preprint: not peer-reviewed; medRχiv http://dx.doi.org/10.1101/2020.07.22.20159707

Age and sex disaggregation of crude excess deaths during the 2020 spring COVID–19 outbreak in Spain with an estimate of infection fatality rate

José María Martín-Olalla[1, a)]
Universidad de Sevilla. Facultad de Física. Departamento de Física de la Materia Condensada. ES41012 Sevilla. Spain

(Dated: 2020-08-21 10:33:26+02:00)

Spanish official records of mortality and population during the 21st century are analyzed to determine the age-sex specific crude death rate in the 2020 spring (week 10 to week 21) COVID–19 outbreak .

Age-sex specific cumulative death rates can be modelled by a Poission regression with rate linearly variying with calendar year from which age-sex specific reference value for 2020 are obtained.

Excess death rate increases exponentially with age showing a doubling time $[4.1, 4.9]$a (female) and $[4.8, 5.4]$a (male). Age specific infection fatality rate doubling times below age 70 years are reported as $[4.7, 8.8]$a (female) and $[4.8, 6.6]$a (male).

Infection fatality rate for people aged more than 80 is discussed in relation to the shares of people living in institutionalized long term care facilities.

Keywords: mortality excess; z-score; EuroMoMo; INE; Instituto de Salud Carlos III

[a)] Electronic mail: olalla@us.es; Twitter: @MartinOlalla_JM;

CC BY NC ND

CONTENTS

1. Introduction 2

2. Data sets 2

3. Results 3

4. Discussion 5
4.1. Doubling times 5
4.2. IFR in aged groups of population 6

5. Conclusions 10

6. Conflict of interest statement 10

7. Data availability 10

References 10

Acknowledgments 11

1. INTRODUCTION

The illness designated COVID–19 caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) caught worldwide attention since its identification in late 2019[1]. Spain is one of the European countries most impacted by the disease during the spring of 2020[2]. Confirmed COVID–19 cases climbed up to several hundred thousands —some few thousands per one million population— and confirmed COVID–19 deaths to some 28 000 or six hundred deaths per one million population.[3] Lock-down measures came into effect on March 16 to help diminishing the outbreak.

Total excess deaths or crude excess deaths is a key quantity to understand the acute impact of a pandemic[4]. Its determination sensibly requires a reference mortality or baseline and a time interval over which the excess is determined. Baseline is determined from previous records of mortality. European National Statistical Institutes and Eurostat are doing a great effort in disseminating weekly deaths in Europe in the past years.

Age specific and sex specific analyses are of great interest since they addressed the question to whom the diseases impact the most[5]. Age specific analysis of mortality in Spain is heavily influenced by the strong impact of the disease within nursing homes[6]. This manuscript takes official records of weekly crude deaths in Spain during the 21st century and population records to ascertain the impact of the COVID–19 in age-sex specific death rates, including a discussion on age specific infection fatality rates.

2. DATA SETS

In April 2020 Eurostat set up "an exceptional temporary data collection on total week deaths in order to support the policy and research efforts related to COVID–19"[7]. The data collection is disaggregated by sex, five year age group and NUTS regions in several countries of the European Union and elsewhere. Data are provided by National Statistical Institutes on a voluntary basis. The 2020 weekly deaths still have the flag of "estimate".

Age group population values until January 1, 2019 for Spain were also collected from Eurostat `demo_pjan` table. This table disaggregates age year by year which allows to build up five year age groups matching to the weekly death data. Linear interpolation was used to compute population a weekly basis.

Total population values for January 1,2020 can be collected from table `ts00001` in Eurostat. Unluckily there is no age disaggregation. However these figure were obtained in Spain from the table `31304` at the Instituto Nacional de Estadística. For the Dutch set the last two available shares of population for every age group were used to extrapolate the shares of population in 2020.

Eurostat requested the National Statistical Institutes the transmission of "a back series weekly deaths for as many years as possible, recommending as starting point the year 2000". In this manuscript Spanish and Dutch weekly deaths will be analyzed from 2001 on-wards. As Eurostat points out "a long enough time series is necessary for temporal comparisons and statistical modelling".

3. RESULTS

In specific death rate analysis the population N is assumed to be a blending, a mixture, or a multicomponent system composed of several age groups and sex groups of size N_i whose behaviour in relation to death is homogeneous under some circumstances. Age-sex specific weekly death rate[8] d_i is nothing but the ratio between weekly deaths D_i within a group i and its population N_i. All age-sex specific weekly deaths sum up the total weekly deaths $D = \sum_i D_i$. Also considering the definition of age-sex specific death rate this can be written as $D = \sum_i N_i \times d_i$ from which the total weekly death rate $d = D/N$ can be obtained as $d = \sum x_i d_i$ where $x_i = N_i/N$ is the shares of population by age-sex group. Total weekly death rate is then the weighted average of age-sex specific death rates.

Results can also be partially grouped, ie. death rates for both male and female yields the total sex category.

Figure 1 shows the male (top) and female (bottom) age specific weekly crude death rate d_i in Spain for 19 age groups (color from a gradient palette) the total —age unspecific— weekly crude death rate d (orange). Death rates are distributed exponentially following the Gompertz law[9] $d(a) \propto 2^{a/\tau}$ where a is the age and τ is the doubling time. This empirical law suggests the predominance of age specific contributions to the cause of death instead of external causes like wars, murders, plagues or the like.

Following the ideas in Ref. [10] this work will analyze the cumulative weekly death rate from W10 to W21 in the 21st century. This interval is noted by strip bands in Figure 1. In 2020 W10 to W21 extends from Monday March 2, 2020 to Sunday May 24, 2020. In this period of time weekly deaths exhibited an extremely large anomaly in Spain ($z-$ score equal to 16). The data collection allows to compute the cumulative deaths recorded in this period for the past years and for every age group and sex. Scaling by the group population at the given year gives us the cumulative death rate in this period and year.

Figure 2 shows the results of this statistics in Spain. Every panel shows age specific male (blue) and female (orange) death rates against calendar year. Vertical axis is normalized by removing the both sex (total) average death rate and by scaling by the total standard deviation. Black lines the central tendency of a full model Poisson regression for data 2001 to 2019 ($n = 19$) with a response linearly depending on calendar year. The tendency is extended until 2020 to give the predicted reference value from the model in that year.

It is worthy to note that the youngest age group exhibits an anomalously high increase in 2020 (see also Table 1 in Ref. [6]). However a close inspection of the signal in Spain only reveals a sudden upraise in mid 2019 after which the signal fluctuates normally, including W10 to W21 in 2020. As an hypothesis it might be possible that fake 0 age deaths are initially reported and attributed to this age group. Later they are cleansed by the National Statistics Institute after close inspection of records. Data after mid 2019 still lacks this upgrade.

Table 1 summarizes the results displayed in Figure 2. It first lists the shares of population in 2020. Then the observed cumulative death rate o in 2019 followed by the male to female death rate ratio. Then the observed cumulative death rate in 2020. Next group of columns displays the results in Figure 2: first the predicted value or reference r for 2020 followed by the death rate excess $e = o - r$ and the P-score e/r. Then the male to female ratio. With few exceptions in the middle ages 2020 excess death rate male to female ratio and 2019 death rate male to female ratio are roughly coincidental.

This is followed by excess age-sex excess deaths $e \times x \times N$ and in the final column data by the age-sex specific infection fatality rate IFR. Values for IFR are computed as the ratio of excess death rate to the shares of IgG+ seroprevalence[6, 11] with confidence intervals given by the delta method. Aged groups of population are not considered because seroprevalence is only representative of non-institutionalized population (see Refs [6, 11]) and large shares of aged population live in institutionalized long-term care facilities (see also Section 4.2).

There are two totals in Table 1. First three rows report results for the age unspecific total group. This is the analysis of the sum of age specific deaths and it is equivalent to that reported in Ref. [10] except for the fact that there excess were taken relative to the average death rate in 2001–2019. It should be noted that age-unspecific male cumulative death rate sustains the null hypothesis of no relationship with calendar year ($p > 0.4$), whereas the age-unspecific female cumulative death rate does not ($p = 0.02$). However, to keep consistency with age-specific analysis this work considers a full linear model in every Poisson regression. This brings age unspecific Spanish excess deaths from $\sim 50\,000$[10] to $\sim 49\,000$ (Table 1).

Last three rows report the weight average results, which is the weighted sum of excess for age specific analysis. Both totals differ in 1 %.

Age/Sex		Shares	W10–21 death rate			Computed from Poisson regression					
			2019		2020	Predicted rate	Excess rate [95 %CI]	P-score	M/F	Deaths	IFR
		%	10^{-6}	M/F	10^{-6}	10^{-6}	10^{-6}	%			%
Total	F	51.0	1953		2954	1957	996 [982, 1010]	51		24 041	
	M	49.0	2092	1.1	3125	2046	1079 [1065, 1093]	53	1.1	25 027	
	T	100.0	2021		3038	2001	1037 [1023, 1051]	52		49 078	
0 to 4	F	2.1	123		177	103	74.7 [65.6, 83.8]	80		73	
	M	2.2	142	1.2	183	122	60.5 [50.6, 70.4]	54	0.8	63	
	T	4.2	133		181	113	67.9 [58.4, 77.4]	66		137	
5 to 9	F	2.4	11		11	13	−2.7 [−6.1, 0.7]	−21.7		−3	
	M	2.5	14	1.2	8	15	−7.1 [−10.6, −3.6]	−54	2.6	−9	
	T	4.9	13		9	14	−4.7 [−8.1, −1.3]	−37		−11	
10 to 14	F	2.6	18		9	17	−8.3 [−12.3, −4.3]	−50		−10	
	M	2.8	20	1.1	21	17	4.5 [0.9, 8.1]	31	−0.5	6	
	T	5.4	19		15	17	−1.5 [−5.2, 2.2]	−9.6		−4	
15 to 19	F	2.4	27		21	22	−1.1 [−5.2, 3]	−6.0		−1	
	M	2.6	40	1.5	30	35	−5.3 [−10.1, −0.5]	−24.3	4.6	−7	
	T	5.1	34		25	28	−3.1 [−7.6, 1.4]	−15.0		−7	
20 to 24	F	2.4	27		31	28	3.2 [−1.6, 8]	12.5		4	
	M	2.5	67	2.5	53	59	−6.2 [−12.8, 0.4]	−13.3	−1.9	−7	
	T	5.0	47		44	44	−0.1 [−5.8, 5.6]	−0.2		0	
25 to 29	F	2.7	41		45	33	12.4 [7.2, 17.6]	42		16	0.022 [0.012, 0.032]
	M	2.7	92	2.3	82	73	9 [1.5, 16.5]	14.4	0.7	12	0.016 [0.003, 0.029]
	T	5.4	66		65	53	11.7 [5.3, 18.1]	26		30	0.02 [0.008, 0.032]
30 to 34	F	2.9	60		69	45	23.9 [17.9, 29.9]	59		33	0.046 [0.032, 0.06]
	M	2.9	114	1.9	120	89	31.1 [23, 39.2]	43	1.3	43	0.066 [0.044, 0.088]
	T	5.9	87		95	67	28 [20.9, 35.1]	49		78	0.057 [0.039, 0.075]
35 to 39	F	3.5	79		102	76	26.2 [18.4, 34]	38		44	0.05 [0.033, 0.067]
	M	3.5	162	2.1	167	122	45.7 [36.4, 55]	49	1.7	76	0.097 [0.069, 0.125]
	T	7.0	120		134	99	35.3 [26.7, 43.9]	43		117	0.071 [0.049, 0.093]
40 to 44	F	4.1	131		139	133	6 [−5, 17]	4.9		12	0.012 [−0.008, 0.032]
	M	4.2	200	1.5	247	212	35 [22, 48]	19.7	5.7	70	0.067 [0.041, 0.093]
	T	8.4	166		194	173	21 [9, 33]	13.6		83	0.04 [0.017, 0.063]
45 to 49	F	4.1	237		275	245	30 [15, 45]	12.6		58	0.057 [0.028, 0.086]
	M	4.1	428	1.8	490	431	59 [40, 78]	14.5	1.9	115	0.111 [0.072, 0.15]
	T	8.2	333		383	339	45 [28, 62]	13.8		173	0.084 [0.049, 0.119]
50 to 54	F	3.9	394		460	427	34 [14, 54]	7.9		62	0.065 [0.025, 0.105]
	M	3.9	778	2.0	919	787	132 [106, 158]	17.3	3.9	241	0.249 [0.187, 0.311]
	T	7.7	585		688	605	83 [60, 106]	13.9		303	0.157 [0.107, 0.207]
55 to 59	F	3.6	618		832	627	204 [179, 229]	33		349	0.392 [0.318, 0.466]
	M	3.5	1319	2.1	1588	1304	285 [251, 319]	22.2	1.4	472	0.54 [0.44, 0.64]
	T	7.1	963		1204	960	245 [215, 275]	26		824	0.466 [0.377, 0.555]
60 to 64	F	3.2	921		1232	915	317 [287, 347]	35		476	0.63 [0.5, 0.76]
	M	3.0	2100	2.3	2730	2070	661 [617, 705]	32	2.1	936	1.35 [1.09, 1.61]
	T	6.2	1493		1959	1477	482 [445, 519]	33		1406	0.97 [0.78, 1.16]
65 to 69	F	2.7	1346		1848	1239	609 [575, 643]	50		774	1.22 [0.99, 1.45]
	M	2.5	3168	2.4	4408	3010	1398 [1346, 1450]	47	2.3	1622	2.85 [2.32, 3.38]
	T	5.1	2216		3070	2089	981 [937, 1025]	48		2385	1.98 [1.61, 2.35]
70 to 74	F	2.5	2118		3181	2031	1150 [1108, 1192]	58		1376	2.5 [1.92, 3.08]
	M	2.2	5012	2.4	7326	4686	2641 [2576, 2706]	58	2.3	2716	5.6 [4.2, 7]
	T	4.7	3457		5099	3268	1831 [1777, 1885]	57		4072	3.94 [3, 4.88]
75 to 79	F	2.1	4020		6062	3922	2140 [1950, 2330]	56		2093	4.7 [3.6, 5.8]
	M	1.7	8004	2.0	12 713	7813	4900 [4640, 5160]	64	2.3	3833	10.4 [7.8, 13]
	T	3.7	5786		9017	5645	3372 [3301, 3443]	61		5936	7.3 [5.6, 9]
80 to 84	F	1.6	8584		13 580	8412	5170 [4900, 5440]	63		3943	
	M	1.1	14 592	1.7	23 142	13 975	9170 [8820, 9520]	67	1.8	4776	
	T	2.7	11 022		17 461	10 704	6757 [6659, 6855]	64		8676	
85 to 89	F	1.3	18 147		27 676	17 781	9900 [9500, 10 300]	57		6282	
	M	0.8	25 679	1.4	39 842	25 430	14 410 [13 930, 14 890]	57	1.5	5291	
	T	2.1	20 897		32 133	20 576	11 560 [11 420, 11 700]	57		11 579	
≥ 90	F	0.8	42 713		64 776	41 908	22 870 [22 240, 23 500]	55		9093	
	M	0.4	48 827	1.1	75 371	48 537	26 830 [26 150, 27 510]	56	1.2	4518	
	T	1.2	44 517		67 924	43 812	24 110 [23 470, 24 750]	55		13 647	
W.A.	F		996		1506	1899	1022	54		24 671	
	M		1026		1531	2012	1068	53	1.0	24 767	
	T		2021		3037	1958	1044	53		49 424	

TABLE 1. Age specific cumulative death rate by sex (F, female; M, male; T, total) in Spain from W10 to W21. First column shows the age-sex specific shares of population x in 2020 when $N = 47\,329\,981$. Then the observed cumulative death rate in 2019 and male to female ratio, folowed by the observed cumulative death rate in 2020. Then results from a full model linear Poisson regression (see Figure 2): the predicted value r and the excess $e = o - r$ followed by 95 % confidence interval, $P-$score e/r, and the male to female ratio. Then age-sex specific total deaths $e \times x \times N$ and the infection fatality rate IFR as the ratio of excess death rate to the shares of seroprevalence from Ref. [6] First three rows report overall age unspecific excesses. Last three rows report weight averages and grand totals of age specific excesses. Both differ in 1 %.

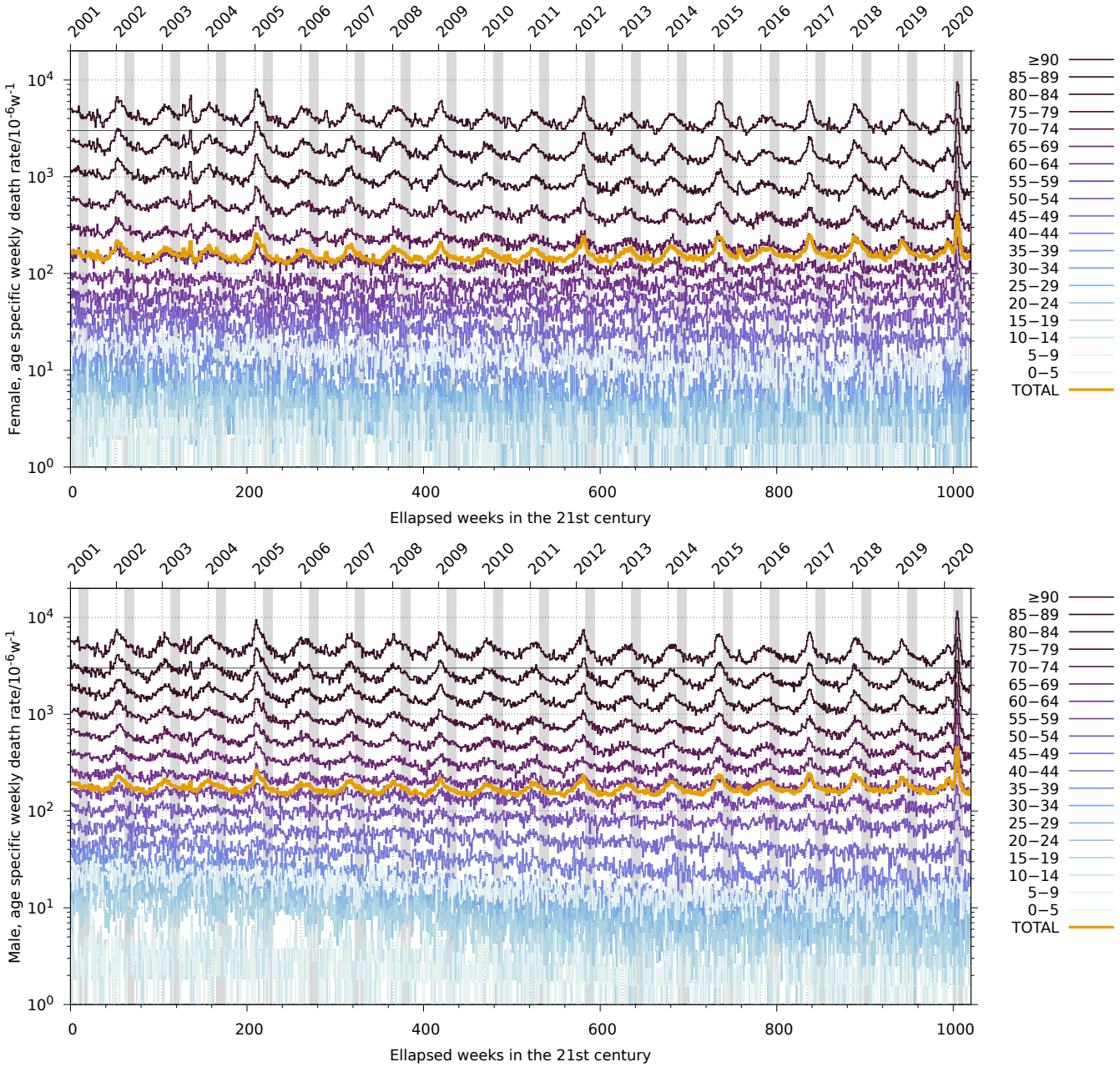

FIG. 1. Age specific, sex specific —female (top) and male (bottom)—, weekly crude death rate in Spain during the 21st century in a gradient color. Orange line shows the total weekly crude death rate.Gray strip bands locate W10 to W21.

4. DISCUSSION

4.1. Doubling times

Mimicking Figure 2 one may think in plotting age-sex specific cumulative death rates as a function of age group for every year in the collection. As an example first two columns in Table 1 show a steadily progressive increase with age group for every sex. Same happens for excess death rates and for IFR in Table 1.

Figure 3 shows these tendencies in a semilogarithmic plot. Every panel shows exponential increase (Gompertz law) with doubling time slightly larger than 6 (death rates) and 5 (excess death rates). Since seroprevalence[11] only changes smoothly with age, IFR replicates the logarithmic behaviour of age specific excess deaths.

Table 2 summarizes the results of the exponential fitting $y \propto 2^{a/\tau}$, where a is the low bound of age group and y is

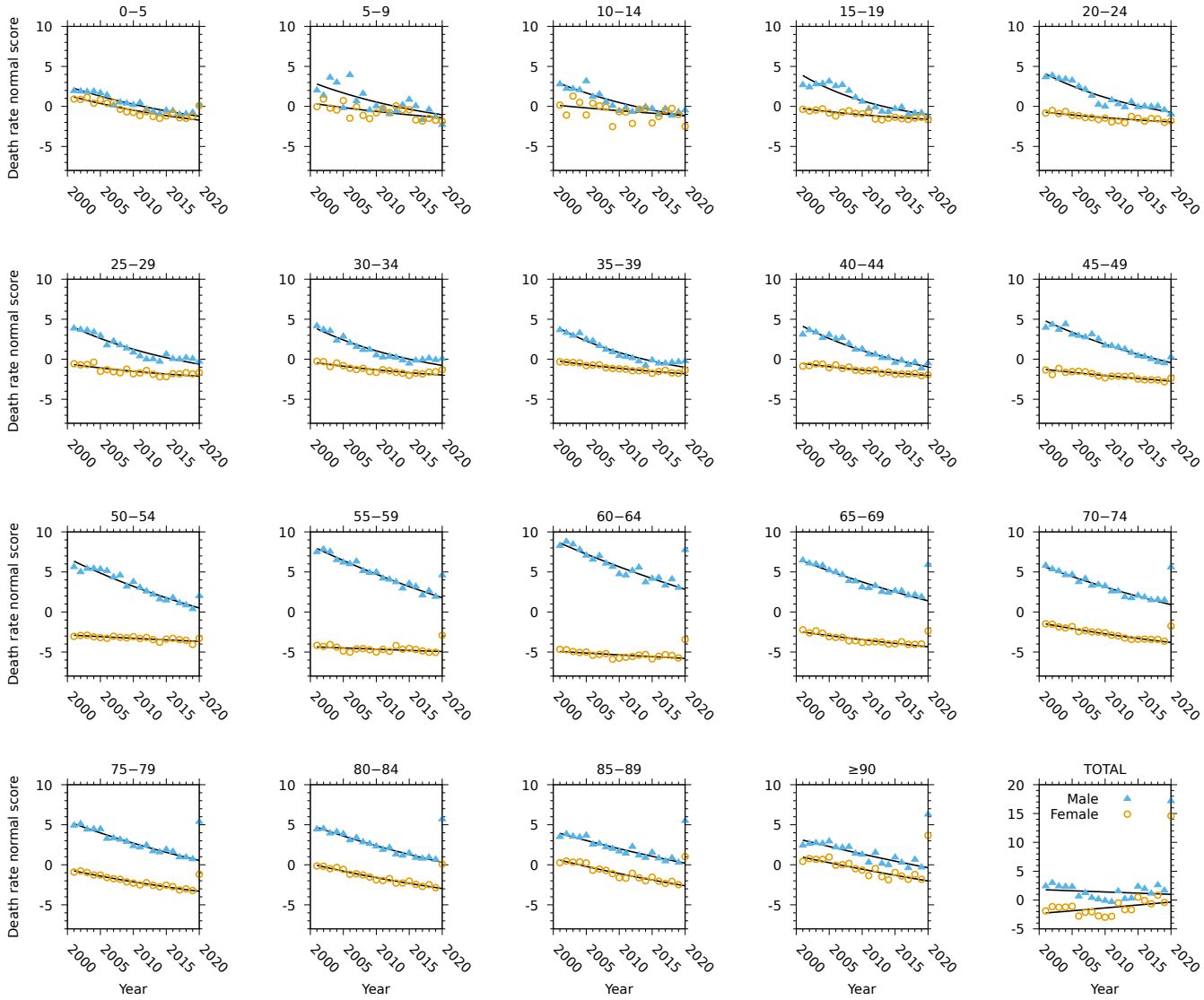

FIG. 2. Age and sex specific death rates from W10 to W21 in Spain. For every age group male death rate is noted in blue (triangles); female in orange (circles). Black line shows the central tendency from a Poisson regression that fitted death rates and calendar year from 2001 to 2019. The predicted value in 2020 is the reference for the model. Notice 2020 observed values lying outside the prediction interval in the two bottom-most rows. Bottom-right most panel displays the total age group death rates. The normal score was obtained removing the total sex sample average value 2001–2019 and scaling by the total sex standard deviation.

any of the responses. A doubling rate close to 5 a means that excess death rates doubles in every step of increase in the age group staircase from 40–45 on and it translates into a log-linear slope equal to $\sim \log(2)/(5\,\mathrm{a}) \sim 0.126\,\mathrm{a}^{-1}$, in accordance with Figure 3 in Ref. [5].

4.2. IFR in aged groups of population

Following the ideas by Pastor-Barriuso *et al.* [6] the ratio of all-cause excess deaths to seroprevalence is not representative of the IFR in the most aged groups of population due to shares of population living in institutionalized long-term care facilities which did not enter in the seroprevalence study.

In order to give some context Abellán García *et al.* [12] estimated 322 000 persons living in Spanish long term care facilities. It is only a 3 % of Spanish population aged more than 65 years. However, $N_1 = 19\,681$ deaths[6] related to COVID–19 were reported in institutionalized long-term care facilities, which makes 42 % of the 46 295 excess deaths

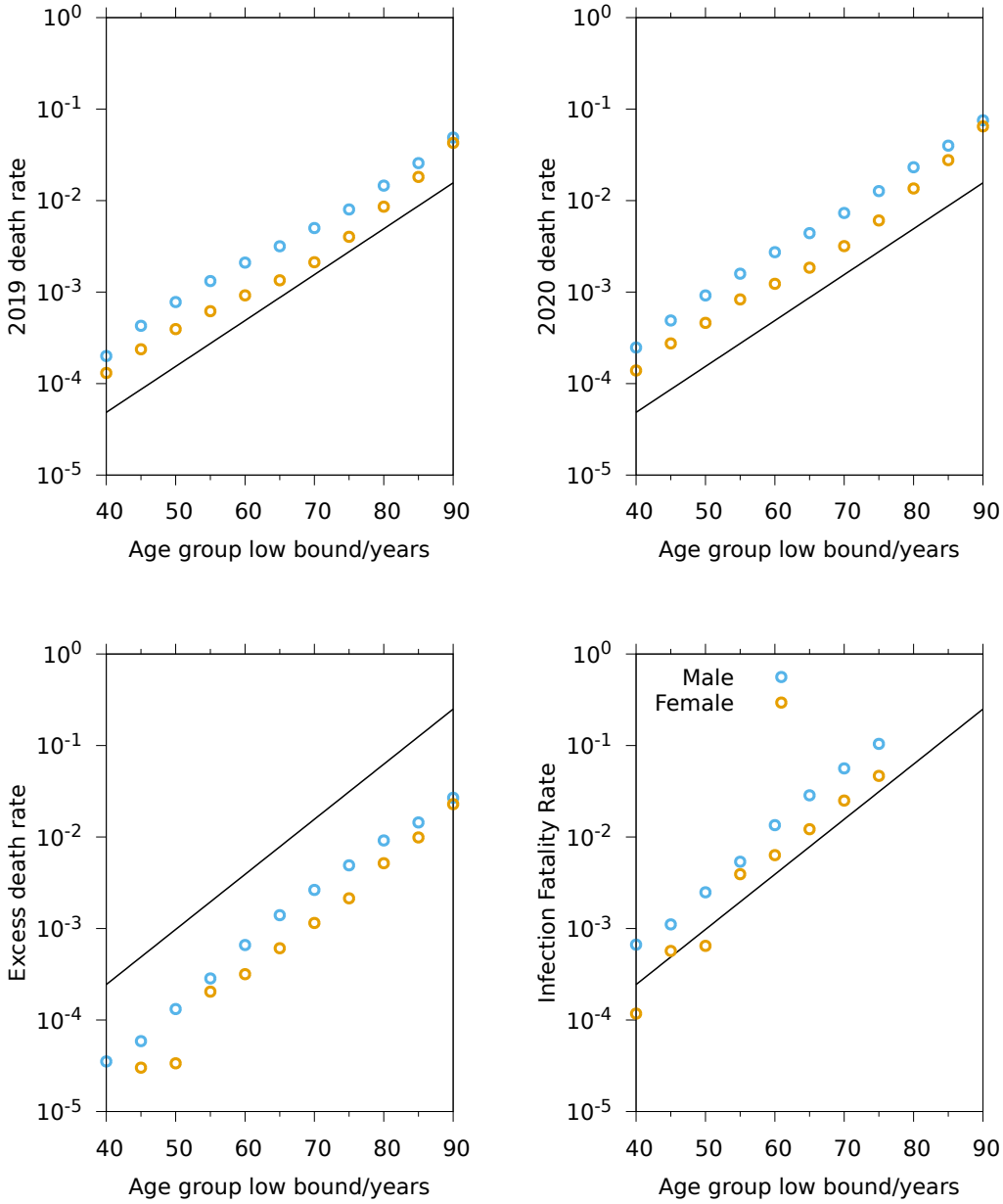

FIG. 3. Semilogarithmic plot of the observed age specific death rates in 2019 and 2020; excess death rate in Spain. Blue circles shows male statistics; orange circles, female statistics. Straight line shows doubling times $\tau = 6\,\mathrm{a}$ (death rates) and $\tau = 5\,\mathrm{a}$ (excess death rate). Excess death rate confidence intervals shown in Table 1 are close to point size in this figure. They are not printed for the sake of clarity. IFR confidence intervals are exceedingly large except for the last four data points due to the logarithmic axis. Doubling times are reported in Table 2.

in population aged more than 65 years (see Table 1).

Pastor-Barriuso *et al.* [6] discussed in detail this issue. They considered $N_2 = 44\,459$ excess death as reported from the Monitor of Mortality[13]) from which they excluded N_1 to give an estimate of excess deaths in Spanish non-institutionalized population. IFR was then computed as d_n/s_n where s_n is the shares of seroprevalence within non-institutionalized population and d_n is the excess death rate in non-institutionalized population.

Infection fatality rate reported in Ref. [6] (Table 1) is listed in the first column data of Table 3. The original IFR can be transformed into $d_n/(s_n + d_n)$. This metric is relevant if $s_n \sim d_n$ as it happens in the most aged groups when d climbs up to few percent.

Response y		Spain $\tau[95\,\%\mathbf{CI}]$ years	R^2
2019 death rate	F	$6.3[5.8, 7.1]$	0.982
	M	$6.7[6.4, 7.0]$	0.996
	T	$6.6[6.2, 7.0]$	0.993
2020 death rate	F	$5.9[5.5, 6.5]$	0.989
	M	$6.2[6.1, 6.4]$	0.999
	T	$6.2[5.9, 6.5]$	0.997
Excess death rate	F	$4.5[4.1, 4.9]$	0.984
	M	$5.1[4.8, 5.4]$	0.995
	T	$4.9[4.7, 5.2]$	0.996
IFR	F	$6.1[4.7, 8.8]$	0.857
	M	$5.6[4.8, 6.6]$	0.959
	T	$5.9[5.0, 7.3]$	0.941

TABLE 2. Doubling times $y \propto 2^{a/\tau}$ —where a is age group low bound— of death rates, excess death rate and IFR and Pearson's correlation coefficient in Spain and Netherlands. Only population aged 40 or more enters in this analysis ($n = 11$), see Figure 3.

Pastor-Barriuso *et al.* started with an overall excess death in non-institutionalized population equal to $D_1 = N_2 - N_1 = 24\,778$. In this work age-sex specific excess deaths adds up 49 424, which makes $D_2 = 49\,424 - 19\,681 = 29\,743$ overall excess deaths in non-institutionalized population. They differ in $D_2/D_1 = 1.20$. As a result IFR in Ref. [6] must be augmented in a factor 1.20 to be compared.[15]

On the other hand overall IFR can be computed as the ratio of excess deaths (see Table 1) to seroprevalance reported in Ref. [6]. This seroprevalence must be corrected to account for the shares of institutionalized population which did not enter in the seroprevalence study. IFR can then be expressed as $d/(d + s_n(1 - x))$. The share x can be obtained from the estimates for Spanish population living in long-term care facilities[12, 14]. The $(1 - x)$ accounts for the fact that d is scaled by resident population while s_n is scaled by non-institutionalized population.

Taking into account all this, Table 3 shows that the only relevant difference is observed for the most aged group. On the first hand this group exhibits 16.9 % (male) and 7.3 % (female) for non-institutionalized IFR, and on the other hand, it makes 24.8 % (male) and 19.1 % (female) for overall IFR. Understandably the issue arises only if the shares of population living in long-term care facilities is large enough. The difference is huge —a factor $\times 2.6$— for female population which highlights the larger shares of population at these ages and the larger shares of people living in nursing homes.

IFR reported in this work is larger because it still lacks the shares of infected survivors living in long-term care facilities. To the best of my knowledge this is yet an unknown quantity and it is the central point of Ref. [6]. In other words, IFR reported in Table 3 would be accurate only if every infection in long-term care facilities had ended in a death, which is not the case.

More sensibly the IFR could be computed as $d/((1 - x) \cdot s_n + d + y)$, where y is the (unknown) ratio of infected survivors within institutionalized population to the overall population in the group. Figure 4 shows this function for male and female population. The figure highlights the striking difference in the female group: not even when every surviving woman aged ≥ 80 living in long-term care facilities were infected, the overall IFR would fall to the value reported for non-institutionalized women. This suggests an specific impact of COVID–19 in institutionalized females that goes beyond of the size of the infection.

On the other, if 30 % of male institutionalized population aged ≥ 80 were infected survivors then the overall IFR would fall to the value reported for non-institutionalized population. Arguably this might be a highest bound for the shares of male infected survivors since, otherwise, overall IFR would be smaller than IFR in non-institutionalized population.

Alternatively, IFR can be computed for non-institutionalized population aged more than 80 years. From Table 3 there are $N_3 = 33\,902$ all-sex excess death in this group of population. The overwhelming majority of N_1 must be assigned to this group of population because: (1) death rates increase exponentially as age increases and (2) the shares of population living in institutionalized facilities increase with age. Therefore a rough estimate of non-institutionalized excess death for people aged more than 80 years is $N_4 = N_3 - N_1 \times 0.95 = 14\,200$ or an excess death rate equal to $d_4 = 5800 \times 10^{-6}$. Considering the seroprevalence of this age group the infection fatality rate for non-institutionalized,

Age	Inst Pop[12, 14]	IFR Pastor-Barriuso *et al.* [6]			This work		
	x %	Original %	Reformed %	Augmented %	IFR %	Excess rate %	Deaths
Male							
0–9	0.0	0.0	0.0	0.0	0.1	0.0	54
10–19	0.0	0.0	0.0	0.0	0.0	0.0	−1
20–29	0.0	0.0	0.0	0.0	0.0	0.0	4
30–39	0.0	0.0	0.0	0.0	0.1	0.0	119
40–49	0.0	0.1	0.1	0.1	0.1	0.0	185
50–59	0.0	0.3	0.3	0.4	0.4	0.0	713
60–69	0.2	1.6	1.6	1.9	2.0	0.1	2558
70–79	1.4	6.1	5.8	6.9	7.2	0.4	6549
≥ 80	8.9	16.4	14.1	16.9	24.8	1.4	14 585
Female							
0–9	0.0	0.0	0.0	0.0	0.1	0.0	70
10–19	0.0	0.0	0.0	0.0	0.0	0.0	−12
20–29	0.0	0.0	0.0	0.0	0.0	0.0	19
30–39	0.0	0.1	0.0	0.1	0.0	0.0	77
40–49	0.0	0.0	0.0	0.0	0.0	0.0	70
50–59	0.0	0.2	0.2	0.2	0.2	0.0	411
60–69	0.2	0.6	0.6	0.7	0.9	0.0	1250
70–79	1.4	2.7	2.6	3.1	3.4	0.2	3469
≥ 80	8.9	6.5	6.1	7.3	19.1	1.1	19 317

TABLE 3. Age and sex specific IFR in Spain from Pastor-Barriuso *et al.* [6] and from this work. First column data lists the shares of population living in long-term care facilities x. Original data in Ref. [6] are d_n/s_n where d_n is age-sex specific death rate in non-institutionalized population (overall deaths adds up to 24 778) and S_n is the shares of seroprevalence in non-institutionalized population. Original data are reformed into $d_n/(s_n+d_n)$. Reformed data are augmented into $1.2\times d_n/(s_n+d_n)$. This work IFR displays $d/((1-x)\cdot s_n+d)$ where d is excess death rate. Overall non-institutionalized excess deaths adds up to 29 743, which differs in 20 % from the value used by Pastor-Barriuso *et al.* and justifies the augmented value.

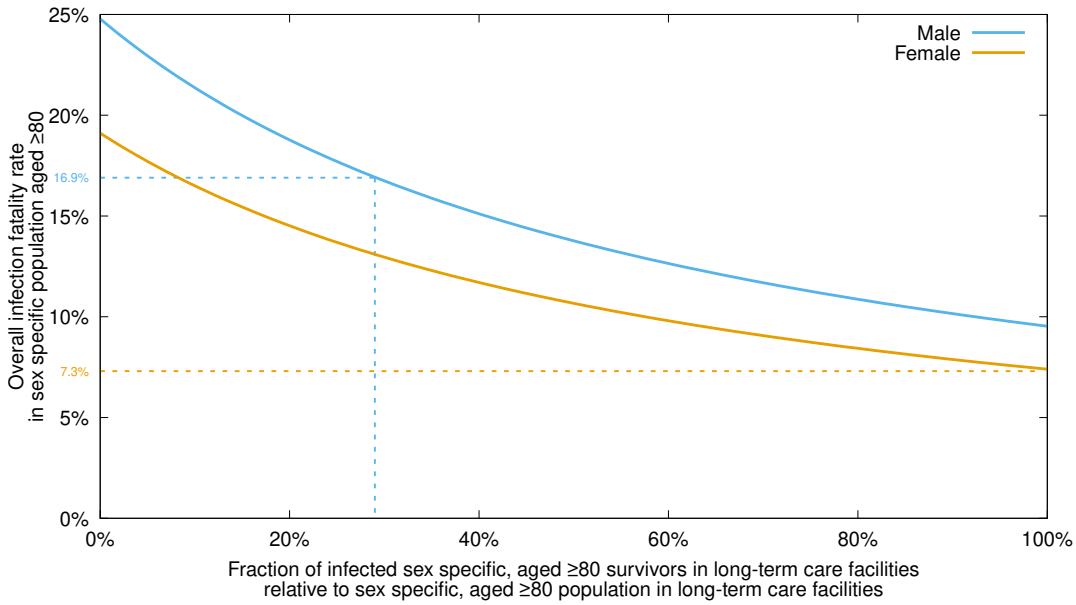

FIG. 4. Overall infection fatality rate in Spain during the spring COVID–19 outbreak for male (blue) and female (orange) population aged ≥ 80 as a function of the fraction of (unknown) infected survivors living in long-term care facilities relative to population living in long-term care facilities. Broken lines show the values reported by Pastor-Barriuso *et al.* [6] for sex specific population aged ≥ 80 not living in long-term care facilities.

all-sex, people aged more than 80 years is 12 %.

5. CONCLUSIONS

Spanish age-sex specific cumulative death rate in the period W10 to W21 during the 21st century can be modelled by a Poisson regression with a rate linearly depending on calendar year. Age-sex specific excess death rates in Spain during the spring 2020 (W10 to W21) COVID–19 outbreak follows Gompertz exponential law with doubling times in the range $[4.1, 4.9]$a (female) and $[4.8, 5.4]$a (male). Age specific infection fatality rate doubling times below age 70 years are reported as $[4.7, 8.8]$a (female) and $[4.8, 6.6]$a (male) in accordance with previous reports[5].

An upper bound for the over-all infection fatality rate in population aged more than 80 is reported as 25 % (male) and 19 % (female). For non-institutionalized population IFR is reported as 12 % (sex unspecific).

6. CONFLICT OF INTEREST STATEMENT

The author declares no conflict of interest.

7. DATA AVAILABILITY

Every piece of data in this work comes from public collection found in Eurostat and the Instituto Nacional de Estadística.

REFERENCES

[1]N. Zhu, D. Zhang, W. Wang, X. Li, B. Yang, J. Song, X. Zhao, B. Huang, W. Shi, R. Lu, P. Niu, F. Zhan, X. Ma, D. Wang, W. Xu, G. Wu, G. F. Gao, and W. Tan, New England Journal of Medicine **382**, 727 (2020).

[2]A. García-Basteiro, C. Alvarez-Dardet, A. Arenas, R. Bengoa, C. Borrell, M. Del Val, M. Franco, M. Gea-Sánchez, J. J. G. Otero, B. G. L. Valcárcel, I. Hernández, J. C. March, J. M. Martin-Moreno, C. Menéndez, S. Minué, C. Muntaner, M. Porta, D. Prieto-Alhambra, C. Vives-Cases, and H. Legido-Quigley, The Lancet **0** (2020), 10.1016/S0140-6736(20)31713-X.

[3]Instituto de Salud Carlos III COVID–19 in Spain `https://cnecovid.isciii.es/covid19` (accessed July 2020).

[4]J. Aaron, J. Mullbauer, C. Giattino, and H. Ritchie, "A pandemic primer on excess mortality statistics and their comparability across countries," (2020).

[5]A. T. Levin, K. B. Cochran, and S. P. Walsh, medRxiv , 2020.07.23.20160895 (2020).

[6]R. Pastor-Barriuso, B. Pérez-Gómez, M. A. Hernán, M. Pérez-Olmeda, R. Yotti, J. Oteo, J. L. Sanmartín, I. León-Gómez, A. Fernández-García, P. Fernández-Navarro, I. Cruz, M. Martín, C. Delgado-Sanz, N. Fernández De Larrea, J. León Paniagua, J. F. Muñoz-Montalvo, F. Blanco, A. Larrauri, and M. Pollán, medRxiv , 2020.08.06.20169722 (2020).

[7]`https://ec.europa.eu/eurostat/cache/metadata/en/demomwk_esms.htm`.

[8]This manuscript will print death rates as a number times 10^{-6} where the number stands for the mortality per one million population.

[9]B. Gompertz, Philosophical Transactions of the Royal Society of London **115**, 513 (1825).

[10]J. M. Martín-Olalla, medRxiv , 2020.07.22.20159707 (2020).

[11]M. Pollán, B. Pérez-Gómez, R. Pastor-Barriuso, J. Oteo, M. A. Hernán, M. Pérez-Olmeda, J. L. Sanmartín, A. Fernández-García, I. Cruz, N. Fernández de Larrea, M. Molina, F. Rodríguez-Cabrera, M. Martín, P. Merino-Amador, J. León Paniagua, J. F. Muñoz-Montalvo, F. Blanco, R. Yotti, F. Blanco, R. Gutiérrez Fernández, M. Martín, S. Mezcua Navarro, M. Molina, J. F. Muñoz-Montalvo, M. Salinero Hernández, J. L. Sanmartín, M. Cuenca-Estrella, R. Yotti, J. León Paniagua, N. Fernández de Larrea, P. Fernández-Navarro, R. Pastor-Barriuso, B. Pérez-Gómez, M. Pollán, A. Avellón, G. Fedele, A. Fernández-García, J. Oteo Iglesias, M. T. Pérez Olmeda, I. Cruz, M. E. Fernandez Martinez, F. D. Rodríguez-Cabrera, M. A. Hernán, S. Padrones Fernández, J. M. Rumbao Aguirre, J. M. Navarro Marí, B. Palop Borrás, A. B. Pérez Jiménez, M. Rodríguez-Iglesias, A. M. Calvo Gascón, M. L. Lou Alcaine, I. Donate Suárez, O. Suárez Álvarez, M. Rodríguez Pérez, M. Cases Sanchís, C. J. Villafáfila Gomila, L. Carbo Saladrigas, A. Hurtado Fernández, A. Oliver, E. Castro Feliciano, M. N. González Quintana, J. M. Barrasa Fernández, M. A. Hernández Betancor, M. Hernández Febles, L. Martín Martín, L.-M. López López,

T. Ugarte Miota, I. De Benito Población, M. S. Celada Pérez, M. N. Vallés Fernández, T. Maté Enríquez, M. Villa Arranz, M. Domínguez-Gil González, I. Fernández-Natal, G. Megías Lobón, J. L. Muñoz Bellido, P. Ciruela, A. Mas i Casals, M. Doladé Botías, M. A. Marcos Maeso, D. Pérez del Campo, A. Félix de Castro, R. Limón Ramírez, M. F. Elías Retamosa, M. Rubio González, M. S. Blanco Lobeiras, A. Fuentes Losada, A. Aguilera, G. Bou, Y. Caro, N. Marauri, L. M. Soria Blanco, I. del Cura González, M. Hernández Pascual, R. Alonso Fernández, P. Merino-Amador, N. Cabrera Castro, A. Tomás Lizcano, C. Ramírez Almagro, M. Segovia Hernández, N. Ascunce Elizaga, M. Ederra Sanz, C. Ezpeleta Baquedano, A. Bustinduy Bascaran, S. Iglesias Tamayo, L. Elorduy Otazua, R. Benarroch Benarroch, J. Lopera Flores, and A. Vázquez de la Villa, The Lancet **0** (2020), 10.1016/S0140-6736(20)31483-5.

[12] A. Abellán García, P. Aceituno Nieto, I. Fernández Morales, D. Ramiro Fariñas, and R. Pujol Rodríguez, "Una estimación de la población que vive en residencias de mayores," (2020).

[13] See https://www.isciii.es/QueHacemos/Servicios/VigilanciaSaludPublicaRENAVE/EnfermedadesTransmisibles/MoMo/Documents/informesMoMo2020/MoMo_Situacion%20a%2019%20de%20julio_CNE.pdf (accessed August 2020).

[14] A. Abellán García, P. Aceituno Nieto, I. Fernández Morales, D. Ramiro Fariñas, B. Castillo, and R. Pujol Rodríguez, "Una nueva estimación de población en residencias de mayores," (2020).

[15] This difference arise, among other things, because Pastor-Barriuso *et al.* [6] extracted deaths from the Spanish Monitor of Mortality which only registers deaths coming from 93 % of Spanish population. On the contrary, this work extracted deaths from Eurostat and the National Institute of Statistics which are reporting on the 100 % of population.

ACKNOWLEDGMENTS

The palette used in Figure 1 is named "dense", from `cmocean` palettes by Kristen Thyng https://matplotlib.org/cmocean/.

The author thanks National Statistics Institutes and Eurostat for releasing public records of weekly deaths. I knew of Eurostat effort from a tweet posted by Kiko Llaneras https://twitter.com/kikollan/status/1288830168925188096 on July 30, 2020.

This work was not founded.

This work benefited from free software running on `xubuntu 18.04.1LTS`. Data bases were imported into `GNU octave-4.2.2` (https://www.gnu.org/software/octave/). Poisson regressions were performed by `R-3.6.3` (https://www.r-project.org/). Pictures were developed thanks to `gnuplot-5.2.2` (http://www.gnuplot.info/). The manuscript was typeset in `GNU emacs-25.2.2` (https://www.gnu.org/software/emacs/) assisted by AUCTEX(https://www.gnu.org/software/auctex/). Data in pictures and in tables have been exported directly from `octave`.

This project started on July 31, 2020.