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:

$$\begin{aligned}\log \mathrm{E}[Y_t/\theta_t\,] = \;& \beta_0 + \beta_{\text{day}\times\text{sex}\times\text{age}}(\text{day}\times\text{sex}\times\text{age}) + \beta_{\text{time}\times\text{sex}\times\text{age}}(\text{time}\times\text{sex}\times\text{age}) + \\ & f_1(\text{age}|\text{sex}) + f_2(\text{week}|\text{sex}) + f_3(\text{week},\text{age}), \\ & Y_t \sim \text{Neg.Bin},\end{aligned}$$

   Where $\mathrm{E}[Y_t]$ are the expected deaths in a given week, θ_t are the exposure times, and f_i are smooth functions.

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:

$$\begin{aligned}\log \mathrm{E}[Y_t/\theta_t\,] = \;& \beta_0 + \beta_{\text{day}\times\text{sex}\times\text{age}}(\text{day}\times\text{sex}\times\text{age}) + \beta_{\text{time}\times\text{sex}\times\text{age}}(\text{time}\times\text{sex}\times\text{age}) + \\ & \beta^1_{\text{week}\times\text{age}\times\text{sex}}\left(\sin\frac{\text{week}2\pi}{52}\times\text{age}\times\text{sex}\right) + \beta^2_{\text{week}\times\text{age}\times\text{sex}}\left(\cos\frac{\text{week}2\pi}{52}\times\text{age}\times\text{sex}\right) + \\ & \beta^3_{\text{week}\times\text{age}\times\text{sex}}\left(\sin\frac{\text{week}2\pi}{26}\times\text{age}\times\text{sex}\right) + \beta^4_{\text{week}\times\text{age}\times\text{sex}}\left(\cos\frac{\text{week}2\pi}{26}\times\text{age}\times\text{sex}\right), \\ & Y_t \sim \text{Pois.},\end{aligned}$$

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. Total excess deaths by the end of week 33 estimated with 4 different models with 95% prediction intervals in England and Wales.

	Female			Male			Total		
Model	**Excess**	**Lower PI**	**Upper PI**	**Excess**	**Lower PI**	**Upper PI**	**Excess**	**Lower PI**	**Upper PI**
GAM Negative Binomial	23,685	21,768	25,709	27,649	25,945	29,336	51,334	48,641	54,056
GAM Poisson	23,639	22,967	24,314	27,619	26,946	28,316	51,258	50,332	52,227
GLM Poisson (Serfling)	23,255	22,584	23,931	27,180	26,426	27,913	50,436	49,482	51,408
Average mortality	21,203	20,492	21,881	24,437	23,682	25,131	45,640	44,620	46,656

Table 2. Total excess deaths by the end of week 33 estimated with 4 different models by age and sex with 95% predictive intervals in England and Wales.

		Female			Male		
Model	**Age group**	**Excess**	**Lower PI**	**Upper PI**	**Excess**	**Lower PI**	**Upper PI**
GAM Negative Binomial	0	-10	-63	41	-43	-108	15
	15	202	90	309	129	-42	296
	45	1,986	1,599	2,381	4,044	3,484	4,543
	65	2,808	2,256	3,321	5,167	4,515	5,840
	75	6,687	5,788	7,614	9,417	8,300	10,502
	85	12,010	10,365	13,538	8,934	7,800	10,095
GAM Poisson	0	-15	-69	36	-45	-107	13
	15	204	103	299	127	-4	265
	45	2,003	1,784	2,228	4,051	3,812	4,310
	65	2,769	2,520	3,011	5,098	4,790	5,400
	75	6,671	6,319	7,031	9,347	8,971	9,735
	85	12,005	11,547	12,448	9,039	8,653	9,427
GLM Poisson (Serfling)	0	-8	-58	42	-45	-101	13
	15	183	85	287	126	0	261
	45	1,997	1,801	2,204	4,003	3,755	4,263
	65	2,663	2,403	2,904	5,026	4,734	5,334
	75	6,624	6,272	6,988	9,203	8,829	9,588
	85	11,796	11,331	12,268	8,866	8,471	9,260

Average mortality	0	-73	-124	-23	-143	-209	-82
	15	148	44	247	15	-114	134
	45	1,749	1,519	1,964	3,620	3,347	3,871
	65	2,373	2,111	2,627	4,400	4,094	4,708
	75	5,303	4,925	5,653	7,791	7,418	8,164
	85	11,701	11,233	12,165	8,753	8,379	9,145

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. Shaded areas indicate 95% prediction intervals.

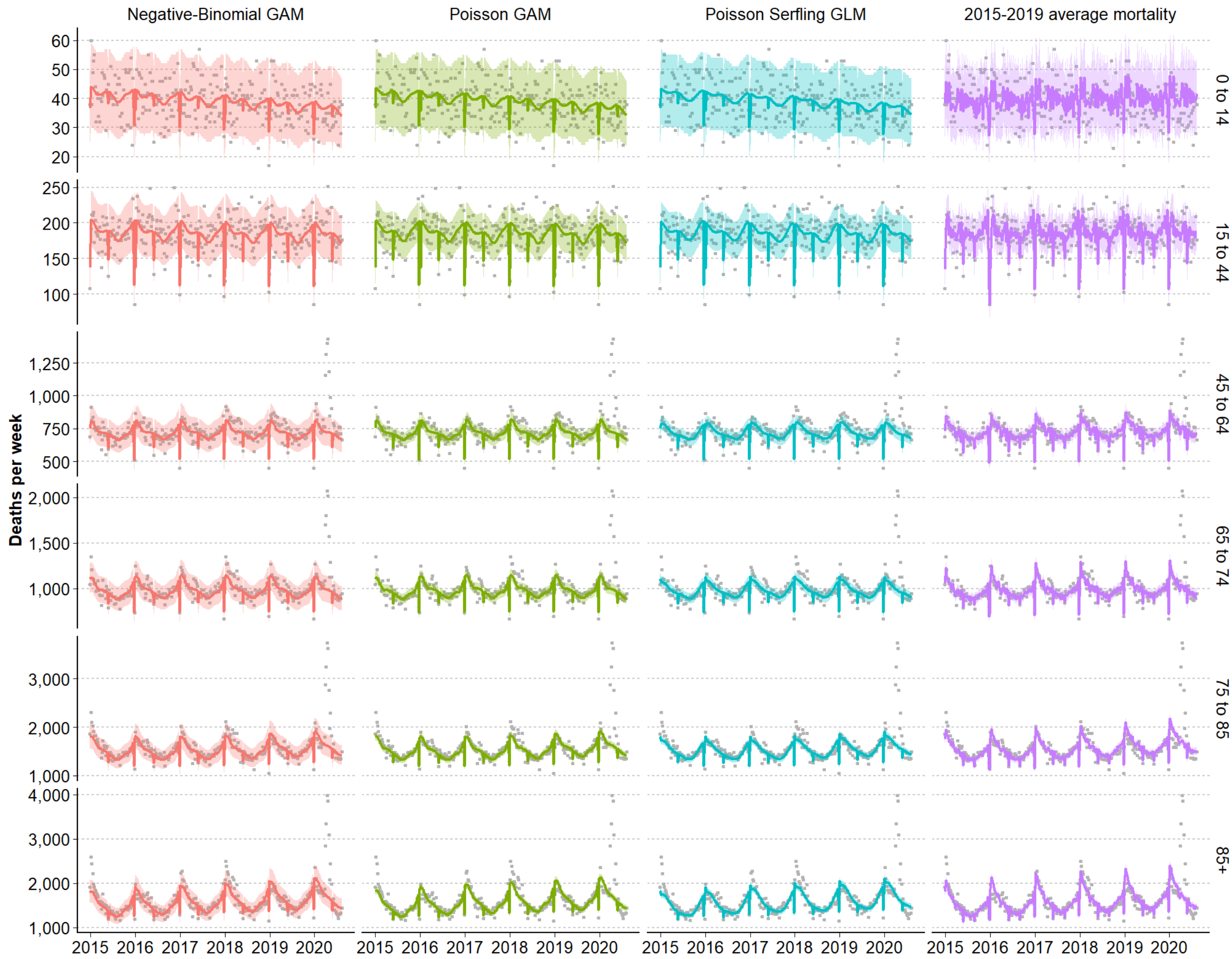

Figure 2. Expected (lines) vs Observed deaths (points) counts based on the 4 approaches described above for females by age groups (rows) 0-14, 15-44, 45-64, 65-74, 75-84 and 85-older years of age. Shaded areas indicate 95% prediction intervals.

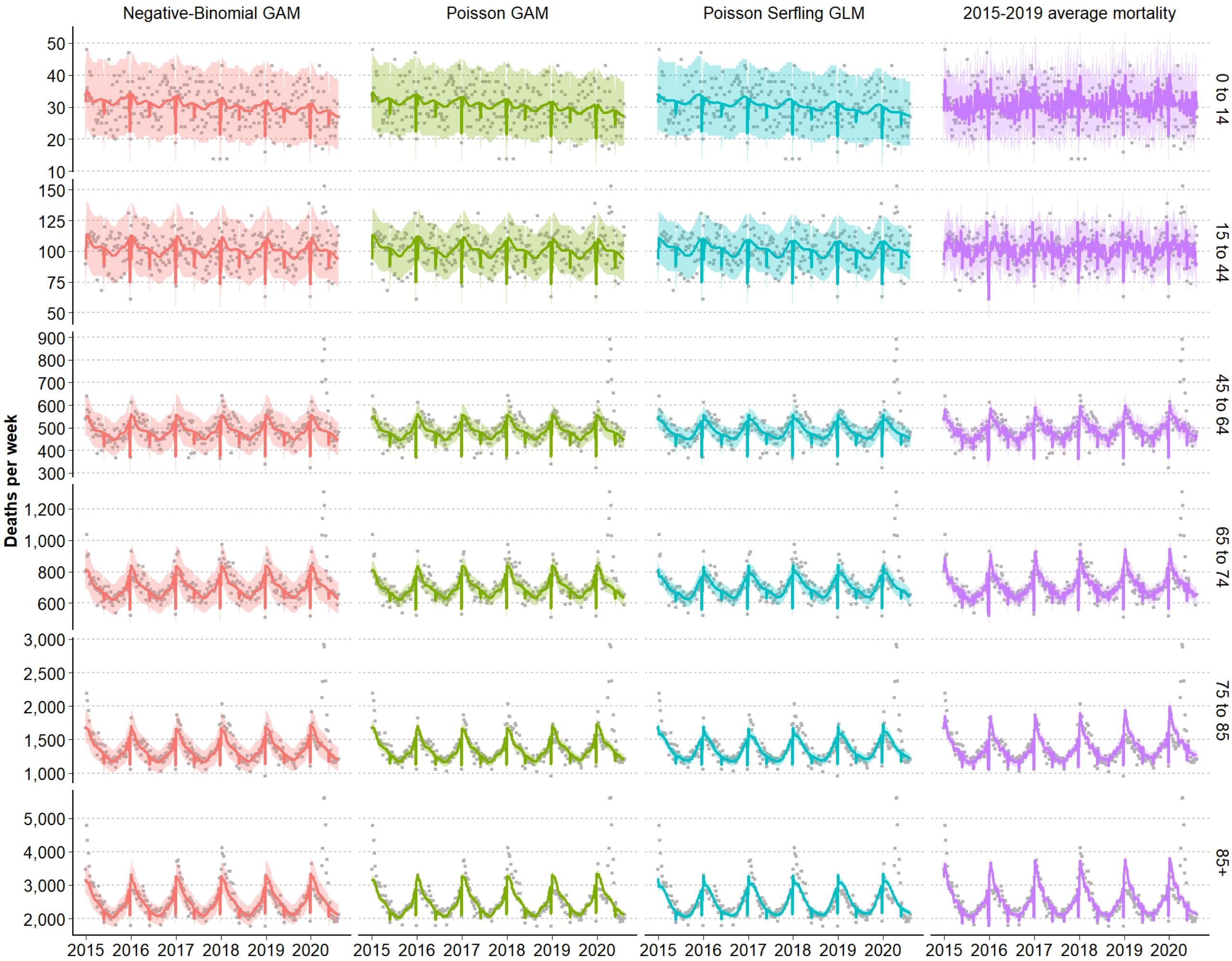

Figure 3. Sex ratio males/females of death rates during the course of the pandemic by age groups (rows) 0-14, 15-44, 45-64, 65-74, 75-84 and 85-older years of age.