Title: Estimating the burden of COVID-19 on mortality, life expectancy and lifespan inequality in England and Wales: A population-level analysis

Authors: José Manuel Aburto, Newton fellow[1,2], Ridhi Kashyap, associate professor[1], Jonas Schöley, postdoctoral fellow[2], Colin Angus, senior research fellow[3], John Ermisch, professor[1], Melinda C. Mills, professor[1], Jennifer Beam Dowd, associate professor[1]

Affiliations:
[1] Leverhulme Centre for Demographic Science, Department of Sociology and Nuffield College, University of Oxford, 42-43 Park End Street, OX1 1JD Oxford, UK.

[2] Interdisciplinary Centre on Population Dynamics, University of Southern Denmark, Odense 5000, Denmark.

[3] School of Health and Related Research, University of Sheffield, Regent Court, Regent Street, S1 4DA Sheffield, UK

Correspondence to:
José Manuel Aburto
42-43 Park End Street, OX1 1JD Oxford, UK.
Email: jose-manuel.aburto@sociology.ox.ac.uk
Tel: +45 31712122
ORCID: 0000-0002-2926-6879

OR

Ridhi Kashyap
Nuffield College, New Road, Oxford OX1 1NF
Email: ridhi.kashyap@sociology.ox.ac.uk
ORCID: 0000-0003-0615-2868

Section 1. Estimation of the baseline mortality risk using 4 different approaches using training data from 2010 to week 10 of 2020 by age and sex.

1. Generalized Additive Model assuming a Negative Binomial distribution to account for overdispersion of deaths during the period we study[1]. The model includes smooth effects for the long term trend, age and seasonality, and an interaction between age and seasonality. The smooth effects are stratified by sex. The basic structure of the model is as follows:

$$\log(E(Y_t)/\theta_t) = \alpha + \text{sex} + f_1(\text{time}_t) + f_2(\text{age}_t) + f_3(\text{seasonality}_t) + f_4(\text{seasonality}_t, \text{age}_t)$$

Where $E(Y_t)$ are the expected deaths in a given week, θ_t are the offset, f_i are smooth functions. f_1 and f_2 are penalized splines for the long term trend and the age effect, respectively. f_3 is a penalized cyclic spline for the seasonality and f_4 is a smooth interaction between age and seasonality.

2. The second approach is a Generalized Additive Model assuming a Poisson distribution with the same structure as above.
3. The third approach is a Generalized Linear Model assuming a Poisson distribution used to estimate baseline mortality during influenza epidemics and known as Serfling model[2,3]. The basic structure of the model is as follows:

$$\log(E(Y_t)/\theta_t) = \alpha + \beta t + \gamma_2 \sin\left(\frac{2\pi t}{52}\right) + \gamma_3 \cos\left(\frac{2\pi t}{52}\right) + \gamma_4 \sin\left(\frac{2\pi t}{26}\right) + \gamma_5 \cos\left(\frac{2\pi t}{26}\right)$$

4. We constructed an empirical baseline mortality based on the average mortality rate over the previous five years 2015-19 within each week and stratum. The associated deaths from this approach result from multiplying the average death rates by the population exposed to the risk.

Excess deaths produced with different models.

Table 1. Excess deaths by the end of week 26 for ages 15 and above estimated with 4 different models with 95% predictive intervals in England and Wales.

Model	Excess deaths	Lower bound	Upper bound
GAM- Negative Binomial	58,982	56,329	61,827
GAM- Poisson	53,938	53,093	54,746
GLM- Poisson	54,150	53,281	54,928
Average Mortality	49,429	48,589	50,210

Table 2. Excess deaths by the end of week 26 for ages 15 and above estimated with 4 different models by age and sex with 95% predictive intervals in England and Wales.

		Females			Males		
Model	**Age group**	**Excess deaths**	**Lower bound**	**Upper bound**	**Excess deaths**	**Lower bound**	**Upper bound**
GAM Negative Binomial	15-44	198	95	304	216	53	379
	45-64	2,191	1,823	2,534	4,125	3,613	4,626
	65-74	3,026	2,504	3,532	5,286	4,591	5,993
	75-84	6,966	6,016	7,898	9,721	8,644	10,872
	85+	15,870	14,217	17,485	11,383	10,232	12,456
GAM Poisson	15-44	135	57	214	155	44	259
	45-64	1,896	1,711	2,083	3,891	3,679	4,098
	65-74	2,587	2,380	2,785	4,928	4,670	5,203
	75-84	6,024	5,693	6,326	9,067	8,731	9,375
	85+	14,408	13,995	14,808	10,848	10,489	11,167
GLM Poisson	15-44	113	32	187	93	-11	203
	45-64	2,020	1,833	2,192	3,786	3,588	3,992
	65-74	2,767	2,558	2,981	5,177	4,919	5,419
	75-84	7,038	6,752	7,348	9,722	9,413	10,031
	85+	13,249	12,841	13,662	10,186	9,868	10,521
Average Mortality	15-44	89	7	170	55	-50	161
	45-64	1,781	1,601	1,953	3,473	3,248	3,702
	65-74	2,438	2,223	2,642	4,601	4,345	4,854
	75-84	5,923	5,618	6,228	8,482	8,169	8,807
	85+	12,811	12,411	13,186	9,776	9,457	10,107

References

1 Wood SN. *Generalized Additive Models: An Introduction with R, Second Edition*. CRC Press 2017.

2 Nielsen J, Krause TG, Mølbak K. Influenza-associated mortality determined from all-cause mortality, Denmark 2010/11-2016/17: The FluMOMO model. *Influenza Other Respir Viruses* 2018;**12**:591–604. doi:10.1111/irv.12564

3 Serfling RE. Methods for Current Statistical Analysis of Excess Pneumonia-Influenza Deaths. *Public Health Rep 1896-1970* 1963;**78**:494–506. doi:10.2307/4591848

Figure 1. Expected (lines) vs Observed deaths (points) counts based on the 4 approaches described above for males by age groups (rows) 0-14, 15-44, 45-64, 65-74, 75-84 and 85-older years of age.

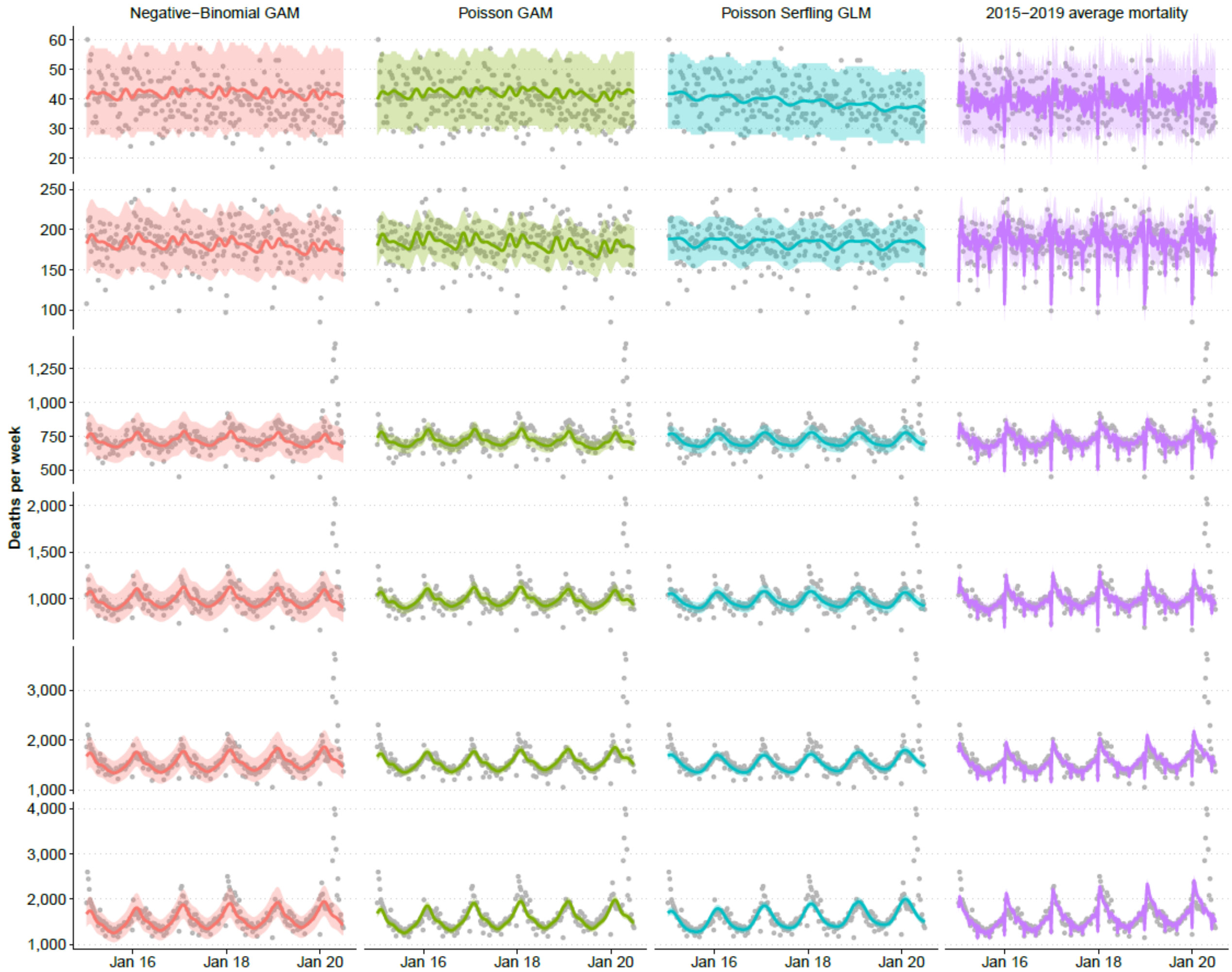

Figure 2. Expected (lines) vs Observed deaths (points) counts based on the 4 approaches described above for females by age groups (rows) 0-14, 15-44, 45-64, 65-74, 75-84 and 85-older years of age.

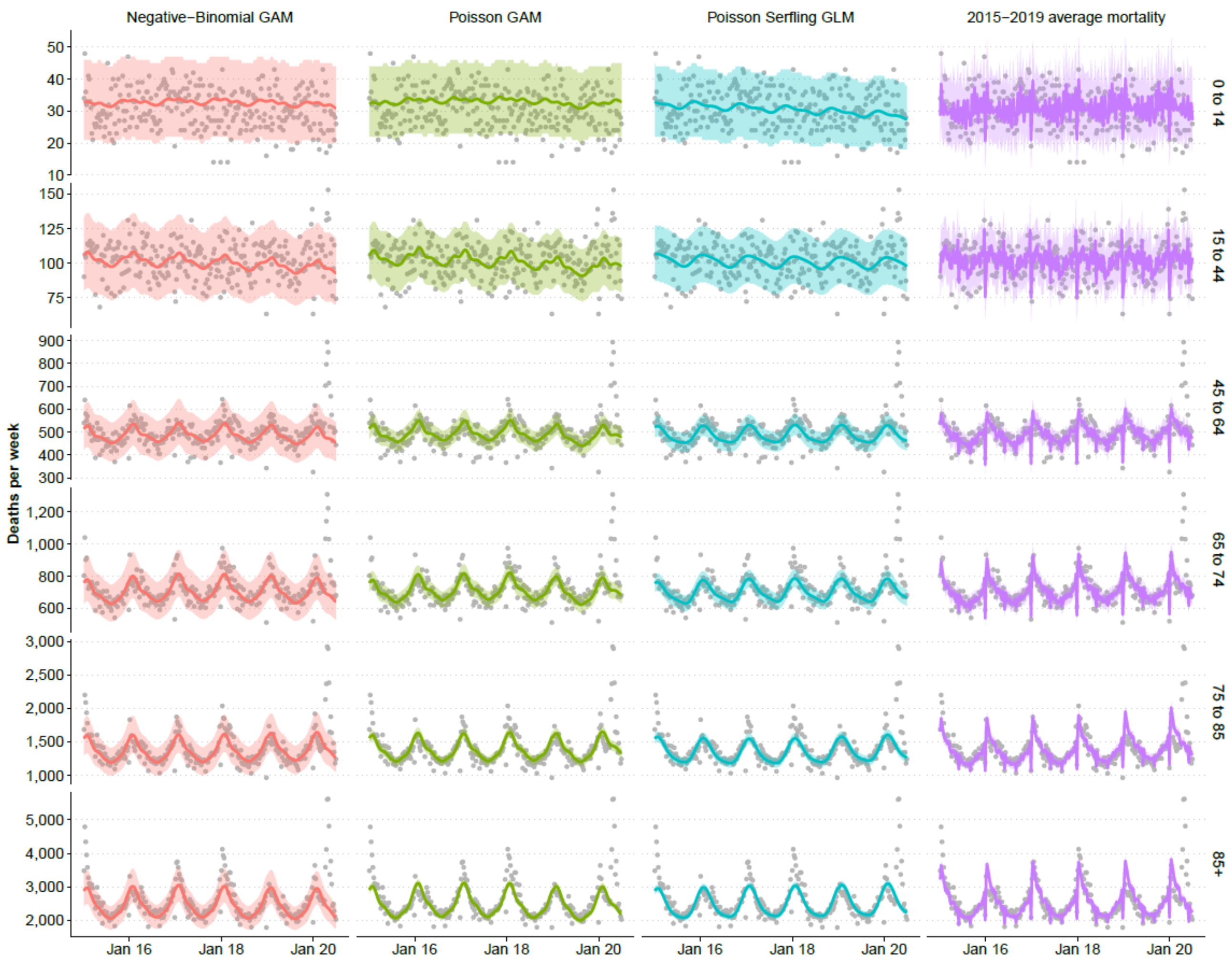

Figure 3. Sex ratio males/females of death rates during the course of the pandemic by age groups (rows) 0-14, 15-44, 45-64, 65-74, 75-84 and 85-older years of age.