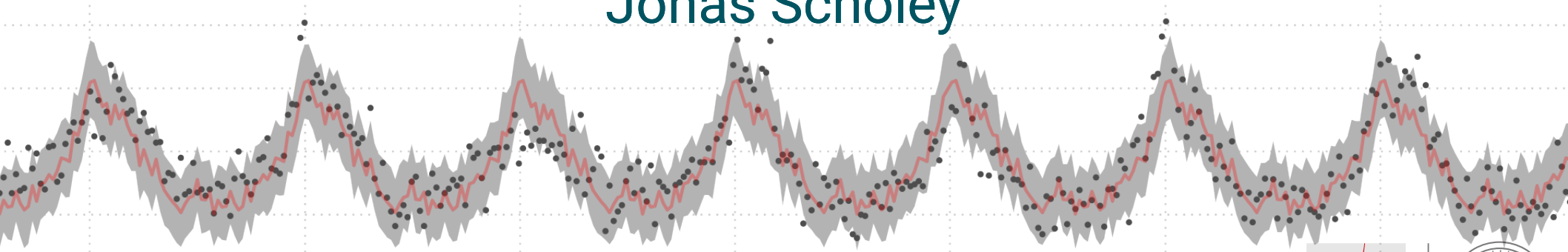
Robustness and **bias** of **excess death estimates** under varying model specifications

Jonas Schöley

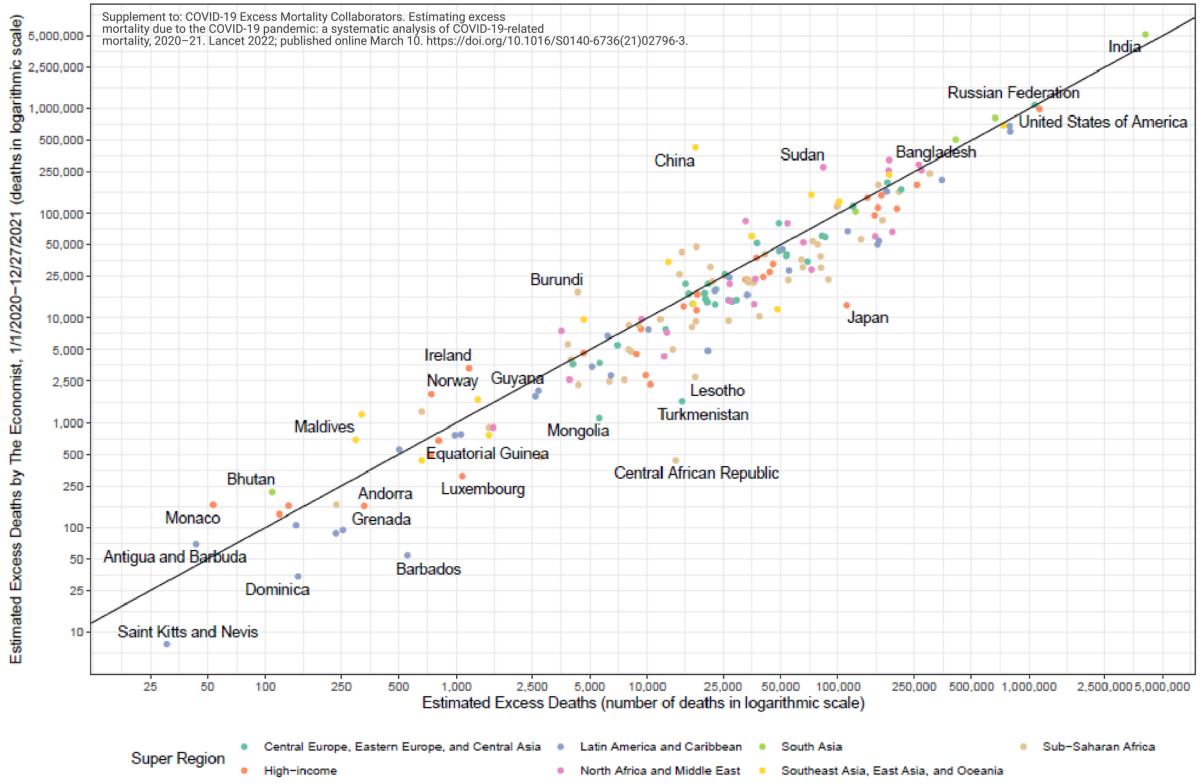


@jschoeley
0000-0002-3340-8518
schoeley@demogr.mpg.de

MAX PLANCK INSTITUTE
FOR DEMOGRAPHIC RESEARCH

Excess disagreement

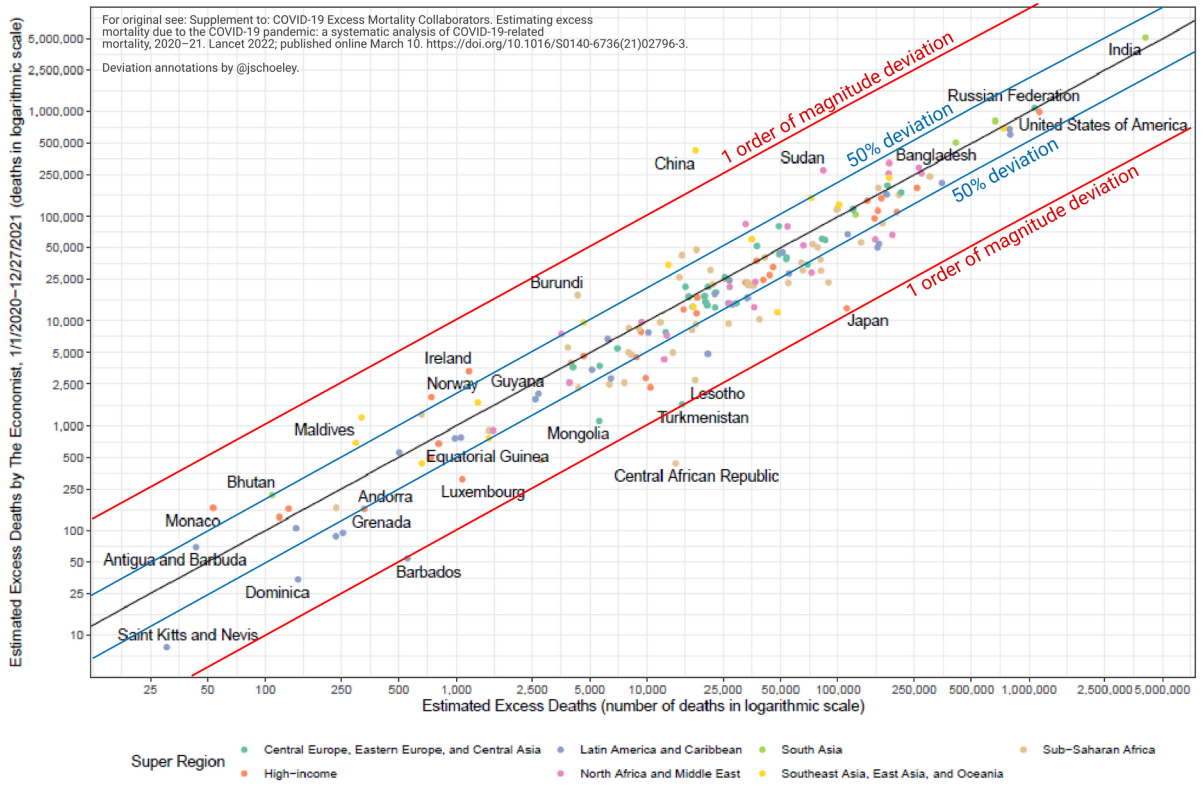
Figure S5. Comparison to *The Economist* excess mortality estimates



10.1016/S0140-6736(21)02796-3

Excess disagreement

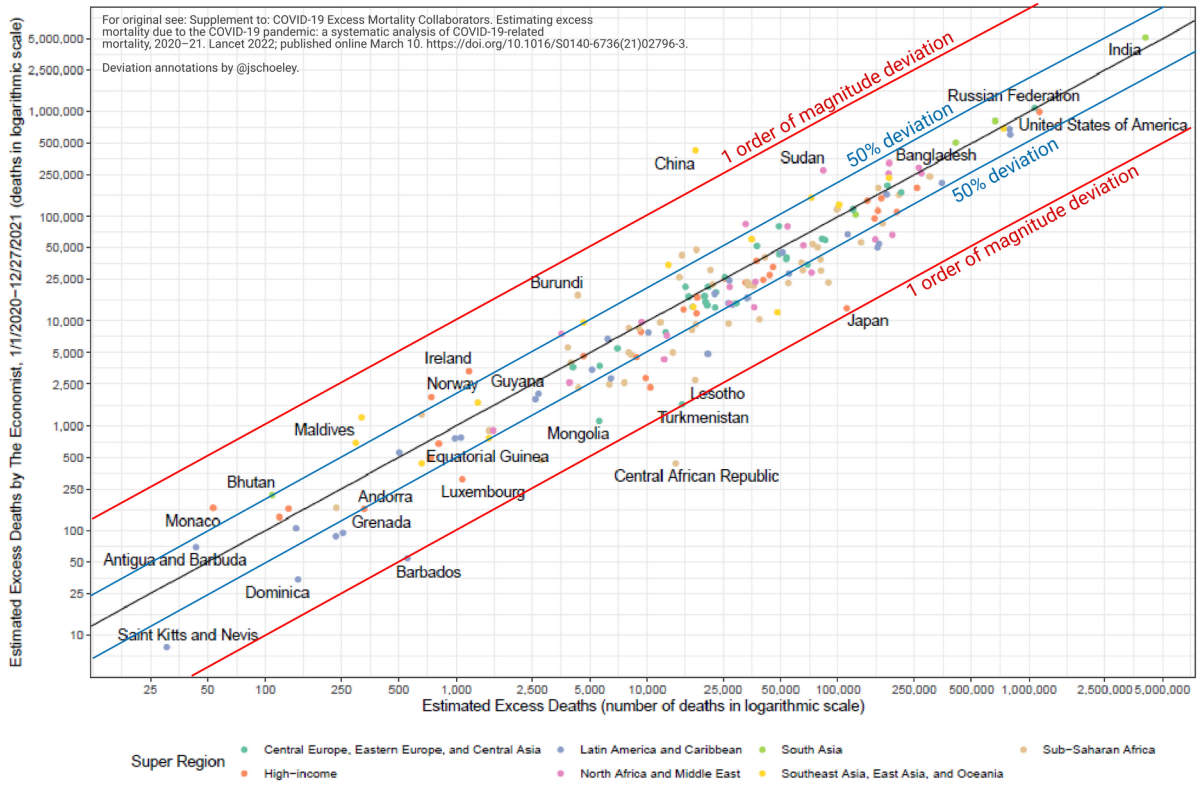
Figure S5. Comparison to *The Economist* excess mortality estimates



10.1016/S0140-6736(21)02796-3

Excess disagreement

Figure S5. Comparison to *The Economist* excess mortality estimates

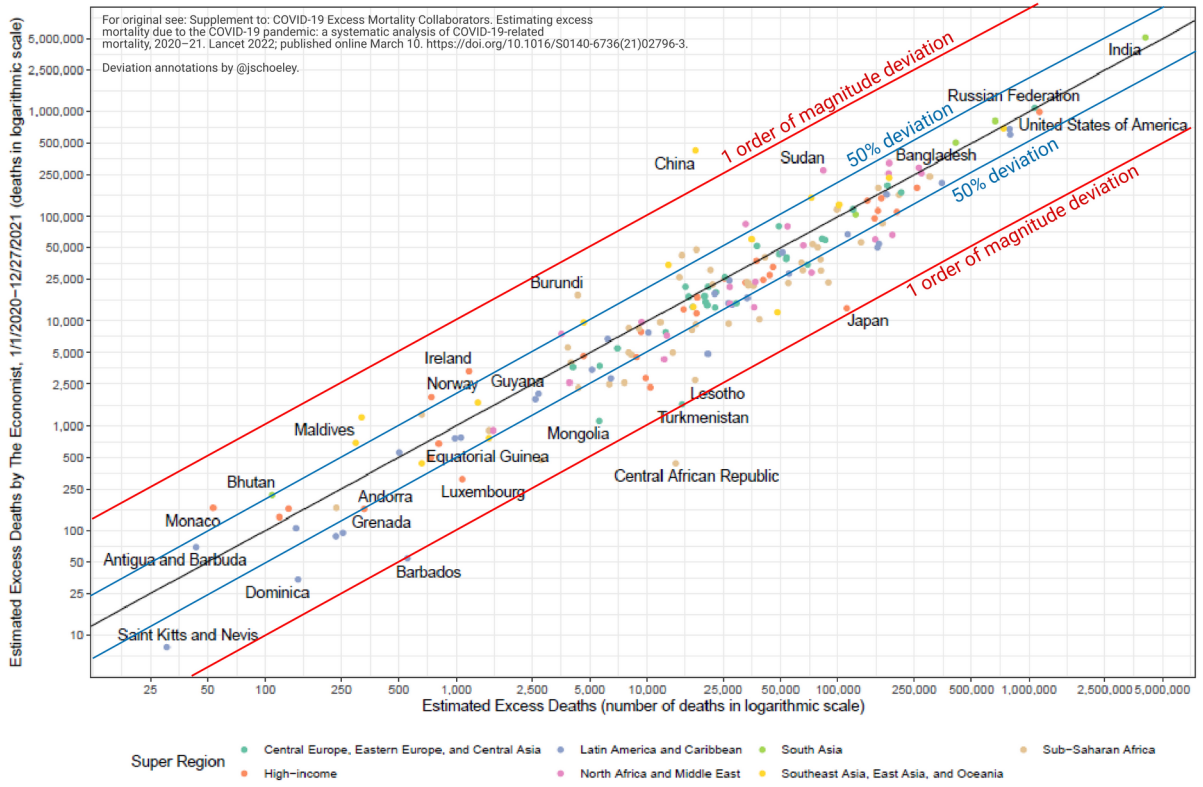


Robustness does model specification change our conclusions on excess deaths?

10.1016/S0140-6736(21)02796-3

Excess disagreement

Figure S5. Comparison to *The Economist* excess mortality estimates

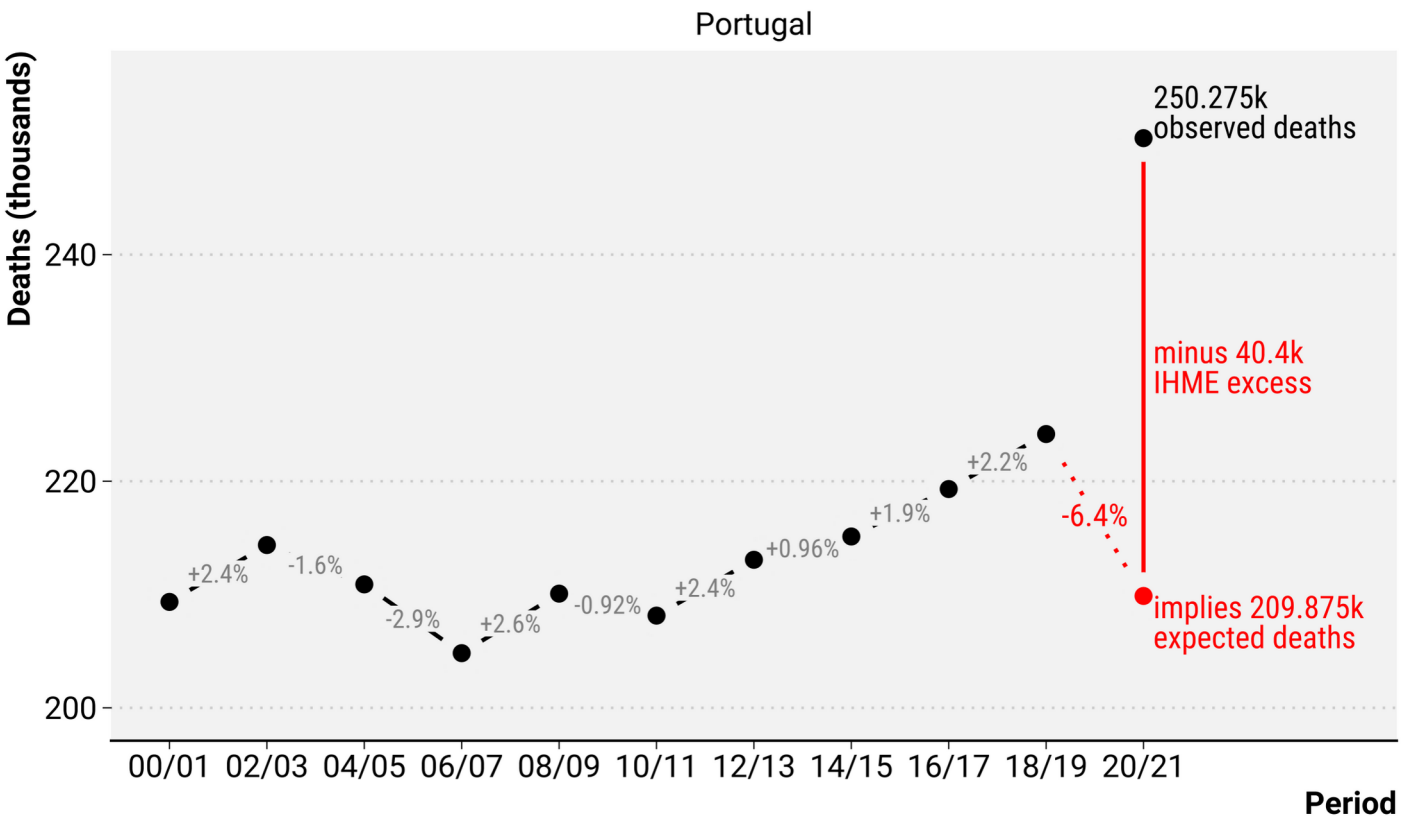


Robustness does model specification change our conclusions on excess deaths?

Bias Are some models prone to predict too many, too few excess deaths?

10.1016/S0140-6736(21)02796-3

Excess disagreement



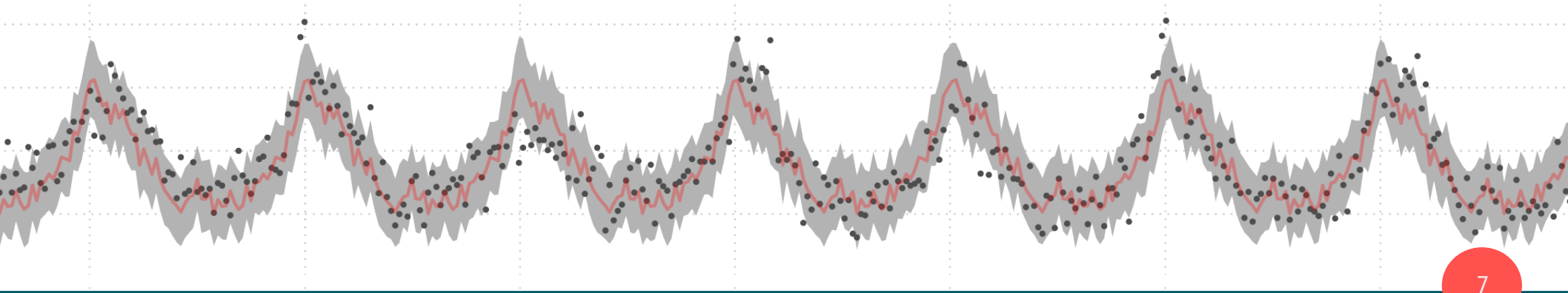
Robustness does model specification change our conclusions on excess deaths?

Bias Are some models prone to predict too many, too few excess deaths?

What should we expect?

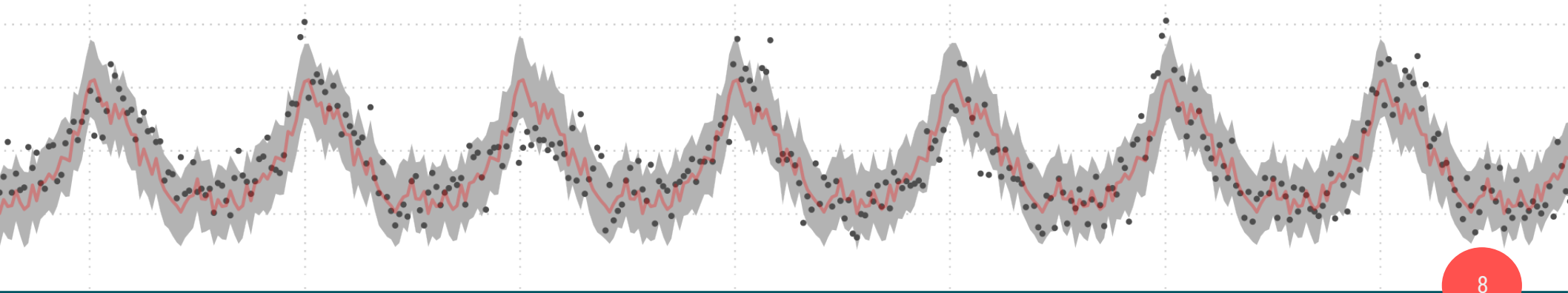
How many deaths should we expect in the absence of COVID-19?

Excess deaths =
Observed deaths -
Expected deaths



Expected deaths

$$D_i \sim \text{Pois}(\lambda_i, \phi)$$
$$\lambda_i = \exp(\eta_i),$$



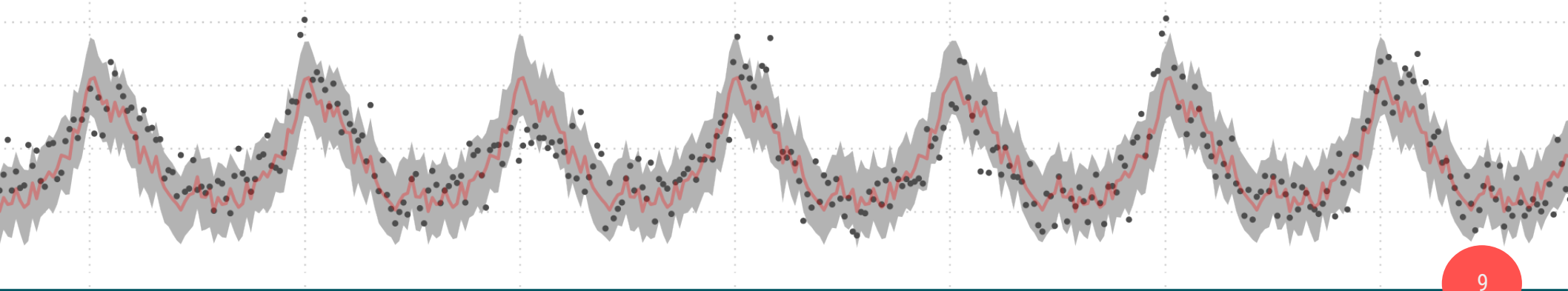
A zoo of models

Expected deaths

- (time-varying) seasonality
- population structure
- time trend
- temperature
- temporal auto-correlation
- public holidays

$$D_i \sim \text{Pois}(\lambda_i, \phi)$$

$$\lambda_i = \exp(\eta_i),$$



A zoo of models

Expected deaths

- (time-varying) **seasonality**
- population structure
- time trend
- temperature
- temporal auto-correlation
- public holidays

Table 1: Models for weekly expected deaths.

Model	Description	References
Weekly averages (AVGc5) 5 year average death counts (AVGr5) 5 year average death rates	Average over the weekly deaths counts or rates in the preceding 5 years implemented as quasi-Poisson GLM with week-of-year coefficients	[20], [29], [5]
Serfling Model (SRFc) without exposures (SRFcem) Euromomo style, i.e. no exposures, fitted only on weeks without flu-activity (SRFr) with exposures	Quasi-Poisson regression on death counts with log-linear long term trend, AIC selected Fourier-term seasonality, and public-holiday coefficients	[3], [31], [32], [9], [12]
Generalized Additive Model (GAMr) without temperature anomaly predictor (GAMrt) with temperature anomaly predictor as smoothly varying coefficient over week of year	Quasi-Poisson regression on death counts with log-linear long term trend, penalized cyclical spline seasonality, and public-holiday coefficients	[1], [26]
Latent Gaussian Model (LGMr) without temperature anomaly predictor (LRMrt) with temperature anomaly predictor as varying coefficient over week of year implemented as cyclic random walk	Bayesian Poisson regression on death counts with autoregressive trend, non-parametric time-varying seasonality, and public-holiday coefficients	[14]

10.1101/2021.06.04.21258353

A zoo of models

Expected deaths

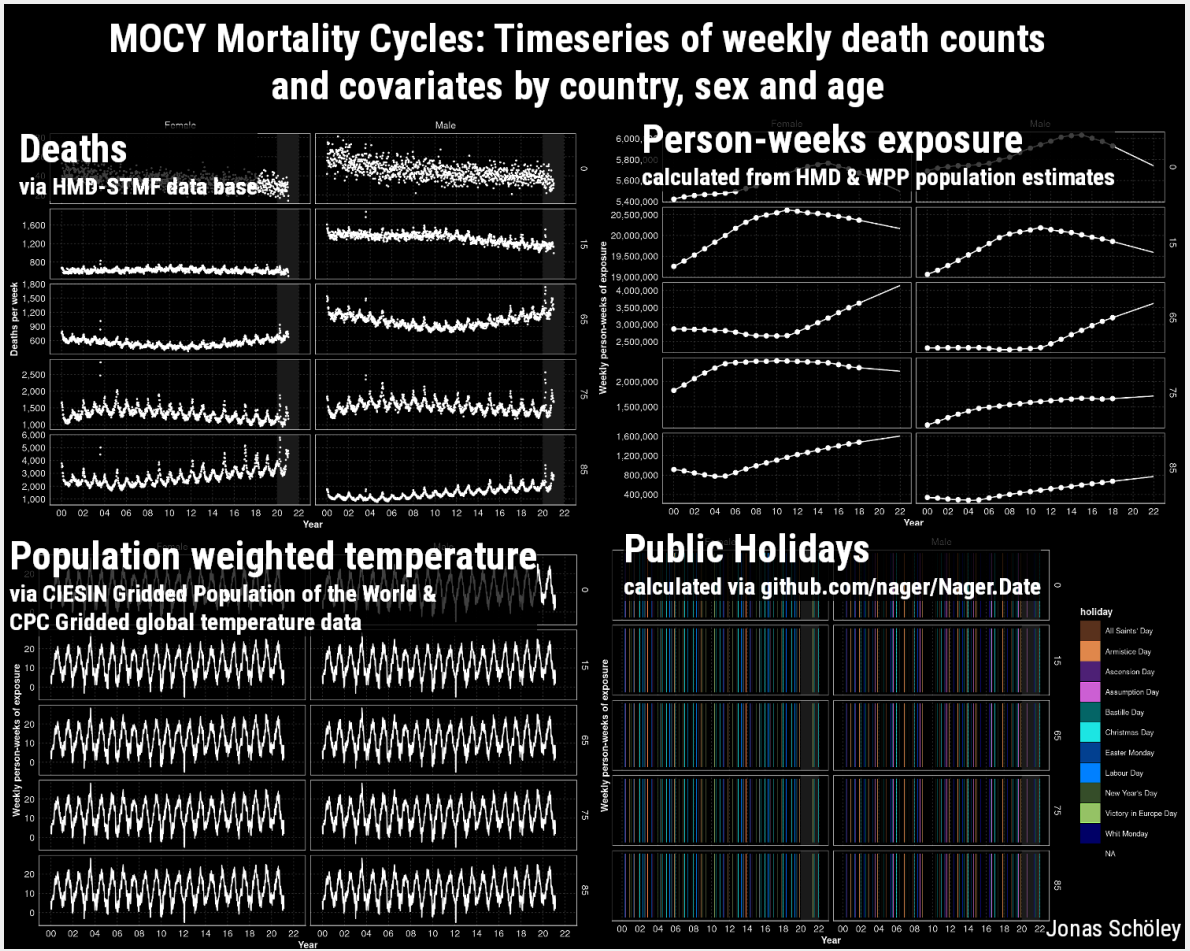
- (time-varying) seasonality
- population structure
- time trend
- temperature
- temporal auto-correlation
- public holidays

Table 1: Models for weekly expected deaths.

Model	Description	References
Weekly averages (AVGc5) 5 year average death counts (AVGr5) 5 year average death rates	Average over the weekly deaths counts or rates in the preceding 5 years implemented as quasi-Poisson GLM with week-of-year coefficients	[20], [29], [5]
Serfling Model (SRFc) without exposures (SRFcem) Euromomo style, i.e. no exposures, fitted only on weeks without flu-activity (SRFr) with exposures	Quasi-Poisson regression on death counts with log-linear long term trend, AIC selected Fourier-term seasonality, and public-holiday coefficients	[3], [31], [32], [9], [12]
Generalized Additive Model (GAMr) without temperature anomaly predictor (GAMrt) with temperature anomaly predictor as smoothly varying coefficient over week of year	Quasi-Poisson regression on death counts with log-linear long term trend, penalized cyclical spline seasonality, and public-holiday coefficients	[1], [26]
Latent Gaussian Model (LGMr) without temperature anomaly predictor (LRMrt) with temperature anomaly predictor as varying coefficient over week of year implemented as cyclic random walk	Bayesian Poisson regression on death counts with autoregressive trend, non-parametric time-varying seasonality, and public-holiday coefficients	[14]

10.1101/2021.06.04.21258353

The data



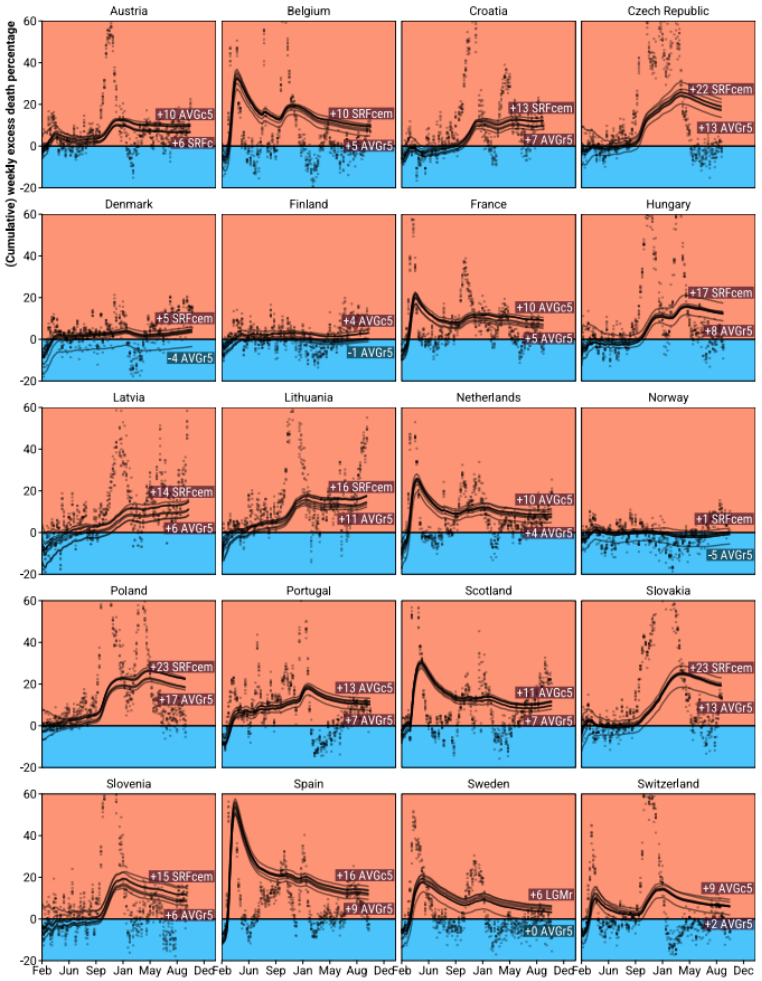
Given timeseries on weekly death counts and covariates predict:

- weekly expected death counts March 2020 to Nov 2021
- across 20 European countries
- by sex and age group (0-65, 65-75, 75-85, 85+)

Download from osf.io/k84rz/

Robustness

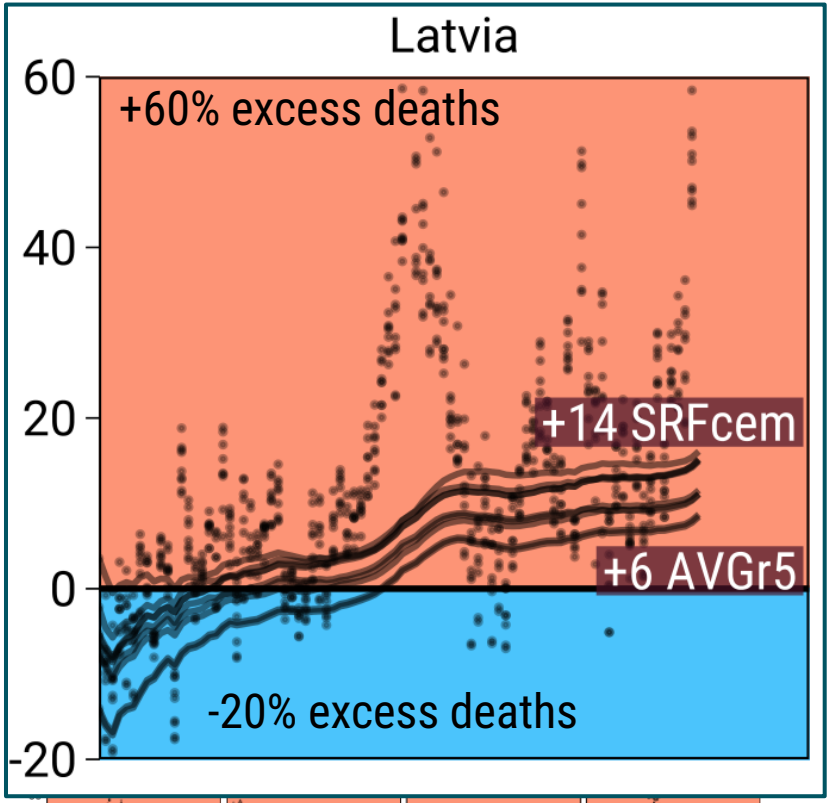
Robustness



Weekly and cumulative excess death percentage since March 2020 as predicted from 9 different models

10.1101/2021.06.04.21258353

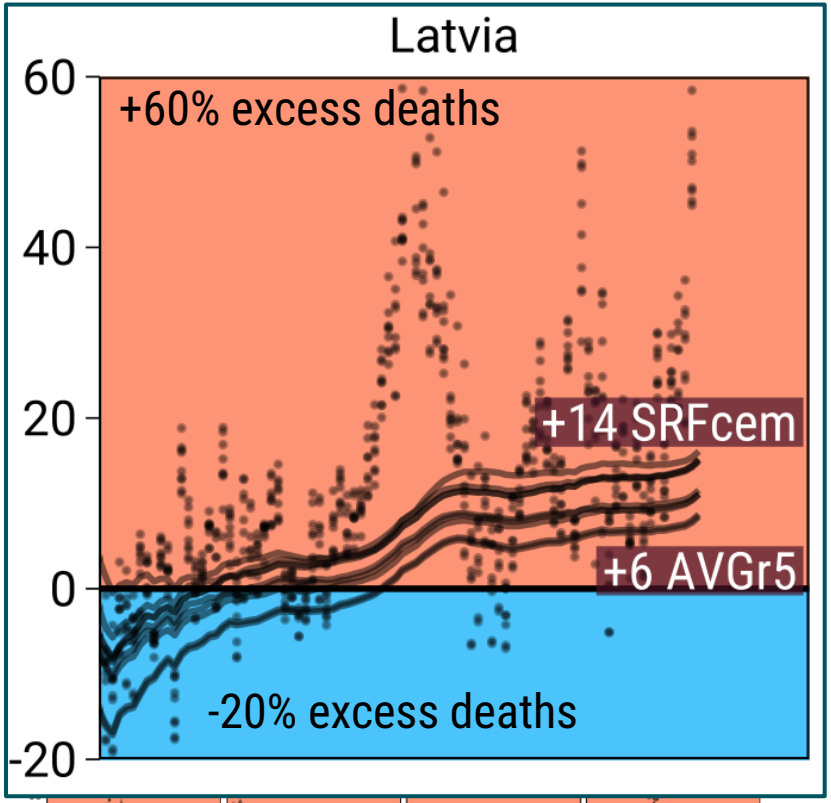
Robustness



Weekly and cumulative excess death percentage as predicted from 9 different models during the year 2020.

10.1101/2021.06.04.21258353

Robustness

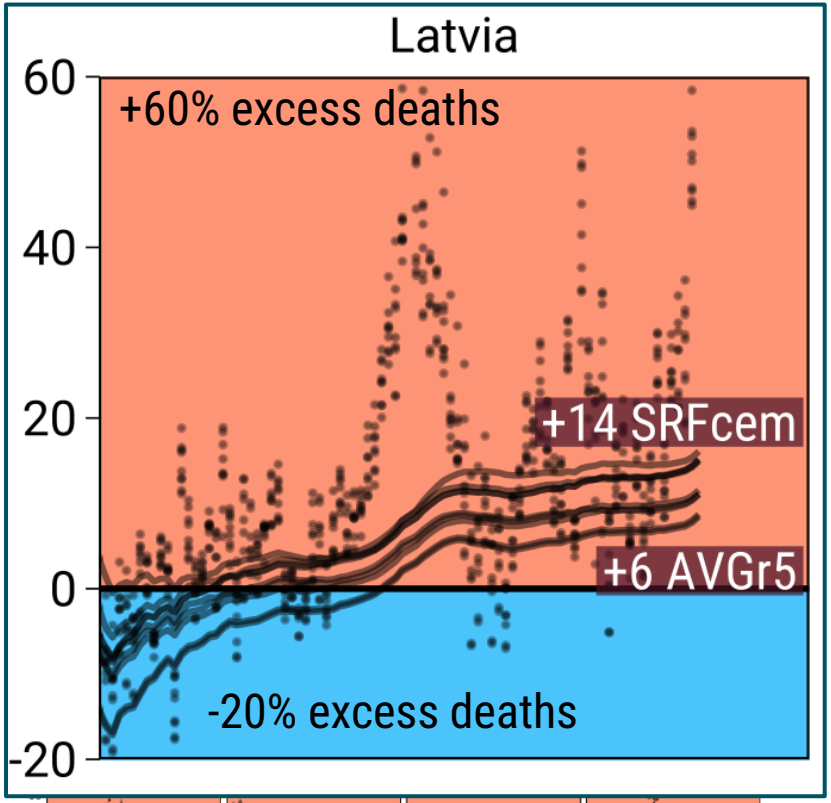


Weekly and cumulative excess death percentage as predicted from 9 different models during the year 2020.

- models **agree on weekly pattern** of excess but **disagree on level**

10.1101/2021.06.04.21258353

Robustness

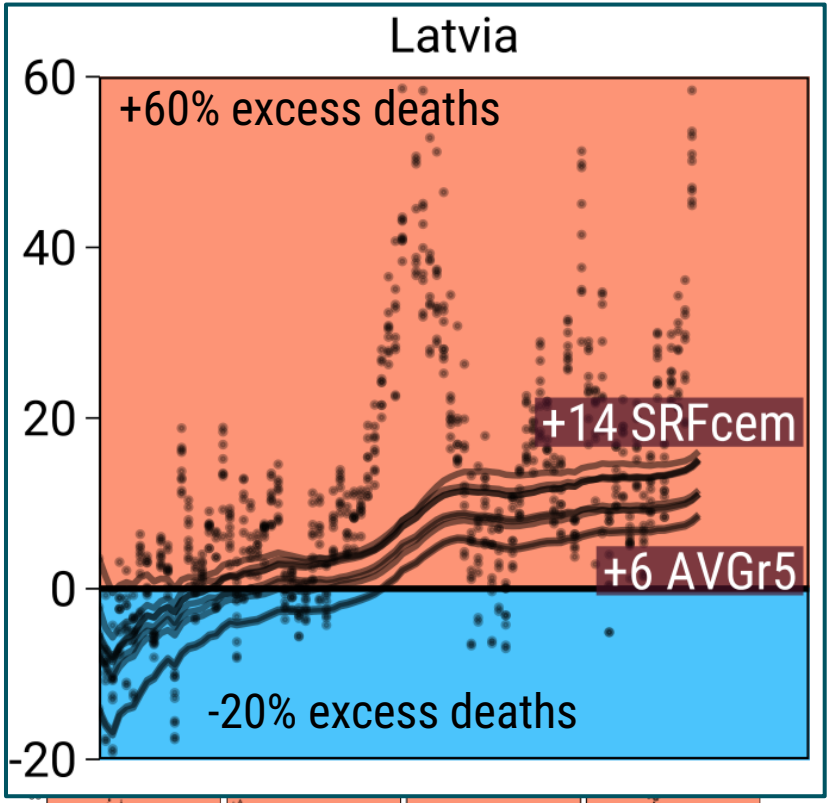


Weekly and cumulative excess death percentage as predicted from 9 different models during the year 2020.

- models **agree on weekly pattern** of excess but **disagree on level**
- model **disagreement increases over time**

10.1101/2021.06.04.21258353

Robustness

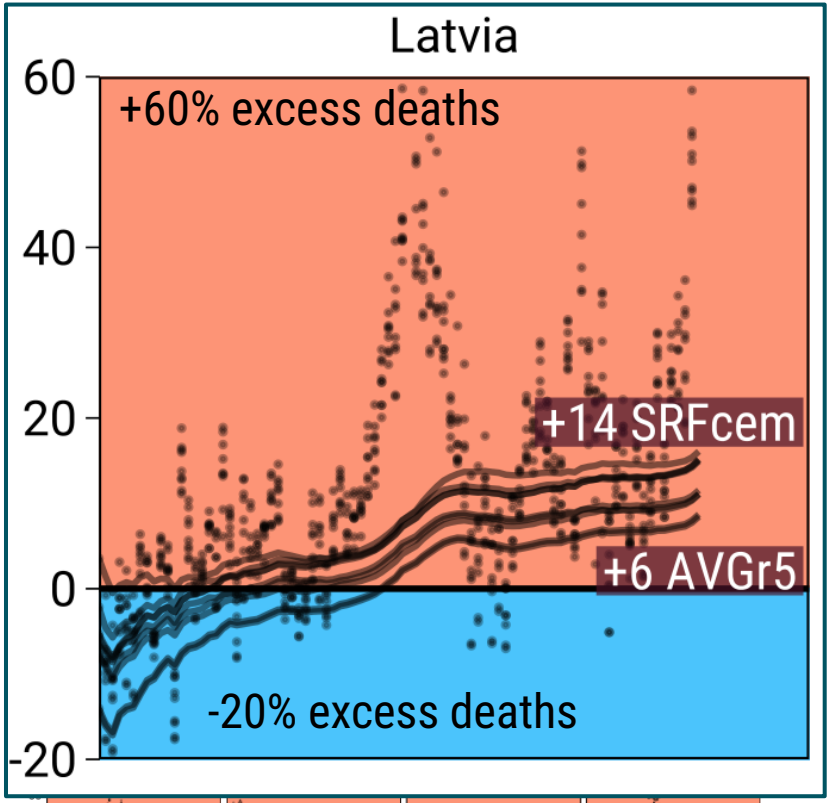


Weekly and cumulative excess death percentage as predicted from 9 different models during the year 2020.

- models **agree on weekly pattern** of excess but **disagree on level**
- model **disagreement increases over time**
- **5-year average death rates** tend towards **low excess**

10.1101/2021.06.04.21258353

Robustness

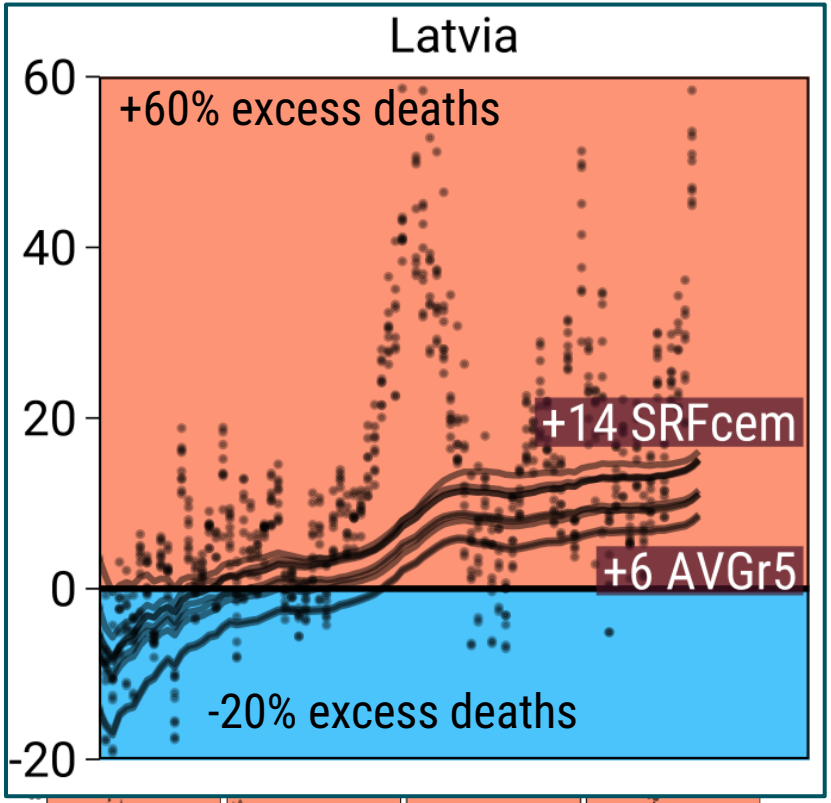


Weekly and cumulative excess death percentage as predicted from 9 different models during the year 2020.

- models **agree on weekly pattern** of excess but **disagree on level**
- model **disagreement increases over time**
- **5-year average death rates** tend towards **low excess**
- **Serfling models** tend towards **high excess**

10.1101/2021.06.04.21258353

Robustness

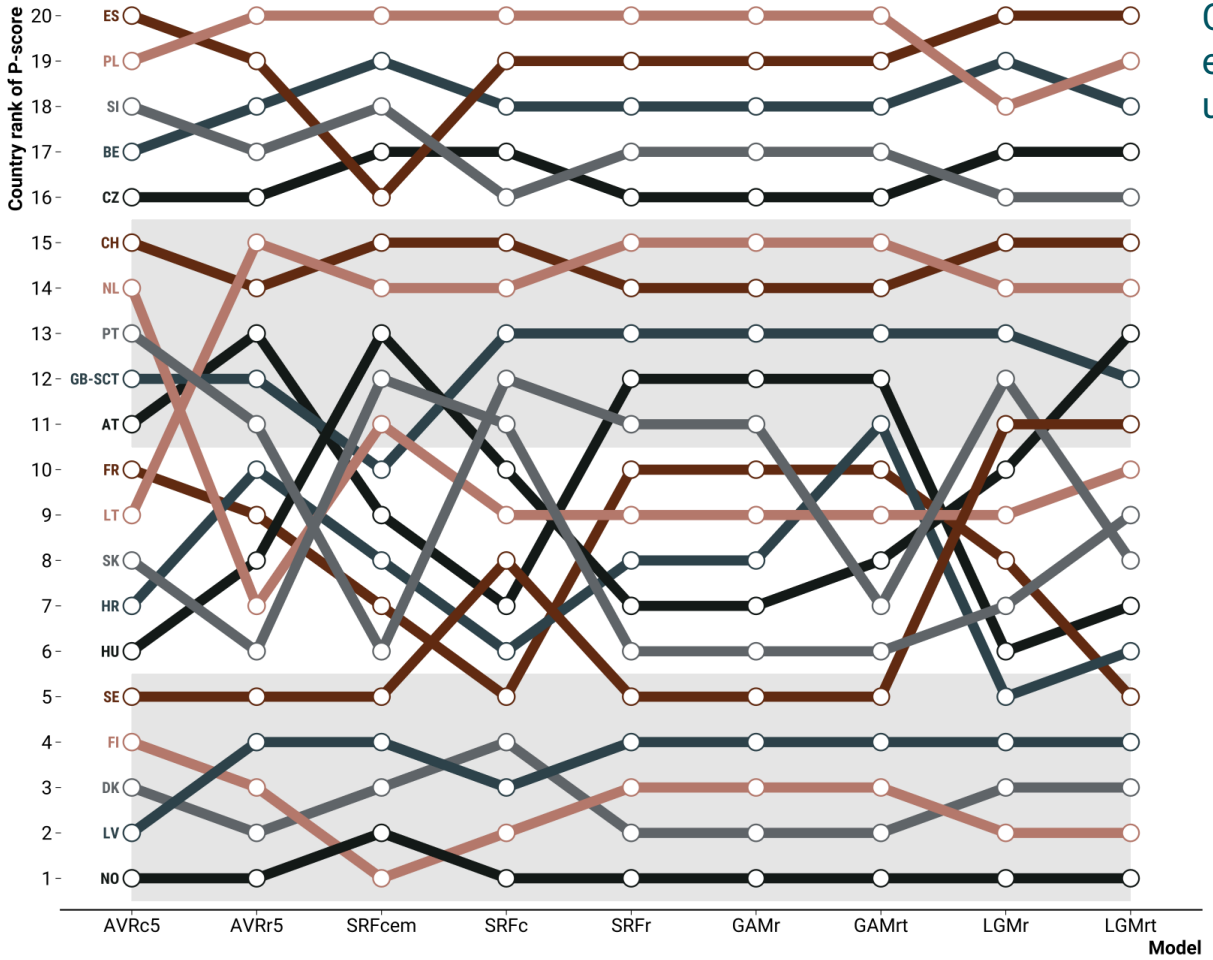


Weekly and cumulative excess death percentage as predicted from 9 different models during the year 2020.

- models **agree on weekly pattern** of excess but **disagree on level**
- model **disagreement increases over time**
- **5-year average death rates** tend towards **low excess**
- **Serfling models** tend towards **high excess**
- **broad agreement** about **existence of** significantly elevated **annual excess deaths**. Exceptions: Denmark, Finland, Latvia and Norway

10.1101/2021.06.04.21258353

Robustness

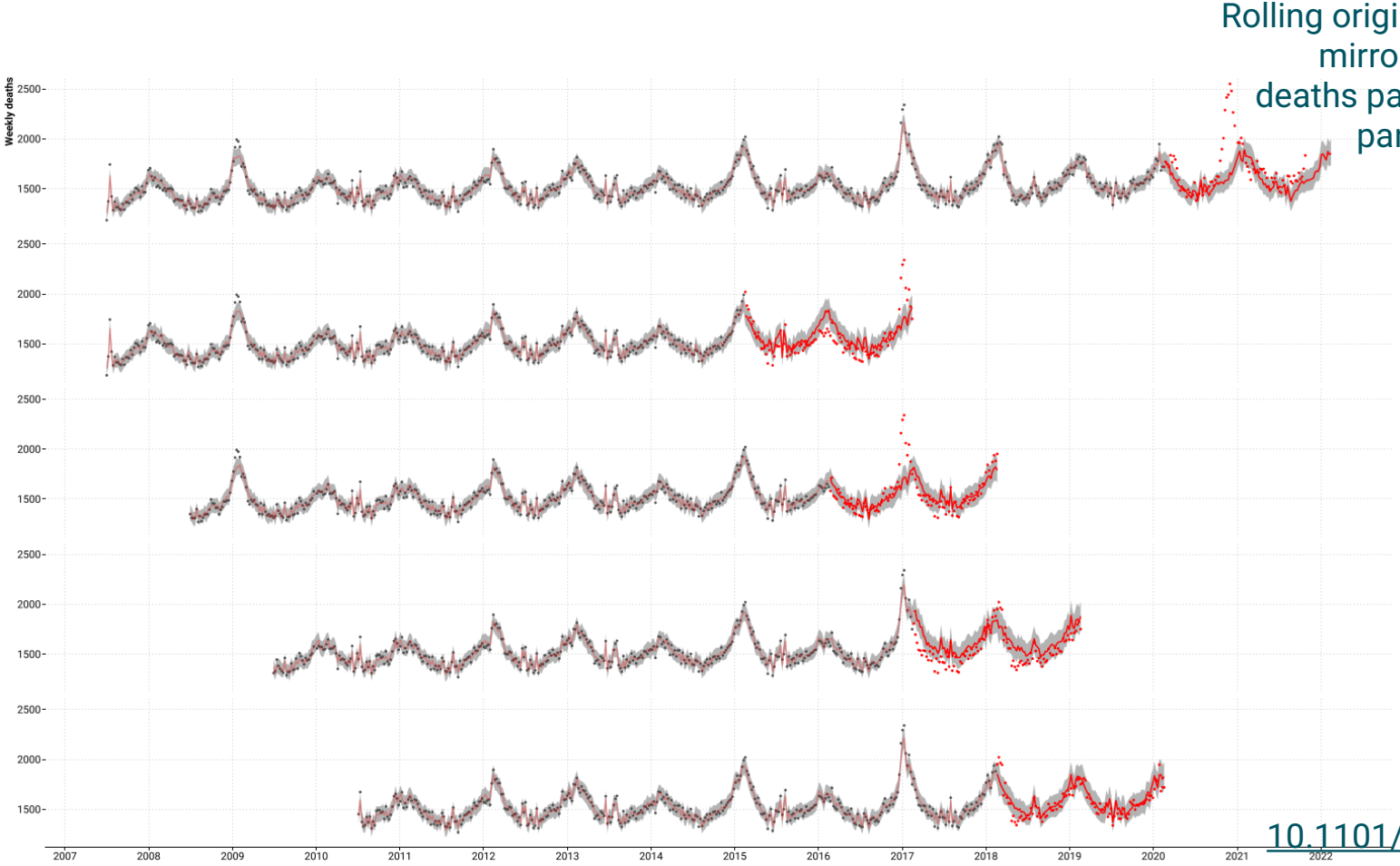


Country rank of percent excess deaths in 2020 under different models

10.1101/2021.06.04.21258353

Error & Bias

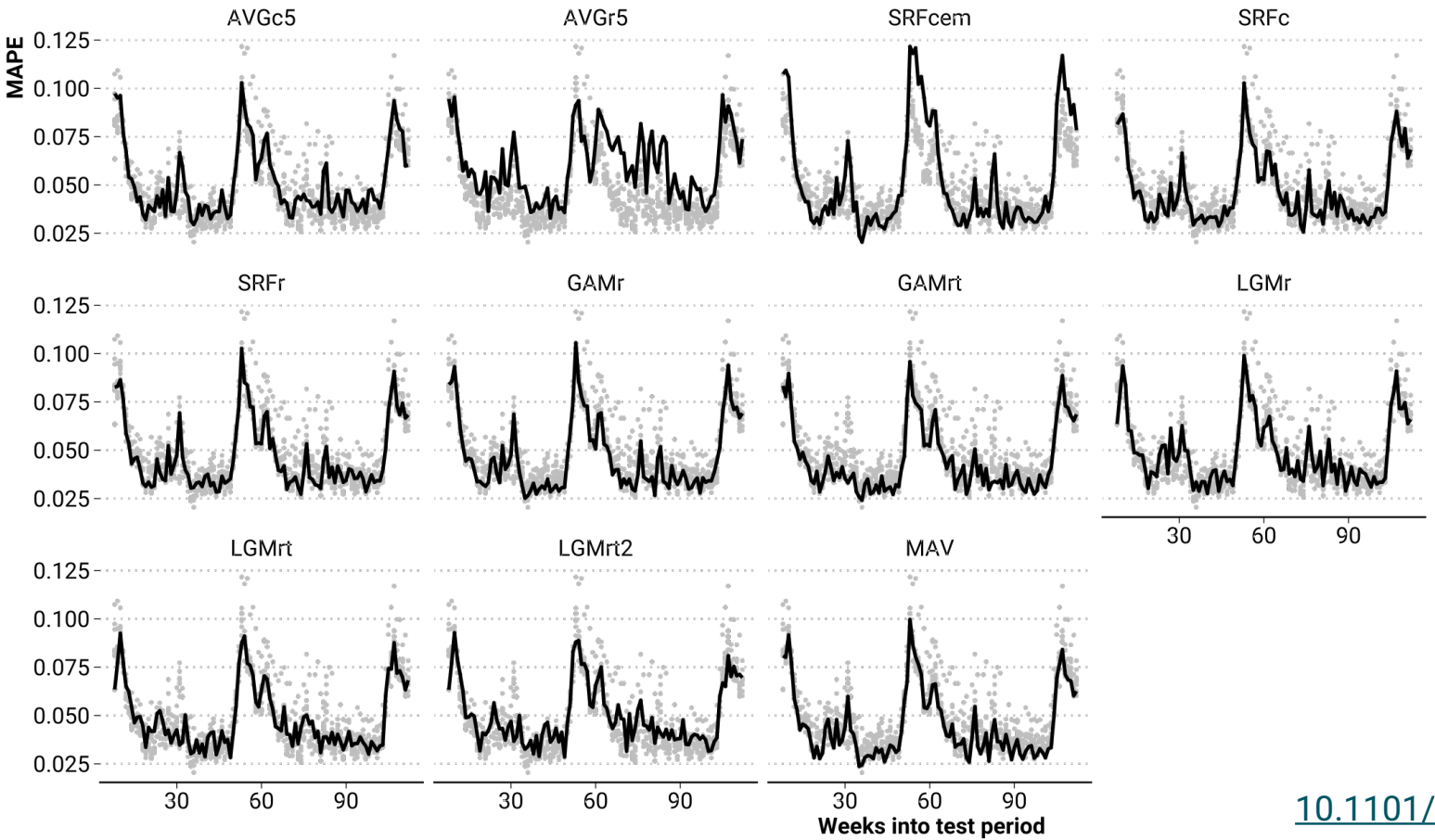
Error & Bias



Rolling origin four-fold cross-validation setup mirroring the task of predicting weekly deaths past the beginning of the COVID-19 pandemic given pre-pandemic data.

10.1101/2021.06.04.21258353

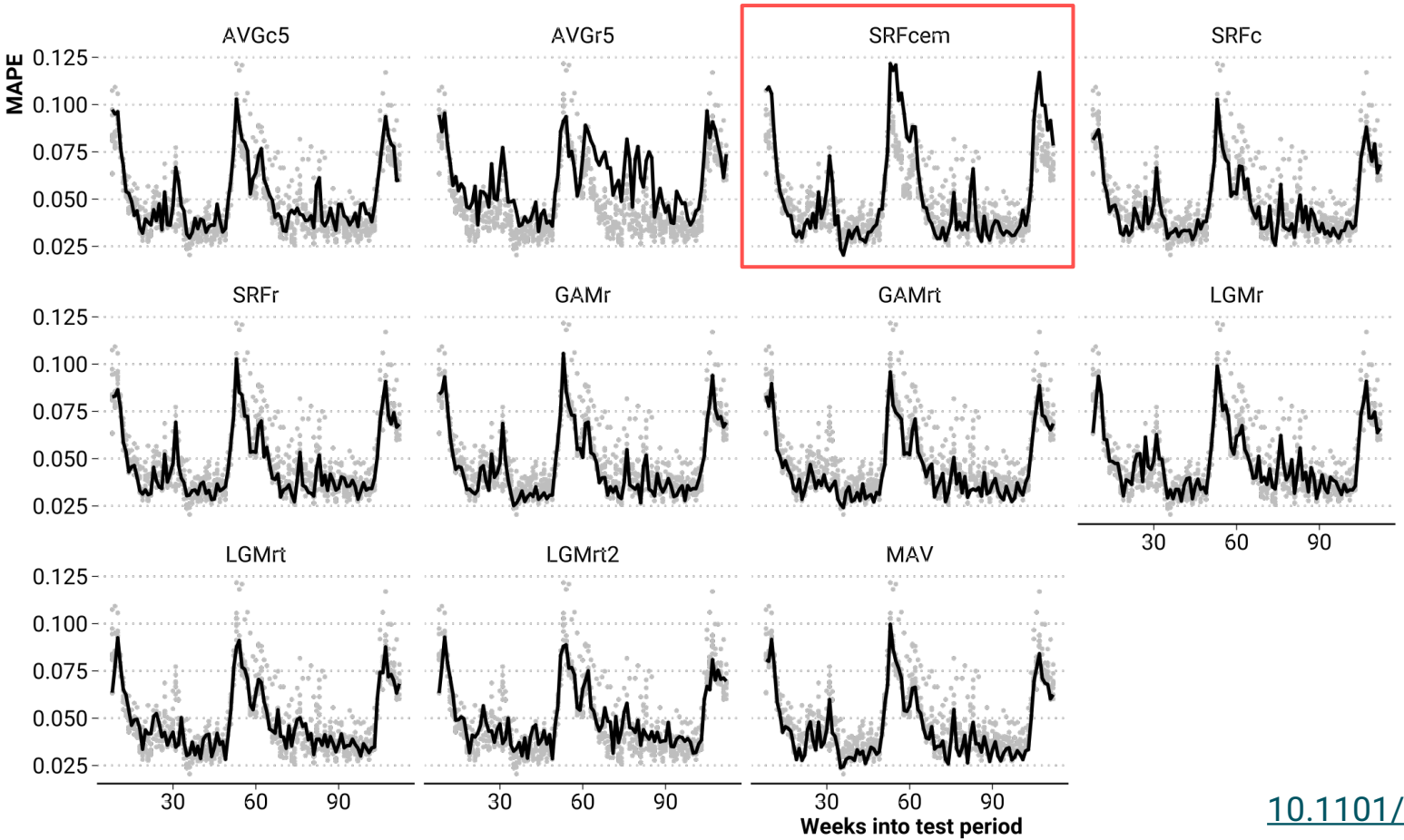
Error & Bias



Mean absolute percentage error of weekly death counts for different models. Grey points show MAPE of all other models

10.1101/2021.06.04.21258353

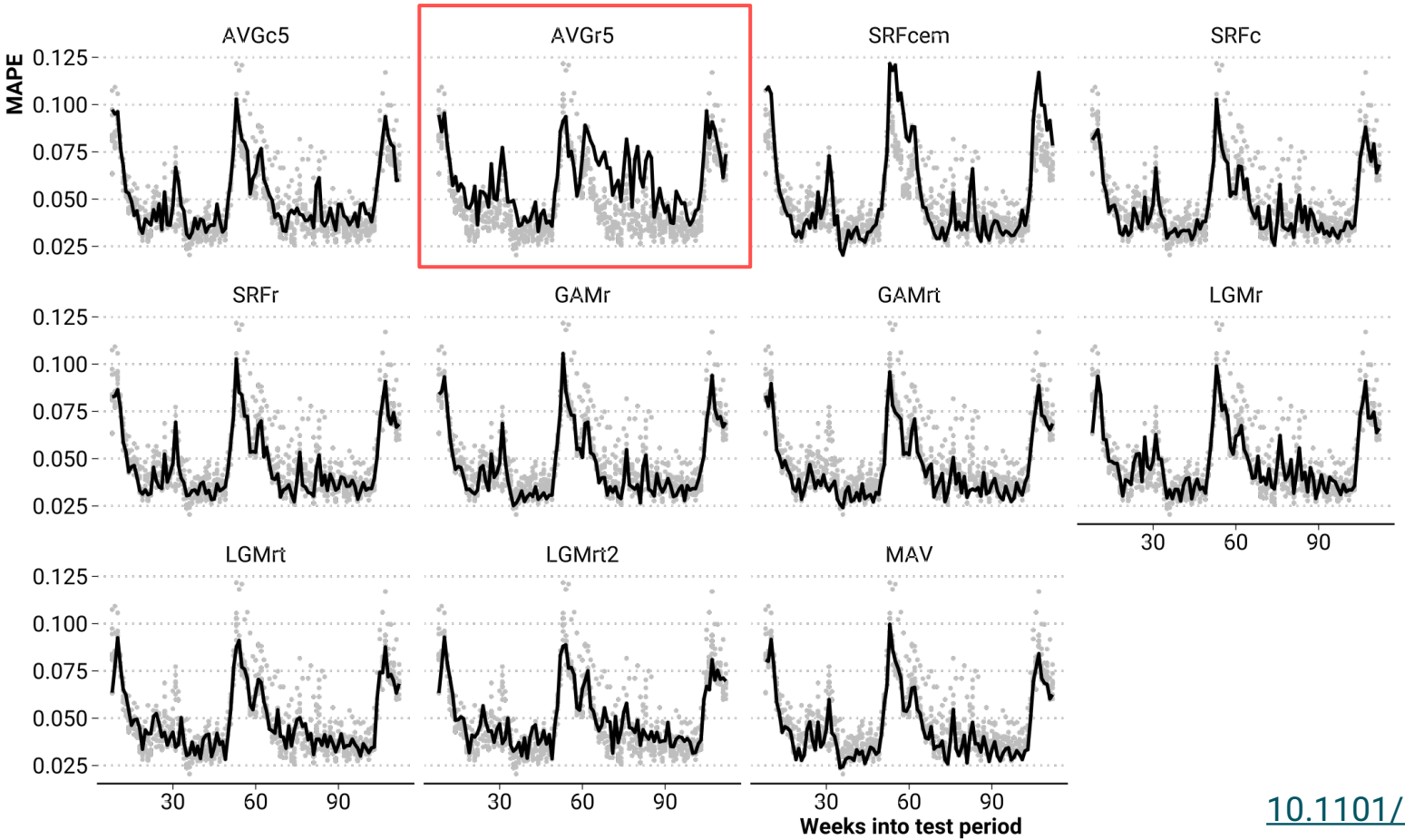
Error & Bias



Mean absolute percentage error of weekly death counts for different models. Grey points show MAPE of all other models

10.1101/2021.06.04.21258353

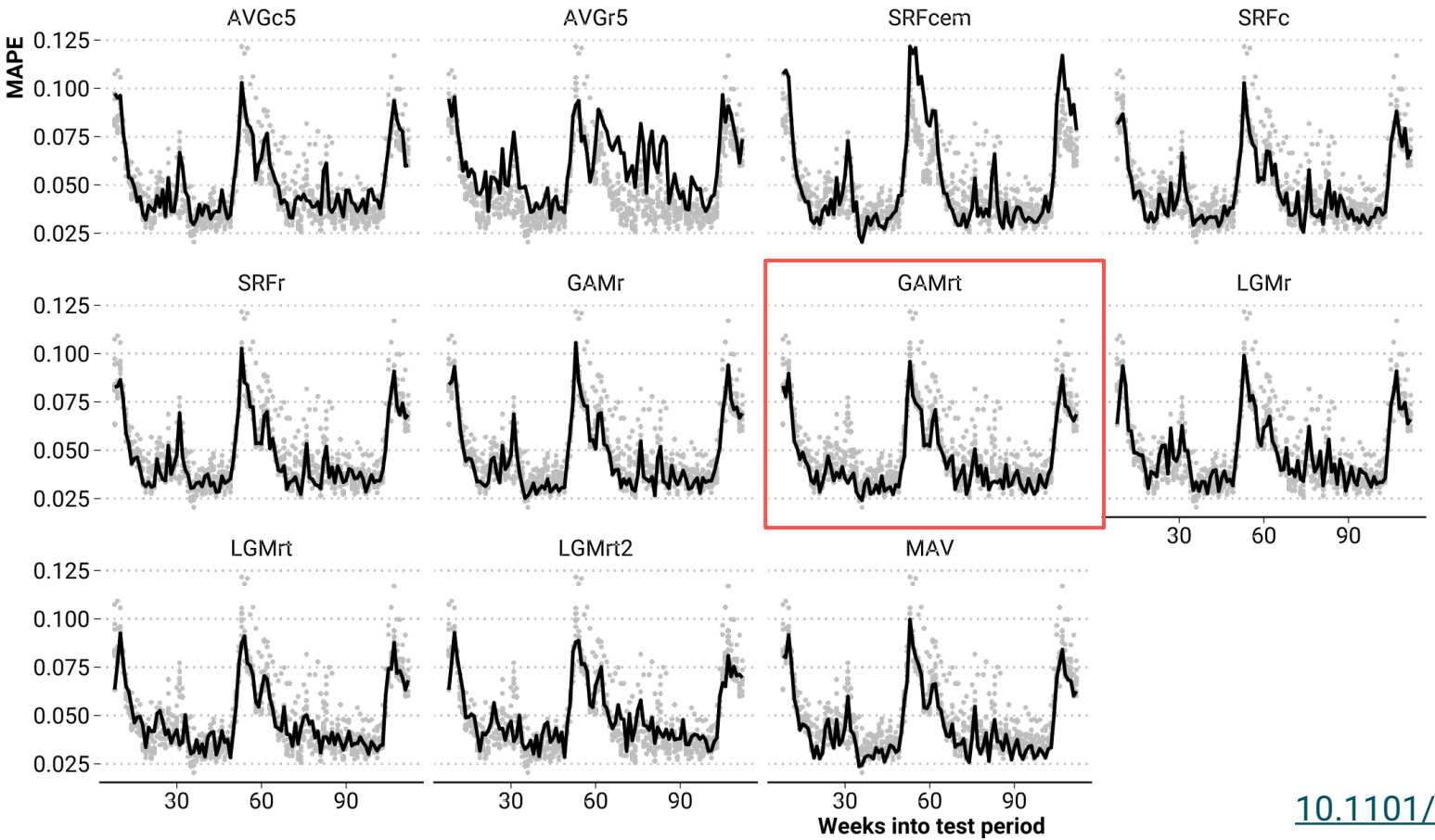
Error & Bias



Mean absolute percentage error of weekly death counts for different models. Grey points show MAPE of all other models

10.1101/2021.06.04.21258353

Error & Bias

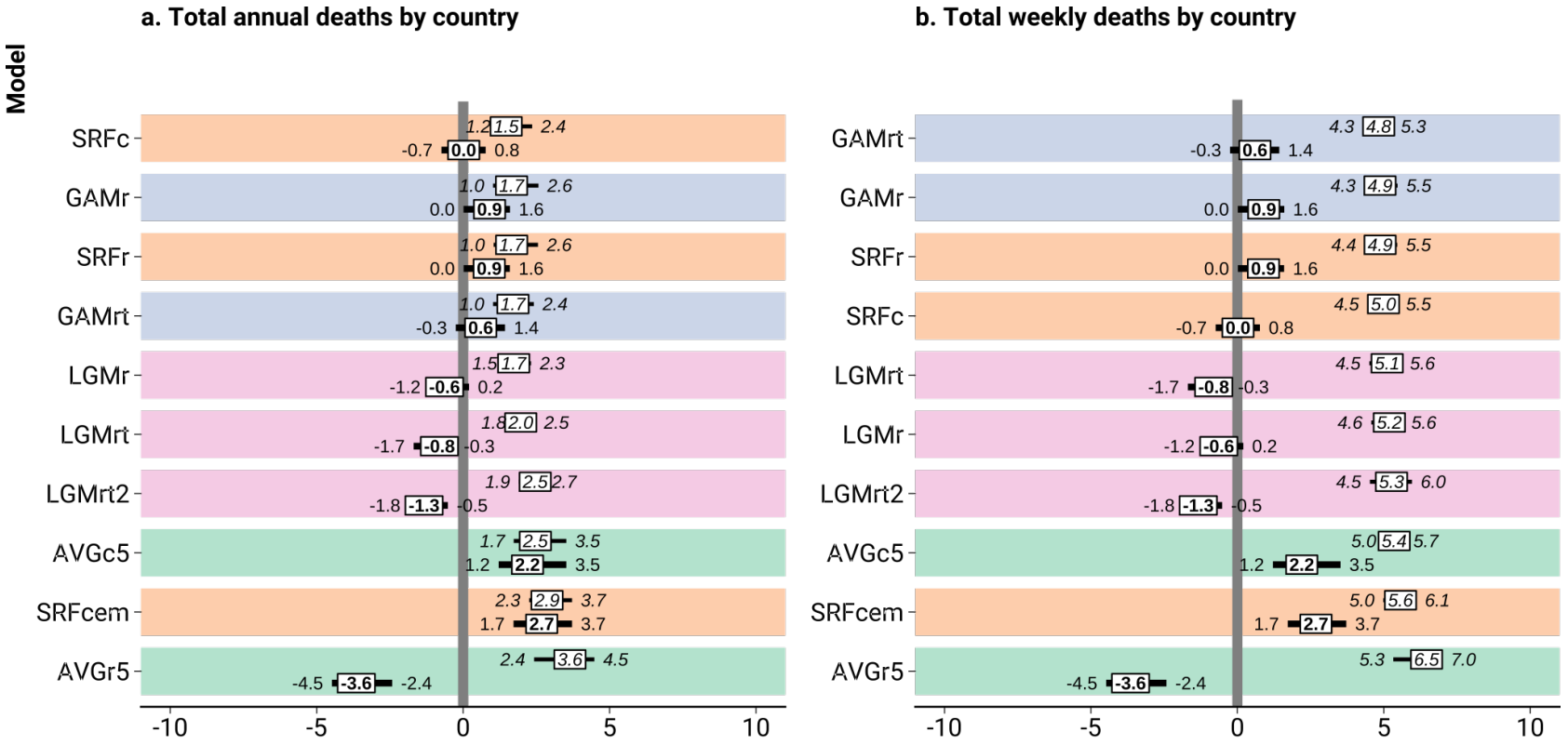


Mean absolute percentage error of weekly death counts for different models. Grey points show MAPE of all other models

10.1101/2021.06.04.21258353

Error & Bias

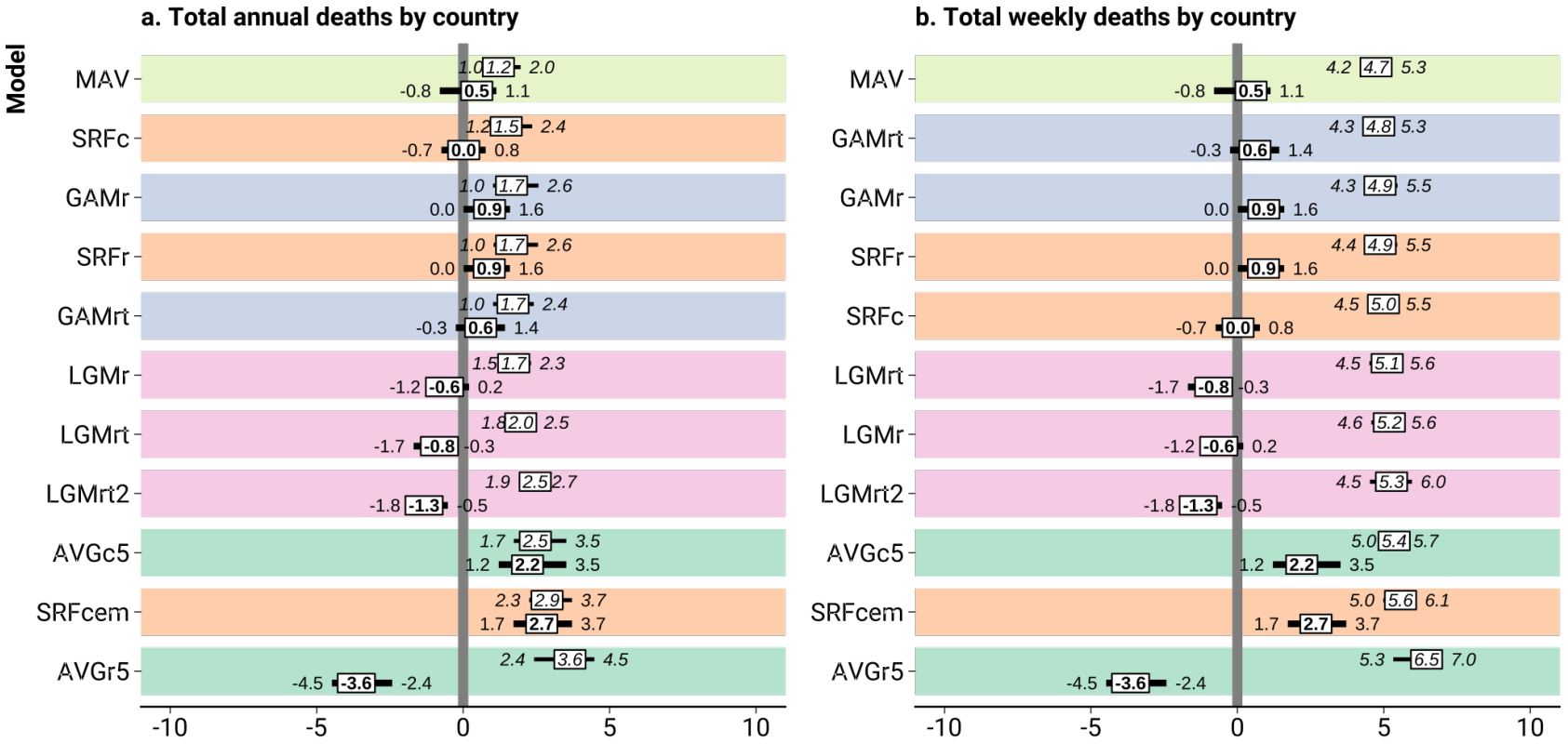
Bias (MPE, bold) and error (MAPE, italic) by model when predicting death counts on test data.



10.1101/2021.06.04.21258353

Error & Bias

Bias (MPE, bold) and error (MAPE, italic) by model when predicting death counts on test data.



10.1101/2021.06.04.21258353

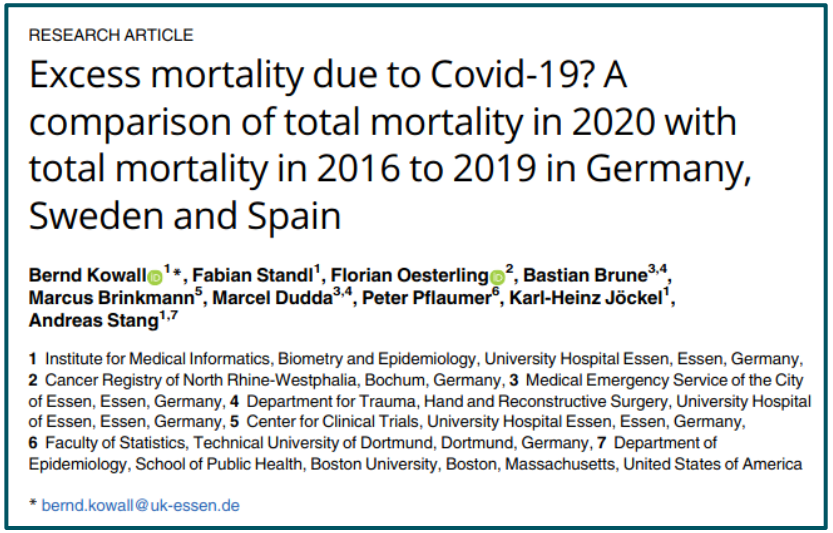
Suggestions

Suggestions

Provide **robustness** checks for
country rankings of excess deaths

Suggestions

Avoid the **average death rate** model



RESEARCH ARTICLE

Excess mortality due to Covid-19? A comparison of total mortality in 2020 with total mortality in 2016 to 2019 in Germany, Sweden and Spain

Bernd Kowall[1]*, Fabian Standl[1], Florian Oesterling[2], Bastian Brune[3,4], Marcus Brinkmann[5], Marcel Dudda[3,4], Peter Pflaumer[6], Karl-Heinz Jöckel[1], Andreas Stang[1,7]

1 Institute for Medical Informatics, Biometry and Epidemiology, University Hospital Essen, Essen, Germany, **2** Cancer Registry of North Rhine-Westphalia, Bochum, Germany, **3** Medical Emergency Service of the City of Essen, Essen, Germany, **4** Department for Trauma, Hand and Reconstructive Surgery, University Hospital of Essen, Essen, Germany, **5** Center for Clinical Trials, University Hospital Essen, Essen, Germany, **6** Faculty of Statistics, Technical University of Dortmund, Dortmund, Germany, **7** Department of Epidemiology, School of Public Health, Boston University, Boston, Massachusetts, United States of America

* bernd.kowall@uk-essen.de

Suggestions

For a **simple** and **sensitive** **robustness check** present excess estimates under the **5-year average death rate** model and a **Euromomo style Serfling model**

Put in appendix only!

Suggestions

Don’t sweat over **temperature**

Suggestions

Consider an **ensemble model** approach

Reproducible analysis

github.com/jschoeley/rbx2020

Jonas Schöley

@jschoeley

0000-0002-3340-8518

schoeley@demogr.mpg.de

MAX PLANCK INSTITUTE
FOR DEMOGRAPHIC RESEARCH