Forecasting the demand for care home places in the UK

OR58: Portsmouth University September 6th-8th 2016

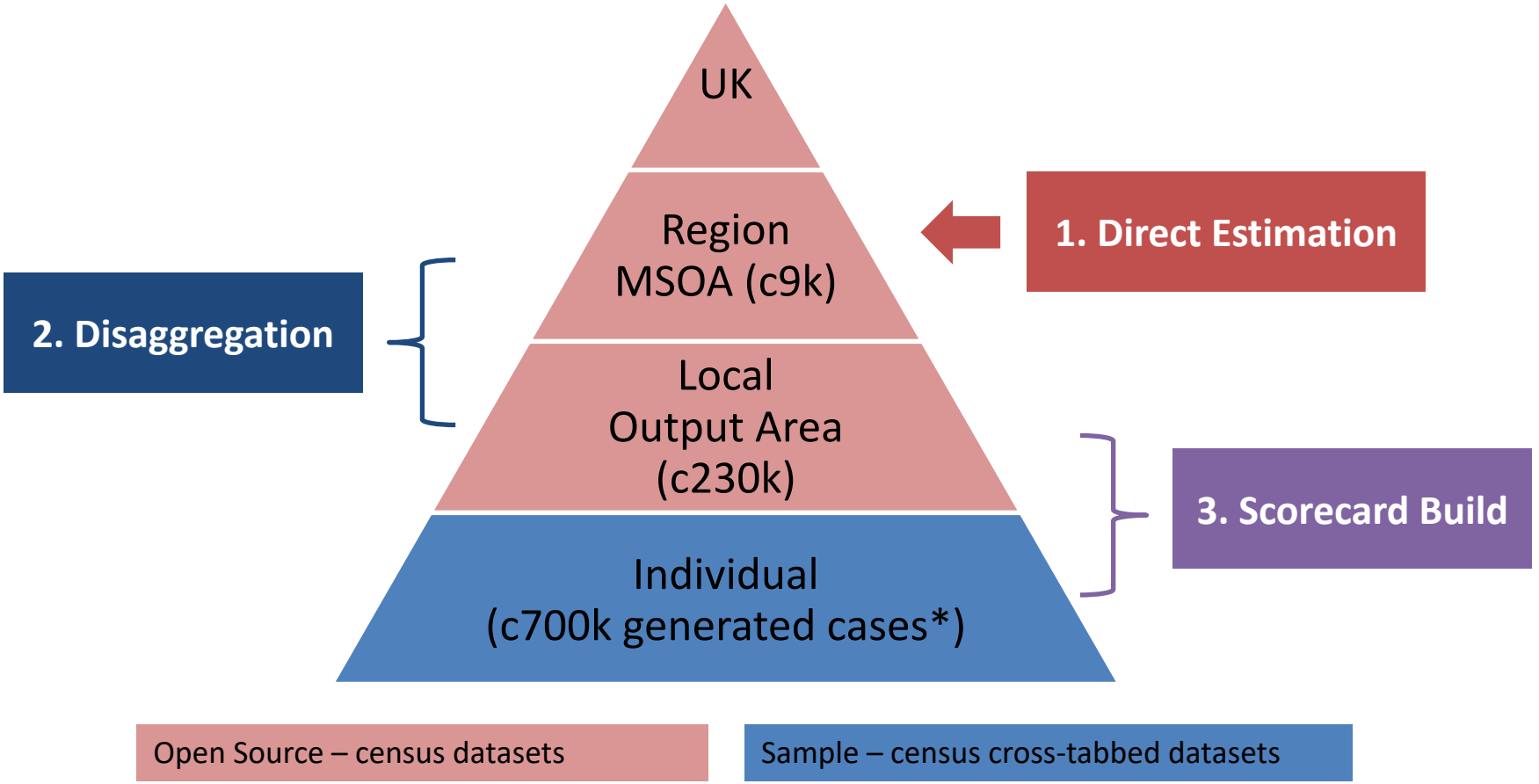
Authors:
Colin Stewart, More Metrics Ltd colin.stewart@moremetrics.co.uk
Peter Mannion Redmayne Consulting Ltd p.mannion@RedmayneConsulting.co.uk

Acknowledgements:
All of the model data we use is National Statistics data © Crown copyright and database right (2001 to 2016) and is used under the terms of the Open Government Licence.
Maps have been produced using QGIS (2.6 Brighton) with the OpenStreetMap plugin
Comparisons of our model outputs have been made to data published by PSSRU and ISD Scotland (National Statistics Publication)

Introduction to today's talk

- More Metrics build mortality risk models using open source data
- A particular interest of ours is to quantify the death rates of care home residents
  - Their death rates are significantly higher than household residents
  - Mortality risk models can be improved if "residency type" is included as a factor
- Including residency type in mortality models requires specific information:
  - Knowledge of who is a care home resident (i.e. an address match)
  - The variation in life expectancy of care home residents by sex, age, and type of care home (i.e. a set of hazard functions)
  - The rate at which household residents become care home residents by sex and age (i.e. a set of stock and flow rates)
- Today I will share with you
  - How we have obtained our own estimates for hazard functions, stock and flow rates using open source data
  - How our results can be used in other areas. I will use care home demand forecasting as an example of this

How do we build our mortality risk scorecards?



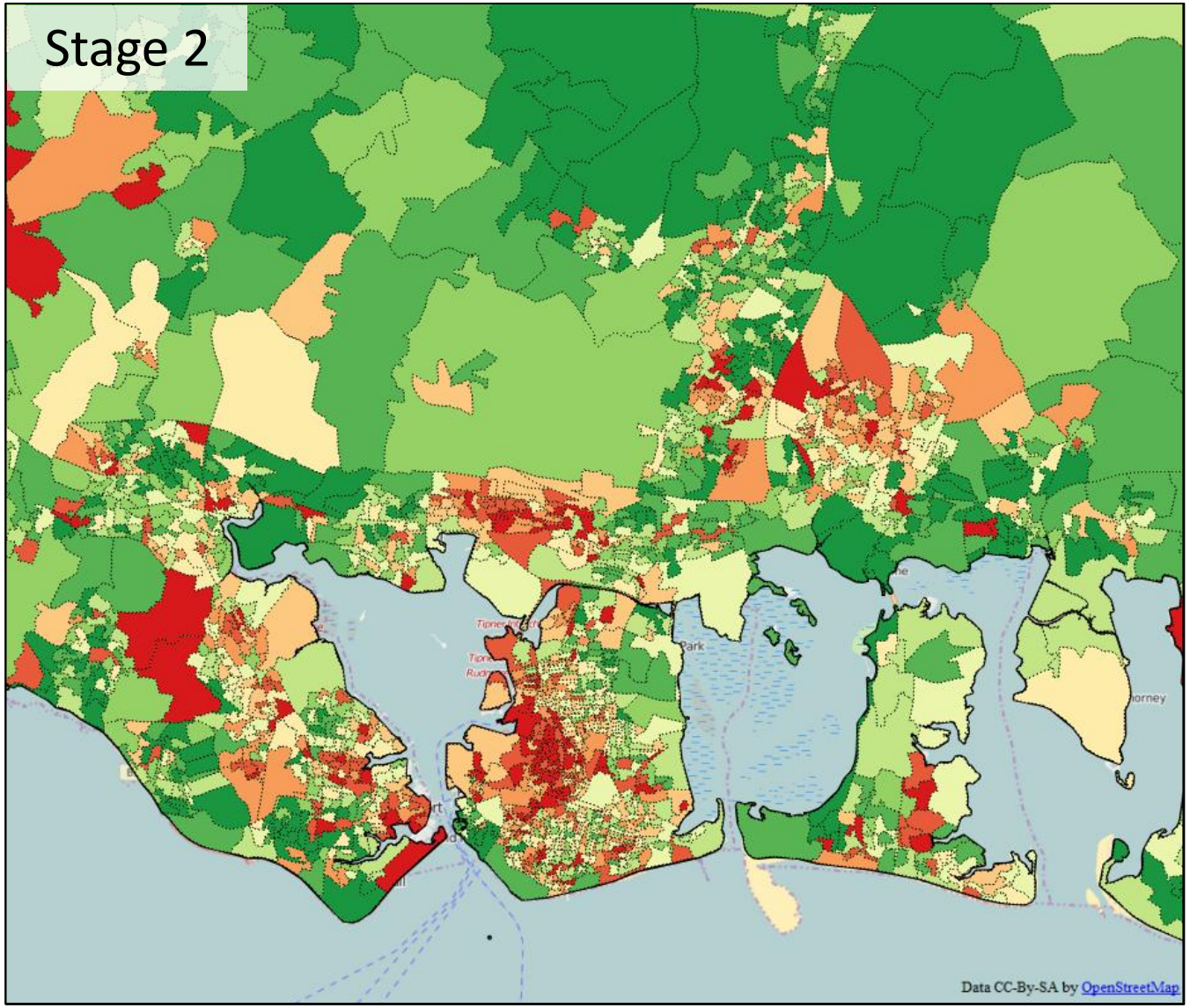
*Northern Ireland data not included in stage 3 but this region can be "scored up" to provide complete UK coverage

Stages 1 and 2: Direct estimation and disaggregation

Creation of local area estimates of the average mortality risk

Mapped data is for Portsmouth and surrounding areas for males of all ages and all residency types

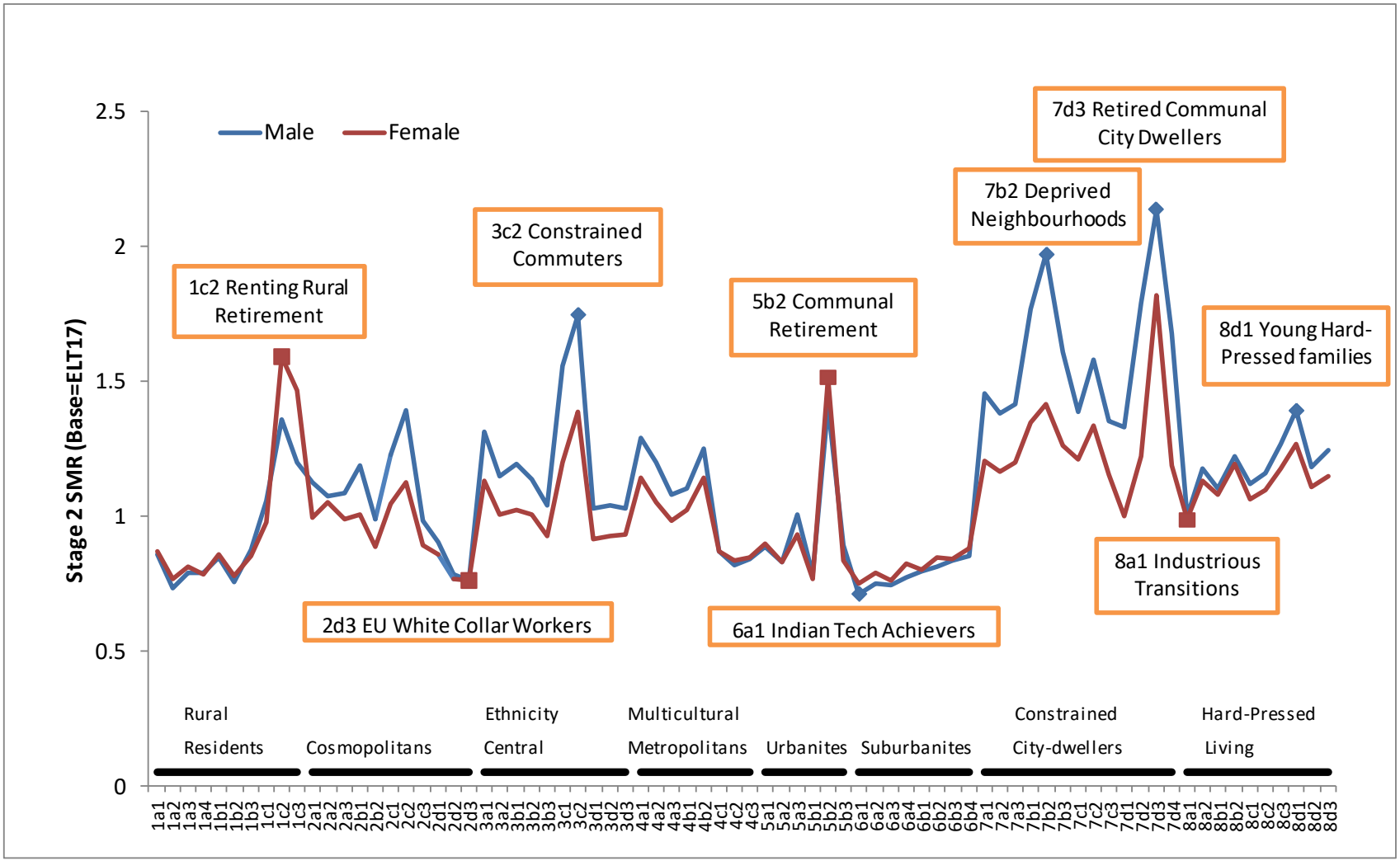
Map Key	Decile
	1 Lowest Risk
	2
	3
	4
	5
	6
	7
	8
	9
	10 Highest Risk



Stage 2: Disaggregation- by Output Area Subgroup Classification

Note that some communal retirement areas are featured in the peaks across the 76 OACs

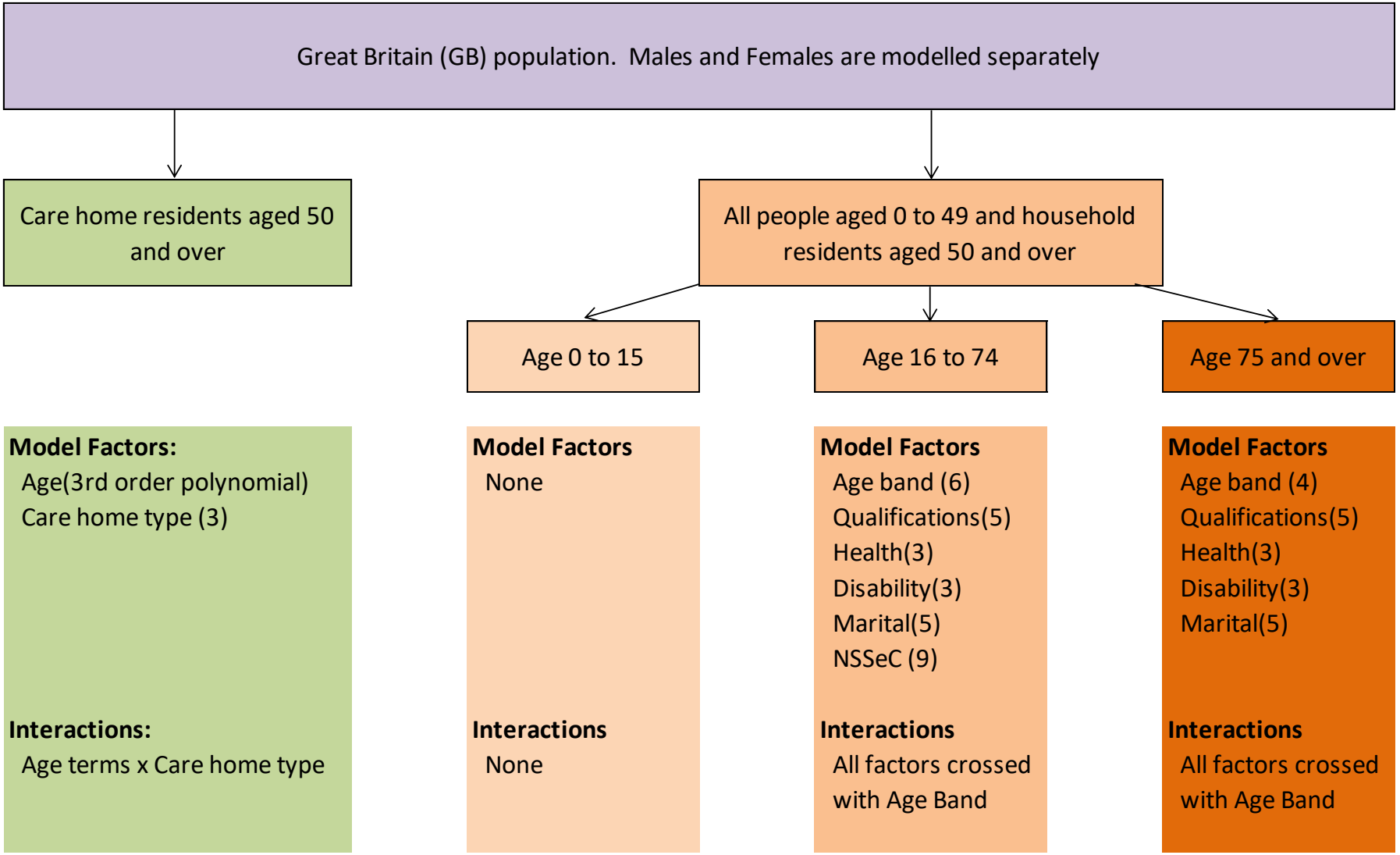
We will use these Stage 2 OAC values to ensure the scorecard we build in Stage 3 is properly calibrated



These risk values are the average Stage 2 score for all male / female residents of households and care homes combined in each OAC

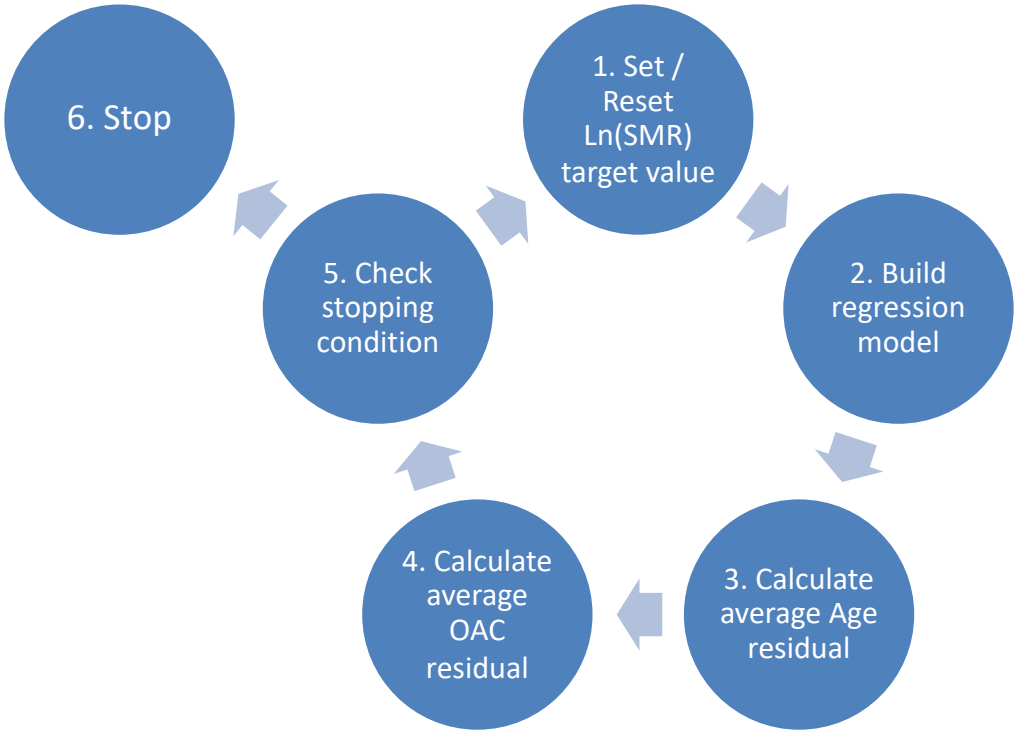
Stage 3: Scorecard Build – Creating the modelling dataset

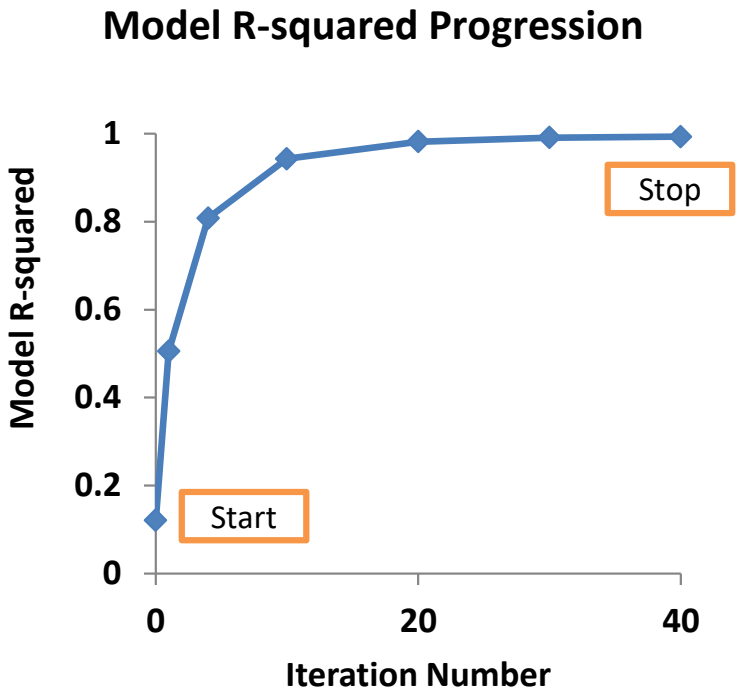
Stratified sampling is used to generate a dataset for 1% of household residents and 10% of care home residents in GB
Cross-tabbed census datasets help us do this accurately. The dependent variable is Ln(SMR)



Stage 3: Scorecard Build – Sharing out risk fairly to individuals

We do this iteratively using our stage 2 Output Area Ln(SMR) values as the initial target value

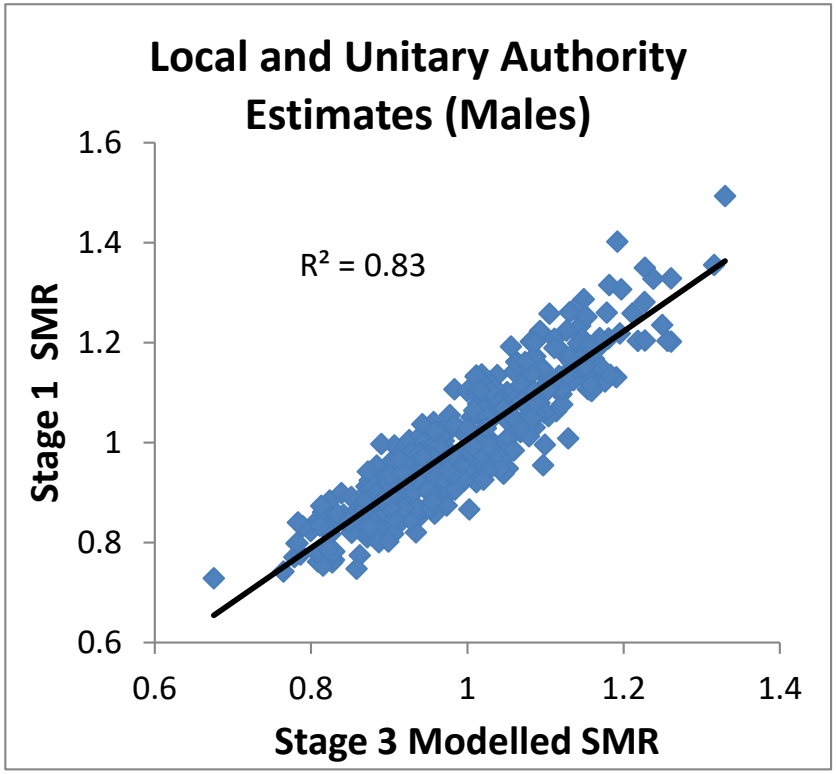


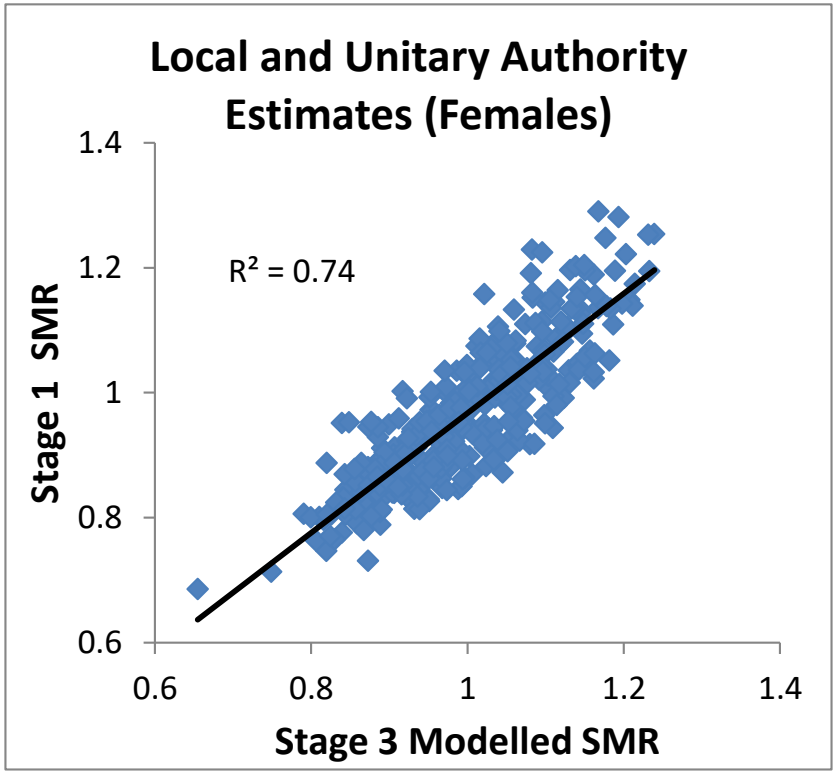


- Individual target values are reset at each iteration by applying calculated residuals
  - Age: Total Modelled deaths = the Life table value for every age from 0 to 100
  - OAC : Total Modelled deaths = the Stage 2 value for every OAC

Scorecard results: Overall Local and Unitary Authority predictions

Stage 3 national scorecard explains over 70% of mortality variation across LAs and UAs



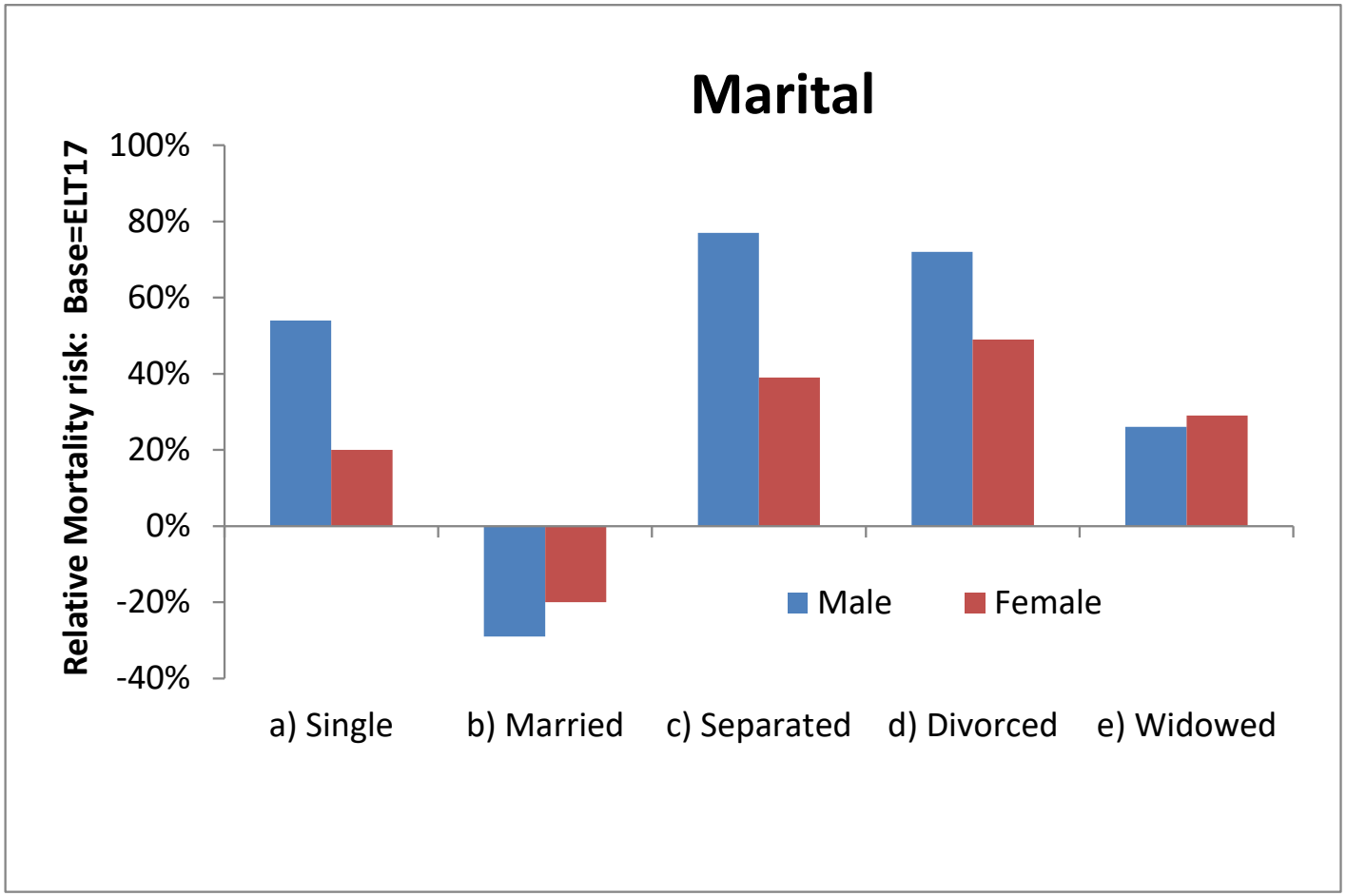


This assessment is a high level check to see that our stage 3 national scorecard accurately predicts the variation seen at a local authority / unitary authority level found at stage 1 (direct estimation).

When the stage 3 national scorecard is used in practice, adjustments are made to the national scorecard to account for regional variations which further improves the fit. This is done using nearest neighbour datasets to robustly apportion stage 3 model residuals. More details on this aspect can be obtained from the authors on request.

Scorecard results: Household residents aged 50 to 64

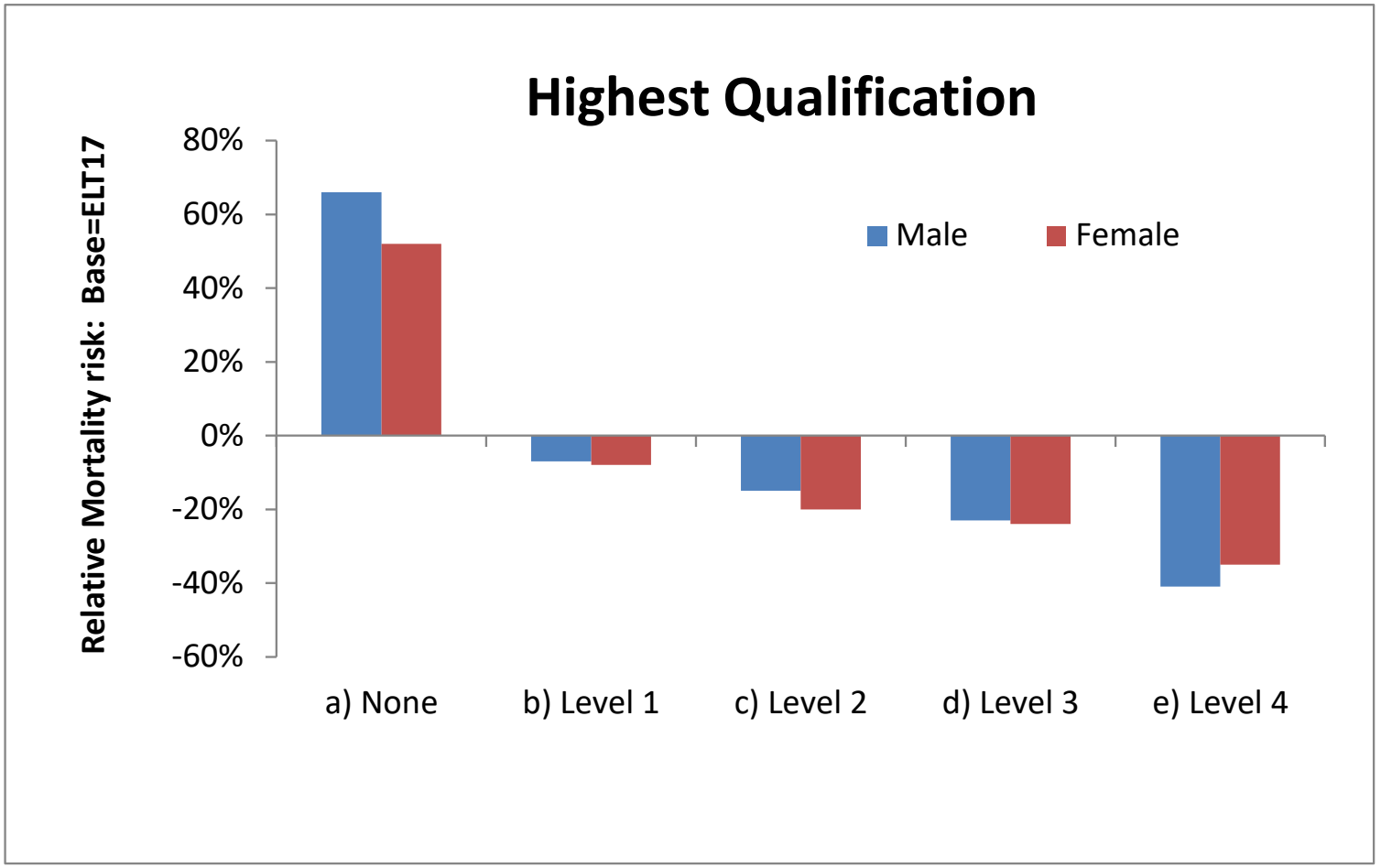
Marital Status



These risk values are the average score for all household residents aged 50 to 64 in each category. They are therefore not model coefficients but the mean model score by sex, age band and marital status category.

Scorecard results: Household residents aged 50 to 64

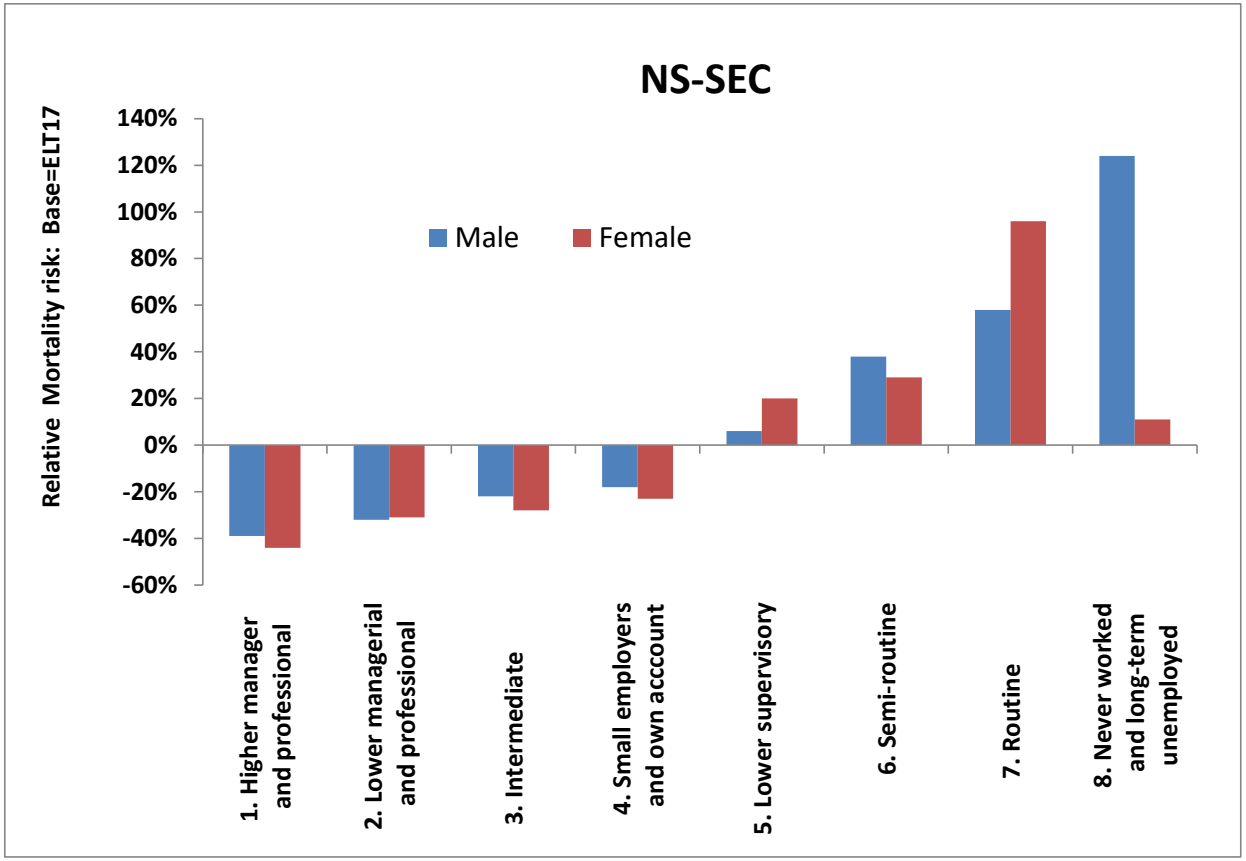
Highest Qualification



These risk values are the average score for all household residents aged 50 to 64 in each category. They are therefore not model coefficients but the mean model score by sex, age band and highest qualification category.

Scorecard results: Household residents aged 50 to 64

National Statistics Socio-economic Classification (NS-SEC)

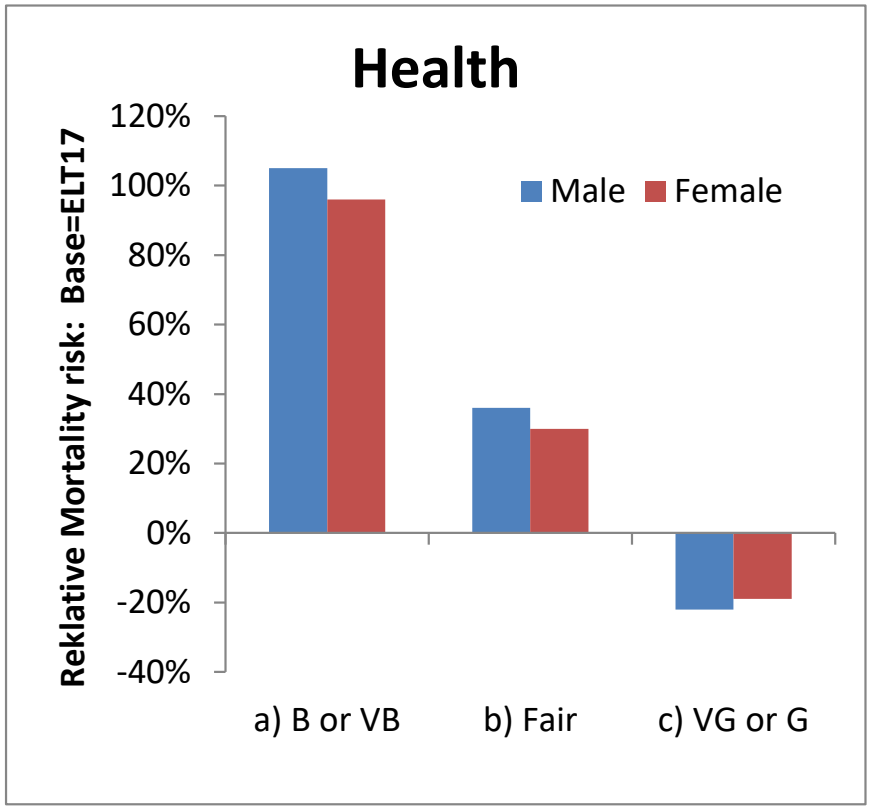


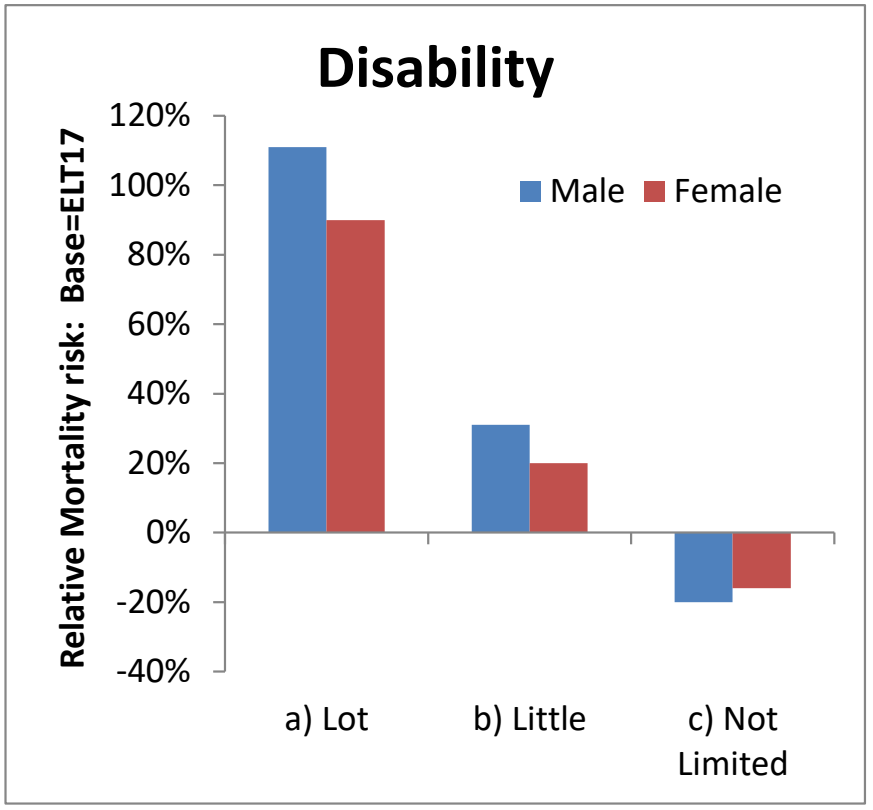
A marked difference in the sexes is seen for the "Never worked and long-term unemployed" classification. This is believed to be because a higher proportion of females in this group have chosen to "stay at home" and are lower risk.

These risk values are the average score for all household residents aged 50 to 64 in each category. They are therefore not model coefficients but the mean model score by sex, age band and NS-SEC category.

Scorecard results: Household residents aged 50 to 64

Health and Disability

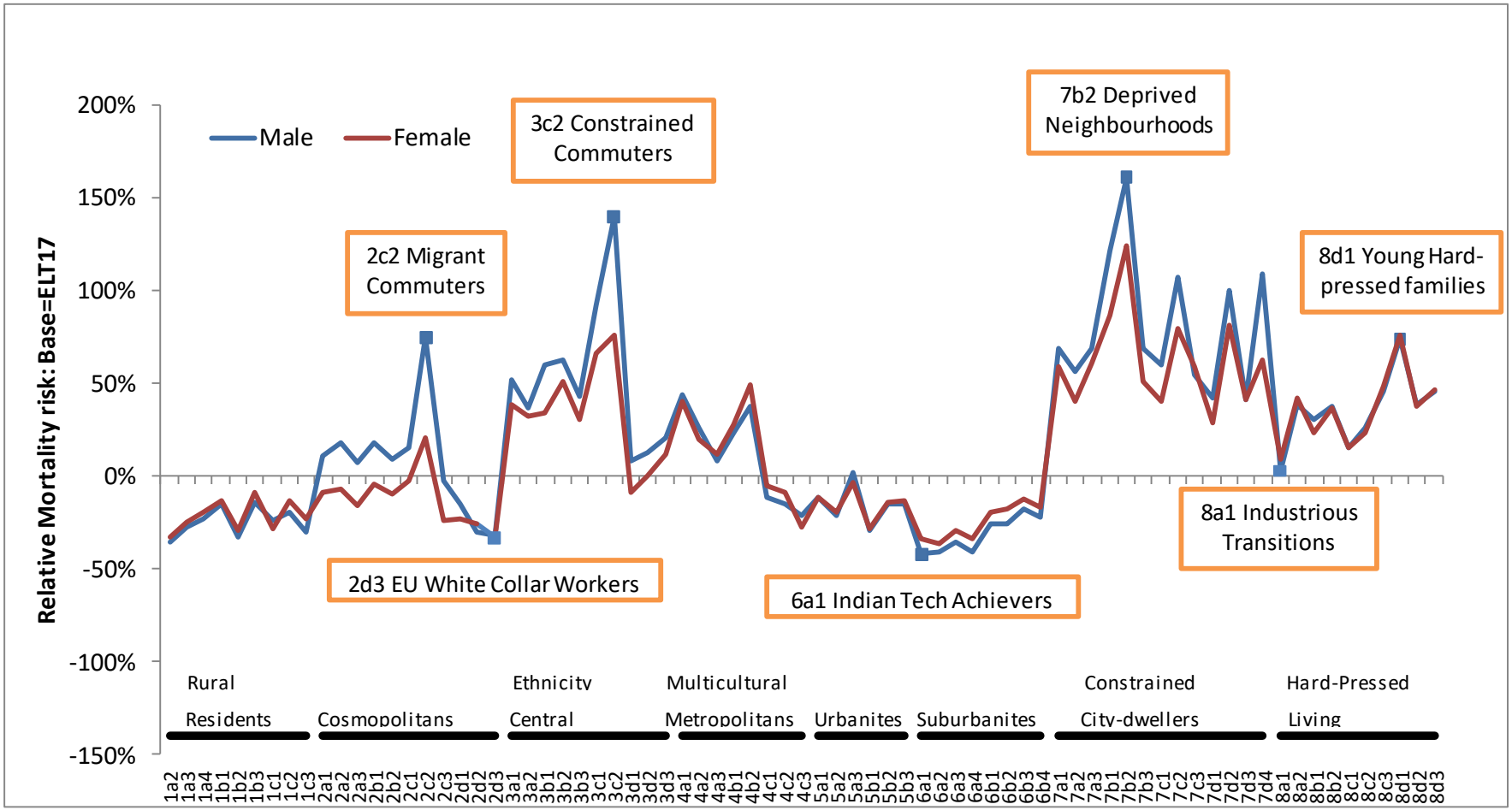




These risk values are the average score for all household residents aged 50 to 64 in each category. They are therefore not model coefficients but the mean model score by sex, age band and health or disability category.

Scorecard results: Household residents aged 50 to 64

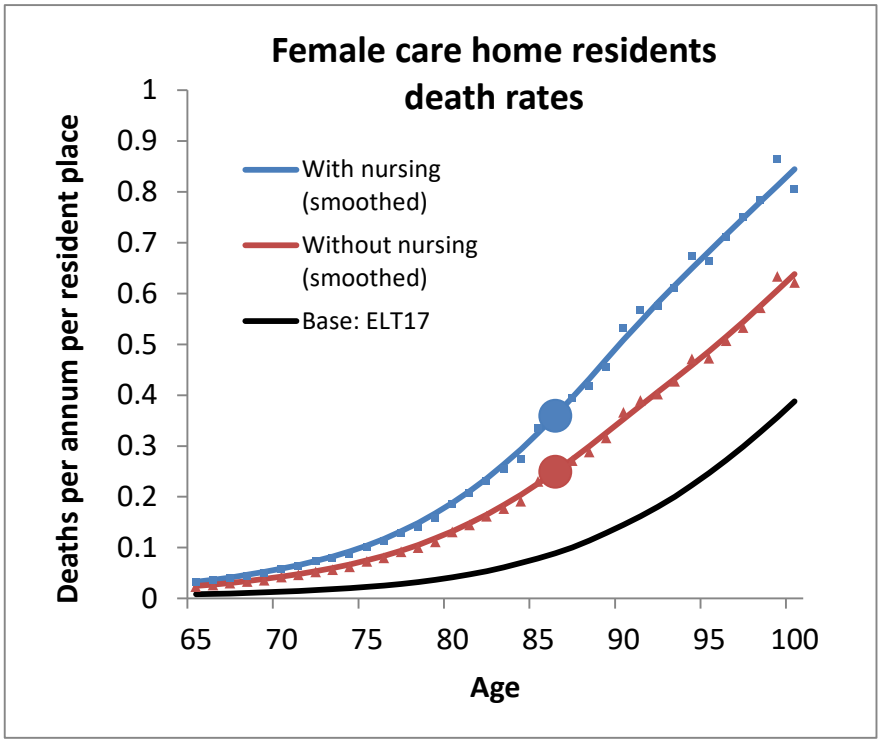
Output Area Subgroup Classification. Note the absence of the peaks for "retirement areas" for this age group

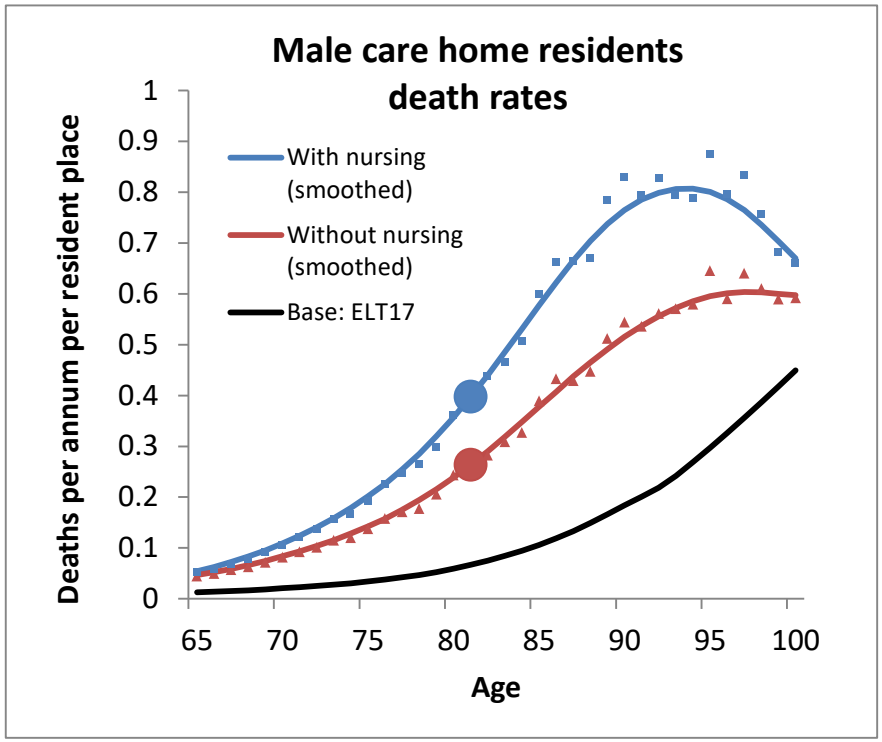


These risk values are the average score for all household residents aged 50 to 64 in each category. They are therefore not model coefficients but the mean model score by sex, age band and output area classification.

Scorecard results: Care Home Population – Scorecard death rates

The death rate used here is the number of deaths per year per resident place, assuming residents are replaced on death
A value of 0.5 means that on average there will be one death and one admission every 2 years to the care home system

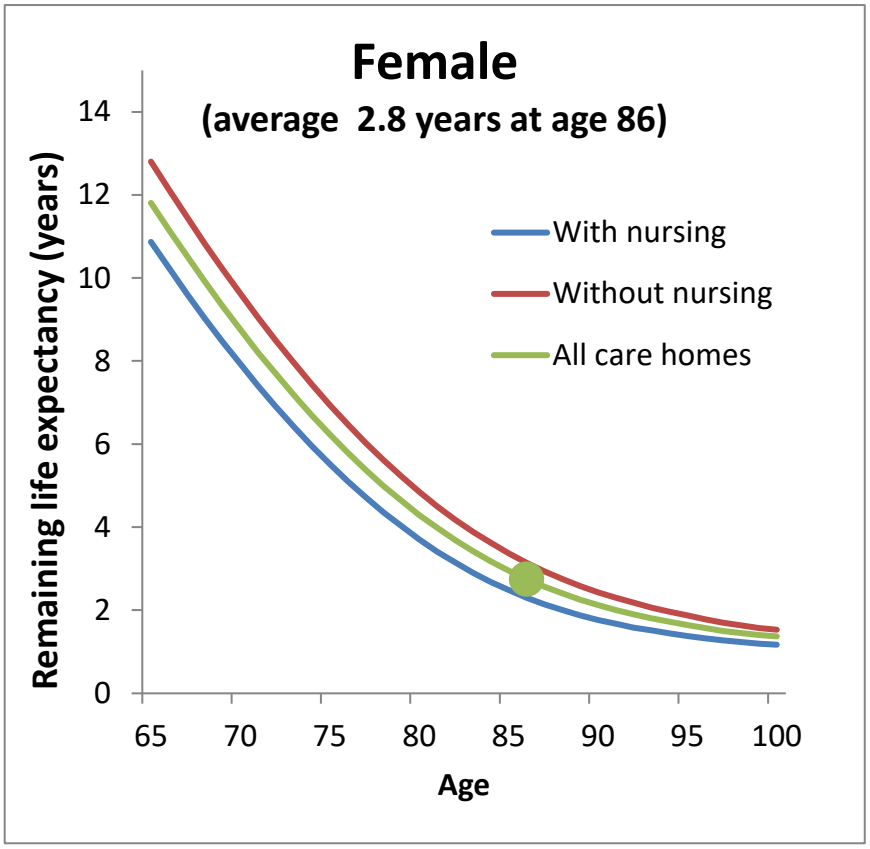


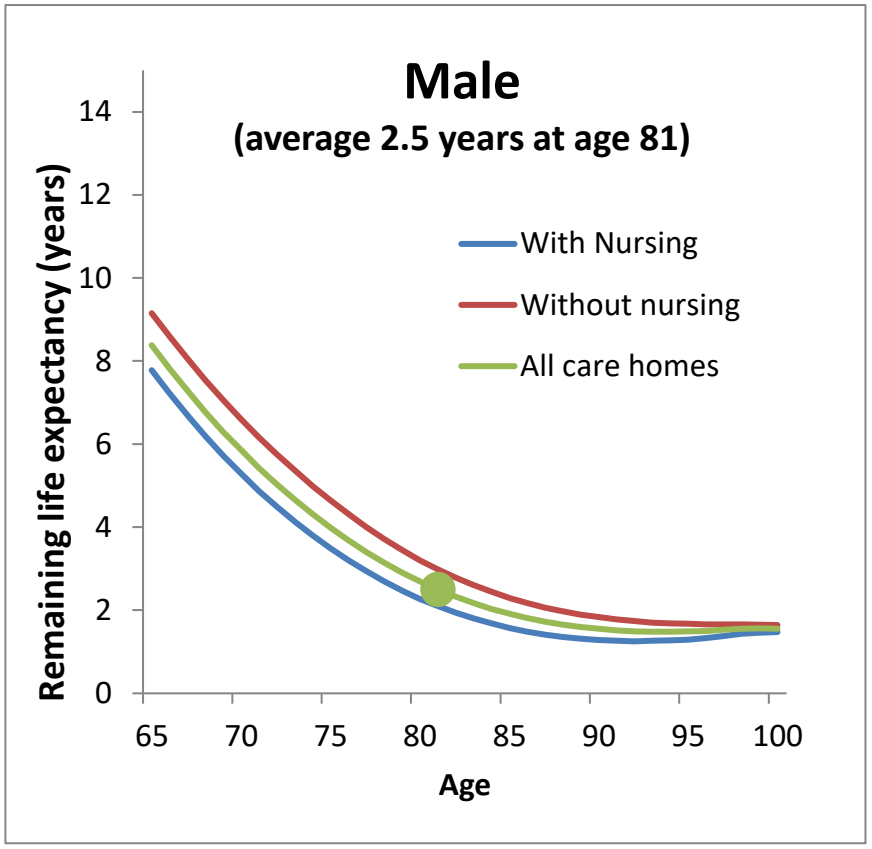


Key points to note:

- Nursing homes have higher death rates than those without nursing
- Care home death rates increase with age but rates flatten / peak at 90 to 95 for males
- Males have higher death rates than females across the range, but are on average 5 years younger than females, 81 versus 86 in our sample.
- The large dots on the graphs represent these average ages which are for residents aged ≥ 65

Scorecard results: Care Home Population – life expectancy





These life expectancy estimates have been calculated from the scorecard hazard functions assuming that a care home resident stays in the same care home type till death and is therefore not discharged

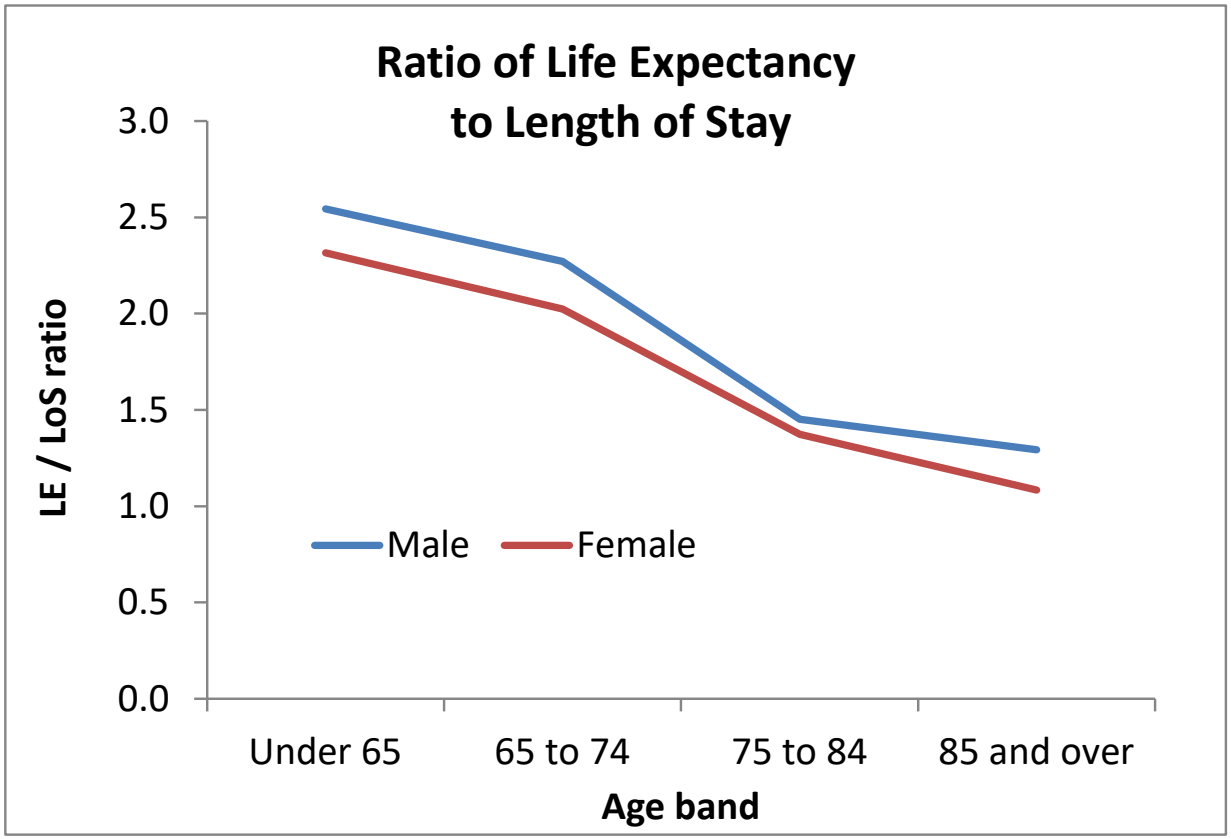
Scorecard results: Comparison of Life expectancy and length of stay

Our life expectancy results look high when compared to length of stay values reported by Forder et al*
However the measures may not be comparable – so other ways of validating our results are needed

Our Life expectancy values are between 2.5x and 1.2x higher than the unadjusted length of stay values reported by Forder et al.

However the measures are difficult to compare:

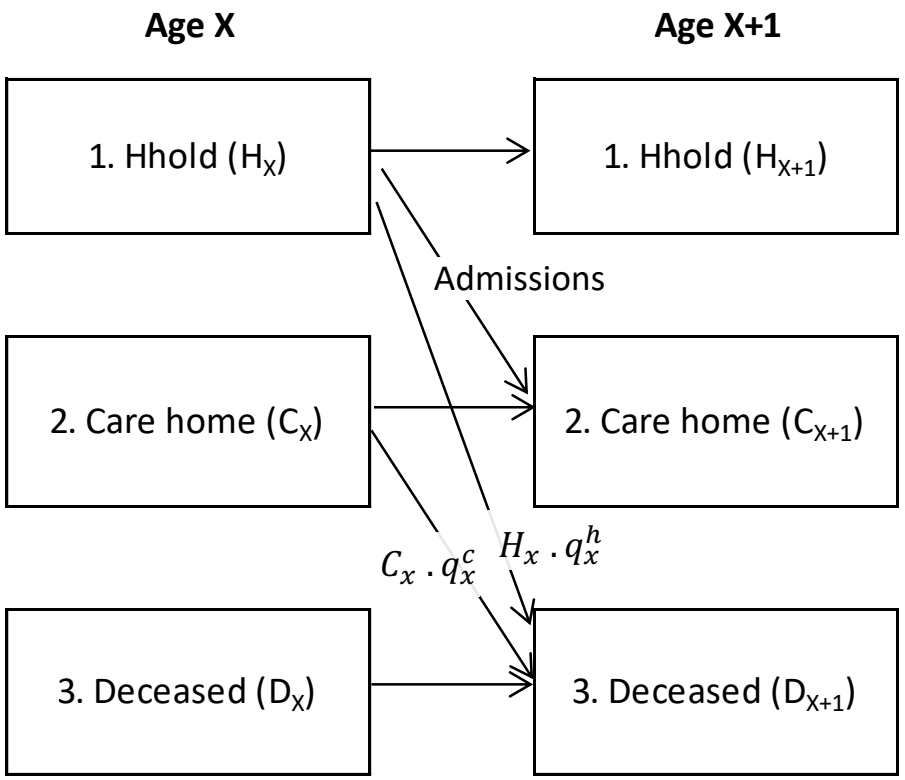
- *Forder et al:* Observed Length of Stay (LoS) for individuals admitted at age x.
- *More Metrics (MM):* Calculated Life Expectancy (LE) for survivors at age x, admitted at any age ≤ x
- Data taken at different time points. Forder et al is for deaths 11/2008 to 5/2010 compared to More Metrics at census day 2011



*Forder, J and Fernandez, J-L (2011) Length of stay in care homes, Report commissioned by Bupa Care Services, PSSRU Discussion Paper 2769, Canterbury: PSSRU, published Jan 2011. Points from table 7 (unadjusted figures) were used to generate ratios in the graph.

Stock and flow calculations (as at 2011)

We use the care home hazard function in conjunction with life table and census data to calculate flow rates



- We set total stock values at each age to equal the relevant life table values
  - $H_x + C_x = l_x$ (ELT17 lives)
  - $D_x = \sum_0^x d_x$ (ELT17 deaths)
- Care home values as a proportion of total lives , C_x/l_x is known from 2011 census data. From this, all stock values can be found at every age
- Sum of death flows at each age sum to the relevant life table values
  - $H_x \,.\, q_x^h + C_x \,.\, q_x^c = l_x . q_x$. (ELT17)

where q_x^c and q_x^h are known allowing us to estimate a value for admissions we can compare to other data sources

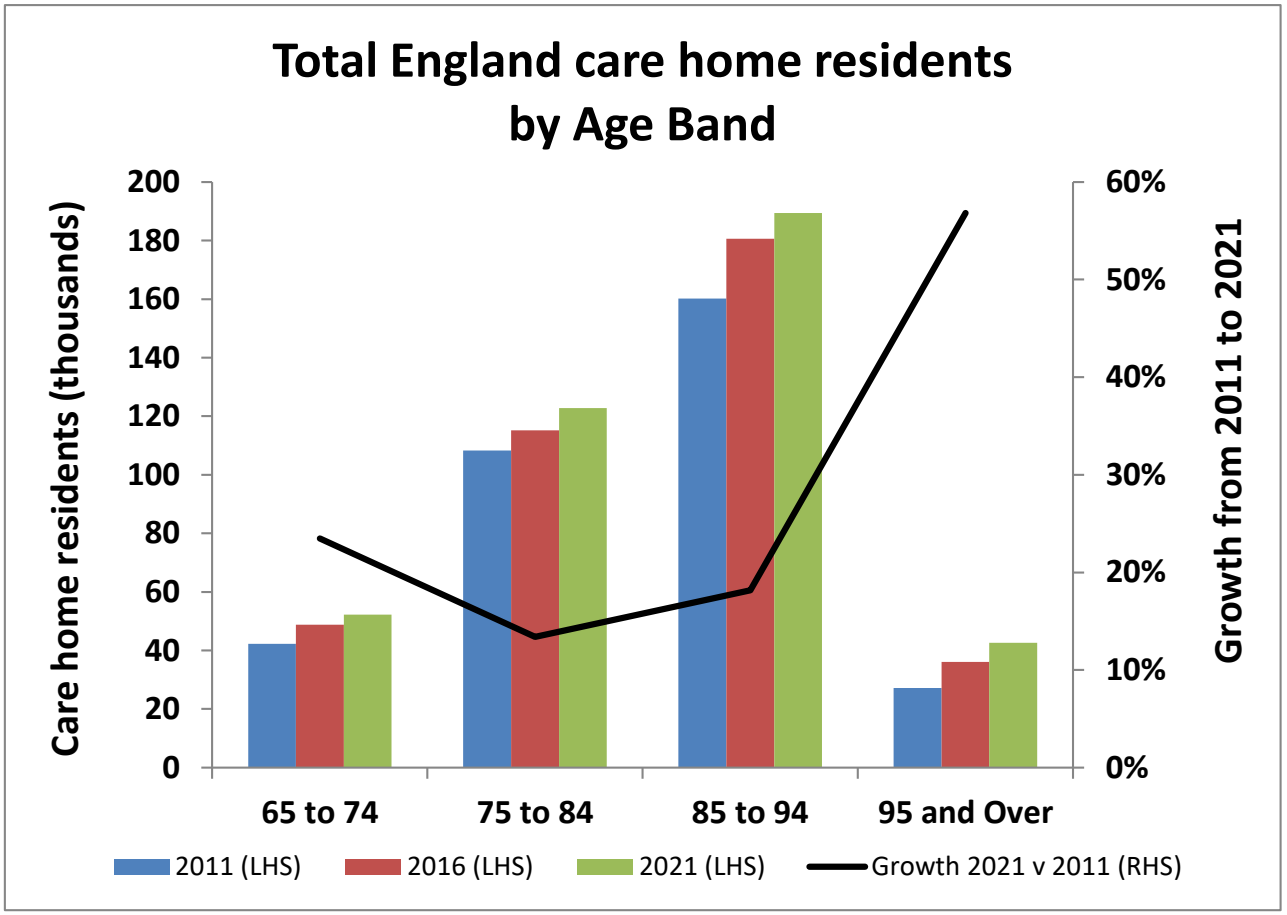
- Using Scotland as an example we can compare the More Metrics (MM) modelled estimates to the relevant care home census figures published by the Information Services Division (ISD) of National Statistics Scotland*
  - Number of older long stay residents (2011) = 33.6k (ISD) versus 36.1k (MM sampled to reflect 2011 census)
  - Number of deaths (2011 / 12) = 10.1k(ISD) versus 10.8k (MM applying scorecard death rates).

This is a close match assuming that 3.4k "long term discharges" do not count as care home deaths in the MM figures

**Care Home Census, Scottish Statistics on Adults Resident in Care Homes, 2006 - 2015*

Simulation results: Demand based on 2011 stock and flow rates

Figures are totals for both sexes in England and show the impact of an ageing population



How do we generate forecasts?

- An initial set of stock values at every age is obtained from 2011 census data for year 1
- The flow rates are then applied to calculate new stock values for 2012 at every age
- This process is repeated until 2021 updating the stock figures each year
- The forecast assumes that the propensity to be admitted to a care home and the death rates stay the same (as at 2011) for all ages and sex

Summary and concluding remarks

- We have shared with you today our approach to mortality modelling and how we are using this work to come up with our own estimates for care home death rates

- Our interest in this topic has to date been prompted by improving our offering to potential clients in the Life and Pensions Industry

- We believe that our analysis may also be of use in medical and care modelling. For example to forecast future demand for care home places

- We have more work underway to validate our model outputs
  - comparisons of our results to length of stay data is so far inconclusive
  - comparison to national admissions data is looking more promising

- We are very aware that many of you may already have the answers we are seeking or better ways of doing things – so would welcome any thoughts.

Thanks and follow up

Thanks goes to the following organisations

Our work is reliant on having access to open source data. More Metrics would therefore like to thank all the staff at ONS, NRS, NISRA, PSSRU, ISD (amongst others); and the contributors to mapping data and tools such as Ordnance Survey (OS), QGIS, OpenStreetMap (amongst others) for their continuing efforts.

Follow up

If you can make use of our care home analysis in your own work or want to do your own validation of our results, please get in contact. We will provide you with additional working papers and spreadsheet outputs where possible. In return we would welcome constructive feedback and your help in making sense of a complex subject.

Our contact details

Colin: colin.stewart@moremetrics.co.uk

Peter: p.mannion@RedmayneConsulting.co.uk