Clubbing together

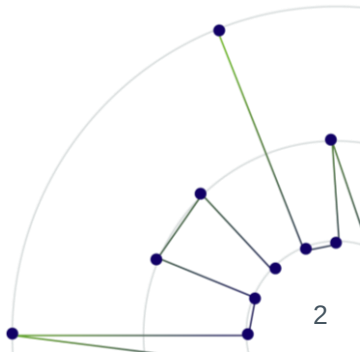
How Club Vita helps pension schemes estimate their members' life expectancy

Conor O'Reilly FFA

4 January 2019

Agenda

- Mortality 101
- Setting assumptions
- Introducing Club Vita
- Constructing VitaCurves



Mortality 101

Mortality 101

- Mortality/Longevity are often used interchangeably… but they are subtly different

  Mortality How likely you are to *die* at a certain age.

  Longevity How long you are expected to *live*

- q_x: the probability of life aged x dying before age $x + 1$
- A mortality curve is a table of q_x by age
  - Mortality rates increase by age (in general)
- Pension schemes need to estimate **when** and for **how long** they will pay pension benefits
  - What are 'current' mortality rates?
  - How will they change in the future?

What affects mortality/longevity?

Age

Gender

Smoker?

Health

Wealth

Fitness

Genetics

Marital status

Blue or white collar?

Where you live

CLUB VITA

Setting assumptions

How to set longevity assumptions

Longevity assumption

=

Baseline assumption

- How long people are **currently living** for.
- Can be measured **objectively** by looking at historical death rates.

+

Future improvements

- How **longevity will change** in the future.
- Typically would expect mortality rates to decrease in the future and life expectancy to go up.
- Informed by views on future medical advances and generational differences in lifestyle etc.
- Recent longevity trends will influence the assumptions you set, but it is important to understand the reasons behind recent experience before relying on it to adjust assumptions.
- This is a **subjective** assumption, and uncertainty will remain.

Introducing Club Vita

Club Vita's dataset

Records for over 2.8 million pensioners

equivalent to 1 in 4 DB pensioners

Relationships with over 220 large pension schemes

with over £300 billion of liabilities

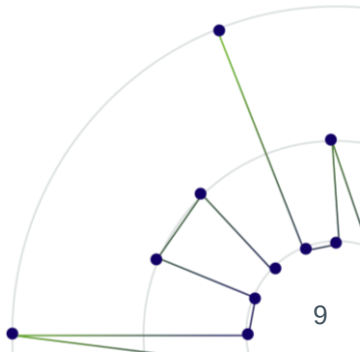
Over 1.4m death records

stretching back 25+ years

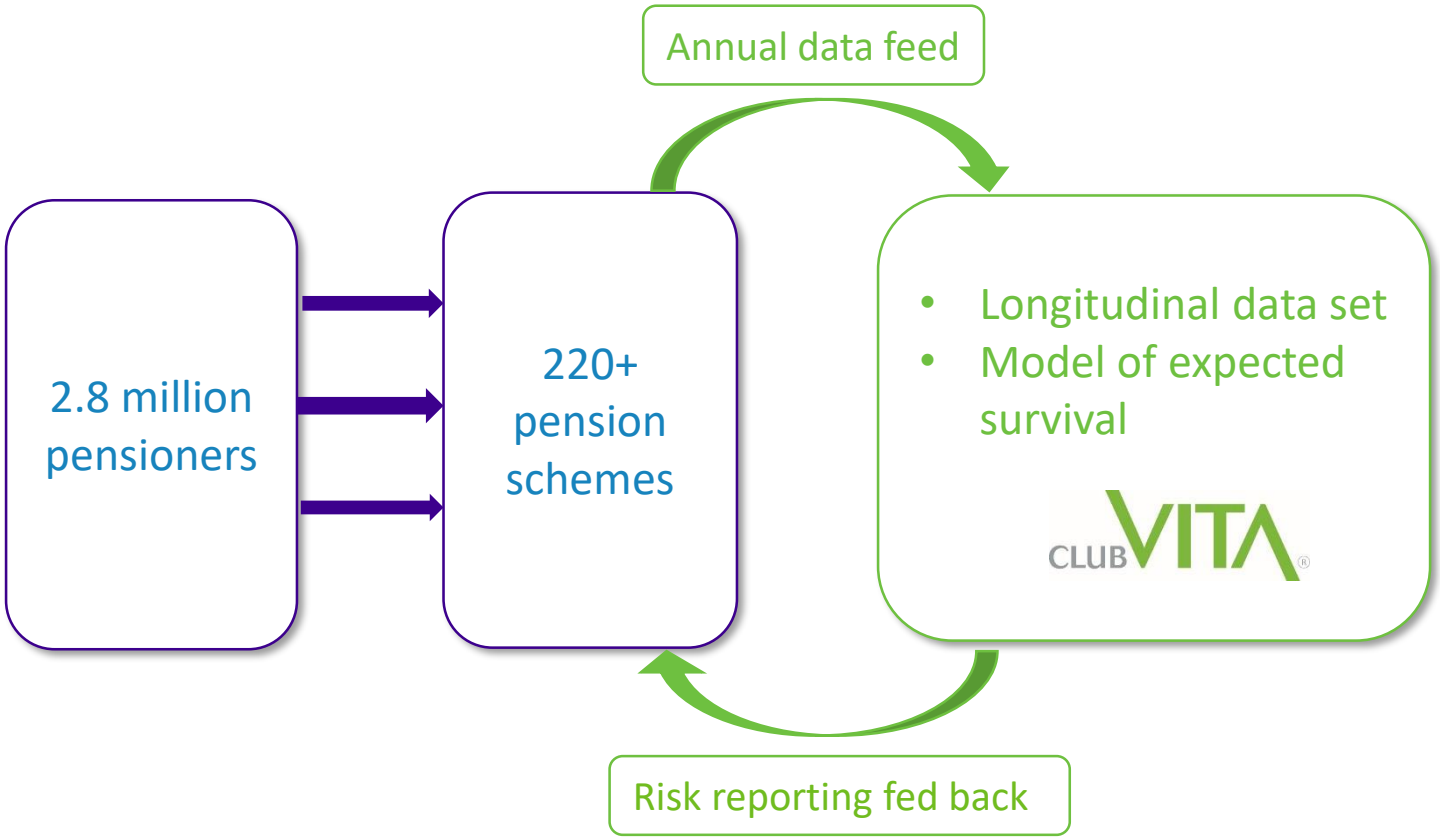
Segmented by affluence, postcode, health and more

Data collected annually

Richest, most flexible dataset of its kind in the market



Supporting pension schemes



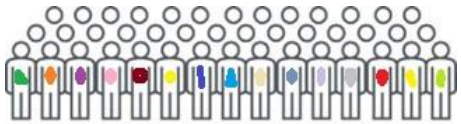
Improved appreciation of nature of longevity risk for pension schemes

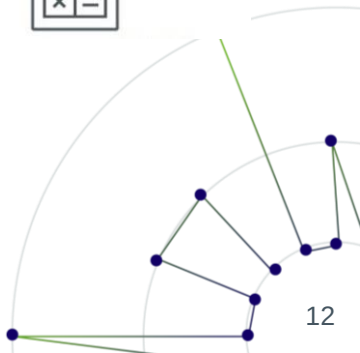
CLUB VITA

Comprehensive suite of analytics

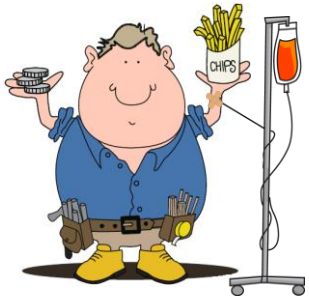
What are VITACURVES®?

- Our **statistical postcode longevity model** of the diverse range of survival patterns that we observe in the recent past.
- **Different pathways** depending on gender, lifestyle, affluence and occupation.
- Captures more **intricacies** in the diverse nature of individuals within a scheme, not just one-size-fits-all.
- Based on Club Vita's **uniquely rich dataset** and the m**ost accurate** way of setting baseline assumptions in the market.





VitaCurve rating factors



Life expectancy from 65: **12.5 years**

Life expectancy from 65: **22.9 years**

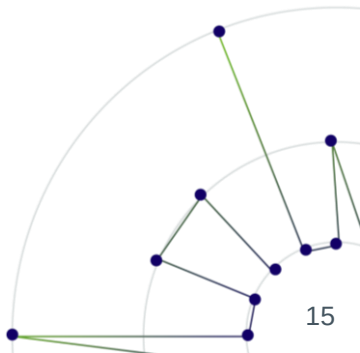
Unhealthy lifestyle postcode	Lifestyle 4½ years	Healthy lifestyle postcode
Low affluence	Affluence 3 years	High affluence
Ill health retirement	Health 2¾ years	Normal health retirement
Manual worker	Occupation < ½ year	Non-manual worker

No other approach uses all of this information

Constructing VitaCurves

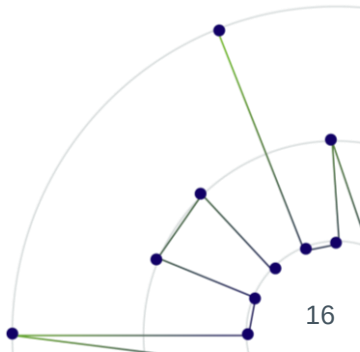
Calibration process - overview

- Steady flow of data throughout the year
  - Subject to data checks and cleansing on input
  - Identify ‘quality flags’ to ensure data quality
- Take ‘snap shot’ extract for annual calibration
- Segment data into similar groups
  - By gender, retirement health, and pensioner type
  - Identify function of mortality with age
- Identify available rating factors
  - Predictors of mortality
- Fit model to underlying data
  - Balance between responding to data and overfitting

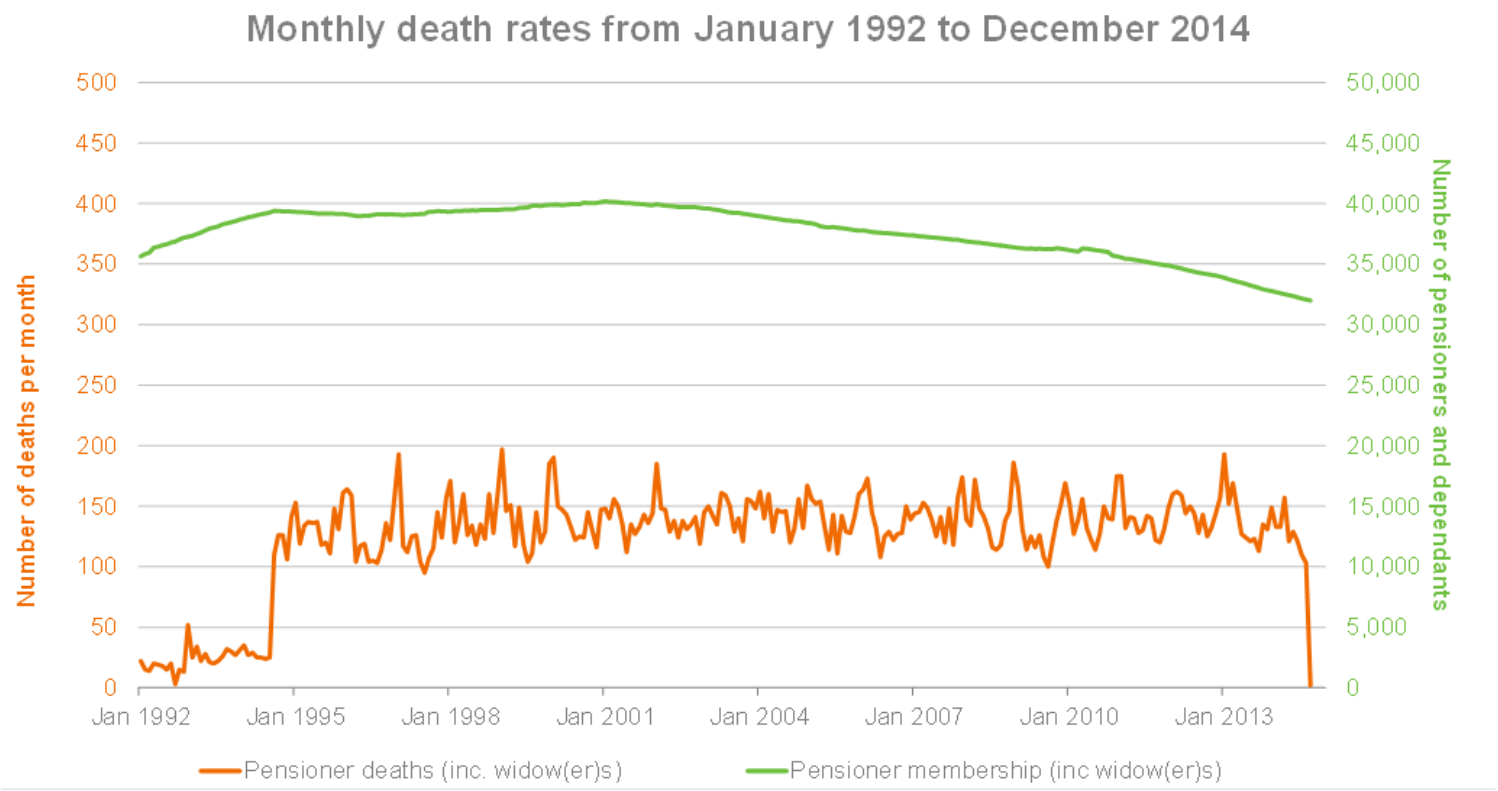


Data quality checks

- Data cleansed where possible
  - Postcodes
  - Missing dead?
- Apply range of ‘quality flags’ to data
- At individual level, check data is sensible
  - Reasonable dates, in expected order
  - Salary/pension look ‘reasonable’
- At scheme level, look at levels of missing/suspect data
  - Important that deaths and lives are consistent
- Some are fixed checks, others time based



Data quality checks



- Identify time period over which scheme data is reliable

Data extract

- Take extract from database
  - Generally around end January
- Calibration period of 3 years
  - Balance between smoothing and responsiveness
  - CV19 calibrated to 2015-2017

	Initial exposed to risk (2014-2016)	Deaths (2014-2016)
Pensioner men	2,434,287	74,281
Pensioner women	2,006,470	38,033
Widows and dependant women	723,577	44,728
Widowers and dependant men	97,438	4,359
Total	**5,261,772**	**161,401**

Choosing a model

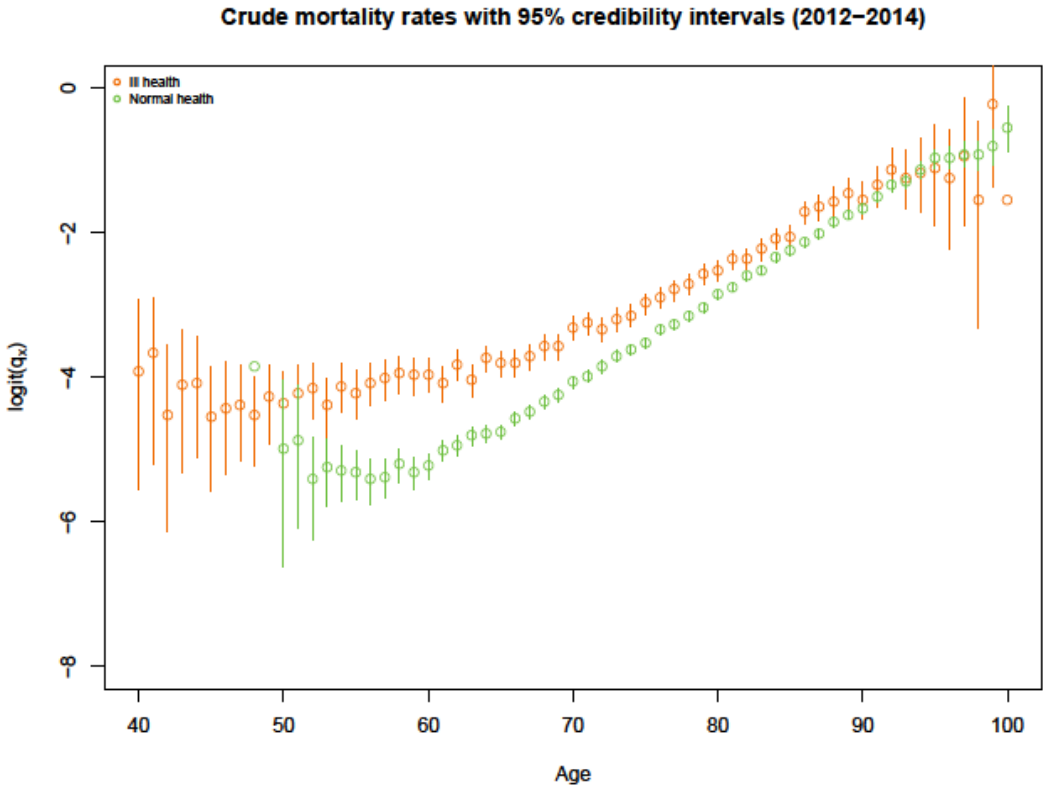
- Adopt a logistic Generalised Linear Model
- Each life either dies in the period or survives to the end
- Model assumes a binomial parametric relationship between age and probability of death
- Consider the linear age only case:

$$q_x = \frac{e^{\alpha+\beta x}}{1+e^{\alpha+\beta x}}$$

$$logit(q_x) = \log\frac{q_x}{1-q_x} = \alpha + \beta x$$

- Extend model:
  - more complex functions with age
  - Allow for other predictors of mortality

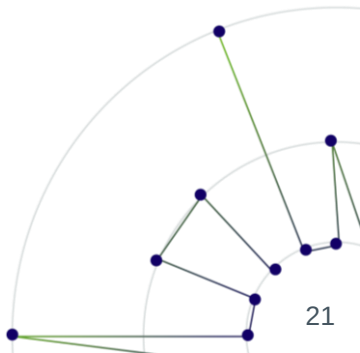
Stratifying the data



- Identify stratifiers to segment data, where:
  - Shape of mortality with age is different
  - Data covers different age ranges
  - Other rating factors have different meanings

Stratifying the data

- Use three stratifiers:
  - Gender
  - Retirement health
  - Pensioner type
- Gives 6 distinct ‘strata’
  - Male/Female normal health pensioners
  - Male/Female ill health pensioners
  - Male/Female dependants
- Other rating factors treated as ‘covariates’

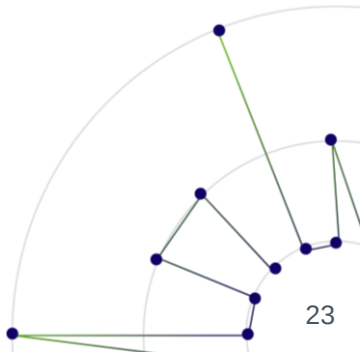


Fitting mortality by age

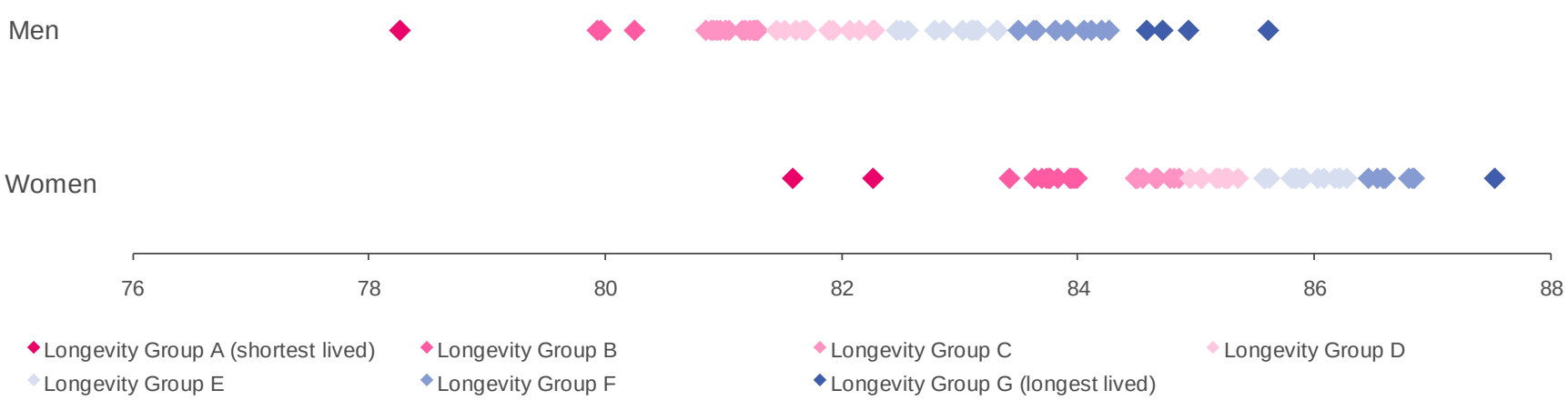
- For each of the six strata, consider:
  - Age range
  - Shape of mortality by age
- Age range selected based on providing credible experience data
- Mortality assumed to be a polynomial of the reciprocal of age
  - Better allowance for reducing impact of rating factors with increasing age

Rating factors

- 3 rating factors used to form strata
  - Gender
  - Retirement health
  - Pensioner type
- Remaining rating factors treated as covariates
  - Lifestyle – proxied by postcode
  - Affluence – either salary or pension
  - Occupation – nature of employment (where known)
- Other rating factors could be added in future
  - Marital status?
  - Duration?



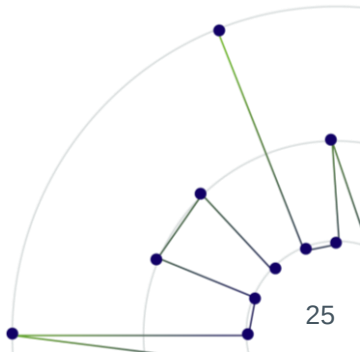
Lifestyling



- Postcode analysis used as a proxy for lifestyle/socio-economic group
- External provider supplies mapping of every postcode to one of 62 geo-demographic types
- Group these into 7 ‘longevity groups’
  - Separately for men and women

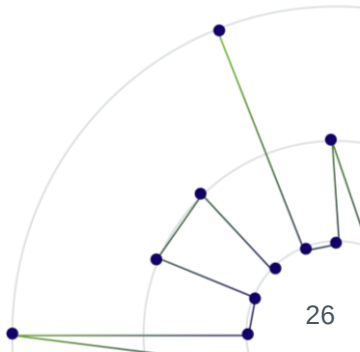
Affluence

- Salary or pension amount
- Revalue to common date
  - Currently July 2016
  - Use RPI for lives
  - Deaths adjusted to allow for expectations of pension increases
- Group into distinct bands
  - Separate for men and women
  - Separate for pensioners and dependants (pension only)
- Consider whether to use pension or salary



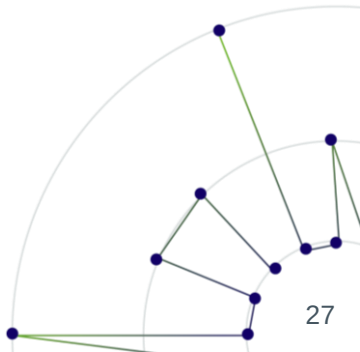
Occupation

- Historically LGPS schemes identified whether a job was manual or officer type
  - Different contribution rates were payable
- Some evidence that manual occupations have higher mortality rates than officer
- From April 1998 joiners no longer classified by occupation (as contribution rates standardised)



Clustering approach

- Use ‘top down’ hierarchical clustering approach
  - Recursive Partitioning and Regression Tree (RPART)
- k-fold cross validation approach to determine optimal numbers of groups
  - Divide into k subsets
  - Remove each in turn
  - Fit model to remaining k-1 subsets
  - Use to predict subset removed
- Verify chosen structure meets criteria
  - At least 2% of data
  - At least ½ year difference in life expectancy
  - Life expectancies are statistically different
- Can ‘group’ adjacent covariates if too similar



Functional forms

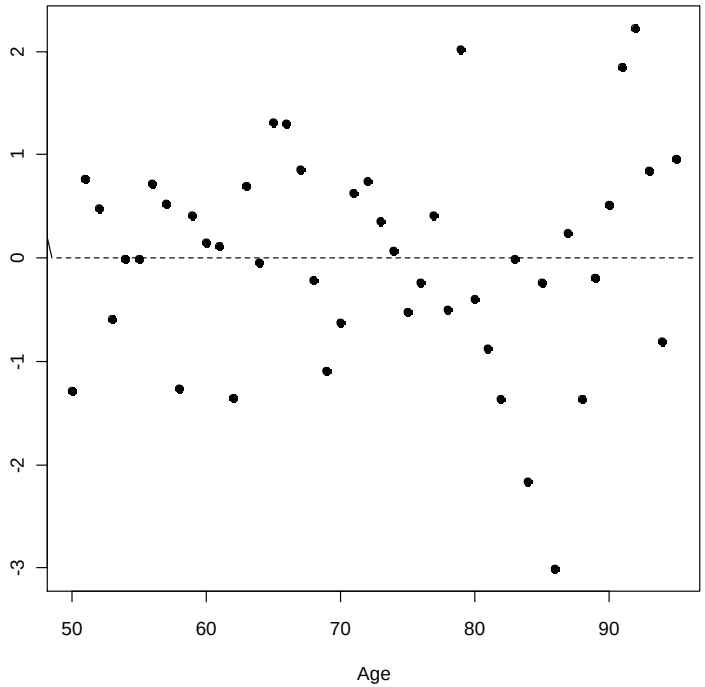
- As above, determine appropriate functional forms
  - Start from age only
  - Add other rating factors
- Male normal health – age only

$$Age^{-1} + Age^{-2} + Age^{-3} + Age^{-4} + Year_{Central}$$

- Male normal health – age, postcode, pension

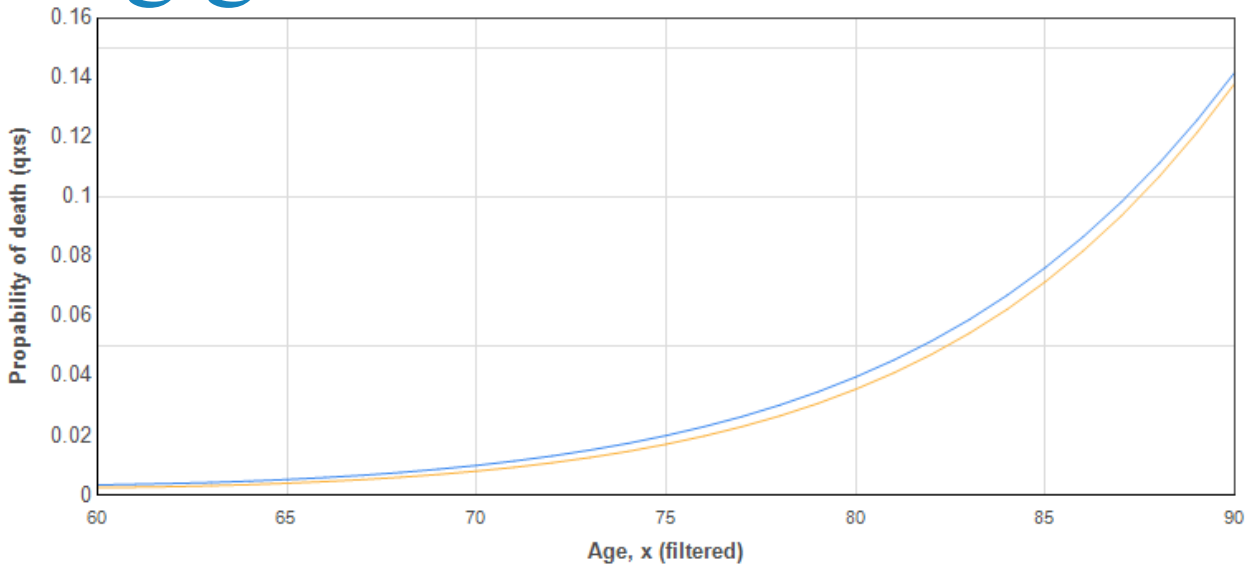
$$Age^{-1} + Age^{-2} + Age^{-3} + Age^{-4} + Year_{Central} + Age^{-4}{:}\,Postcode + Age^{-4}{:}\,Pension$$

Checking goodness of fit



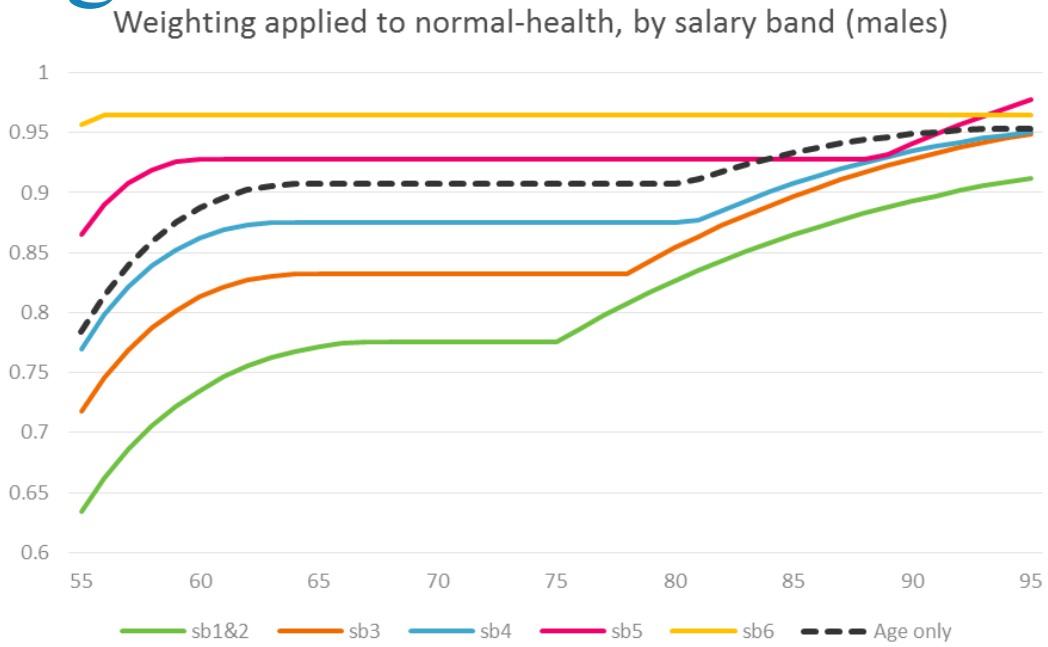
- Apply range of tests to fitted curves
  - Statistical tests (e.g. AIC, BIC)
  - Actuarial tests (e.g. Chi-squared, Signs test, runs test)
  - Validation tests (e.g. compare crude and fitted life expectancy)
- Look at overall picture
  - Apply weights to different tests

Checking goodness of fit



- Consistency checks
  - Increasing mortality rates with age
  - Curves in right ‘order’ (e.g. pension band 1 above pension band 2)
  - Less granular curves within extremes of granular curves (e.g. pension band 1 should be between pension band 1, longevity group A, and pension band 1, longevity group G curves)

Extending curves

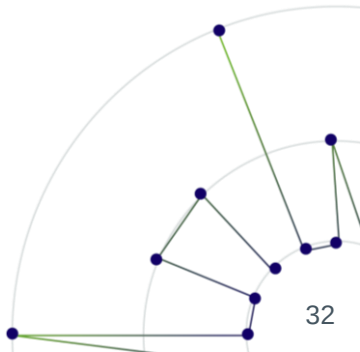


- Generate ‘all health’ curves
  - Weighted average of normal and ill health
  - Weights vary by gender, and (for men) by affluence
- Extend curves to older and younger ages
  - Down to age 16
  - Up to age 125

CLUB VITA

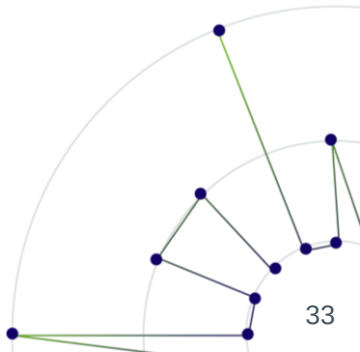
Final curves

- Ultimately end up with full suite of curves
  - Male/female
  - Pensioner/Dependant
  - Normal/ill-health/all-health
  - 7 Longevity groups (plus unknown)
  - 6 (men) / 3 (women) salary bands (plus unknown)
  - 5 (pensioner men) / 4 (pensioner women) / 2 (dependant men) / 4 (dependant women) pension bands (plus unknown)
  - Manual/Officer occupation (plus unknown)
- Total of 1,504 distinct combinations
- Each member allocated to curve based on their characteristics



Annual updates

- Fresh cut of data
  - Latest scheme data
  - New joiners to Club
- Check existing bands and longevity groups
  - Prefer not to change
- Check existing functional forms remain good fit
  - Again prefer not to change



Thank you