# Estimating the impact of COVID-19 on mortality using granular data

Frank van Berkum, Bertrand Melenberg & Michel Vellekoop

Longevity 18, 7-8 September 2023

University of Amsterdam

RCLR

# Outline

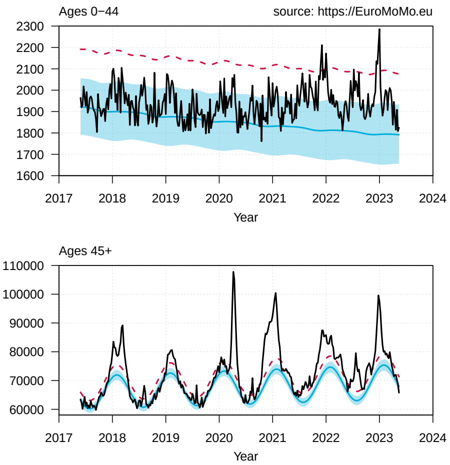
# A mortality shock

## Challenging times for life actuaries

The spread of the corona-virus throughout the world since December 2019 resulted in periods with elevated levels of mortality (excess mortality):

- In the Netherlands, mortality during 2020-2022 was $>$ 10% higher than expected (CBS.nl)
- Actuaries must understand the impact of COVID on mortality for pricing and reserving purposes

**Observed deaths in various EU countries**



# A mortality shock

## Challenging times for life actuaries

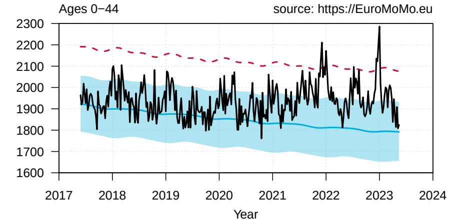
The spread of the corona-virus throughout the world since December 2019 resulted in periods with elevated levels of mortality (excess mortality):

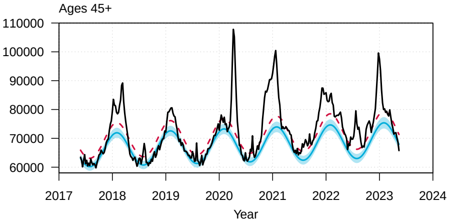
- In the Netherlands, mortality during 2020-2022 was $>$ 10% higher than expected (CBS.nl)
- Actuaries must understand the impact of COVID on mortality for pricing and reserving purposes

There are various elements to consider when analyzing the impact of COVID on mortality:

- On short term: how many COVID-related deaths will we continue to observe?
- On long term: what is the net effect of delayed health care and long COVID vs. positive effects on longevity?

**Observed deaths in various EU countries**



# Literature overview

## What approaches have others taken?

The literature on COVID-19 and mortality (forecasting) is growing:

Cairns et al. [2020]:

- Analyze mortality at higher ages during the early stages of the pandemic (May 2020) using a Gompertz model
- Show that deaths early in pandemic were mostly related to less healthy people (accelerated deaths)
- If deaths are accelerated, then there is limited impact on life expectancy of survivors

# Literature overview
## What approaches have others taken?

The literature on COVID-19 and mortality (forecasting) is growing:

### Cairns et al. [2020]:

- Analyze mortality at higher ages during the early stages of the pandemic (May 2020) using a Gompertz model
- Show that deaths early in pandemic were mostly related to less healthy people (accelerated deaths)
- If deaths are accelerated, then there is limited impact on life expectancy of survivors

### Schnürch et al. [2022]:

- Analyze weekly and annual mortality observations using STMF data
- Consider LC and CBD model and assess how extra deaths in 2020/2021 affect mortality predictions
- Focus is on changes in time series parameters; no additional parameters are estimated

# Literature overview

## What approaches have others taken?

The literature on COVID-19 and mortality (forecasting) is growing:

### Cairns et al. [2020]:

- Analyze mortality at higher ages during the early stages of the pandemic (May 2020) using a Gompertz model
- Show that deaths early in pandemic were mostly related to less healthy people (accelerated deaths)
- If deaths are accelerated, then there is limited impact on life expectancy of survivors

### Schnürch et al. [2022]:

- Analyze weekly and annual mortality observations using STMF data
- Consider LC and CBD model and assess how extra deaths in 2020/2021 affect mortality predictions
- Focus is on changes in time series parameters; no additional parameters are estimated

### Robben et al. [2022]:

- Use weekly STMF data to construct annual mortality observations
- Calibrate Li-Lee model without and with COVID observations
- Introduce weights in the time series model to limit the impact from period effects during COVID years
- Mitigate the impact of pandemic data points by adjusting the starting point of the projection

# Outline

# A COVID-extension for the Li-Lee model

## Analyzing weekly death counts using a pre-COVID baseline

We model weekly death counts using the following structure:

$$D^{c,g}_{x,t,w} \sim \text{Poisson}\Big( \overbrace{E^{c,g}_{x,t,w}\mu^{c,g}_{x,t}}^{\substack{\text{pre-COVID}\\ \text{expected mortality}}} \times \underbrace{\phi^{c,g}_{x,w}}_{\substack{\text{seasonal}\\ \text{effect}}} \times \overbrace{\exp(\mathfrak{B}^{c,g}_x \mathfrak{K}^{c,g}_{t,w})}^{\text{COVID-term}} \Big),$$

with

$$\ln \mu^{c,g}_{x,\,t} = B^g_x K^g_t + \alpha^{c,g}_x + \beta^{c,g}_x \kappa^{c,g}_t.$$

We calibrate the parameters $\mathfrak{B}^{c,g}_x$ and $\mathfrak{K}^{c,g}_{t,w}$ by optimizing the Poisson-likelihood assuming $E^{c,g}_{x,t,w}$, $\mu^{c,g}_{x,t}$ and $\phi^{c,g}_{x,w}$ are known.

*We impose the constraint $\sum_x \mathfrak{B}^2_x = 1$ to uniquely identify the parameters.*

Meaning of sub- and superscripts:

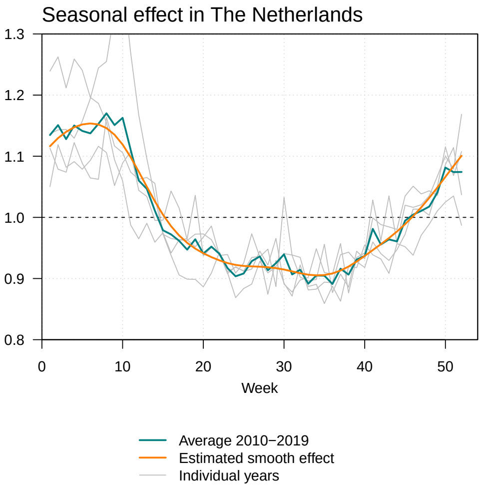
$$\begin{aligned} c &= \text{country} \\ g &= \text{gender} \\ x &= \text{age} \\ t &= \text{year} \\ w &= \text{week number} \end{aligned}$$

Death counts available:

- Statistics Netherlands (NL only): per age, gender, week
- STMF (for various countries): per age group, gender, week

# Seasonal patterns in mortality

## Two possible approaches to reflect in the model




Recall the expression for the $E(D^{c,g}_{x,t,w})$:

$$\overbrace{E^{c,g}_{x,t,w}\mu^{c,g}_{x,t}}^{\text{pre-COVID expected mortality}} \times \underbrace{\phi^{c,g}_{x,w}}_{\text{seasonal effect}} \times \overbrace{\exp(\mathfrak{B}^{c,g}_{x}\mathfrak{K}^{c,g}_{t,w})}^{\text{COVID-term}}$$

We consider two approaches for treating the seasonal effect:

**Method 1:** Set $\phi^{c,g}_{x,w} = 1$, i.e., do not impose any seasonal effect.

**Method 2:** Set $\phi^{c,g}_{x,w} = \bar{\phi}^{c,g}_{x,w}$, i.e., impose the historically observed seasonal effect.

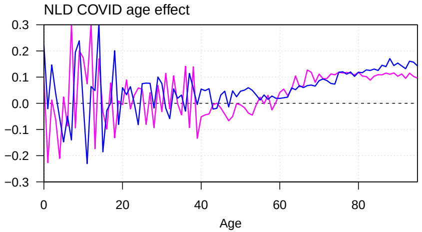
# Outline

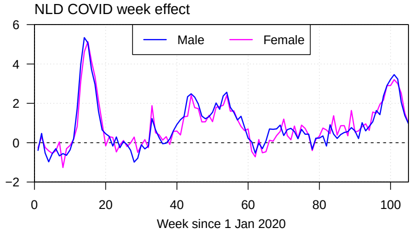
# Initial calibration

## Calibrated using Dutch data, ages 0-95, both genders, Method 2 for the seasonal effect

Our first results are in line with observations during the pandemic:

- Very volatile age effects at lower ages indicates mortality at those ages was hardly affected by COVID
- For higher ages ($> 40$), the age effect is more stable and increasing with age indicating mortality at higher ages is (more) severely affected
- The week effects show clear spikes which follow the various COVID waves (1st peak is around March '20, 2nd peak during winter '20/'21, 3rd peak during winter '21/'22)



# Sensitivity analysis

## Comparison of two approaches for seasonal effect

We described two methods to treat the seasonal effect:

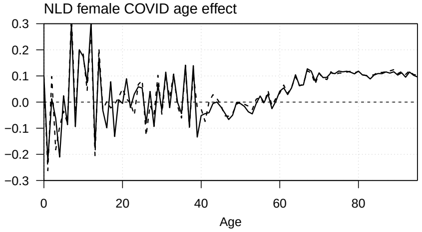
Method 1: Do not impose any seasonal effect upfront.

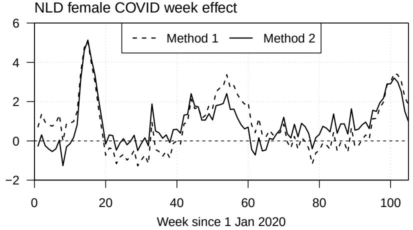
Method 2: Impose the historically observed seasonal effect.

Comparing the results of the two approaches, we observe the following:

- The age effects are nearly the same.
- The week effects show important differences:
  - The week effect of Method 1 captures the impact of both seasonal effects and COVID on mortality
  - The week effect of Method 2 represents the impact of only COVID on mortality

When the aim is to analyze the impact of COVID on mortality, Method 2 is to be preferred.



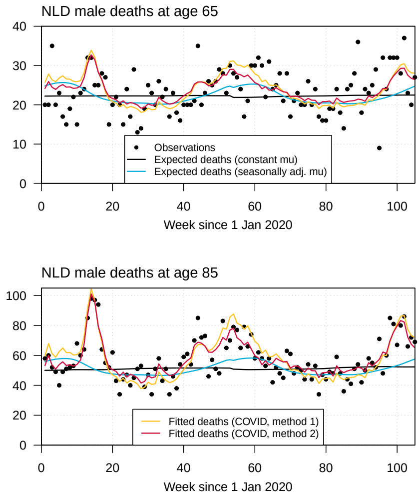



# Inspection of calibration results

## Fitted mortality rates

Comparing observed, expected and fitted death counts, we observe:

- Death counts at individual ages are volatile, but especially at higher ages show clear wave patterns.
- Seasonally adjusted expected deaths clearly show seasonal patterns, but the COVID waves occur at different points in time and have different magnitude
- The fitted death counts (both methods for seasonal effect) closely resemble the weekly developments in observed mortality.



# Sensitivity analysis

## Importance of granular data

For forecasting purposes we prefer observations for individual ages, but often we only have weekly death counts available for age groups

Consider the following three datasets with different levels of granularity:

Level 1: Death counts for individual ages:
$D^{c,g}_{x,t,w}$ for $x = \{0\}, \{1\}, ..., \{95\}$

Level 2: Death counts for age groups of five years:
$x = \{0\text{-}4\}, \{5\text{-}9\}, ..., \{90\text{-}94\}, \{95+\}$

Level 3: Death counts for large age groups:
$x = \{0\text{-}14\}, \{15\text{-}64\}, \{65\text{-}74\}, \{75\text{-}84\}, \{85+\}$

# Sensitivity analysis

## Importance of granular data

For forecasting purposes we prefer observations for individual ages, but often we only have weekly death counts available for age groups

Consider the following three datasets with different levels of granularity:

Level 1: Death counts for individual ages:
$D^{c,g}_{\chi,t,w}$ for $\chi = \{0\}, \{1\}, ..., \{95\}$

Level 2: Death counts for age groups of five years:
$\chi = \{0\text{ - }4\}, \{5\text{ - }9\}, ..., \{90\text{ - }94\}, \{95+\}$

Level 3: Death counts for large age groups:
$\chi = \{0\text{ - }14\}, \{15\text{ - }64\}, \{65\text{ - }74\}, \{75\text{ - }84\}, \{85+\}$

We use annual death counts for individual ages ($D^{c,g}_{x,t}$) to allocate weekly death counts per age group ($D^{c,g}_{\chi,t,w}$) to weekly death counts for individual ages ($D^{c,g}_{x,t,w}$):

$$D^{c,g}_{x,t,w} = D^{c,g}_{\chi,t,w} \cdot \frac{\sum_{t=2015}^{2019} D^{c,g}_{x,t}}{\sum_{x\in\chi}\sum_{t=2015}^{2019} D^{c,g}_{x,t}}.$$

We compare the COVID parameters using the resulting datasets (with death counts for individual ages)

# Sensitivity analysis

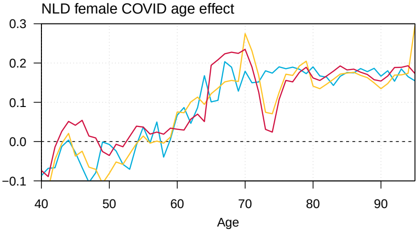
## Importance of granular data (cont.)

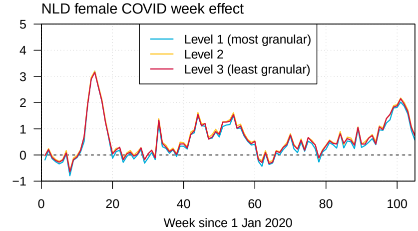
We note the following:

- The COVID-parameters using Level 1 are the same as the parameters shown before.
- The COVID age effects using Level 2 and Level 3 are substantially different, which is caused by cohort effects in observed deaths.
- The COVID week effects are practically the same for all three levels of granularity.

We conclude:

- For analyzing the development over time, one can use data from large age groups.
- To understand how different ages are affected, one should use data on individual ages.



# Outline

# Annalizing the COVID-parameters

## From week to annual parameters

We intend to develop scenarios that reflect the impact of COVID on mortality forecasts:

- The COVID parameters are estimated on weekly scale, but mortality forecasts are on annual basis
- We need to transform the COVID parameters on week basis to parameters on annual basis

# Annualizing the COVID-parameters

## From week to annual parameters

We intend to develop scenarios that reflect the impact of COVID on mortality forecasts:

- The COVID parameters are estimated on weekly scale, but mortality forecasts are on annual basis
- We need to transform the COVID parameters on week basis to parameters on annual basis

We first annualize the week effects by equating the annual survival probability to the product of weekly survival probabilities,

$$\exp\left(-\mu_{x,t}^{\mathrm{c,g}} \cdot \exp[\mathfrak{V}_x^{\mathrm{c,g}} \mathfrak{X}_t^{\mathrm{c,g}}]\right) = \prod_{w=1}^{w_t} \exp\left(-\frac{1}{w_t} \cdot \mu_{x,t}^{\mathrm{c,g}} \cdot \phi_{x,w}^{\mathrm{c,g}} \cdot \exp\left[\mathfrak{B}_x^{\mathrm{c,g}} \mathfrak{K}_{t,w}^{\mathrm{c,g}}\right]\right),$$

and when temporarily assuming $\sum_x \mathfrak{B}_x^{\mathrm{c,g}} = 1$, we get

$$\mathfrak{X}_t^{\mathrm{c,g}} = \sum_{x \in \mathcal{X}_a} \ln \left\{ \frac{1}{w_t} \sum_{w=1}^{w_t} \phi_{x,w}^{\mathrm{c,g}} \cdot \exp[\mathfrak{B}_x^{\mathrm{c,g}} \mathfrak{K}_{t,w}^{\mathrm{c,g}}] \right\}.$$

# Annalizing the COVID-parameters

## From week to annual parameters

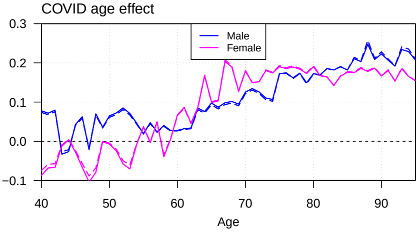
Next, we annualize the age effects by making survival over the years 2020 and 2021 equal to surviving over all weeks in those years:

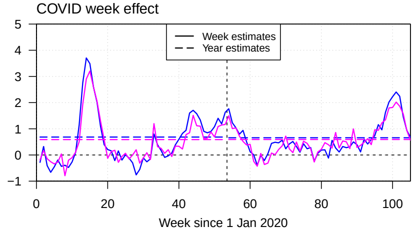
$$\prod_{t=2020}^{2021} \exp\left(-\mu_{x,t}^{c,g} \cdot \exp[\mathfrak{V}_x^{c,g}\mathfrak{X}_t^{c,g}]\right) =$$

$$\prod_{t=2020}^{2021} \prod_{w=1}^{w_t} \exp\left(-\frac{1}{w_t} \cdot \mu_{x,t}^{c,g} \cdot \phi_{x,w}^{c,g} \cdot \exp[\mathfrak{B}_x^{c,g}\mathfrak{K}_{t,w}^{c,g}]\right).$$

This equation can be solved numerically for each age $x$ separately

Note that the annualized COVID period effect is close to the average of the week effects in a year



# The impact of COVID on future mortality
## Forecasting the COVID period effect

We analyze how mortality rates develop under varying assumptions for the future impact of COVID-19, all of which can be written as:

$$\begin{aligned}\ln \mu_{x,2021+h}^{\text{scen,c,g}} &= \ln \mu_{x,2021+h}^{\text{c,g}} + \mathfrak{B}_x^{\text{c,g}} \mathfrak{X}_{2021+h}^{\text{scen,c,g}} \\ \mathfrak{X}_{2021+h}^{\text{scen,c,g}} &= \mathfrak{X}_{\text{start}}^{\text{c,g}} \eta^h + (1-\eta^h)\mathfrak{X}_{\infty}^{\text{c,g}},\end{aligned}$$

# The impact of COVID on future mortality

## Forecasting the COVID period effect

We analyze how mortality rates develop under varying assumptions for the future impact of COVID-19, all of which can be written as:

$$\ln \mu^{\text{scen,c,g}}_{x,2021+h} = \ln \mu^{\text{c,g}}_{x,2021+h} + \mathfrak{Y}^{\text{c,g}}_{x} \mathfrak{X}^{\text{scen,c,g}}_{2021+h}$$
$$\mathfrak{X}^{\text{scen,c,g}}_{2021+h} = \mathfrak{X}^{\text{c,g}}_{\text{start}} \eta^{h} + (1-\eta^{h}) \mathfrak{X}^{\text{c,g}}_{\infty},$$

To generate forecasts for all ages, we assume the following:

- For $\mathfrak{X}^{\text{scen,c,g}}_{2021+h}$, we consider the scenarios as defined on the right-hand side with $0 \leq \eta \leq 1$
- $\mathfrak{Y}^{\text{c,g}}_{x} = \mathfrak{Y}^{\text{c,g}}_{40}$ for $x < 40$ and $\mathfrak{Y}^{\text{c,g}}_{x} = \mathfrak{Y}^{\text{c,g}}_{95}$ for $x > 95$.

**Completely incidental:**

$$\mathfrak{X}^{\text{c,g}}_{\text{start}} = \mathfrak{X}^{\text{c,g}}_{\infty} = 0$$

**Completely structural:**

$$\mathfrak{X}^{\text{c,g}}_{\text{start}} = \mathfrak{X}^{\text{c,g}}_{\infty} = \mathfrak{X}^{\text{c,g}}_{2021}$$

**Decreasing impact:**

$$\mathfrak{X}^{\text{c,g}}_{\text{start}} = \mathfrak{X}^{\text{c,g}}_{2021},\ \mathfrak{X}^{\text{c,g}}_{\infty} = 0$$

**Growing impact:**

$$\mathfrak{X}^{\text{c,g}}_{\text{start}} = \mathfrak{X}^{\text{c,g}}_{2021},\ \mathfrak{X}^{\text{c,g}}_{\infty} = 1.25\mathfrak{X}^{\text{c,g}}_{2021}$$

**New normal:**

$$\mathfrak{X}^{\text{c,g}}_{\text{start}} = \mathfrak{X}^{\text{c,g}}_{2021},\ \mathfrak{X}^{\text{c,g}}_{\infty} = 0.25\mathfrak{X}^{\text{c,g}}_{2021}$$

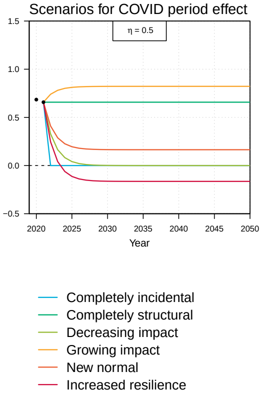
**Increased resilience:**

$$\mathfrak{X}^{\text{c,g}}_{\text{start}} = \mathfrak{X}^{\text{c,g}}_{2021},\ \mathfrak{X}^{\text{c,g}}_{\infty} = -0.25\mathfrak{X}^{\text{c,g}}_{2021}$$

# Outline

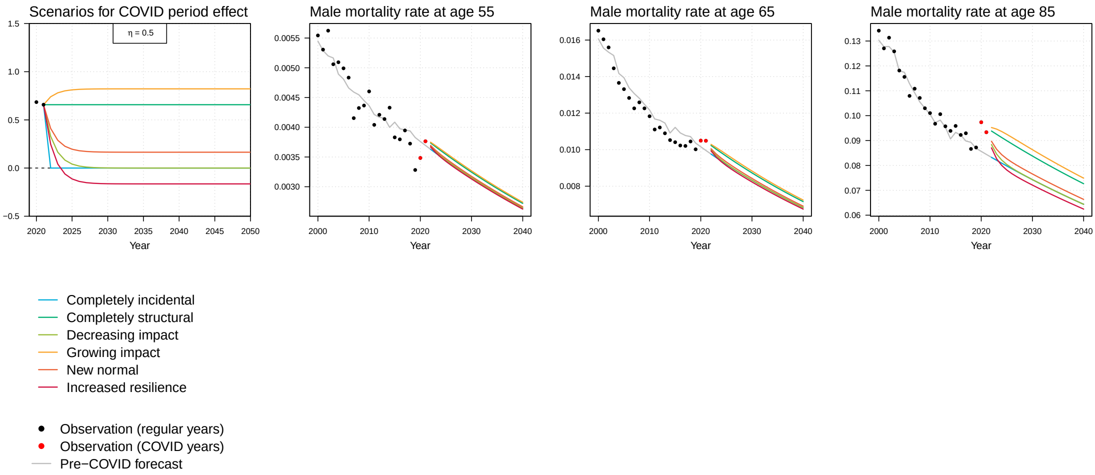
# The impact of COVID on future mortality

## Forecasts of the COVID period effect



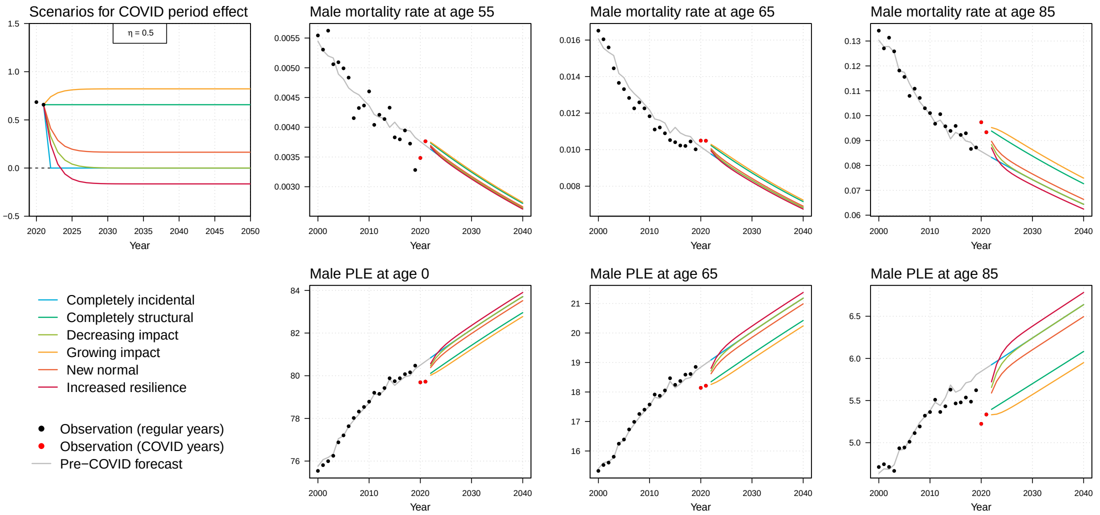
# The impact of COVID on future mortality

## Forecasts of the COVID period effect, mortality rates



# The impact of COVID on future mortality

## Forecasts of the COVID period effect, mortality rates, and period life expectancies



# Outline

# Conclusion

Using weekly mortality observations, we have quantified the impact of COVID-19 relative to pre-COVID expectations

The estimated COVID-factors are used to define various scenarios for the future course of the pandemic

For more details, see the working paper:
van Berkum et al. [2022]

# Conclusion

Using weekly mortality observations, we have quantified the impact of COVID-19 relative to pre-COVID expectations

The estimated COVID-factors are used to define various scenarios for the future course of the pandemic

For more details, see the working paper: van Berkum et al. [2022]

The big challenge for life actuaries is how to update mortality predictions:

- Should the long term mortality trend be adjusted, and if so, how?

At **RCLR**, we continue with research on this important topic by:

- Analyzing mortality by cause-of-death before and during the pandemic
- Using socio-economic factors, neighborhood information and vaccination/infection data to quantify differences between individuals

These insights may be used in setting new long term mortality assumptions

# References

Andrew J. G. Cairns, David P. Blake, Amy Kessler, and Marsha Kessler. The impact of COVID-19 on future higher-age mortality. Available at `dx.doi.org/10.2139/ssrn.3606988`, 2020.

J. Robben, K. Antonio, and S. Devriendt. Assessing the impact of the COVID-19 shock on a stochastic multi-population mortality model. *Risks*, 10(2):26, 2022.

S. Schnürch, T. Kleinow, R. Korn, and A. Wagner. The impact of mortality shocks on modelling and insurance valuation as exemplified by COVID-19. *Annals of Actuarial Science*, 16(3):498–526, 2022.

F. van Berkum, B. Melenberg, and M.H. Vellekoop. Estimating the impact of COVID-19 on mortality using granular data. Available at `https://doi.org/10.48550/arXiv.2209.06473`, 2022.

# Appendix - additional sensitivity analysis

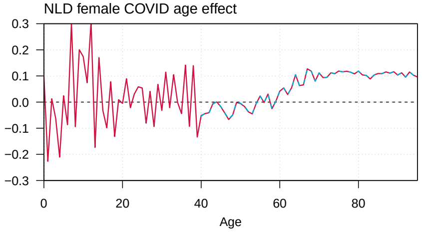
## Different age ranges for calibration

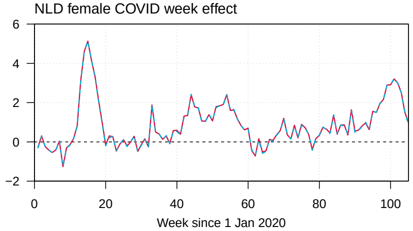
The COVID age effects are volatile at lower ages, so we might consider excluding these ages from the calibration

We consider the age ranges (0-95) and (40-95) and compare the resulting parameter estimates:

- For the ages that are included in both age ranges, the COVID age effects are near identical (after appropriate rescaling).
- The COVID week effects are also near identical.

We conclude that mortality at lower ages is hardly affected by COVID, and these ages are therefore excluded from further analyses.



# Appendix - additional sensitivity analysis

## Comparison of five countries

Using weekly mortality data on five-year age groups from STMF, we compare the COVID parameter estimates for five countries

While accounting for the seasonal effect up front, we observe:

- COVID waves occurred at different points in time due to different measures taken by governments
- COVID age effects are quite similar for NLD, DEU, FRA and BEL, but GBR is markedly different with age effect substantially different from 0 at lower ages
- Volatile COVID age effects at higher ages are probably caused by allocation from age groups to individual ages

*Graphs for males are similar.*