



Continuous Mortality Investigation

Institute and Faculty of Actuaries

Data submissions to the CMI Income Protection investigation

Introduction

Historically the CMI specified detailed data requirements for its investigations. More recently it has adopted a flexible approach to data collection, to make it easier for insurers to submit data.

This document outlines the data we are seeking from insurers for the income protection investigation; in particular it incorporates changes to take account of the introduction of the General Data Protection Regulation (GDPR) in May 2018. These changes encourage firms to:

- Use a secure file transfer site for data submission;
- Round dates of birth to month and year; and
- Provide capped benefit amounts, above £75,000 per annum.

We are also introducing a measure based on the Index of Multiple Deprivation (IMD) for each record, to facilitate analysis of claims experience by socio-economic status.

What is the scope of the investigation?

The Income Protection investigation covers individual income protection cover, focussing on policies that are not medically rated. It does not currently cover group income protection cover; this part of the investigation was discontinued following the withdrawal of a number of offices leading to unacceptably high concentrations of data between offices.

What data do we need to supply?

The CMI is adopting a flexible approach to keep the number of compulsory fields to a minimum, thereby making data submission as easy as possible for offices whilst still providing valuable aggregated results.

Details of the data that the CMI needs are contained in the Appendix to this note.

What will the CMI do with our data?

Initial checks will be applied to ensure that we understand the data you have supplied; this will be done as soon as possible and we will then let you know whether we are able to proceed. In some cases we will need to check our understanding of your data but we will seek to keep the number of queries to a minimum by focussing only on issues material to the results.

When the processing of your data is complete, we will send you a “data report”, summarising the assumptions and mappings that we have applied to your data, and office-specific results. These office-specific results will cover each individual year and each period for which data is submitted. You will be asked to review the data report and the results to ensure that we have not misinterpreted your data.

Once data collection is deemed complete, the CMI will produce aggregated (“all offices”) results for each calendar year and for each quadrennium. These results will be available to all CMI subscribers; as will commentary on key features contained in a Working Paper.

The CMI may also make high-level results available more widely, for example via articles in *The Actuary* magazine or presentations open to all members of the IFoA.

Correspondence address: Cheapside House, 138 Cheapside, London, EC2V 6BW

Tel 020 7776 3820 Fax 020 7776 3810 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

Who will have access to our data?

Data submitted to the CMI, in any form, is subject to the [Terms and Conditions for Data Contributors](#), and is processed in accordance with CMI's Data Handling Protocols which address both data protection considerations and protecting the confidentiality of individual contributors' data. These are available on request.

Contact Details

If you have any questions, please contact us on incomeprotection@cmilimited.co.uk or 020 7776 3820.

April 2018

Appendix: CMI Income Protection investigation – Data submission guidelines

Flexibility

We recognise that the resources necessary to convert data from an office's own database(s) into a rigidly specified format may discourage a potential contributor. Clearly though, each office's data will have to be converted into a standard format before being combined with other offices' data; so the CMI Secretariat needs enough explanatory information regarding your data to enable us to format the data appropriately. An accompanying explanation of the data fields and approach used would be greatly appreciated.

We suggest that offices consider the data they can submit and then contact the Secretariat at an early stage in their data preparation to agree whether the data meets the CMI's needs. In particular, it may be helpful if offices submit a small volume of "test" data to allow the Secretariat to provide an opportunity for any issues to be resolved in advance of preparing a full data submission.

Personal data

Data supplied to the CMI must not include name, address or postcode of any policyholder. Consequently, 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 capping benefit amounts (see below) in the data we use; and
- 2 Treat all such data as if it were personal data from a legal and data security perspective.

Principles

The CMI is seeking to analyse experience in two areas:

- Claim inceptions (i.e. the frequency of claims); and
- Claim terminations (i.e. the length of claims).

This is done by comparing claim events (inceptions and terminations) with exposure. In order to do this we require:

- Details of each policyholder under each distinct benefit;
- Details of when each policy went on-risk and off-risk (if census data, off-risk details are required only for claims); and
- Additional details regarding claims; the definition of these is considered further below.

In addition, we seek to analyse the experience:

- by deferred period;
- by CMI occupation class, so require details of insurers' occupation classes and how these are mapped from occupations;
- by year, so the data should allow us to calculate exposure and allocate deaths by calendar year.

We would also like to extend our standard analyses in future to incorporate some or all of the following factors:

- smoker status;
- socio-economic status, using IMD measures (see below);
- amounts weighting, which requires a measure of current benefit amount;
- sum insured band, requiring a measure of initial benefit amount;
- distribution channel; and
- policy duration, so require benefits that are subject to new underwriting to be separated from the original policy (but non-underwritten increments should not generate a separate record).

The principal analyses are intended to cover:

- Individual business only (not group).
- Income protection policies only.
- UK business, but small volumes of non-domestic business may be included.
- Policies issued at normal premium rates; however the Committee may in future consider producing high-level results for medically-rated lives – it would therefore prefer ALL data to be submitted, with either an indication of whether the policy is subject to a medical rating, or the medical rating itself.

- Policies subject to exclusions can be included in data submissions, or excluded; however offices are requested to apply this consistently – in particular between in force records and claims. Please advise which approach you use.
- Direct business, gross of reinsurance.

Structure of data files

The CMI is happy to accept data files in a number of structures, including:

- “Movements” data, indicating when policies go on-risk and off-risk within each calendar year.
- “Census” data, comprising in force data at the end of each 31 December (including the prior year) and details of claims.
- “Policy history” files, including the complete history of a policy in a single record.

Other approaches may also be acceptable – but please discuss them with us first!

Essential data items

The data items that we regard as essential are:

- Record identifier, to enable ready identification of a record should a query arise.
- Gender.
- Date of birth (see below).
- Record type (i.e. In force, exit or claim, dependent on the structure of data that is supplied).
- Policy number (or other unique identifier).
- Deferred period.
- Occupation class.
- Policy commencement date.
- Whether a policy has a medical rating (unless you only supply data on “standard rates” cases).
- Date and type of exit, if applicable (for census data, only claims will be supplied).

For claim records, the following additional fields are required:

- Date of falling sick.
- Payment commencement date.
- Payment cessation date.
- Mode of cessation, e.g. through recovery, death, lump sum payment or expiry.

Please talk to us if any of these data items cannot be supplied; other fields may provide the information that we need.

Desirable data items

The following additional data fields are particularly desirable:

- Smoker status.
- Territory of the life insured at outset (unless the data can be regarded as UK).
- Initial benefit amount, ideally expressed as an annual amount (to allow analysis by amount band).
- Current benefit amount, ideally expressed as an annual amount (to allow analysis weighted by amounts).
- Whether the policy has a benefit guarantee and, if so, how much it is.
- Type of benefit (level/increasing/decreasing/lump sum/other).
- Level of rating, where a policy/life is rated (if these are included).
- Distribution channel – distinguishing between Bancassurance, IFA/Multi-tie, Single-tie, Direct sales and Unknown (unless the data can be regarded as being from a single channel, in which case please advise us of this alongside the data extract).
- Whether the policy is sold in connection with a mortgage.
- Whether the policy has a maximum benefit period and, if so, the period in years.
- Policy expiry date (or year)
- Sickness definition, e.g. own occupation, any occupation, ADLs.
- Cause of disability (for claims): this should be an indicator of the condition triggering the claim. This should be based on ICD codes, which we group into around 70 cause groups, a mapping for which we can provide.
- Benefit rate percentage (for claims).
- Socio-economic indicators – we are keen to receive a measure based on the Index of Multiple Deprivation (IMD) for each record. We will issue an IMD mapping, from full postcode, and supporting documentation soon.

Dates of birth

To reduce the likelihood of individuals being identifiable, we encourage firms to always use “16” as the value for “dd”. 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.

Note: We do not think that the other dates would be well-known, so this should not make a person more identifiable. We therefore request the exact date for these.

Benefit amounts

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 initial and the current amount, we encourage firms to provide a value of “999999” for any record with an annual benefit over £75,000 and we will then re-weight these values in future 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.

Investigation years

Our current focus is to complete the collection of data relating to experience up to the end of 2016, and are urgently seeking outstanding data for the years 2011 to 2016.

Our intention is to then return to regular annual reporting of experience for individual income protection, as well as incorporating additional data fields in our standard analyses, where data allows.

Methods of submitting data

We are happy to accept data from your secure file transfer site or you can use the Barnett Waddingham site – please email us on incomeprotection@cmilimited.co.uk to agree the more convenient approach – in the following file formats:

- Text/CSV – commas should separate each field or the file should be arranged in fixed width format.
- Spreadsheet.
- Database – we request that database formats are Microsoft Access. For offices using different database software, we suggest that data be converted into Text/CSV.

Alternatively, data can be submitted by email but this should ONLY be done if you are confident that the data submission does not include any personal data. Where email is used, the attaching file should, where possible, be compressed. We encourage data contributors to encrypt and password-protect all data files. Passwords should be provided separately (by telephone) from the data.

The Secretariat may be able to accept other commonly used form of electronic data transmission but please contact us beforehand to check