# Interpreting mortality trends in the presence of heterogeneity: A population dynamics approach

Sarah Kaakaï

Laboratoire Manceau de Mathématiques, Le Mans Université

Based on joint works with S. Arnold, N. El Karoui, H. Labit-Hardy and A. Villegas

Longevity Heterogeneity and Pension Design workshop, 28 january 2020

# Outline

# Socioeconomic gradient in mortality

- Research on the relationship between socioeconomic status (SES) and mortality is longstanding (Villermé (1830), General Register Office (1851)) $\Rightarrow$ broad consensus on the strong correlation between SES and mortality.
- New trends observed in the past decades: increasing of socioeconomic gaps in health and mortality.
  - US female life expectancy gaps at birth between less and more educated women: 7.7 years in 1990, 10.3 years in 2008 (Olshansky et al. (2012)).
  - Gap in male life expectancy at 65 between higher managerial and routine occupations (England Wales): 2.4 years 1982-1986, 3.9 years 2007-2011, ONS).

Widening gaps: socioeconomic subgroups experience rather different mortality than national mortality rates.

# Taking heterogeneity into account

- SES inequities declared key public issues by the World Health Organization in its last report on ageing and health.
- Not taking into account heterogeneity can lead to:
  - Increased inequalities due public health reforms (Alai et al. (2017)) or "unfair" redistribution properties of pensions systems.
  - Errors in funding of annuity and pension obligation (Meyricke and Sherris (2013), Villegas and Haberman (2014)).
- Better understanding of heterogeneity allows for a better understanding of basis risk (Longevity basis risk report (2014))

# Modeling heterogeneous mortality rates

- Growing literature in the joint modeling and forecasting of the mortality of socioeconomic subgroups Bensusan (2010), Jarner and Kryger (2011), Villegas and Haberman (2014), Cairns et al. (2016) ...
- Remaining questions:
  - Interpreting targets set by institutions (Department of Health, WHO) (Alai et al. (2017)).
  - Consistency of sub-national and national estimates/forecasts (Shang and Hyndman (2017), Shang and Haberman (2017)).

# Modeling heterogeneous mortality rates

- Growing literature in the joint modeling and forecasting of the mortality of socioeconomic subgroups Bensusan (2010), Jarner and Kryger (2011), Villegas and Haberman (2014), Cairns et al. (2016) ...
- Remaining questions:
  - Interpreting targets set by institutions (Department of Health, WHO) (Alai et al. (2017)).
  - Consistency of sub-national and national estimates/forecasts (Shang and Hyndman (2017), Shang and Haberman (2017)).

---

- Standard tool for modelling and forecasting longevity: mortality rates.
- Approach: take into account all population data rather that just mortality data.

## What can we learn from population dynamic?

Population divided in $p$ risk classes:

- One year central death rate in the global population for individuals age $x$ during calendar year $t$:

$$\mu_{xt} = \frac{D_{xt}}{E_{xt}} = \sum_{j=1}^{p} \frac{D^{j}{}_{xt}}{E_{xt}}$$

# What can we learn from population dynamic?

Population divided in $p$ risk classes:

- One year central death rate in the global population for individuals age $x$ during calendar year $t$:

$$\mu_{xt} = \frac{D_{xt}}{E_{xt}} = \sum_{j=1}^{p} \frac{{D^j}_{xt}}{E_{xt}} = \sum_{j=1}^{p} \frac{D^j_{xt} E^j_{xt}}{E^j_{xt} E_{xt}}$$

# What can we learn from population dynamic?

Population divided in $p$ risk classes:

- One year central death rate in the global population for individuals age $x$ during calendar year $t$:

$$\mu_{xt} = \frac{D_{xt}}{E_{xt}} = \sum_{j=1}^{p} \frac{D^{j}{}_{xt}}{E_{xt}} = \sum_{j=1}^{p} \mu^{j}_{xt} \frac{E^{j}_{xt}}{E_{xt}}$$

  - The relative exposure $W^{j}_{xt} = \frac{E^{j}_{xt}}{E_{xt}}$ is linked to the proportion of individual of age $x$ in risk class $j$.
  - Evolution of these quantities are determined by the population dynamics and can vary a lot depending on the age $x$ and time $t$.

*How changes in the socioeconomic composition of the population affect aggregated indicators? Could we miss a cause-of-death reduction in presence of heterogeneity?*

# Outline

# Data

- **Two datasets**:
  - 1981-2007: Department of Applied Health Research, UCL.
  - 2001-2015: Office for National Statistics, UK .
- English cause-specific number of deaths and mid-year population estimates per socioeconomic circumstances, age and gender.

---

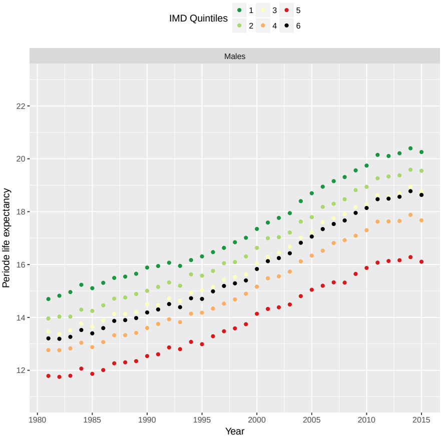
Socioeconomic circumstances are measured by the Index of multiple deprivation (IMD), based on the postcode of individuals.

- Small areas (LSOA) are ranked based on seven broad criteria: income, employment, health, education, barriers to housing and services, living environment and crime.
- This ranking permits to divide the population in 5 quintiles with about same number of individuals in each quintile.

# Life expectancy

Figure: Evolution of period life expectancy at age 65


(a) Females

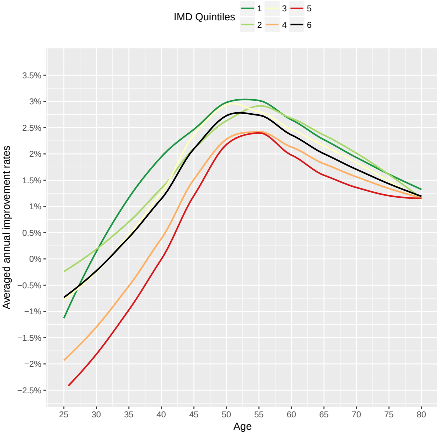



(b) Males

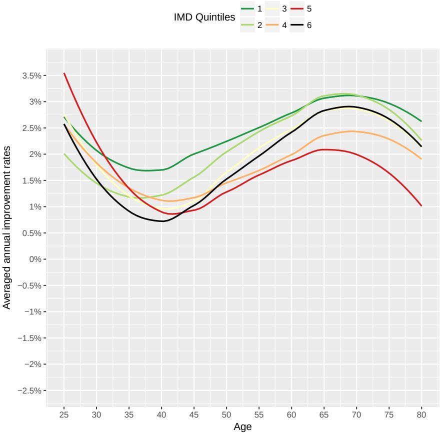
- Increase in gap: 2.2 ⇒ 4.2 years (females), 2.9 ⇒ 3.9 (males).

# Mortality by deprivation

Figure: Males average annual mortality improvement rates by age




(a) 1981-1995

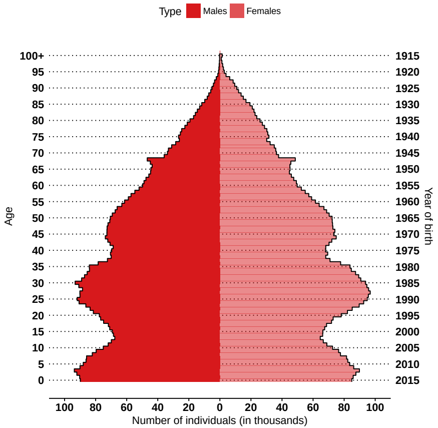



(b) 2001-2015

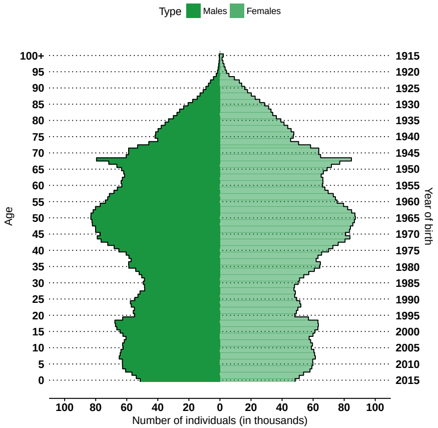
- 2001-2015: widening of mortality improvement rates gap above 60.

# Population composition, 2015

Age-pyramids by IMD quintile, 2015




(a) Most deprived quintile ($)
Median age: 33y

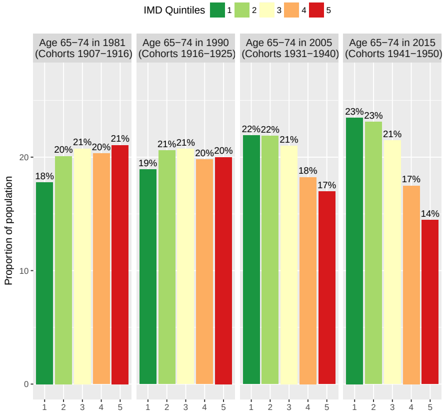



(b) Least deprived quintile ($$$$$)
Median age: 44.2y

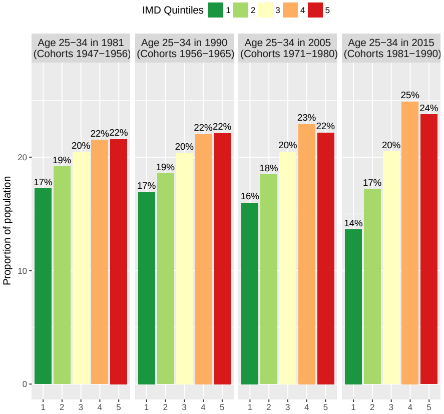
- Baby-boom cohort less deprived than younger/older cohorts.

# Changes in the population composition

Figure: Compostion of males age classes in years **1981, 1990, 2005, 2015**.




(a) Age class **65-74**




(b) Age class **25-34**

- Decrease of deprivation over time for older age classes.
- Increase of deprivation for younger age classes.

# Heterogeneous population dynamics

- Simple age-structured population dynamics framework to illustrate different impacts of heterogeneity on the aggregated mortality.
- **Deterministic** evolution of each subgroup is described by a McKendrick (1926) -Von Foerster (1959).

---

- Equation for each gender $\epsilon = m$ or $f$ and subgroup:
  - Aging law:
    $$(\partial_a + \partial_t) g_j^{\epsilon}(a,t) = -\mu_j^{\epsilon}(a,t) g_j^{\epsilon}(a,t)$$
  - Birth law:
    $$g_j^{\epsilon}(0,t) = \int_0^{a^\dagger} p^{\epsilon} g_j^{f}(a,t) b_j(a,t) da$$
  - Initial Pyramid:
    $$g_j^{\epsilon}(a,0)$$

# Aggregated population

- Aggregated population:
  - $g^\epsilon(a,t) = \sum_{j=1}^{p} g_j^\epsilon(a,t)$
  - Aging law: $(\partial_a + \partial_t) g^\epsilon(a,t) = -d^\epsilon(a,t) g^\epsilon(a,t)$
- Aggregated death rate:
  - Weighted sum of the subpopulations death rates:

$$d^\epsilon(a,t) = \sum_j w_j^\epsilon(a,t)\mu_j^\epsilon(a,t), \quad w_j^\epsilon(a,t) = \frac{g_j^\epsilon(a,t)}{g^\epsilon(a,t)} \tag{1}$$

  - $d$ depends non-linearly on the population inputs: $g_j^0$, $\mu_j$, and $b_j$.
- Even with time-independent rates $\mu_j^\epsilon(a, \cancel{t})$

$\Rightarrow$ the aggregate death rate $d^\epsilon(a, \mathbf{t})$ depends on time, due to changes of composition of the heterogeneous population.

# Numerical results

- **Two applications**
  1. Impact of the age-pyramid heterogeneity.
     ↳ Compare order of magnitude of mortality changes induced by compositional changes to constant mortality improvements.
  2. Cause specific mortality reduction vs "reverse" cohort effect.
     ↳ Compensation of cause-specific mortality reduction due to adverse compositional changes in some cohorts.

---

- We consider a synthetic population composed of the most and least deprived IMD quintile (illustrative purpose).

# Three demographic scenarios

**A** **Scenario A**: Population evolution with time-invariant mortality

Compositional changes isolated $\Rightarrow$ death rates in each subpopulation do not depend on time:

$$d^{\epsilon}(a,t) = \mu_1^{\epsilon}(a) w_1^{\epsilon}(a,t) + \mu_5^{\epsilon}(a) w_5^{\epsilon}(a,t).$$

**B** **Scenario B**: Population evolution with mortality improvement

Constant annual mortality improvement rates of $r = 0.5\%$:

$$d^{\epsilon}(a,t) = \mu_1^{\epsilon}(a)(1-r)^t w_1^{\epsilon}(a,t) + \mu_5^{\epsilon}(a)(1-r)^t w_5^{\epsilon}(a,t).$$

**C** **Scenario C**: Mortality improvements without composition changes

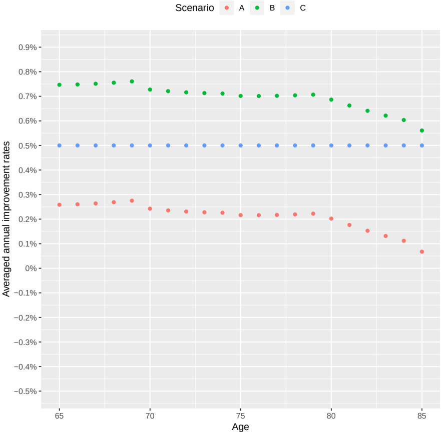
$$d^{\epsilon}(a,t) = \mu_1^{\epsilon}(a)(1-r)^t w_1^{\epsilon}(a) + \mu_5^{\epsilon}(a)(1-r)^t w_5^{\epsilon}(a).$$

---

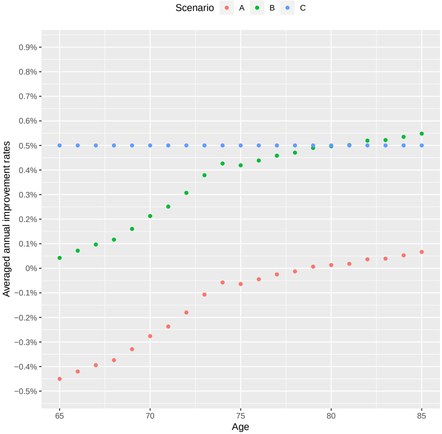
Mortality rates and initial age-pyramid fitted to the data for year 1981 and 2015.

# Mortality improvement rates (males)

Figure: Average annual mortality improvement rates over years 0-30




(a) 1981 Inputs




(b) 2015 Inputs

# Summary

- 1981 initial population: positive contribution of changes in the composition of the 65+ age class.
- 2015 initial population: negative contribution of composition changes $\Rightarrow$ might offset future mortality improvement rates.
- Order of magnitude of age-pyramid heterogeneity impact in annual mortality improvement rates of 0.2%- 0.5%.

# Composition changes and compensation effect

**Example** of scenario illustrating impact of changes of demographic rates:

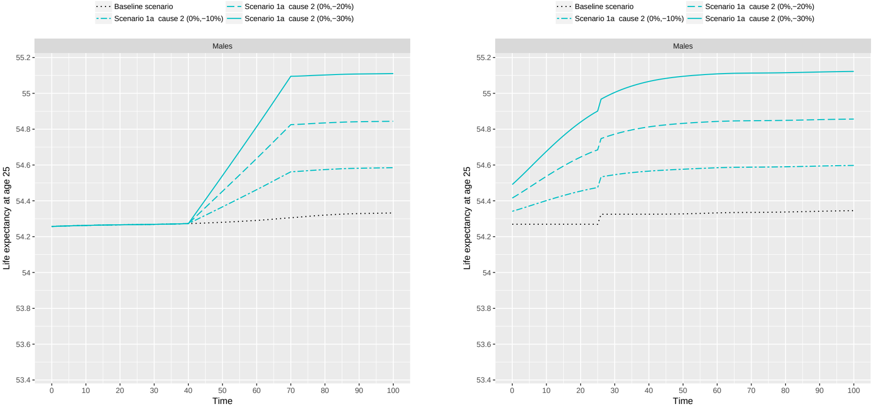
- Cause specific reduction of mortality vs "reverse" cohort effect (adverse composition changes quantified by changes in birth patterns).

---

- **Baseline ("neutral") scenario**:
  - **Death rates**: time-invariant death rates in each population (year 2015).
  - **Birth rates** same birth rates in each population (England, 2015).
  - **Initial pyramid** our aim is to limit influence of initial pyramids evolution: ~~Same initial pyramid~~ $\Rightarrow$ Stable pyramids.

# Scenario 1a: cause of death reduction

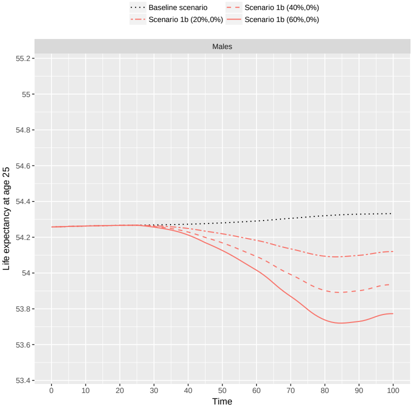
Figure: Reduction of mortality rates from cardiovascular disease (CVD)



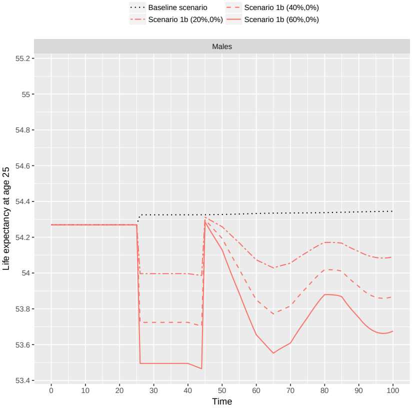

(a) Male period life expectancy at 25 (b) Male cohort life expectancy at 25

- Reduction of CoD mortality from CVD (10, 20 and 30%) over a period of 30 years, starting at $t = 40$.

# Scenario 1b: "Reverse" Cohort effect




(a) Male period life expectancy at 25




(b) Male cohort life expectancy at 25

- Reverse cohort effect: increase of birth rates in most deprived subgroup over period $[0, 20]$.
- $\nearrow$ 60% $\Rightarrow$ cohorts composed of 63% of most deprived subgroup.

# Scenario 2: combined CoD reduction and cohort effect

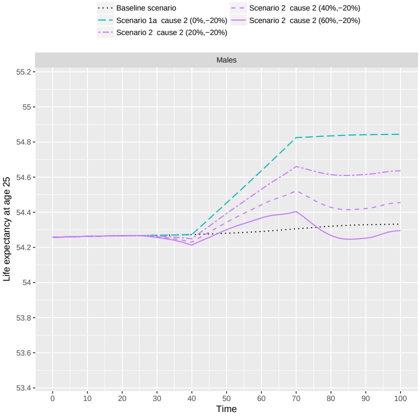



Figure: Male period life expectancy at 25

*When the population heterogeneity is not taken into account, cause-of-death mortality reduction could be compensated for and/or misinterpreted depending on the population composition evolution.*

# Outline

# The consistency issue

- Population divided in $p$ risk classes.

  Example: national population divided in socioeconomic or geographical subgroups...
- Many forecasts are made simultaneously at the national level and at the subgroup level.
- **Issue** Are forecasts consistent?

$$\textit{(Aggregation relation)} \quad \mu_{xt} = \sum_{j=1}^{p} \mu_{xt}^{j} W_{xt}^{j} \quad ?$$

- Problem when comparing policies, question robustness of forecasts.

# Existing approaches

- **Relative** approach (e.g.Jarner and Kryger (2011), Cairns et al. (2011), Villegas and Haberman (2014), Li et al. (2015))
  - Method: First forecast of the national population and *then* mortality differentials $\mu^j_{xt} - \mu_{xt}$.
  - Advantages: national mortality data available for longer period, more precise age-groups, estimations less volatile.
  - Drawback: No consistency!

# Existing approaches

- **Relative** approach (e.g.Jarner and Kryger (2011), Cairns et al. (2011), Villegas and Haberman (2014), Li et al. (2015))
  - Method: First forecast of the national population and *then* mortality differentials $\mu^j_{xt} - \mu_{xt}$.
  - Advantages: national mortality data available for longer period, more precise age-groups, estimations less volatile.
  - Drawback: No consistency!
- **Bottom-up** approach:
  - Method: First forecast subnational mortality and *then* pose $\mu_{xt} = \sum_{j=1}^{p} \mu^j_{xt} W^j_{xt}$.
  - Advantages: consistent forecasts.
  - Drawbacks: poor/missing data quality, small datasets.

# Existing approaches

- **Relative** approach (e.g.Jarner and Kryger (2011), Cairns et al. (2011), Villegas and Haberman (2014), Li et al. (2015))
  - Method: First forecast of the national population and *then* mortality differentials $\mu^j_{xt} - \mu_{xt}$.
  - Advantages: national mortality data available for longer period, more precise age-groups, estimations less volatile.
  - Drawback: No consistency!
- **Bottom-up** approach:
  - Method: First forecast subnational mortality and *then* pose $\mu_{xt} = \sum_{j=1}^{p} \mu^j_{xt} W^j_{xt}$.
  - Advantages: consistent forecasts.
  - Drawbacks: poor/missing data quality, small datasets.
- **Optimal combination** : Combines the advantages of the two approaches.

# Optimal combination

- Forecast reconciliation method used in economics, operation research and recently for mortality rates (Shang and Hyndman (2017), Shang and Haberman (2017)).
- Vector notations: $\beta_n(x) = \begin{pmatrix} \mu_n^1 \\ \vdots \\ \mu_n^p \end{pmatrix}(x)$, $Y_t(x) = \begin{pmatrix} \mu_t \\ \beta_t \end{pmatrix}(x)$.
- Consistency equation:

$$Y_n = W_n \beta_n, \quad W_n = \begin{pmatrix} W_n^1 & W_n^2 & \cdots & W_n^p \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & & 0 \\ 0 & \cdots & & 1 \end{pmatrix}.$$

## Optimal combination

**Method**:

1. First forecast mortality rates $(Y_n)_{n=t..T}$ using any method:

$$Y_n \neq W_n \beta_n.$$

2. Revise the subpopulations forecasts $\beta_n = {}^t(\hat{\mu}_n^1, \cdots, \hat{\mu}_n^p)$ so that the forecast $\tilde{Y}_n = W_n \tilde{\beta}_n$ obtained by bottom-up is as close as possible as the initial forecast $Y$.

$$Y_n = W_n \tilde{\beta}_n + \epsilon_n.$$

3. Replace $Y$ by the consistent forecast $\tilde{Y}_n = W_n \tilde{\beta}_n$.

# Optimal combination

**Method**:

1. First forecast mortality rates $(Y_n)_{n=t..T}$ using any method:

$$Y_n \neq W_n \beta_n.$$

2. Revise the subpopulations forecasts $\beta_n = {}^t(\hat{\mu}_n^1, \cdots, \hat{\mu}_n^p)$ so that the forecast $\tilde{Y}_n = W_n \tilde{\beta}_n$ obtained by bottom-up is as close as possible as the initial forecast $Y$.

$$Y_n = W_n \tilde{\beta}_n + \epsilon_n.$$

3. Replace $Y$ by the consistent forecast $\tilde{Y}_n = W_n \tilde{\beta}_n$.

---

**How can the relative exposures $W_n^j(x) = \frac{E_n^j(x)}{E_n(x)}$ be forecasted?**

# Population composition forecast

**Proposed methods**

- Forecast Exposure to Risk at each age $x$ $(E_n^j(x))_{n=t..T}$ using independent time series.

# Population composition forecast

**Proposed methods**

- Forecast Exposure to Risk at each age $x$ $(E_n^j(x))_{n=t..T}$ using independent time series.
- Constant Exposure to risk $E_{t+h}^j(x) = E_t^j(x-h)$.

---

**Reseach program**

- Population dynamics model provides a natural framework for the projection of relative exposures, using all the age-pyramid information.
- Adapt the bottom-up and optimal combination method in this context.
- Challenge:

$$Y_n = W_n((\beta_s)_{s=t...n})\beta_n,$$

# Conclusion and perspectives

**Population dynamics framework** ⇒ study of interactions between changes in socioeconomic composition and mortality indicators:

- Evidence from data: future evolution of the population socioeconomic composition could have a different impact on mortality than 30y ago.
- Cause-of-death mortality changes can be compensated and/or misinterpreted in the presence of heterogeneity.

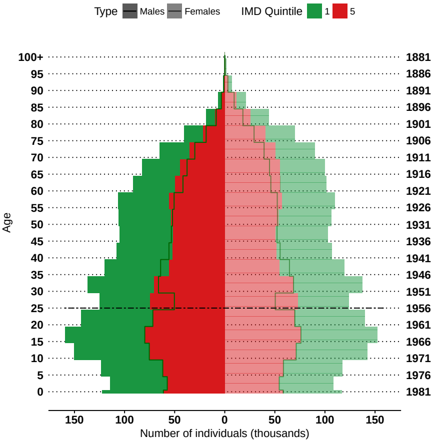
---

**Perspectives**

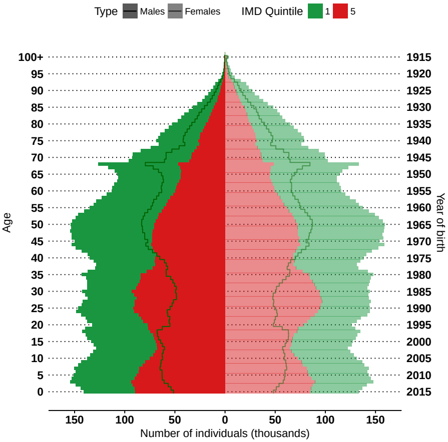
- Take into account internal migrations: ⇒ non-linear mortality models.
- Stochastic modelling
  ↳ R package IBMPopSim in preparation,(with D. Giorgi and V. Lemaire).
- Consistency of mortality forecasts under different demographic scenarios.

Thank you for you attention!

# Initial age-pyramids



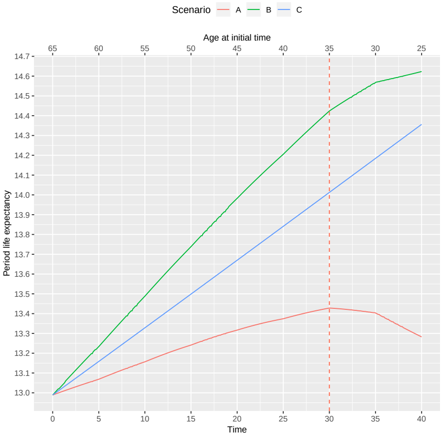

(a) 1981 Inputs



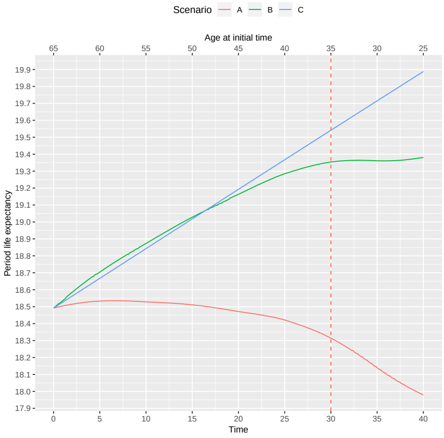

(b) 2015 Inputs

# Period life expectancy

Figure: Evolution of male life expectancy at age 65




(a) 1981 Inputs




(b) 2015 Inputs

- Life expectancy at $t = 30$: surviving individuals initially 35+

# Birth-Death-Swap processes

*Pathwise construction of Birth-Death-Swap systems leading to an averaging result in presence of two timescales, with N. El Karoui*

- **Goal**: To study the random evolution of an heterogeneous population including:
  - A time-varying random environment.
  - Model changes in the population's composition induced by interacting individuals changing characteristics.

---

- Main contributions:
  - General mathematical framework and tools to study such processes.
  - Study of the aggregated "macro" dynamic produced by such models.
- **Averaging result**: aggregated mortality rates are approximated by "averaged" rates depending non-trivially on the number of individuals in the population.