 days, which is off the scale of this chart. In contrast, office F settles very few claims within a short period, taking close to 100 days to settle just 10% of claims.

Figure 6.7. Cumulative time-intervals from diagnosis to settlement for each decile of claims to be settled using Burr models of the claim development distribution for selected offices in the set of claims settled in 1999-2004 that was derived in section 4. Figures in brackets are the numbers of claims on which each model is based.



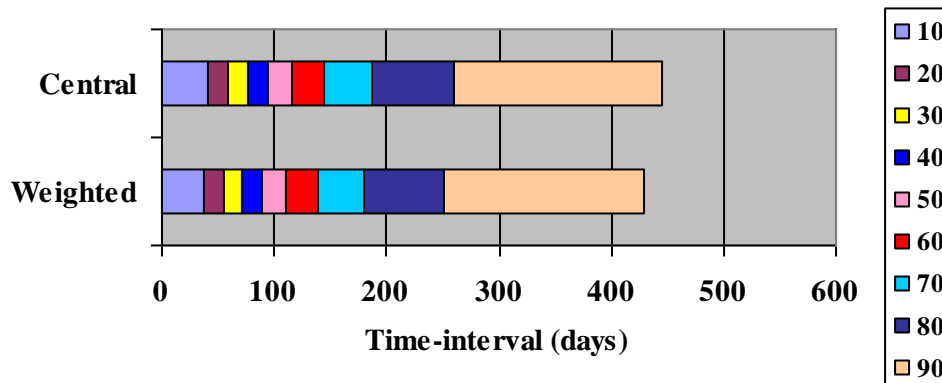
6.24. These variations by office appear significant, and actuaries using this methodology should consider whether it is reasonable to use the all offices distribution. We believe it is reasonable for us to use it to produce all offices results and, in particular, have no real alternative with regard to the assumptions used for offices who have not submitted any dates of claim.

*Choice of claim development distribution for use in application of the methodology*

6.25. We have chosen to use the central Burr distribution fitted to the dataset derived in section 4 to derive results for 1999-2002 in section 9. However as noted in paragraph 4.21, there is a difference in the mix of claims by cause between that dataset and all accelerated claims in 1999-2004. We have not made any adjustment in this regard as, conveniently, the under-statement of death claims and the under-statement of TPD claims act in opposite directions and hence tend to offset each other.

6.26. Figure 6.8 illustrates the insensitivity of the overall claim development distribution in this regard, as it compares the overall distribution with an adjusted distribution in which the weight afforded to each death claim is more than doubled, to achieve a similar proportion within the modelled dataset to that in the full accelerated critical illness claims dataset. It will be observed that the additional weighting given to death claims shortens the distribution, but we consider that the difference is certainly no greater than that arising from other assumptions within our work.

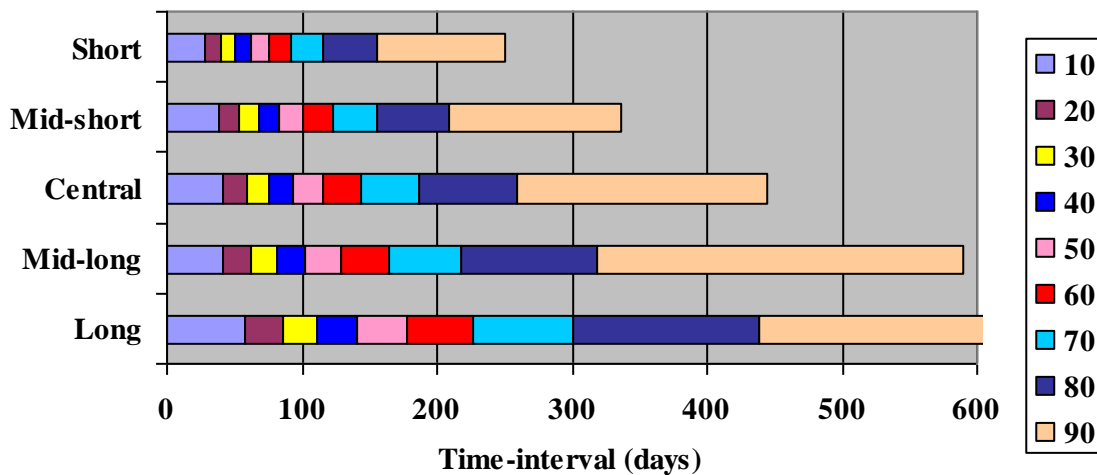
Figure 6.8. Cumulative time-intervals from diagnosis to settlement for each decile of claims to be settled using Burr models of the claim development distribution for the set of claims settled in 1999-2004 that was derived in section 4 and for a set of claims with additional weight given to death claims.



*Claim development distributions used for sensitivity tests*

6.27. As noted earlier, we consider the sensitivity of the 1999-2002 results to alternative claim development distributions in section 10. The distributions used are illustrated in Figure 6.9. The short and long distributions were derived using the 95% confidence intervals around the parameters, as described in paragraph 5.45. As these are relatively extreme scenarios, we also use intermediate distributions which we have termed ‘mid-short’ and ‘mid-long’.

Figure 6.9. Cumulative time-intervals from diagnosis to settlement for each decile of claims to be settled using the central Burr model of the claim development distribution and four variations of the claim development distribution



6.28. These four scenarios do not relate to particular subsets of the data, but are intended to reasonably cover the range of distributions that have been derived for subsets within this section. They will be considered again in section 10 when we look at the sensitivity of the 1999-2002 results to alternative claim development distributions.



## **7. DEVELOPMENT OF THE METHODOLOGY**

- 7.1. Working Paper 28 contained a description of proposed revisions to the methodology we use for analysing critical illness experience. As outlined in Working Paper 28, known in force data were used to estimate the exposure that will give rise to settled claims during the period for which we have data (referred to as the 'investigation period'). This exposure can be multiplied by a set of claim rates to produce expected diagnosed claims and a claim development distribution applied to transform these expected diagnosed claims into expected settled claims, which we can compare to the known settled claims.
- 7.2. The application illustrated in Working Paper 28 was implemented using spreadsheets and involved some generalising assumptions, for example that dates of commencement occur in the middle of a calendar year. We have since developed a more robust application which we believe generates more accurate exposure. In this section we describe the differences between these two applications, which include the use of actual dates of commencement of policies and a day-count calculation of exposure.
- 7.3. Some approximations have however been retained in this application to avoid excessive run-times; we believe the most significant of these is that we have only modelled duration in quarter-years.

### ***Exposure during the Investigation Period***

- 7.4. In the initial application, the exposure during the investigation period was calculated using a census approach, averaging the start and end in force. Given the growth in business that occurred throughout much of 1999-2004, this may overstate the exposure and introduce distortions by age and duration.
- 7.5. In the revised application we have used the actual dates of commencement, which are known for all data submitted to the CMI. Note that this not only affects the amount of exposure in the first year of a policy, but also affects the allocation of exposure by duration in subsequent calendar years.
- 7.6. The CMI critical illness investigation uses an age definition of "age nearest" and curtate duration; within this application we group exposure by year of age and by quarter of duration (Note: the reason for splitting duration by quarter is explained later, in paragraph 7.51). Within these calculations we calculate the number of days' exposure but retain this by calendar month, for ease of computation; we do not believe this results in any significant loss of accuracy, compared to retaining exposure by calendar day.
- 7.7. Note that in summing the exposure within each category of age, quarterly duration and calendar month, we retain office and other risk factors such as gender, smoker status, sales channel and benefit type, although the split by sales channel and benefit type is not considered further within this paper.

### ***New business***

- 7.8. Where a policy has curtate duration 0 at the end-year in force, exposure is calculated from the actual date of commencement to the end of that year. For example, consider a record that is in force on 31/12/1999 with:

- Date of birth 13/5/1969 and
- Date of commencement 4/10/1999.

This record is assumed to generate exposure by age (nearest) and (curtate) duration as shown in Table 7.1.

Table 7.1. Exposure generated in 1999 by a record with a date of birth of 13/5/1969 and a date of commencement of 4/10/1999.

Calendar Month	Age Nearest	Curtate Duration	Exposure (days)
October 1999	30	0	28
November 1999	30	0	11*
November 1999	31	0	19
December 1999	31	0	31

\* The date of birth is 13/5/1969, and the move from age nearest 30 to age nearest 31 is assumed to occur on 12/11/1999, i.e. 183 days after the birthday, so that the life is treated as age 30 nearest on 11/11/1999 and age 31 nearest on 12/11/1999.

7.9. This calculation is undertaken for every record at duration 0 in every set of end-year in force data (i.e. from 31/12/1999 to 31/12/2004 in respect of 1999-2004 data).

#### *Non-new business*

7.10. The approach adopted for other policies is slightly different and is intended to allow for lapses and reinstatements (as well as policies that remain in force throughout). It can be summarised as follows:

- For policies which exit during a calendar year we do not have the date of exit. Note that the CMI has only recently begun to collect data on policies exiting the investigation before the subsequent year-end (under the Per Policy initiative – see the CMI website for more details). This information is not available for the data relating to years up to 2004.
- Matching policies between successive year ends is at best problematic, as we do not always have consistent policy identifiers.
- Given the above, one approach to calculating exposure would be a traditional census approach, in which we could assume that all movements occurred midway through a calendar year. Exposure in respect of lapses during a calendar year would be allocated to the first half of that calendar year.
- However, for any given policy, the split of exposure by age, duration and calendar year would then be biased; a policy reaching duration  $t$  after the mid-year point would have no exposure allocated to duration  $t$ , for example. At best one might hope that different biases would cancel out when exposures were summed over policies.
- We have therefore used an alternative method, which assumes that the probabilities of exit are equal for each day of the year, with exposure being derived and allocated accordingly. This method avoids the need to match records but nonetheless allocates exposure by age, duration and calendar year without bias, unlike the census approach.

7.11. The following description, which we base on calendar year 1999, is intended to illustrate the approach. For each record in the 1/1/1999 in force data, we assume that a

day's exposure is generated for 1/1/1999, and we use the actual date of commencement and date of birth to allocate this day's exposure to the appropriate age and duration.

- 7.12. The records in the 1/1/1999 in force data may relate to policies that remain in force throughout 1999 or to policies that leave the investigation during 1999 (through lapsation, maturity, etc). From the 1/1/1999 data alone, we do not know into which of these categories any particular record falls. Hence we cannot be sure whether or not exposure is generated in respect of 2/1/1999 and subsequent days, so we ratio down the exposure accordingly and allocate  $364/365$ ths of a day's exposure for 2/1/1999,  $363/365$ ths of a day's exposure for 3/1/1999,  $362/365$ ths of a day's exposure for 4/1/1999, etc down to  $1/365$ th of a day's exposure for 31/12/1999.
- 7.13. A similar process is followed in respect of records in the 31/12/1999 in force data in relation to 1999, without seeking to take account of whether policies have remained in force throughout 1999 or have been reinstated during 1999 (new business having been addressed earlier – see 7.8 above). From the 31/12/1999 data alone, we do not know into which of these categories any particular record falls. Hence we allocate  $364/365$ ths of a day's exposure for 31/12/1999,  $363/365$ ths of a day's exposure for 30/12/1999,  $362/365$ ths of a day's exposure for 29/12/1999, etc down to  $1/365$ th of a day's exposure for 2/1/1999.
- 7.14. This approach to calculating the exposure from the start- and end- in force should give the same outcome as if we matched records between the two datasets, given that we do not have the date of exit (or date of reinstatement), but is computationally easier.
- 7.15. For a policy that does feature in both the start- and end-in force data, note that the effect of summing the exposure from the two calculations is a day's exposure for each day in the year. Exits and reinstatements are assumed to be equally likely to occur on any day throughout the year, which is an appropriate assumption in the absence of information on the exact dates; this weights the exposure for exits towards the start of 1999 and the exposure for reinstatements towards the end of 1999.
- 7.16. The following examples hopefully illustrate the calculation of exposure for the calendar year 1999. Consider a record that exists in the in force data for both 1/1/1999 and 31/12/1999 with:
- Date of birth 13/5/1969 and
  - Date of commencement 1/7/1997.

This record is assumed to generate the following exposure by age (nearest) and (curtate) duration as shown in Table 7.2.

Table 7.2. Exposure generated in 1999 by a record with a date of birth of 13/5/1969 and a date of commencement of 1/7/1997.

Date	Age Nearest	Curtate Duration	Exposure from	
			1/1/1999 in-force	1/1/2000 in-force
1/1/1999	30	1	365/365	0
2/1/1999	30	1	364/365	1/365
3/1/1999	30	1	363/365	2/365
4/1/1999	30	1	362/365	3/365
etc				
28/12/1999	31	2	4/365	361/365
29/12/1999	31	2	3/365	362/365
30/12/1999	31	2	2/365	363/365
31/12/1999	31	2	1/365	364/365
<b>Total</b>			66795/365 = 183	66430/365 = 182

- 7.17. Note that the exposure sums to 365 days, and that the age and duration reflect the actual date of birth and date of commencement.
- 7.18. In the example above, it was assumed that the record was in force throughout 1999. This may not be the case because policies have entered the investigation during 1999 (either as new business or as a reinstatement) or exited (by lapse, maturity, etc).
- 7.19. Suppose now that the record featured only in the 1/1/1999 in force data, because it has exited the investigation at an unknown point during 1999. Only the first column of exposure will be generated. Whilst this sums to close to half-a-year of exposure, the categorisation by age and duration is weighted towards the start of the year, reflecting an equal likelihood of exit at any date during the year.
- 7.20. Similarly, suppose now that the record featured only in the 31/12/1999 in force data, because it has joined the investigation at an unknown point during 1999, presumably as a result of a reinstatement, since it does not arise from new business. Only the second column of exposure will be generated. Whilst this also sums to close to half-a-year of exposure, the categorisation by age and duration is weighted towards the end of the year, reflecting an equal likelihood of reinstatement at any date during the year.
- 7.21. In relation to the approach outlined in the preceding paragraphs, note that:
- The number of exits is likely to substantially exceed the number of reinstatements, hence the balance of the exposure during the year is affected by the weighting towards the start of the year from the exits.
  - The approach ignores exposure in respect of policies that enter and exit in the same calendar year, and thus are never captured in year-end in force data. This almost exclusively affects exposure at curtate duration 0.
  - The approach overstates the exposure in respect of policies that exit and are then reinstated in the same calendar year, as a full year's exposure is generated from the combination of the start- and end-year in force data, so are effectively assumed to have been in force throughout. This may affect the exposure at all

durations, although in practice one might expect reinstatements to be weighted towards short durations.

We believe this approach will produce the most accurate estimate of exposure possible in the absence of information on the actual date of exit or reinstatement, both in relation to overall exposure and to the allocation by age and duration. Access to these dates and use of an exact method of calculating exposure would alter the exposure calculated (and its categorisation by age and duration).

7.22. While the above calculations are carried out on a daily basis, the data are summed by calendar month and stored by age, duration (in quarters), etc.

### ***Exposure before the Investigation Period***

#### *Known prior year exposure*

7.23. In Working Paper 28, we described the generation of known exposure for the years prior to the investigation period (which gives rise to diagnosed claims that may be settled in the investigation period). In our current application, we use the actual commencement dates for those records that are in the earliest in force data for that office, rather than a census approach, as in the initial application.

7.24. Each record in the initial in force data is assumed to have been in force continuously since its commencement date, and the exposure generated is calculated on a daily basis with regard to age (nearest) and curtate duration (in quarters). As an example, the record described in 7.16 generates exposure as shown in Table 7.3. (Note that in Tables 7.3 and 7.4 we have illustrated curtate quarterly duration changing on the 1<sup>st</sup> of the relevant month, for simplicity. In practice we have actually assumed that the first quarter lasts for 91 days from commencement, etc.)

Table 7.3. Exposure generated in 1997 and 1998 by a record with a date of birth of 13/5/1969 and a date of commencement of 1/7/1997.

<b>Period</b>	<b>Age Nearest</b>	<b>Curtate Duration</b>	<b>Exposure (days)</b>
12/11/1998 – 31/12/1998	30	1.25	50
1/10/1998 – 11/11/1998	29	1.25	42
1/7/1998 – 30/9/1998	29	1.00	92
1/4/1998 – 30/6/1998	29	0.75	91
1/1/1998 – 31/3/1998	29	0.50	90
12/11/1997 – 31/12/1997	29	0.25	50
1/10/1997 – 11/11/1997	28	0.25	42
1/7/1997 – 30/9/1997	28	0	92

7.25. This calculation is undertaken for each record in the start in force data for the earliest year for each office, projecting it back to its date of commencement. Note that whilst we calculate the exposure on a daily basis, it is again summed (by age, duration, etc) on a calendar monthly basis.

#### *Synthetic prior year exposure*

7.26. In Working Paper 28, we described the use of off rates to estimate the 'synthetic' in force data for the years prior to the investigation period. These data arise from policies that are no longer in force at the date of the first in force data available to us, but as with

the 'known' prior year exposure may have given rise to diagnosed claims that are settled in the investigation period.

- 7.27. The derivation of off rates was described in paragraph 3.5 of Working Paper 28, and some further analysis to help derive a suitable assumption is described in section 8 of this paper. However instead of applying these to the known prior year in force data, as in the initial application, we now apply them to the known exposure to estimate the additional synthetic exposure.
- 7.28. This can be most easily explained with an example. Using an off rate of 9% p.a. and assuming that offs occur evenly over the year, the additional exposure generated is calculated as follows:
- 31 December 1998: A policy exiting on 31 December is assumed to generate exposure for that day. We estimate this additional exposure (in days) as  $[1/(1-0.09)^{(1/365)}-1]$ , i.e. an additional 0.0258419%.
- 30 December 1998: The additional exposure from policies exiting on 31 December is also estimated as 0.0258419% for 30 December. A policy exiting on 30 December will generate additional exposure (in days) of  $1.000258419*[1/(1-0.09)^{(1/365)}-1]$ , i.e. an additional 0.0258486%, giving an additional 0.0516905% in total.
- 29 December 1998: The additional exposure from policies exiting on 31 and 30 December is again estimated as 0.0258419% and 0.0258486% respectively for 29 December. A policy exiting on 29 December will generate additional exposure (in days) of  $1.000258486*[1/(1 - 0.09)^{(1/365)}-1]$ , i.e. an additional 0.0258552%, giving an additional 0.0775457% in total.
- 7.29. Similar considerations apply to all earlier dates. Note that the additional exposure gradually increases as one works back in time. This is due to the increasing likelihood that a policy will have exited before 1 January 1999 and hence the exposure from these "future exits" increases.
- 7.30. These percentage additions apply to all known exposure at any given date, assuming that the profile of business exiting is identical to that of business remaining in force.
- 7.31. As described above, the total exposure, known and synthetic, at any given date (by age, duration and other risk factors) can be derived by ratioing up the known exposure by the appropriate percentage. Since the known exposure already takes account of actual date of commencement for any particular record, the additional synthetic exposure is only generated back to that date.
- 7.32. To illustrate this by way of an example, suppose there is a group of 100 records in the 1/1/1999 in force data identical to the record described in 7.16. The exposure generated (by calendar month) is as shown in Table 7.4.

Table 7.4. Known and synthetic exposure generated in 1997 and 1998 by 100 records all with a date of birth of 13/5/1969 and a date of commencement of 1/7/1997 assuming an off rate of 9% p.a.

Calendar Month	Calendar Year	Age Nearest	Curtate Duration	Exposure (days)			% Synthetic
				Known	Synthetic	Total	
December	1998	30	1.25	3,100	12.85	3,112.85	0.4%
November	1998	30	1.25	1,900	20.24	1,920.24	1.1%
November	1998	29	1.25	1,100	16.03	1,116.03	1.4%
October	1998	29	1.25	3,100	62.30	3,162.30	2.0%
September	1998	29	1.00	3,000	84.51	3,084.51	2.7%
August	1998	29	1.00	3,100	112.54	3,212.54	3.5%
July	1998	29	1.00	3,100	138.38	3,238.38	4.3%
June	1998	29	0.75	3,000	158.71	3,158.71	5.0%
May	1998	29	0.75	3,100	189.82	3,289.82	5.8%
April	1998	29	0.75	3,000	208.89	3,208.89	6.5%
March	1998	29	0.50	3,100	242.09	3,342.09	7.2%
February	1998	29	0.50	2,800	241.75	3,041.75	7.9%
January	1998	29	0.50	3,100	293.42	3,393.42	8.6%
December	1997	29	0.25	3,100	320.72	3,420.72	9.4%
November	1997	29	0.25	2,000	221.50	2,221.50	10.0%
November	1997	28	0.25	1,000	115.06	1,115.06	10.3%
October	1997	28	0.25	3,100	375.06	3,475.06	10.8%
September	1997	28	0	3,000	389.57	3,389.57	11.5%
August	1997	28	0	3,100	430.26	3,530.26	12.2%
July	1997	28	0	3,100	458.65	3,558.65	12.9%
<b>Total</b>				<b>54,900</b>	<b>4,092.34</b>	<b>58,992.34</b>	<b>6.9%</b>

- 7.33. Note that the approach outlined in the preceding paragraphs only generates a different amount of exposure in total from the initial application set out in Working Paper 28 if dates of commencement do not occur at the middle of the year, on average. However even if this assumption holds true, the allocation of this exposure by age and duration does alter from the initial application as a result of using exact dates of birth and commencement.
- 7.34. Note also that whilst we have continued to use a single off rate within this example, this is not a necessary assumption. Different assumptions regarding off rates would result in different uplifts at each prior date, which would then be applied to the appropriate subset of the data.
- 7.35. One final point to note is that the application of this process did highlight an issue with dates of commencement supplied by one office. The data included twenty-four policies with commencement dates stated to be as far back as 1972; application of the above process would generate exposure for every year back to 1972, as well as distorting the exposure by duration. The subsequent application of the (truncated) claim development distribution effectively means that any exposure prior to 1992 is irrelevant to claims settled in 1999 and later, so the practical effect is limited. As a consequence, these policies have not been investigated further with the office concerned but, for the purposes of generating the prior year exposure, the commencement dates of these policies have been arbitrarily reset to 1<sup>st</sup> January 1988.

7.36. Whilst the example above may appear trivial and amusing, it does illustrate the dependency of the methodology on the accuracy of dates of commencement supplied to the CMI, which would previously only have been used for the purpose of calculating duration. One possible explanation for the issue noted above is that date of policy commencement has been supplied, and a critical illness benefit added at a later date; if this has been widespread practice it could distort our results, but would not be detected for cases where the original policy had been written in the 1990s (or later).

7.37. In summary, the differences in the calculation of exposure during the investigation period from that calculated for Working Paper 28 are:

- In the initial application and in the released results, exposure was calculated using a census approach, averaging the start and end in force. In the revised application we have used the actual dates of commencement, which are known for all data submitted to the CMI. This changes the amount of exposure and also allocates it accurately by duration.
- Similarly we have used actual dates of birth for each record. Whilst this does not change the amount of exposure, it does allocate it accurately by age.
- We do not have details of date of exit for individual records. In Working Paper 28, and in the released results, we have assumed that exits all occur mid-year; in the revised application we have assumed an equal probability of exit, for those records that exit, on each day of the year.

Similar differences apply to the calculation of exposure prior to the investigation period.

7.38. A reconciliation of the exposure in 1999-2002 using our current approach with that calculated in the released results and in Working Paper 28 is contained in Appendix B.

### ***Expected Diagnosed Claims***

7.39. The combination of known and synthetic exposure gives rise to the total exposure required to produce expected diagnosed claims. These are the claims that we expect to be diagnosed during the investigation period and the relevant preceding years using CIBT93 to calculate the expected.

7.40. In the current application of the methodology we have retained the exposure by calendar month, and assumed that the expected diagnosed claims occur on the 15<sup>th</sup> of each month. The calculation is straightforward; however it may be helpful to re-state some details of our use of CIBT93.

7.41. CIBT93 was originally contained in the paper “A Critical Review”. Separate rates are provided for males and females, but they are aggregate rates, i.e. they are not adjusted for smoking status. The tables only allow for certain critical illnesses (cancer, heart attack, stroke, coronary artery bypass surgery, multiple sclerosis, kidney failure, major organ transplant, total and permanent disability and – for accelerated cover – death) but we have compared claims from all causes with these, i.e. including claims from causes not specifically included within CIBT93.

7.42. CIBT93 provides rates for ages from 20 to 80 and in all our work to date we have used the rates at age 20 for ages below that, and the rate at age 80 for ages above that. The volume of exposure for 1999-2002 at ages below 20 is around 40,000 life-years but, as the vast majority of this relates to ages 16-19, the Committee does not believe this to be



a material assumption. The volume of exposure at ages over 80 is around 10 life-years so is clearly not material. The Committee will keep this assumption under review in future years but, for now, believes the simplicity of the current approach is preferable to an extrapolation of CIBT93 or any other alternative.

- 7.43. 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 A/E values in the paper 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.
- 7.44. The released results use an initial exposed to risk and an age definition of age nearest. This means that lives are exposed from age exact  $x-1/2$  to  $x+1/2$  and we have therefore used the rates from CIBT93 unadjusted.
- 7.45. As noted in paragraph 2.4 regarding the feedback to Working Paper 28, we referred (in paragraph 4.16 of Working Paper 28) to the definition of our exposure calculation as indeterminate. This arose because although we made no explicit allowance in the exposure calculation for exposure after the date of diagnosis for claims; we suspected that policies will often remain premium-paying until (around) the date of settlement and exist in the in force data at year-end(s) following the date of diagnosis.
- 7.46. Whilst we have adopted a more accurate calculation of exposure within this paper, in particular with regard to actual dates of commencement, no change has been made with regard to exposure beyond the date of diagnosis which we only intend to address when we have more complete submission of dates of diagnosis.

### ***Expected Settled Claims***

- 7.47. These are the claims that we expect to be settled during the investigation period using:
- The methodology set out earlier in this section to calculate the exposure (using the off rates specified in the next section);
  - CIBT93 to calculate the expected diagnosed claims from the exposure, noting the points above regarding the interpretation of the CIBT93 rates; and
  - The Burr model of the claim development distribution with parameters shown in Table 5.2, subject to truncation of the tail as set out in paragraph 5.50, to transform the expected diagnosed claims to expected settled claims.
- 7.48. For each category of gender and smoker status, we have calculated expected diagnosed claims by age (nearest), by duration (curtate, in quarters) and by calendar month.
- 7.49. As noted above, the expected diagnosed claims are assumed to occur on the 15<sup>th</sup> of each month. These are assumed to be at age exact and the duration is assumed to be halfway through the quarter (measured in days), e.g. expected diagnosed claims at curtate duration 0.5 are assumed to all occur with exact duration 0.625 of a year, which has been taken to equal 229 days.

- 7.50. The claim development distribution is then applied to each of these sub-divisions of expected diagnosed claims to calculate the expected settled claims on each (future) day. These are then summed according to:
- Age nearest at date of settlement;
  - Curtate duration at date of settlement; and
  - Calendar year.
- 7.51. Within the current application of the methodology outlined in this section we have summed exposure, expected diagnosed claims and expected settled claims by integer age (nearest) and by (curtate) duration in quarters. More refined categorisation may have produced more accurate results but at the expense of increased run-times compared to the approach that we adopted. We did not consider it necessary to sub-divide age, beyond annual categories: exposure builds up gradually by age and – when the claim development distribution is applied to the exposure – claims moving to a higher age (at settlement) are, to an extent, offset by claims moving from a lower age (at diagnosis).
- 7.52. In contrast, exposure at curtate duration zero is very high (indeed in 1999-2002 it is the duration with the greatest amount of exposure) and whilst the application of the claim development distribution to the exposure moves claims to a higher duration (at settlement) there is no offsetting movement of claims from a lower duration (at diagnosis). We undertook checks on different assumptions and believe that quarterly segmentation of duration provides sufficiently accurate results without the additional storage and run-time issues associated with, say, monthly segmentation.
- 7.53. A comparison of these expected settled claims with the actual settled claims, for calendar years 1999 to 2002, is considered in section 9.

## 8. ANALYSIS OF OFF RATES

- 8.1. Working Paper 28 introduced the concept of ‘off rates’, which are a necessary feature of the revised methodology. We have undertaken a limited amount of further analysis of off rates since Working Paper 28 was published and this is summarised in this section.
- 8.2. In this further analysis we are only looking to take account of major features. Whilst the methodology is not dependent on a single assumption regarding the off rate, the *application of* the methodology is undoubtedly complicated if more complex assumptions are used. Furthermore, as stated in Working Paper 28 (and repeated in paragraph 2.14 above),
- “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” [of Working Paper 28].
- 8.3. Our additional analysis is not intended to be thorough or complete, but is intended:
- To provide a better estimate of the off rates we should assume in our principal estimate of past exposure, which drives our estimate of expected settled claims in our derivation of 1999-2002 results in section 9; and
  - To inform the choice of scenarios that we will use to illustrate the sensitivity of the results to off rates in section 10.
- 8.4. In particular it is worth noting at this point that the derivation of the single assumption (of 9% p.a.) used to illustrate the methodology in Working Paper 28 was based on consideration of 1999 and 2000 only, whereas in this analysis we use experience from 1999-2004. Given that our intention is to use these off rates to produce results for accelerated business only, we have also restricted our analysis to this subset of the business.
- 8.5. The off rate is simply the percentage of business in a particular subset of the start in force data that does not exist in the end in force data. 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.
- 8.6. Note also the need for consistency of the start- and end- in force data; five ‘submission groups’<sup>1</sup> were indicated for the 1999-2002 data in paragraph 4.3 of Working Paper 28, this has increased to nine for the 1999-2004 data, due to additional offices (and portfolios) starting and ceasing to submit data. The values in this section (and indeed the application of the methodology in section 7) have therefore been calculated within submission groups, but then aggregated for ease of presentation.
- 8.7. Table 8.1 shows a high-level summary of off rates during the period by calendar year and duration. Note that the volume of data varies considerably between cells; in

---

<sup>1</sup> The concept of ‘submission groups’ is described in Working Paper 28. There were numerous changes in the composition of business included in the CMI critical illness investigation during 1999-2002 and submission groups divide the data into subsets according to the years for which data is thought to have been submitted continuously. See paragraphs 4.2 to 4.4 of Working Paper 28 for more detail.

particular volumes tend to increase by calendar year and reduce with increasing duration.

Table 8.1. Calculated off rates for full acceleration business during 1999-2004 by calendar year and curtate duration.

Duration	Calendar Year						
	1999	2000	2001	2002	2003	2004	ALL
<b>0</b>	8.3%	9.1%	11.7%	13.1%	11.8%	11.9%	<b>11.4%</b>
<b>1</b>	10.6%	12.1%	15.3%	19.2%	16.6%	13.8%	<b>14.9%</b>
<b>2</b>	11.2%	12.6%	16.1%	18.7%	19.6%	14.4%	<b>15.6%</b>
<b>3</b>	9.2%	11.5%	14.8%	17.3%	18.6%	15.1%	<b>14.9%</b>
<b>4</b>	8.2%	10.3%	14.1%	16.8%	19.0%	16.7%	<b>14.9%</b>
<b>5</b>	8.3%	9.0%	10.8%	13.7%	16.0%	14.3%	<b>12.7%</b>
<b>6</b>	8.0%	8.4%	10.8%	10.9%	14.4%	12.7%	<b>11.6%</b>
<b>7</b>	5.4%	7.9%	10.6%	10.6%	13.0%	11.9%	<b>11.1%</b>
<b>8</b>	5.0%	6.5%	9.3%	10.8%	12.3%	10.6%	<b>10.3%</b>
<b>9</b>		8.5%	10.0%	14.2%	15.5%	14.6%	<b>13.6%</b>
<b>10+</b>			6.3%	8.0%	11.1%	12.1%	<b>10.6%</b>
<b>ALL</b>	<b>9.4%</b>	<b>10.6%</b>	<b>13.4%</b>	<b>15.5%</b>	<b>15.2%</b>	<b>13.5%</b>	<b>13.5%</b>

8.8. It will be noted that there is considerable variation by calendar year, with off rates increasing from 1999 to 2002 and then reducing slightly to 2004. In view of the size of the differences by year, the Committee decided it was appropriate to use different off rates by calendar year, as shown in Table 8.2 in deriving the 1999-2004 results.

Table 8.2. Assumed off rates by calendar year for subsequent modelling.

Calendar Year	1998 & prior	1999	2000	2001	2002	2003
Off Rate	9%	10%	11%	13%	15%	15%

8.9. There is also variation by duration, with off rates increasing over the first year or two years, then generally reducing. The variation by duration is often of similar magnitude to the variation by calendar year, and there is considerable variation within each duration by calendar year, so that – to take account of variation by duration – we would need to use a two-way matrix by duration and calendar year. This adds complexity to the implementation of the methodology and the Committee decided not to vary off rates by duration in our principal estimate, despite the pattern apparent above, at this stage of our work.

8.10. Note that “duration” in this context means “curtate duration at the start of a calendar year”. This means that, for example, the off rate for “duration 0” applies (on average) from duration 0.5 to duration 1.5, which differs from its normal interpretation.

8.11. Table 8.3 shows the variation in off rates by other selected factors. These values are for all calendar years combined, but in all cases the same pattern by calendar year that was noted above exists.

Table 8.3. Calculated off rates for selected subsets of full acceleration business during 1999-2004.

<b>Factor</b>	<b>Subset</b>	<b>% of data</b>	<b>Off rate</b>
<b>Smoker Status</b>	Non-smoker	80%	12.7%
	Smoker	19%	16.8%
<b>Age group</b> (age nearest as at 1/1/2003)	< 30	24%	16.0%
	31-40	47%	13.0%
	41-50	22%	12.1%
	51+	7%	13.6%
<b>Benefit Amount</b> (current Sum Insured at each 01/01/YY)	< £25,000	19%	12.5%
	£25,000 - £50,000	32%	14.5%
	£50,000 - £100,000	33%	14.1%
	£100,000 - £200,000	13%	11.4%
	£200,000+	2%	10.5%
<b>Gender and Single Life or Joint Life policy</b>	Male/Single	22%	13.3%
	Male/Joint	29%	13.9%
	Female/Single	19%	12.5%
	Female/Joint	29%	13.9%

8.12. The variations by age, benefit amount, gender and policy type are comfortably inside the range of sensitivities considered in section 10, so the Committee did not consider that these warranted further attention at this stage of our work.

8.13. The variation by smoker status does warrant further consideration. Members of the Committee have seen evidence of higher lapse experience amongst smokers (than non-smokers) on term assurance-type products. This might be explained by the higher premium leading to a reconsideration of the value of the cover in a greater number of cases, or by the trend away from smoking, allowing individuals to re-apply for cheaper cover, after a 12-month period without tobacco use. Whatever the reason, this appears to be evidenced by higher off rates for smokers in Table 8.3.

8.14. However, term assurance-type critical illness products only began to feature prominently in the UK market from about 1997/1998, so a significant part of the business to which we are applying the off rates is likely to be other product types, e.g. unit-linked whole life and mortgage endowments. Such products may have higher morbidity deductions for smokers, but the premiums have an investment element (especially for mortgage endowments) which is identical between smokers and non-smokers. Hence even though we have found higher off rates for smokers in 1999-2004, if these arise from term assurance-type products we cannot necessarily assume these also apply to the other product types that predominate in earlier years.

8.15. As a result we consider that a more complete analysis of off rates (and incorporation into the methodology) would be required to consider differences by product type, as well as by smoker status. Unfortunately, product type is not a field that we have yet been able to analyse. The data requirements allow offices to input any identifier into this field, and whilst for some this makes the product-type readily apparent to the CMI, for others we need offices to provide an indication of what their codes mean. We will continue to pursue this during 2008.

8.16. In addition, it must be recognised that the further analysis of off rates may lead to more complex differentiation between off rates in our estimation of the prior years' exposure. This is unlikely to be a trivial exercise.

8.17. It is also important to note that considerable variation in off rates exists between offices; this is illustrated in Table 8.4 which shows the highest and lowest off rate at each duration by office. Whilst some offices have consistently lower-than-average off rates (and others higher), the figures in Table 8.4 may be for different offices at each duration.

Table 8.4. The lowest and highest calculated off rates by office for curtate duration 0 to 5 years for full acceleration business during 1999-2004.

<b>Duration</b>	<b>Lowest</b>	<b>Highest</b>	<b>ALL</b>
<b>0</b>	6.6%	15.8%	<b>11.4%</b>
<b>1</b>	8.3%	21.1%	<b>14.9%</b>
<b>2</b>	8.1%	22.5%	<b>15.6%</b>
<b>3</b>	7.5%	22.1%	<b>14.9%</b>
<b>4</b>	4.3%	21.6%	<b>14.9%</b>
<b>5</b>	7.5%	17.4%	<b>12.7%</b>

8.18. Table 8.4 contains a very simplistic comparison (e.g. the business from some offices will be heavily weighted towards the more recent calendar years, where off rates appear higher) but nevertheless illustrates that detailed analysis of variations in off rates may not produce more accurate results since the data from which off rates are calculated is necessarily either from a different set of offices or from a later time period than that for which they are being used to estimate exposure.

8.19. It is also possible that off rates by office may be correlated with new business volumes, if they reflect policyholders taking advantage of reducing premium rates (as was common in the UK until 2003). This too remains an area for future investigation.

## **9. 1999-2002 RESULTS**

- 9.1. In this section, we set out key features of results for all causes experience on a lives basis for All Office full acceleration business in 1999-2002.
- 9.2. These results are contained in Appendix C, together with details of their derivation.

### ***Previous results***

- 9.3. Results for 1999-2002 were sent to member offices in May 2005, at the same time as Working Paper 14 was published. As previously noted, these 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 indicated the need for an overall grossing-up factor of around 15%.
- 9.4. A set of results using our initial application of the revised methodology was also included within Working Paper 28. These were intended to illustrate the methodology but incorporated a number of simplifying assumptions so did not necessarily reflect the true underlying experience by age and duration.

### ***Adjusted results***

- 9.5. Within this paper we have calculated results based on expected settled claims derived using the following:
- The methodology set out in section 7 of this paper to calculate the exposure including the off rates specified in Table 8.2;
  - CIBT93 to calculate the expected diagnosed claims from the exposure, noting the explanation regarding our use of CIBT93 in paragraphs 7.41 to 7.44; and
  - The Burr model of the claim development distribution with parameters shown in Table 5.2, subject to truncation of the tail as set out in paragraph 5.50, to transform the expected diagnosed claims to expected settled claims.
- 9.6. A reconciliation of these values with the released results and those in Working Paper 28 is contained in Appendix B.

### ***Key features of results***

- 9.7. We refer to the values of  $100 \times \text{Actual Settled Claims} / \text{Expected Settled Claims}$  (both based on date of settlement) as 'adjusted' results in the following paragraphs, and to values of  $100 \times \text{Actual Settled Claims} / \text{Expected Diagnosed Claims}$  (both based on date of diagnosis) as 'released' results. Both sets of results are shown in Appendix C.
- 9.8. Care is required in interpreting the adjusted results. The methodology has been developed to seek to make best use of the data available to us but the results depend on a substantial number of assumptions, summarised in paragraph 10.2. We expect to undertake further work that may lead to better estimates for some of the assumptions. Hence we do not regard the results as necessarily definitive.

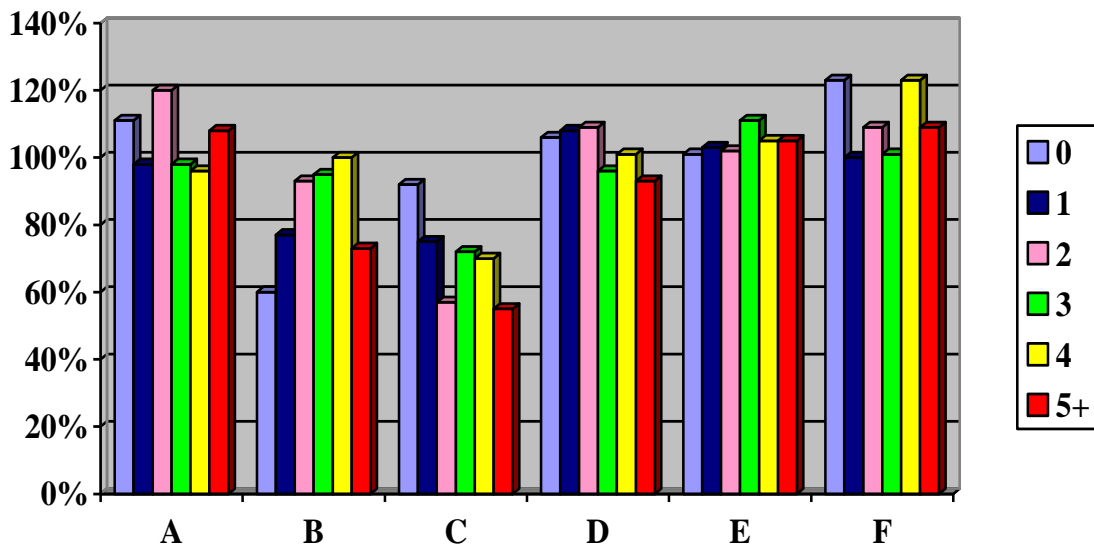
- 9.9. It is also important to note that the adjusted results apply to experience in terms of settled claims. Patterns in the experience of settled claims will not be the same as the patterns that would be observed in diagnosed claims; this is especially true of duration (and, to a lesser extent, age) as experience in settled claims at any duration other than zero is necessarily a function of experience in diagnosed claims at that and earlier durations.
- 9.10. The first feature to note regarding the adjusted results is that they significantly exceed the released results. This is to be expected. The overall level of increase is around 12%, which is close to the grossing-up factor of 15% indicated in Working Paper 14. (Indeed our original estimate of the grossing-up factor was 13%, which we chose to state as “around 15%” to avoid it being regarded as an unduly accurate estimate.) We therefore believe that, overall, the revised methodology produces a realistic representation of true experience of settled claims.
- 9.11. The final column of each table in Appendix C shows a comparison of the adjusted results with the released results. At the all durations and all ages level this is very similar to the grossing-up factor introduced in Working Paper 14, as it moves Actual/Expected values from (unmatched) released results to (matched) adjusted results. It is not, however, identical. As described in Working Paper 14, grossing-up factors attempted to estimate the adjustment from released results to Actual Diagnosed Claims/Expected Diagnosed Claims whereas the adjusted results within this paper are expressed as Actual Settled Claims/Expected Settled Claims. At an overall level, the Committee expects that these two values would normally be close, as the difference is solely a timing difference (between claims diagnosed in 1999-2002 and claims settled in 1999-2002) and the underlying time periods overlap significantly.
- 9.12. When considering results for a specific duration or age, the comparison of the adjusted results with the released results will not necessarily be close to the grossing-up factor, as we defined it, because the Actual Settled Claims in the adjusted results are based on date of settlement whereas those in the released results are based on date of diagnosis, hence there is a more significant difference than just timing.
- 9.13. The comparison is still of interest below the overall level though as it shows the degree of adjustment required to the released results to get to a matched comparison.
- 9.14. The percentage differentials between the adjusted results and the released results are very similar between males and females and between non-smokers and smokers at an all ages and all durations level. This is unsurprising given that:
- As noted originally in Working Paper 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 depend on the off rates and claim development distribution used to derive the adjusted results and we have not differentiated these by gender or by smoker status.
- 9.15. There is however considerable variation by duration. The most consistent features of these results are that the released results understate experience at duration 0 and at



durations 5+. The released results also understate the adjusted results at most other durations but with exceptions, for example at duration 2 for male non-smokers.

- 9.16. The adjusted results show less of a select effect than the released results, but still appear to demonstrate some positive selection: overall experience at duration 0 is 48% of CIBT93, increasing to 49% at duration 1, 51% at each of durations 2 to 4 and then to 54% at durations 5+. However this is not true of each of the categories for business – for example, it is difficult to discern any pattern by duration for male smokers.
- 9.17. Note that it can be misleading to infer select effects from All Office results, because of the variation in weight given to each office by duration. Figure 9.1 shows the relative adjusted results by duration for each of six large offices (for males and females, non-smokers and smokers combined) and for most offices the select effect is far from obvious.
- 9.18. Furthermore the positive selection apparent in the adjusted results – based on settled claims – says nothing directly about the shape of the select effect in underlying rates of claim diagnoses. 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. Indeed the Committee debated not including results by duration for fear that they might be misinterpreted.

Figure 9.1. Adjusted results by curtate duration for six large offices for full acceleration business during 1999-2002, expressed relative to the adjusted results for these six offices combined, all durations.



- 9.19. Note that the adjusted results in Figure 9.1 do not necessarily demonstrate the variability that exists between offices. These results compare Actual Settled Claims for an individual office with Expected Settled Claims, where the latter takes account of the profile of business (by age and duration) for that office but using the same assumptions regarding off rates and claim development distribution for all offices. We have noted variations between offices in both of these assumptions in earlier sections.

9.20. Comparison of the adjusted results for smokers with those for non-smokers provides an indication of the smoker differentials, given that the sets of results are both derived from CIBT93, which is not differentiated. Differentials by age band and duration are shown in Table 9.1.

Table 9.1. Ratio of smoker to non-smoker adjusted results by age band and curtate duration for full acceleration business during 1999-2002.

Age Band	Duration						
	0	1	2	3	4	5+	All
<b>Males</b>							
< 30	146%	88%	104%	104%	104%	98%	<b>110%</b>
31-40	161%	175%	146%	181%	192%	102%	<b>154%</b>
41-50	309%	233%	217%	216%	203%	187%	<b>216%</b>
51-60	323%	225%	201%	192%	236%	187%	<b>212%</b>
61+	74%	122%	247%	111%	165%	166%	<b>152%</b>
<b>All ages</b>	<b>221%</b>	<b>187%</b>	<b>177%</b>	<b>183%</b>	<b>198%</b>	<b>161%</b>	<b>182%</b>
<b>Females</b>							
< 30	136%	104%	128%	103%	120%	117%	<b>118%</b>
31-40	109%	125%	94%	102%	74%	124%	<b>107%</b>
41-50	109%	119%	124%	136%	148%	153%	<b>133%</b>
51-60	144%	141%	131%	179%	139%	151%	<b>147%</b>
61+	0%	1185%	236%	68%	0%	361%	<b>260%</b>
<b>All ages</b>	<b>118%</b>	<b>126%</b>	<b>117%</b>	<b>127%</b>	<b>114%</b>	<b>147%</b>	<b>127%</b>

9.21. Whilst there are small volumes of actual claims in some cells (especially for female smokers), it would appear that smoker differentials are substantially higher for males than females. It would also appear that differentials reduce with increasing duration for males, but may increase with increasing duration for females.

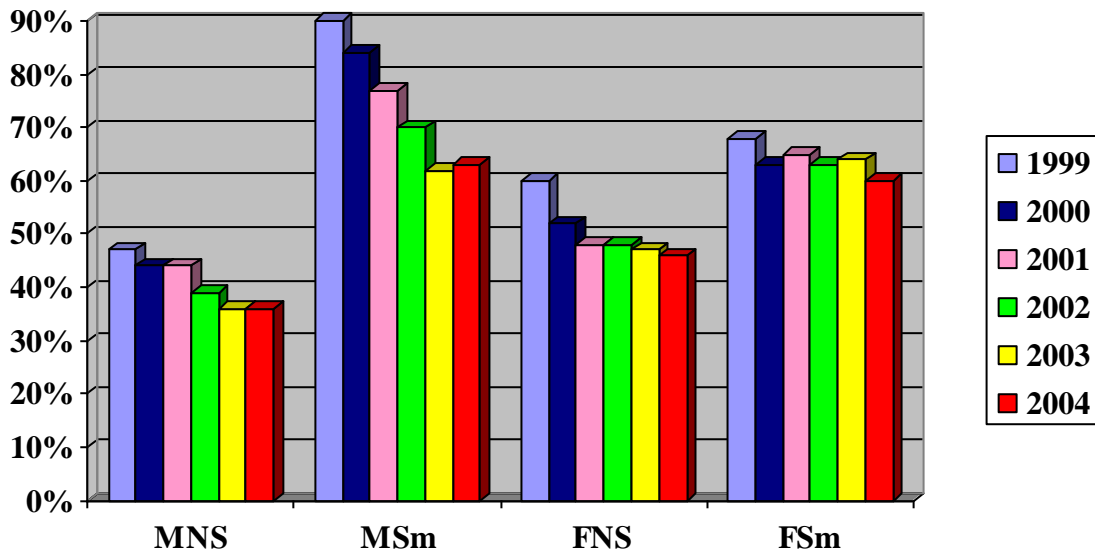
9.22. Unsurprisingly, differentials are low for both males and females in the youngest age category, but then tend to increase rapidly.

9.23. Although the smoker differentials apply to settled claims, the Committee does not expect the patterns to be significantly different in diagnosed claims, except perhaps by duration.

9.24. Figure 9.1 also provides an indication of the variation in adjusted results between offices (subject to the comment in paragraph 9.19). There is relatively little variation between four of the offices shown (A, D, E and F) each of which has overall results of between 101% and 110% of the All Office results contained in Appendix C. Offices B and C show much lighter overall experience, but this may reflect the immaturity of their portfolios, as each has over 70% of their settled claims at durations 0 to 2, whereas offices A, D, E and F have less than 50%.

9.25. Although the results in this paper are limited to 1999-2002, we have also produced adjusted results for 2003 and 2004 and these will be sent to CMI member offices shortly. A high-level summary for all these years is included in Figure 9.2.

Figure 9.2. Adjusted results for full acceleration business during 1999-2004 by calendar year, all offices, all ages and all durations combined.



- 9.26. Note that Figure 9.2 illustrates adjusted results, in terms of settled claims, and does not therefore indicate the experience for diagnosed claims, still less by underwriting year. In addition, this is not a true reflection of experience by calendar year as the same claim development distribution, derived from claims settled in 1999-2004, has been used throughout.
- 9.27. For males, experience appears to be improving by year, whereas the position for females is less clear with little variation between 2000 and 2004 in particular. However as with most CMI All Office results, different offices feature in different years, so any comparison across time is not definitive.
- 9.28. Furthermore even illustrating experience for a restricted group of offices (sometimes termed “loyal offices”) that have contributed data in all years could still present a misleading picture. In particular the relative weights of offices alter significantly during the period, according to business growth rates, so that even a group of “loyal offices” would not represent a constant dataset.

## 10. ASSUMPTIONS AND SENSITIVITIES IN THE RESULTS

10.1. In this section we summarise the assumptions that underlie the adjusted results in the preceding section and illustrate the sensitivity of the results to some of the key assumptions.

10.2. A substantial number of assumptions underlie the adjusted results. It is important to recognise that a degree of uncertainty attaches to each of these, and hence a considerable degree of uncertainty surrounds the results. The areas where assumptions are required are set out below:

### *EXPOSURE DURING THE INVESTIGATION PERIOD*

- a) Dates of exit are unknown for all records and hence are estimated to calculate the exposure during the investigation period.
- b) Exposure during the investigation period makes no allowance for policies entering and exiting within a calendar year (nor is exposure reduced for policies exiting and being reinstated within a year).
- c) Exposure does not stop at the date of diagnosis.
- d) The current implementation of the methodology is not exact; in particular we have used time in months, age in years and duration in quarters. These were adopted to avoid excessive run times but have considerable significance when the claim development distribution is applied.

### *EXPOSURE PRIOR TO THE INVESTIGATION PERIOD*

- e) We have used a simple structure of off rates to estimate past exposure, even though these almost certainly vary by product type, by duration and by office.
- f) Off rates are applied to an earlier period than the data from which they are estimated; in particular we have no data for 1998 and prior years, yet are making an assumption regarding off rates during those years.

### *EXPECTED DIAGNOSED CLAIMS*

- g) Our interpretation of CIBT93 needs to be noted.
- h) We have assumed that diagnoses occur on the 15<sup>th</sup> of the calendar month, on a birthday (as we are using age nearest) and at the mid-point of duration in quarter-years.

### *ESTIMATING THE CLAIM DEVELOPMENT DISTRIBUTION*

- i) The date of diagnosis is not well-defined for many CI events. This leads to uncertainty over the timing of the event we are seeking to measure.
- j) We have used a single claim development distribution, even though it might vary with a number of factors.
- k) This claim development distribution is only based on around half the settled claims within our analysis. In particular a number of offices have not submitted any claims that were included in the subset of the data used for modelling, so we are using the average claim development distribution for these offices without being able to gauge its appropriateness.
- l) The claim development distribution derived from 1999-2004 claims data is assumed to apply throughout the period under investigation (and, in particular, to 1999-2002).

- m) We have made decisions (some of them arbitrary) as to which claims to include in the dataset for modelling the claim development distribution, for example excluding claims with equal dates of diagnosis and settlement.
- n) We have also made arbitrary decisions on how to allow for an increasing proportion of claims with the relevant dates of claim.
- o) We have assumed that a Burr model reasonably represents the claim development distribution.
- p) We have also assumed that it is reasonable to truncate the fitted Burr distribution to achieve a finite distribution.

*GENERAL*

- q) We assume the data fields used are accurate. In some cases we have sought to correct for instances where data values are highly unlikely to be accurate, but there are probably other instances that have gone undetected.

10.3. The novel assumptions within our methodology relate to the claim development distribution and the level and structure of off rates. We have therefore carried out sensitivity tests to check whether these assumptions might have introduced any serious distortions in our results. These tests and our conclusions from them are set out in the following sections.

*Sensitivity to off rates*

- 10.4. Firstly, we consider two sets of alternative assumptions regarding the off rates that are used to estimate the exposure in prior years. These reflect the variation by duration and the variation between smokers and non-smokers noted in section 8, however we believe that a wider inference can be drawn regarding the insensitivity of the results to the assumed off rates.
- 10.5. In the first scenario, we have used off rates 3% higher for male smokers and 1% lower for male non-smokers than those in Table 8.2. This is consistent with the calculated off rates for 1999-2004. The overall effect is to reduce the number of expected settled claims for non-smokers from 9,923.2 to 9,915.9 (-0.07%) and to increase the number of expected settled claims for smokers from 2,419.7 to 2,424.9 (+0.21%).
- 10.6. The impact of the amended assumptions increases as duration increases and there is some variation by age, but for the cell showing the greatest impact (ages up to 30, durations 5+) the effect on the adjusted results shown in Appendix C is to reduce the smoker value from 51.8% to 51.6% and to increase the non-smoker value from 53.0% to 53.1%.
- 10.7. In the second scenario, we have varied the off rates by duration. To restrict run-times, we have undertaken this scenario for male smokers only and have varied the off rates as shown in Table 10.1. These are intended to be reasonably consistent with the calculated off rates for 1999-2004 shown in Table 8.1.
- 10.8. The overall effect is to increase the number of expected settled claims for smokers from 2,419.7 to 2,423.5 (+0.15%).

Table 10.1. Assumed off rates by calendar year and duration for scenario test.

Duration	Calendar Year				
	1998 & prior	1999	2000	2001	2002
0	15.0%	16.0%	17.0%	19.0%	21.0%
1	12.5%	13.5%	14.5%	16.5%	18.5%
2	10.0%	11.0%	12.0%	14.0%	16.0%
3	7.5%	8.5%	9.5%	11.5%	13.5%
4	6.0%	7.0%	8.0%	10.0%	12.0%
5+	5.0%	6.0%	7.0%	9.0%	11.0%

10.9. The impact of the amended assumptions reduces as age increases and there is some variation by duration, but for the cell showing the greatest impact (ages up to 30, duration 2) the effect on the adjusted results shown in Appendix C is to reduce the value from 43.6% to 43.5%.

10.10. We consider the differences arising under each scenario to be small. As the increase in off rates for smokers in the first scenario exceeds the variation observed for any other factor in Table 8.3, we believe these results demonstrate the robustness of the approach to off rates, especially in the context of the other assumptions implicit in our work.

*Sensitivity to the claim development distribution*

10.11. In order to investigate the sensitivity of the adjusted results to the claim development distribution we have used four alternative distributions. The more extreme variants are the short and long distributions derived using the 95% confidence intervals around the parameter values, as described in paragraph 5.44. In addition we have used the mid-short and mid-long distributions depicted in Figure 6.9.

10.12. These alternative distributions are not intended to correspond to any particular subsets of data investigated in section 6. In most cases, the fits of the Burr model obtained for subsets of the data lie between the mid-short and mid-long distributions; the principal exceptions to this generality are:

- The distributions for claims settled in 1999, 2000, 2001 and 2003 are all generally shorter than the mid-short distribution; 2002 sits between mid-short and mid-long and, as noted in paragraph 6.18, no fit was obtained for claims settled in 2004;
- Some of the distributions fitted for the larger offices are generally shorter than the mid-short distribution, or longer than the mid-long distribution; and
- The distribution for stroke claims is generally longer than the mid-long distribution.

The Committee does not believe that these exceptions invalidate the use of the mid-short and the mid-long distribution as useful tests on the sensitivity of the 1999-2002 All Office results for all causes of claim combined.

10.13. Of the exceptions noted above, only some of the individual office fits and the latter part of the 1999 distribution sit outside the short and long distributions (specifically, the part of the 1999 distribution is shorter than the short distribution).

10.14. In addition many of the distributions derived for subsets of data lie outside the mid-short and mid-long distributions for the early part of the distribution, but these are usually only up to the second decile. This early part of the curve has limited impact on the adjusted results, so again the Committee is comfortable that the variations illustrated in this section are reasonably robust tests to apply within this paper.

10.15. To save run-time, these alternative distributions have been tested on the male smoker data only.

10.16. Table 10.2 shows the effect of the alternative distributions on the expected settled claims and the values of the adjusted results by duration.

10.17. The overall impact of the shorter distributions is that the mid-short distribution increases the expected settled claims by 2.3% to 2,475.3 whilst the short distribution increases it by 4.7% to 2,533.1. As a result the overall adjusted results reduce from 78.2% to 76.4% using the mid-short distribution and to 74.7% using the short distribution.

10.18. Conversely, the mid-long distribution reduces the expected settled claims by 2.4% to 2,362.6, whilst the long distribution reduces it by 6.5% to 2,261.8. As a result the overall adjusted results increase to 80.1% using the mid-long distribution and to 83.6% using the long distribution.

Table 10.2. (A) Expected settled claims by duration using the short, central and long claim development distributions; male smokers only.

Duration	Expected Settled Claims				
	Short	Mid-short	Central	Mid-long	Long
<b>0</b>	467.9	412.8	379.9	354.5	288.9
<b>1</b>	507.2	499.9	485.2	467.9	443.2
<b>2</b>	405.8	405.7	400.4	392.2	384.2
<b>3</b>	321.9	322.9	320.6	316.4	313.4
<b>4</b>	254.7	255.6	254.7	252.7	251.6
<b>5+</b>	575.7	578.3	578.9	578.9	580.7
<b>ALL</b>	<b>2,533.1</b>	<b>2,475.3</b>	<b>2,419.7</b>	<b>2,362.6</b>	<b>2,261.8</b>

Table 10.2. (B) Adjusted results by duration using the short, central and long claim development distributions; male smokers only

Duration	Adjusted Results				
	Short	Mid-short	Central	Mid-long	Long
<b>0</b>	68.2%	77.3%	84.0%	90.0%	110.4%
<b>1</b>	72.9%	74.0%	76.3%	79.1%	83.5%
<b>2</b>	71.5%	71.5%	72.4%	73.9%	75.5%
<b>3</b>	80.8%	80.5%	81.1%	82.2%	83.0%
<b>4</b>	82.8%	82.5%	82.9%	83.5%	83.9%
<b>5+</b>	76.8%	76.4%	76.3%	76.3%	76.1%
<b>ALL</b>	<b>74.7%</b>	<b>76.4%</b>	<b>78.2%</b>	<b>80.1%</b>	<b>83.6%</b>

10.19. The sensitivity of the results to the claim development distribution is clearly greater than the sensitivity to off rates considered earlier, and hence is worth considering in further detail.

10.20. The greatest impact of the different distributions is on the experience at duration 0. The increase in exposure means that using a shorter distribution generates additional expected settled claims at duration 0 in the investigation period, and hence reduces the adjusted results. The effect on the other durations is lower, because a change to the distribution moves some claims into a given duration but others out.

10.21. There is less variation between the age bands arising from the use of the alternative distributions; Table 10.3 shows the overall variations for the mid-short and mid-long distributions only, but much of the apparent variation arises from the different mix by duration.

Table 10.3. Expected settled claims and adjusted results by age band using the mid-short, central and mid-long claim development distributions; male smokers only.

Age band	Expected Settled Claims			Adjusted Results		
	Mid-short	Central	Mid-long	Mid-short	Central	Mid-long
< 30	243.7	235.6	227.7	55.8%	57.7%	59.7%
31-40	747.1	728.4	709.5	66.4%	68.1%	69.9%
41-50	817.4	799.2	780.6	84.4%	86.3%	88.4%
51-60	599.6	589.2	577.9	87.1%	88.6%	90.3%
61+	67.6	67.4	66.9	71.0%	71.2%	71.7%
<b>ALL</b>	<b>2,475.3</b>	<b>2,419.7</b>	<b>2,362.6</b>	<b>76.4%</b>	<b>78.2%</b>	<b>80.1%</b>

10.22. The Committee chose to use all of the 1999-2004 data for the central fit because of the substantially greater volume of claims data involved. It does not, therefore, necessarily consider that the 1999-2002 distribution is more appropriate for the 1999-2002 results in this paper, but the Committee felt it was important to also demonstrate the impact of using the claim development distribution derived from the 1999-2002 claims only. The distribution derived from the 1999-2002 claims was illustrated in Figure 6.5.

10.23. Table 10.4 shows the effect of using the 1999-2002 distribution on the expected settled claims and the adjusted results by duration, again for male smokers only.

Table 10.4. Expected settled claims and adjusted results by duration using the claim development distribution derived from 1999-2002 settled claims compared with the central distribution; male smokers only.

Duration	Expected Settled Claims		Adjusted Results	
	1999-2002	Central	1999-2002	Central
<b>0</b>	382.5	379.9	83.4%	84.0%
<b>1</b>	474.6	485.2	78.0%	76.3%
<b>2</b>	392.9	400.4	73.8%	72.4%
<b>3</b>	316.1	320.6	82.3%	81.1%
<b>4</b>	252.3	254.7	83.6%	82.9%
<b>5+</b>	578.1	578.9	76.5%	76.3%
<b>ALL</b>	<b>2,396.5</b>	<b>2,419.7</b>	<b>78.9%</b>	<b>78.2%</b>



- 10.24. The overall impact of using the 1999-2002 distribution is to reduce the number of expected settled claims in the investigation period, and hence increase the adjusted results by 1%. The impact varies by duration: the adjusted results are lower at duration 0 using the 1999-2002 distribution, with varying increases at other durations, reflecting different growth patterns.
- 10.25. The impact on the adjusted results of using only the 1999-2002 data is therefore considerably less than the impact of the mid-short and mid-long distributions, illustrated earlier, and probably introduces no greater difference in adjusted results than the more material of the other assumptions and approximations involved in their calculation.
- 10.26. Whereas we were able to demonstrate earlier in this section that the adjusted results are relatively insensitive to off rates, they most certainly are dependent on the claim development distribution used. Growth in business means that a change to the distribution alters how many claims are assumed to be settled in the investigation period. The shorter the distribution that is assumed, the higher the number of expected settled claims in 1999-2002 and the lower the adjusted results. This sensitivity does not invalidate the methodology, but does demonstrate the need to model the distribution as accurately as possible.
- 10.27. Finally, the Committee has also considered the effect of variations in both off rates and the claim development distribution, but has not included the results here since the combined impact is broadly similar to the sum of the separate effects. In particular, the impact on the adjusted results of varying the off rates in addition to the claim development distribution is not dissimilar from the effect of varying only the claim development distribution, which is the dominant sensitivity.

## **11. SUMMARY AND FURTHER WORK**

- 11.1. We believe that the developments outlined in this paper represent significant progress in our quest to provide a realistic assessment of UK critical illness claims experience.
- 11.2. The methodology that we have developed may appear quite different from the grossing-up factors introduced in Working Paper 14 that we first used to try to adjust the released results to properly match claims to exposure. It is not. The key area where we have developed the previous approach is to introduce modelling based on known dates of commencement and estimates of off rates to reduce the uncertainty inherent in the estimate of the growth in business that drove the grossing-up factor. This methodology was introduced in Working Paper 28 and developed further in this paper. Some further analysis of off rates is also included in this paper.
- 11.3. Another new area in this paper is the use of a parametric model of the interval between the diagnosis of a claim and its settlement, which we refer to as the claim development distribution. At an overall level, the claim development distribution used in this paper differs from that in Working Paper 14 principally because we have used data from 1999-2004 in its derivation, not just 1999-2002. Within this paper we have used the Burr model which appears to fit the overall set of claims well, but does not necessarily fit all subsets of the data, where the use of only three parameters may prevent a reasonable fit. The primary advantage of the parametric approach over the model used in Working Paper 14 is the additional flexibility, in particular providing us with an estimate for every day of the interval.
- 11.4. We have used the methodology and our best estimates of the claim development distribution and off rates to produce results for accelerated critical illness experience on a lives basis for the years 1999-2004. These results are referred to as 'adjusted' results in this paper. These are the first results that we have calculated that properly match claims to exposure, but they do so in terms of settled claims, not diagnosed claims, and so need careful interpretation.
- 11.5. In aggregate, the adjustment from the released results to the adjusted results is little different from our original estimate of the grossing-up factor. The ratio of the adjusted results to the released results is not directly comparable to the grossing-up factor in Working Paper 14, as there is a slight timing difference. However given that the vast majority of claims diagnosed in 1999-2002 are also settled during those years, the overall experience should not differ significantly whether measured in terms of diagnosed claims or settled claims. The same is not true for results for specific ages and durations however.
- 11.6. The refinements to the methodology have allowed us to also arrive at adjusted results for subsets of the data. These results require some care in their interpretation, since they are expressed in terms of settled claims. We do not expect patterns within the results by age or between smoker statuses to be materially different from those that would feature in results expressed in terms of diagnosed claims. However the pattern by duration may be fundamentally different since the claims settled at any particular duration are necessarily a mix of claims diagnosed at that and earlier durations.

- 11.7. We have also sought to test the sensitivity of the adjusted results to some of the assumptions. As shown in section 10, the list of assumptions is extensive and we have limited our sensitivity tests to variations in the two novel assumptions in this methodology – off rates and the claim development distribution. It is important to note the different conclusions for these – whereas the adjusted results are relatively insensitive to quite significant variations in off rates, they most certainly are dependent on the claim development distribution used.
- 11.8. Growth in business means that a change to the claim development distribution alters how many claims are expected to be settled in the investigation period. The shorter the distribution that is assumed, the higher the number of expected settled claims in 1999-2002 and the lower the adjusted results. This sensitivity does not invalidate the methodology, but does demonstrate the need to model the distribution as accurately as possible.
- 11.9. The Committee believes that it is important to consider this sensitivity in context. In particular there is inherent uncertainty within the CMI critical illness dataset arising from changes in the offices contributing data, different growth rates of offices, incomplete data on dates of claim, changing business practices and a myriad of other factors. This uncertainty is independent of the choice of methodology.
- 11.10. Whilst we have not tried to attach an exact probabilistic interpretation to the alternative claim development distributions illustrated within this paper, the Committee does believe that the short and long distributions represent quite extreme variations, for all office experience in 1999-2004 (though not necessarily for other groups of offices or other periods). The impact of the short and long distributions is to reduce the overall adjusted results by 4% and increase them by 7%, respectively. As with much of our analysis, this variation is most pronounced at duration 0 where the short distribution produces adjusted results of around 80%, and the long 130%, of the results using the central distribution.
- 11.11. Within Working Paper 14 we provided a table indicating the grossing-up factors for various rates of growth of business. This was intended to give individual offices an indication of the adjustment required to their released results to adjust for the understatement arising from the mis-match between exposure and claims. We have not provided anything similar in this paper, not least as we have moved away from simple estimates of the growth in business to modelling based on known dates of commencement. Such modelling ought to be unnecessary for an individual office as it can access its past data; it is also likely to need to consider carefully the claim development distribution it uses, since there is considerable variation between offices and the adjustment to the results is sensitive to this assumption. We have also observed significant variation in off rates between offices, but the adjustment appears to be less sensitive to the assumption here.
- 11.12. The absence of a table comparable with that in Working Paper 14 should not be misinterpreted as meaning that we no longer consider growth in business to be the key determinant of the adjustment required from the released results for an individual office. Our conviction on this remains unchanged. Our further work has only marginally altered

the aggregate adjustment and we do not believe that the adjustments for other growth rates would be significantly affected either.

- 11.13. Within this paper we have included high-level All Office adjusted results for 1999-2002. In addition we are releasing corresponding results to CMI member offices for the years 2003 and 2004. These results have only been produced for accelerated business on a “policies” basis (not amounts) and at an “All Causes” level, so clearly we have further work to do.
- 11.14. However our immediate priority is to try to use the methodology we have now developed to generate claim diagnosis rates, which may be suitable to graduate and produce a published table. It is our intention to undertake this at both an “All Causes” level and for key individual causes separately. The two should of course prove consistent.
- 11.15. We are also keen to apply the methodology to other areas, including amounts experience and considering how experience varies by factors such as sales channel, product type, benefit amount, commencement year and office.
- 11.16. We will also, in time, seek to develop a more robust underlying statistical model for critical illness business. This should increase clarity but we do not anticipate it will significantly alter our best estimate of the adjusted rates and hence we do not currently regard this as a priority.
- 11.17. In Working Paper 28, we noted an alternative development of the methodology in which an adjusted claim development distribution would be used to estimate dates of diagnosis for claims where these are unknown in order to estimate the diagnosed claims within each period. This might allow results to be derived for (estimated) Actual Diagnosed Claims / Expected Diagnosed Claims. The Committee has not investigated this approach further at this stage, but has made 1999-2002 data available to a PhD student who is investigating this, and other aspects of critical illness analysis. We look forward to hearing of her progress and hope to share this more widely as and when appropriate.
- 11.18. Many of the calculations underlying the work described in this paper are complex. We have undertaken extensive checking of the results and are confident that no material errors remain, however we note that most of the calculations can be undertaken from the 1999-2002 dataset that the Committee first made available in 2005 and we would welcome feedback from any actuary who attempts to reproduce the methodology using the CMI data.
- 11.19. Feedback on this paper is welcomed by the CMI. Views on the prioritisation of further work would be particularly appreciated. Please e-mail any feedback, preferably by 30 September 2008, to [ci@cmib.org.uk](mailto:ci@cmib.org.uk).

## **REFERENCES**

ABI Statement of Best Practice for Critical Illness Cover (April 2006)

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)

CMI Working Paper 28 : Progress towards an improved methodology for analysing CMI critical illness experience (July 2007)

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

Health Claims Forum : Recording of ‘Date of Claim’ for Critical Illness claims (November 2006) [see [http://www.actuaries.org.uk/knowledge/cmi/cmi\\_faqs/critical\\_illness\\_faqs](http://www.actuaries.org.uk/knowledge/cmi/cmi_faqs/critical_illness_faqs)]

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

## Appendix A: Defining the date of diagnosis

- A1. The CMI Critical Illness Committee has adopted the date of diagnosis as the key date for its analysis. The rationale for this is explained in section 5.3 of Working Paper 14.
- A2. That section also noted that the date of diagnosis was not well-defined and different dates could be assigned to the same claim by different offices or even by different claims assessors within an office. This can most easily be illustrated by means of examples. Some critical illness events are traumatic in nature so that the date of diagnosis is already reasonably well-defined – examples include Heart Attack and Coronary Artery Bypass Surgery. For some other causes of claim it is less well-defined. For example, the current ABI definition for Multiple Sclerosis requires both a definite diagnosis and 6 months' continuous impairment; so do claims assessors consider the date of diagnosis to be the date of the definitive diagnosis or after the 6 months of continuous impairment? Indeed this issue is likely to be most pronounced for claims under TPD (Total and Permanent Disability), which are notoriously difficult to settle, due to delays in establishing the permanence of the disability.
- A3. As a result, the Committee launched an initiative in conjunction with the Health Claims Forum (an organisation of claims assessors and claims managers whose purpose is to improve the management of health claims) to draw up guidelines for claims assessors on what constitutes the 'date of diagnosis' for Critical Illness claims. It is important to note that the initiative was intended to promote standardisation of reporting of the date of claim and was not intended to alter companies' decisions as to whether a critical illness claim is paid or not.
- A4. A consultation document was issued to Health Claims Forum members in May 2006 and guidelines, taking into account the feedback received, were issued by the Health Claims Forum in November 2006. This document can be obtained from the Health Claims Forum and can also be accessed via the CMI pages of the Profession's website (see 'References').
- A5. In essence, we now regard the date of diagnosis to be defined as “**the date at which the critical illness definition was fulfilled**”.
- A6. The interpretation of the date of diagnosis for each of the ABI definitions of critical illness events is set out in the appendix to the document. To illustrate these, the interpretations for some of the key critical illness events are shown in Table A1.
- A7. Where a company covers events not included in the ABI Statement of Best Practice or where definitions differ from the ABI model definitions, claims assessors are asked to follow the principle that the date of diagnosis is the date at which the critical illness definition was fulfilled, in particular:
- If the illness has a clear event date, then that date should be used; and
  - If the illness is a degenerative disease, then the date of diagnosis should allow for permanence to be established or for such other delay as is needed to ensure the definition is met.

Table A1. Definition of the date of diagnosis for selected critical illness claim events.

<b>Critical Illness claim event</b>	<b>Definition of date of diagnosis</b>
Cancer – <i>excluding less advanced cases</i>	Date of biopsy or other test that generates a definite diagnosis of Cancer that satisfies the policy definition
Coronary artery by-pass grafts – <i>with surgery to divide the breastbone</i>	Date of operation
Heart attack – <i>of specified severity</i>	Date of event
Multiple sclerosis – <i>with persisting symptoms</i>	The later of: <ul style="list-style-type: none"> <li>• Date of diagnosis of MS, and</li> <li>• The date 6 months after the start of a period of continuous symptoms</li> </ul>
Stroke – <i>resulting in permanent symptoms</i>	Date of event OR, if any uncertainty, date of evidence confirming the presence of permanent neurological deficit.

A8. Note that:

- Where an office requires permanence (of neurological deficit, for example) then the date of diagnosis will be later, reflecting the interval whilst permanence is established;
- There is currently no ABI definition for TPD, and hence the date of diagnosis is not explicitly defined within the HCF guidance. Following the principles noted above, we consider that the date of diagnosis for TPD should be the date when the permanence (and totality) of disability is clearly established; and
- Where the definition is dependent on medical evidence, then the date that the definition was fulfilled is the date on which the relevant evidence exists, not when it has been advised to the claimant, or to the claims assessor.

A9. Companies were asked to adopt this guidance from 1 January 2007 and the CMI hopes that this guidance has become established practice for claims assessors, thereby:

- Improving consistency within offices;
- Improving consistency between offices;
- Improving consistency over time; and
- Further increasing the proportion of claims where the date of diagnosis is recorded.

A10. The guidance will not have affected the data considered in this paper, but may materially alter future data. In particular if the HCF guidelines have been adopted to a substantial extent then we expect to see a significant reduction in the number of claims with long intervals between diagnosis and settlement in the years from 2007.

#### *Date of notification*

A11. The guidance issued by the HCF did not extend to date of notification.

A12. To date, we have defined this in the data submission requirements as “the date the office was notified of the claim”. The interpretation of this is ambiguous, in a similar way to date of diagnosis and, with hindsight, it would have been preferable to have clarified the interpretation of both dates through the HCF at the same time.

A13. Alternative approaches to the definition of date of notification can be justified, for example:

- (i) Consistency with the HCF definition of date of diagnosis would lead to date of notification being defined as:
- Date of notification of event for event-based claims (e.g. heart attack, stroke, death), and
  - Date of notification of final evidence for other claims, e.g. those demanding permanence.

Consistency with date of diagnosis gives this approach the benefit of acting as a data quality check on the various dates of claim supplied to the CMI.

- (ii) Such an approach could lead to the loss of interesting information regarding the overall claim process though, since we would no longer have details of when the claim was first submitted. To counteract this, date of notification could be defined as “the date a claim form is received by the insurer”. Such a definition has the additional benefit of being objective, but is likely to produce quite different relationships between date of diagnosis and date of notification from the previous example.

A14. Given that date of notification is not as important to our analysis as date of diagnosis, the Committee does not see standardisation of its interpretation as a key priority, but we welcome views on this.

A15. The Committee would also welcome views on whether further work to understand the claims process, for the different causes of claim, would be worthwhile.



## Appendix B: Reconciliation of results

B1. In this appendix we summarise the differences between the results for 1999-2002 that were released to CMI member offices in May 2005, the summary results in Working Paper 28 and those in this paper and explain how the differences arise.

### *Actual settled claims*

B2. A total of 10,310 claims on accelerated business were included in the May 2005 results, which included 21 claims where the smoker status was not provided. The analyses in Working Paper 28 and in this paper only use smoker-differentiated records and hence show a total of 10,289 claims.

B3. The allocation of claims by age and duration does alter however. The May 2005 results separated settled claims by age and duration at date of diagnosis (if known or, if not, using the date of diagnosis derived as set out in Table 5 of Working Paper 14). Within Working Paper 28, settled claims are shown by age and duration at date of settlement (if known or, if not, using 1<sup>st</sup> July in the year of settlement). As a result, claims tended towards older ages and longer durations in Working Paper 28, as illustrated for male non-smokers in Tables B1 and B2 below. Claims are shown by age and duration at both date of diagnosis and date of settlement in Appendix C (ASCd in column 1 and ASCs in column 5, respectively).

Table B1. Comparison of the allocation of actual settled claims by age band, according to whether age (nearest) is calculated at date of diagnosis or at date of settlement. Male non-smoker claims settled in 1999-2002.

Age	Actual Settled Claims by age at diagnosis	Actual Settled Claims by age at settlement	
	(1)	(2)	(2) / (1) %
< 30	460	416	90%
31-40	1,197	1,179	98%
41-50	1,319	1,306	99%
51-60	1,097	1,150	105%
61+	188	210	112%
<b>All</b>	<b>4,261</b>	<b>4,261</b>	<b>100%</b>

Table B2. Comparison of the allocation of actual settled claims by duration, according to whether curtate duration is calculated at date of diagnosis or at date of settlement. Male non-smoker claims settled in 1999-2002.

Curtate Duration	Actual Settled Claims by duration at diagnosis	Actual Settled Claims by duration at settlement	
	(1)	(2)	(2) / (1) %
0	760	513	68%
1	725	724	100%
2	658	623	95%
3	526	564	107%
4	394	440	112%
5+	1,198	1,397	117%
<b>All</b>	<b>4,261</b>	<b>4,261</b>	<b>100%</b>

B4. For both age and duration, it will be noted that the impact is greatest in the end categories. This is because claims can only go out of the lowest category (as one progresses from diagnosis to settlement), and only come into the highest category, whereas there are transitions in both directions in the intermediate categories. The impact on duration 0 is especially pronounced.

*Expected diagnosed claims*

B5. As noted in section 7, the same claim rates (from CIBT93) have been used in all three sets of results. The differences in the expected diagnosed claims between the three sets of results arise solely from differences in the calculation of exposure. These differ both in aggregate and in the allocation by age and duration.

B6. Let us consider first the overall change in the exposure. In the released results, exposure has been calculated using a census method based on start and end in-force data, with an addition of half the claims during each year. A total of 6,388,816 life-years of exposure on accelerated business was included in the May 2005 results. This included a small volume of data where the smoker status was not provided, totalling 41,266 life-years of exposure (corresponding to the 21 claims noted earlier in paragraph B2). The analyses in Working Paper 28 and in this paper only use smoker-differentiated records and hence the corresponding total exposure in the released results is 6,347,550 life-years.

B7. A census method of calculation was also adopted for Working Paper 28 but no addition was made in respect of the claims. As a result the total exposure should have been slightly lower at 6,344,533 life-years, a difference of about 0.03%. However the initial application included within Working Paper 28 was undertaken in spreadsheets and inadvertently, no exposure was included for durations in excess of 10 years, reducing the exposure by a further 1.1%.

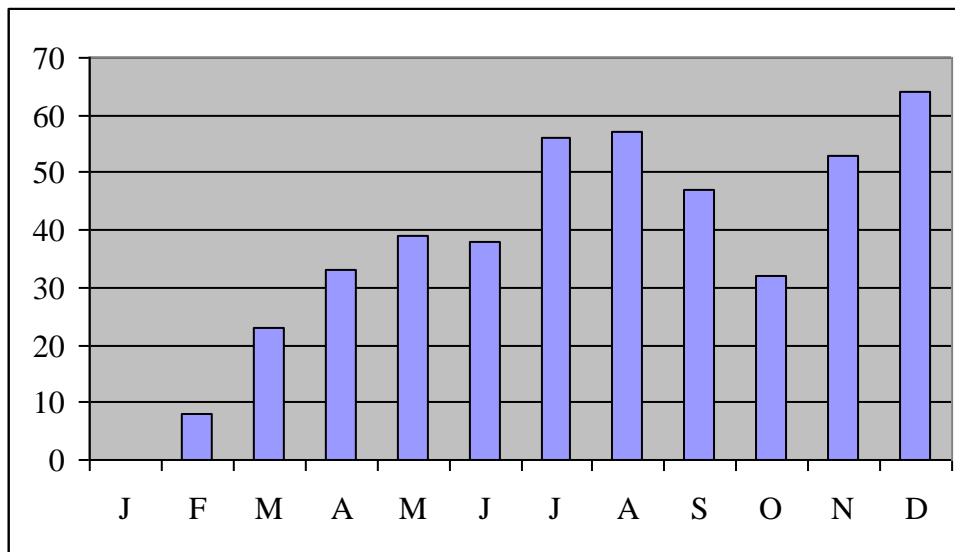
B8. Within this paper, exposure has been calculated taking account of dates of commencement of each record. This produces a more accurate value for exposure, although as dates of exit are not known, a simplifying assumption is still required in this regard. The effect of this change is to reduce the total exposure from that in the released results by about 0.6% to 6,308,558 life-years. Note that again, no addition in respect of the actual claims has been made in this paper and this explains a small part of the reduction, however most of the reduction arises from the use of actual commencement dates in the calculation of exposure.

B9. Whilst the overall differences in exposure between the results are small, there are significant differences when one looks in more detail. A comparison by calendar year and duration is shown in Table B3. The volumes of exposure generally reduce as duration increases, and hence the use of actual dates of commencement can have a more pronounced effect at longer durations. Note that diagonal patterns emerge, for example duration 10 in 1999 through to duration 13 in 2002. Exposure at such long durations arises from a small number of offices who started issuing policies in the early years of critical illness insurance in the UK and, in this example, the dates of commencement are heavily skewed towards the end of the year (see Figure B1); perhaps as a result of products being launched during the year. Similar effects, both positive and negative of smaller magnitude, can be observed at shorter durations.

Table B3. Comparison of exposure, by calendar year and duration (up to and including 12), all ages combined, calculated using actual dates of commencement (as in this paper) as a percentage of that calculated using a census method (as in the released results). All accelerated business in 1999-2002.

Curtate Duration	Calendar Year				
	1999	2000	2001	2002	All
0	-1.4%	-0.2%	-4.6%	1.1%	-1.2%
1	0.0%	-1.2%	0.9%	-3.5%	-1.2%
2	1.4%	0.2%	-1.2%	1.3%	0.4%
3	-2.3%	0.4%	0.5%	-0.3%	-0.3%
4	2.5%	-1.5%	-0.7%	-0.6%	-0.3%
5	-5.8%	4.5%	-0.1%	1.3%	0.7%
6	4.6%	-5.7%	4.9%	-0.6%	0.6%
7	-11.3%	2.2%	-6.2%	5.5%	0.0%
8	-3.8%	-12.4%	0.6%	-8.0%	-5.8%
9	11.2%	3.4%	-8.6%	2.3%	1.0%
10	-30.3%	9.8%	3.8%	-6.9%	-1.1%
11	-11.4%	-36.8%	5.6%	3.4%	-1.3%
12	-	-21.0%	-37.5%	1.9%	-8.7%
13	-	-29.3%	-13.8%	-29.6%	-28.5%
All	-0.6%	-0.4%	-1.3%	-0.2%	-0.6%

Figure B1. Number of records in first in force data submission from each office with date of commencement in 1988, by calendar month of commencement. All accelerated business submitted in 1999-2002.



B10. The overall difference in the expected diagnosed claims, between the released results and this paper, is very similar to the difference in the exposure. Table B4 shows the differences in exposure and in expected diagnosed claims by age band.

B11. The difference in the two sets of values for expected diagnosed claims is also apparent within Appendix C, where the released results are shown in column 2 (labelled EDC) and the figures using the exposure calculated in this paper are shown in column 4 (labelled EDC').

Table B4. Comparison of exposure, by age band for all durations combined, calculated using a census method (as in the released results) or using actual dates of commencement (as in this paper). The consequent effect on expected diagnosed claims is also shown. All accelerated business in 1999-2002.

Age Band	Exposure			Expected diagnosed claims		
	Census method	WP33 method		Census method	WP33 method	
	(1)	(2)	(2)/(1)%	(3)	(4)	(4)/(3)%
< 30	1,726,495	1,721,540	99.7%	2,204	2,185	99.1%
31-40	2,865,374	2,851,184	99.5%	6,821	6,781	99.4%
41-50	1,325,678	1,314,566	99.2%	7,542	7,472	99.1%
51-60	405,115	400,737	98.9%	5,477	5,409	98.8%
61+	24,888	24,425	98.1%	732	718	98.1%
<b>All</b>	<b>6,347,550</b>	<b>6,312,452</b>	<b>99.4%</b>	<b>22,776</b>	<b>22,565</b>	<b>99.1%</b>

*Expected settled claims*

B12. As noted previously, the results for 1999-2002 sent to member offices in May 2005 compared actual settled claims with expected diagnosed claims. No values were provided for expected settled claims other than an indication that the overall 1999-2002 All Office experience needed to be adjusted by 15%. Hence no further reconciliation with the released results is contained in this appendix.

B13. We have not attempted a detailed reconciliation of the values for expected settled claims in this paper with those included in Working Paper 28. Differences between the overall values arise from a number of sources, including:

- (i) The calculation of exposure, as we are now using exact dates of commencement;
- (ii) The inadvertent omission of exposure at durations 11+ from the spreadsheet model used for Working Paper 28;
- (iii) The use of different off rates for calendar years 1999 to 2002 in this paper, compared to the single off rate of 9% used in Working Paper 28;
- (iv) The use of a slightly “shorter” claim development distribution as a result of basing the distribution in this paper on claims settled in 1999-2004, rather than just 1999-2002;
- (v) The use of an inaccurate approximation to the Working Paper 14 claim development distribution in Working Paper 28; and
- (vi) The use of different groupings of expected diagnosed claims, most significantly by duration, which will affect whether expected settled claims occur inside or outside 1999-2002.

B14. Whilst we have included a reconciliation of the exposure during 1999-2002 between Working Paper 28 and this paper above; points (i) and (iii) also affect the exposure for the preceding years: (i) affects both known and synthetic exposure and (iii) affects synthetic exposure only. (Point (ii) does not affect the prior exposure, because no business had reached durations 11+ in those years). Points (iv) to (vi) only affect the translation of expected diagnosed claims into expected settled claims.

## **Appendix C: Accelerated critical illness experience 1999-2002**

C1. In this appendix we show summary results for 1999-2002. These are contained in four tables, as follows:

- Table C1: Male Non-Smokers
- Table C2: Male Smokers
- Table C3: Female Non-Smokers
- Table C4: Female Smokers

C2. These results are for Full Acceleration business on a Lives basis, covering all causes of claim (including mortality).

C3. The derivation of the values contained in the eight columns is explained in the following paragraphs. In all cases, claims are aggregated by age (nearest birthday) and (curtate) duration.

### *Column 1 – Actual Settled Claims (ASCd)*

C4. 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). The age and duration are as at the date of diagnosis (known or, if not, estimated as set out in Table 5 of Working Paper 14).

### *Column 2 – Expected Diagnosed Claims (EDC)*

C5. These are the claims that we expect to be diagnosed during 1999-2002 using CIBT93 to calculate the expected. The claim rates are unaltered from previous work; however it is important to note the explanation regarding our use of CIBT93 in paragraphs 7.41 to 7.44. The volume of exposure is as in the released results, rather than as estimated in this paper and hence the number of expected diagnosed claims is identical to those in the released results

### *Column 3 – 100 x Actual Settled Claims / Expected Diagnosed Claims (ASCd/EDC)*

C6. This is the comparison that we provided in the released results.

### *Column 4 – Expected Diagnosed Claims (EDC')*

C7. These figures differ from those in column 2 only because we have used slightly different measures of exposure in each case. These figures are based on the exposure estimated in this paper. A reconciliation with the exposure in the released results is contained in Appendix B. In aggregate the exposure is little changed and, as the claim rates are identical, expected diagnosed claims are also little affected.

### *Column 5 – Actual Settled Claims (ASCs)*

C8. 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). The age and duration are as at the date of settlement (known or, if not, estimated as 1<sup>st</sup> July in the year of settlement). The number of claims equals that in column 1 overall, but differs markedly in the categorisation by age and duration.

### *Column 6 – Expected Settled Claims (ESC)*

C9. These are the claims that we expect to be settled during 1999-2002 using:

- The methodology set out in section 7 of this paper to calculate the exposure including the off rates specified in Table 8.2;

- CIBT93 to calculate the expected diagnosed claims from the exposure, noting the explanation regarding our use of CIBT93 in paragraphs 7.41 to 7.44;
- The Burr model of the claim development distribution with parameters shown in Table 5.2, subject to truncation of the tail as set out in paragraph 5.50, to transform the expected diagnosed claims to expected settled claims.

C10. Note that these expected settled claims are derived using a mixture of population and insured data (claim rates and claim development distribution, respectively) and hence are not “expected” in the normal use of the term.

*Column 7 – 100 x Actual Settled Claims / Expected Settled Claims (ASCs/ESC)*

C11. This is a comparison of the values specified above. These are referred to as “adjusted” results in section 9.

*Column 8 – Ratio of ASCs/ESC to ASCd/EDC*

C12. This is a comparison of the adjusted results and the released results. This is discussed in section 9.

**Table C1: Male Non-Smokers.**

Full Acceleration business; Lives basis; All Causes (incl. mortality); 1999–2002.

Age Band	Based on age and duration at date of diagnosis				Based on age and duration at date of settlement			Ratio of ASCs/ESC to ASCd/EDC
	Actual Settled Claims (ASCd)	Expected Diagnosed Claims (EDC)	100 ASCd/ EDC	Expected Diagnosed Claims (EDC')	Actual Settled Claims (ASCs)	Expected Settled Claims (ESC)	100 ASCs/ ESC	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	<b>Diagnosed at duration 0</b>				<b>Settled at duration 0</b>			
< 30	133	320.9	41%	314.5	96	179.0	54%	129%
31-40	255	720.4	35%	713.0	171	398.4	43%	121%
41-50	209	785.7	27%	775.7	132	431.3	31%	115%
51-60	138	569.4	24%	560.0	96	312.4	31%	127%
61+	25	58.4	43%	57.3	18	32.0	56%	131%
<b>All</b>	<b>760</b>	<b>2,454.7</b>	<b>31%</b>	<b>2,420.6</b>	<b>513</b>	<b>1,353.1</b>	<b>38%</b>	<b>122%</b>
	<b>Diagnosed at duration 1</b>				<b>Settled at duration 1</b>			
< 30	115	218.1	53%	215.7	109	208.8	52%	98%
31-40	226	559.3	40%	553.5	228	514.5	44%	110%
41-50	184	615.1	30%	603.5	203	558.3	36%	120%
51-60	183	487.0	38%	477.6	164	438.7	37%	97%
61+	17	62.2	27%	61.2	20	55.8	36%	133%
<b>All</b>	<b>725</b>	<b>1941.7</b>	<b>37%</b>	<b>1,911.5</b>	<b>724</b>	<b>1,776.1</b>	<b>41%</b>	<b>111%</b>
	<b>Diagnosed at duration 2</b>				<b>Settled at duration 2</b>			
< 30	85	149.0	57%	149.8	65	155.1	42%	74%
31-40	193	448.2	43%	450.3	195	442.9	44%	102%
41-50	202	491.2	41%	491.4	179	481.4	37%	90%
51-60	153	394.7	39%	392.8	163	385.0	42%	108%
61+	25	60.5	41%	59.4	21	58.1	36%	88%
<b>All</b>	<b>658</b>	<b>1543.7</b>	<b>43%</b>	<b>1,543.6</b>	<b>623</b>	<b>1,522.4</b>	<b>41%</b>	<b>95%</b>
	<b>Diagnosed at duration 3</b>				<b>Settled at duration 3</b>			
< 30	66	99.4	66%	99.0	68	105.2	65%	98%
31-40	137	370.4	37%	369.1	143	368.8	39%	105%
41-50	173	415.2	42%	415.2	182	412.3	44%	105%
51-60	127	332.4	38%	331.6	146	330.0	44%	116%
61+	23	52.4	44%	53.1	25	53.6	47%	107%
<b>All</b>	<b>526</b>	<b>1269.8</b>	<b>41%</b>	<b>1,267.9</b>	<b>564</b>	<b>1,269.8</b>	<b>44%</b>	<b>107%</b>
	<b>Diagnosed at duration 4</b>				<b>Settled at duration 4</b>			
< 30	27	62.6	43%	63.0	39	68.1	57%	133%
31-40	115	301.3	38%	300.3	137	301.7	45%	118%
41-50	119	359.6	33%	356.5	122	354.6	34%	103%
51-60	122	290.7	42%	285	125	284.3	44%	105%
61+	11	43.9	25%	42.9	17	44.1	39%	156%
<b>All</b>	<b>394</b>	<b>1058.2</b>	<b>37%</b>	<b>1,047.7</b>	<b>440</b>	<b>1,052.6</b>	<b>42%</b>	<b>114%</b>
	<b>Diagnosed at durations 5+</b>				<b>Settled at durations 5+</b>			
< 30	34	65.9	52%	66.3	39	73.6	53%	102%
31-40	271	625.2	43%	623.7	305	633.9	48%	112%
41-50	432	1040.6	42%	1,033.9	488	1,032.1	47%	112%
51-60	374	1019.7	37%	1,009.3	456	1,006.4	45%	122%
61+	87	205.7	42%	200.7	109	203.1	54%	129%
<b>All</b>	<b>1,198</b>	<b>2957.2</b>	<b>41%</b>	<b>2,933.9</b>	<b>1,397</b>	<b>2,949.1</b>	<b>47%</b>	<b>115%</b>
	<b>Diagnosed at all durations</b>				<b>Settled at all durations</b>			
< 30	460	915.9	50%	908.3	416	789.8	53%	106%
31-40	1,197	3024.8	40%	3,009.9	1,179	2,660.1	44%	110%
41-50	1,319	3707.5	36%	3,676.2	1,306	3,270.0	40%	111%
51-60	1,097	3094.0	35%	3,056.3	1,150	2,756.7	42%	120%
61+	188	483.1	39%	474.6	210	446.6	47%	121%
<b>All</b>	<b>4,261</b>	<b>11,225.3</b>	<b>38%</b>	<b>11,125.2</b>	<b>4,261</b>	<b>9,923.2</b>	<b>43%</b>	<b>113%</b>

**Table C2: Male Smokers.**

Full Acceleration business; Lives basis; All Causes (incl. mortality); 1999–2002.

Age Band	Based on age and duration at date of diagnosis				Based on age and duration at date of settlement			Ratio of ASCs/ESC to ASCd/EDC
	Actual Settled Claims (ASCd)	Expected Diagnosed Claims (EDC)	100 ASCd/EDC	Expected Diagnosed Claims (EDC')	Actual Settled Claims (ASCs)	Expected Settled Claims (ESC)	100 ASCs/ESC	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	<b>Diagnosed at duration 0</b>				<b>Settled at duration 0</b>			
< 30	63	105.6	60%	102.8	46	58.6	78%	130%
31-40	124	226.9	55%	222.4	86	124.4	69%	125%
41-50	159	218.4	73%	213.9	113	119.4	95%	130%
51-60	91	132.7	69%	129.4	72	72.6	99%	143%
61+	4	8.9	45%	8.6	2	4.8	41%	91%
<b>All</b>	<b>441</b>	<b>692.4</b>	<b>64%</b>	<b>677.1</b>	<b>319</b>	<b>379.9</b>	<b>84%</b>	<b>131%</b>
	<b>Diagnosed at duration 1</b>				<b>Settled at duration 1</b>			
< 30	29	67.1	43%	66.5	30	65.4	46%	107%
31-40	105	166.2	63%	164.1	120	154.8	77%	122%
41-50	129	166.2	78%	162.7	129	152.5	85%	109%
51-60	87	113.0	77%	111.0	87	103.3	84%	109%
61+	5	10.4	48%	10.1	4	9.1	44%	92%
<b>All</b>	<b>355</b>	<b>522.9</b>	<b>68%</b>	<b>514.2</b>	<b>370</b>	<b>485.2</b>	<b>76%</b>	<b>112%</b>
	<b>Diagnosed at duration 2</b>				<b>Settled at duration 2</b>			
< 30	25	43.1	58%	43.5	20	45.9	44%	76%
31-40	94	125.6	75%	126.3	81	126.2	64%	85%
41-50	101	127.8	79%	128.2	103	127.6	81%	103%
51-60	77	91.4	84%	91.2	77	90.6	85%	101%
61+	3	10.2	29%	10.3	9	10.1	89%	307%
<b>All</b>	<b>300</b>	<b>398.1</b>	<b>75%</b>	<b>399.5</b>	<b>290</b>	<b>400.4</b>	<b>72%</b>	<b>96%</b>
	<b>Diagnosed at duration 3</b>				<b>Settled at duration 3</b>			
< 30	21	27.7	76%	27.7	20	29.9	67%	88%
31-40	70	98.7	71%	98.5	70	100.0	70%	99%
41-50	95	104.5	92%	103.9	100	104.7	96%	104%
51-60	70	76.6	91%	75.6	65	76.4	85%	93%
61+	5	9.8	51%	9.6	5	9.7	52%	102%
<b>All</b>	<b>262</b>	<b>317.3</b>	<b>83%</b>	<b>315.3</b>	<b>260</b>	<b>320.6</b>	<b>81%</b>	<b>98%</b>
	<b>Diagnosed at duration 4</b>				<b>Settled at duration 4</b>			
< 30	6	16.6	36%	16.8	11	18.4	60%	167%
31-40	45	76.7	59%	76.4	68	78.0	87%	147%
41-50	58	85.4	68%	85.0	60	86.0	70%	103%
51-60	48	64.6	74%	63.6	67	64.5	104%	141%
61+	7	8.2	85%	7.7	5	7.8	64%	75%
<b>All</b>	<b>164</b>	<b>251.6</b>	<b>65%</b>	<b>249.5</b>	<b>211</b>	<b>254.7</b>	<b>83%</b>	<b>128%</b>
	<b>Diagnosed at durations 5+</b>				<b>Settled at durations 5+</b>			
< 30	8	15.2	53%	15.4	9	17.4	52%	98%
31-40	60	141.2	42%	140.6	71	144.9	49%	117%
41-50	151	208.4	72%	206.5	185	209.1	88%	122%
51-60	134	182.9	73%	180.0	154	181.7	85%	116%
61+	17	25.9	66%	25.1	23	25.8	89%	135%
<b>All</b>	<b>370</b>	<b>573.6</b>	<b>65%</b>	<b>567.5</b>	<b>442</b>	<b>578.9</b>	<b>76%</b>	<b>117%</b>
	<b>Diagnosed at all durations</b>				<b>Settled at all durations</b>			
< 30	152	275.3	55%	272.6	136	235.6	58%	105%
31-40	498	835.3	60%	828.2	496	728.4	68%	114%
41-50	694	910.7	76%	900.2	690	799.2	86%	113%
51-60	507	661.3	77%	650.8	522	589.2	89%	116%
61+	41	73.3	56%	71.3	48	67.4	71%	127%
<b>All</b>	<b>1,892</b>	<b>2,755.9</b>	<b>69%</b>	<b>2,723.1</b>	<b>1,892</b>	<b>2,419.7</b>	<b>78%</b>	<b>114%</b>



**Table C3: Female Non-Smokers.**

Full Acceleration business; Lives basis; All Causes (incl. mortality); 1999–2002.

Age Band	Based on age and duration at date of diagnosis				Based on age and duration at date of settlement			Ratio of ASCs/ESC to ASCd/EDC
	Actual Settled Claims (ASCd)	Expected Diagnosed Claims (EDC)	100 ASCd/EDC	Expected Diagnosed Claims (EDC')	Actual Settled Claims (ASCs)	Expected Settled Claims (ESC)	100 ASCs/ESC	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	<b>Diagnosed at duration 0</b>				<b>Settled at duration 0</b>			
< 30	99	273.3	36%	268.1	60	151.7	40%	109%
31-40	253	568.5	45%	563.0	158	313.7	50%	113%
41-50	196	503.4	39%	496.0	123	276.2	45%	114%
51-60	107	261.5	41%	256.9	72	142.9	50%	123%
61+	4	18.0	22%	17.9	3	9.9	30%	136%
<b>All</b>	<b>659</b>	<b>1,624.7</b>	<b>41%</b>	<b>1,601.9</b>	<b>416</b>	<b>894.4</b>	<b>47%</b>	<b>115%</b>
	<b>Diagnosed at duration 1</b>				<b>Settled at duration 1</b>			
< 30	79	191.6	41%	189.5	81	180.6	45%	109%
31-40	222	443.3	50%	437.8	205	405.9	51%	101%
41-50	189	398.7	47%	391.3	182	362.1	50%	106%
51-60	96	225.2	43%	222.5	97	204.3	47%	111%
61+	7	19.0	37%	18.6	2	16.8	12%	32%
<b>All</b>	<b>593</b>	<b>1,277.9</b>	<b>46%</b>	<b>1,259.6</b>	<b>567</b>	<b>1,169.7</b>	<b>48%</b>	<b>104%</b>
	<b>Diagnosed at duration 2</b>				<b>Settled at duration 2</b>			
< 30	71	134.5	53%	135.2	71	137.5	52%	98%
31-40	186	358.0	52%	359.7	192	352.7	54%	105%
41-50	164	322.7	51%	323.3	180	316.6	57%	112%
51-60	78	187.5	42%	187.4	95	183.5	52%	124%
61+	8	18.5	43%	18.0	10	17.4	58%	133%
<b>All</b>	<b>507</b>	<b>1,021.2</b>	<b>50%</b>	<b>1,023.5</b>	<b>548</b>	<b>1,007.7</b>	<b>54%</b>	<b>110%</b>
	<b>Diagnosed at duration 3</b>				<b>Settled at duration 3</b>			
< 30	42	93.3	45%	92.8	39	96.6	40%	90%
31-40	161	299.4	54%	298.6	147	297.4	49%	92%
41-50	139	275.8	50%	274.6	130	272.7	48%	95%
51-60	57	159.4	36%	158.9	72	158.3	45%	127%
61+	6	16.2	37%	16.5	9	16.6	54%	146%
<b>All</b>	<b>405</b>	<b>844.0</b>	<b>48%</b>	<b>841.3</b>	<b>397</b>	<b>841.6</b>	<b>47%</b>	<b>98%</b>
	<b>Diagnosed at duration 4</b>				<b>Settled at duration 4</b>			
< 30	29	61.7	47%	62.0	31	65.4	47%	101%
31-40	120	247.4	49%	246.3	142	246.6	58%	119%
41-50	119	238.4	50%	236.5	134	235.6	57%	114%
51-60	57	139.0	41%	136.9	52	136.8	38%	93%
61+	4	14.4	28%	13.7	5	14.0	36%	129%
<b>All</b>	<b>329</b>	<b>700.8</b>	<b>47%</b>	<b>695.4</b>	<b>364</b>	<b>698.4</b>	<b>52%</b>	<b>111%</b>
	<b>Diagnosed at durations 5+</b>				<b>Settled at durations 5+</b>			
< 30	38	68.2	56%	68.7	43	74.3	58%	104%
31-40	238	530.0	45%	527.3	310	534.3	58%	129%
41-50	276	663.5	42%	658.5	325	658.9	49%	119%
51-60	193	447.1	43%	441.3	259	442.8	58%	136%
61+	19	66.9	28%	65.5	28	66.3	42%	149%
<b>All</b>	<b>764</b>	<b>1,775.7</b>	<b>43%</b>	<b>1,761.3</b>	<b>965</b>	<b>1,776.6</b>	<b>54%</b>	<b>126%</b>
	<b>Diagnosed at all durations</b>				<b>Settled at all durations</b>			
< 30	358	822.6	44%	816.2	325	706.1	46%	106%
31-40	1180	2446.5	48%	2,432.7	1,154	2,150.6	54%	111%
41-50	1083	2402.5	45%	2,380.1	1,074	2,122.1	51%	112%
51-60	588	1419.8	41%	1,403.9	647	1,268.7	51%	123%
61+	48	152.9	31%	150.1	57	141.0	40%	129%
<b>All</b>	<b>3,257</b>	<b>7,244.3</b>	<b>45%</b>	<b>7,183.1</b>	<b>3,257</b>	<b>6,388.4</b>	<b>51%</b>	<b>113%</b>

**Table C4: Female Smokers.**

Full Acceleration business; Lives basis; All Causes (incl. mortality); 1999–2002.

Age Band	Based on age and duration at date of diagnosis				Based on age and duration at date of settlement			Ratio of ASCs/ESC to ASCd/EDC
	Actual Settled Claims (ASCd)	Expected Diagnosed Claims (EDC)	100 ASCd/EDC	Expected Diagnosed Claims (EDC')	Actual Settled Claims (ASCs)	Expected Settled Claims (ESC)	100 ASCs/ESC	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	<b>Diagnosed at duration 0</b>				<b>Settled at duration 0</b>			
< 30	26	67.7	38%	65.7	20	37.1	54%	140%
31-40	63	135.8	46%	133.4	41	74.6	55%	118%
41-50	55	123.7	44%	121.4	33	67.8	49%	109%
51-60	31	57.8	54%	56.4	23	31.6	73%	136%
61+	1	2.9	34%	2.9	0	1.6	0%	-
<b>All</b>	<b>176</b>	<b>387.9</b>	<b>45%</b>	<b>379.7</b>	<b>117</b>	<b>212.8</b>	<b>55%</b>	<b>121%</b>
	<b>Diagnosed at duration 1</b>				<b>Settled at duration 1</b>			
< 30	27	45.0	60%	44.5	20	43.0	47%	78%
31-40	56	100.5	56%	99.4	59	93.3	63%	113%
41-50	53	96.0	55%	94.4	53	88.3	60%	109%
51-60	38	50.5	75%	49.9	31	46.4	67%	89%
61+	3	3.3	91%	3.1	4	2.8	141%	157%
<b>All</b>	<b>177</b>	<b>295.4</b>	<b>60%</b>	<b>291.3</b>	<b>167</b>	<b>273.8</b>	<b>61%</b>	<b>102%</b>
	<b>Diagnosed at duration 2</b>				<b>Settled at duration 2</b>			
< 30	14	30.5	46%	30.7	21	31.7	66%	144%
31-40	42	76.6	55%	76.9	39	76.5	51%	93%
41-50	51	73.8	69%	74.2	52	73.7	71%	102%
51-60	26	41.7	62%	41.7	28	41.2	68%	109%
61+	3	3.0	102%	3.0	4	2.9	136%	138%
<b>All</b>	<b>136</b>	<b>225.6</b>	<b>60%</b>	<b>226.5</b>	<b>144</b>	<b>226.1</b>	<b>64%</b>	<b>106%</b>
	<b>Diagnosed at duration 3</b>				<b>Settled at duration 3</b>			
< 30	12	20.7	58%	20.6	9	21.7	41%	72%
31-40	27	61.0	44%	61.1	31	61.7	50%	114%
41-50	42	60.3	70%	59.7	39	60.1	65%	93%
51-60	19	35.8	53%	35.7	29	35.7	81%	153%
61+	1	2.7	38%	2.7	1	2.7	37%	100%
<b>All</b>	<b>101</b>	<b>180.4</b>	<b>56%</b>	<b>179.7</b>	<b>109</b>	<b>182.0</b>	<b>60%</b>	<b>107%</b>
	<b>Diagnosed at duration 4</b>				<b>Settled at duration 4</b>			
< 30	8	13.1	61%	13.2	8	14.1	57%	93%
31-40	23	48.9	47%	48.6	21	49.4	43%	90%
41-50	36	49.7	72%	49.5	42	50.0	84%	116%
51-60	14	30.5	46%	30.0	16	30.4	53%	115%
61+	0	2.3	0%	2.3	0	2.3	0%	-
<b>All</b>	<b>81</b>	<b>144.9</b>	<b>56%</b>	<b>143.5</b>	<b>87</b>	<b>146.1</b>	<b>60%</b>	<b>107%</b>
	<b>Diagnosed at durations 5+</b>				<b>Settled at durations 5+</b>			
< 30	7	13.3	53%	13.4	10	14.8	68%	128%
31-40	54	91.2	59%	90.6	67	93.0	72%	122%
41-50	72	117.2	61%	116.5	89	118.1	75%	123%
51-60	62	86.1	72%	84.8	76	85.8	89%	123%
61+	13	8.8	147%	8.1	13	8.5	153%	104%
<b>All</b>	<b>208</b>	<b>316.6</b>	<b>66%</b>	<b>313.4</b>	<b>255</b>	<b>320.1</b>	<b>80%</b>	<b>121%</b>
	<b>Diagnosed at all durations</b>				<b>Settled at all durations</b>			
< 30	94	190.2	49%	188.1	88	162.4	54%	110%
31-40	265	514.2	52%	509.9	258	448.5	58%	112%
41-50	309	520.8	59%	515.6	308	458.0	67%	113%
51-60	190	302.3	63%	298.4	203	271.0	75%	119%
61+	21	23.0	91%	22.0	22	20.9	105%	115%
<b>All</b>	<b>879</b>	<b>1,550.6</b>	<b>57%</b>	<b>1,534.0</b>	<b>879</b>	<b>1,360.8</b>	<b>65%</b>	<b>114%</b>