Institute
and Faculty
of Actuaries

Graduating mortality base tables Theory and practice (part 1)

Matt Fletcher
Piero Cocevar
Chair: Chris Reynolds

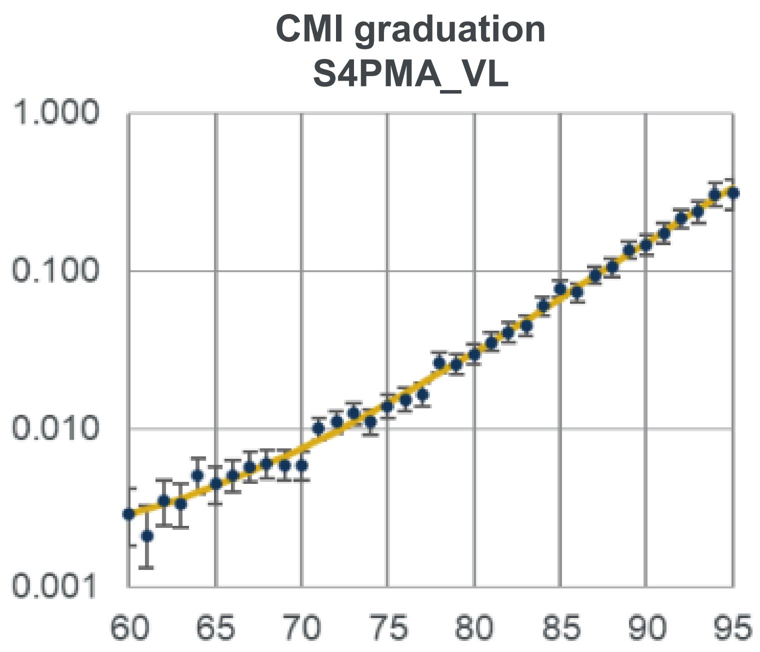
Expressions of individual views by members of the Institute and Faculty of Actuaries and its staff are encouraged.

The views expressed in this presentation are those of the presenters.

10 September 2025

Introduction

- This webinar in two parts will set out the theory and practice of producing mortality tables
- This first session is more about the theory, the second next week focusses on practical aspects
- The sessions will help you understand
    - data requirements
    - pros and cons of graduation as well as alternative approaches
    - principles underpinning typical graduation techniques
    - how to produce robust and reliable mortality tables
    - best practice in communication with data providers and table users



Source: Working Paper 181

Institute and Faculty of Actuaries

(Basic) History of graduation

1662 John Graunt's Bills of Mortality

1693 Edmund Halley's Life Tables

Curves drawn "by eye" through jagged crude rates

1924 CMI Bureau established to regularly collect and graduate insurer, then pension scheme data

CMI adopted Gompertz-Makeham models

2015 CMI Graduation & Modelling Working Party produced a working paper and software for subscribers

Institute and Faculty of Actuaries

Mortality analysis – some key terms

Exposure period (R_x)

- For an individual, the amount of time they were alive and in the dataset at age x
- For a dataset, the sum over the population of individual exposures

Force of mortality (μ_x)

- The probability that an individual aged x dies in the next instant

Actual deaths (A_x)

- The sum of deaths at age x over the dataset

Crude mortality rate

- Actual deaths divided by exposure period

Expected deaths (E_x)

- The integral of mortality μ_x across the exposure period

Institute and Faculty of Actuaries

Source: Working Paper 185

What is graduation?

- Mortality data is noisy, even for a large population
- Underlying force of mortality applying to the population will be smoother than observed data
  - over time
  - age by age
- **Graduation** is the process of producing a smooth set of mortality rates from observed data

Chart 5E: Crude and (initial) fitted rates by age for female non-smokers

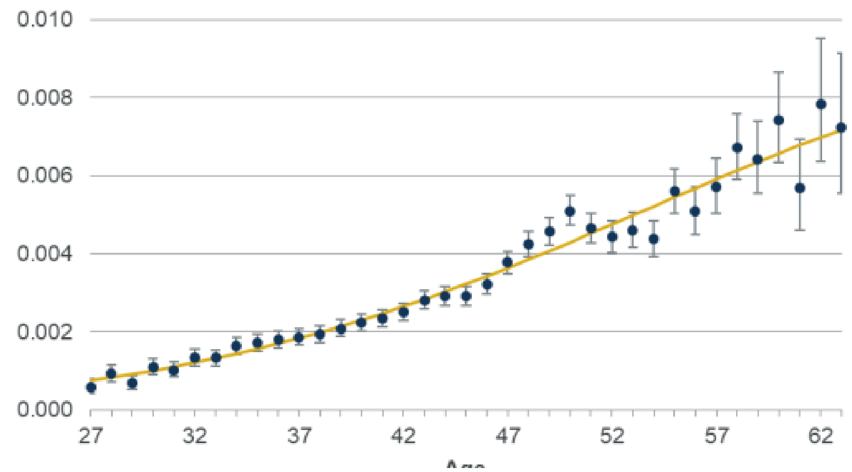
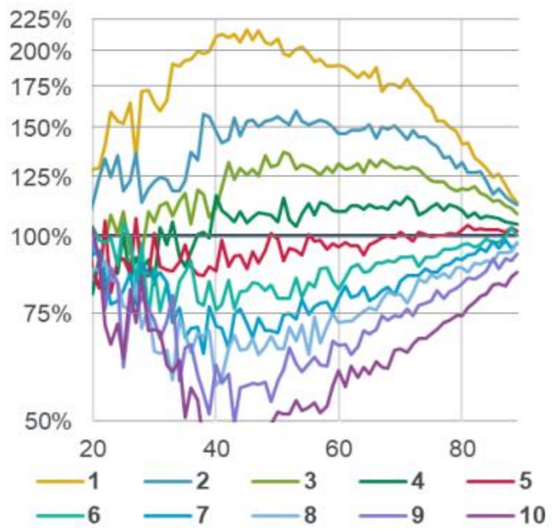


Chart 2B Ratio of crude mortality rates for England & Wales by IMD decile relative to "all decile" rates, logarithmic scale



Institute and Faculty of Actuaries

Source: Working Paper 185

Why should we graduate mortality?

- Graduating mortality rates for
  - **whole populations** (e.g. England & Wales males) gives us a set of **baseline mortality rates** by age for that population
  - **sub-populations** (e.g. female pensioners with pensions over £16,000 p.a.) allows differences in **shape and level** of mortality in those populations to be reflected
- Graduated tables
  - give a more **stable and reliable** set of rates
  - give actuaries a ‘**common currency**’
  - provide a basis against which to **benchmark experience**

Period life expectancies at 65

Data	Subpopulation	Male	Female
UK population	All	18.5	21.0
Pensioners	All	20.6	22.7
Pensioners	High pension	21.8	23.0
Pensioners	Low pension	18.6	21.9
Annuitants	All	20.8	22.9
Annuitants	Individual	21.1	23.3

Institute and Faculty of Actuaries

Source: ONS, various Working Papers

Why shouldn’t we graduate mortality?

Insufficient data
(volume or quality)

Suitable tables already exist

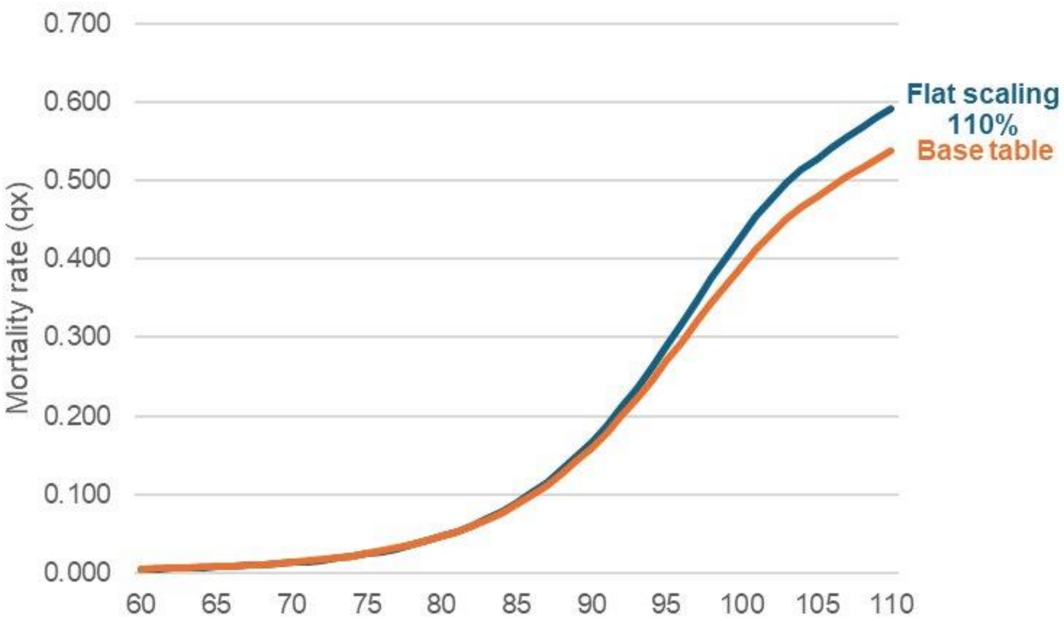
Avoid tables being treated as recommendations

Avoid complexity from having too many tables

Institute and Faculty of Actuaries

Alternatives to graduation

- If the population of interest has a different mortality profile to existing tables, we should reflect this in the mortality assumption used
- Options for adjusting existing tables include
  - **Flat scaling** e.g. 110% of SAPS S4 Males Light
    - Simple to describe and understand
    - Changes assumed mortality at high ages – is this desirable?
  - **Structured scaling** e.g. 90% of SAPS S4 Males Heavy up to age 60, tapering linearly to 100% at age 90
    - More complex – harder to communicate
    - May give a more appropriate shape



Institute and Faculty of Actuaries

Graduation objectives

The graduation process is not just a statistical exercise – judgement comes into play and the following should be assessed / considered:

- Capture underlying trend
- Smoothness vs goodness of fit
- Preserve key features
- Practicality for users

Institute and Faculty of Actuaries

Graduation Process

Data preparation / analysis

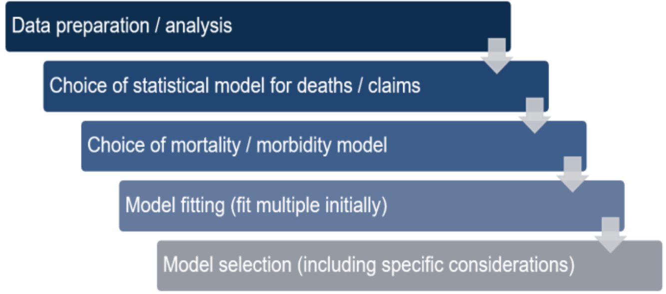
Choice of statistical model for deaths / claims

Choice of mortality / morbidity model

Model fitting (fit multiple initially)

Model selection (including specific considerations)

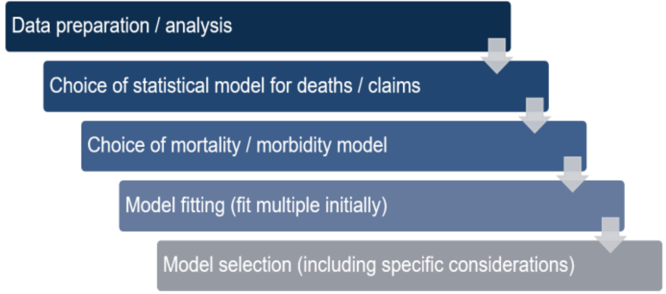
Institute and Faculty of Actuaries



Data preparation and analysis

- Objective is to calculate crude mortality / morbidity: $\frac{Deaths\ or\ Claims}{Exposure} = \frac{A_x}{R_x}$
- Data is collected and processed for individual lives:
  - Deaths / claims
  - Details of when each life went on-risk and off-risk, during period of interest, so the exposure can be calculated
- Data could come from:
  - A single data source, e.g. population data or data within an organisation
  - Multiple data sources, e.g. industry data, as for the CMI
- Considerable data validation is necessary – validity and reasonability of the data
- Data for the individual lives, i.e. deaths / claims and exposure, (and from the multiple sources) is then aggregated by **age (x)**

Institute and Faculty of Actuaries



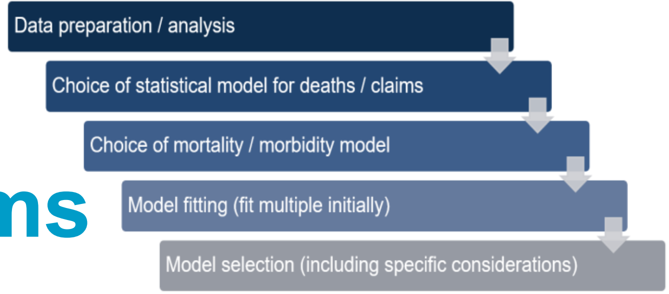
Data preparation and analysis

- **Key objective:** Create homogeneous subsets, i.e. where you would expect the mortality / morbidity experience to be similar (regular experience analyses help inform this decision)
- The data by age can be sub-divided into data subsets, depending on availability of data fields, for example:

  - Sex
  - Smoker status
  - Socio-economic measure (e.g. IMD)
  - Policy or Pensioner type
  - Benefit or pension amount

- Considerations:
  - **Volumes of data:** low data volumes may apply limits to the level of granularity
  - **Age range:** The age range to include will likely be driven by the volume and consistency of data
  - **Time period:** The time period of the data should balance consistency of data and the need to have sufficient volumes of data
  - **Needs of users:** What subsets are of interest?

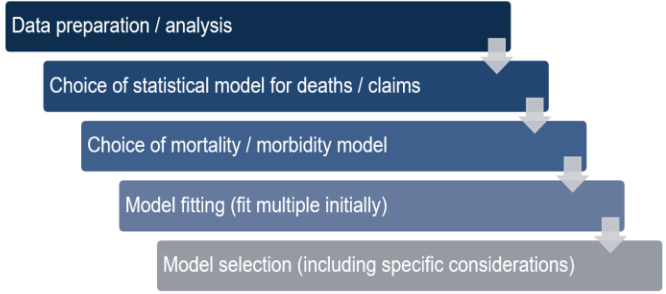
Institute and Faculty of Actuaries



Choice of statistical model for deaths / claims

- We need a statistical model for the deaths / claims (A_x) in our experience data
- The common assumption is that A_x is Poisson distributed with mean and variance both equal to R_x μ_x , where R_x is the (central) exposure and μ_x is the force of mortality (or morbidity)
- The Poisson model:
  - Arises naturally from a survival model for independent lives
  - Allows Maximum Likelihood Estimation (MLE), a relatively simple approach, to be applied
- MLE is a principled, reproducible, approach to fit parametric laws
- Alternative models could be considered where the Poisson model does not fit the data well

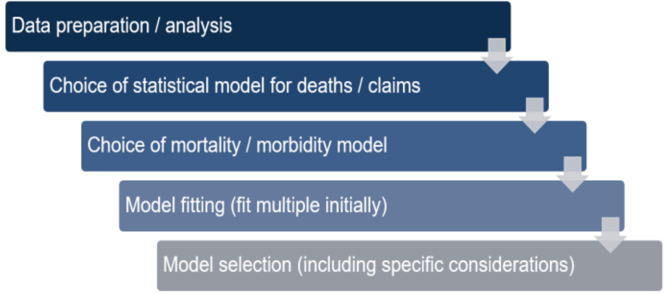
Institute and Faculty of Actuaries



Choice of mortality / morbidity model

- Early graduation approaches included “graphical graduation” – drawing a plausible curve, reading off values and hand smoothing
- Now we tend to fit a mathematical formula to the crude mortality / morbidity rates
- Choosing the right parametric family of models is a core decision in graduation. This is where theory meets actuarial judgment!
- The parametric formula needs to fit the data adequately – many formulae may do this, so there is a need consider the graduation objectives and to achieve the appropriate balance
- The choice of parametric formula is a balance between biological realism, statistical fit and actuarial usability. No single law is universally "best": the right choice depends on the age range, data quality and the ultimate purpose of the graduated table.

Institute and Faculty of Actuaries



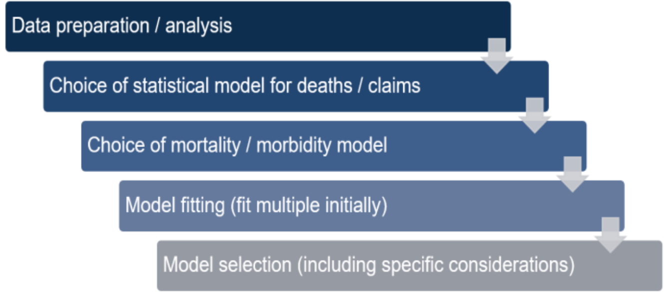
Choice of mortality / morbidity model

- **Gompertz** observed that adult mortality increases roughly exponentially with age or equivalently, that the logarithm of death rates rises about linearly with age: $\mu_x = exp(b_0+b_1x)$ *or log* $\mu_x = b_0+b_1x$
- **Makeham** observed that a better fit, continuing down to younger ages, was achieved by adding a non-exponential element: $\mu_x = a_0+exp(b_0+b_1x)$
- A generalisation, that is widely used in CMI graduations, is the **Gompertz-Makeham** GM(r,s) series of formulae, allowing higher order polynomials in age:

$$\mu_x = (a_0+ a_1x+\ldots+ a_{r-1}x^{r-1}) + exp(b_0+ b_1x+\ldots+ b_{s-1}x^{s-1})$$

- More complex laws with many parameters (e.g. Heligman-Pollard with 8 parameters) can fit well almost any pattern, but often at the cost of interpretability and simplicity.

Institute and Faculty of Actuaries



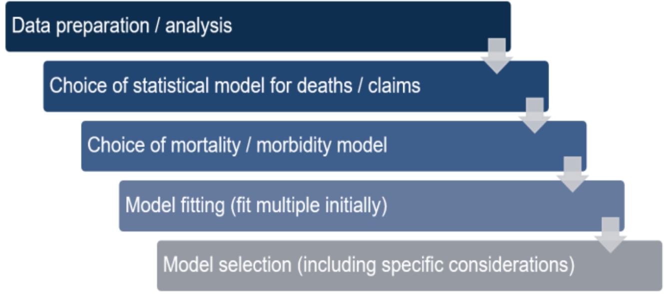
Choice of mortality / morbidity model

- In practice, CMI Committees have tended to use simple Gompertz G(s) or GM(r,s) models, with the exception of the income protection graduations, where GLMs were used:

Committee	Type	Tables	Working Paper	Graduation formula
SAPS	Mortality	S4	WP181	G(4) and G(5)
Annuities	Mortality	16 Series	WP130	G(3) and G(4)
Assurances	Mortality	16 Series	WP150	G(5), GM(1,2), GM (3,2)
Assurances	Critical Illness	16 Series	WP150	G(3) and G(4)
Income Protection	Income Protection	IP11	WP131	GLM

- The choice is generally driven by balancing:
  - Simplicity: ease of estimating, interpreting and communicating
  - Flexibility: better fit to the data
  - Consistency: retain similar formulae for similar dataset

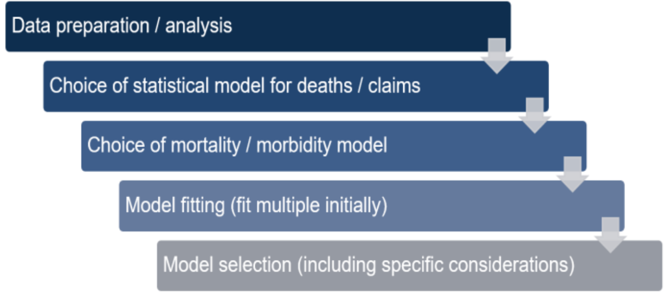
Institute and Faculty of Actuaries



Model fitting

- Recap of where we have got to:
  - Have a dataset, or multiple datasets, of deaths / claims and central exposures
  - An assumption that deaths / claims follow the Poisson distribution
  - Chosen parametric formula, say GM(r,s) model
- The next step is to calculate the fitted values, i.e. the $a_0 \dots a_{r-1}$ and $b_0 \dots b_{s-1}$ parameters, using the Maximum Likelihood Estimation (MLE) process
- The MLE process finds the set of model parameters that make the observed death count most probable under the assumed Poisson distribution, weighting each age in proportion to its exposure.
- The CMI committees use the CMI Graduation Software developed by the Graduation and Modelling Working party for this purpose.

Institute and Faculty of Actuaries

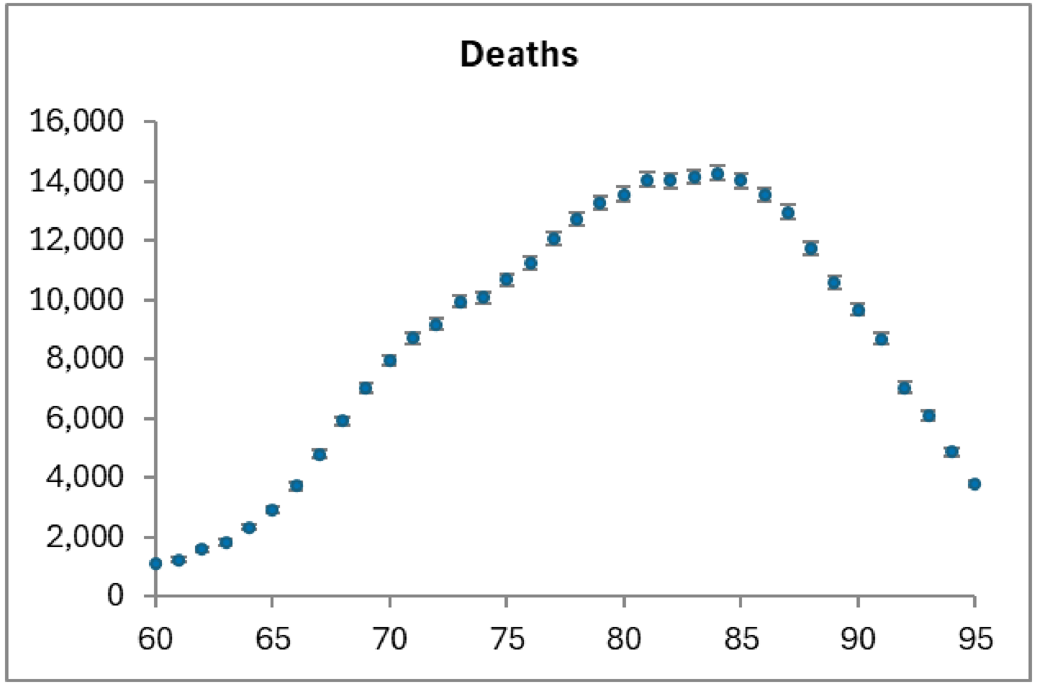


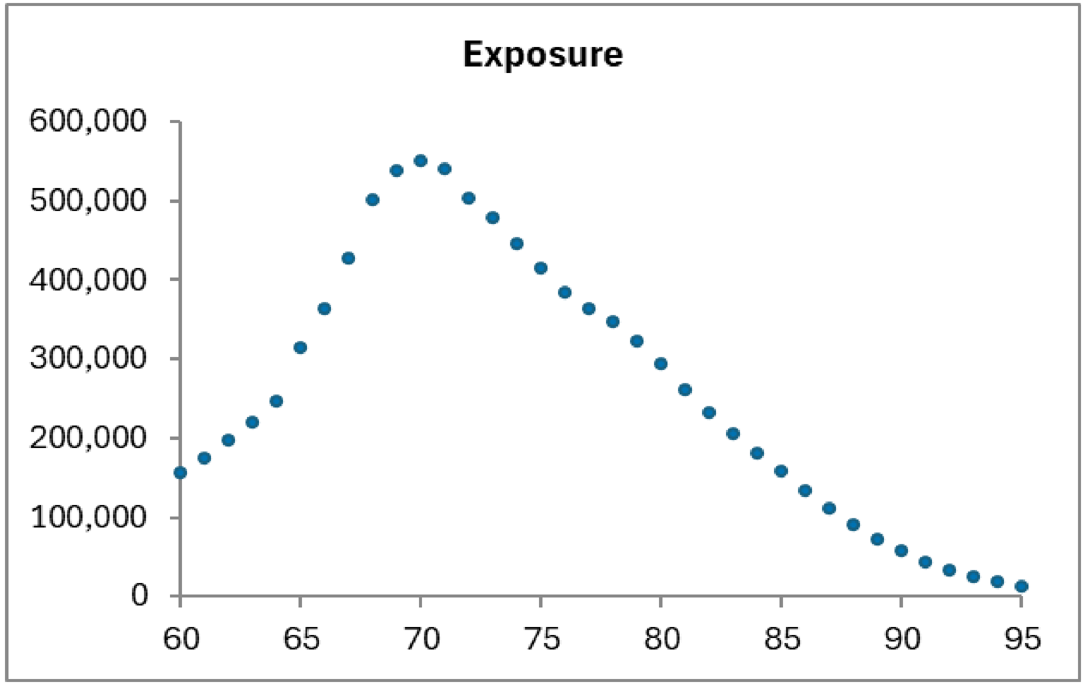
Model selection

- The model selection stage is an iterative process:
  - Start with a simple Gompertz G(s), e.g. setting s=2 or 3
  - Test whether adding extra parameters (GM(r,s) with higher s and r) significantly improves fit
  - Use statistical tests and actuarial judgment to decide when to stop
- The "sweet spot" is often found when there are no more parameters we can remove without compromising the fit, i.e. "as simple as possible, but no simpler".
- Within the CMI the GoLD (Gompertz log-difference) chart is often used in this step, by plotting the logarithmic difference between a simple Gompertz model and those implied by a more flexible GM(r,s) model against age.

Institute and Faculty of Actuaries

Model selection – Example

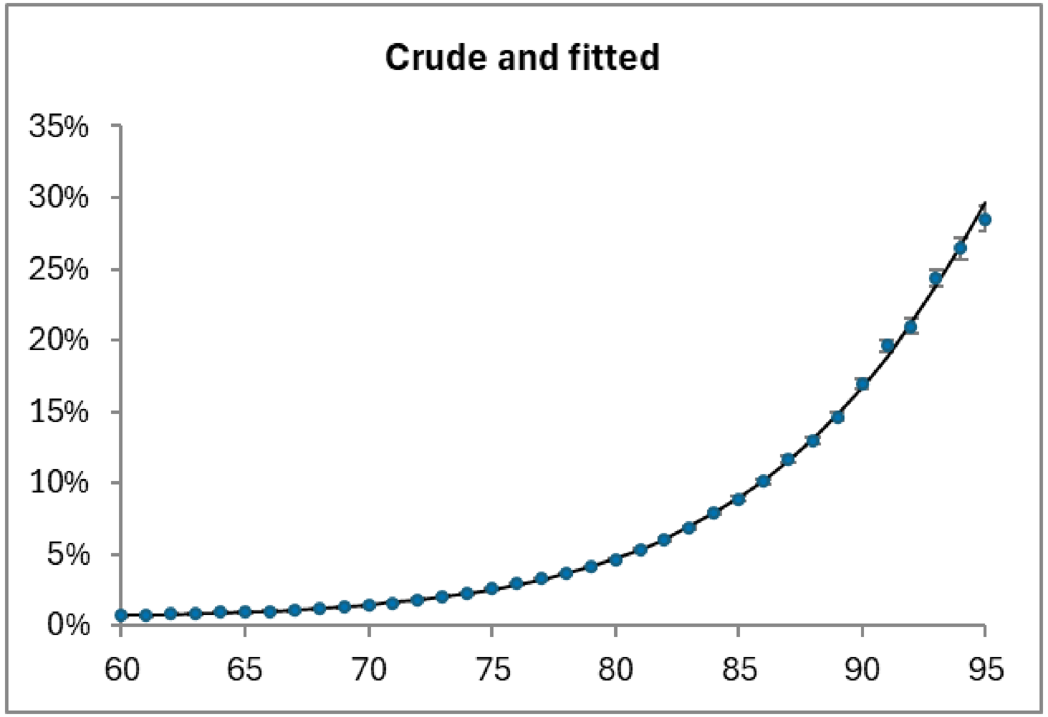




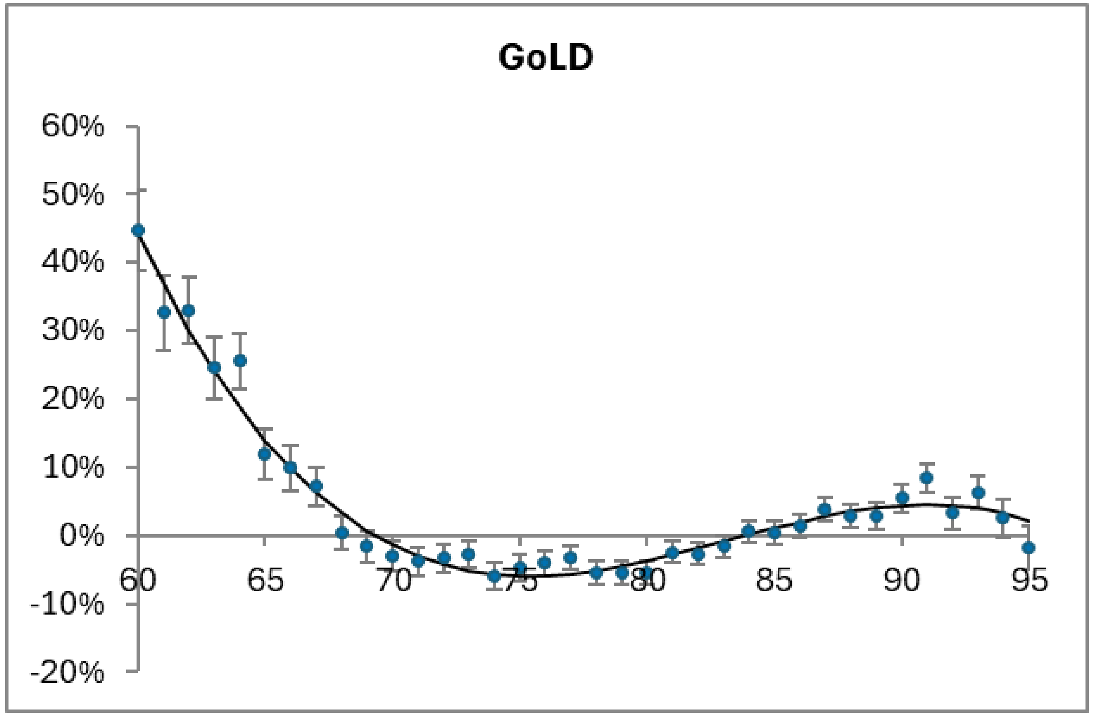
Pensioners male – Lives, Ages 60-95, Years 2015-2019: Death and Exposures

Institute and Faculty of Actuaries

Model selection – Example



G(4) fitted to Pensioners male – Lives Ages 60-95, Years: 2015-2019



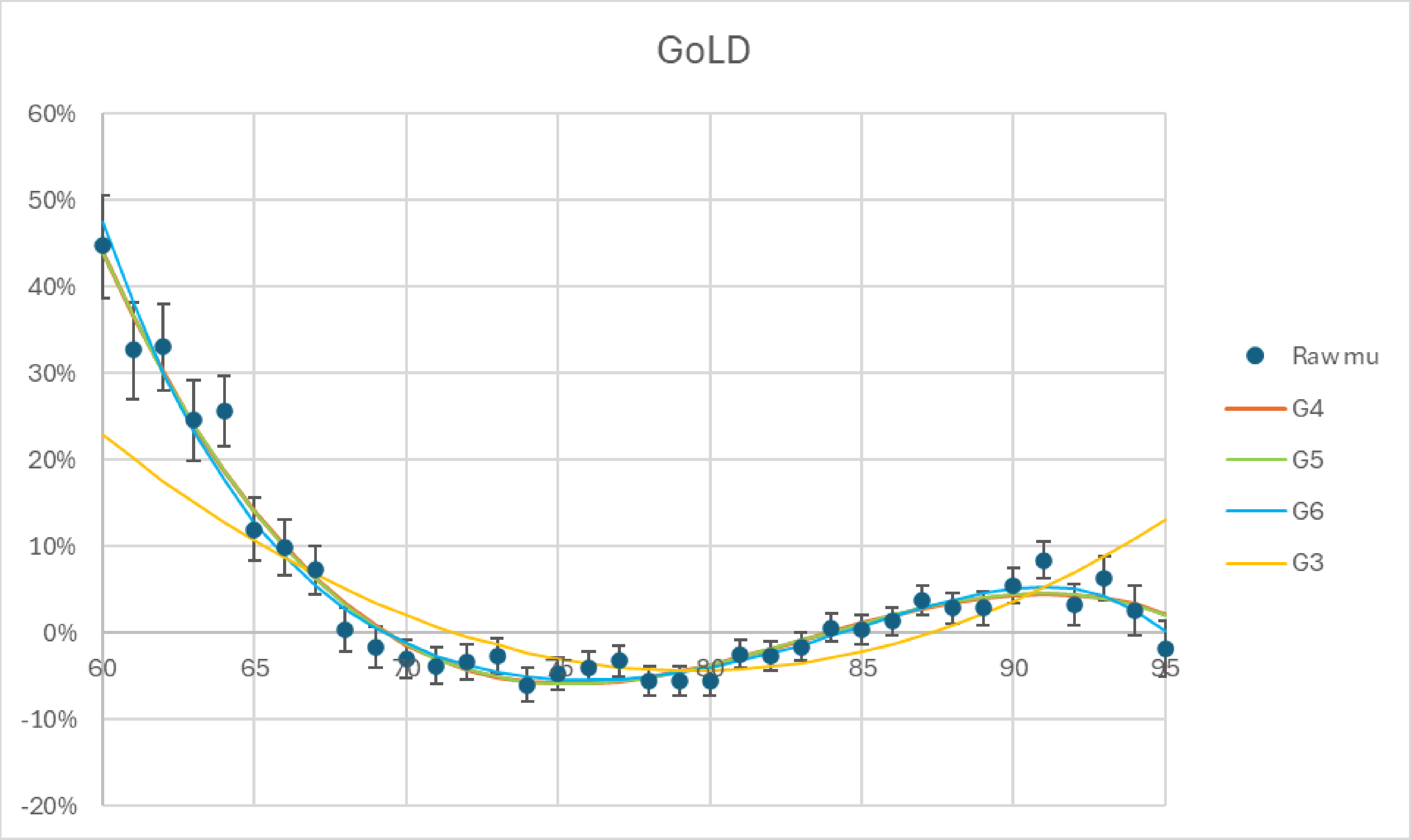
GoLD (Gompertz log-differences) chart:

- The solid line is the log difference between the Gompertz model G(2) and the more complex G(4).
- The dots represent the log difference between the Gompertz model G(2) fit and the crude mortality rates.

Same data as on the left-hand side chart.

Institute and Faculty of Actuaries

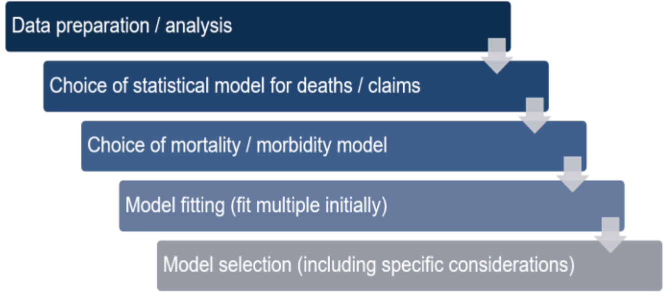
Model Selection – Example



The GoLD chart helps us see if added parameters capture real features or just introduce noise or false patterns.

G(3), G(4), G(5) and G(6) models fitted to Pensioners male – Lives, Ages 60-95, Years: 2015-2019

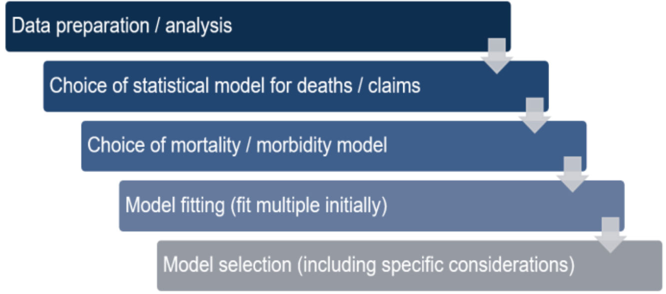
Institute and Faculty of Actuaries



Model Selection – considerations

- Statistical tests are used to compare potential graduations:
  - Information criteria can "score" the graduations – awarding points for good fit but losing points for complexity, e.g. AIC (Akaike) and BIC (Bayesian): lower AIC/BIC should be generally preferred
  - Deviance residual-based tests:
    - Signs test: count how often crude rates are above vs below graduated rates (should be 50/50)
    - Run test: randomness of residuals – too many runs = too jumpy, too few runs = systematic lack of fit
    - Standardised residuals: should behave like N(0,1). Large outliers suggest local misfit.
- Judgment considerations:
  - Smoothness: mortality should increase monotonically with age in adulthood
  - Tail behaviour: extrapolation at high ages must be sensible
  - Comparability: consistent with previously graduated tables, where relevant
  - Usability: overall requirement that the rates are stable over time and credible for the intended use.

Institute and Faculty of Actuaries

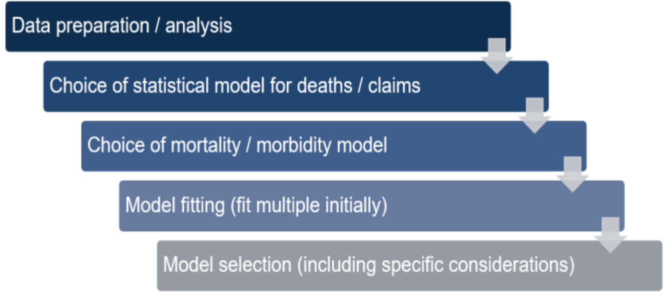


Consistency across sub-groups

- If we are graduating multiple datasets, we need to carry out consistency checks to ensure that multiple graduations are not only statistically sound in isolation, but also coherent, stable, and plausible relative to each other.
- Items that are generally considered are:
  - Relative levels: ensure that subgroup rates maintain plausible ordering (e.g. high-amount lives consistently lighter mortality than low-amount) cross checking external references where possible
  - Smoothness of ratios: graduated rates between groups should change gradually not oscillate
  - Crossovers: avoid situations where one subgroup overtakes another at an implausible age, with particular attention at financially material ages
  - Convergence: check that subgroup of populations converge reasonably to each other and to the aggregate dataset.

These aspects will be covered further in Part 2.

Institute and Faculty of Actuaries

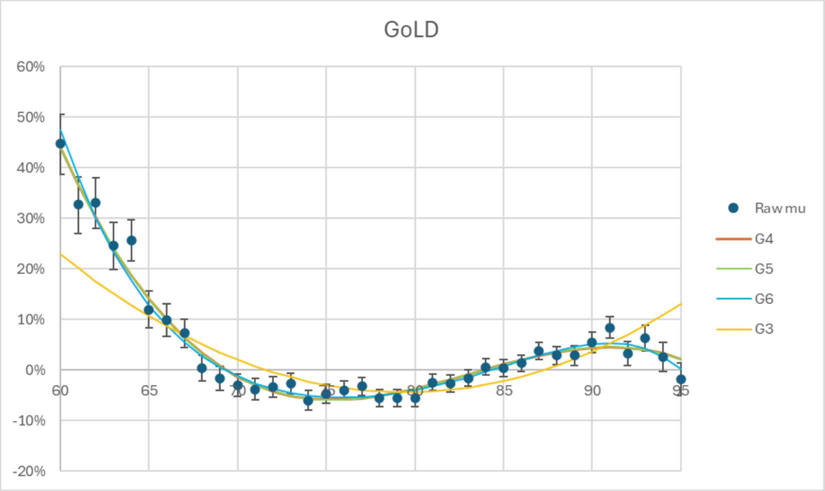


Extensions to young and old ages

- At young and old ages, data can be sparse and unreliable. Typically, the model fitting and selection process is restricted to ages where the data are deemed to be reliable (e.g. 60-90 for a pensioner data set). The fitted mortality curve is then extended to younger and older ages based on expert judgment.
- Extension should balance statistical fit with biological plausibility, often requiring external references (e.g. HAMWP) or parametric laws designed for the tails. Smooth, credible transitions are essential for practical use.
- Extension at young ages is less material and generally involves smoothing to a reference rate from external tables.
- Extension at oldest ages can have a big financial impact on annuities. Approaches used involves often converging smoothly using generally accepted laws of convergence at old ages.
- These aspects will also be covered further in Part 2.

Institute and Faculty of Actuaries

Conclusion



- We've provided an overview of
  - data requirements
  - principles
  - process
- Principles are simple and remain fairly constant – process has developed over many years
- The process set out in this session gets us to a set of graduations for a dataset
  - Multiple decision points along the way – means there is art to the process as well as science
  - Second session will expand on this aspect as well as practical considerations

Institute and Faculty of Actuaries

Questions

Comments

Expressions of individual views by members of the Institute and Faculty of Actuaries and its staff are encouraged.

The views expressed in this presentation are those of the presenters.

Institute and Faculty of Actuaries

Institute
and Faculty
of Actuaries

SCAN ME

Feedback