medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

The infection fatality rate of COVID-19 inferred from seroprevalence data

John P.A. Ioannidis

Departments of Medicine, of Epidemiology and Population Health, of Biomedical Data Science, and of Statistics and Meta-Research Innovation Center at Stanford (METRICS), Stanford University, Stanford, CA, USA

E-mail: jioannid@stanford.edu

Funding: METRICS has been supported by a grant from the Laura and John Arnold Foundation

Disclosures: I am a co-author (not principal investigator) of one of the 23 seroprevalence studies.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

ABSTRACT

Objective To estimate the infection fatality rate of coronavirus disease 2019 (COVID-19) from data of seroprevalence studies.

Methods Population studies with sample size of at least 500 and published as peer-reviewed papers or preprints as of July 11, 2020 were retrieved from PubMed, preprint servers, and communications with experts. Studies on blood donors were included, but studies on healthcare workers were excluded. The studies were assessed for design features and seroprevalence estimates. Infection fatality rate was estimated from each study dividing the number of COVID-19 deaths at a relevant time point by the number of estimated people infected in each relevant region. Correction was also attempted accounting for the types of antibodies assessed. Secondarily, results from national studies were also examined from preliminary press releases and reports whenever a country had no other data presented in full papers of preprints.

Results 36 studies (43 estimates) were identified with usable data to enter into calculations and another 7 preliminary national estimates were also considered for a total of 50 estimates. Seroprevalence estimates ranged from 0.222% to 47%. Infection fatality rates ranged from 0.00% to 1.63% and corrected values ranged from 0.00% to 1.31%. Across 32 different locations, the median infection fatality rate was 0.27% (corrected 0.24%). Most studies were done in pandemic epicenters with high death tolls. Median corrected IFR was 0.10% in locations with COVID-19 population mortality rate less than the global average (<73 deaths per million as of July 12, 2020), 0.27% in locations with 73-500 COVID-19 deaths per million, and 0.90% in locations exceeding 500 COVID-19 deaths per million. Among people <70 years old, infection fatality rates ranged from 0.00% to 0.57% with median of 0.05% across the different locations (corrected median of 0.04%).

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Conclusions The infection fatality rate of COVID-19 can vary substantially across different locations and this may reflect differences in population age structure and case-mix of infected and deceased patients as well as multiple other factors. Estimates of infection fatality rates inferred from seroprevalence studies tend to be much lower than original speculations made in the early days of the pandemic.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

The infection fatality rate (IFR), the probability of dying for a person who is infected, is one of the most critical and most contested features of the coronavirus disease 2019 (COVID-19) pandemic. The expected total mortality burden of COVID-19 is directly related to the IFR. Moreover, justification for various non-pharmacological public health interventions depends crucially on the IFR. Some aggressive interventions that potentially induce also more pronounced collateral harms[1] may be considered appropriate, if IFR is high. Conversely, the same measures may fall short of acceptable risk-benefit thresholds, if the IFR is low.

Early data from China, adopted also by the World Health Organization (WHO),[2] focused on a crude case fatality rate (CFR) of 3.4%; CFR is the ratio of COVID-19 deaths divided by the number of documented cases, i.e. patients with symptoms who were tested and found to be PCR-positive for the virus. The WHO envoy who visited China also conveyed the message that there are hardly any asymptomatic infections.[3] With a dearth of asymptomatic infections, the CFR approximates the IFR. Other mathematical models suggested that 40-70%,[4] or even[5] 81% of the global population would be infected. Influential mathematical models[5,6] eventually dialed back to an IFR of 1.0% or 0.9%, and these numbers long continued to be widely cited and used in both public and scientific circles. The most influential of these models, constructed by Imperial College estimated 2.2 million deaths in the USA and over half a million deaths in the UK in the absence of lockdown measures.[5] Such grave predictions justifiably led to lockdown measures adopted in many countries. With 0.9% assumed infection fatality rate and 81% assumed proportion of people infected, the prediction would correspond to a global number of deaths comparable with the 1918 influenza, in the range of 50 million fatalities.

Since late March 2020, many studies have tried to estimate the extend of spread of the virus in various locations by evaluating the seroprevalence, i.e. how many people in population samples have developed antibodies for the virus. These studies can be useful because they may

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

inform about the extend of under-ascertainment of documenting the infection based on PCR testing. Moreover, they can help obtain estimates about the IFR, since one can divide the number of observed deaths by the estimated number of people who are inferred to have been infected.

At the same time, seroprevalence studies may have several caveats in their design, conduct, and analysis that may affect their results and their interpretation. Here, data available as of July 11, 2020 were collected, scrutinized, and used to infer estimates of IFR in different locations where these studies have been conducted.

METHODS

Seroprevalence studies

The input data for the calculations of IFR presented here are studies of seroprevalence of COVID-19 that have been done in the general population, or in samples that might approximate the general population (e.g. with proper reweighting) and that have been published in peer-reviewed journals or have been submitted as preprints as of July 11, 2020. Only studies with at least 500 assessed samples were considered, since smaller datasets would entail extremely large uncertainty for any calculations to be based on them. When studies focused on making seroprevalence assessments at different time interval, they were eligible if at least one time interval assessment had a sample size of at least 500 participants; among different eligible time points, the one with the highest seroprevalence was selected, since seroprevalence may decrease over time as antibody titers wane. Studies with data collected over more than a month, and that could not be broken into at least one eligible time interval that did not exceed one month in duration were excluded, since it would not be possible to estimate a point seroprevalence with any reliability. Studies were eligible regardless of the exact age range of included participants, but studies including only children were excluded.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Studies where results were only released through press releases were not considered here, since it is very difficult to tell much about their design and analysis, and this is fundamental in making any inferences based on their results. Nevertheless, secondarily, results from national studies were also examined from preliminary press releases and reports whenever a country had no other data presented in full papers of preprints as of July 11, 2020. This allowed these countries to be represented in the collected data, but extra caution is required given the preliminary nature of this information. Preprints should also be seen with caution since they have not been yet fully peer-reviewed (although some of them have already been revised based on very extensive comments from the scientific community). However, in contrast to press releases, preprints typically offer at least a fairly complete paper with information about design and analysis.

Studies done of blood donors were eligible, although it is possible they may underestimate seroprevalence and overestimate IFR due to healthy volunteer effect. Studies done on health care workers were not eligible, since they deal with a group at potentially high exposure risk which may lead to seroprevalence estimates much higher than the general population and thus implausibly low IFR. For a similar reason, studies focused on communities (e.g. shelters or religious or other shared-living communities) were also excluded. Studies were eligible regardless of whether they aimed to evaluate seroprevalence in large or small regions, provided that the population of reference in the region was at least 5000 people.

Searches were made in PubMed (LitCOVID), medRxiv, bioRxiv, and Research Square using the terms “seroprevalence” and “antibodies” with continuous updates (last update July 11, 2020). Communication with colleagues who are field experts sought to ascertain if any major studies might have been missed.

Information was extracted from each study on location, recruitment and sampling strategy, dates of sample collection, sample size, types of antibody used (IgG, IgM, IgA), estimated crude

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

seroprevalence (positive samples divided by all samples test), and adjusted seroprevalence and features that were considered in the adjustment (sampling process, test performance, presence of symptoms, other).

Calculation of inferred IFR

Information on the population of the relevant location was collected from the papers. Whenever it was missing, it was derived based on recent census data trying to approximate as much as possible the relevant catchment area (e.g. region(s) or county(ies)), whenever the study did not pertain to an entire country. Some studies targeted specific age groups (e.g. excluding elderly people and/or excluding children) and some of them made inferences on number of people infected in the population based on specific age groups. For consistency, the entire population, as well as, separately, only the population with age <70 years were used for estimating the number of infected people. It was assumed that the seroprevalence would be similar in different age groups, but significant differences in seroprevalence according to age strata that had been noted by the original authors were also recorded to examine the validity of this assumption.

The number of infected people was calculated multiplying the relevant population with the adjusted estimate of seroprevalence. Whenever an adjusted seroprevalence estimate had not been obtained, the unadjusted seroprevalence was used instead. When seroprevalence estimates with different adjustments were available, the analysis with maximal adjustment was selected.

For the number of COVID-19 deaths, the number of deaths recorded at the time chosen by the authors of each study was selected, whenever the authors used such a death count up to a specific date to make inferences themselves. If the choice of date had not been done by the authors, the number of deaths accumulated until after 1 week of the mid-point of the study period was chosen. This accounts for the differential delay in developing antibodies versus dying from the infection. It should be acknowledged that this is an averaging approximation, because some

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

patients may die very soon (within <3 weeks) after infection (and thus are overcounted), and others may die very late (and thus are undercounted due to right censoring).

The inferred IFR was obtained by dividing the number of deaths by the number of infected people for the entire population, and separately for people <70 years old. The proportion of COVID-19 deaths that occurred in people <70 years old was retrieved from situational reports for the respective countries, regions, or counties in searches done in June 3-7 for studies published until June 7 and in July 3-11 for studies published later. A corrected IFR is also presented, trying to account for the fact that only one or two types of antibodies (among IgG, IgM, IgA) might have been used. Correcting seroprevalence upwards (and inferred IFR downwards) by 1.1-fold for not performing IgM measurements and similarly for not performing IgA measurements may be reasonable, based on some early evidence,[7] although there is uncertainty about the exact correction factor.

Data synthesis considerations

Inspection of the IFR estimates across all locations showed vast heterogeneity with heterogeneity I^2 exceeding 99.9% and thus a meta-analysis would be inappropriate to report across all locations. Quantitative synthesis with meta-analysis across all locations would also be misleading since locations with high seroprevalence would tend to carry more weight than locations with low seroprevalence; locations with more studies (typically those that have attracted more attention because of high death tolls and thus high IFRs) would be represented multiple times in the calculations; and more sloppy studies with fewer adjustments would get more weight, because they would have spuriously tighter confidence intervals than more rigorous studies with more careful adjustments allowing for more uncertainty. Finally, with a highly skewed IFR distribution and with extreme between-study heterogeneity, synthesis with a typical random effects model would tend to produce an erroneously high summary IFR that approximates the

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

mean of the study-specific estimates (also heavily driven by hotbed high-mortality locations with more studies done), while for a skewed distribution the median is more appropriate.

Therefore, at a first step, IFR estimates from studies done in the same country (or in the US, the same state) were grouped together and a single IFR was obtained for that location, weighting the study-specific IFRs by the sample size of each study. This allowed to avoid giving inappropriately more weight to studies with higher seroprevalence estimates and those with seemingly tighter confidence intervals because of poor or no adjustments, while still giving more weight to larger studies. Then, a single summary estimate was used for each location and the median of the distribution of location-specific IFR estimates was calculated. Finally, it was explored whether the location-specific IFRs were associated with the COVID-19 mortality rate in the population (COVID-19 deaths per million people) in each location as of July 12, 2020; this allowed to assess whether IFR estimates tend to be higher in harder hit locations.

RESULTS

Seroprevalence studies

36 studies with a total of 43 eligible estimates were published either in the peer-reviewed literature or as preprints as of July 11, 2020.[8-43] Dates and processes of sampling and recruitment are summarized in Table 1, sample sizes, antibody types assessed and regional population appear in Table 2, estimated prevalence, and number of people infected in the study region are summarized in Table 3, and number of COVID-19 and inferred IFR estimates are found in Table 4. Several studies performed repeated seroprevalence surveys at different time points, and only the time point with the highest seroprevalence estimate is considered in the calculations. With three exceptions, this is also the latest time point. Furthermore, another 7 preliminary national estimates were also considered (Table 5)[44-50] from countries that had no other seroprevalence study published as a full paper or preprint. This yielded a total of 50 eligible estimates.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

At least seven studies found some statistically significant, modest differences in seroprevalence rates across some age groups (Oise: decreased seroprevalence in age 0-14, increased in age 15-17; Geneva: decreased seroprevalence in age >50; Netherlands: increased seroprevalence in age 18-30; New York state: decreased seroprevalence in age >55; Brooklyn: decreased seroprevalence in age 0-5, increased in age 16-20; Tokyo: increased seroprevalence in age 18-34, Spain: decreased seroprevalence in age 0-10, Belgium: higher seroprevalence in age >90). The patterns are not strong enough to suggest major differences in extrapolating across age groups, although higher values in adolescents and young adults and lower values in children cannot be excluded.

As shown in Table 1, these studies varied substantially in sampling and recruitment designs. The main issue is whether they can offer a representative picture of the population in the region where they are performed. A generic problem is that vulnerable people who are at high risk of infection and/or death may be more difficult to recruit in survey-type studies. COVID-19 infection seems to be particularly widespread and/or lethal in nursing homes, among homeless people, in prisons, and in disadvantaged minorities. Most of these populations are very difficult, or even impossible to reach and sample from and they are probably under-represented to various degrees (or even entirely missed) in surveys. This would result in an underestimation of seroprevalence and thus overestimation of IFR. Eleven of the 36 studies that are available as full papers (Iran,[8] Geneva,[10] Gangelt,[16] Rio Grande do Sul,[17] Luxembourg,[20] Los Angeles county,[22] three Brazil studies,[25,34,42], Spain,[36] and Louisiana[37]) explicitly aimed for random sampling from the general population. In principle, this is a stronger design. However, even with such designs, people who cannot be reached (e.g. by e-mail or phone or even visiting them at a house location) will not be recruited, and these vulnerable populations are likely to be missed. Moreover, 5 of these 11 studies[8,10,16,42,37] focused on studying geographical locations that had extreme numbers of

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

deaths, higher than other locations in the same city or country, and this would tend to select eventually for higher IFR on average.

Seven studies assessed blood donors in Denmark,[12] Netherlands,[15] Scotland,[18] the Bay Area in California,[24] Zurich/Lucerne,[28] Apulia[31] and Rio De Janeiro.[41] By definition these studies include people in good health and without symptoms, at least recently, and therefore may markedly underestimate COVID-19 seroprevalence in the general population. A small set of 200 blood donors in Oise, France[13] showed 3% seroprevalence, while pupils, siblings, parents, teachings and staff at a high school with a cluster of cases in the same area had 25.9% seroprevalence; true population seroprevalence may be between these two values.

For the other studies, healthy volunteer bias may lead to underestimating seroprevalence and this is likely to have been the case in at least one case (the Santa Clara study)[19] where wealthy healthy people were rapidly interested to be recruited when the recruiting Facebook ad was released. The design of the study anticipated correction with adjustment of the sampling weights by zip code, gender, and ethnicity, but it is likely that healthy volunteer bias may still have led to some underestimation of seroprevalence. Conversely, attracting individuals who might have been concerned of having been infected (e.g. because they had symptoms) may lead to overestimation of seroprevalence in surveys. Finally studies of employees, grocery store clients, or patient cohorts (e.g. hospitalized for other reasons, or coming to the emergency room, or studies using residual lab samples) may have sampling bias with unpredictable direction.

As shown in Table 2, all studies have tested for IgG antibodies, but only about half have also assessed IgM, 4 have assessed IgA. Only three studies assessed all three types of antibodies and one more used a pan-Ig antibody. Studies typically considered the results to be “positive” if any tested antibody type was positive, but one study (Luxembourg) that considered the results to be “positive” only if both IgG and IgA were detected. The ratio of people sampled versus the total

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

population of the region was better than 1:1000 in 11 studies (Idaho,[9] Denmark blood donors,[12] Gangelt,[16] Santa Clara,[19] Luxembourg,[20] Brooklyn,[27] Zurich,[28] San Francisco,[33] Espirito Santo,[34] Spain,[36] and Vitacura[43]).

Seroprevalence estimates

As shown in Table 3, prevalence ranged from as little as 0.222% to as high as 47%. Studies varied a lot on whether they tried or not to adjust their estimates for test performance, sampling (striving to get closer to a more representative sample), and clustering effects (e.g. when including same household members) as well as other factors. The adjusted seroprevalence occasionally differed substantially from the crude, unadjusted value. In principle adjusted values are likely to be closer to the true estimate, but the exercise shows that each study alone may have some unavoidable uncertainty and fluctuation, depending on the analytical choices preferred. In studies that sampled people from multiple locations, large between-location heterogeneity could be seen (e.g. 0-25% across 133 Brazilian cities).[25]

Inferred IFR

Inferred IFR estimates varied a lot, from 0.00% to 1.63%. Corrected values also varied extensively, from 0.00% to 1.31%. For 10 locations, more than one IFR estimate was available and thus IFR from different studies evaluating the same location could be compared. As shown in figure 1, the IFR estimates tended to be more homogeneous within each location, while they differed a lot across locations. The sample size-weighted summary was used to generate a single estimate to represent each location. Data were available for 32 different locations. The median IFR across all 32 locations was 0.27% (0.24% using the corrected values). Most data came from locations with high death tolls and 23 of the 32 locations had a population mortality rate (deaths per million population) higher than the global average (73 deaths per million population as of July 12) (Figure 2). The uncorrected IFR estimates had a range of 0.01-0.16% (median 0.13%) across

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

the 9 locations with population mortality rate below the global average, 0.07-0.73% (median 0.27%) across the 15 locations with population mortality rate above the global average but below 500 deaths per million population, and 0.59-1.63% (median 1.12%) across the 8 extreme hotbed locations with over 500 deaths per million population. The corrected IFR estimates had medians of 0.10%, 0.25%, and 0.90%, respectively, for the three groups of locations.

The proportion of COVID-19 deaths that occurred in people <70 years old varied substantially across locations. All deaths in Gangelt were in elderly people while in Wuhan half the deaths occurred in people <70 years old and the proportion might have been higher in Iran, but no data could be retrieved for this country. When limited to people <70 years old, IFR ranged from 0.00% to 0.57% with median of 0.05% (corrected, 0.00-0.46% with median of 0.04%). IFR estimates in people <70 years old were lower than 0.1% in all but 7 locations that were hard-hit hotbeds (Belgium, Wuhan, Italy, Spain, Connecticut, Louisiana, New York).

DISCUSSION

IFR is not a fixed physical constant and it can vary substantially across locations, depending on the population structure, the case-mix of infected and deceased individuals and other, local factors. Inferred IFR values based on emerging seroprevalence studies typically show a much lower fatality than initially speculated in the earlier days of the pandemic.

The studies analyzed here represent 50 different estimates of IFR, but they are not fully representative of all countries and locations around the world. Most of them come from locations with overall COVID-19 mortality rates exceeding the global average (73 deaths per million people as of July 12). The median inferred IFR in locations with COVID-19 mortality rate below the global average is low (0.13%, corrected 0.10%). For hotbed countries with COVID-19 mortality rates above the global average but lower than 500 deaths per million, the median IFR is still not that high (median 0.27%, corrected 0.25%). Very high IFR estimates have been documented

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

practically in locations that had devastating experiences with COVID-19. Such epicenters are unusual across the globe, but they are overrepresented in the 50 seroprevalence estimates available for this analysis. Therefore, if one could sample equally from all countries and locations around the globe, the median IFR might be even lower than the one observed in the current analysis.

Several studies in hard-hit European countries inferred modestly high IFR estimates for the overall population, but the IFR was still low in people <70 years old. Some of these studies were on blood donors and may have underestimated seroprevalence and overestimated IFR. One study in Germany aimed to test the entire population of a city and thus selection bias is minimal: Gangelt[16] represents a situation with a superspreader event (in a local carnival) and 7 deaths were recorded, all of them in very elderly individuals (average age 81, sd 3.5). COVID-19 has a very steep age gradient of death risk.[51] It is expected therefore that in locations where the infection finds its way into killing predominantly elderly citizens, the overall, age-unadjusted IFR would be higher. However, IFR would still be very low in people <70 in these locations, e.g. in Gangelt IFR is 0.00% in non-elderly people. Similarly, in Switzerland, 69% of deaths occurred in people >80 years old[51] and this explains the relatively high overall IFR in Geneva and Zurich. Similar to Germany, very few deaths in Switzerland have been recorded in non-elderly people, e.g. only 2.5% have occurred in people <60 years old and IFR in that age-group would be ~0.01%. The majority of deaths in most of the hard hit European countries have happened in nursing homes[52] and a large proportion of deaths also in the US[53] also follow this pattern. Moreover, many nursing home deaths have no laboratory confirmation and thus should be seen with extra caution in terms of the causal impact of SARS-CoV-2.

Locations with high burdens of nursing home deaths may have high IFR estimates, but the IFR would still be very low among non-elderly, non-debilitated people. The average length of stay in a nursing home is slightly more than 2 years and people who die in nursing homes die in a

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

median of 5 months[54] so many COVID-19 nursing home deaths may have happened in people with life expectancy of only a few months. This needs to be verified in careful assessments of COVID-19 outbreaks in nursing homes with detailed risk profiling of fatalities. If COVID-19 happened in patients with very limited life expectancy, this pattern may even create a dent of less than expected mortality in the next 3-6 months after the coronavirus excess mortality wave. As of July 12 (week 28), preliminary Euromonitor data[55] indeed already show a substantial dent below baseline mortality in France, and a dent below baseline mortality is seen also for the aggregate European data.

Within China, the much higher IFR estimates in Wuhan versus other areas may reflect the wide spread of the infection to hospital personnel and the substantial contribution of nosocomial infections to a higher death toll in Wuhan;[56] plus unfamiliarity with how to deal with the infection in the first location where COVID-19 arose. Massive deaths of elderly individuals in nursing homes, nosocomial infections, and overwhelmed hospitals may also explain the very high fatality in specific locations in Italy[57] and in New York and neighboring states. Seroprevalence studies in health care workers and administrative hospital staff in Lombardy[58] found 8% seroprevalence in Milan hospitals and 35-43% in Bergamo hospitals, supporting the scenario for widespread nosocomial infections among vulnerable patients. The high IFR values in New York metropolitan area and neighboring states are also not surprising, given the vast death toll witnessed. A very unfortunate decision of the several state governors was to have COVID-19 patients sent to nursing homes. Moreover, some hospitals in New York City hotspots reached maximum capacity and perhaps could not offer optimal care. Use of unnecessarily aggressive management (e.g. mechanical ventilation) and hydroxychloroquine may also have contributed to worse outcomes. Furthermore, New York City has an extremely busy, congested public transport system that may have exposed large segments of the population to high infectious load in close contact

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

transmission and, thus, perhaps more severe disease. A more aggressive viral clade has also been speculated, but this needs further verification.[59]

IFR may reach very high levels among disadvantaged populations and settings that have the worst combination of factors predisposing to higher fatalities. Importantly, such hotspot locations are rather uncommon exceptions in the global landscape. Moreover, even in these locations, the IFR for non-elderly individuals without predisposing conditions may remain very low. E.g. in New York City only 0.65% of all deaths happened in people <65 years without major underlying conditions.[51] Thus the IFR even in New York City would probably be lower than 0.01% in these people.

Studies with extremely low inferred IFR are also worthwhile discussing. Possible overestimation of seroprevalence and undercounting of deaths need to be considered. E.g., for Kobe, the authors of the study[11] raise the question whether COVID-19 deaths have been undercounted in Japan. Both undercounting and overcounting of COVID-19 deaths may be a caveat in different locations and this is difficult to settle in the absence of very careful scrutiny of medical records and autopsies. The Tokyo data,[29] nevertheless, also show similarly very low IFR. Moreover, evaluation of all-cause mortality in Japan has shown no excess deaths during the pandemic, consistent with the possibility that somehow the Japanese population was spared. Very low IFRs seem common in Asian countries, including China (excluding Wuhan), Iran, Israel and India. Former immunity from exposure to other coronaviruses, genetic differences, hygienic etiquette, lower infectious load, and other unknown factors may be speculated. IFR seems to be very low also in Singapore where extensive PCR testing was carried out. As of July 12, 2020, in Singapore there were only 26 deaths among 46,283 cases, suggesting an upper bound of 0.06% for IFR, even if no cases had been missed.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Some surveys have also been designed to assess seroprevalence repeatedly spacing out measurements in the same population over time. A typical pattern that seems to emerge is that seroprevalence may increase several fold within a few weeks, but plateau or even decline may follow.[10,28] A more prominent decline of seropositivity was seen in a study in Wuhan.[32] Genuine decrease may be difficult to differentiate from random variation. However, some preliminary data[60,61] suggest that decrease in antibody titers may be fast. Decrease in seropositivity over time means that the numbers of infected people may be underestimated and IFR overestimated.

The only data from a low-income country among the 23 studies examined here come from Iran[8] and India[49] and the IFR estimates appears to be very low. Iran has a young population with only slightly over 1% of the age pyramid at age >80 and India's population is even younger. Similar considerations apply to almost every less developed country around the world. Given the very sharp age gradient and the sparing of children and young adults from death by COVID-19, one may expect IFR to be fairly low in the less developed countries. However, it remains to be seen whether comorbidities, poverty and frailty (e.g. malnutrition) and congested urban living circumstances may have adverse impact on risk and thus increase IFR also in these countries.

One should caution that the extent of validation of the antibody assays against positive and negative controls differs across studies. Specificity has typically exceeded 99.0%, which is reassuring. However, for very low prevalence rates, even 99% specificity may be problematic. Sensitivity also varies from 60-100% in different validation exercises and for different tests, but typically it is closer to the upper than the lower bound. One caveat about sensitivity is that typically the positive controls are patients who had symptoms and thus were tested and found to be PCR-positive. However, it is possible that symptomatic patients may be more likely to develop antibodies than patients who are asymptomatic or have minimal symptoms and thus had not sought PCR testing.[61-65] For example, one study found that 40% of asymptomatic patients became

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

seronegative within 8 weeks.[61] Since the seroprevalence studies specifically try to unearth these asymptomatic/mildly symptomatic missed infections, a lower sensitivity for these mild infections could translate to substantial underestimates of the number of infected people and substantial overestimate of the inferred IFR.

The corrected IFR estimates are trying to account for undercounting of infected people when not all 3 antibodies (IgG, IgM, and IgA) are assessed.[7] However, the magnitude of the correction is uncertain and may also vary in different circumstances. Moreover, it is possible that an unknown proportion of people may have handled the virus using immune mechanisms (mucosal, innate, cellular) that did not generate any serum antibodies.[66-69] This may lead to substantial underestimation of the frequency of infection and respective overestimation of the IFR. One study has found indeed that mild SARS-CoV-2 infections may lead to nasal release of IgA, without serum antibody response.[68] Another study has found that 6 of 8 interfamilial contacts of index cases remained seronegative despite developing symptoms and 6 of 8 developed persisting T cell responses[69] and the important role of cellular immune responses even in seronegative patients has been documented also by other investigators.[70]

An interesting observation is that even under congested circumstances, like cruise ships, aircraft carriers or homeless shelter, the proportion of people detected positive typically does not get to exceed 20-45%.[71,72] Similarly, at a wider population level, values ~47% are the maximum values documented to-date and most values are much lower, yet epidemic waves seem to wane. It has been suggested[73,74] that differences in host susceptibility and behavior can result in herd immunity at much lower prevalence of infection in the population than originally expected. COVID-19 spreads by infecting certain groups more than others because some people have much higher likelihood of exposure. People most likely to be exposed also tend to be those most likely to spread for the same reasons that put them at high exposure risk. In the absence of random

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

mixing of people, the epidemic wave may be extinguished even with relatively low proportions of people becoming infected. Seasonality may also play a role in the dissipation of the epidemic wave. It has also been observed that many people have CD4 cellular responses to SARS-CoV-2 even without being exposed to this virus, perhaps due to prior exposure to other coronaviruses.[75] It is unknown whether this proportion varies in different populations around the world and whether this immunity may contribute to SARS-CoV-2 epidemic waves waning without infecting a large share of the population.

A major limitation of the current analysis is that the calculations presented in this paper include several studies that have not yet been fully peer-reviewed. Moreover, there are several studies that are still ongoing. New emerging data may offer more insights and updated estimates. Given that the large majority of studies have been done in locations that were hard hit from COVID-19, it would be useful to do more studies in less hit locations, so as to have a more balanced global perspective.

A comparison of COVID-19 to influenza is often attempted, but many are confused by this comparison unless placed in context. Based on the IFR estimates obtained here, COVID-19 may have infected as of July 12 approximately 300 million people (or more), far more than the ~13 million PCR-documented cases. The global COVID-19 death toll is still evolving, but it is still not much dissimilar to a typical death toll from seasonal influenza (290,000-650,000),[76] while "bad" influenza years (e.g. 1957-9 and 1968-70) have been associated with 1-4 million deaths.[77] Notably, influenza devastates low-income countries, but is more tolerant of wealthy nations, probably because of the availability and wider use of vaccination in these countries.[58] Conversely, in the absence of vaccine and with a clear preference for elderly debilitated individuals, COVID-19 may have an inverse death toll profile, with more deaths in wealthy nations than in low-income countries. However, even in the wealthy nations, COVID-19 seems to affect predominantly the

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

frail, the disadvantaged, and the marginalized – as shown by high rates of infectious burden in nursing homes, homeless shelters, prisons, meat processing plants, and the strong racial/ethnic inequalities against minorities in terms of the cumulative death risk.[78,79]

While COVID-19 is a formidable threat, the fact that its IFR is typically much lower than originally feared, is a welcome piece of evidence. The median IFR found in this analysis is very similar to the estimate recently adopted by CDC for planning purposes.[80] The fact that IFR can vary substantially also based on case-mix and settings involved also creates additional ground for evidence-based, more precise management strategies. Decision-makers can use measures that will try to avert having this lethal virus infect people and settings who are at high risk of severe outcomes. These measures may be possible to be more precise and tailored to specific high-risk individuals and settings than blind lockdown of the entire society. Of course, uncertainty remains about the future evolution of the pandemic, e.g. the presence and height of subsequent waves.[81] However, it is helpful to know that SARS-CoV-2 has relatively modest IFR overall and that possibly IFR can be made even lower with appropriate, precise non-pharmacological choices.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table 1. Seroprevalence studies on COVID-19 published or depositing preprints as of July 11, 2020: dates, sampling and recruitment process

Location	Dates	Sampling and recruitment
Iran (Guilan)[8]	April (until April 21)	Population-based cluster random sampling design through phone call invitation, household-based.
Idaho (Boise)[9]	Late April	People from the Boise, Idaho metropolitan area, part of the Crush the Curve initiative.
Switzerland (Geneva)[10]	April 6-May 9 (5 consecutive weeks)	Randomly selected previous participants of the Bus Santé study with an email (or phone contact, if e-mail unavailable); participants were invited to bring all members of their household, aged 5 years and older.
Japan (Kobe)[11]	March 31-April 7	Randomly selected patients who visited outpatient clinics and received blood testing for any reason. Patients who visited the emergency department or the designated fever consultation service were excluded.
Denmark blood donors[12]	April 6-May 3	All Danish blood donors aged 17-69 years giving blood. Blood donors are healthy and must comply with strict eligibility criteria; they must self-defer for two weeks if they develop fever with upper respiratory symptoms.
France (Oise)[13]	March 30-April 4	Pupils, their parents and siblings, as well as teachers and non-teaching staff of a high-school.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

China (Wuhan)[14]	April 3-15	People applying for a permission of resume (n=1,021) and hospitalized patients during April 3 to 15 (n=381).
Netherlands blood donors[15]	April 1-15	Blood donors. Donors must be completely healthy, but they may have been ill in the past, provided that they recovered at least two weeks before.
Germany (Gangelt)[16]	March 30-April 6	600 adult persons with different surnames in Gangelt were randomly selected, and all household members were asked to participate in the study.
Brazil (Rio Grande do Sul)[17]	May 9-11 (third round, after April 11-13, and 25-27)	Multi-stage probability sampling was used in each of 9 cities to select 500 households, within which one resident was randomly chosen for testing.
Scotland blood donors[18]	March 21-23	Blood donors. Donors should not have felt unwell in the last 14 days, also some other deferrals applied regarding travel and COVID-19 symptoms.
California (Santa Clara)[19]	April 2-3	Facebook ad with additional targeting by zip code.
Luxembourg[20]	April 16-May 5	Representative sample (no details how ensured), 1807 of 2000 contacted provided data, were <79 years and had serology results.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Germany (Frankfurt)[21]	April 6-14	Employees of Infraserv Höchst, a large industrial site operator in Frankfurt am Main. No exclusion criteria.
California (Los Angeles)[22]	April 10-14	Proprietary database representative of the county. A random sample of these residents was invited, with quotas for enrollment for subgroups based on age, sex, race, and ethnicity distribution.
New York[23]	April 19-28	Convenience sample of patrons ≥18 years and residing in New York State, recruited consecutively upon entering 99 grocery stores and via an in-store flyer.
California (Bay Area)[24]	March	1,000 blood donors in diverse Bay Area locations (excluding those with self-reported symptoms or abnormal vital signs)
Brazil[25]	May 15-22	Sampling from 133 cities (the main city in each region), selecting 25 census tracts with probability proportionate to size in each sentinel city, and 10 households at random in each tract. Aiming for 250 participants per city.
Croatia[26]	April 23-28	DIV factory workers in Split and Sibenik-Knin invited for voluntary testing
New York (Brooklyn)[27]	Early May	Patients seen in urgent care facility in Brooklyn
Switzerland (Zurich)[28]	Prepandemic until June (patients)	Patients at the University Hospital of Zurich and blood donors in Zurich and Lucerne

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

	and May (blood donors)	
Japan (Tokyo)[29]	April 21-May 20	Two community clinics located in the major railway stations in Tokyo (Navitas Clinic Shinjuku and Tachikawa)
Spain (Barcelona)[30]	April 14-May 5	Consecutive pregnant women for first trimester screening or delivery in two hospitals
Italy (Apulia) blood donors[31]	May 1-May 31	Blood donors 18-65 years old free of recent symptoms possibly related to COVID-19, no close contact with confirmed cases, symptoms free during the preceding 14 days, no contacts with suspected cases
China (Wuhan B)[32]	March 26-April 28	Age 16-64, going back to work, with no fever, headache, or other symptoms of COVID-19
California (San Francisco)[33]	April 25-April 28 (and n=40 in May)	U.S. census tract 022901 population-dense area (58% Latinx) in San Francisco Mission district, expanded to neighboring blocks on April 28
Brazil (Espirito Santo)[34]	May 13-15	Cross-sectional of major municipalities structured over houses as the sampling units
USA (six states)[35]	March 23-April 1 (WA Puget Sound and NYC), April 6-10 (south Florida), April 20-	Convenience samples using residual sera obtained for routine clinical testing (screening or management) by two commercial laboratory companies

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

	26 (MO), April 20-May 3 (UT), April 26-May 3 (CT)	
Spain[36]	April 27-May 11	35883 households selected from municipal rolls using two-stage random sampling stratified by province and municipality size, with all residents invited to participate (75.1% of all contacted individuals participated)
Louisiana (Orleans and Jefferson Parish)[37]	May 9-15	Pool of potential participants reflective of the demographics of the Parishes was based on 50 characteristics, then a randomized subset of 150,000 was selected, then 25,000 were approached with digital aps, and 2640 recruited.
Belgium[38]	March 30-April 5 and April 20-26	Residual sera from ten private diagnostic laboratories in Belgium, with fixed numbers per age group, region and periodical sampling, and stratified by sex
France (Crepy-en-Valois)[39]	April 28-30	Pupils, their parents and relatives, and staff of primary schools exposed to SARS-CoV-2 in February and March 2020 in a city north of Paris
China (several regions)[40]*	March 30-April 10	Voluntary participation by public call for hemodialysis patients (n=979 in Zingzhou, Ubei and n=563 in Guangzhou/Foshun, Guangdong)

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

		and outpatients in Chingqing (n=993), and community residents in Chengdu, Sichuan (n=9442), and required testing for factory workers in Guangzhou, Guandong (n=442)
Brazil (Rio de Janeiro) blood donors[41]	April 14-27 (eligible: April 24-27)	Blood donors could not have had flulike symptoms within the 30 days before donation; had close contact with suspected or confirmed covid-19 cases in the 30 days before donation; or traveled abroad in the past 30 days.
Brazil (Sao Paulo)[42]	May 4-12	Randomly selected adults and their cohabitants sampled form 6 districts of Sao Paulo City with high number of cases
Chile (Vitacura)[43]	May 4-19	Classroom stratified sample of children and all staff in a community placed on quarantine after school outbreak

Two of the studies included additional datasets of <500 participants that are not presented here (n=200 blood donors in Oise and n=387 patients in the California (Bay Area) study)
*not considered here are some sub-cohorts from this study, including healthcare workers, and staff from hotel for healthcare workers

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table 2. Sample size, types of antibodies, and population in relevant region

Location	Sample size	Antibody	Population in region*	Population <70 years (%)
Iran (Guilan)[8]	551	IgG/IgM	2354848	95
Idaho (Boise)[9]	4856	IgG	481587 (Ada county)	92
Switzerland (Geneva)[10]	577 (4/20-27)	IgG	500000	88
Japan (Kobe)[11]	1000	IgG	1518870	79 (Japan)
Denmark blood donors[12]	20640	IgG/IgM	5771876	86
France (Oise)[13]	661	IgG	5978000 (Hauts-de-France)	89
China (Wuhan)[14]	1401	IgG/IgM	11080000	93 (China)
Netherlands blood donors[15]	7361	IgG/IgM/IgA	17097123	86
Germany (Gangelt)[16]	919	IgG/IgA	12597	86
Brazil (Rio Grande do Sul)[17]	4500	IgG	11377239	91
Scotland blood donors[18]	500	IgG	5400000	88
California (Santa Clara)[19]	3300	IgG/IgM	1928000	90
Luxembourg[20]	1807	IgG/IgA**	615729	90
Germany (Frankfurt)[21]	1000	IgG	2681000***	84 (Germany)
California (Los Angeles)[22]	863	IgG/IgM	7892000	92
New York[23]	15101	IgG	19450000	90
California (Bay Area)[24]	1000	IgG	7753000	90
Brazil (133 cities)[25]	24995	IgG/IgM	74656499	94 (Brazil)

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Croatia[26]	1494	IgG/IgM	4076000	86
New York (Brooklyn)[27]	11092	IgG	2559903	91
Switzerland (Zurich)[28]	1644 patients (4/1-15) 1640 blood donors (May)	IgG	1520968 (canton Zurich) 1930525 (Zurich+Lucerne)	88
Japan (Tokyo)[29]	1071	IgG	13902077	79 (Japan)
Spain (Barcelona)[30]	874	IgG/IgM/IgA	7566000 (Catalonia)	86
Italy (Apulia) blood donors[31]	909	IgG/IgM	4029000	84
China (Wuhan B)[32]	1196 (4/4-8)	IgG/IgM	11080000	93 (China)
California (San Francisco)[33]	3953	IgG (also PCR testing)	5174 (census 022901)	95
Brazil (Espirito Santo)[34]	4608	IgG/IgM	4018650	94 (Brazil)
USA (six states)[35]		Pan-Ig		
WA Puget Sound	3264		4273548	90 (WA)
UT	1132		3282120	92
NYC	2482		9260870	89
MO	1882		6110800	88
South FL	1742		6345345	86 (FL)
CT	1431		3562989	88
Spain[36]	61075	IgG	46940000	85
Louisiana (Orleans and Jefferson Parish)[37]	2640	IgG	825057	92 (LA)

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Belgium[38]	3391 (4/20-26)	IgG	11589623	86
France (Crepy-en-Valois)[39]	1340	IgG	5978000 (Hauts-de-France)	89
China (several regions)[40]		IgG/IgM		
Hubei (not Wuhan)	979		48058000	93 (China)
Chongqing	993		31243200	93 (China)
Sichuan	9442		83750000	93 (China)
Guangdong	1005		115210000	93 (China)
Brazil (Rio de Janeiro) blood donors[41]	669 (4/24-27)	IgG/IgM	17264943	94 (Brazil)
Brazil (Sao Paulo)[42]	517	IgG/IgM	298240 (6 districts)	94 (Brazil)
Chile (Vitacura)[43]	1244	IgG/IgM	85000	92 (Chile)

*The authors of some studies preferred to focus on age-restricted populations: 17-70 years old in the Denmark blood donor study (n=3800000), those 18-79 years old in the Luxembourg study (n=483000); those <70 years old in Netherlands blood donor study (n=13745768); those >=18 years old in the New York state study (n=15280000); and those >19 years old in the UT population of the six states US study (n=2173082).
**considered positive if both IgG and IgA were positive
***participants were recruited from a large number of districts, but most districts had very few participants; here the population of the 9 districts with >1:10,000 sampling ratio is included (846/1000 participants came from these 9 districts).

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table 3. Prevalence of infection and estimated number of infected people

Location	Seroprevalence (%)		Estimated
	Crude	Adjusted (adjustments)	infected
Iran (Guilan)[8]	22.0	33.0 (test, sampling)	770000
Idaho (Boise)[9]	1.79	ND	8620
Switzerland (Geneva)[10]	10.6	10.9 (test, age, sex)	54500
Japan (Kobe)[11]	3.3	2.7 (age, sex)	40999
Denmark blood donors[12]	2.0	1.9 (test)	109665
France (Oise)[13]	25.9	ND	1548000
China (Wuhan)[14]	10.0	ND	1108000
Netherlands blood donors[15]	2.7	ND	461622
Germany (Gangelt)[16]	15.0	20.0 (test, cluster, sym)	2519
Brazil (Rio Grande do Sul)[17]	0.222	0.222 (sampling)*	25283
Scotland blood donors[18]	1.2	ND	64800
California (Santa Clara)[19]	1.5	2.6 (test, sampling, cluster)	51000
Luxembourg[20]	1.9	2.06 (age, sex, district)	12684
Germany (Frankfurt)[21]	0.6	ND	16086
California (Los Angeles)[22]	4.06	4.65 (test, sex, race/ethnicity, income)	367000
New York[23]	12.5	14.0 (test, sex, age race/ethnicity, region)	2723000
California (Bay Area)[24]	0.4 (blood donors)	0.1 (test/confirmation)	7753

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Brazil (133 cities)[25]	1.39	1.62 overall, varying from 0 to 25.0 across 133 cities (test, design)	1209435**
Croatia[26]	1.27***	ND	51765
New York (Brooklyn)[27]	47.0	ND	1203154
Switzerland (Zurich)[28]	unclear	1.3 in patients in April 1-15 and 1.6 in blood donors in May (multivariate Gaussian conditioning)	19773 (Zurich) 30888 (Zurich+Lucerne)
Japan (Tokyo)[29]	3.83	ND	532450
Spain (Barcelona)[30]	14.3	ND	1081938
Italy (Apulia) blood donors[31]	0.99	ND	39887
China (Wuhan B)[32]	8.36 (3.53 for entire period)	ND (2.80 (age, gender, test) for entire period)	926288
California (San Francisco)[33]	4.3 in the census track	6.1 (age, sex, race/ethnicity, test)	316
Brazil (Espirito Santo)[34]	2.1	ND	84391
USA (six states)[35]		(age, sex, test)	
WA Puget Sound	1.3	1.13	48291
UT	2.4	2.18	71550
NYC	5.7	6.93	641778
MO	2.9	2.65	161936

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

South FL	2.2	1.85	117389
CT	4.9	4.94	176012
Spain[36]	ND	5.0**** (sampling, age, sex, income)	2347000
Louisiana (Orleans and Jefferson Parish)[37]	6.9 (IgG or PCR)	6.9 for IgG (census weighting, demographics)	56578
Belgium[38]	5.7	6.0 (sampling, age, sex, province)	695377
France (Crepy-en-Valois)[39]	10.4	ND	620105
China (several regions)[40]			
Hubei (not Wuhan)	3.6	ND	1718110
Chongqing	3.8	ND	11956109
Sichuan	0.6	ND	487847
Guangdong	2.2	ND	2522010
Brazil (Rio de Janeiro) blood donors[41]	6.0	4.7 (age, sex, test)	811452
Brazil (Sao Paulo)[42]	5.2	4.7 (sampling design)	14017
Chile (Vitacura)[43]	11.2	ND	9500

Among studies with multiple consecutive time points where seroprevalence was evaluated, the seroprevalence estimate was the highest in the most recent time interval with the exception of Switzerland (Geneva) where the highest value was seen two weeks before the last time interval, Switzerland (Zurich) where the highest was seen in April 1-15 for patients at the university hospital and in May for blood donors, and the China (Wuhan B) study where the highest value was seen about 3 weeks before the last time interval.

*an estimate is also provided adjusting for test performance, but the assumed specificity of 99.0% seems inappropriate, since as part of the validation process the authors also found that several of the test-positive individuals had household members who were also infected, thus the estimated specificity was deemed to be at least 99.95%

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

**the authors calculate 760000 infected in the 90 cities that had 200-250 samples tested, but many of the other 43 cities with <200 samples may be equally or ever better represented, since they tended to be smaller than the 90 (mean population 356213 versus 659326)
***1.20% in workers in Split without mobility restrictions, 3.37% in workers in Knin without mobility restrictions, 1.57% for all workers without mobility restrictions; Split and Knin tended to have somewhat higher death rates than nation-wide Croatia, but residence of workers is not given, so the entire population of the country is used in the calculations
**** 5.0% by point of care test, 4.6% with immunoassay, 3.7% by having both tests positive, 6.2% by having at least one test positive
test: test performance; ND: no data available

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table 4. Inferred infection fatality rates

Location	COVID-19 deaths (date)	Inferred IFR (corrected), %	COVID-19 deaths <70 years,*** %	IFR in <70 years (corrected), %
Iran (Guilan)[8]	617 (4/23)	0.08 (0.07)	No data	No data
Idaho (Boise)[9]	14 (4/24)	0.16 (0.13)	14 (Idaho)	0.02 (0.02)
Switzerland (Geneva)[10]	243 (4/30)	0.45 (0.36)	8	0.04 (0.03)
Japan (Kobe)[11]	10 (mid-April)	0.02 (0.02)	21 (Japan)	0.01 (0.01)
Denmark blood donors[12]	370 (4/21)	0.34 (0.27)	12	0.05 (0.04)
France (Oise)[13]	932 (4/7)#	0.06 (0.05)	7 (France, <65 years)	0.01 (0.01)
China (Wuhan)[14]	3869 (5/2)	0.35 (0.31)	50	0.19 (0.15)
Netherlands blood donors[15]	3134 (4/15)	0.68 (0.68)	11	0.09 (0.09)
Germany (Gangelt)[16]	7 (4/15)	0.28 (0.25)	0	0.00 (0.00)
Brazil (Rio Grande do Sul)[17]	124 (5/14)	0.49 (0.39)	31 (Brazil, <60 years)	0.19 (0.15)
Scotland blood donors[18]	47 (4/1)	0.07 (0.06)	9 (<65 years)	0.01 (0.01)
California (Santa Clara)[19]	94 (4/22)	0.18 (0.17)	35	0.07 (0.06)
Luxembourg[20]	92 (5/2)	0.73 (0.58)	9	0.07 (0.06)
Germany (Frankfurt)[21]	42*(4/17)	0.26 (0.21)	14 (Germany)	0.04 (0.03)
California (Los Angeles)[22]	724 (4/19)	0.20 (0.18)	24 (<65 years)	0.06 (0.05)
New York[23]	18610 (4/30)^	0.68 (0.54)^	34	0.26 (0.23)^

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

California (Bay Area)[24]	12 (3/22)	0.15 (0.12)	25	0.04 (0.03)
Brazil (133 cities)[25]	**	Median 0.30 (0.27)	31 (<60 years)	0.10 (0.9)
Croatia[26]	79 (5/3)	0.15 (0.14)	13	0.02 (0.02)
New York (Brooklyn)[27]	4894 (5/19)^	0.41 (0.33)^	34 (NY State)	0.15 (0.14)^
Switzerland (Zurich)[28]	107 (4/15 Zurich), 147 (5/22, Zurich+Lucerne)	0.51 (0.41)	8 (Switzerland)	0.05 (0.04)
Japan (Tokyo)[29]	189 (5/11)	0.04 (0.03)	21 (Japan)	0.01 (0.01)
Spain (Barcelona)[30]	5137 (5/2)	0.48 (0.48)	13 (Spain)	0.07 (0.07)
Italy (Apulia) blood donors[31]	530 (5/22)	1.33 (1.20)	15 (Italy)	0.24 (0.22)
China (Wuhan B)[32]	3869 (4/13)	0.42 (0.38)	50	0.23 (0.21)
California (San Francisco)[33]	0 (5/4)	0.00 (0.00)	0	0.00 (0.00)
Brazil (Espirito Santo)[34]	363 (5/21)	0.43 (0.39)	31 (Brazil, <60 years)	0.14 (0.13)
USA (six states)[35]				
WA Puget Sound	207 (4/4)	0.43 (0.43)	10 (WA, <60)	0.05 (0.05)
UT	58 (5/4)	0.08 (0.08)	28 (<65)	0.03 (0.03)
NYC	4146 (4/4)	0.65 (0.65)	34 (NY state)	0.25 (0.25)
MO	329 (4/30)	0.20 (0.20)	23	0.05 (0.05)
South FL	295 (4/15)	0.25 (0.25)	28 (FL)	0.08 (0.08)
CT	2718 (5/6)	1.54 (1.54)	18	0.31 (0.31)

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Spain[36]	26920 (5/11)	1.15 (0.92)	13	0.18 (0.14)
Louisiana[37]	925 (5/16)	1.63 (1.31)	32	0.57 (0.46)
Belgium[38]	7594 (4/30)	1.09 (0.87)	10	0.13 (0.10)
France (Crepy-en-Valois)[39]	2325 (5/5)#	0.37 (0.30)	7 (France, <65 years)	0.04 (0.03)
China (several regions)[40]				
Hubei (not Wuhan)	643 (4/12)	0.04 (0.03)	About 50? (if similar to Wuhan)	0.02 (0.02)
Chongqing	6 (4/12)	0.00 (0.00)		0.00 (0.00)
Sichuan	3 (4/12)	0.00 (0.00)		0.00 (0.00)
Guangdong	8 (4/12)	0.00 (0.00)		0.00 (0.00)
Brazil (Rio de Janeiro) blood donors[41]	1019 (5/3)	0.12 (0.11)	31 (Brazil, <60 years)	0.04 (0.04)
Brazil (Sao Paulo)[42]****	Unknown (5/15)	Unknown, but likely >0.4	31 (Brazil, <60 years)	Unknown, but likely >0.1
Chile (Vitacura)[43]****	Unknown (5/18)	Unknown, but likely <0.2	36 (Chile)	Unknown, but likely <0.1

*approximated from number of deaths in the Hesse province on 4/17 times the proportion of deaths in the 9 districts with key enrollment in the study among all Hesse province deaths.
**data are provided by the authors for deaths per 100,000 population in each city along with inferred IFR in each city, with wide differences across cities; the IFR shown here is the median across the 36 cities with 200-250 samples and at least one positive sample (the interquartile range for the uncorrected IFR is 0.20-0.60 and the full range across all cities is 0-2.4%, but with very wide uncertainty in each city). A higher IFR is alluded in the preprint, but the preprint also shows a scatter diagram for survey-based seroprevalence versus reported deaths per population with a regression slope that agrees with IFR of 0.3.
*** whenever the number/proportion of COVID-19 deaths at age <70 years was not provided in the paper, the proportion of these deaths was retrieved from situational reports of the relevant location; when this was not possible to find for the specific location, a larger geographic entity was used. For Brazil, the closest information that could be found was from a news report (https://www.thejakartapost.com/news/2020/05/22/in-brazil-covid-19-hitting-young-people-harder.html). For Croatia, data on age for 45/103 deaths were retrieved through Wikipedia.
**** Information on deaths not available for the specific locations. In the Sao Paulo study, the 6 districts were selected as being the most hit of Sao Paulo, but it is not stated which are the districts and one cannot

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

retrieve what the number of deaths was as of mid-May, but even using data for death rates across all Sao Paulo would lead to IFR >0.4 overall. In the Vitacura study, similarly one can infer from the wider Santiago Metropolitan area that in the Vitacura area the IFR would probably be <0.2 overall.
^Confirmed deaths; inclusion of probable deaths would increase the IFR estimates by about a quarter.
#for France, governmental situational reports provide per region the number of deaths only for in-hospital deaths, therefore the number of in-hospital deaths has been multiplied by a factor equal to the total deaths/in-hospital deaths across France.
IFR: infection fatality rate. The inferred IFR is derived by dividing the number of accumulated deaths (at the time chosen by the authors of each study, or until after 1 week of the mid-point of the study dates, whenever the authors had not arbitrated on a date for the death count) by the estimated number of infected people. The corrected IFR is obtained from the inferred IFR assuming that, as compared with assessing IgG, IgM, and IgA, 20% of the infections are missed when only IgG is assessed, and 10% of the infections are missed when two of the three antibodies are assessed.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table 5. Additional seroprevalence data from nation-wide studies that have been announced to the press and/or in preliminary reports, but are not presented yet as full articles. Only countries not represented in Tables 1-4 are considered.

Country	Sample size (antibody)	Date	Seroprevalence (%)	Population	Deaths (date)	IFR (corrected)
United Kingdom[44]	885 (IgG)	April 26-May 24^	6.78	67897000	34636 (5/17)	0.75 (0.60)
Finland[45]	674 (IgG)	April 20-26^	2.52	5541000	211 (4/30)	0.15 (0.12)
Sweden[46]	1200 (IgG)	May 18-24	6.3	10101000	4501 (5/28)	0.71 (0.57)
Czechia[47]	26549 (IgG)	April 23-May 1	0.40	10710000	252 (5/4)	0.59 (0.47)
Israel[48]	1709 (IgG?)	May 2020	2-3	9198000	299 (6/10)	0.13 (0.10)*
India[49]	26400 (IgG?)	May 2020	0.73**	1380382000	8107 (6/10)	0.08 (0.06)
Slovenia[50]	1368 (IgG?)	April 2020	3.1	2079000	92 (5/1)	0.14 (0.11)

^ slightly lower seroprevalence recorded in subsequent weeks
*Assuming seroprevalence 2.5%. Another ongoing study in Beni Brak (population n=198900, 0 deaths as of mid-June) has shown seroprevalence of 1.4% among the first 2933 samples. A much larger seroprevalence study based on HMOs has also been launched.
**based on data from part of the sample (65 of the 83 districts)

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

References:

1. Melnick ER, Ioannidis JPA. Head to Head: Should governments continue lockdown to slow the spread of covid-19? BMJ 2020;369:m1924. doi: 10.1136/bmj.m1924.
2. https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---3-march-2020. Accessed May 10, 2020.
3. World Health Organization. Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19) 16-24 February 2020. Published 2020-02-28. Available from: https://www.who.int/publications-detail/report-of-the-who-china-joint-mission-on-coronavirus-disease-2019-(covid-19). Accessed May 10, 2020.
4. McGinty JC. How many people might one person with coronavirus infect? Wall Street Journal. February 14, 2020. https://www.wsj.com/articles/how-many-people-might-one-person-with-coronavirus-infect-11581676200. Accessed February 27, 2020.
5. https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf
6. Meyerowitz-Katz G, Merone L 2020 A systematic review and meta-analysis of published research data on COVID-19 infection-fatality rates doi: https://doi.org/10.1101/2020.05.03.20089854
7. Garcia-Basteiro AL, Moncunill G, Tortajada M, Vidal M, Guinovart C, Jimenez A, et al. Seroprevalence of antibodies against SARS-CoV-2 among health care workers in a large Spanish reference hospital. Nature Communications 2020;11:3500
8. Shakiba M, Nazari S, Mehrabian F, et al. Seroprevalence of COVID-19 virus infection in Guilan province, Iran. medRxiv doi: https://doi.org/10.1101/2020.04.26.20079244

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

9. Bryan A, Pepper G, Wener MH, Fink SL, Morishima C, Chaudhary A, et al. Performance characteristics of the Abbott Architect SARS-CoV-2 IgG assay and seroprevalence testing in Idaho. J Clin Microbiol. 2020 doi:10.1128/JCM.00941-20
10. Stringhini S, Wisniak A, Piumatti G, Azman AS, Lauer SA, Baysson H, et al. Seroprevalence of anti-SARS-CoV-2 IgG antibodies in Geneva, Switzerland (SEROCoV-POP): a population-based study. Lancet 2020; https://doi.org/10.1016/S0140-6736(20)31304-0
11. Doi A, Iwata Kuroda H, Hasuike T, Nasu S, Kanda A, et al. Seroprevalence of novel coronavirus disease (COVID-19) in Kobe, Japan. medRxiv https://doi.org/10.1101/2020.04.26.20079822
12. Erikstrup C, Hother CE, Pedersen OB, Mølbak K, Skov RL, Holm DK, et al. Estimation of SARS-CoV-2 infection fatality rate by real-time antibody screening of blood donors. Clin Infect Dis 2020; https://doi.org/10.1093/cid/ciaa849.
13. Fontanet A, Tondeur L, Madec Y, Grant R, Besombes C, Jolly N, et al. Cluster of COVID-19 in northern France: A retrospective closed cohort study. medRxiv 2020 https://doi.org/10.1101/2020.04.18.20071134
14. Wu X, Fu B, Chen L, Feng Y. Serological tests facilitate identification of asymptomatic SARS-CoV-2 infection in Wuhan, China. J Med Virol 2020 https://onlinelibrary.wiley.com/doi/10.1002/jmv.25904
15. Slot E, Hogema BM, Reusken CBEM, Reimerink JH, Molier M, Karregat HM, et al. Herd immunity is not a realistic exit strategy during a COVID-19 outbreak. Research Square 2020. https://dx.doi.org/10.21203/rs.3.rs-25862/v1

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

16. Streeck H, Schulte B, Kümmerer BM, Richter E, Höller T, Fuhrmann C, et al. Infection fatality rate of SARS-CoV-2 infection in a German community with a super-spreading event. medRxiv 2020 https://doi.org/10.1101/2020.05.04.20090076
17. Silveira MF, Barros AJD, Horta BL, Pellanda LC, Dellagostin O, Struchiner CJ, et al. Repeated population-based surveys of antibodies against SARS2-CoV-2 in Southern Brazil. Nature Medicine 2020 (in press) https://doi.org/10.1038/s41591-020-0992-3
18. Thompson C, Grayson N, Paton RS, Lourenco J, Penman BS, Lee L. Neutralising antibodies to SARS coronavirus 2 in Scottish blood donors – a pilot study of the value of serology to determine population exposure. medRxiv https://doi.org/10.1101/2020.04.13.20060467
19. Bendavid E, Mulaney B, Sood N, et al. COVID-19 Antibody Seroprevalence in Santa Clara County, California. medRxiv doi: https://doi.org/10.1101/2020.04.14.20062463
20. Snoeck CJ, Vaillant M, Abdelrahman T, Satagopam VP, Turner JD, Beaumont K, et al. Prevalence of SARS-CoV-2 infection in the Luxembourgish population: the CON-VINCE study. medRxiv 2020.05.11.20092916; doi: https://doi.org/10.1101/2020.05.11.20092916
21. Kraehling V, Kern M, Halwe S, Mueller H, Rohde C, Savini M, et al. Epidemiological study to detect active SARS-CoV-2 infections and seropositive persons in a selected cohort of employees in the Frankfurt am Main metropolitan area. medRxiv 2020.05.20.20107730; doi: https://doi.org/10.1101/2020.05.20.20107730
22. Sood N, Simon P, Ebner P, Eichner D, Reynolds J, Bendavid E, Bhattacharya J. Seroprevalence of SARS-CoV-2-Specific Antibodies Among Adults in Los Angeles County, California, on April 10-11, 2020. JAMA. 2020 May 18:e208279.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

23. Rosenberg ES, Tesoriero JM, Rosenthal EM, Chung R, Barranco MA, Styer LM, et al. Cumulative incidence and diagnosis of SARS-CoV-2 infection in New York. Ann Epidemiol 2020 (in press). DOI: https://doi.org/10.1016/j.annepidem.2020.06.004

24. Ng D, Goldgof G, Shy B, Levine A, Balcerek J, Bapat SP, et al. SARS-CoV-2 seroprevalence and neutralizing activity in donor and patient blood from the San Francisco Bay Area. medRxiv 2020.05.19.20107482; doi: https://doi.org/10.1101/2020.05.19.20107482

25. Hallal PC, Hartwig FP, Horta BL, Victora GD, Silveira MF, Struchiner CJ, et al. Remarkable variability in SARS-CoV-2 antibodies across Brazilian 1 regions: nationwide serological household survey in 27 states. medRxiv 2020 doi: https://doi.org/10.1101/2020.05.30.20117531

26. Jerkovic I, Ljubic T, Basic Z, Kruzic I, Kunac N, Bezic J, et al. SARS-CoV-2 antibody seroprevalence in industry workers in Split-Dalmatia and Sibenik-Knin County, Croatia. medRxiv 2020.05.11.20095158; doi: https://doi.org/10.1101/2020.05.11.20095158

27. Reifer J, Hayum N, Heszkel B, Klagsbald I, Streva VA. SARS-CoV-2 IgG antibody responses in New York City. medRxiv 2020.05.23.20111427; doi: https://doi.org/10.1101/2020.05.23.20111427

28. Emmenegger M, De Cecco E, Lamparter D, Jacquat RPB, Ebner D, et al. Early plateau of SARS-CoV-2 seroprevalence identified by tripartite immunoassay in a large population. medRxiv 2020.05.31.20118554; doi: https://doi.org/10.1101/2020.05.31.20118554

29. Takita M, Matsumura T, Yamamoto K, Yamashita E, Hosoda K, Hamaki T, Kusumi E. Regional difference in seroprevalence of SARS-CoV-2 in Tokyo: Results from the community point-of-care antibody testing. medRxiv 2020.06.03.20121020; doi: https://doi.org/10.1101/2020.06.03.20121020

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

30. Crovetto F, Crispi F, Llurba E, Figueras F, Gomez-Roig MD, Gratacos E. Seroprevalence and clinical spectrum of SARS-CoV-2 infection in the first versus third trimester of pregnancy. medRxiv 2020; doi: https://doi.org/10.1101/2020.06.17.

31. Fiore JR, Centra M, De Carlo A, Granato T, Rosa A, Sarno M, et al. Far away from herd immunity to SARS-CoV-2: results from a survey in healthy blood donors in South Eastern Italy. medRxiv 2020; doi: https://doi.org/10.1101/2020.06.17.20133678

32. Ling R, Yu Y, He J, Zhang J, Xu S, Sun R, et al. Seroprevalence and epidemiological characteristics of immunoglobulin M and G antibodies against SARS-CoV-2 in asymptomatic people in Wuhan, China. medRxiv 2020; doi: https://doi.org/10.1101/2020.06.16.20132423.

33. Chamie G, Marquez C, Crawford E, Peng J, Petersen M, Schwab D, et al. SARS-CoV-2 community transmission during Shelter-in-place in San Francisco. medRxiv 2020; doi: https://doi.org/10.1101/2020.06.15.20132233

34. Gomes CC, Cerutti C, Zandonade E, et al. A population-based study of the prevalence of COVID-19 infection in Espírito Santo, Brazil: methodology and results of the first stage. medRxiv 2020; doi: https://doi.org/10.1101/2020.06.13.20130559

35. Havers FP, Reed C, Lim T, Montgomery JM, Klena JD, Hall AJ, et al. Seroprevalence of antibodies to SARS-CoV-2 in six sites in the United States, March 23-May 3, 2020. medRxiv 2020; doi: https://doi.org/10.1101/2020.06.25.20140384

36. Pollán M, Pérez-Gómez B, Pastor-Barriuso R, Oteo J, Hernán MA, Pérez-Olmeda M, et al. Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study. Lancet 2020; https://doi.org/10.1016/S0140-6736(20)31483-5

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

37. Feehan AK, Fort D, Garcia-Diaz J, et al. Point prevalence of SARS-CoV-2 and infection fatality rate in Orleans and Jefferson Parish, Louisiana, May 9-15, 2020. medRxiv 2020; doi: https://doi.org/10.1101/2020.06.23.20138321

38. Herzog S, De Bie J, Abrams S, et al. Seroprevalence of IgG antibodies against SARS coronavirus 2 in Belgium – a prospective cross sectional study of residual samples. medRxiv 2020; doi: https://doi.org/10.1101/2020.06.08.20125179

39. Fontanet A, Grant R, Tondeur L, et al. SARS-CoV-2 infection in primary schools in northern France: A retrospective cohort study in an area of high transmission. medRxiv 2020; doi: https://doi.org/10.1101/2020.06.25.20140178

40. Xu X, Sun J, Nie S, et al. Seroprevalence of immunoglobulin M and G antibodies against SARS-CoV-2 in China. Nature Medicine 2020; https://doi.org/10.1038/s41591-020-0949-6

41. Filhol LA, Scwarcwald CL, Garcia Mateos S, et al. Seroprevalence of anti-SARS-CoV-2 among blood donors in Rio de Janeiro, Brazil. Rev Saude Publica 2020;54:69

42. Tess BH, Granato CFH, Alves MC, et al. SARS-CoV-2 seroprevalence in the municipality of São Paulo, Brazil, ten weeks after the first reported case. medRxiv 2020 doi: https://doi.org/10.1101/2020.06.29.20142331

43. Torres JP, Pinera C, De La MAza V, et al. SARS-CoV-2 antibody prevalence in blood in a large school community subject to a Covid-19 outbreak: a cross-sectional study. Clin Infect Dis 2020 (in press)

44. https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/coronaviruscovid19infectionsurveypilot/28may2020, last accessed July 12, 2010

45. https://www.thl.fi/roko/cov-vaestoserologia/sero_report_weekly.html, last accessed July 20, 2020

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

46. https://www.folkhalsomyndigheten.se/nyheter-och-press/nyhetsarkiv/2020/juni/forsta-resultaten-om-antikroppar-efter-genomgangen-covid-19-hos-blodgivare/, last accessed July 20, 2020

47. https://covid-imunita.uzis.cz/res/file/prezentace/20200506-dusek.pdf, last accessed July 20, 2020

48. https://www.jpost.com/israel-news/coronavirus-herd-immunity-not-in-israel-according-to-a-serological-study-630059, last accessed July 20, 2020

49. https://indianexpress.com/article/india/india-not-in-community-transmission-stage-yet-health-ministry-6454183/, last accessed July 20, 2020

50. https://www.gov.si/en/news/2020-05-06-first-study-carried-out-on-herd-immunity-of-the-population-in-the-whole-territory-of-slovenia/, last accessed July 20, 2020

51. Ioannidis JPA, Axfors C, Contopoulos-Ioannidis DG. Population-level COVID-19 mortality risk for non-elderly individuals overall and for non-elderly individuals without underlying diseases in pandemic epicenters. Environmental Research 2020 (in press)

52. https://www.theguardian.com/world/2020/apr/13/half-of-coronavirus-deaths-happen-in-care-homes-data-from-eu-suggests. Accessed April 27, 2020.

53. American Geriatrics Society. American Geriatrics Society Policy Brief: COVID-19 and Nursing Homes. J Am Geriatr Soc. 2020 Apr 8. doi: 10.1111/jgs.16477.

54. Kelly A, Conell-Price J, Covinsky K, Stijacic Cenzer I, Chang A, Boscardin WJ, et al. Lengths of stay for older adults residing in nursing homes at the end of life. J Am Geriatr Soc. 2010;58(9):1701-6.

55. EUROMOMO. Graphs and maps: Excess mortality. Updated week 28, 2020. Accessed July 12, 2020. https://www.euromomo.eu/graphs-and-maps/

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

56. Wang D, Hu B, Hu C, et al. Clinical characteristics of 138 hospitalized patients with 2019 novel coronavirus-infected pneumonia in Wuhan, China. JAMA. 2020 Feb 7:e201585.

57. Boccia S, Ricciardi W, Ioannidis JPA. What Other Countries Can Learn From Italy During the COVID-19 Pandemic. JAMA Intern Med. 2020 Apr 7. doi: 10.1001/jamainternmed.2020.1447.

58. Sandri MT, Azzolini E, Torri V, Carloni S, Tedeschi M, Castoldi M, Mantovani A, Rescigno M. IgG serology in health care and administrative staff populations from 7 hospital representative of different exposures to SARS-CoV-2 in Lombardy, Italy. medRxiv doi: https://doi.org/10.1101/2020.05.24.20111245

59. Brofsky A. Distinct clades of SARS-CoV-2: implications for modeling of viral spread. J Med Virol. 2020 Apr 20. doi: 10.1002/jmv.25902.

60. Rosado J, Cockram C, Merkling SH, Demeret C, Meola A, Kerneis S, Terrier B, Fafi-Kremer S, et al. Serological signatures of SARS-CoV-2 infection: Implications for antibody-based diagnostics. medRxiv 2020. doi: https://doi.org/10.1101/2020.05.07.20093963

61. Long Q, Tang X, Shi Q, et al. Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. Nat Med 2020; https://doi.org/10.1038/s41591-020-0965-6

62. Wu F, Wang A, Liu M, et al. Neutralizing antibody responses to SARS-CoV-2 in a COVID-19 recovered patient cohort and their implications. medRxiv 2020;2020.03.30.20047365.

63. Seow J, Graham C, Merrick B, et al. Longitudinal evaluation and decline of antibody responses in SARS-CoV-2 infection. medRxiv 2020; doi: https://doi.org/10.1101/2020.07.09.20148429.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

64. Edouard S, Colson P, Melenotte C, De Pinto F, Thomas L, La Scola B, et al. Evaluating the serological status of COVID-19 patients using an indirect immunofluorescent assay, France. medRxiv 2020; doi: https://doi.org/10.1101/2020.05.05.20092064

65. Solbach W, Schiffner J, Backhaus I, Burger D, Staiger R, Tiemer B, et al. Antibody profiling of COVID-19 patients in an urban low-incidence region in Northern Germany. medRxiv 2020.05.30.20111393; doi: https://doi.org/10.1101/2020.05.30.20111393

66. Kikkert M. Innate immune evasion by human respiratory RNA viruses. J Innate Immun 2020;12:4–20

67. Krammer F. The human antibody response to influenza A virus infection and vaccination. Nature Reviews Immunology 2019;19:1.

68. Cervia C, Nilsson J, Zurbuchen Y, Valaperti A, Schreiner J, Wolfensberger A, et al. Systemic and mucosal antibody secretion specific to SARS-CoV-2 during mild versus severe COVID-19. bioRxiv 2020 doi: https://doi.org/10.1101/2020.05.21.108308

69. Gallais F, Velay A, Wendling M-J, Nazon C, Partisani M, Sibilia J, et al. Intrafamilial exposure to SARS-CoV-2 induces cellular immune response without seroconversion. medRxiv 2020; doi: https://doi.org/10.1101/2020.06.21.20132449

70. Sekine T, Rivera-Ballesteros O, Stralin K, et al. Robust T cell immunity in convalescent individuals with asymptomatic or mild COVID-19. bioRxiv 2020; doi: https://doi.org/10.1101/2020.06.29.174888

71. Moriarty LF, Plucinski MM, Marston BJ, Kurbatova EV, Knust B, Murray EL, et al. Public health responses to COVID-19 outbreaks on cruise Ships - Worldwide, February-March 2020. MMWR Morb Mortal Wkly Rep. 2020 Mar 27;69(12):347-352.

72. Maxmen A. Coronavirus is spreading under the radar in US homeless shelters. Nature. 2020 May;581(7807):129-130.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

73. Britton T, Ball F, Trapman P. A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2. Science 23 Jun 2020:eabc6810. DOI: 10.1126/science.abc6810

74. Gomes GM, Corder RM, King JG, Langwig KE, Souto-Maior C, Carneiro J, et al. Individual variation in susceptibility or exposure to SARS-CoV-2 lowers the herd immunity threshold. medRxiv https://doi.org/10.1101/2020.04.27.20081893

75. Grifoni A, Weiskopf D, Ramirez SI, et al. Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals. Cell 2020;181:1-13.

76. Paget J, Spreeuwenberg P, Charu V, Taylor RJ, Iuliano AD, Bresee J, et al. Global mortality associated with seasonal influenza epidemics: New burden estimates and predictors from the GLaMOR Project. J Glob Health. 2019 Dec; 9(2): 020421.

77. Pandemic Influenza Preparedness and Response: A WHO Guidance Document. Geneva, World Health Organization, 2009.

78. Webb Hooper M, Nápoles AM, Pérez-Stable EJ. COVID-19 and racial/ethnic disparities. JAMA. 2020 May 11. doi: 10.1001/jama.2020.8598.

79. Centers for Disease Control and Prevention. COVID-19 in Racial and Ethnic Minority Groups. Updated April 24, 2020. Accessed April 30, 2020. https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/racial-ethnic-minorities.html

80. https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html, Accessed June 6, 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

81. Kissler SM, Tedijanto C, Goldstein E, Grad YH, Lipsitch M. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. Science. 2020 Apr 14:eabb5793. doi: 10.1126/science.abb5793.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

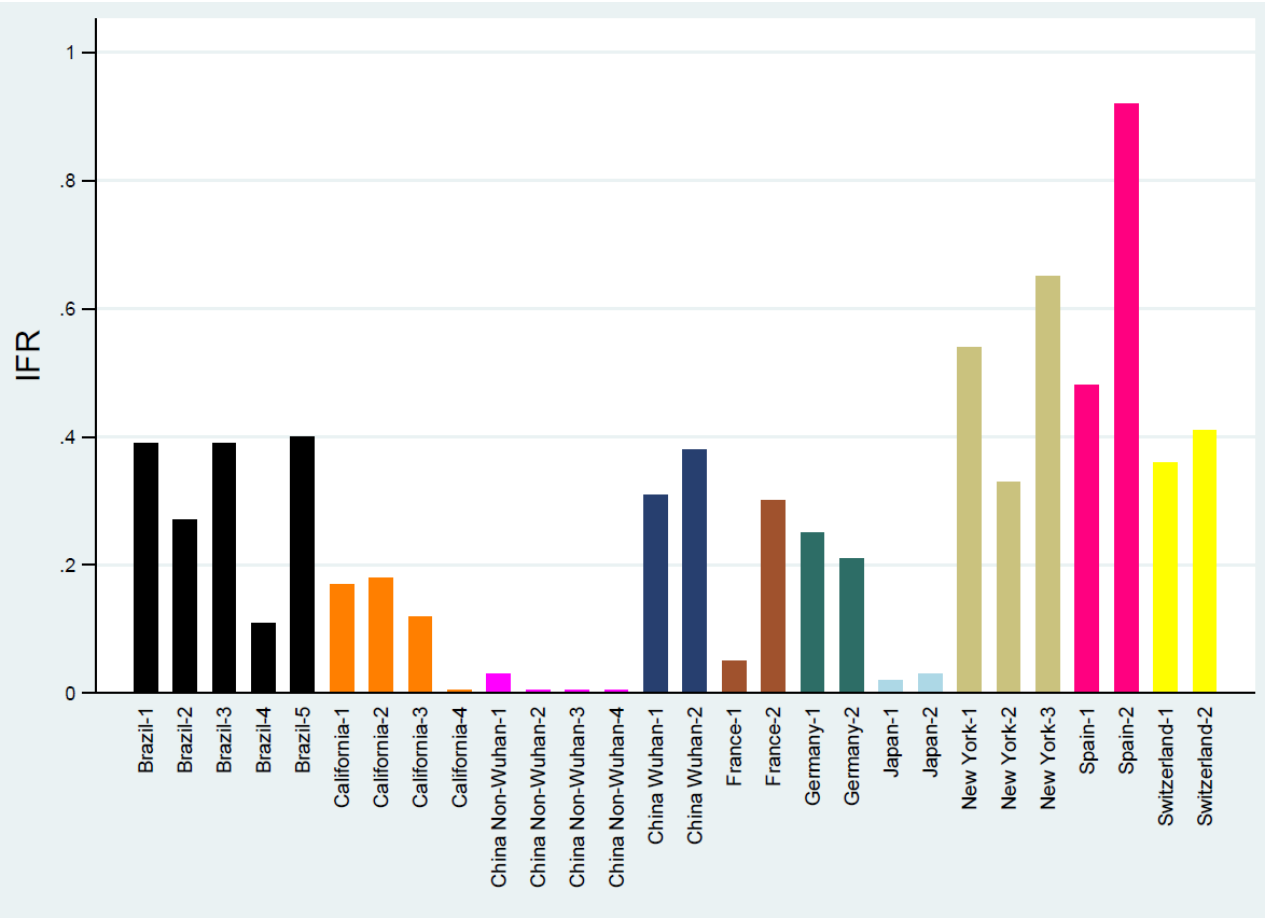
Figure legends

Figure 1. Locations that had two or more IFR estimates from different studies. Locations are defined at the level of countries, except for the USA where they are defined at the level of states and China is separated into Wuhan and non-Wuhan areas. Corrected IFR estimates are shown. Within the same location, IFR estimates tend to have only modest differences, even though it is possible that different areas within the same location may also have genuinely different IFR. France is one exception where differences are large, but both estimates come from population studies of outbreaks from schools and thus may not provide good estimates of population seroprevalence and may lead to underestimated IFR.

Figure 2. Scatterplot of corrected IFR estimates (%) in each location plotted against the COVID-19 mortality rate as of July 12, 2020 in that location (in deaths per million population). Locations are defined at the level of countries, except for the USA where they are defined at the level of states and China is separated into Wuhan and non-Wuhan areas. When several IFR estimates are available from multiple studies for a location, the sample size-weighted mean has been used. Not shown are two locations with >1000 deaths per million population, both of which have high IFR (New York and Connecticut).

medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure 1



medRxiv preprint doi: https://doi.org/10.1101/2020.05.13.20101253; this version posted July 14, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

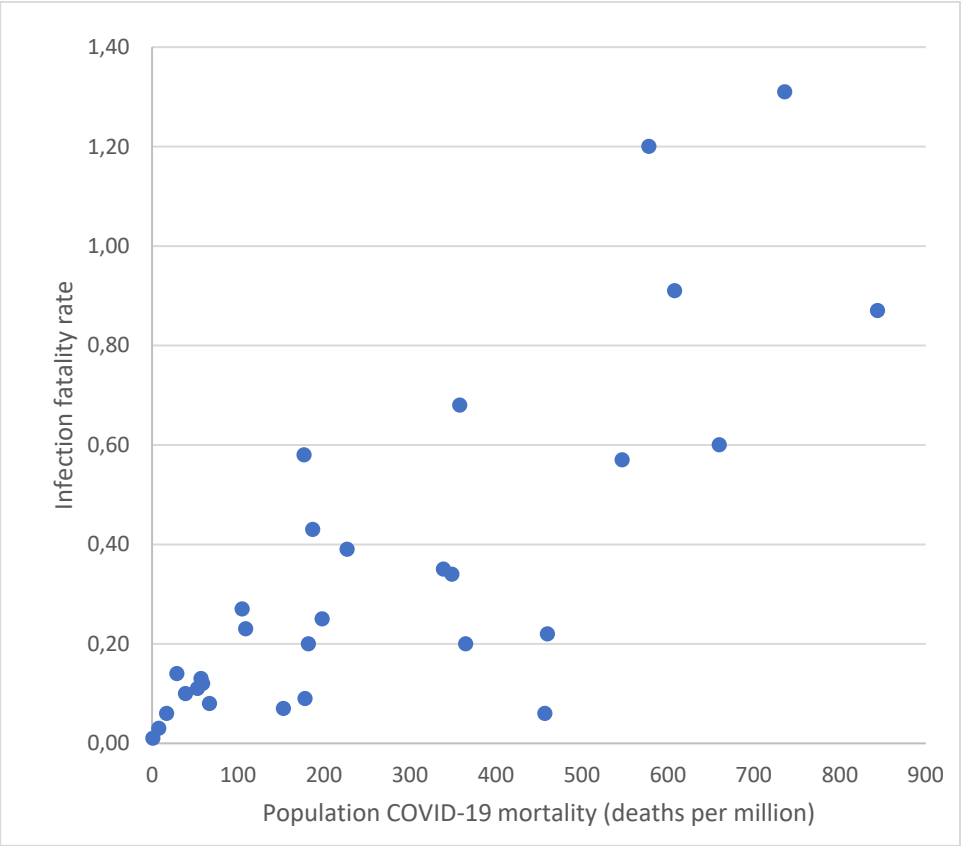



Figure 2