Trends in Mortality
Socio-Economics and Causes of Deaths

Torsten Kleinow
based on joint research with:
Frank van Berkum, Michel Vellekoop, Andrew J.G. Cairns, Alex Yiu, Jie Wen, George Streftaris

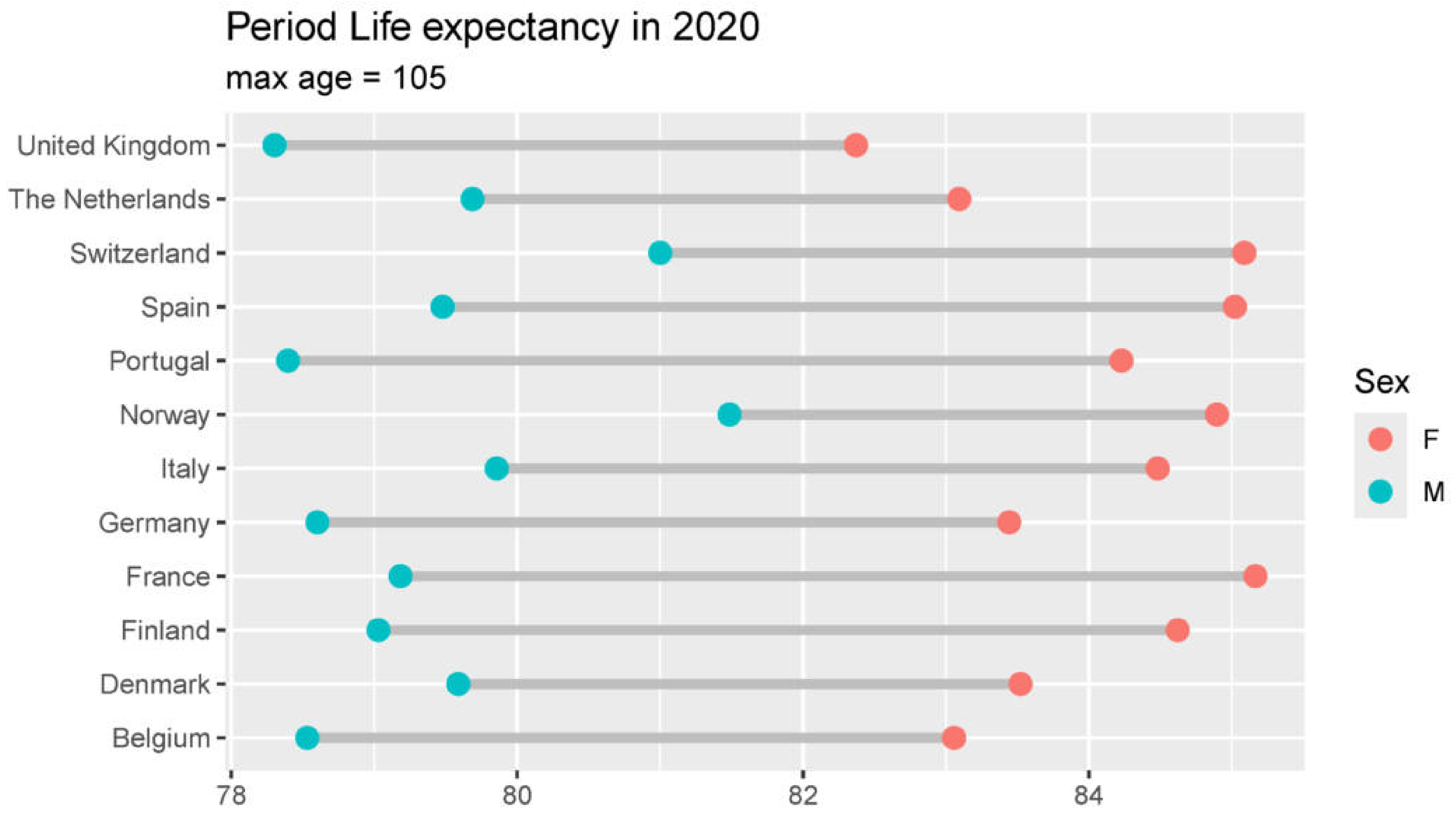
30 April 2025
DAV-/DGVFM-Jahrestagung

University of Amsterdam
RCLR Research Centre for Longevity Risk

Outline

Introduction
Mortality in Europe - two examples

Life expectancy in European countries



Basic Definitions

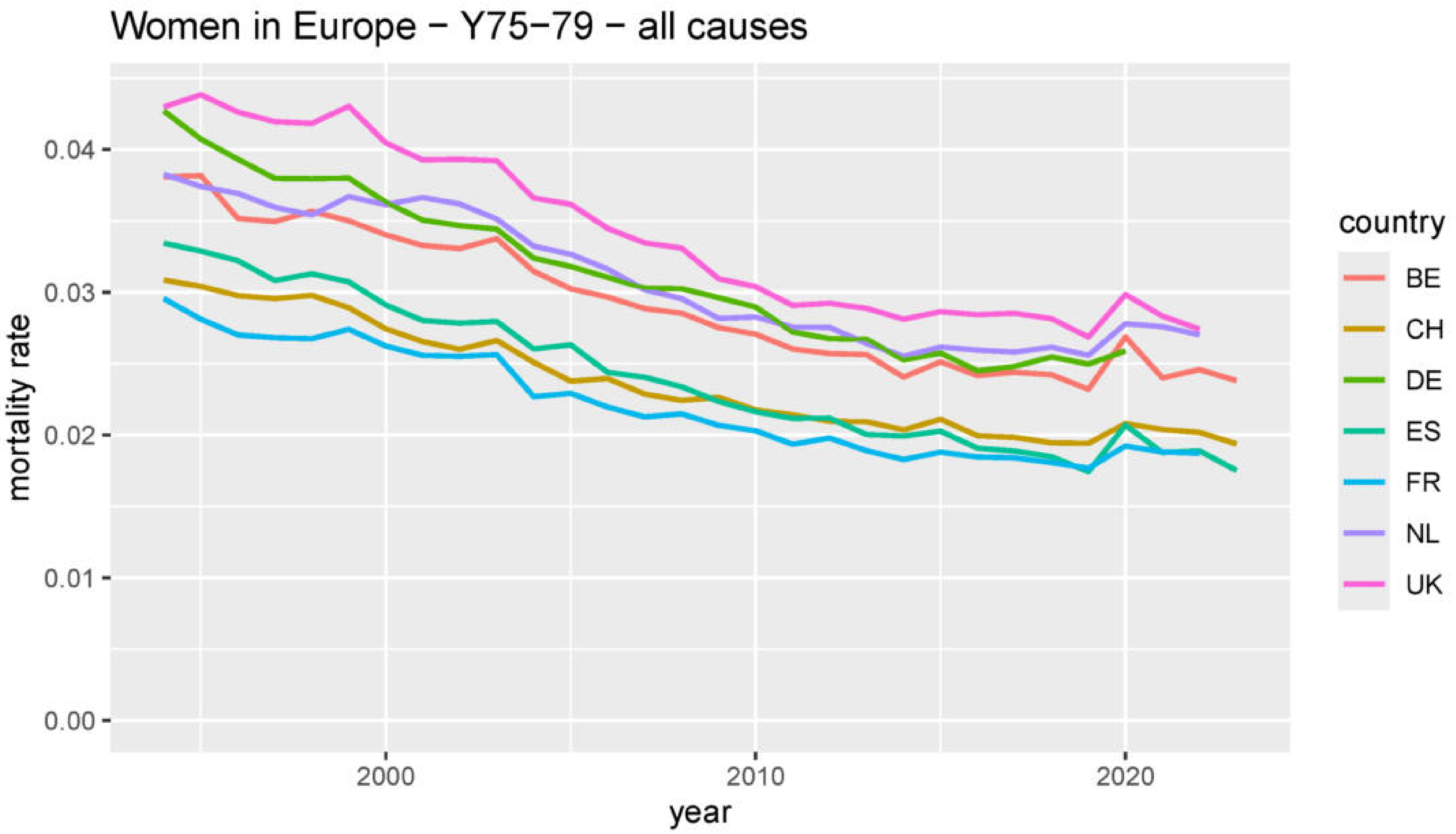
For any given population, each age (group) x and year t we define

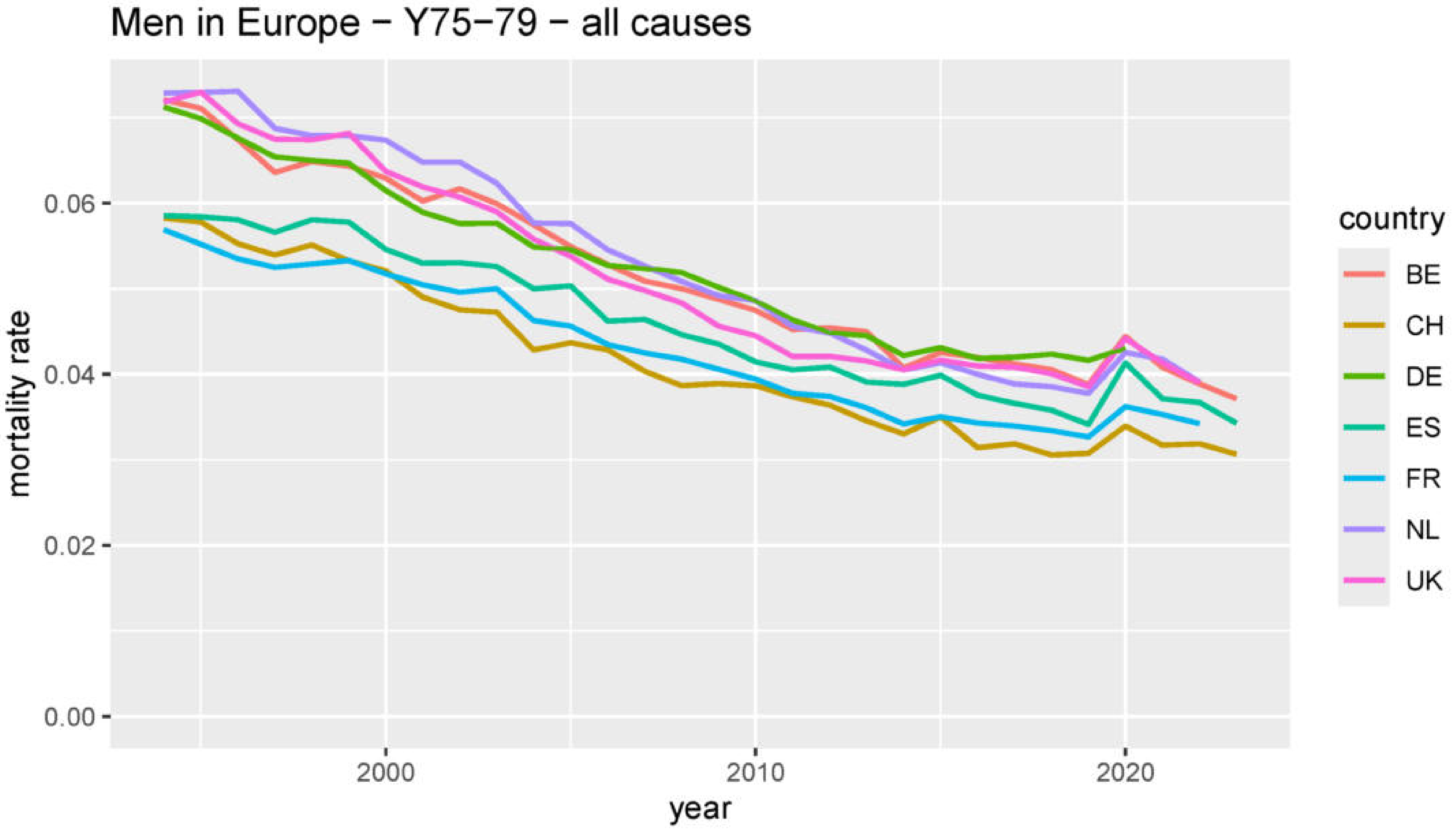
- Observed death d_{xt} as realisations of RVs D_{xt}
- Central Exposure to risk E_{xt}: mid-year population estimates
- For a given force of mortality μ_{xt} we assume

$$\mathbb{E}[D_{xt}] = \mu_{xt} E_{xt}$$

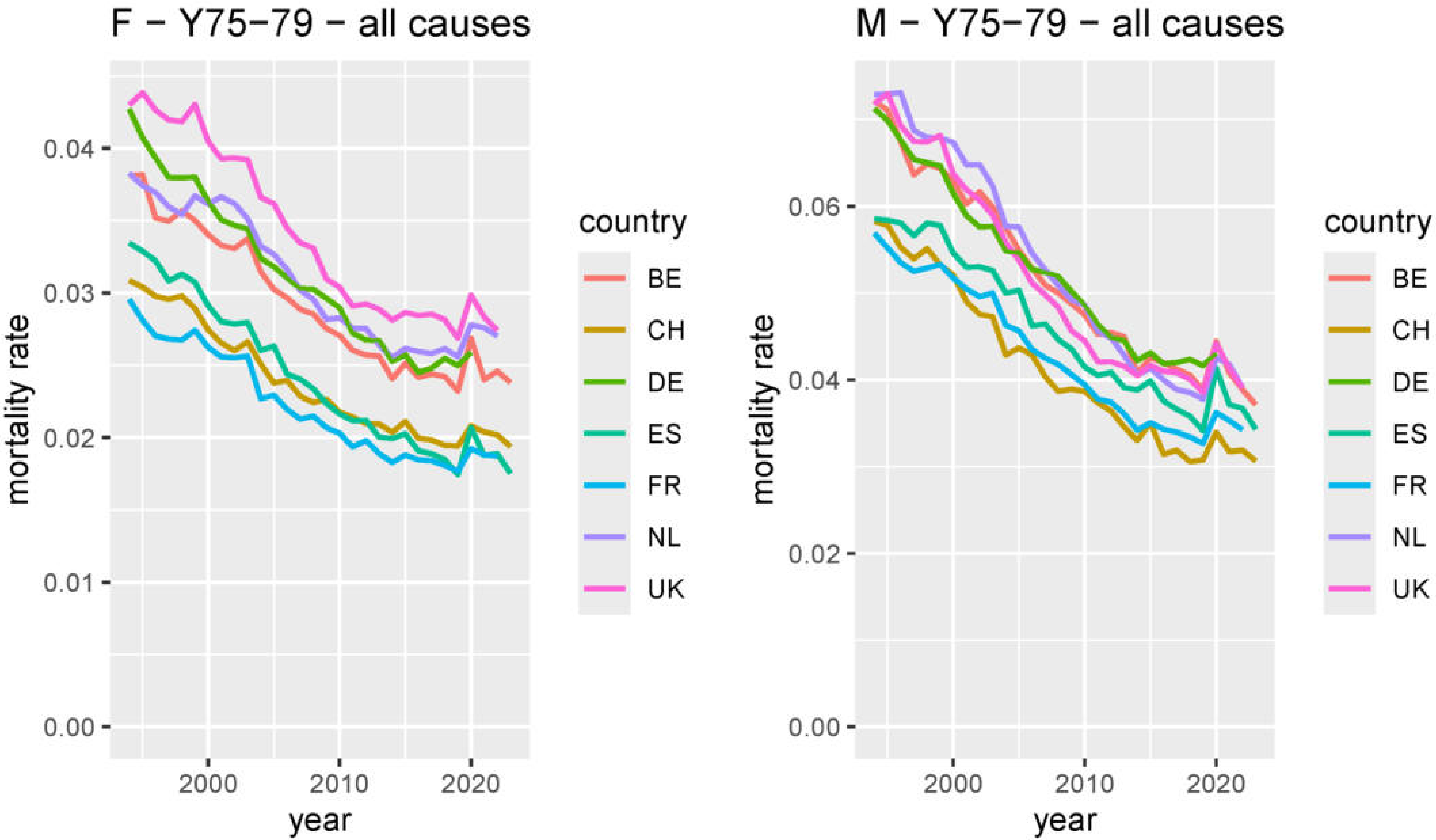
  for example, $D_{xt} \sim \text{Poisson}(\mu_{xt} E_{xt})$
- For observed mortality we use the same notation:

$$\mu_{xt} = \frac{D_{xt}}{E_{xt}}$$

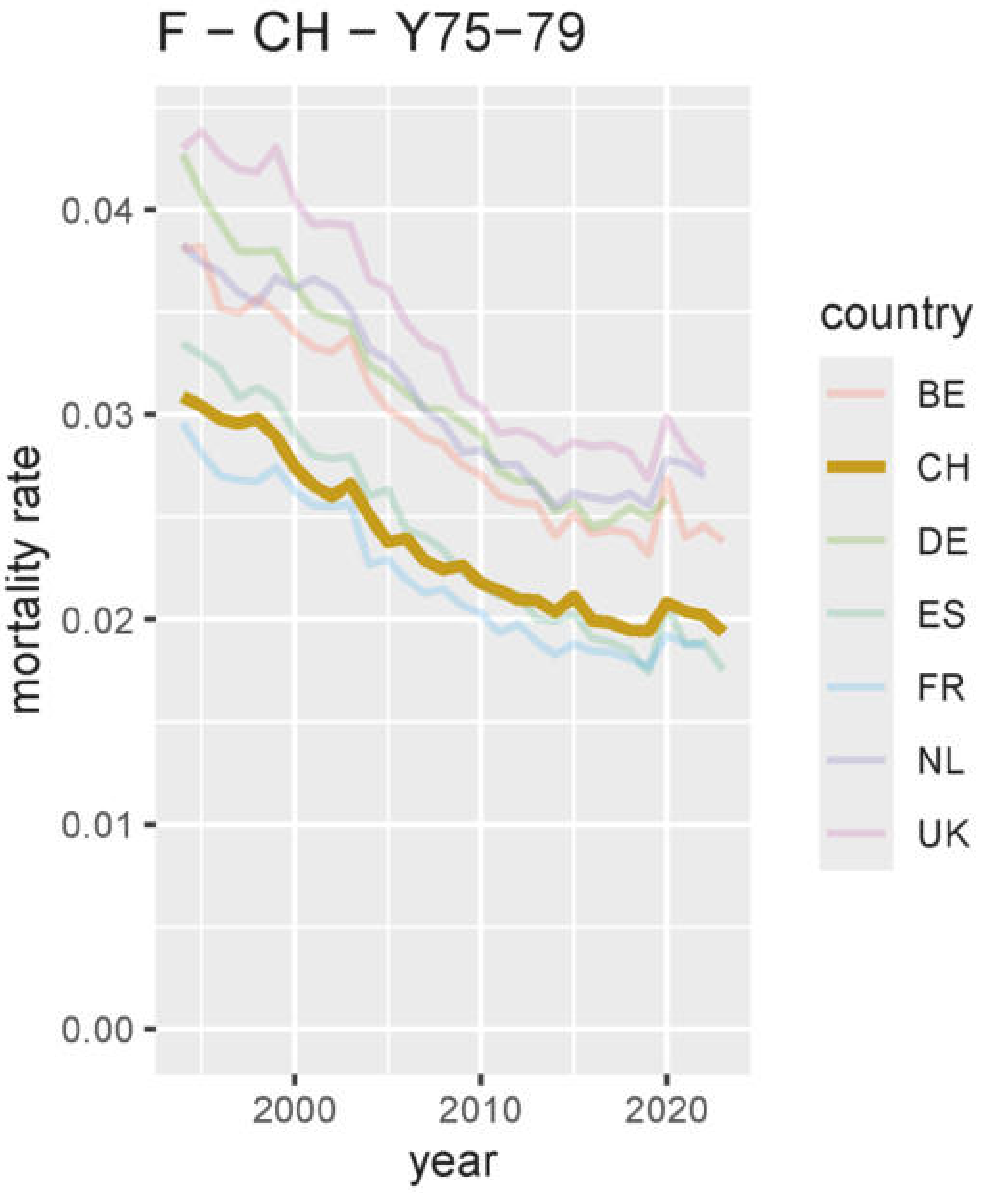

Women in Europe − Y75−79 − all causes
mortality rate
0.00
0.01
0.02
0.03
0.04
year
2000
2010
2020
country
BE
CH
DE
ES
FR
NL
UK

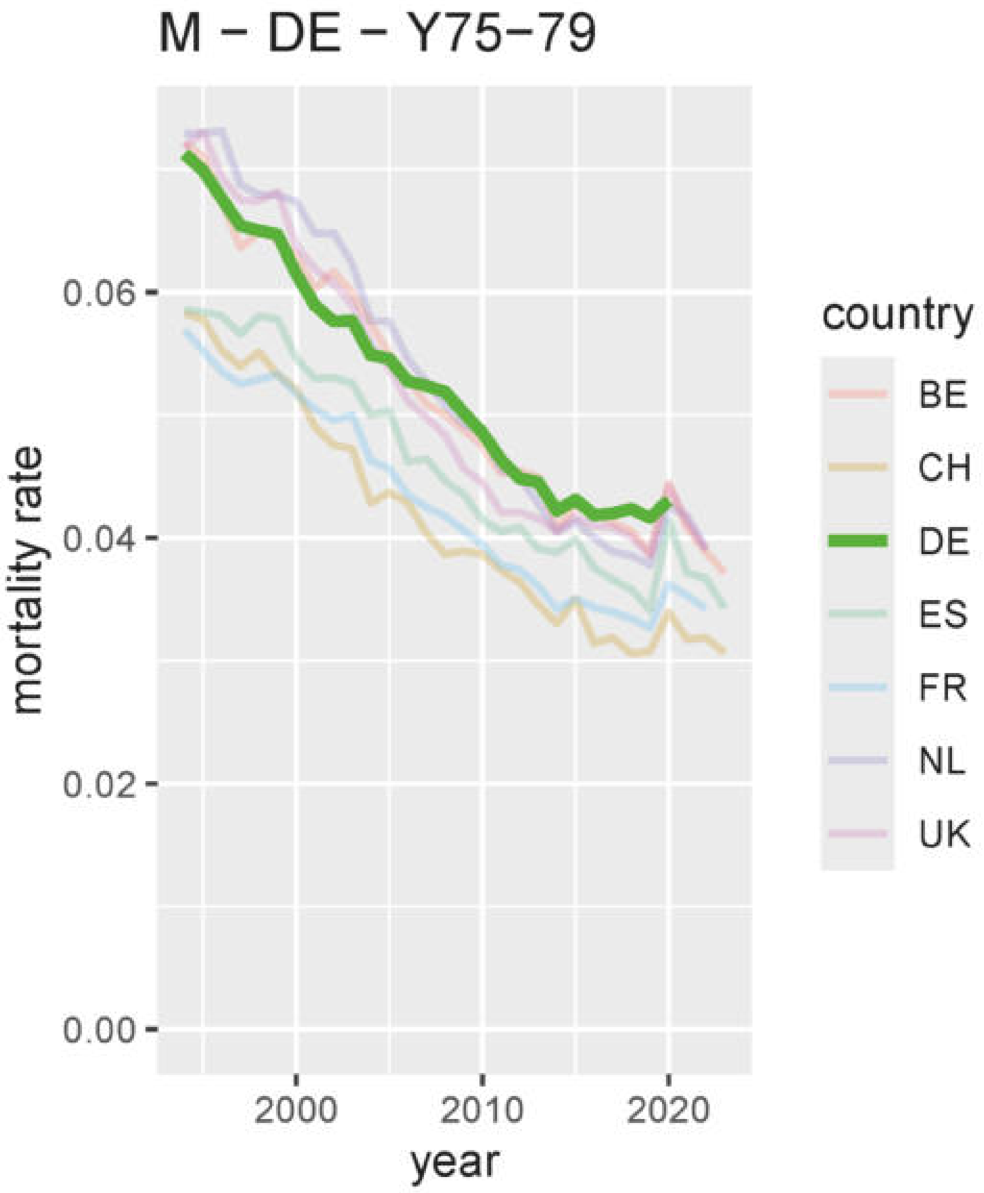

Men in Europe − Y75−79 − all causes
mortality rate
0.00
0.02
0.04
0.06
year
2000
2010
2020
country
BE
CH
DE
ES
FR
NL
UK

Mortality rates in Europe

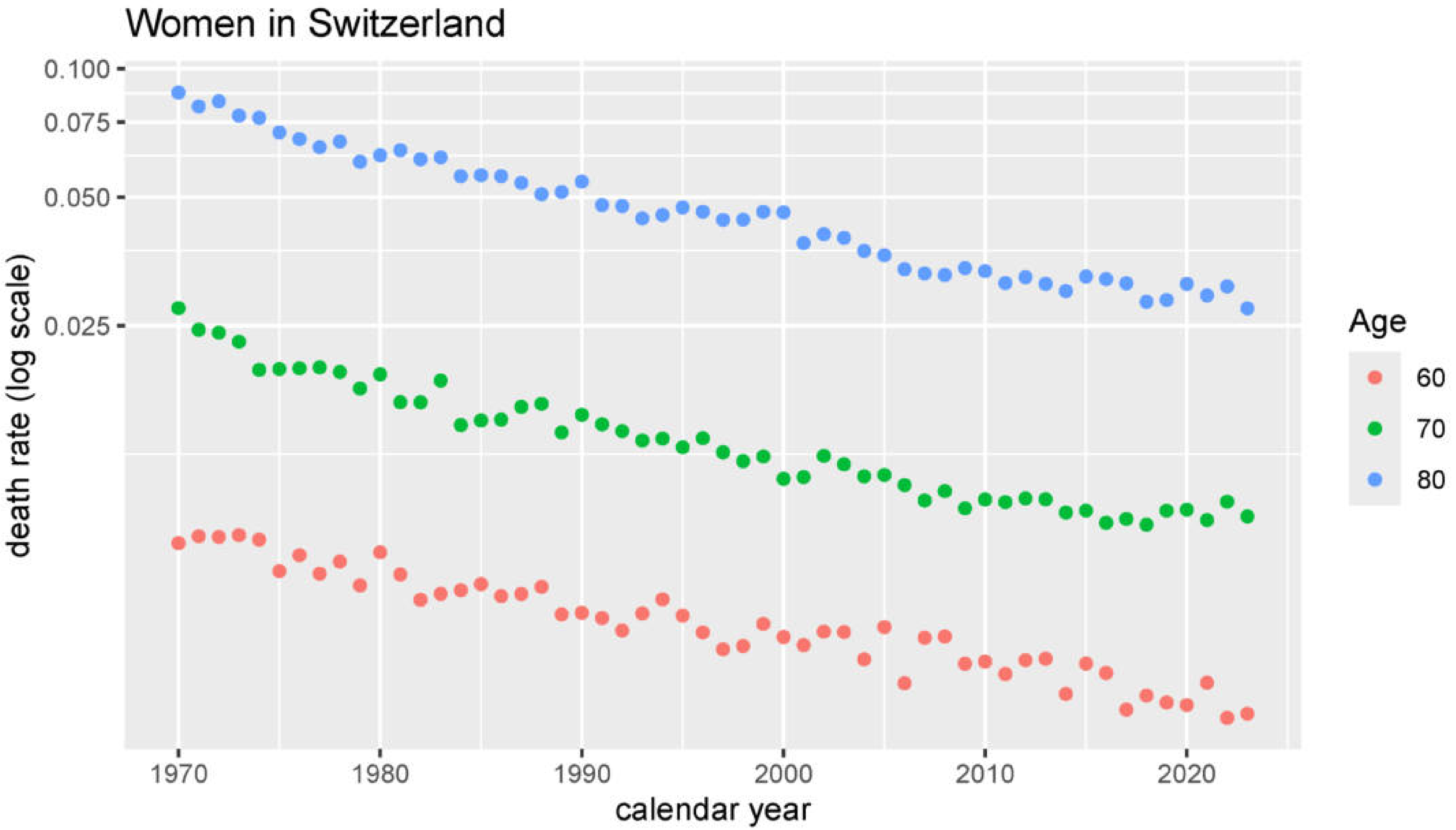


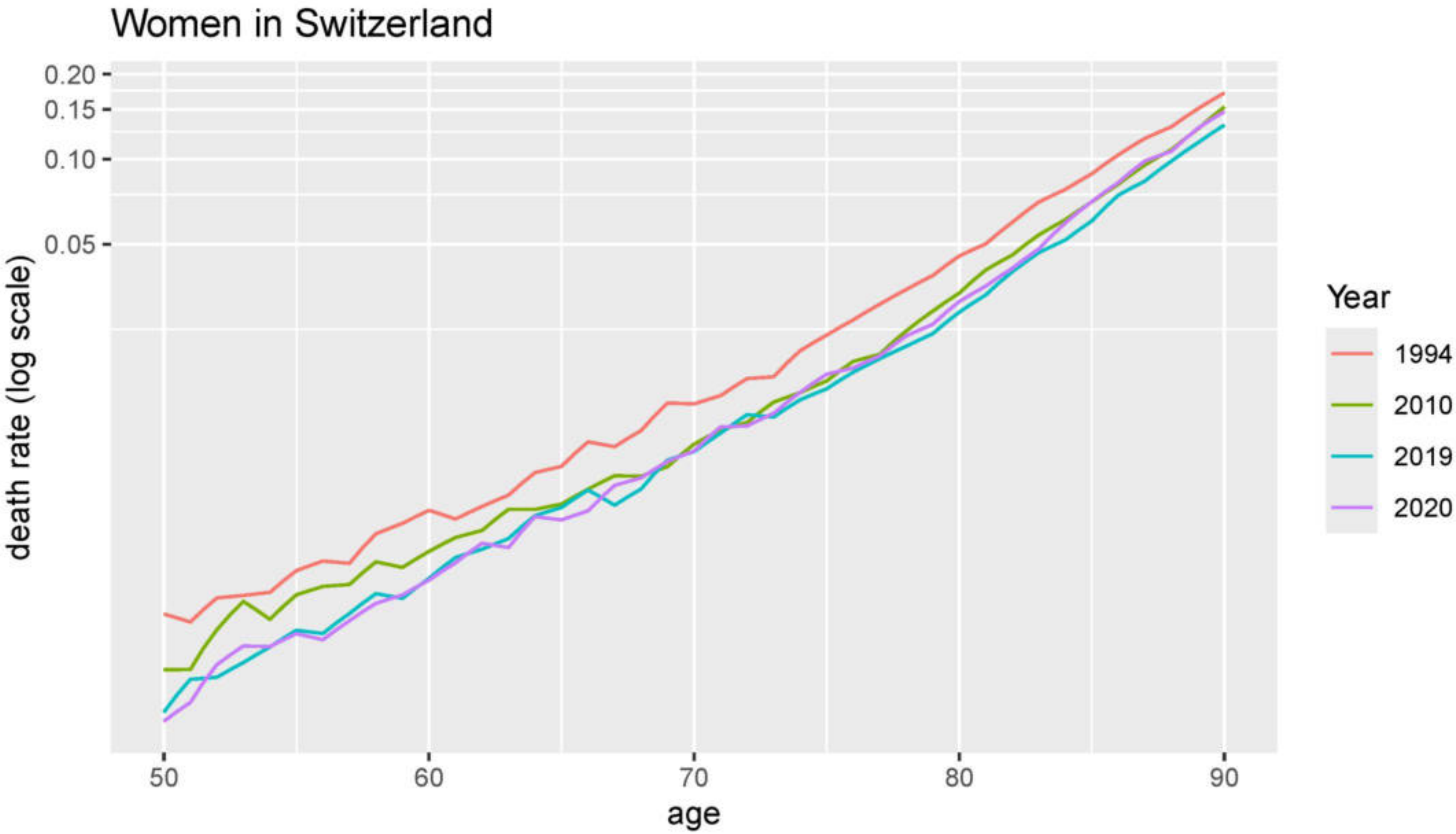
Mortality rates in Europe

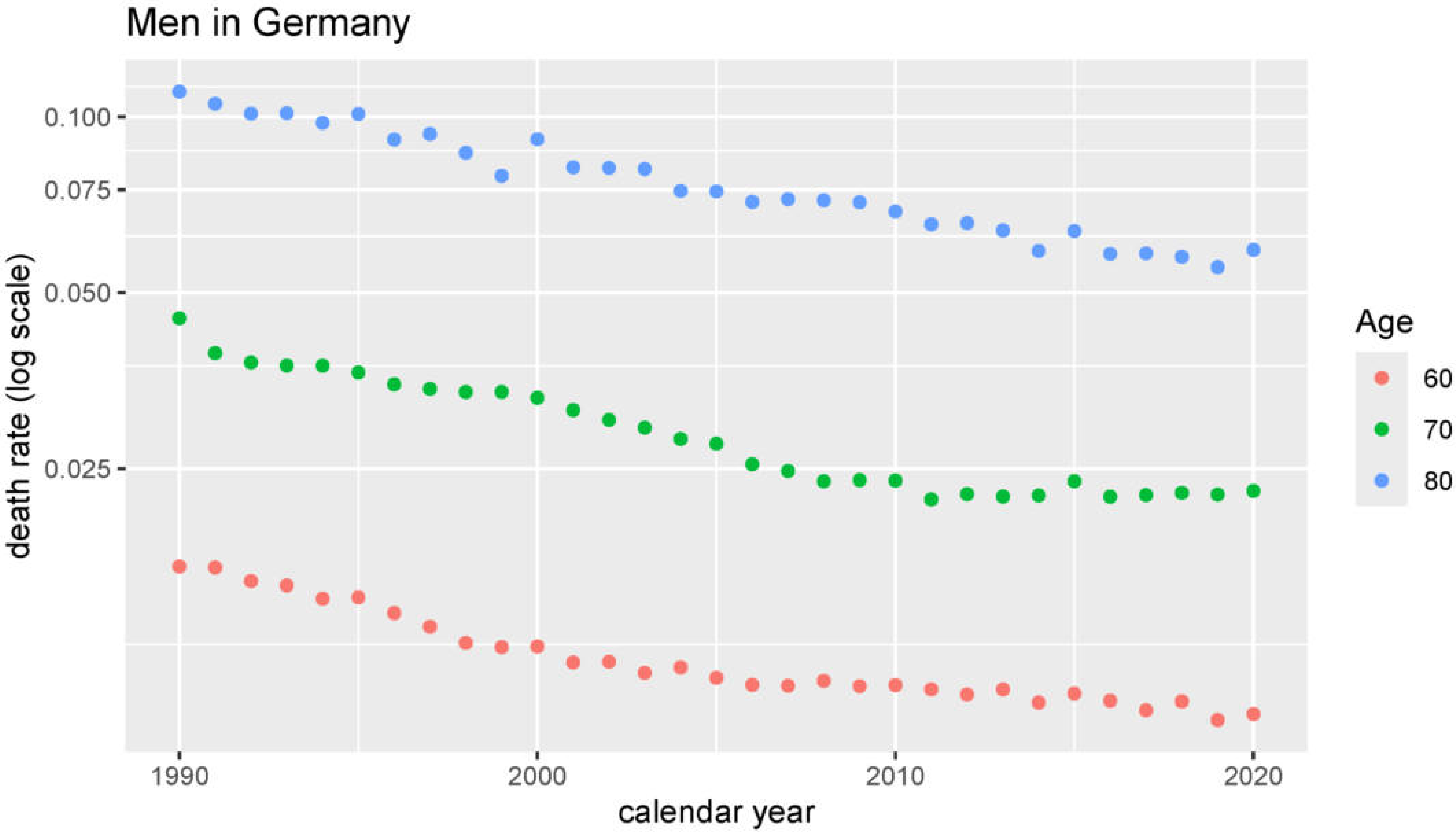




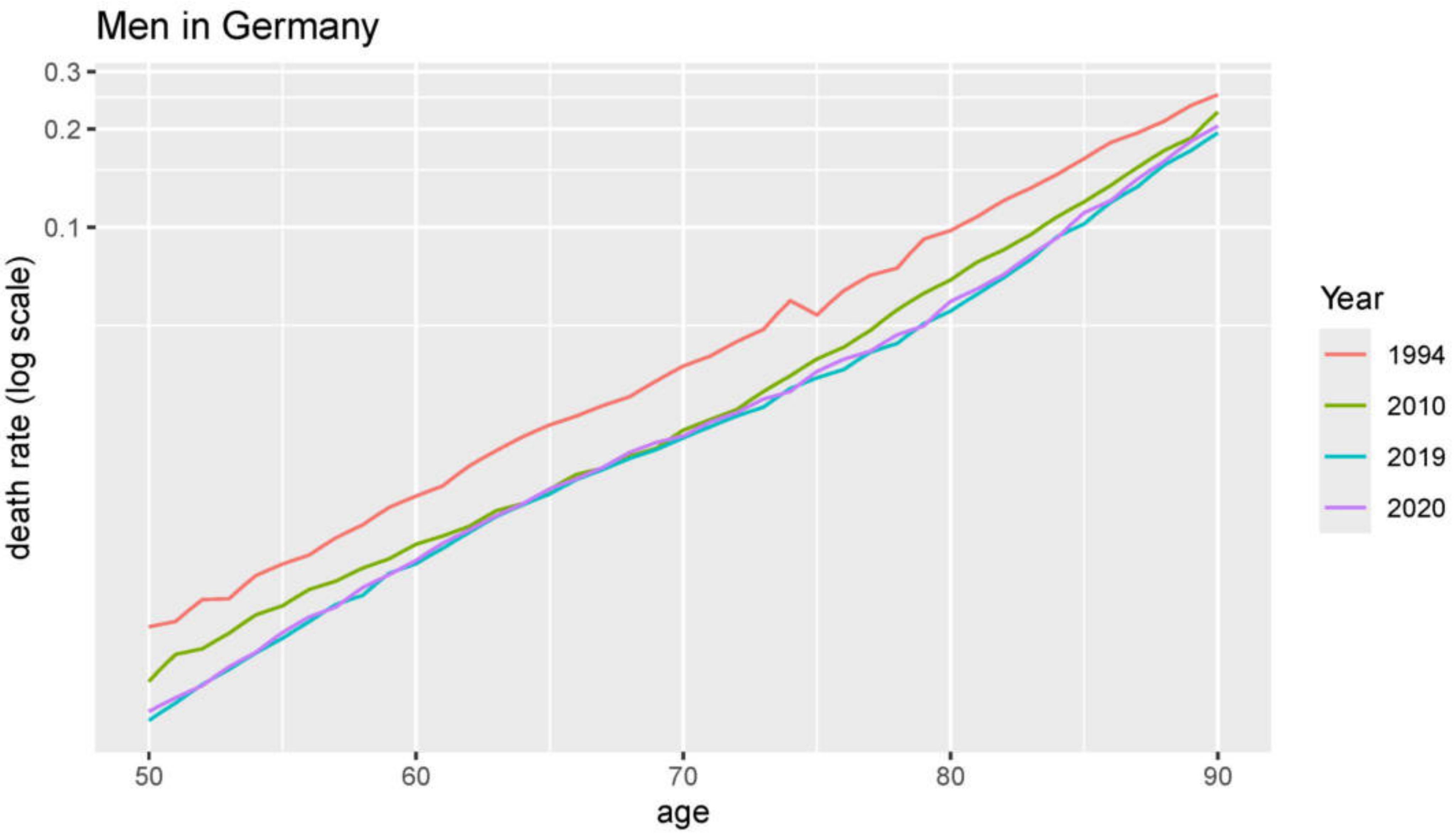
Women in Switzerland




Women in Switzerland
0.20
0.15
0.10
0.05
death rate (log scale)
50
60
70
80
90
age
Year
1994
2010
2019
2020


Men in Germany
death rate (log scale)
0.100
0.075
0.050
0.025
1990
2000
2010
2020
calendar year
Age
60
70
80

Men in Germany



Modelling - Single Population

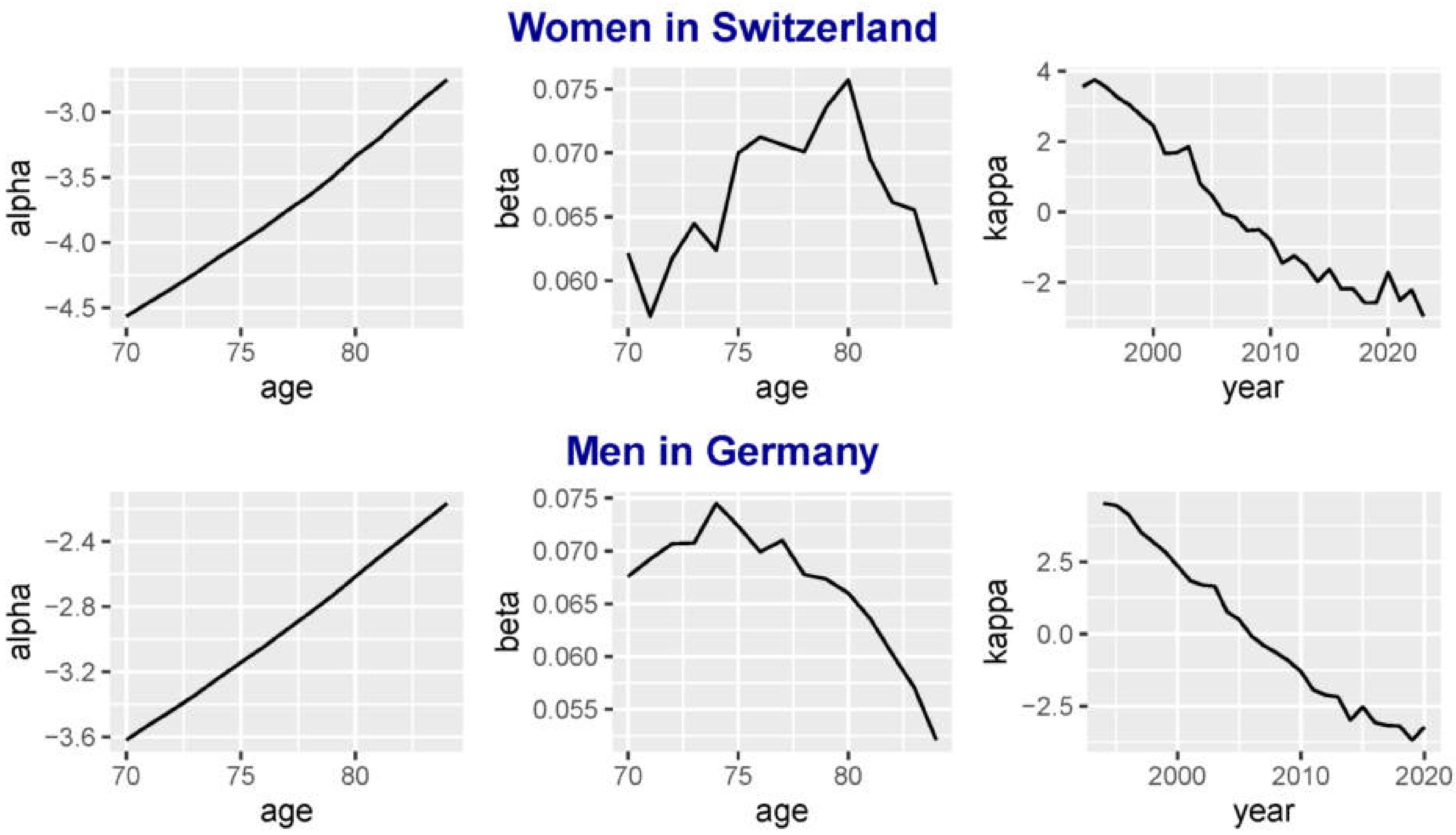
For age x and year t:

$$\begin{aligned} D_{xt} &\sim \text{Poisson}\left(\mu_{xt} E_{xt}\right) \\ \mu_{xt} &= \exp\left(\alpha_x + \beta_x \kappa_t\right) \end{aligned}$$

And a time series model for the period effects κ.
Extensions are often applied:

$$\mu_{xt} = \exp\left(\alpha_x + \beta_x \kappa_t + \ldots\right)$$

where the extra terms are related to covariates (socio-economics), externally given terms (mortality in Europe), temporary period effects (pandemic), cohort effects, extra age-period effects, etc.,


Women in Switzerland
alpha
-3.0
-3.5
-4.0
-4.5
70
75
80
age
beta
0.075
0.070
0.065
0.060
70
75
80
age
kappa
4
2
0
-2
2000
2010
2020
year
Men in Germany
alpha
-2.4
-2.8
-3.2
-3.6
70
75
80
age
beta
0.075
0.070
0.065
0.060
0.055
70
75
80
age
kappa
2.5
0.0
-2.5
2000
2010
2020
year

For each cause i, age (group) x and year t we define

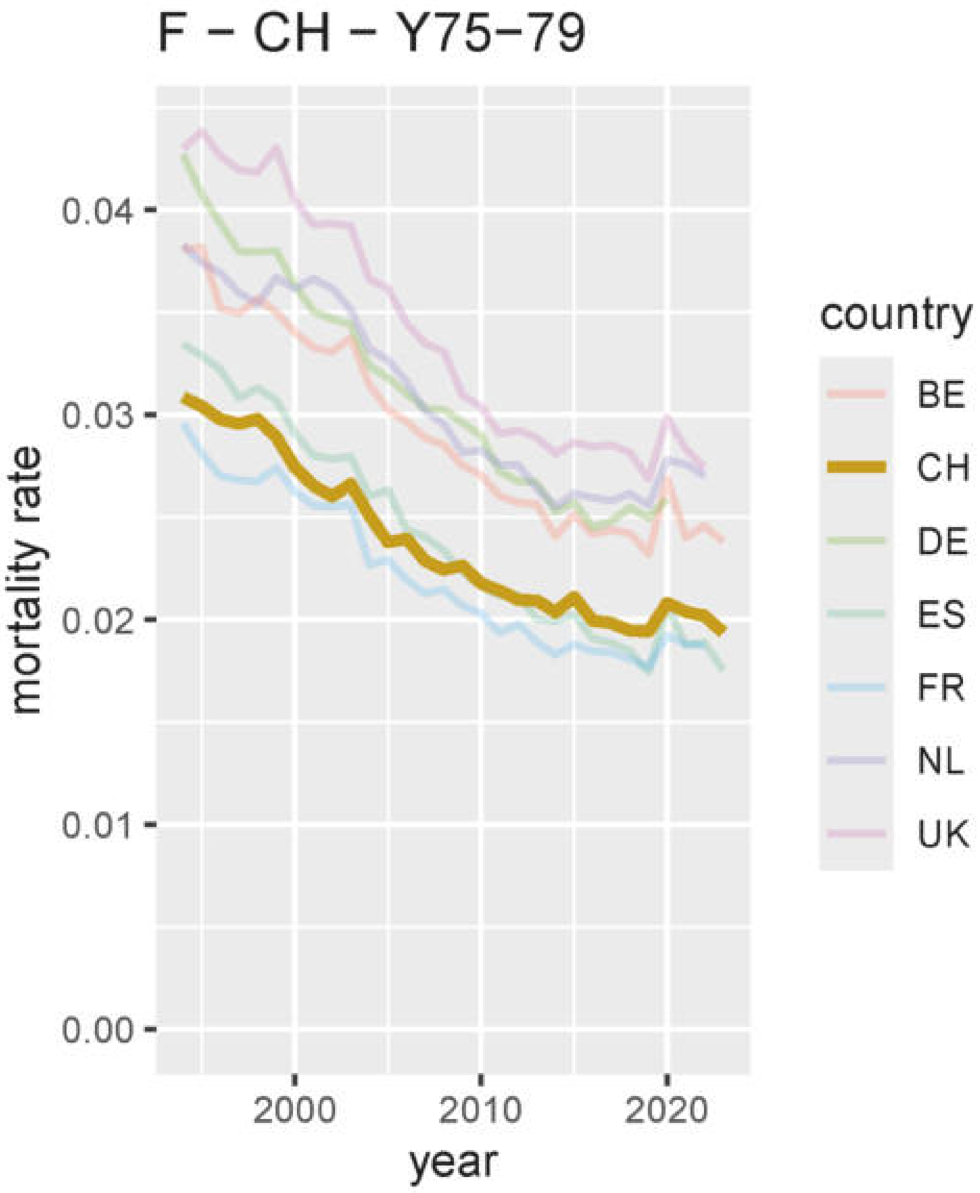
- Observed death $d_{xt}^{(i)}$ are realisations of RV $D_{xt}^{(i)}$
- Central Exposure to risk E_{xt}: mid-year population estimates
- For a given force of mortality $\mu_{xt}^{(i)}$ we assume

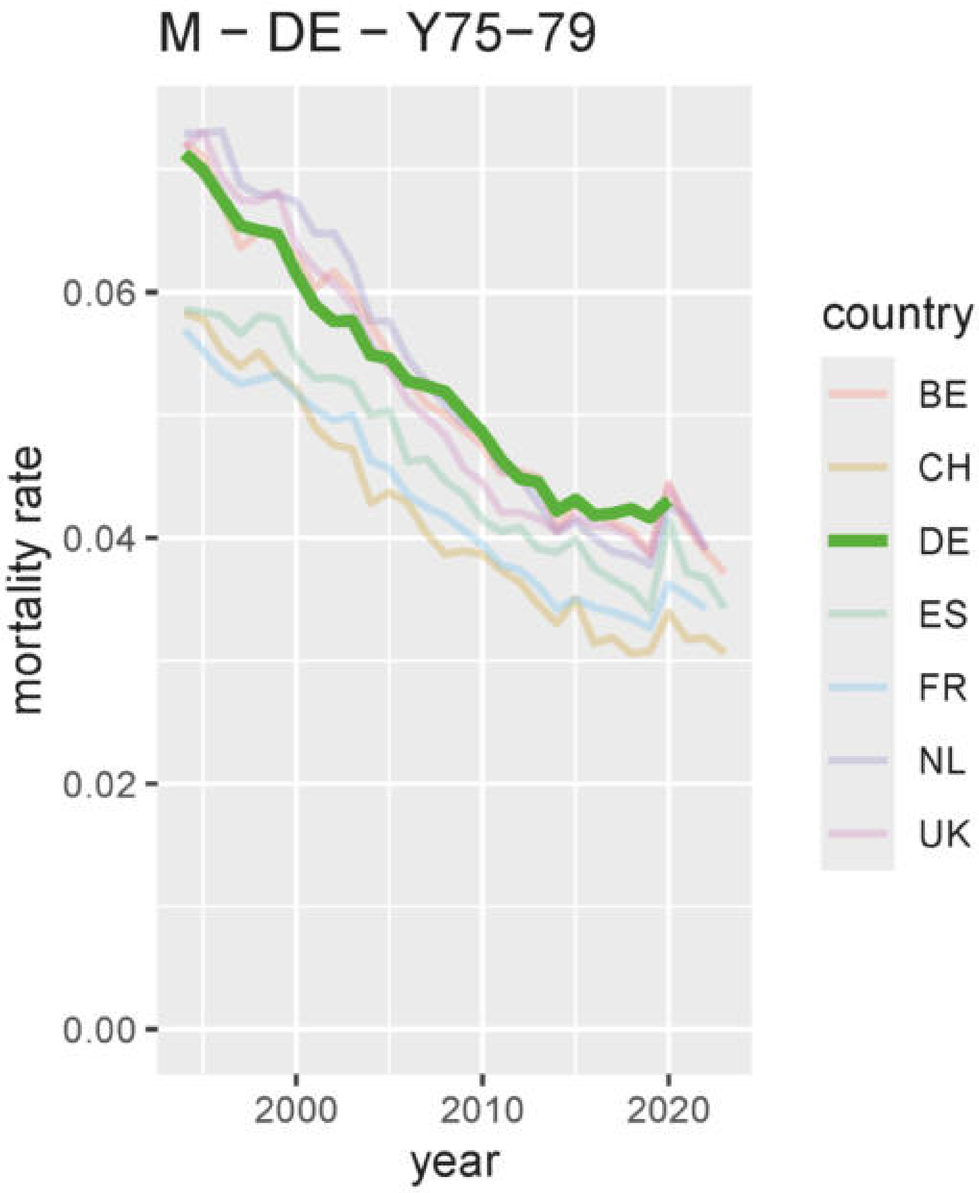
$$\mathbb{E} D_{xt}^{(i)} = \mu_{xt}^{(i)} E_{xt}$$

  for example, $D_{xt}^{(i)} \sim \text{Poisson}\left(\mu_{xt}^{(i)} E_{xt}\right)$
- Observed mortality for cause i:

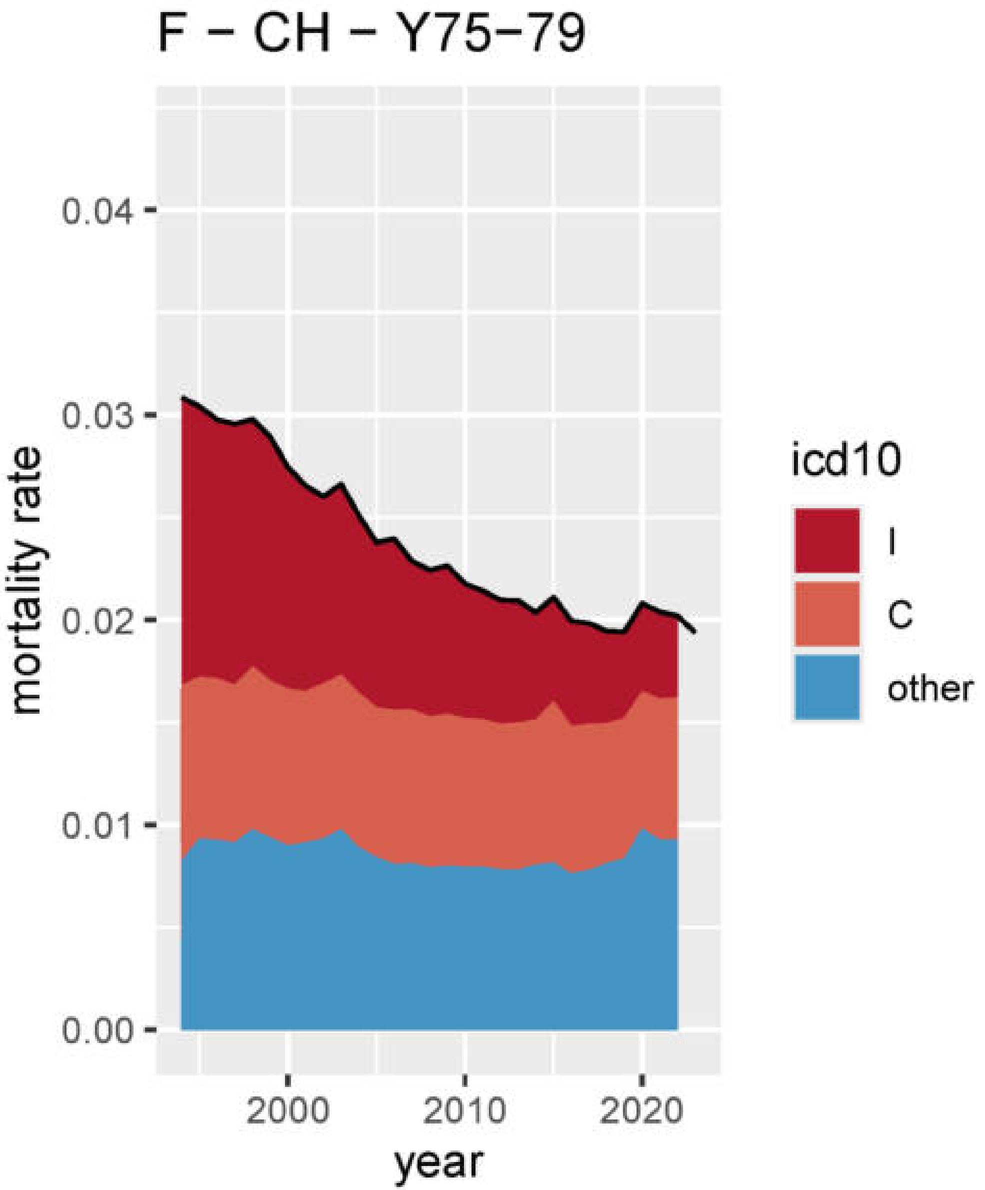
$$\hat{\mu}_{xt}^{(i)} = \frac{D_{xt}^{(i)}}{E_{xt}}$$

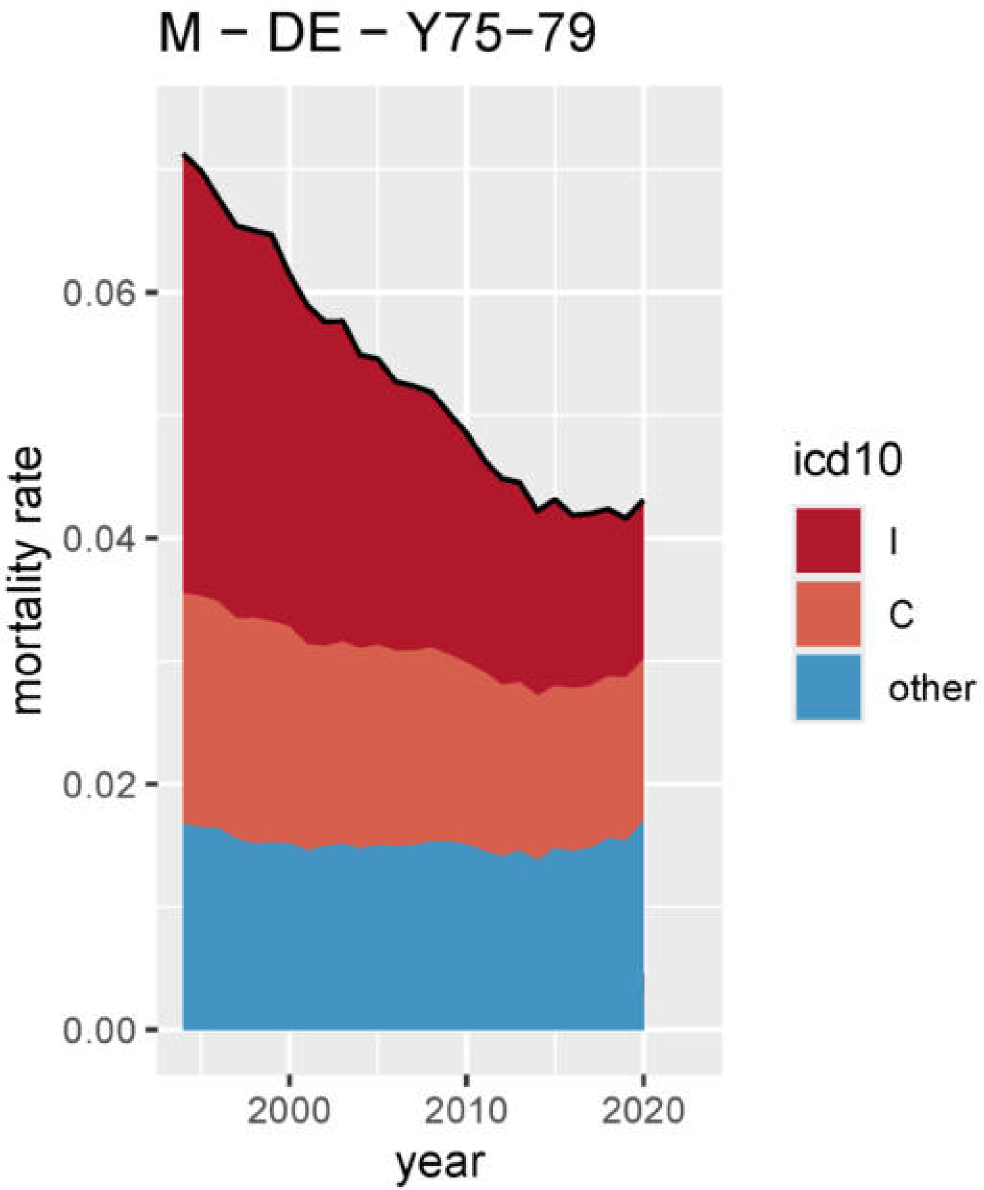
Mortality rates in Europe



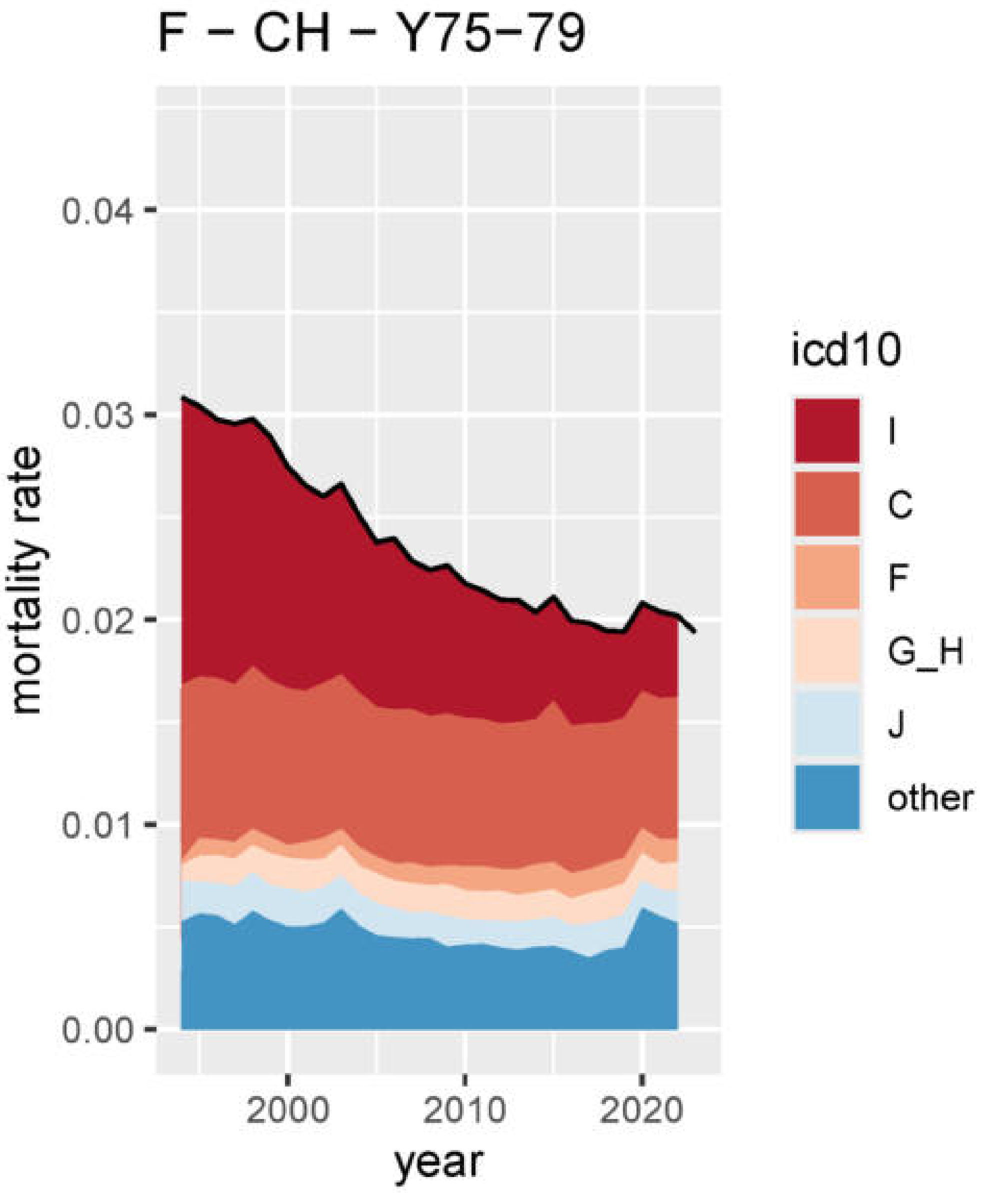


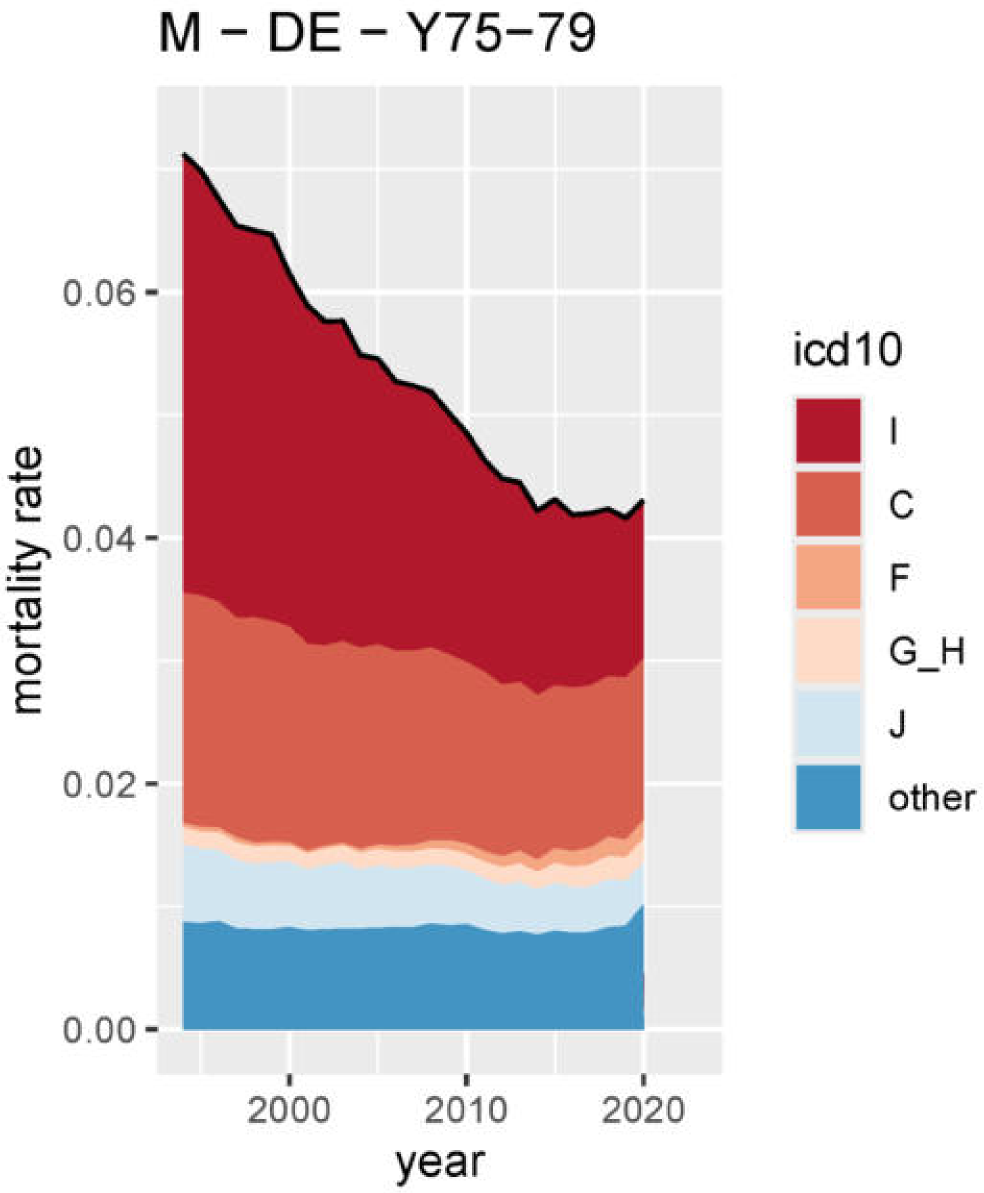
Mortality rates by Cause





Mortality rates by Cause





Some Definitions

- Set of causes that people have died of during year t: $\mathcal{I}_t$.
- Deaths from all causes: $D_{xt} = \sum_{i \in \mathcal{I}_t} D_{xt}^{(i)}$
- All-cause mortality

$$\hat{\mu}_{xt} = \sum_{i \in \mathcal{I}_t} \hat{\mu}_{xt}^{(i)}$$

- Relative importance of cause i:

$$w_{xt}^{(i)} = \frac{D_{xt}^{(i)}}{D_{xt}}$$

Observed mortality for cause i: $\mu_t^{(i)} = \frac{D_t^{(i)}}{E_t}$

Deaths from all causes: $D_t = \sum_i D_t^{(i)}$

Set of causes that people have died of during year t: $\mathcal{C}_t$. We assume $\mathcal{C}_t \subset \mathcal{C}_{t+1}$.

$$\mu_{t+1} = \sum_{i \in \mathcal{C}_{t+1}} \mu_{t+1}^{(i)} = \sum_{i \in \mathcal{C}_t} \mu_{t+1}^{(i)} + \sum_{k \notin \mathcal{C}_t} \mu_{t+1}^{(k)}$$

Observed mortality for cause i: $\mu_t^{(i)} = \frac{D_t^{(i)}}{E_t}$
Deaths from all causes: $D_t = \sum_i D_t^{(i)}$
Set of causes that people have died of during year t: $\mathcal{C}_t$. We assume $\mathcal{C}_t \subset \mathcal{C}_{t+1}$.

$$\mu_{t+1} = \sum_{i \in \mathcal{C}_{t+1}} \mu_{t+1}^{(i)} = \sum_{i \in \mathcal{C}_t} \mu_{t+1}^{(i)} + \sum_{k \notin \mathcal{C}_t} \mu_{t+1}^{(k)}$$

All-cause improvement factor:

$$\begin{aligned} \frac{\mu_{t+1}}{\mu_t} &= \frac{\sum_{i \in \mathcal{C}_t} \mu_{t+1}^{(i)} + \sum_{k \notin \mathcal{C}_t} \mu_{t+1}^{(k)}}{\mu_t} \\ &= \sum_{i \in \mathcal{C}_t} w_t^{(i)} \frac{\mu_{t+1}^{(i)}}{\mu_t^{(i)}} + \frac{1}{\mu_t} \sum_{k \notin \mathcal{C}_t} \mu_{t+1}^{(k)}, \qquad w_t^{(i)} = \frac{D_t^{(i)}}{D_t} \end{aligned}$$

Decomposition of improvement rates $\rho_{t+1}^{(i)} = (\mu_t^{(i)} - \mu_{t+1}^{(i)})/\mu_t^{(i)}$:

$$\rho_{t+1} = \sum_{i \in \mathcal{C}_t} w_t^{(i)} \rho_{t+1}^{(i)} - \frac{1}{\mu_t} \sum_{k \notin \mathcal{C}_t} \mu_{t+1}^{(k)} \quad \text{with} \quad w_t^{(i)} = \frac{D_t^{(i)}}{D_t}$$

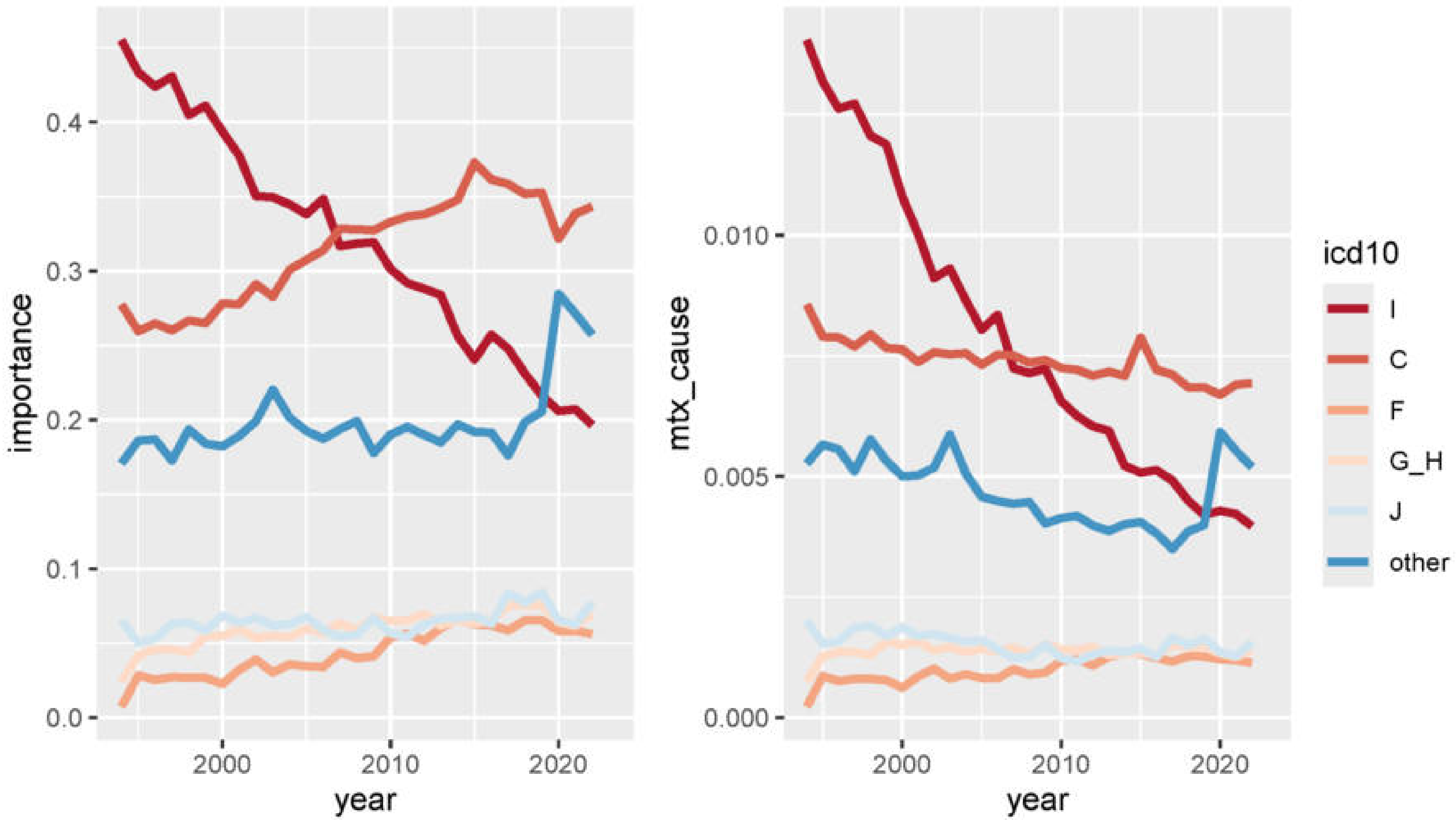
Decomposition of improvement rates $\rho_{t+1}^{(i)} = \big(\mu_t^{(i)} - \mu_{t+1}^{(i)}\big)/\mu_t^{(i)}$:

$$\rho_{t+1} = \sum_{i \in \mathcal{C}_t} w_t^{(i)} \rho_{t+1}^{(i)} - \frac{1}{\mu_t} \sum_{k \notin \mathcal{C}_t} \mu_{t+1}^{(k)} \quad \text{with} \quad w_t^{(i)} = \frac{D_t^{(i)}}{D_t}$$

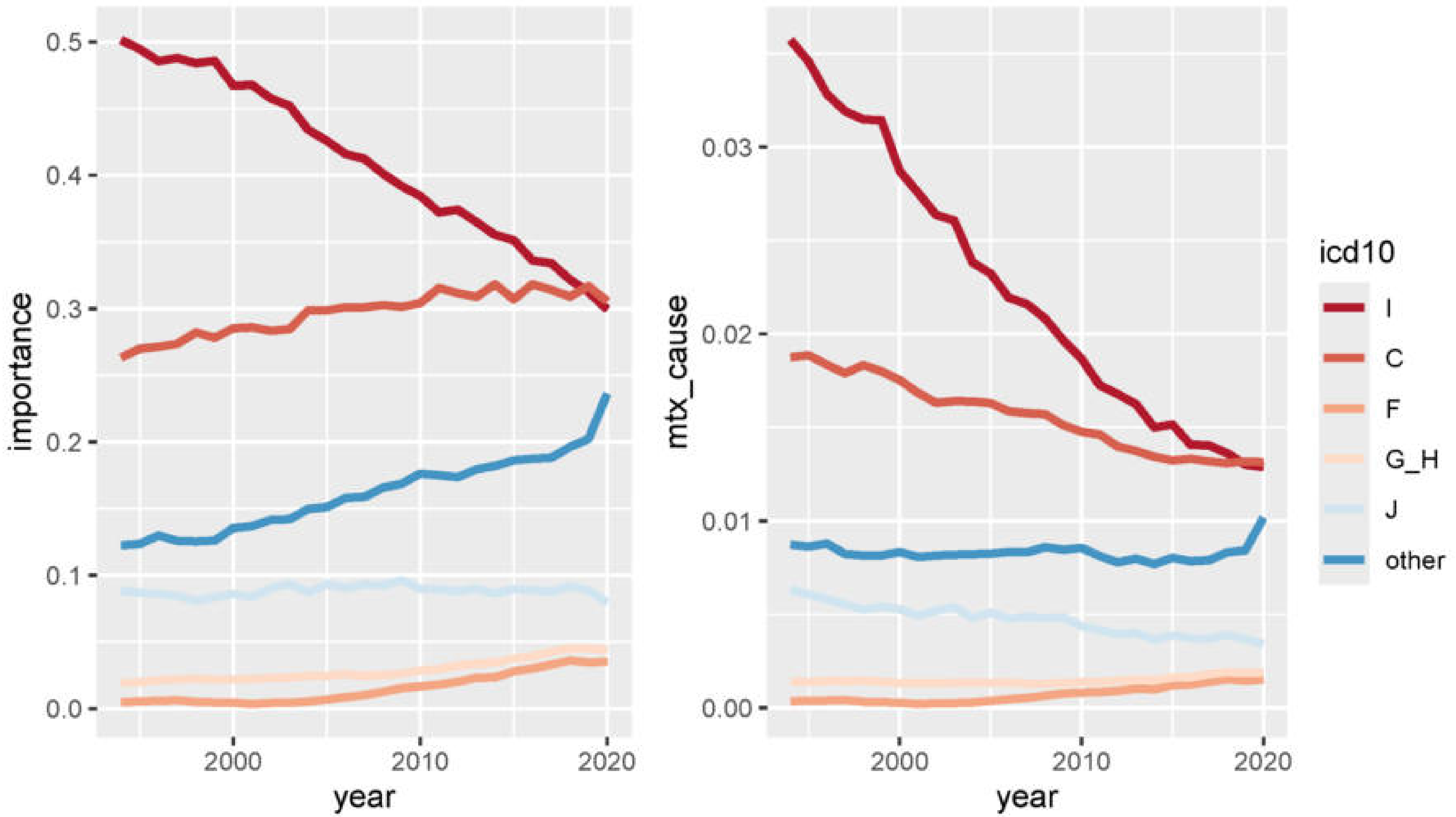
For example, $t = 2019$:

$$\rho_{2020} = \sum_{i \in \mathcal{C}_{2019}} w_{2019}^{(i)} \rho_{2020}^{(i)} - \frac{\mu_{2020}^{\text{Covid}}}{\mu_{2019}}$$

Causes of Death - Relative Importance - Swiss Women



Percentage of all deaths due to certain causes and cause-specific mortality rates



Percentage of all deaths due to certain causes and cause-specific mortality rates

Classification of Causes and Data Sources

"the disease or injury which initiated the train of morbid events leading directly to death, or the circumstances of the accident or violence which produced the fatal injury"

World Health Organization

International Classification of Diseases - ICD

- International Statistical Classification of Diseases and Related Health Problems
- Created and maintained by the WHO - "allows the systematic recording, analysis, interpretation and comparison of mortality and morbidity data collected in different countries or regions and at different times"
- Under the auspices of WHO since the 6th Revision adopted in 1948
- 11th edition in effect from January 2022
- Often ICD-10 is used - in effect from 1993 - latest version 2019 with emergency codes for Covid-19 added in 2020

Classification and Data Sources

- ICD-10 Version 2019
- Alternative classification: Global Burden of Disease
- WHO database for cause-specific death counts

- Cause specific death counts are available from Eurostat, and some national statistics offices
- Eurostat uses ICD-10 codes
- data browser at Eurostat

- Cause specific death counts are also available the Human Mortality Database

Causes are usually grouped:

- by ICD-10 chapter
- Global burden of disease study distinguishes: Communicable diseases, Non-communicable diseases, Injuries
- according to life-style related causes (lung cancer) and causes with unclear risk factors (prostate cancer)
- grouping cannot be too fine to allow for a sufficiently high number of deaths per group

Conclusions - Cause of Death

Conclusions

Reasons to use cause-specific mortality data

- good fit of standard models for single and multiple populations
- provides a more detailed explanation of observed trends
- Covid-19 shock can be "filtered out"
- allows for a more detailed description of projected mortality rates and life expectancy, and the quantification of uncertainty about projected rates
- Expert judgement more easily incorporated than in all-cause models
- Projections based on plausible cause-specific scenarios are easily generated.

Conclusions

Reasons to use cause-specific mortality data

- good fit of standard models for single and multiple populations
- provides a more detailed explanation of observed trends
- Covid-19 shock can be "filtered out"
- allows for a more detailed description of projected mortality rates and life expectancy, and the quantification of uncertainty about projected rates
- Expert judgement more easily incorporated than in all-cause models
- Projections based on plausible cause-specific scenarios are easily generated.

Challenges

- Main challenge is the choice of appropriate time series models for period effects per population and cause.
- Careful consideration of dependencies between rates for causes and populations.

Empirical observations

- the reduced importance of mortality from diseases of the circulatory system leads to reduced mortality improvements
- Diseases of the nervous system and mental diseases are becoming increasingly more common causes of death, and developments in those will have a strong impact on overall mortality rates
- Improvement rates for those "new" causes and cancer will need to be at least as strong as those observed for deaths from diseases of the circulatory system for further strong all-cause improvemnts

Mortality by Socio-Economic Group (data from England)

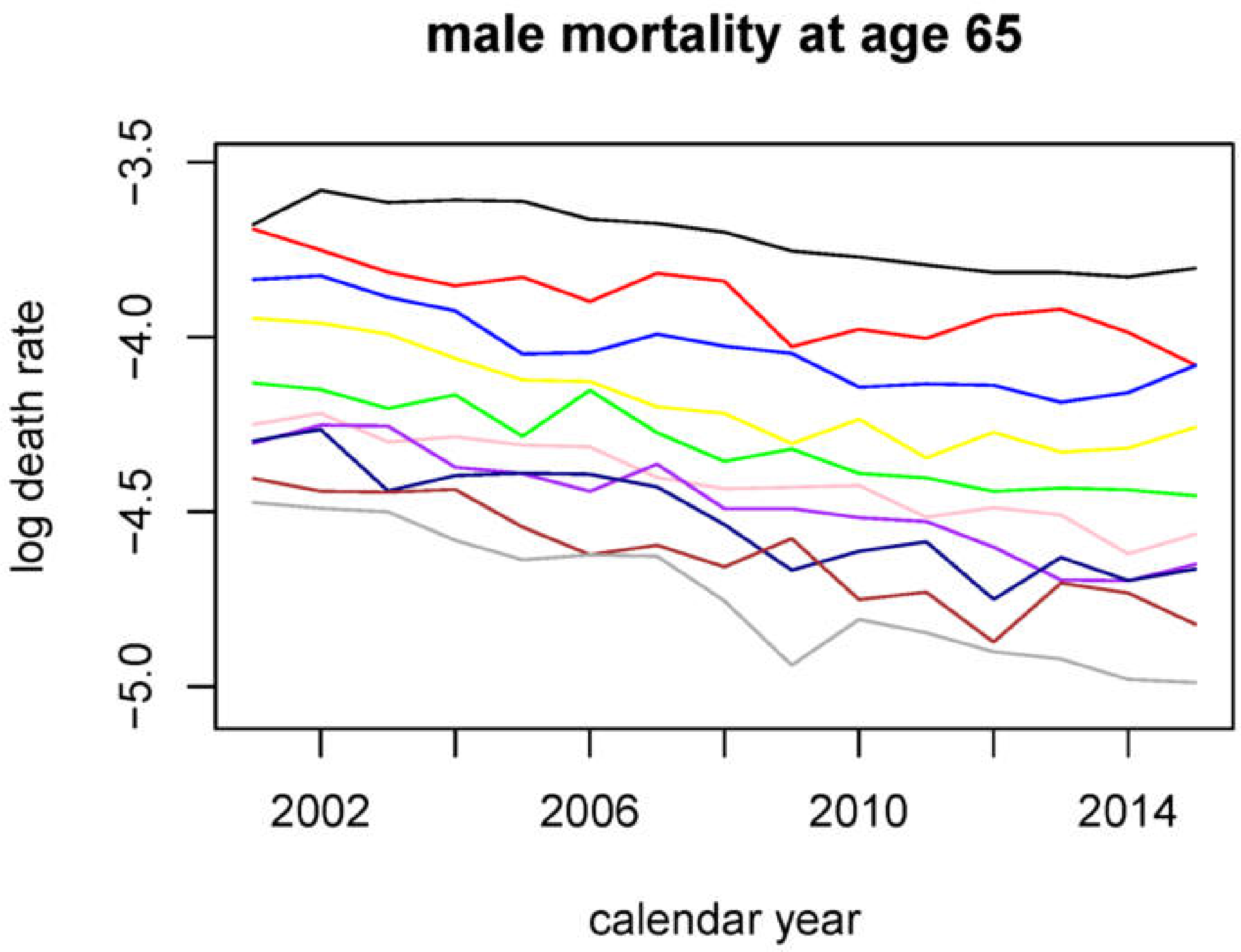
- Ten sub-populations in England based on IMD deciles in 2015 (index of multiple deprivation)
- no individual, neighbourhood or geographical characteristics
- similar data are more easily available for other countries than individual records or neighbourhood information
- Based on: Wen J, Cairns AJG, Kleinow T. Fitting multi-population mortality models to socio-economic groups. Annals of Actuarial Science. 2021;15(1)

Number of deaths per 1,000 lives by IMD decile

	Deaths per 1,000 lives		
	most deprived	least deprived	ratio
2001	25.3	11.4	2.219
2005	27.0	9.7	2.784
2010	23.0	8.2	2.805
2015	22.3	6.8	3.279

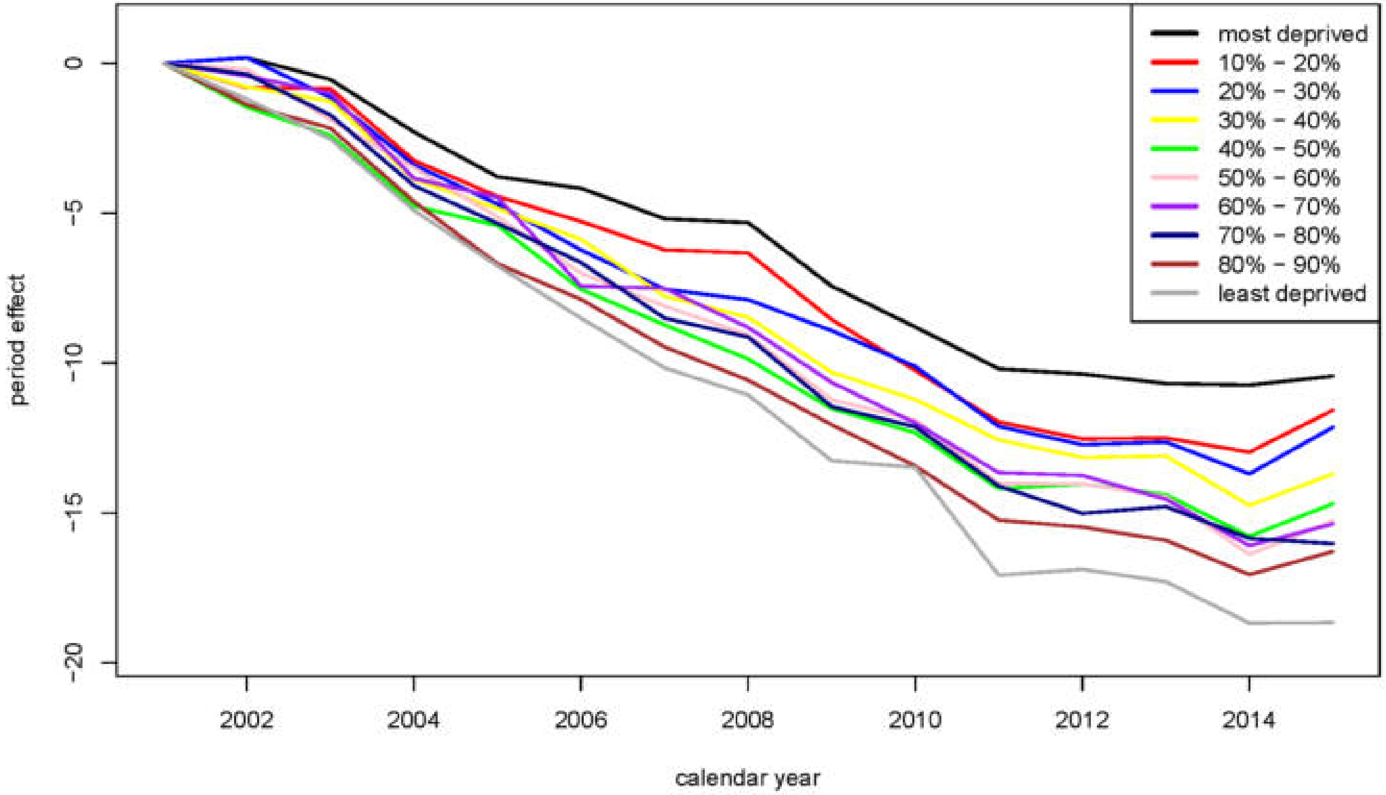
Males aged 65

Mortality rates by IMD decile for males aged 65



- downward trend strongest for least deprived

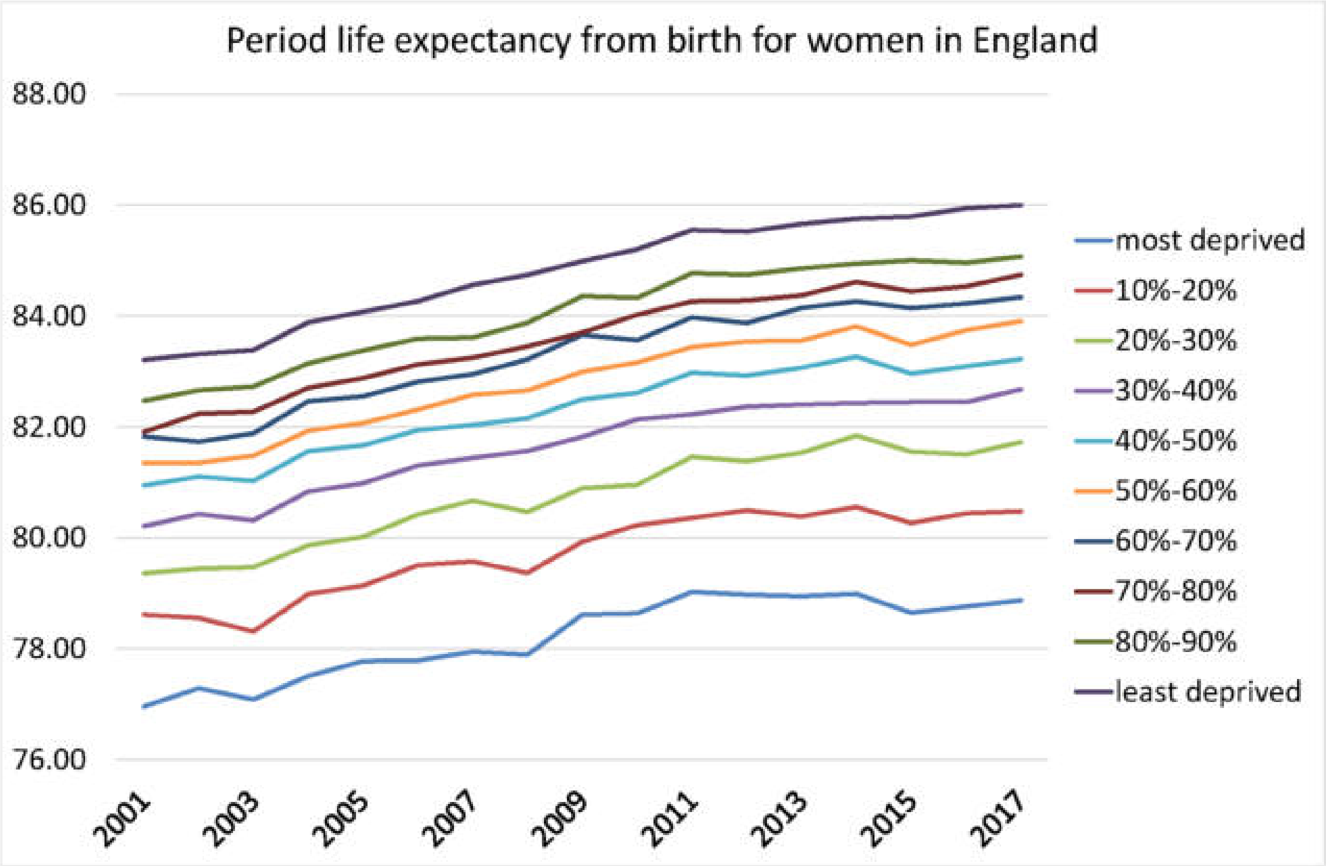
Lee-Carter model fitted to individual IMD decile



- downward trend strongest for least deprived
- no improvements for most deprived since 2011
- slowdown of improvements for least deprived since 2011

- mortality improvements have been smaller for the most deprived groups leading to a widening gap between the levels of mortality in different socio-economic groups
- This conclusion must be treated with caution since we only considered data based on the 2015 deciles
- Neighbourhoods migrate between deciles over time - some get more deprived, others get less deprived

Life Expectancy Compared to Germany



In Germany: "Bei der Lebenserwartung ab Geburt beträgt die Differenz zwischen der niedrigsten und höchsten Einkommensgruppe für Frauen 4,4 Jahre und für Männer 8,6 Jahre." (Robert Koch-Instituts)

More detailed approach: Modelling Socio-Economic Mortality at Neighbourhood Level

Mortality at Neighbourhood Level - Introduction

Based on Wen J, Cairns AJG, Kleinow T. Modelling socio-economic mortality at neighbourhood level. ASTIN Bulletin. 2023;53(2):285-310. doi:10.1017/asb.2023.12

- strong association between mortality and socio-economic status
- we develop a socio-economic mortality index that we call the Longevity Index for England - LIFE
- the LIFE index models the relative risk of dying within a small population in England as a function of certain socio-economic variables
- example for a location/postcode based relative risk index
- alternative to IMD and other postcode measures of relative risk
- The index is created based on a regression analysis using a Random Forest algorithm to estimate a non-parametric regression function.

Neighbourhoods

- Small homogeneous areas
- We use data for Lower Layer Super Output areas (LSOA)
- An LSOA is a small area in England with a population of about 1,500 people.
- For the 2015 indices of deprivation, there were 32,844 LSOAs in England
- The existence of care homes has the potential to artificially increase the mortality in an LSOA - this has to be taken into account

- Estimating the effect of socio-economic factors on the deviation of LSOA-specific mortality from national mortality is a regression problem.
- We use a random forest algorithm to estimate the regression function
- other non-parametric estimators can be used
- parametric models like GLMs could also be used, but they impose a structure on the relationship between s/e variables and relative mortality risk - non-linear effects and joint effects will not be captured by GLMs

- mid-year population estimates (exposure size) E_{ita} by single LSOA $i = 1, \ldots, N$, year t and age a;
- death counts D_{ita} by single LSOA $i = 1, \ldots, N$, year t and age a;
- a vector of K predictive variables $X_i = (X_{i,1}, \ldots, X_{i,K})$ for each LSOA $i = 1, \ldots, N$.
  Not year or age specific
  describe socio-economic characteristics of the entire population of an LSOA measured at a specific point in time
- E_{ita} and D_{ita}, are available for calendar years 2001 to 2018 by single year of age
- we group ages for the construction of the mortality index: 60-69, 70-79 and 80-89

- baseline death rate m_{ta}^b for year t and age a for the whole of England:

$$m_{ta}^b = \frac{\sum_{i=1}^N D_{ita}}{\sum_{i=1}^N E_{ita}}. \qquad (1)$$

- Expected total number of deaths $\hat{D}_i^0$ across all ages $a \in \mathcal{A}$ and years $t \in \mathcal{T}$ in LSOA i is given by

$$\hat{D}_i^0 = \sum_{t \in \mathcal{T},\, a \in \mathcal{A}} m_{ta}^b E_{ita} \text{ for all } i = 1, \ldots, N,$$

- The observed relative risk of death R_i^0 for an individual living in LSOA i is the ratio of the actual number of deaths to the expected number of deaths in that LSOA:

$$R_i^0 = \frac{\sum_{t \in \mathcal{T},\, a \in \mathcal{A}} D_{ita}}{\hat{D}_i^0} \text{ for all } i = 1, \ldots, N. \tag{2}$$

- the relative risk R_i^0 is not age and year specific...
- ... but can, of course, be fitted to different age groups and observation periods

x_1	old age income deprivation
x_2	employment deprivation (i.e. unemployment)
x_3	proportion of the age-65+ population with no qualifications
x_4	crime rate
x_5	average number of bedrooms
x_6	proportion of the population born in the UK
x_7	wider barriers to housing (affordability, homelessness)
x_8	employment/occupation: proportion in a management position
x_9	proportion working more than 49h per week (ages 16-74)
x_{10}	urban-rural classification
x_{11}	proportion of population aged 60+ in a care home with nursing care
x_{12}	proportion of population aged 60+ in a care home without nursing care

- The socio-economic characteristics of any neighbourhood are given as a vector taking values in the $K = 12$ dimensional space

$$L_0 = \mathbb{R}^9 \times \{1, \ldots, 5\} \times [0, 1]^2. \tag{3}$$

- The socio-economic characteristics of any neighbourhood are given as a vector taking values in the $K = 12$ dimensional space

$$L_0 = \mathbb{R}^9 \times \{1, \ldots, 5\} \times [0,1]^2. \tag{3}$$

Urban/Rural (UR) class	**Definition**
1	Urban conurbation (except London)
2	Urban city and town
3	Rural town and village
4	Rural hamlet and isolated dwellings
5	Urban conurbation (in London)

- We model the conditional expectation of the relative mortality risk R^0:

$$f(x) := \mathbb{E}[R^0|x] \text{ for any } x \in L_0 \tag{4}$$

where $x = (x_1, \ldots, x_K)$ is the vector of predictive socio-economic variables.

- Care homes: in our sample of LSOAs there are some with care homes and some without care homes. But this is not one of the main socio-economic characteristics of the LSOA $(x_1, \ldots, x_{10})$.
- What would be the relative risk of dying in LSOA i if we kept all socio-economic variables to the values observed in that LSOA, but changed the proportion of people living in care homes to the average for the whole of England, or to some other chosen value?

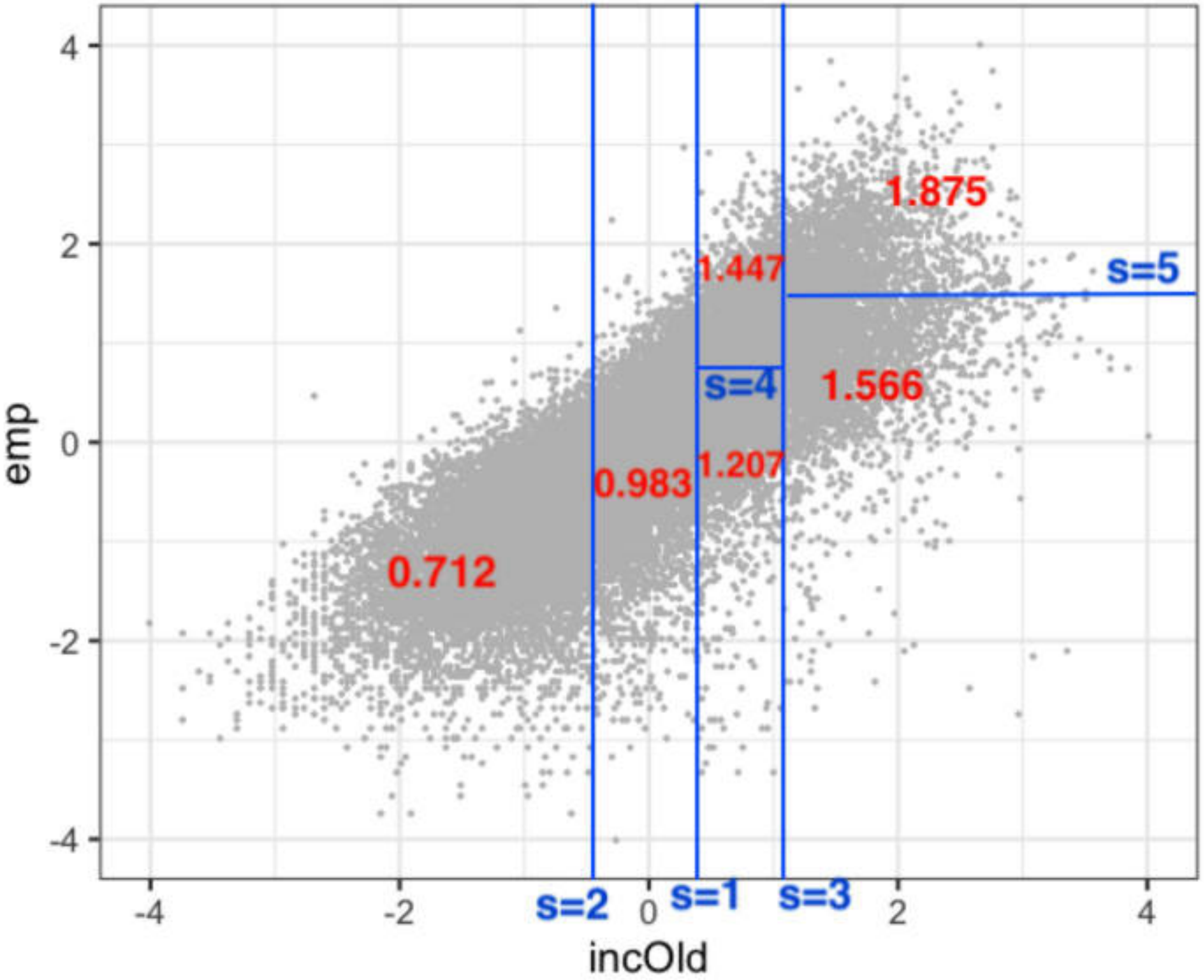
The Socio-Economic Index

- The LIFE index: the value of f for specific neighbourhoods using the socio-economic characteristics of this neighbourhood but replacing the proportion of people living in care homes with the average for the whole of England.

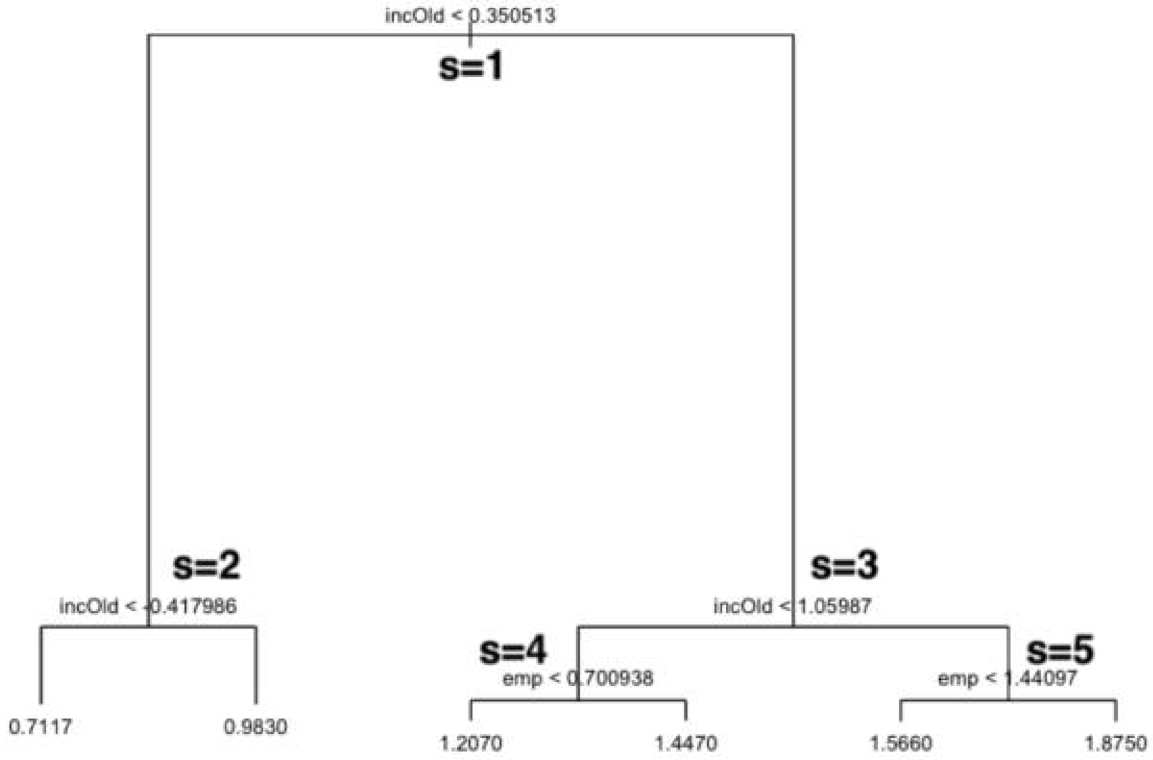
$$R_i = f(\tilde{X}_i) \text{ with } \tilde{X}_i = (X_{i,1}, \ldots, X_{i,9}, X_{i,10}, \bar{X}_{11}, \bar{X}_{12}) \tag{5}$$

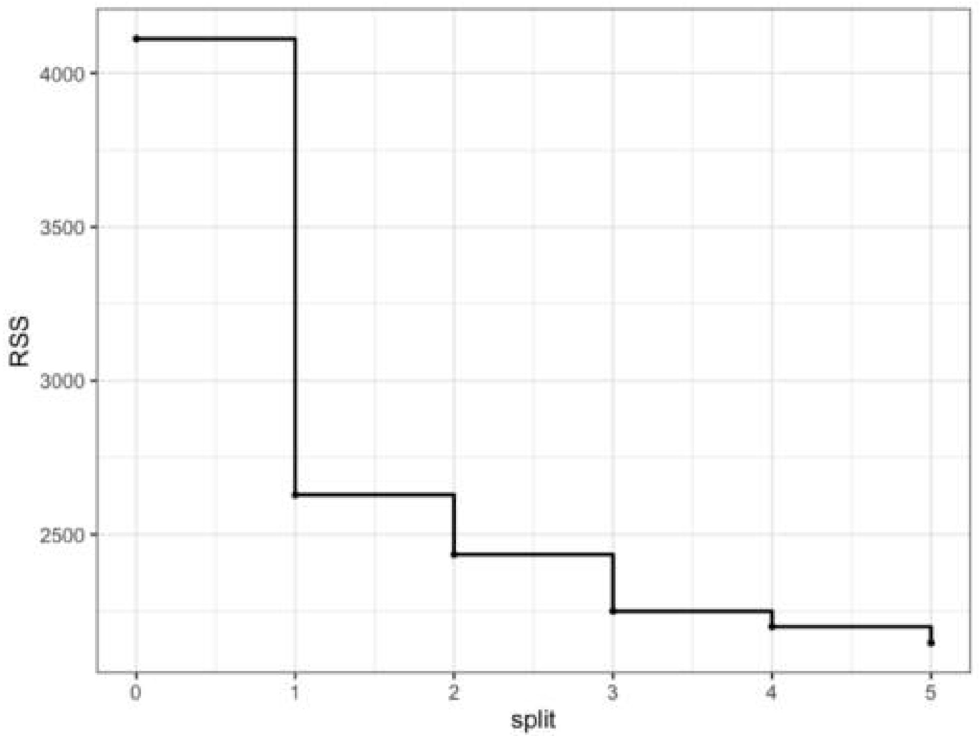
  where $\bar{X}_{11}$ and $\bar{X}_{12}$ denote the average values of the proportion of an LSOA's population living in care homes.
- Low correlation between proportion of people living in care homes and other s/e variables.
- Other adjustments to the care home variables are possible, for example, $X_{i,11} = X_{i,12} = 0$.

One Tree from the Forest



One Tree from the Forest



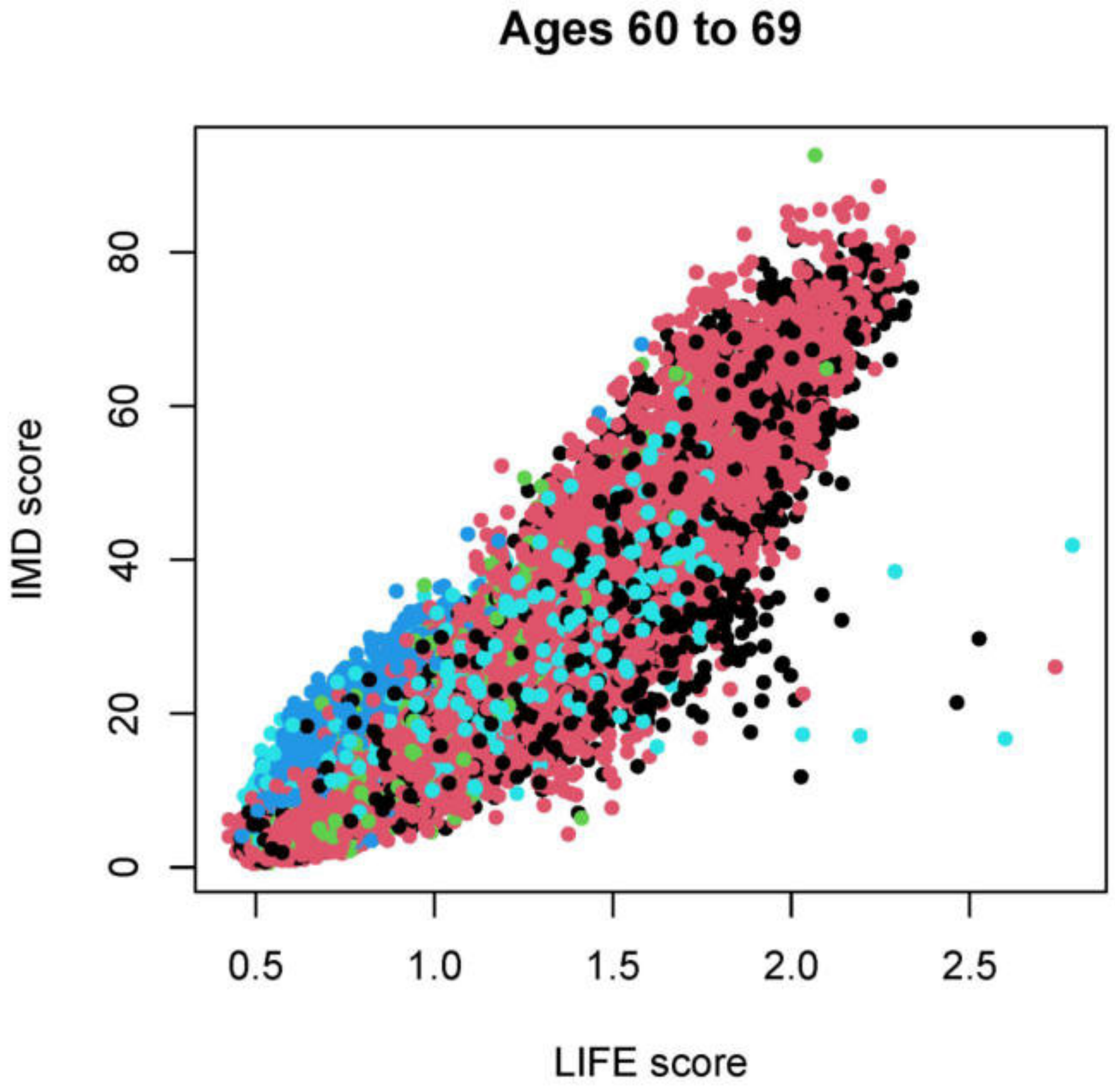


Construct many trees, each with

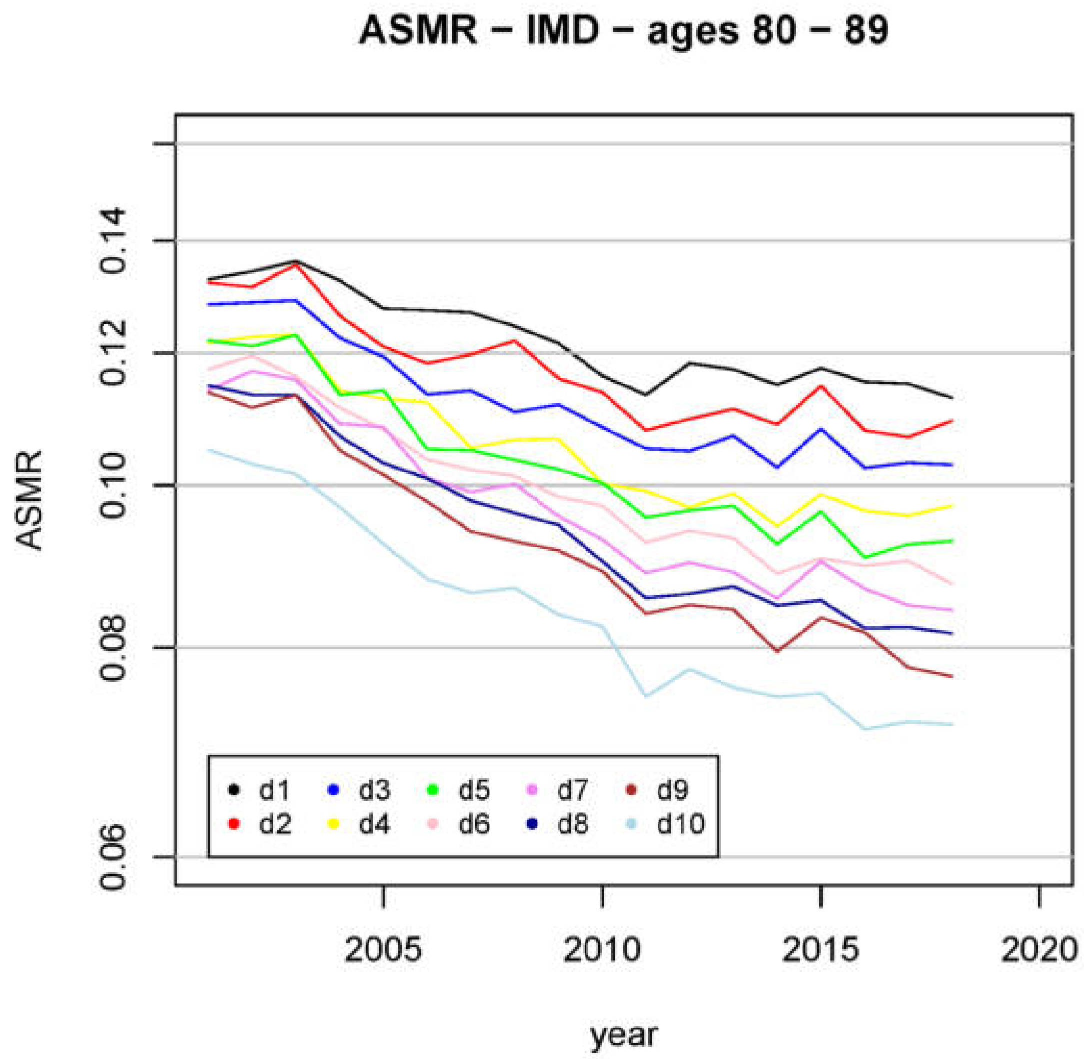
- a random sample from the full data set (sample size is a hyper-parameter)
- a random choice of a small number of predictive variables (number of predictors is a hyper-parameter)
- Our final random forest estimator $\hat{f}^{\mathrm{RF}}$ for the regression function f is the average over all individual regression trees $\hat{f}^{(b)}$:

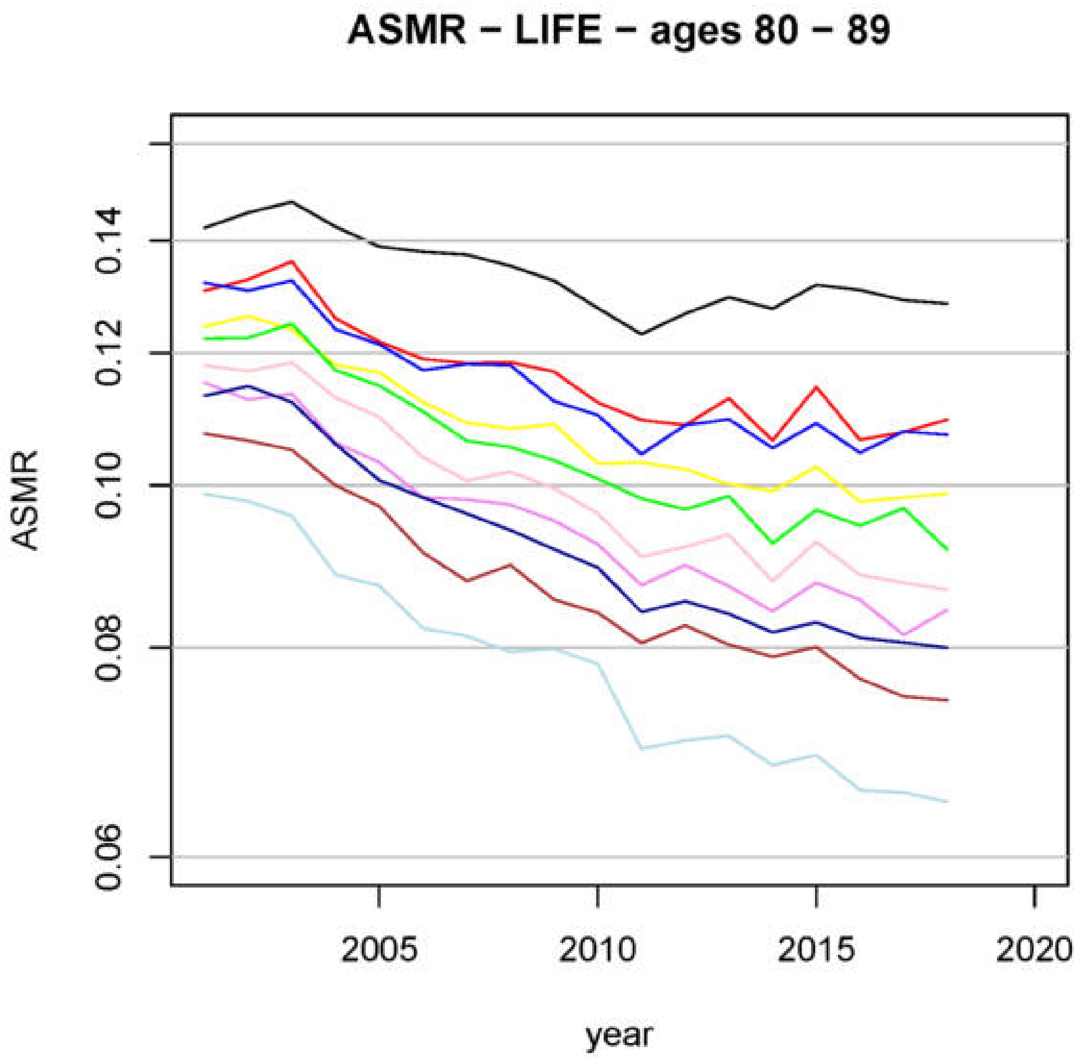
$$\hat{f}^{\mathrm{RF}}(x) = \frac{1}{B} \sum_{b=1}^{B} \hat{f}^{(b)}(x) \text{ for any } x \in L_0 \qquad (6)$$

- Note that $\hat{f}^{\mathrm{RF}}$ is piecewise constant over the full range of values of $x \in L_0$. However, $\hat{f}^{\mathrm{RF}}$ can take many more values compared to any individual tree $\hat{f}^{(b)}$.


Ages 60 to 69
IMD score
0
20
40
60
80
0.5
1.0
1.5
2.0
2.5
LIFE score

The Socio-Economic Index





Distribution of Relative Risk

UR class	all LSOAs	Relative Risk R_k fitted to ages 60 - 69			
		$G^l_{0.05}$	$G^u_{0.05}$	$G^l_{0.5}$	$G^u_{0.5}$
1	7921	220 (2.8)	944 (11.9)	2681 (33.9)	5240 (66.2)
2	14515	734 (5.1)	610 (4.2)	7205 (49.6)	7310 (50.4)
3	3056	162 (5.3)	8 (0.3)	2216 (72.5)	840 (27.5)
4	2542	326 (12.8)	0 (0.0)	2444 (96.1)	98 (3.9)
5	4810	201 (4.2)	81 (1.7)	1876 (39.0)	2934 (61.0)
Total	32844	1643 (5.0)	1643 (5.0)	16422 (50.0)	16422 (50.0)

Distribution of Relative Risk

Region	all LSOAs	$G^{l}_{0.05}$	$G^{u}_{0.05}$	$G^{l}_{0.5}$	$G^{u}_{0.5}$
East	3614	197 (5.4)	59 (1.6)	2198 (60.8)	1416 (39.2)
East Midlands	2774	69 (2.5)	103 (3.7)	1442 (52.0)	1332 (48.0)
London	4835	205 (4.2)	81 (1.7)	1895 (39.2)	2940 (60.8)
North East	1657	35 (2.1)	180 (10.9)	557 (33.6)	1100 (66.4)
North West	4497	114 (2.5)	556 (12.4)	1832 (40.7)	2665 (59.3)
South East	5382	616 (11.4)	76 (1.4)	3531 (65.6)	1851 (34.4)
South West	3281	179 (5.5)	75 (2.3)	2040 (62.2)	1241 (37.8)
West Midlands	3487	107 (3.1)	234 (6.7)	1482 (42.5)	2005 (57.5)
Yorkshire	3317	121 (3.6)	279 (8.4)	1445 (43.6)	1872 (56.4)
total	32844	1643 (5)	1643 (5)	16422 (50)	16422 (50)

ADSMR vs ASMR

- The standard ASMR adjusts mortality for different age profiles in population

$$ASMR_t = \sum_a w_a m_{ta} = \sum_a w_a \frac{D_{ta}}{E_{ta}} \text{ with } w_a = \frac{E_a^S}{\sum_a E_a^S}.$$

- The Age and Deprivation Standardised Mortality Rate - ADSMR - is designed to adjust for different deprivation profiles.

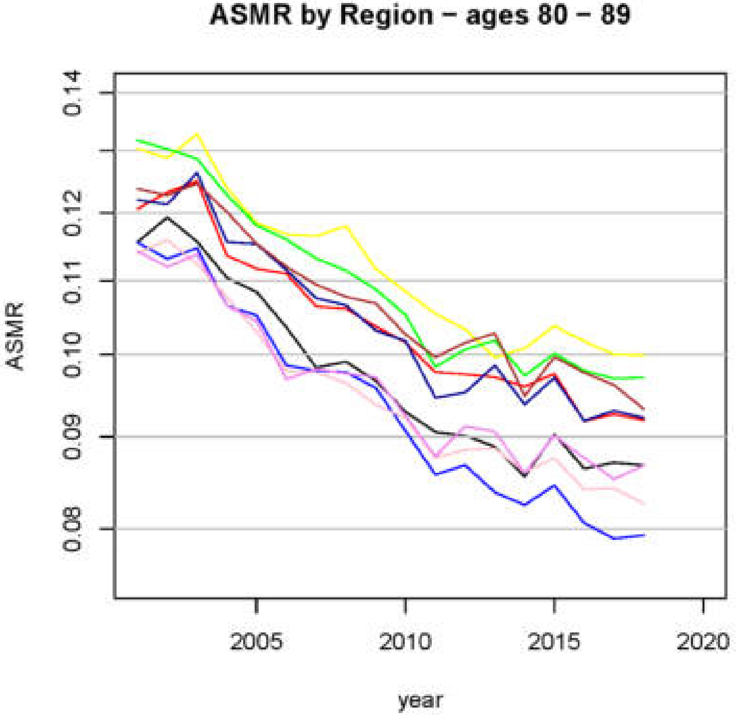
ADSMR vs ASMR

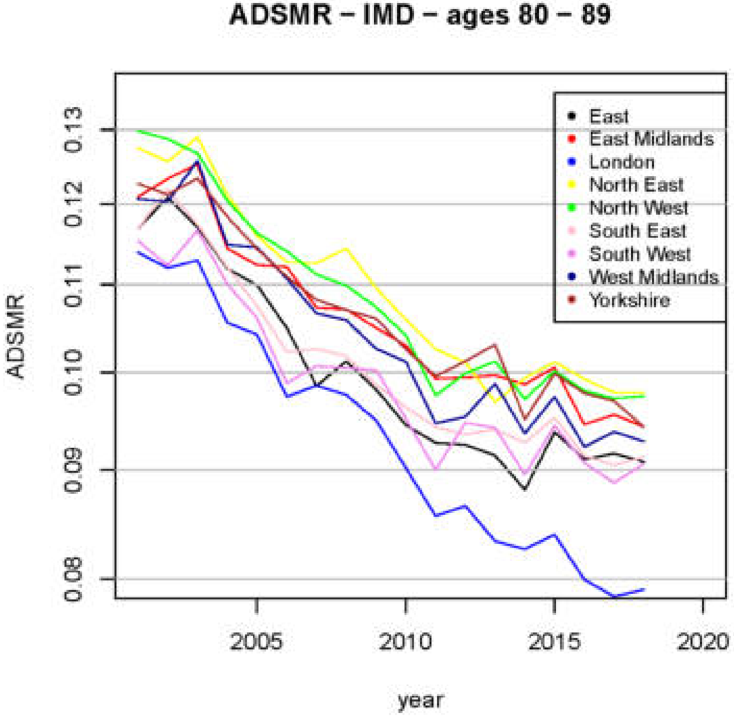
- ADSMR in region r in year t:

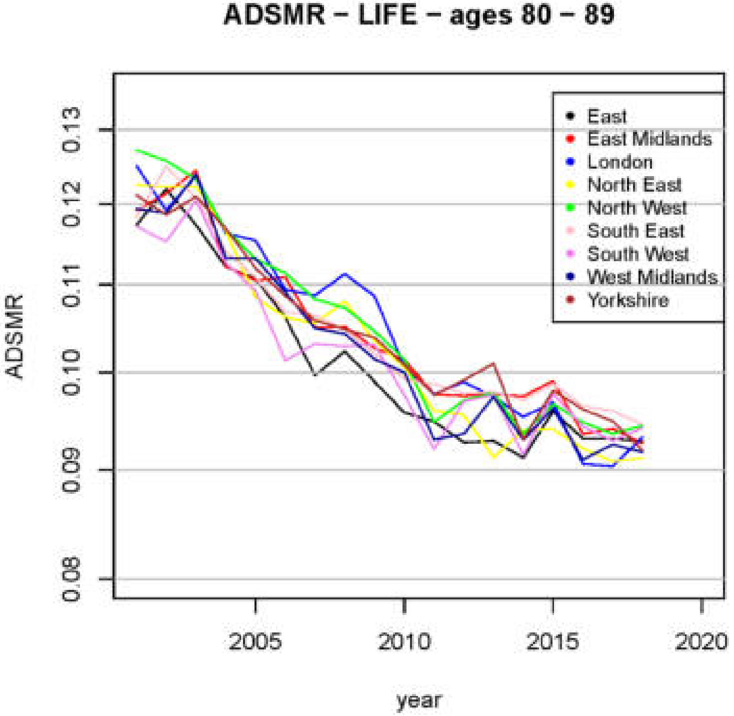
$$ADSMR_{rt} = \frac{1}{10} \sum_{k=1}^{10} ASMR_{rkt} \tag{7}$$

  where k is the deprivation decile.
- Therefore, adjust the regional deprivation profile to the national profile of 10% of LSOAs in each decile.
- … and adjust the age structure in each decile and region separately.

ADSMR vs ASMR







LIFE R Shiny App

RCLR Research Centre for Longevity Risk

www.rclr.nl