SARS-CoV-2 infection fatality risk in a nationwide seroepidemiological study

Roberto Pastor-Barriuso†a,b, Beatriz Pérez-Gómez†a,b, Miguel A. Hernán†c, Mayte Pérez-Olmeda†d, Raquel Yotti†e, Jesús Oteo†f, Jose L. Sanmartín†g, Inmaculada León-Gómez†a,b, Aurora Fernández-García†bd, Pablo Fernández-Navarro†a,b, Israel Cruz†h, Mariano Martín†e, Concha Delgado-Sanz†a,b, Nerea Fernández de Larrea†a,b, Jose León Paniagua†e, Juan F. Muñoz-Montalvo†f, Faustino Blanco†f, Amparo Larrauri†a,b, Marina Pollán†a,b§*, on behalf of the ENE-COVID Study Group‡

a. National Centre for Epidemiology, Institute of Health Carlos III, Madrid, Spain
b. Consortium for Biomedical Research in Epidemiology and Public Health (CIBERESP), Institute of Health Carlos III, Madrid, Spain
c. Departments of Epidemiology and Biostatistics, Harvard T.H. Chan School of Public Health; Harvard-MIT Division of Health Sciences and Technology, Boston, MA, USA
d. National Centre for Microbiology, Institute of Health Carlos III, Majadahonda, Madrid, Spain
e. Institute of Health Carlos III, Madrid, Spain
f. Spanish Network for Research in Infectious Diseases (REIPI), Institute of Health Carlos III, Madrid, Spain
g. Ministry of Health, Madrid, Spain
h. National School of Public Health. Institute of Health Carlos III, Madrid, Spain

† Equal contribution (co-first authors)
§ Equal contribution (co-senior authors)
‡ Collaborators are listed in the Supplementary Material

* Correspondence to: Prof. Marina Pollán, National Centre for Epidemiology, Institute of Health Carlos III, Monforte de Lemos 5, 28029 Madrid, Spain

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
ABSTRACT

The magnitude of the infection fatality risk (IFR) of SARS-CoV-2 remains under debate. Because the IFR is the number of deaths divided by the number of infected, serological studies are needed to identify asymptomatic and mild cases. Also, because ascertainment of deaths attributable to COVID-19 is often incomplete, the calculation of the IFR needs to be complemented with data on excess mortality. We used data from a nation-wide seroepidemiological study and two sources of mortality information—deaths among laboratory-confirmed COVID-19 cases and excess deaths—to estimate the range of IFR, both overall and by age and sex, in Spain.

The overall IFR ranged between 1.1% and 1.4% in men and 0.58% to 0.77% in women. The IFR increased sharply after age 50, ranging between 11.6% and 16.4% in men ≥80 years and between 4.6% and 6.5% in women ≥80 years. Our IFR estimates for SARS-CoV-2 are substantially greater than IFR estimators for seasonal influenza, justifying the implementation of special public health measures.
The infection fatality risk (IFR)—the proportion of infected individuals who die from the infection—is a key indicator to design public health policies to control infectious diseases. Because the magnitude of the infection fatality risk (IFR) of SARS-CoV-2 remains under debate,1,2 lockdowns and other extreme forms of social distancing have been questioned as appropriate responses to the COVID-19 pandemic.

An accurate estimation of the IFR of SARS-CoV-2 is difficult. Even if all symptomatic infections were diagnosed, something that so far has not occurred in most countries, asymptomatic infections cannot be clinically identified. Therefore, estimating the IFR needs to rely on well-designed serosurveys that provide an approximation to the proportion of individuals that has been infected, regardless of symptoms.3

A recent unpublished review of 24 serological reports4, several of them also unpublished, estimated an overall IFR of 0.68% (95% CI 0.53-0.83). However, the methodological quality of many of these studies was questionable, IFR estimates were based only on surveillance-registered deaths, and there was a very high between-study heterogeneity, with estimates ranging from 0.16% to 1.60%. Also, because the IFR for SARS-CoV-2 is expected to increase with age, overall IFR estimates cannot be directly compared between populations (e.g., China and Western Europe) with different age structure. Accurate and reliable age-specific estimates of IFR are urgently needed.

Here, we report overall and age- and sex-specific IFR estimates for SARS-CoV-2 from ENE-COVID, a large nationally representative serosurvey in Spain.

RESULTS

Through July 15, 2020, 19,228 laboratory-confirmed COVID-19 deaths and 24,778 excess all-cause deaths were estimated to occur among individuals residing in Spain outside of nursing homes. The distribution by age and sex was similar for both sources of death data: 64% of the COVID-19 deaths and 62% of the excess deaths occurred among men; 79% of confirmed COVID-19 deaths and 83% of excess deaths occurred among individuals aged 70 years or older.
Overall, the IFR estimate (95% CI) was 0.83% (0.78, 0.89) for confirmed COVID-19 deaths and 1.07% (1.00, 1.15) for excess deaths. The corresponding estimates were 1.11% and 1.40% for men, and 0.58% and 0.77% for women (Table 1). That is, depending on the source of death data, men were between 81% and 93% more likely to die than women.

The IFR estimate varied greatly with age. It was under 1 per 1000 through age 49, with much lower values in younger age groups, and increased sharply in older age groups (Figure 1). Among men aged 80 years or older, the IFR estimate (95% CI) was 11.6% (8.1, 16.5) for confirmed COVID-19 deaths and 16.4% (11.4, 23.2) for excess deaths. Among women aged 80 years or older, the corresponding estimates were 4.6% (3.4, 6.3) and 6.5% (4.7, 8.8).

**DISCUSSION**

We estimated an IFR for SARS-CoV-2 between 0.83% and 1.07% in Spain through July 15, 2020. The IFR was greater in men than in women and increased with age: 11.6% to 16.4% in men aged ≥80 years and 4.6% to 6.5% in women aged ≥80 years. Because incomplete ascertainment of deaths is unavoidable during a large-scale epidemic, we obtained separate IFR estimates based on confirmed COVID-19 deaths and excess all-cause deaths. The latter include mortality directly due to SARS-CoV-2 infection and net mortality due to the societal impact of the epidemic and its control measures, such as delayed care for emergencies and pre-existing chronic conditions, psychological distress, reductions in traffic injuries and other accidents, etc.

Our findings suggest that some of the heterogeneity in published IFR estimates is driven by the different age structure of the population. Our IFR estimates, like others from Italy, are larger than those from countries with a smaller proportion of population in the older age groups. Variations in IFR values may also be explained by the local dynamics of the epidemic (e.g., surge in number of new cases, diffusion of the virus among vulnerable collectives) and the health system capacity to treat severe cases.
The greater mortality in the elderly may result from a greater prevalence of comorbidities (cardiovascular disease, type 2 diabetes, lung and chronic kidney diseases) that are associated with greater COVID-19 mortality, and immunological changes (including a decrease of CD8 T cells) that affect the severity of SARS-CoV-2 infections. Sex differences in cellular immunity may explain the higher mortality among men, who present a poorer T-cell activation and an increase in pro-inflammatory cytokines. A negative correlation of T cell response with patients’ age was found in males but not in female patients.

Because the ENE-COVID serosurvey was conducted among the non-institutionalized Spanish population, we excluded deaths in long-term care facilities from the IFR estimates. However, with an estimated 333,920 people living in nursing homes (76% of them aged 80 or older) and more than 19,000 deaths, the epidemic was particularly serious in these institutions. Further research is needed to characterize the mortality in long-term care facilities with vulnerable populations in which the virus spreads very rapidly. This research, which requires a specific approach, would be helped by the inclusion of specific indicators to monitor these groups in regional and national surveillance systems.

The ENE-COVID serosurvey was timed to provide an IFR estimate for first wave of SARS-CoV-2 infection in Spain. The first round of the study started one month after the peak, which took place around March 20, and the last round ended on June 22. Thus, most participants would have been infected one month before their first participation. As IgG antibodies are detected 2–3 weeks after symptom onset in more than 90% of COVID-19 cases and decrease 2–3 months after infection, ENE-COVID is expected to cover infections through at least the first week of June. To include potentially delayed COVID-19 deaths, we included all deaths registered through July 15th. The median delay between onset of symptoms and death in our series –75% of deaths occurring before the 20th day- is similar to previously reported seroconversion times (14–21 days).

In conclusion, we estimated IFR estimates for SARS-CoV-2 by age and sex in one of the largest serosurveys in the world. Our overall IFR estimates (from 0.83% to 1.07%) are about 10 times larger than those for seasonal influenza, which provides support for strong control measures.
ETHICS
ENE-COVID study was approved by the Institutional Review Board of the Institute of Health Carlos III (Register number: PI 39_2020), and a written informed consent was obtained from all participants.

CONTRIBUTORS
BPG, RPB, MAH & MP are responsible for the conception and design of the study; RY & FB are the executive coordinators of the ENE-COVID study; MPO, JO & AFG are responsible for the serological analysis of the ENE-COVID study, coordinating microbiological labs. JLS, MM, JFM, IC and JLP are responsible for the ENE-COVID study logistics; ILG, CDS, PFN & AL extracted and curated RENAVE and MoMo data; MP, BPG, RPB, NFL and MAH were in charge of statistical analyses and tables and figures design; other authors included in the ENE-COVID group contributed to data acquisition, laboratory analyses and quality control of the ENE-COVID study at their respective regions and/or at national level. The first draft was initially written by MP, BPG, RPB, MAH, RY, AL & MPO. All authors contributed to data interpretation, substantially reviewed the first draft and approved the final version and agreed to be accountable for the work.

DECLARATION OF INTERESTS
We declare no competing interests.

FUNDING
Spanish Ministry of Health, Institute of Health Carlos III & Spanish National Health System.

ACKNOWLEDGMENTS
This work was supported by the Spanish Ministry of Health, the Institute of Health Carlos III (Ministry of Science and Innovation) and the National Health System, including the Health Services of all Autonomous Communities and autonomous cities: Servicio Andaluz de Salud, Servicio Aragonés de Salud, Servicio de Salud Principado de Asturias, Servei de Salut Illes Balears, Servicio Canario de la Salud, Servicio Cántabro de Salud, Servicio de Salud Castilla-La Mancha, Servicio de Salud de Castilla y León, Servei Català de Salut, Conselleria de Sanitat Universal i Salut Pública Generalitat Valenciana, Servicio Extremeño de Salud, Servizo Galego de Saúde, Servicio Riojano de Salud,
Servicio Madrileño de Salud, Servicio Murciano de Salud, Servicio Navarro de Salud-Osasunbidea & Instituto de Salud Pública y Laboral de Navarra, Servicio de Salud del País Vasco, Instituto Gestión Sanitaria. The Spanish Institute of Statistics provided the random selection of households and the information required for participants’ contact. We would like to thank all the nurses, general practitioners, administrative personnel, and other health-care workers who collaborated in this study and all participants. This study is the result of the efforts of many professionals and the trust and generosity of more than 60,000 participants who have understood the interest of providing time, information, and samples to learn about the situation of the COVID-19 epidemic in our country.

METHODS

The IFR was defined as the number of deaths due to COVID-19 divided by the number of individuals with SARS-CoV-2 infection in the non-institutionalized Spanish population.

Estimation of the number of SARS-CoV-2 infections

We calculated the prevalence of IgG antibodies against SARS-CoV-2 in the non-institutionalized Spanish population using data from ENE-COVID, a nationwide population-based serosurvey whose design has been described elsewhere. Briefly, 1,500 census tracts, and 24 households within each tract, were randomly selected using a stratified two-stage sampling. All residents of the 35,883 households were invited to participate in the study, carried out between April 27 and June 22, 2020 in three two-week rounds, with a one-week break between rounds. Epidemiologic questionnaires and serology tests were administered to 68,292 individuals who participated in at least one round. The study used two immunoassays to detect IgG antibodies: a point-of-care test (Orient Gene Biotech COVID-19 IgG/IgM Rapid test Cassette), and a chemiluminiscent microparticle immunoassay (CMI) that required venipuncture (SARS-CoV-2 IgG for use with ARCHITECT; Abbott Laboratories, Abbott Park, IL, USA; reference 06R8620), with better performance characteristics.

We calculated the seroprevalence, overall and in strata defined by age and sex, as the proportion of participants who had detectable IgG antibodies against SARS-CoV-2 in any round in the CMI test (61,092 participants had a valid CMI result). To account for the different sampling selection
probabilities by province and to adjust for non-response to the CMI test based on sex, age, and census tract average income, we assigned sampling weights to each study participant.\textsuperscript{21}

We then calculated the number of seropositive persons in Spain by multiplying the age- and sex-specific prevalences of IgG antibodies times the size of the corresponding non-institutionalized Spanish population groups as of July 15, 2020, provided by the National Institute of Statistics.\textsuperscript{23}

**Estimation of the number of deaths due to COVID-19**

Given the practical difficulties in reporting and adjudicating deaths from COVID-19 during the epidemic, we estimated the IFR separately using confirmed COVID-19 deaths and excess all-cause deaths.\textsuperscript{7} The two sources of information were the Spanish National Epidemiological Surveillance Network (RENAVE) and the Monitoring Mortality System (MoMo).

RENAVE\textsuperscript{17,24} provided individual data on the 29,137 laboratory-confirmed COVID-19 deaths registered in Spain up to July 15, 2020. The age and sex of 249 records with missing information were imputed based on the total sex and age distribution.

MoMo collects information on deaths from 3,945 municipal civil registries that cover 93\% of the Spanish population.\textsuperscript{25} Using a model described elsewhere,\textsuperscript{26} the data from MoMo is used to quantify excess deaths for a particular period, taking into account the historical series of the last 10 years and incorporating a secular trend and a seasonal component. Between March 1 and July 15, 44,459 excess all-cause deaths were estimated (mainly concentrated between March 13 and May 22).\textsuperscript{25}

Neither RENAVE nor MoMo distinguish between institutionalized and non-institutionalized population. It was estimated that 9,909 deaths with confirmed COVID-19 and 19,681 deaths attributed to suspected cases occurred in long-term care facilities, mainly nursing homes, during the same period (Supplementary Table 1). We subtracted these deaths from those identified by RENAVE and MoMo, respectively, in the population aged 60 years and older (see Supplementary Methods for details).
Estimation of infection fatality risks

We obtained separate estimates of the overall IFR using the COVID-19 deaths from RENAVE (lower bound of deaths, due to limited ascertainment in surveillance) and the excess all-cause deaths from MoMo (a possible upper bound because of the inclusion of deaths that may not result from direct or indirect effects of the epidemic). We then repeated the above analyses in each stratum defined by sex and 10-year age group. We calculated 95% confidence intervals based on delta methods that accounted for both the binomial variance in the number of deaths and the estimated design-based variance in the number of infections. Analyses were performed using survey commands in Stata, version 16 and survey package in R, version 3.

DATA AVAILABILITY STATEMENT

The manuscript includes all figures needed to replicate the IFR estimations and indicate the sources used. ENE-COVID seroprevalence figures are provided here for all sex and age groups. Data on deaths come from RENAVE and MoMo, two Spanish National Surveillance Systems. Anonymized data from these systems are available under request. The specific formulary for this purpose is provided by the Department of Communicable Diseases at the National Center for Epidemiology. Instituto de Salud Carlos III. C/ Monforte de Lemos nº 5 28029 Madrid. (e-mail: vigilancia.cne@isciii.es and mortalidad@isciii.es respectively). Population figures have been provided by the National Institute of Statistics and are publicly available at their website (www.ine.es).

CODE AVAILABILITY STATEMENT

The code for IFR calculation can be requested to rpastor@isciii.es and will be available at https://portalcne.isciii.es/enecovid19/

REFERENCES

1. Faust JS, Del Rio C. Assessment of Deaths From COVID-19 and From Seasonal Influenza. JAMA Intern Med 2020;


FIGURE LEGENDS

Figure 1. Infection fatality risk for SARS-CoV-2 based on (A) confirmed COVID-19 deaths and (B) excess deaths from all causes in non-institutionalized population, ENE-COVID study, April 27–June 22, 2020, Spain.
Table 1. Infection fatality risk for SARS-CoV-2 in non-institutionalized population by sex and age, ENE-COVID study, April 27–June 22, 2020, Spain.

<table>
<thead>
<tr>
<th>Sex, age (years)</th>
<th>Individuals in population, thousands</th>
<th>SARS-CoV-2 seroprevalence*, % (95% CI)</th>
<th>Individuals with SARS-CoV-2 antibodies, thousands (95% CI)</th>
<th>Confirmed COVID-19 deaths</th>
<th>Excess all-cause deaths</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>46,887.1</td>
<td>4.9 (4.6–5.3)</td>
<td>2,306.7 (2,153.6–2,470.1)</td>
<td>19,228</td>
<td>24,778</td>
</tr>
<tr>
<td>Men</td>
<td>23,006.9</td>
<td>4.8 (4.4–5.2)</td>
<td>1,106.0 (1,017.6–1,201.6)</td>
<td>12,317</td>
<td>15,480</td>
</tr>
<tr>
<td>0–9</td>
<td>2,205.5</td>
<td>3.2 (1.9–5.4)</td>
<td>71.7 (42.5–119.7)</td>
<td>3</td>
<td>32</td>
</tr>
<tr>
<td>10–19</td>
<td>2,557.9</td>
<td>3.7 (2.8–4.8)</td>
<td>93.5 (71.2–122.5)</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>20–29</td>
<td>2,479.1</td>
<td>5.8 (4.7–7.1)</td>
<td>142.9 (116.2–175.3)</td>
<td>18</td>
<td>0</td>
</tr>
<tr>
<td>30–39</td>
<td>2,978.7</td>
<td>4.7 (3.8–5.7)</td>
<td>139.7 (114.0–170.9)</td>
<td>48</td>
<td>3</td>
</tr>
<tr>
<td>40–49</td>
<td>3,916.7</td>
<td>5.3 (4.6–6.2)</td>
<td>209.0 (180.0–242.4)</td>
<td>192</td>
<td>168</td>
</tr>
<tr>
<td>50–59</td>
<td>3,493.8</td>
<td>5.3 (4.5–6.1)</td>
<td>184.0 (157.8–214.3)</td>
<td>705</td>
<td>601</td>
</tr>
<tr>
<td>60–69</td>
<td>2,598.2</td>
<td>4.9 (4.1–5.9)</td>
<td>127.2 (105.3–153.3)</td>
<td>1,904</td>
<td>2,065</td>
</tr>
<tr>
<td>70–79</td>
<td>1,783.7</td>
<td>4.7 (3.7–6.0)</td>
<td>83.7 (65.5–106.7)</td>
<td>4,145</td>
<td>5,114</td>
</tr>
<tr>
<td>≥80</td>
<td>993.3</td>
<td>4.6 (3.2–6.5)</td>
<td>45.6 (31.8–64.9)</td>
<td>5,299</td>
<td>7,497</td>
</tr>
<tr>
<td>Women</td>
<td>23,880.1</td>
<td>5.0 (4.7–5.4)</td>
<td>1,200.5 (1,110.5–1,297.4)</td>
<td>6,911</td>
<td>9,298</td>
</tr>
<tr>
<td>0–9</td>
<td>2,078.3</td>
<td>4.2 (2.7–6.7)</td>
<td>88.0 (55.1–139.0)</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>10–19</td>
<td>2,396.7</td>
<td>4.4 (3.4–5.6)</td>
<td>105.1 (81.7–134.7)</td>
<td>3</td>
<td>22</td>
</tr>
<tr>
<td>20–29</td>
<td>2,404.1</td>
<td>5.7 (4.6–7.0)</td>
<td>137.4 (111.2–169.3)</td>
<td>17</td>
<td>10</td>
</tr>
<tr>
<td>30–39</td>
<td>3,012.4</td>
<td>5.2 (4.4–6.2)</td>
<td>156.7 (132.0–185.8)</td>
<td>29</td>
<td>71</td>
</tr>
<tr>
<td>40–49</td>
<td>3,877.8</td>
<td>5.3 (4.6–6.2)</td>
<td>206.8 (177.9–240.0)</td>
<td>103</td>
<td>91</td>
</tr>
<tr>
<td>50–59</td>
<td>3,563.5</td>
<td>5.2 (4.5–6.0)</td>
<td>184.4 (158.8–213.8)</td>
<td>318</td>
<td>369</td>
</tr>
<tr>
<td>60–69</td>
<td>2,803.4</td>
<td>5.0 (4.2–6.0)</td>
<td>140.4 (117.2–167.9)</td>
<td>749</td>
<td>875</td>
</tr>
<tr>
<td>70–79</td>
<td>2,138.1</td>
<td>4.6 (3.7–5.8)</td>
<td>98.9 (79.0–123.4)</td>
<td>1,986</td>
<td>2,646</td>
</tr>
<tr>
<td>≥80</td>
<td>1,605.8</td>
<td>5.0 (3.7–6.8)</td>
<td>80.2 (58.7–108.9)</td>
<td>3,704</td>
<td>5,203</td>
</tr>
</tbody>
</table>

* Based on a chemiluminiscent microparticle immunoassay for IgG antibodies
A Deaths with confirmed COVID−19

Infection fatality risk for SARS−CoV−2 (%)

- Men
- Women

B Excess deaths from all causes

Infection fatality risk for SARS−CoV−2 (%)