

Continuous Mortality Investigation

SAPS Mortality Committee

WORKING PAPER 34

Methodology and assumptions used for CMI Self-Administered Pension Schemes mortality experience analyses

October 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.

The Institute of Actuaries, the Faculty of Actuaries, the CMI and its Committee Members do not accept or assume any responsibility for the use of this document by any party in any context. This document does not provide any form of guidance and should not be relied on as such.

Continuous Mortality Investigation

SAPS Mortality Committee

WORKING PAPER 34

Methodology and assumptions used for CMI Self-Administered Pension Schemes mortality experience analyses

CONTENTS

1. Introduction	3
2. Results production process	5
3. Assumptions and approximations	8
4. Exposure Calculation	12
Appendix A: Overview of processing stages	15
Appendix B: Data checks within SAPS program	17
Appendix C: Analysis of effect of methodology changes	19

**CMI Working Paper 34:
Methodology and assumptions used for CMI Self-Administered
Pension Schemes mortality experience analyses**

1. Introduction

- 1.1. The SAPS Mortality investigation has so far produced five ‘All Schemes’ experience analyses, presented in Working Papers 4, 9, 17, 29 & 31. The next analysis, based on data collected to June 2008, will be released in a forthcoming report.
- 1.2. Draft graduations of the SAPS dataset were released in consultation form in Working Paper 32 (January 2008). Working Paper 35, containing feedback to this consultation, the SAPS Mortality Committee’s response to the feedback, and the revised graduations themselves, is released alongside this document.
- 1.3. As well as these published analyses of combined experience, the SAPS mortality investigation produces ongoing output in the form of individual scheme results. These are produced for every set of pension scheme data submitted, and have been provided to the scheme actuary that submitted the data since data collection began in early 2003.
- 1.4. The methodology and assumptions underlying both the individual and ‘All Schemes’ results have not previously been publicly documented. Work on the graduations resulted in a change to the exposure calculation method, and this presented a good opportunity to review other aspects of the methodology and revise where appropriate. This Working Paper documents the approach used for all outputs (except the graduations methodology which is given separately in Working Paper 35). It contains all recent changes and the reasoning behind them, but is intended as a ‘static’ paper that can be used to provide further detail on methodology in use at any point in time up to and including Working Paper 35.
- 1.5. To provide context for the methodology and assumptions involved, the results production process, including data validation and results checking, is outlined in section 2. Changes to this process are given where they have evolved over time.
- 1.6. Section 3 then considers the assumptions and approximations made in the SAPS mortality analyses. These have not previously been comprehensively documented and, in some cases, the approach has been amended from that agreed when the SAPS mortality investigation began. Changes to the previous approach are described, together with the rationale for the change.
- 1.7. The most significant change in methodology is to calculate central exposures directly for the final graduations, whereas these were estimated from initial exposures for the draft graduations. A detailed explanation of the approach now used is contained in section 4, whilst the limitations with the previous approach and background to the change are provided in a separate document (entitled ‘Comparison of approaches for calculating initial exposure’ and available on request).
- 1.8. The aggregated impact on the graduations dataset of all retrospective changes is shown in Appendix C.

1.9. It is also intended that central rather than initial exposed to risk will be used for all future analyses as well as for the graduations. The SAPS Mortality Committee will give consideration to when this change will be made and what analysis will be made available to demonstrate its impact.

1.10. Comments and feedback on this Working Paper, as well as requests for the document 'Comparison of approaches for calculating initial exposure', should be sent to:

Simon Spencer, CMI, Cheapside House, 138 Cheapside, London, EC2V 6BW

Email: self-admin@cmib.org.uk

2. Results production process

- 2.1. This section outlines the production of individual scheme results and how it has evolved since the outset of the investigation. The process starts with the submission of data for a particular scheme and covers validation of that data and the checking of results once they have been produced by the SAPS program. The final stage is the inclusion, or otherwise, of the scheme's data within the SAPS 'All Schemes' dataset. A diagrammatical representation of this process is contained in Appendix A.
- 2.2. When the SAPS Mortality investigation was set up, budget constraints minimised the time spent by the Secretariat on the manual elements of processing. This means that the investigation has been heavily reliant on the data providers in terms of:
 - Formatting of data according to the requirements of the SAPS Coding Guide
 - Data checking before submission
 - Checking of results
 - Correcting data and resubmitting where necessary
- 2.3. However, as the investigation has progressed, data issues that have come to light have necessitated wider checking of the 'All Schemes' dataset by the Secretariat.
- 2.4. This means that the process has evolved as the dataset has increased in size. The sections below document the changes, starting with the basic process and then specifying the checks performed for the datasets underlying each of the Working Papers so far produced.

Basic process

- 2.5. Data should be submitted following its receipt and acceptance by the scheme actuary for use in a scheme valuation and therefore it is expected that it will have already been checked by the scheme actuary. Once submitted, the validation process is largely automated.
- 2.6. A full list of the checks performed by the SAPS program as it uploads the data is given in Appendix B. First are a limited number of basic checks on the data for the scheme as a whole (see B.1 and B.2). Provided these are fulfilled, checks are then performed on each member record in turn (see B.3 to B.5). A scheme's data can be analysed if at least one record is uploaded. Results are therefore run on the validated records and all rejected records are listed in an error report.
- 2.7. This error report also provides some summary figures such as the number of validated pensioners, deaths, 'other' exits and new pensioners during the investigation period, plus average pensions and pension increases or decreases over the period for different pensioner groups. If postcode data has been provided then invalid postcodes are listed, though member records without valid postcodes are still included (i.e. only the postcode field is omitted).
- 2.8. The results and error report are sent to the data provider with a request that they check that results look as expected. They are also asked to correct and resubmit the data should these checks reveal any data issues. Such resubmissions then overwrite the previous submission for that scheme.

Datasets for Working Papers 4, 9 and 17

2.9. The datasets for these Working Papers comprised all scheme data which had not been overwritten by a resubmission. The onus at this stage of the investigation was very much on the data providers to check the results and take the appropriate action.

Dataset for Working Paper 29

2.10. The analysis leading to Working Paper 29 (which was based on data collected to June 2006) revealed various inconsistencies in the data. This led to a thorough investigation of the entire dataset and the resulting exclusion of a number of schemes where there were significant questions as to the reliability of the data.

- Nine schemes were removed with results that looked odd or data issues that were highlighted, such as no deaths for all records or for a particular subset.
- Thirteen schemes were excluded as amounts data had been provided for a low proportion or none of their death records. The program uses scheme-specific averages to estimate missing amounts, and the concern was that the amounts analysis may be skewed if this estimation was used for most of the deaths.

This had the effect of reducing the exposures in the years 2000 and 2001 compared with Working Paper 17, but this was more than offset by the increased confidence in the quality of the data used for analysis.

Dataset for Working Papers 31 and 32

2.11. These were based on data collected to June 2007. All checks performed on the Working Paper 29 data were put in place for future data submissions and were therefore used for the data underlying Working Papers 31 and 32.

2.12. In particular, following this data cleaning exercise, schemes are now marked as 'Not Live' if suspect features are identified from the results checks by the Secretariat, and 'Live' otherwise. Those assigned 'Not Live' status are queried with data providers, and are only included if the provider verifies or subsequently corrects the data. For 'Live' schemes, confirmation from the data provider is still sought, so as to encourage checking of results, but it is not deemed essential for the inclusion of that scheme in the 'All Schemes' list.

Dataset for Working Paper 35

2.13. Working Paper 35 is also based on the data collected to June 2007. The graduation work and other analysis revealed further data issues which had not previously been identified. The entire dataset was checked for the three issues noted below and data for a number of schemes was queried. In several cases the data was amended or resubmitted but there were thirteen schemes for which issues were not resolved by the time of finalising the dataset:

- Three schemes were removed because the given extract date for their data was more than 30 days before the investigation end date. It was agreed that some schemes may genuinely extract data before the valuation date but that extracting it over a month before was more likely to be a data error. The reliability of this data was therefore questionable.
- Eight schemes were removed because the same date of death had been used for a high proportion of records. The assigned dates did not spread the deaths evenly over the period of investigation, leading to concern that this may skew the analysis by age.

- Two schemes had an unusually high number of deaths at certain ages. These were removed due to concern that 100% commutations may have been miscoded as deaths.

2.14. Additionally, amendments to certain records (rather than entire schemes) were required:

- A small number of negative pension amounts were found within the dataset. The affected records were removed (with the remaining records for the associated schemes left intact) and a check was added to ensure that negative pension values for future submissions would not be uploaded.
- A coding error meant that missing pension amounts given as a string of zeroes (rather than being blank or a single “0”) were uploaded as zero pensions and no amount estimation was performed (see 3.7 - 3.13 for details on this estimation). To correct this, the program was amended to recognise strings of zeroes as missing amounts, and these records were then reloaded. Their amounts were estimated on reload so that no zero pensions remained in the dataset.

The aggregate effect on the graduations dataset of all data amendments is shown in Appendix C.

2.15. A further investigation was subsequently performed into schemes that provided no dependant data due to concern that dependants may have been miscoded as pensioners in some cases. 65 submissions included no dependant data, and a further eight contained pensioner codes that might be inconsistent with the other data (for example, a scheme containing data for Normal Health pensioners but none for Ill-health pensioners). All were queried but as the graduation work was almost complete no action was taken for these schemes. They have all been included within the graduations dataset, and the majority of feedback so far indicates that the data has been correctly classified.

2.16. These more extensive checks are to be performed on all future data submissions and incorporated into the decisions regarding the schemes’ ‘Live’ status.

3. Assumptions and approximations

3.1. The SAPS program makes several assumptions and approximations when uploading and analysing a scheme's data. An approach was agreed when the SAPS Mortality investigation began, and this remained broadly unchanged until the recent decision to move from initial to central exposure (see section 4) prompted a review of other aspects of the methodology. Where it was felt that improvements could be made, it seemed sensible to make these at the same time as the switch in exposure calculation. For individual results production, all the changes in this section were implemented for schemes submitted from 1 July 2008. Some changes apply retrospectively so affect the data underlying the final graduations, whilst other changes will only apply to future data submissions. The aggregated impact on the graduations dataset of all retrospective changes is shown in Appendix C.

3.2. Assumptions are required in the following areas:

- Start and end dates of the investigation period
- Pension amounts at the start and/or end of the investigation period (if missing for any record)
- Pension review dates during investigation period (if missing for any record)
- Inter-valuation pension amounts for all records (used for amounts exposure calculation plus analysis by pension band)
- Late-reported deaths adjustment (for any scheme with extract date within a month of end date)
- Allocation of exposure and deaths for pension bands analysis

For each assumption, the new methodology is detailed below along with the rationale behind any changes made.

Start and End of investigation period

3.3. The coding guide does not clearly specify whether the investigation start and end dates should be expressed as e.g. 6/4/2000-5/4/2003 or 5/4/2000-5/4/2003, where the most recent valuation date is 5/4/2003. Both are used by data contributors with the majority providing the latter. (The coding guide will be amended to encourage consistent coding of investigation start and end dates across schemes, in line with the latter approach in the near future.)

3.4. The previous approach assumed that exposure ran from the day after the given start date until the given end date (inclusive). This is not an ideal treatment in all cases, however. If, for example, 6/4/2000-5/4/2003 dates are provided then exposure is counted from 7/4/2000 which is likely to be one day's less exposure than the data provider intended.

3.5. In addition deaths on the given start date were included in the analysis, even though exposure was not deemed to start until the next day.

3.6. The new approach deals separately with the different types of date range given. It is hoped this allocates the correct number of days' exposure regardless of which way the valuation period is specified. There are three different cases which are treated as follows:

1. Where the end date is an exact anniversary of the start date but this is not the 1st of any month or 6th April, exposure starts on the day following the start date

and ends on the end date. E.g. 5/4/2000-5/4/2003 exposure runs from 6/4/2000 to 5/4/2003.

2. Where the end date is an exact anniversary of the start date and either these are both the 1st of any month or both 6th April, exposure starts on the start date and ends on the day before the end date. E.g. 6/4/2000-6/4/2003 exposure runs from 6/4/2000 to 5/4/2003, and 1/1/2000-1/1/2003 exposure runs from 1/1/2000 to 31/12/2002.
3. If the start date is not an exact anniversary of the end date, exposure starts and ends on the dates given. E.g. 6/4/2000-5/4/2003 exposure runs from 6/4/2000 to 5/4/2003.

In all these cases, deaths are now included only if they occur on a day which has been included within the exposure.

Estimation of missing pension amounts

- 3.7. The coding guide requests that annual pension amounts applying at both the start and end of each member's investigation period are provided if possible. If one or both amounts are unavailable, the program estimates these to enable an amounts analysis to be performed.
- 3.8. Note that schemes with no pension amounts data at all are not analysed as the program needs sufficient data for that scheme on which to base its estimates. Additionally, schemes where amounts data is provided for a low proportion of death records are analysed, but they are not included in the 'All Schemes' dataset as explained in 2.10.
- 3.9. The program first estimates missing amounts for all members where either the start or end pension has been provided. The missing start (end) amount is based on the provided end (start) amount, assuming that pensions are increased by RPI. The RPI values applying at member's start and end dates are used for this.
- 3.10. It should be noted that the RPI figures used in the calculations above are assumed to be those applicable at the time the pension amount is being estimated. In practice the RPI figures applicable at a particular date may not yet be known due to the time lag involved in publishing the Index. In this case the most recent available RPI value is used.
- 3.11. A coding error meant that in certain circumstances, the estimation of start amounts used RPI values from the month preceding the start of exposure. The RPI values from the correct months are now used in all cases.
- 3.12. The program then estimates missing pensions for all members for whom neither amount has been provided. These are based on the average start and end pension amounts for "similar" members within the scheme.

"Similar" members are defined initially as those with the same gender and pensioner type who fall into the same 5-year 'age' band (where this 'age' is actually the difference between the year of birth and the start year of the investigation period). Note that for this purpose the average start pension is calculated using both members exposed at the start of the period and those who enter the investigation during the period.

If no "similar" members exist then the 'age' band criterion is dropped.

If there are still no “similar” members then the pension amount is set to 1. This ensures that the member is included in the lives analysis, whilst having minimal impact on the amounts analysis, provided only a small number of lives are affected.

- 3.13. Note that no changes have been made to the method for estimating missing pension amounts, so the only change to the dataset arises from the correction of the coding error noted in 3.11.

Pension review dates during investigation period

- 3.14. To estimate pension amounts during the investigation period a pension review date is needed for each member. On this date in each calendar year it is assumed that the pension amount changes (unless end pension is equal to start pension). If no review date has been provided for a member then this is estimated as the anniversary in each calendar year of the day following the scheme end date. (For increasing pensions, this assumption maximises the amounts exposure and hence generates lower values of mortality rates.)

- 3.15. A coding error meant that before July 2008, some review dates were incorrectly interpreted. These dates are now correctly handled for all analyses.

Inter-valuation pension amounts

- 3.16. The pensions at start and end are used to estimate the inter-valuation amounts attained by a member on each review date. These allow the exposure in any calendar year to be weighted by the pension amount applying at that time, facilitating an amounts analysis. When analysing by pension band, these inter-valuation amounts are also used to place the exposure for each member into the appropriate band.

- 3.17. For increasing pensions (i.e. end pension greater than start) the amount is increased on the member’s review date in each calendar year from the start amount to the end amount:

- If RPI has increased overall during the investigation period, it is assumed that each increase is proportional to the changes in RPI.
- If RPI has not increased over the investigation period, it is assumed that the increase is linear.

- 3.18. This is a slight alteration to the previous method, which based all inter-valuation amounts on the end pension only, taking no account of the start pension. This created a potential discontinuity between the start pension and the amount attained at the first review.

- 3.19. No changes have been made to the approach for decreasing or level pensions but they are given for completeness. The approach taken for decreasing pensions is to linearly reduce the pension on the member’s review date in each calendar year, from the start amount to the end amount. For level pensions, this amount is used throughout the investigation period.

Late-reported deaths

- 3.20. For schemes where the data extract is produced within a month (either side) of the end of the investigation period an adjustment is made to the exposure to allow for late-reported deaths. It is assumed that information on deaths from the end of the period is likely to be missing, and therefore this end date is brought forward by:
$$30 - (\text{Min}[30, \text{Extract date} - \text{End date}])$$
- 3.21. If the extract date is within 30 days before the end date, this adjustment brings the end of the investigation period forward to exactly 30 days prior to the extract date. Note that if the extract date precedes the valuation date by more than 30 days the scheme is not now processed (see 2.13).
- 3.22. This method is slightly altered from the pre-July 2008 approach, and addresses the following shortcomings identified with the previous adjustment:
- The adjustment was previously made only to the exposure, and no adjustment was made to the deaths. Deaths shortly before the valuation date were therefore included even if exposure was deemed to have ceased.
 - The adjustment only applied to the last calendar year. In particular, if a scheme had a valuation date in January, no adjustment was made to the exposure in the preceding December.
 - The adjustment applied to all of the exposure for the calendar year. Hence when a life aged from x to $x+1$ during the year, the exposure at both ages was reduced, even though the adjustment was in respect of the “end” of the period.
 - The approach was not designed to deal with cases where the extract date preceded the end date, and the resulting reduction in exposure in these cases was too great.
- 3.23. The Committee intends to carry out an analysis of the current practice to determine whether 30 days remains appropriate for both purposes for which it is used, although no timetable for this work has yet been set.

Allocation of exposure and deaths for pension bands analysis

- 3.24. Several previous ‘All Schemes’ analyses have split the data into pension amounts bands to analyse separately the experience in each. The approach used for this is unchanged but is given here for completeness.
- 3.25. Note that fixed bands are applied across the period of investigation (which was 2000-2006 for Working Paper 31 for example), with no allowance for inflationary increases.
- 3.26. Each member will have (provided or otherwise estimated) start and end pension amounts from which their inter-valuation amounts are determined. As described in 3.16, it is assumed that these are attained on the member’s review date in each calendar year. These review dates and associated amounts therefore determine the annual pension applying at any point in the member’s period of exposure. When performing an analysis by pension band, each day of exposure is allocated to the correct band according to the pension amount applying on that day.
- 3.27. Deaths are allocated according to the amount of pension in payment at the date of death.

4. Exposure Calculation

- 4.1. Initial exposed to risk was used for all SAPS experience analyses from the start of the investigation in 2003. It was considered that Scheme Actuaries preferred working with q_x rather than the m_x or μ_x consistent with central exposed to risk.
- 4.2. In the draft graduations, initial exposed to risk was also used as the start-point. This was adjusted by the deduction of half the number of deaths at each age to give an estimate of the central exposed to risk from which graduated values of μ_x were derived.
- 4.3. The limitations associated with the use of initial exposed to risk are discussed fully in a separate document (entitled 'Comparison of approaches for calculating initial exposure' and available on request from the address in 1.10).
- 4.4. Given the limitations, the Committee decided to calculate central exposed to risk directly for the final graduations. The detail of how this has been done is set out in the remainder of this section.
- 4.5. The Committee also intends to use central exposed to risk for future individual scheme results and 'All Scheme' experience analyses, with expected deaths estimated as given in 4.22.

Investigation period/calendar year

- 4.6. The investigation period for a particular pension scheme is likely to be three years and cover three or four calendar years, depending on the start date of the investigation period. This section describes how the central exposure is calculated for a single calendar year. The results for each calendar year are then added together to give the results for the overall investigation period.

Age definition

- 4.7. To date the SAPS analyses have all been based on an age definition of age last birthday, and this is also used in the current work.

Duration

- 4.8. At this stage the data provided for the SAPS analyses is insufficient for analysis by duration. Currently pension commencement dates are not provided for all scheme members; this field is only compulsory for those members whose pensions commence during the investigation period. Hence exposure is calculated by age only, and not duration at this stage.
- 4.9. The Committee intends to examine whether the schemes that do contain full data for pension commencement date form a sufficiently large and homogeneous group for meaningful analysis. No timetable for this work has yet been set. The methodology outlined below can easily be extended so that exposure can be analysed by duration as well, but for clarity this is not shown at this stage.

Central exposed to risk – lives analysis

- 4.10. For a particular calendar year, the contribution to the exposed to risk from each member is calculated as the number of days that the life is observed to be at the risk of death during the calendar year (subject to any adjustment for late-reported deaths). For each

record, the contribution to the exposure cell for age x equals the number of days the member is at risk and is age x according to the age definition to be used, age x last birthday in this case. The age on the date of death is used to allocate the deaths to the appropriate cell for age x .

4.11. The key areas of judgement relate to deciding

- the dates at which the age changes;
- the dates at which exposure to a particular exposure cell of age x commences and ends; and
- the pension amount applicable to each date (for amounts investigations).

Age changes

4.12. In a particular calendar year, a member is assumed to be age x for the period up to but not including the day of the age anniversary (for the age last birthday age definition, this would be the member's birthday) in the year and age $x+1$ for the period from, and including, the day of age anniversary in the year.

4.13. For the age last birthday age definition, we are effectively assuming that everyone is born at 00.01 hours on the day of birth.

Commencement and cessation of exposure

4.14. The first day of exposure is the latest of:

1. 1st January of the calendar year;
2. Start of investigation period for scheme in the calendar year; and
3. The date of pension commencement in the calendar year (i.e. we are assuming that all pensions start at 00.01 hours on the given entry date).

The last day of exposure is the earliest of:

1. Date of death or exit in the calendar year (i.e. we are assuming that deaths and other exits take place at 23.59 on the day of death or exit);
2. End of investigation period for scheme in the calendar year; and
3. 31st December of the calendar year.

For a member exposed for the full calendar year the contribution to exposure will be 365 days or 366 days in a leap year.

4.15. In any calendar year of exposure, a pensioner may contribute to one or both of the following periods of exposure:

- E_x Exposure at age x
- E_{x+1} Exposure at age $x+1$

Central exposed to risk – amounts analysis

4.16. The contribution to the amounts exposed to risk from each member is calculated similarly to the lives exposed to risk but weighted by pension amounts.

Each member is assumed to have a pension review date in each calendar year of investigation, before which the pension amount is at its pre-review level and after which it attains its post-review level. This means that for the amounts analysis, exposure for a calendar year is split between age pre- and post-birthday, as for the lives analysis, and also split further into pre- and post-review periods. In any calendar year a pensioner may be exposed to risk in up to three of the following periods:

- $E_{x,pre}$ Exposure at age x , pre-review
- $E_{x,post}$ Exposure at age x , post-review
- $E_{x+1,pre}$ Exposure at age $x+1$, pre-review
- $E_{x+1,post}$ Exposure at age $x+1$, post-review

The amounts exposed to risk for a member for a given period will be their lives exposure for that period multiplied by their pension amount applying throughout that period. If the pension amount for a member is £1 per annum throughout a period, their lives and amounts exposures will be identical.

Calculation of exposure (lives and amounts)

4.17. For each member record the lives exposure for a particular age cell is calculated as the number of days that the record contributes to that cell divided by the number of days in the calendar year (i.e. 365 days or 366 days in a leap year). Thus, for example, exposure throughout December 2003 counts as 31/365 of a year whilst January 2004 counts as 31/366 of a year.

4.18. For amounts exposure, the lives exposure for a particular age cell is multiplied by the (known or estimated) pre- or post-review annual rate of pension applicable to that cell.

Deaths – lives analysis

4.19. Where a member dies in a particular calendar year a contribution of 1 death is added to the deaths cell for age x where x is calculated as at the date of death on the appropriate age definition (i.e. the definition used for exposures), again subject to any adjustment for late-reported deaths, which is considered in 3.20.

Deaths – amounts analysis

4.20. Where a member dies in a particular calendar year, the contribution to the amounts death cell is the (known or estimated) annual pension amount applicable on the date of death.

Estimating μ_x

4.21. Using the central exposed to risk and death figures, estimates of crude μ on both a lives and amounts basis can be obtained. The estimates will relate to $\mu_{x+1/2}$ if age last birthday is used:

$$\mu_{x+1/2} \approx \frac{D_x}{E_x} = m_x$$

Where:

- D_x deaths for age x last birthday
- E_x exposure for age x last birthday

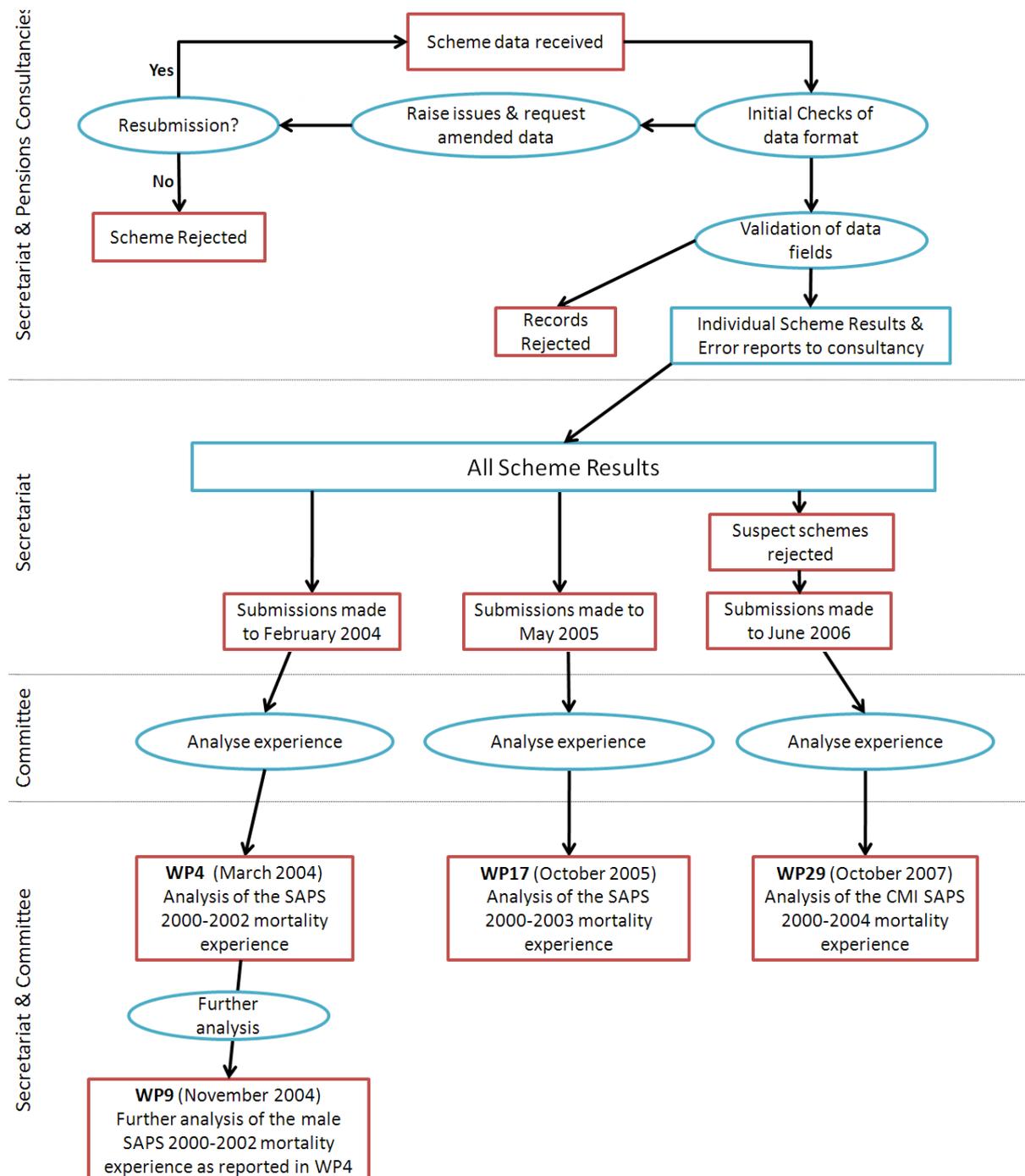
A/E analyses

4.22. The central exposed to risk can be multiplied by values of $\mu_{x+1/2}$ (if an age definition of age last birthday is used) to obtain the expected number of deaths. The actual deaths can then be compared against the expected number of deaths.

Appendix A: Overview of processing stages

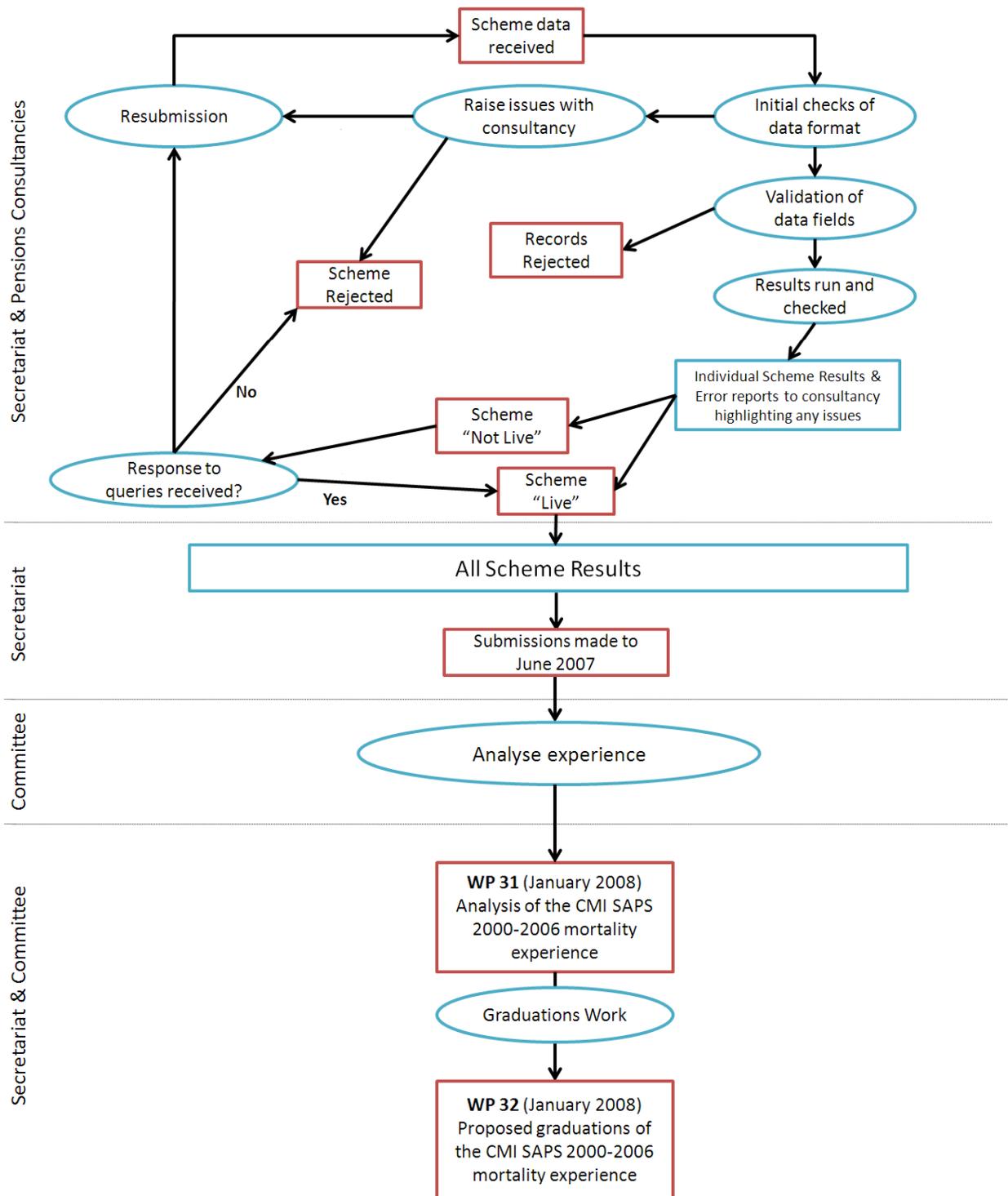
The following diagrams graphically illustrate the different approaches taken to data processing since the SAPS investigation commenced.

Diagram A – Processing stages for data in Working Papers 4, 9, 17 & 29 (see 2.5 - 2.10.)



Suspect schemes were removed from the 'All Scheme' results for WP29 only, as described in 2.10. As described in 2.12 the process has subsequently been amended so that some schemes (marked 'Not Live') do not automatically form part of the dataset, as depicted in Diagram B below:

Diagram B – Processing stages for data in Working Papers 31 & 32 (see 2.5 - 2.8 & 2.11-2.16)



Appendix B: Data checks within SAPS program

B1. The program checks at a scheme level for the following errors and will not upload the scheme at all if any of the following are true:

- The file is empty
- The Header Record does not have the correct number of characters (42)
- The Header Record does not start with an “S” (indicating a scheme record)
- The Office Number is not numeric
- The Industry Classification is not numeric
- The Start, End and Data Extract Dates are invalid dates (for example, 31/2/2006 would be rejected)
- The Record Count is not numeric
- The Header and Footer Records do not match
- The Record Count provided does not match the number of records

If the scheme is not uploaded for any of the above reasons this is queried with the data provider and a correction/resubmission is requested.

B2. An addition to this list of checks was implemented in July 2008:

- The Extract Date is more than 30 days before the End Date

If the Extract is less than or equal to 30 days before the End Date, then this is accepted, but the dates are queried to check that the data has genuinely been extracted prior to the scheme valuation date. The adjustment for late-reported deaths (see 3.20 - 3.23) then deals with this premature extract by bringing the end date forward to 30 days before the Extract Date.

B3. If the scheme is validated for upload, the program checks and uploads each member record in turn. A record is rejected if any of the following hold:

- The record does not start with an “M” (indicating a member record)
- The record does not have the correct number of characters (55)
- The Member Identifier is blank
- The Member Identifier has been used for a previously uploaded record in that submission. (Note that this treats the record as a ‘duplicate’, but these records do not have to match on any other fields.)
- The Type of Exit is something other than blank, “D” or “X” (the codes corresponding to no exit, death and ‘other exit’, respectively)
- The Date became Pensioner is an invalid date
- The Date of Exit is an invalid date
- The Date of Birth is an invalid date
- The Pension Amounts are not numeric
- The Type of Pensioner is not a number between 01 and 05 (the codes corresponding to Normal Health retirement, Ill-health retirement, ‘Combined’ retirement where health status is unknown, Dependant and Unknown status, respectively)
- The Pension Review Date is an invalid date
- The Sex is not a number equal to 1 or 2
- No Type of Exit is supplied when the Date of Exit is provided

- No Date of Exit is supplied when the Type of Exit is provided
- The Date became Pensioner is after the Scheme End Date
- The Date of Birth is before 1/1/1890
- The Date of Exit is before the Date Became Pensioner
- The Date of Exit is before the Scheme Start Date
- The Date of Birth is after the Date of Death
- The Date of Birth is after the Date became Pensioner

B4. An addition to this list of checks was implemented in July 2008:

- Either the Start Pension Amount or End Pension Amount is negative

B5. Member records are flagged by the program as warnings if:

- The End Pension Amount is less than the Start Pension Amount

Such records are accepted without amendment unless a correction is received from the scheme actuary.

Appendix C: Analysis of effect of methodology changes

The table below summarises the effect on the graduations dataset of the changes described in Section 3. For each stage, the percentage change from the previous step is shown in *blue italics*.

	Exposure				Deaths			
	Lives		Amounts (£'000)		Lives		Amounts (£'000)	
	Males	Females	Males	Females	Males	Females	Males	Females
WP 32 Dataset	5,713,143	4,477,881	36,078,421	12,122,793	220,892	161,788	957,670	373,274
Data changes (see Section 2)	5,643,975	4,425,252	35,782,963	12,013,637	217,658	159,443	945,637	369,397
	<i>-1.21%</i>	<i>-1.18%</i>	<i>-0.82%</i>	<i>-0.90%</i>	<i>-1.46%</i>	<i>-1.45%</i>	<i>-1.26%</i>	<i>-1.04%</i>
Minor program changes (see Section 3)	5,620,170	4,401,428	35,540,019	11,906,566	216,509	158,552	941,840	368,066
	<i>-0.42%</i>	<i>-0.54%</i>	<i>-0.68%</i>	<i>-0.89%</i>	<i>-0.53%</i>	<i>-0.56%</i>	<i>-0.40%</i>	<i>-0.36%</i>
Move to central exposed to risk (see Section 4)	5,552,379	4,351,996	35,247,713	11,793,207	216,509	158,552	941,840	368,066
	<i>-1.21%</i>	<i>-1.12%</i>	<i>-0.82%</i>	<i>-0.95%</i>	-	-	-	-
Overall change from WP32 dataset	-160,764	-125,885	-830,709	-329,587	-4,383	-3,236	-15,829	-5,208
	<i>-2.81%</i>	<i>-2.81%</i>	<i>-2.30%</i>	<i>-2.72%</i>	<i>-1.98%</i>	<i>-2.00%</i>	<i>-1.65%</i>	<i>-1.40%</i>

Note that the impact on 100A/Es is negligible due to changes in both exposure and deaths of similar magnitudes and in the same direction.