



Continuous Mortality Investigation

Institute and Faculty of Actuaries

CMI Self-Administered Pension Scheme Mortality Investigation

Coding Guide, Version 3.1 (February 2019)

This Coding Guide provides a description of required data and formatting suggestions for firms submitting data to the CMI SAPS Mortality investigation.

This Coding Guide has been, and may continue to be, subject to further revision from time to time in the light of our experience of data collection and analysis methodology. A list of the previous versions, and the key changes in each, is contained in Appendix 2.

Correspondence address: Two London Wall Place, 123 London Wall, London, EC2Y 5AU

Tel 020 7776 3820 Web www.cmilimited.co.uk Email info@cmilimited.co.uk

Continuous Mortality Investigation Limited ("CMI") is registered in England & Wales (Company number: 8373631) with its Registered Office at: 7th floor, Holborn Gate, 326-330 High Holborn, London, WC1V 7PP

1. GENERAL INFORMATION

1.1 Data submission

All queries regarding data submissions should be sent to:

Tel: 020 7776 3820

Email: self-admin@cmilimited.co.uk

Data submissions should be coordinated within a firm. If data is not submitted via the data coordinator, please ensure that they are informed of the scheme data that has been submitted and any queries that may arise. If you do not know who the data coordinator is for your firm then please contact the CMI.

Details of the format and mode of submission are given in Section 2 below.

1.2 Personal data

Data supplied to the CMI must not include name, address or NI number; however a member identifier or other internal identifier should be included to permit ready identification of a record should a query or the need for correction arise.

The vast majority of records within a CMI dataset could not be related to a particular individual and would not, therefore, constitute “personal data” under data protection legislation. However, a few records may fall within the legal definition and we therefore:

- 1 Seek to minimise this likelihood by using rounded dates of birth and retirement (see 5.11 and 5.12) and capping pension amounts (see 5.15) in the data we use; and
- 2 Treat all such data as if it were personal data from a legal and data security perspective.

1.3 CMI's data processing

Data submitted to the CMI, in any form, is subject to the Terms and Conditions for Data Contributors, which are available at <http://www.actuaries.org.uk/learn-and-develop/continuous-mortality-investigation/cmi-data>. Further details of how the CMI processes data are set out in the Data Handling Protocols, which are available on request but, in brief:

- Data will be received and held by the Secretariat, which is currently outsourced to Barnett Waddingham LLP. Appropriate security measures are taken to retain data securely, with regard to both data protection considerations and to protect the confidentiality of individual contributors' data.
- A firm's data and scheme experience are confidential to that firm and the CMI takes great care that neither data nor details of a firm's scheme experience are distributed to third parties without the permission of the firm concerned other than as part of pooled industry experience. If a subset of the pooled industry data is analysed, e.g. by industry class, then the results of such analyses will only be distributed if the underlying dataset is based on multiple schemes and is not dominated by any single scheme.
- Please also note that the CMI does not seek the name of the scheme, only a unique identifier that is used in correspondence between the CMI and the firm.

2. DATA SUBMISSION

2.1 Physical methods of submission

We can collect data from your secure file exchange site or you can use the Barnett Waddingham site – please email us, using the address provided in Section 1.1, to agree the more convenient approach.

Alternatively, data can be submitted by email, using the address provided in Section 1.1, but this should ONLY be done if you are confident that the data submission does not include any personal data.

2.2 Labelling and accompanying information

The CMI receives a regular stream of data from many sources. In order to ensure proper control of data, firms are requested to send an accompanying email when data is submitted via secure file exchange indicating the scheme identifier, the investigation period, the person(s) to contact in case of queries and any special instructions in order to access data files or return results.

2.3 Submission of data by e-mail

If submitted by email, the data should be attached as a compressed file (e.g. using WINZIP). The covering e-mail should contain the information set out in 2.2.

We encourage all contributors to encrypt and password protect all data files. Passwords should be provided by a separate means (i.e. telephone) from the data.

As with many organisations, filters and firewalls restrict the transmission of emails to the CMI that may contain harmful software. From experience, spreadsheets frequently contain macros, sometimes without the sender's knowledge. For this reason we ask that all files are checked to ensure that they are free of macros before being sent to us.

3. DATA FORMAT

3.1 Types of electronic data file

Text/CSV files

Data submitted as a text file should be in the format as described below in Tables 1 and 2.

Spreadsheets

We are happy to accept data in spreadsheets but we request that these are formatted as for Excel version 5 (or later). Data should contain the same information as the text file, with each field occupying one cell.

3.2 Record types

The data file submitted should consist of:

- One “header” record, which contains scheme data;
- A number of data records, one for each scheme member; and
- An “ender” record (a repeat of the header record) to signify the end of the file.

3.3 Header and Ender Records

The header and ender records must start with the character “S”. The specification of these records is shown below in Table 1.

3.4 Data records

Data records for each member record must start with the character “M”. The specification of these records is shown below in Table 2.

Each member record should occupy one text line or spreadsheet row and should contain the information in Table 2 below.

General guidelines on coding are given in Section 4 while guidelines pertinent to a specific field are given in Section 5.

Table 1 – Scheme record specification

This record is used at the start and end of the file

Field Description	Location (for fixed width text submissions)	Size	Field Type	Paragraph Reference
Record Type	1	1	Must be "S"	
Firm Code	2-5	4	Numeric	5.1
Scheme Code ^a	6-9	4	Alpha-numeric	5.2
Industry Classification: Industry or Supersector	10-11	2	Numeric	5.3
Beginning Date	12-19	8	ddmmyyyy	5.4
End Date	20-27	8	ddmmyyyy	5.5
Data extract date	28-35	8	ddmmyyyy	5.6
Record count	36-42	7	Numeric	5.7

^a The scheme identification code should be unique for each scheme and consistent between years.

Table 2 – Member record specification

Field Description	Location (for fixed width text submissions)	Size	Field Type	Paragraph Reference
Record Type	1	1	Must be "M"	
Member Identifier ^b	2-11	10	Alpha-numeric	5.8
Sex	12	1	Numeric	5.9
Type of Pensioner	13-14	2	Numeric	5.10
Date of Birth	15-22	8	ddmmyyyy	5.11
Date Became Pensioner	23-30	8	ddmmyyyy	5.12
Date of Exit	31-38	8	ddmmyyyy	5.13
Type of Exit	39	1	"D", "X" or blank	5.14
Annual Pension at Beginning	40-45	6	Numeric (££££££)	5.15 & 5.16
Annual Pension at End	46-51	6	Numeric (££££££)	5.15 & 5.17
Date of Pension Reviews	52-55	4	ddmm	5.18
IMD Decile (Region)	56-57	2	Numeric	5.19
Region	58-59	2	Alpha	5.19
IMD Decile (UK)	60-61	2	Numeric	5.19

^b The Member Identifier should be unique for each record.

4. GENERAL CODING GUIDELINES

4.1 Data collection and co-ordination

Normally, data (as specified in Tables 1 and 2 and in Section 5) will be collected via Scheme Actuaries. It will usually be submitted in respect of each scheme as at a valuation date and cover the period since the last valuation or for a period over which the mortality of the scheme has been analysed internally. Each firm employing Scheme Actuaries is asked to coordinate data submissions in respect of that firm.

We are keen to avoid duplication of scheme data submitted to the CMI. Duplication is most likely to occur in the following instances:

1. *You are submitting data for a scheme for which you are not the Scheme Actuary*

Please check that the Scheme Actuary has not already submitted data to the CMI. Please also inform the CMI, when submitting the data, that you are not the Scheme Actuary and that you have received confirmation that data for this scheme has not already been submitted.

2. *You have recently been appointed as the Scheme Actuary*

Guidelines for dealing with a change of Scheme Actuary depend on whether that change involves Scheme Actuaries within the same firm or in different firms. In the first case, it is requested that the firm's data coordinator ensures a smooth transition, and that there be no change in scheme number. If however the change involves two firms, we request that the data coordinator tell us if data is being submitted for the first time, and if not, the previous firm and the scheme reference number used for earlier submissions.

The CMI will allocate firm codes to all contributing firms. The central coordinator in each firm will be responsible for allocating scheme numbers to each of their schemes. These must be unique and should be used, unchanged, for each data submission in respect of each scheme. See Sections 5.1 and 5.2.

4.2 Data to include and exclude

- Data is collected in respect of pensioners who were active during the inter-valuation period.
- Data for deferred pensioners and members who have not retired should be excluded.
- Ill-health pensioners should be included and, if possible, identified as described in 5.10.
- Data in respect of pensioners where a pension scheme has purchased annuities on behalf of the scheme to mirror its liabilities (i.e. buy-ins) should be included. However, data in respect of pensioners where a pension scheme has effected annuities for members (i.e. buy-outs) should be excluded; this data is instead captured from insurance companies for the CMI Annuities investigation.
- Pensions paid in respect of both self-administered DB and DC schemes should be included.
- Top-up schemes should be excluded.

4.3 Deaths and other exits

Each scheme's data submission should include data in respect of all active pensioners during the period, including those who died or retired during that time. It is expected that the most common type of pensioner exit from the investigation will be via a death. However, other exits will occur and it is necessary to identify these so that the exposed to risk can be correctly calculated and these exits can be excluded from the count of deaths. See Sections 5.13 and 5.14.

4.4 Dependant pensions

If a dependant pension (spouse's pension) is paid on the death of a pensioner, a new record should be created in respect of the dependant with the *Date became Pensioner* field being set to the date the new pension commences.

4.5 Minimum size of schemes

Whilst all data submissions are encouraged, in order to keep the data volumes (and hence costs) within reasonable bounds, only schemes with at least 500 current pensioners should be included in the investigation.

5. GUIDELINES FOR CODING DATA FIELDS

Scheme record data specification

5.1 Firm Code

Each firm is allocated a unique firm code by the CMI to be used in communications between us. It will not be visible in any published results. If, when preparing data, the firm code is not known then please contact us and we will be happy to help.

5.2 Scheme Code

The firm should allocate a unique scheme code to each scheme and this should be used for each data submission in respect of that scheme. This will only be used in correspondence between the CMI and the firm.

5.3 Industry Classifications

The industry classifications are broadly in line with the FTSE Actuaries Industry Sectors, which have followed the Industry Classification Benchmark system since 31/12/2005. If the company is quoted, please use the known FTSE classification and code according to the following key:

Key	Industry
51	Oil and Gas
52	Basic Materials
53	Industrials
54	Consumer Goods
55	Health Care
56	Consumer Services
57	Telecommunications
58	Utilities
59	Financials
60	Technology
61	Government/Civil Service
62	Local Authority
63	Miscellaneous

If the company is not quoted and the industry is known then please select the appropriate code from the list above, otherwise use the miscellaneous code. More detailed information about the breakdown of each of these codes is given in Appendix 1 to this document, to assist correct coding.

The CMI is interested in collecting more granular information on industry classification for certain industries where a more detailed breakdown may be of interest. It is therefore requested that data contributors provide information on the industry supersector, instead of the industry classification above, where possible. The keys for relevant supersectors, and how they align with the high level industry classifications, are shown in the table below:

Key	Supersector	Industry
71	Chemicals	Basic Materials
72	Basic Resources	
73	Construction & Materials	Industrials
74	Industrial Goods & Services	
75	Automobiles & Parts	Consumer Goods
76	Food & Beverages	
77	Personal & Household Goods	
78	Retail	Consumer Services
79	Media	
80	Travel & Leisure	
81	Banks	Financials
82	Insurance	
83	Financial Services	

5.4 Beginning Date

Beginning of the investigation period.

5.5 End Date

End of the investigation period.

For example, if data is submitted for a three-year period between valuations carried out as at 31 December 2012 and 31 December 2015 then the *Beginning Date* should be 31 December 2012 and the *End Date* should be 31 December 2015. The exposed to risk will be calculated from 1 January 2013 to 31 December 2015.

5.6 Data Extract Date

The date the data is extracted from the administrator's database. This date will be used to judge if late-reported deaths are missing from the data. This may happen if the extract date is before or soon after the end date.

5.7 Record Count

This field should contain a count of the member records included in the file. The header and ender Scheme Records should not be included in this count.

Member record data specification

5.8 Member Identifier

A unique identifier should be provided up to a maximum of 10 characters. The member's national insurance number and/or name should not be used for this purpose. This information is for use by the CMI only when pursuing data enquiries with the provider.

5.9 Sex

The member's sex should be provided here using the following codes:

- 1 Male
- 2 Female

5.10 Type of Pensioner

The type of pensioner should be recorded using the following codes. Use of codes 01 and 02 is encouraged but if the appropriate split of data between Ill-health and Normal Health pensioners is not available, use code 03:

- 01 Normal Health retirement
- 02 Ill-health retirement
- 03 Combined retirement (where the health of the member at retirement was not known)
- 04 Dependant
- 05 Unknown (if using this code it must be used for all records in the scheme)

The standard combinations of pensioner types included in a submission for a particular scheme are shown in the following diagram:

Normal Health	OR	Combined	OR	Unknown
Ill-health		Dependants		
Dependants				

5.11 Date of Birth

The pensioner's date of birth should be provided here.

This field has the format ddmmyyy, but we encourage firms to always use "16" as the value for "dd" to reduce the likelihood of individuals being identifiable. However we will accept and process data submissions containing:

- Exact values, but will replace "dd" with "16" for the purposes of our analyses.
- Dates with quarterly-or half-yearly rounding; if so, please advise us of the rounding used.

5.12 Date became Pensioner

The date the pensioner's benefits came into payment should be entered here. This field is compulsory for pensioners (including dependants) whose pension commenced after the *Beginning Date*. If the exact date is not known, then an estimated date should be provided. For pensioners (including dependants) whose pension commenced prior to or on the *Beginning Date* this field can be left blank if the date is not available.

Please exclude cases where the *Date became Pensioner* is equal to the *Date of Exit*, i.e. trivial commutations.

As retirement often occurs on the 60th or 65th birthday, use of the exact date could inadvertently reveal the exact Date of Birth. For each submission, Date became Pensioner should be recorded consistently with Date of Birth.

5.13 Date of Exit

Normally this will be the date of death. However, if the pensioner ceases to receive a pension from the scheme for other reasons (e.g. via a buy-out, on re-marriage or cessation of a child's pension) then the date that this happened should be recorded.

We request that you provide the exact date (i.e. no rounding, as per 5.11), if this is known; if the exact date is not known, then an estimated date should be provided. A blank field indicates a "survivor".

5.14 Type of Exit

If the exit is a death then this field should contain the character "D". All other exits should be indicated with the character "X". A blank field indicates a survivor. This field must be consistent with the *Date of Exit* field; i.e. if one field is completed then the other must not be blank.

Where the status of a member record is pending subject to confirmation of a widow(er)s pension, then this field should contain the character "D". If a members' pension becomes suspended for any reason, then this field should contain the character "X".

If it is considered that a significant number of records in respect of exits other than deaths will be omitted from the data file (e.g. if there has been a bulk buyout of annuities and the scheme administrator no longer maintains these records) then the scheme must not be submitted to the investigation. If this is not done then there is a risk that the investigation results will be biased since the exposed to risk in respect of these cases will be omitted but associated deaths occurring while that group of lives were included in the scheme may be recorded.

5.15 Annual Pension amounts at beginning and end

It is essential we are provided with either start or end pension amount and preferable that both are provided.

As with exact date of birth, we recognise that very high amounts potentially increase the chance of certain people becoming identifiable. Consequently, for both the start and end pension amounts, we encourage firms to provide a value of "999999" for any record with an amount over £100,000 and we will then re-weight these values in our analyses. However we will accept and process data submissions containing exact values, but will replace this with the substitute value for the purposes of our analyses.

5.16 Annual Pension at beginning

Data in this field is not essential (if end amount is included – see 5.15) but preferable. Where start pension amounts are missing we would be grateful if Scheme Actuaries estimated the missing pension amounts based on their knowledge of the specific scheme rules and history of pension increases.

For surviving pensioners (and dependants) at the *Beginning Date* (as given in the *Scheme Record*), this should be the annualised pension in payment at that time, including GMP and any pension increase coming into effect at that time.

For new pensioners (including dependants) in the investigation period, this should show the annualised pension on *Date became Pensioner* (or an approximation).

5.17 Annual Pension at end

Again, the data in this field is not essential (if start amount is included – see 5.15) but preferable. Where end pension amounts are missing we would be grateful if Scheme Actuaries estimated the missing pension amounts based on their knowledge of the specific scheme rules and history of pension increases.

For surviving pensioners (and dependants) at the *End Date* (as given in the *Scheme Record*), this should be the annualised pension in payment at the *End Date* including GMP but excluding any pension increase coming into effect at that time.

For deaths (or other exits) during the investigation period, the amount shown here should be the pension in payment at the date of death (or exit). If this is not available, then an approximation should be provided.

5.18 Date of pension reviews

The day and month in each year that this member's pension is reviewed. If there is no fixed review date, or there is more than one review date each year, the actuary should make a best guess if possible, otherwise leave blank.

5.19 IMD and regional measures

These measures are derived from the postcode of each member. The measures should be calculated using a current version of the "CMI postcode mapping tool", which is available to Authorised Users from the [CMI data page](#), where you can also find a document describing the first version of the tool.

The tool maps each postcode to the measures, which are based on the Index of Multiple Deprivation (IMD) decile and the region.

For members with no valid UK postcode, these fields should be left blank.

Please note that provision of IMD and regional measures is preferable but not essential. Although the Committee encourage its inclusion in submissions, you can continue to submit data without providing this data.

Appendix 1: Industry Classification Benchmark from 31/12/2005

Industry	Supersector	Sector	Subsector
Oil & Gas	Oil & Gas	Oil & Gas Producers	Exploration & Production Integrated Oil & Gas
		Oil Equipment & Services	Oil Equipment & Services Pipelines
Basic Materials	Chemicals	Chemicals	Commodity Chemicals Speciality Chemicals
	Basic Resources	Forestry & Paper	Forestry Paper
		Industrial Metals	Aluminium Non-Ferrous Metals
		Mining	Steel Coal Diamonds & Gemstones General Mining Gold Mining Platinum & Precious Metals
Industrials	Construction & Materials	Construction & Materials	Building Materials & Fixtures Heavy Construction
	Industrial Goods & Services	Aerospace & Defence	Aerospace Defence
		General Industrials	Containers & Packaging Diversified Industrials
		Electronic & Electrical Equipment	Electrical Components & Equipment Electronic Equipment
		Industrial Engineering	Commercial Vehicles & Trucks Industrial Machinery
Industrial Transportation	Delivery Services Marine Transportation Railroads Transportation Services		

Industry	Supersector	Sector	Subsector
			Trucking
		Support Services	Business Support Services Business Training & Employment Agencies Financial Administration Industrial Suppliers Waste & Disposal Services
Consumer Goods	Automobiles & Parts	Automobiles & Parts	Automobiles Auto Parts Tyres
	Food & Beverage	Beverages	Brewers Distillers & Vintners Soft Drinks
		Food Producers	Farming & Fishing Food Products
	Personal & Household Goods	Household Goods	Durable Household Products Non-Durable Household Products Furnishings Home Construction
		Leisure Goods	Consumer Electronics Recreational Products Toys
		Personal Goods	Clothing & Accessories Footwear Personal Products
		Tobacco	Tobacco
Health Care	Health Care	Health Care Equipment & Services	Health Care Providers Medical Equipment Medical Supplies
		Pharmaceuticals & Biotechnology	Biotechnology Pharmaceuticals
Consumer Services	Retail	Food & Drug Retailers	Drug Retailers Food Retailers & Wholesalers
			Apparel Retailers

Industry	Supersector	Sector	Subsector
	Media	General Retailers	Broadline Retailers Home Improvement Retailers Specialised Consumer Services Speciality Retailers
		Media	Broadcasting & Entertainment Media Agencies Publishing
	Travel & Leisure	Travel & Leisure	Airlines Gambling Hotels Recreational Services Restaurants & Bars Travel & Tourism
Telecommunications	Telecommunications	Fixed Line Telecommunications	Fixed Line Telecommunications
		Mobile Telecommunications	Mobile Telecommunications
Utilities	Utilities	Electricity	Electricity
		Gas, Water & Multi-Utilities	Gas Distribution Multi-Utilities Water
Financials	Banks	Banks	Banks
	Insurance	Non-Life Insurance	Full Line Insurance Insurance Brokers Property & Casualty Insurance Reinsurance
		Life Insurance	Life Insurance
	Financial Services	Real Estate	Real Estate Holding & Development Real Estate Investment Trusts
General Financial		Asset Managers Consumer Finance Speciality Finance Investment Services Mortgage Finance	

Industry	Supersector	Sector	Subsector
		Equity Investment Instruments	Equity Investment Instruments
		Non-Equity Investment Instruments	Non-Equity Investment Instruments
Technology	Technology	Software & Computer Services	Computer Services Internet Software
		Technology Hardware & Equipment	Computer Hardware Electronic Office Equipment Semiconductors Telecommunications Equipment

Appendix 2: Version history of SAPS Coding Guides

Version	Effective from	Key changes
1.0 ¹	March 2005	First version.
1.1	October 2006	Minor corrections.
2.0	February 2007	Additional data field requested to provide postcode information. Industry classifications updated to reflect changes in FTSE industry classifications.
2.1	March 2009	Minor amendments, including: <ul style="list-style-type: none"> • Additional information to avoid duplication of scheme data; • Provision of additional information to aid industry classification; • Clarification on certain data fields; and • Clarification that trivial commutations should be excluded.
2.2	July 2009	Changes regarding status of data under data protection legislation and security of data transmissions.
2.3	March 2011	Minor amendments to encourage provision of start and end pension amounts.
2.4	November 2016	Minor amendments, including: <ul style="list-style-type: none"> • Remove request to provide postcodes as this information is no longer being collected ²; • Updating industry classification requirements; and • Clarifying certain data fields.
3.0	April 2018	Revisions to take account of the introduction of the General Data Protection Regulation (GDPR) in May 2018; in particular encouraging firms to: <ul style="list-style-type: none"> • Use secure file transfer site for data submission; • Round dates of birth and dates of retirement to month and year; and • Provide capped pension amounts, above £100,000 pa. <p>We also introduced a measure based on the Index of Multiple Deprivation (IMD) for each record.</p>
3.1	February 2019	Minor amendments to describe Index of Multiple Deprivation data specification.

¹ A data specification was provided to SAPS Mortality data providers prior to the Coding Guide (Version 1.0) being created.

² Data protection legislation requires that personal data should not be collected if it is not being used and unfortunately insufficient postcode data had been submitted to perform any credible analyses; therefore it was decided to stop requesting postcode data in regular SAPS data submissions. Postcodes submitted prior to this date have been deleted from the CMI's databases.