medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.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 4.0 International license .

# A systematic review and meta-analysis of published research data on COVID-19 infection-fatality rates

Gideon Meyerowitz-Katz*[1,2], Lea Merone[3,4]

1. Western Sydney Local Health District

2. University of Wollongong

3. James Cook University

4. Apunipima Cape York Health Council

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.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/10.1101/2020.05.03.20089854; this version posted July 7, 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 The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.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 4.0 International license .

**Key messages:**

- COVID-19 infection-fatality rate (IFR) is an important statistic for policy about the disease
- Published estimates vary, with a 'true' fatality rate hard to calculate
- Systematically reviewing the literature and meta-analyzing the results shows an IFR of 0.68% (0.53-0.82%)

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.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 4.0 International license .


## Abstract

An important unknown during the COVID-19 pandemic has been the infection-fatality rate (IFR). This differs from the case-fatality rate (CFR) as an estimate of the number of deaths as a proportion of the total number of cases, including those who are mild and asymptomatic. While the CFR is extremely valuable for experts, IFR is increasingly being called for by policy-makers and the lay public as an estimate of the overall mortality from COVID-19.

## Methods

Pubmed, Medline, SSRN, and Medrxiv were searched using a set of terms and Boolean operators on 25/04/2020 and re-searched 14/05/2020, 21/05/2020, and 16/06/2020. Articles were screened for inclusion by both authors. Meta-analysis was performed in Stata 15.1 using the metan command, based on IFR and confidence intervals extracted from each study. Google/Google Scholar was used to assess the grey literature relating to government reports.

## Results

After exclusions, there were 24 estimates of IFR included in the final meta-analysis, from a wide range of countries, published between February and June 2020.

The meta-analysis demonstrated a point-estimate of IFR of 0.68% (0.53-0.82%) with high heterogeneity ($p<0.001$).

## Conclusion

Based on a systematic review and meta-analysis of published evidence on COVID-19 until May, 2020, the IFR of the disease across populations is 0.68% (0.53-0.82%). However, due to very high heterogeneity in the meta-analysis, it is difficult to know if this represents the 'true' point estimate. It is likely that, due to age and perhaps underlying comorbidities in the population, different places will experience different IFRs due to the disease. Given issues with mortality recording, it is also likely that

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

this represents an underestimate of the true IFR figure. More research looking at age-stratified IFR is urgently needed to inform policy-making on this front.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

## Introduction

2020 saw the emergence of a global pandemic, COVID-19, caused by the SARS-CoV-2 virus, which began in China and has since spread across the world. One of the most challenging questions to answer during the COVID-19 pandemic has been regarding the true infection-fatality rate (IFR) of the disease. While case-fatality rates (CFR) are eminently calculable from various published data sources (1) – CFR being the number of deaths divided by the number of confirmed cases - it is far more difficult to extrapolate to the proportion of all infected individuals who have died due to the infection because those who have very mild, atypical or asymptomatic disease are frequently left undetected and therefore omitted from fatality-rate calculations (2). Given the issues with obtaining accurate estimates, it is not unexpected that there are wide disparities in the published estimates of infection. This is an issue for several reasons, most importantly in that policy is dependent on modelling, and modelling is dependent on assumptions. If we do not have a robust estimate of IFR, it is challenging to make predictions about the true impact of COVID-19 in any given susceptible population, which may stymie policy development and may have serious consequences for decision-making into the future. While CFR is a more commonly-used statistic, and is very widely understood among experts, IFR provides important context for policy makers that is hard to convey, particularly given the wide variation in CFR estimates. While CFR is naturally a function of the denominator – i.e. how many people have been tested for the disease – policy-makers are often most interested in the total burden in the population rather than the biased estimates given from testing only the acutely unwell patients.

This is particularly important when considering the reopening of countries post 'lockdown'. Depending on the severity of the disease, it may be reasonable to reopen services such as schools, bars, and clubs, at different timings. Another salient point is the expected burden of disease in younger age groups – while there are likely long-term impacts other than death, it will be important

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

for future planning to know how many people in various age groups are likely to die if the infection becomes widespread across societies. Age-stratified estimates are also important as it may give countries some way to predict the number of deaths expected given their demographic breakdown.

There are a number of methods for investigating the IFR in a population. Retrospective modelling studies of influenza have successfully predicted the true number of cases and deaths from influenza-like-illness records and excess mortality estimates (3, 4). However, these may not be accurate, in part due to the general difficulty in attributing influenza cases to subsequent mortality, meaning that CFRs may both overestimate and equally underestimate the true number of deaths due to the disease in a population (5). The standard test for COVID-19 involves polymerase chain reaction testing (PCR) of nasopharyngeal swabs from patients suspected of having contracted the virus. This can produce some false negatives (6), with one study demonstrating almost a quarter of patients experiencing a positive result following up to two previous false negatives (7). PCR is also limited in that it cannot test for previous infection. Serology testing is more invasive, requiring a blood sample, however it can determine if there has been previous infection and can be performed rapidly at the point of care (PoC). Serology PoC testing cannot determine if a person is infectious, or if infection is recent and there is risk of misinterpretation of results (8). Serology testing is more sensitive and specific than PCR, but will still likely overestimate prevalence when few people have been infected with COVID-19 and underestimate in populations with more infections (9).

Given the emergence of COVID-19 as a global pandemic, it is somewhat unlikely that these issues are directly mirrored for the newer disease, but there are likely similarities between the two. Some analysis in mainstream media publications and pre-prints has implied that there is a large burden of deaths that remains unattributed to COVID-19. Similarly, serological surveys have demonstrated that there is a large proportion of cases that have not been captured in the case numbers reported in the U.S., Europe, and potentially worldwide (10-12).

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

This paper presents a systematic effort to collate and aggregate these disparate estimates of IFR using an easily replicable method. While any meta-analysis is only as reliable as the quality of included studies, this will at least put a realistic estimate to the IFR given current published evidence.

**Methods**

This study used a simple systematic review protocol. PubMed, MedLine, and Medrxiv were searched on the 25/04/2020 using the terms and Boolean operators: (infection fatality rate OR ifr OR seroprevalence) AND (COVID-19 OR SARS-CoV-2). This search was repeated on 14/05/2020, 25/05/2020, and 16/06/2020. The preprint server SSRN was also searched on 25/05/2020, however as it does not allow this format the Boolean operators and brackets were removed. While Medrxiv and SSRN would usually be excluded from systematic review, given that the papers included are not peer-reviewed, during the pandemic it has been an important source of information and contains many of the most recent estimates for epidemiological information about COVID-19. Inclusion criteria for the studies were:

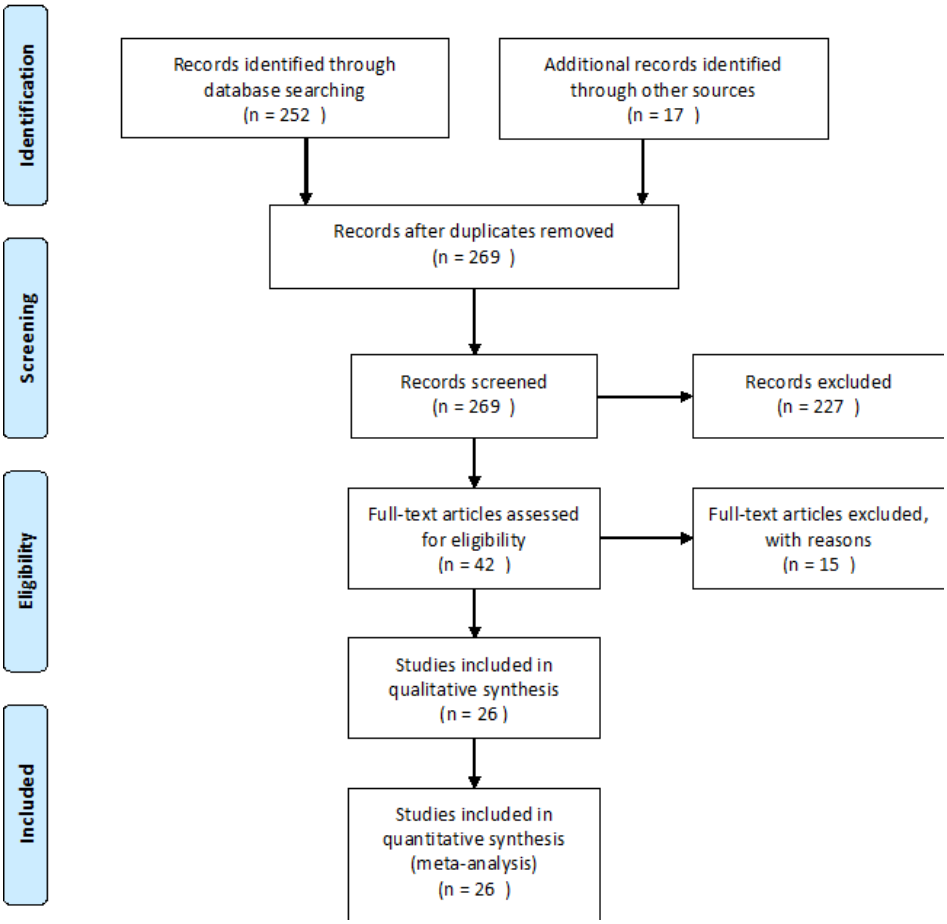
- Regarding COVID-19/SARS-CoV-2 (i.e. not SARS-CoV-1 extrapolations)
- Presented an estimated population infection-fatality rate (or allowed calculation of such from publicly-available data)

Titles and abstracts were screened for eligibility and discarded if they did not meet the inclusion criteria. GMK then conducted a simple Google and Google scholar search using the same terms to assess the grey literature, in particular published estimates from government agencies that may not appear on formal academic databases. LM assessed the articles to ensure congruence. If these met the inclusion criteria, they were included in the systematic review and meta-analysis. Similarly,

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

Twitter searches were performed using similar search terms to assess the evidence available on social media. Estimates for IFR and the confidence interval were extracted for each study.

All analysis and data transformation was performed in Stata 15.1. The meta-analysis was performed using the metan command for continuous estimates, with IFR and the lower/upper bounds of the confidence interval as the variables entered. This model used the DerSimonian and Laird random-effects method. The metan command in Stata automatically generates an $I^2$ statistic that was used to investigate heterogeneity. Histograms were visually inspected to ensure that there was no significant positive or negative skew to the results that would invalidate this methodology. For the studies where no confidence interval was provided, one was calculated.




A PRISMA flow diagram of the search methods

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

Sensitivity analyses were performed stratifying the results into the type of study – serological vs non – by country, and by month of calculation.

The metabias and metafunnel commands were used to examine publication bias in the included research, with Egger's test used for the metabias estimation. It was challenging to formally rate the risk of bias of the included modelling studies, as there was very significant heterogeneity in methodology and implementation, with the result that the risk of bias in these studies was considered to be high across all included research. Serological surveys were rated using a the risk of bias in prevalence tool with a resulting estimate in line with Cochrane GRADE criteria of low, moderate, or high (13). This tool asks a series of 10 questions about the sampling and data collection of prevalence studies, with a final rating based on the previous questions. Each question is answered yes/no, with a lack of information presumed to be no/unclear. A separate sensitivity analysis was conducted using only serological survey results stratified by risk of bias.

Due to a recent surge in the number of serological surveys being published, these were included in the infection-fatality estimate despite not formally calculating an IFR in the study text itself. Regional death rates were taken from the John Hopkins University CSSE dashboard (14) 10 days after the serosurvey completion where no IFR was calculated to account for right-censoring of these estimates (15), and used to estimate the IFR given the population.

All code and data files are available (in .do and .csv format) upon request.

**Results**

Initial searches identified 252 studies across all databases. Later searches on Google and social media, as well as resampling the included databases revealed a further seventeen estimates to include in the study. These came from a variety of sources, with some appearing from blog posts, others posted on

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

Twitter, and some government documents being found through Google. There were no duplicates specifically, however two pre-prints had been published and so appeared in slightly different forms in both databases. In this case, the published study was used rather than the pre-print. Results are collated in table 1.

Studies were excluded for a variety of reasons. Some studies only looked at COVID-19 incidence, rather than prevalence of antibodies, and were thus considered potentially unreliable as population estimates (16). The most common reason for exclusion was selection bias – many studies only looked at targeted populations in their seroprevalence data, and thus could not be used as population estimators of IFR (11, 17-24). For some data, it was difficult to determine the numerator (i.e. number of deaths) associated with the seroprevalence estimate, or the denominator (i.e. population) was not well defined and thus we did not calculate an IFR (25, 26). One study explicitly warned against using its data to obtain an IFR (27). Another study calculated an IFR, but did not allow for an estimate of confidence bounds and thus could not be included in the quantitative synthesis (28).

After screening titles and abstracts, 227 studies were removed. Many of these looked at case-fatality estimates, or discussed IFR as a concept and/or a model input, rather than estimating the figure themselves. 40 papers were assessed for eligibility for inclusion into the study, resulting in a final 25 to be included in the qualitative synthesis.

Studies varied widely in design, with 3 entirely modelled estimates (29-31), 4 observational studies (10, 32-34), 5 pre-prints that were challenging to otherwise classify (2, 35-38), and a number of serological surveys of varying types reported by government agencies (39-51). For the purposes of this research, an estimate for New York City was calculated from official statistics and the serosurvey, however this was correlated with a published estimate (28) to ensure validity.

The main result from the random-effects meta-analysis is presented in Figure 1. Overall, the aggregated estimate across all 24 studies indicated an IFR of 0.68% (95% CI 0.53-0.82%), or 68 deaths

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

per 10,000 infections. Heterogeneity was extremely high, with the overall $I^2$ exceeding 99% ($p<0.0001$).

The sensitivity analysis by month from Figure 3 showed that earlier estimates of IFR were lower, with later estimates showing a higher figure, although this appears to have stabilised in May.

Analysing by region of origin did not appear to have a substantial effect on the findings, although there was a slightly lower estimate seen in Asia. As the Middle East was only represented by one study, this region was excluded from the meta-synthesis by region. Two studies were also excluded as they did not present an IFR for a specific region (i.e. Diamond Princess).

Of note, there was some difference in estimates of IFR between estimates based on serosurveys and those of modelled or PCR-based estimates. The overall estimates from serosurvey studies was 0.60% (0.42-0.77%), although again with very high heterogeneity, as can be seen in Figure 4.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

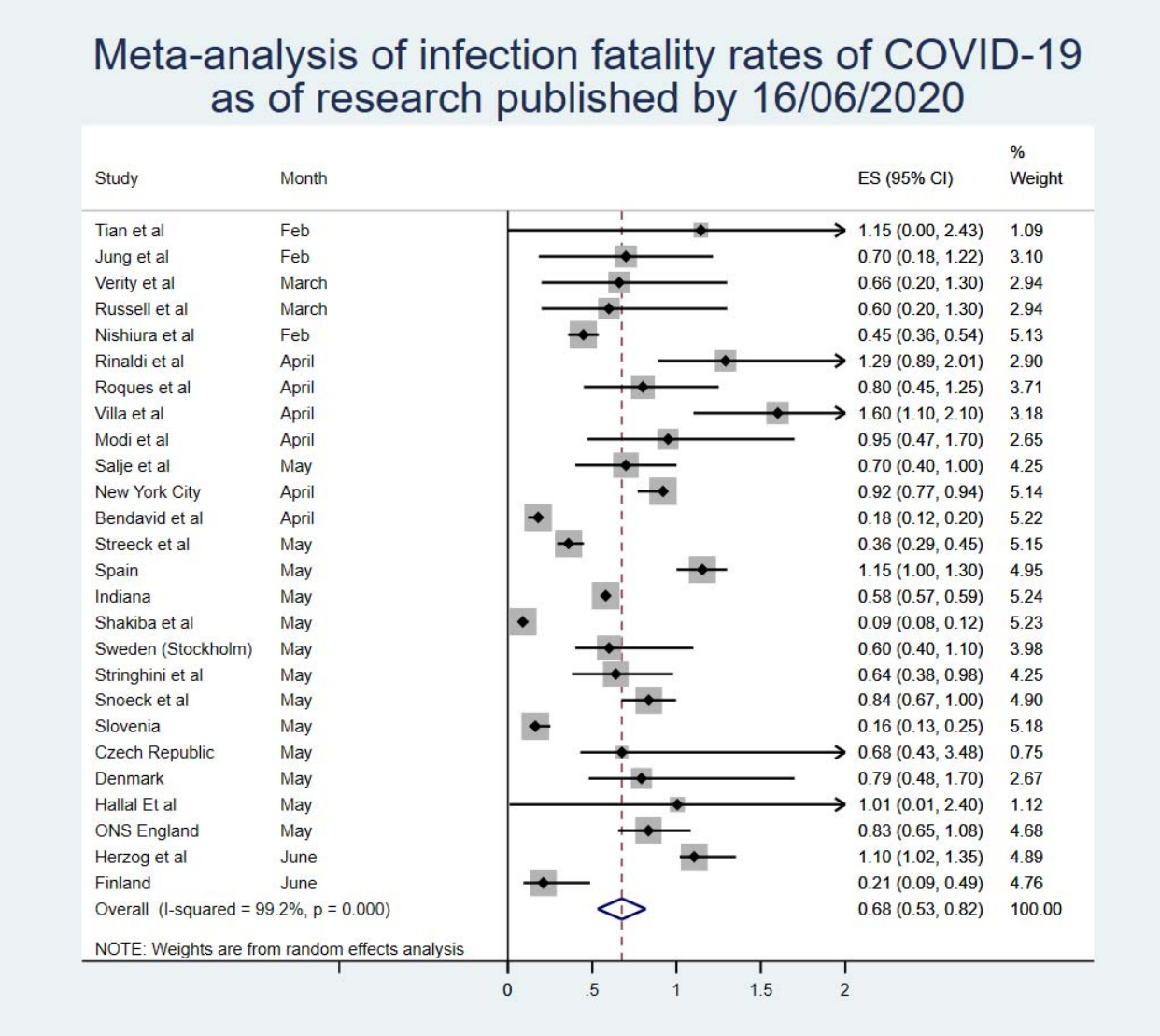



**Figure 1**

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

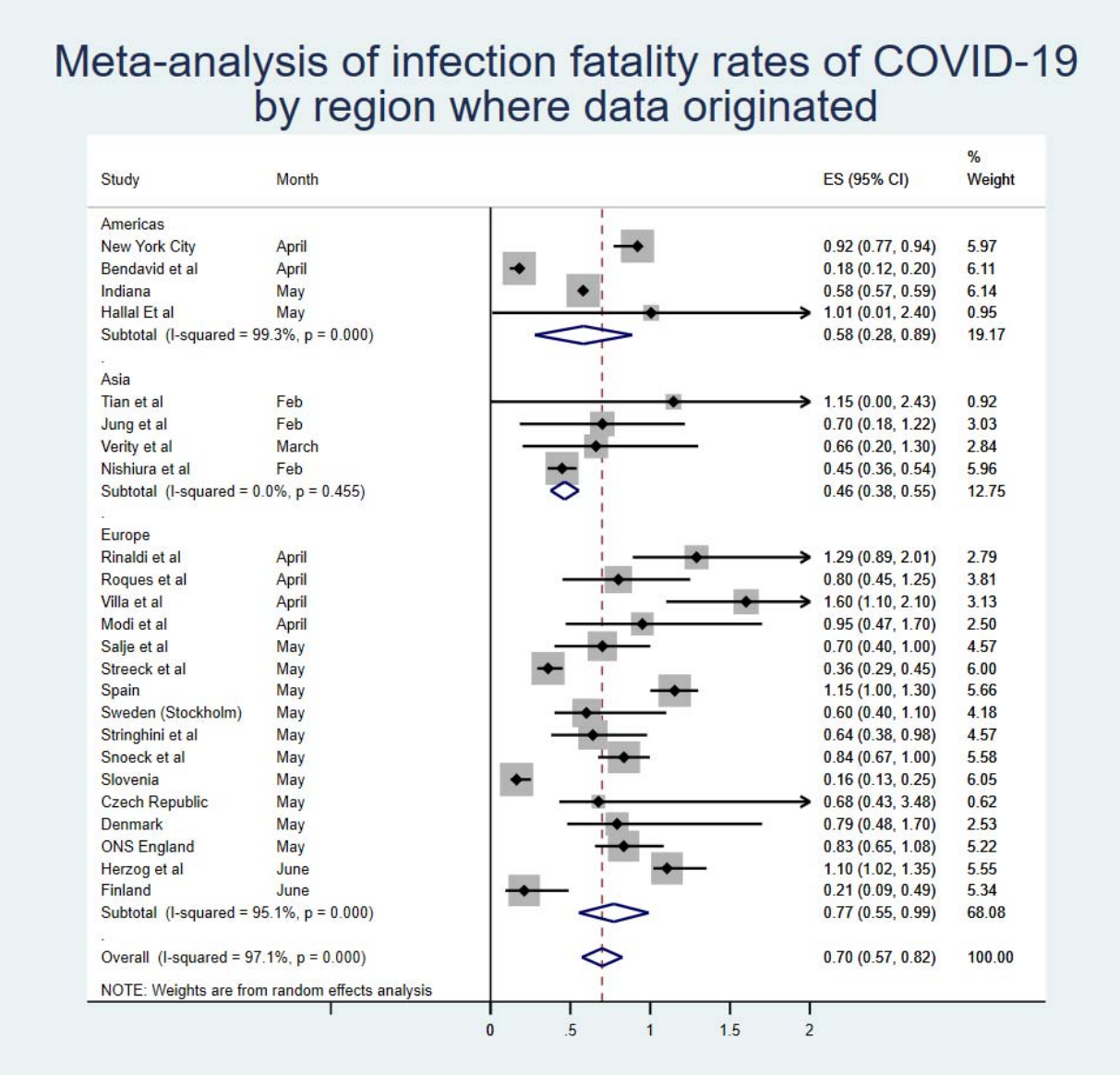



**Figure 2**

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

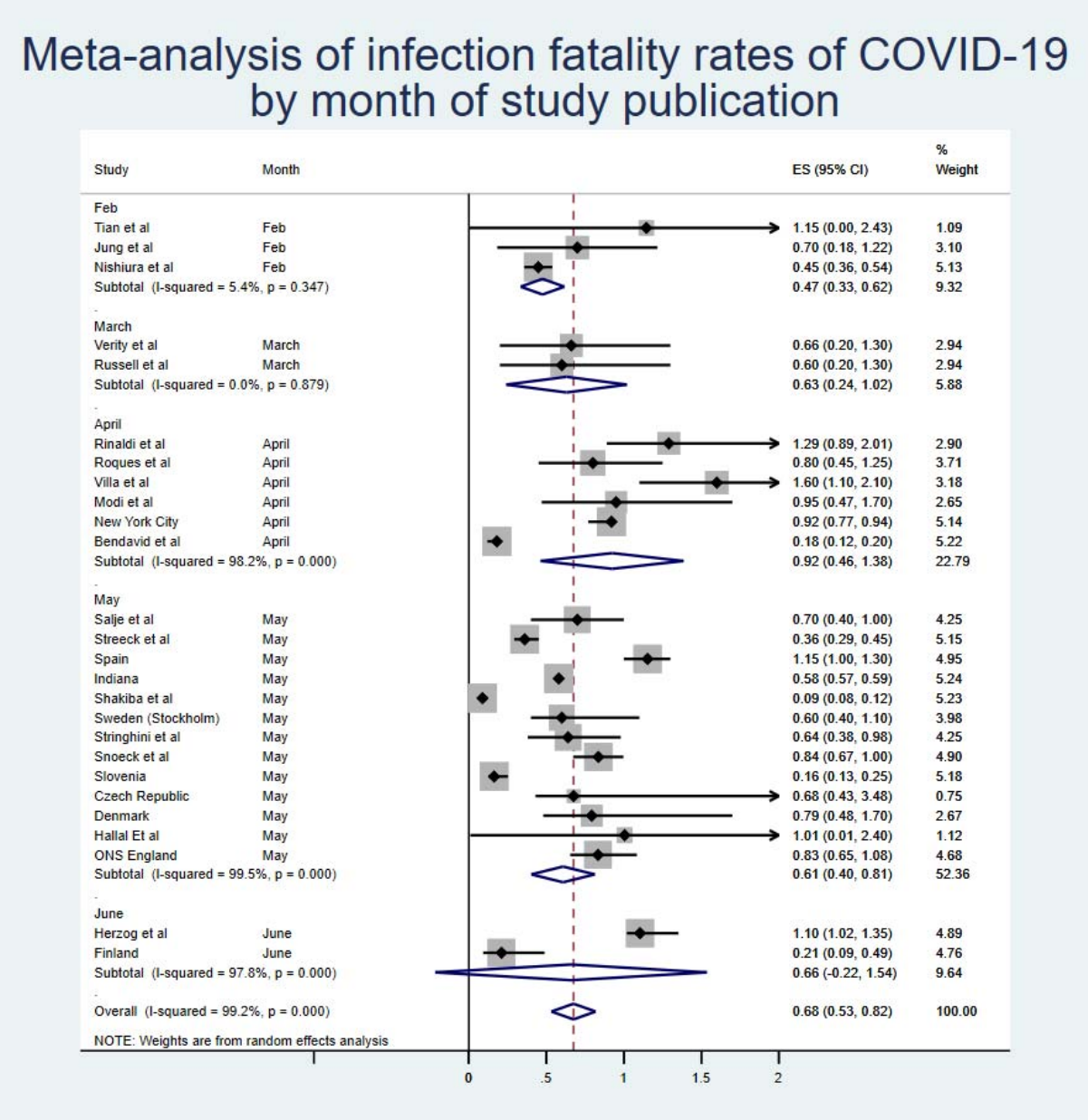



**Figure 3**

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

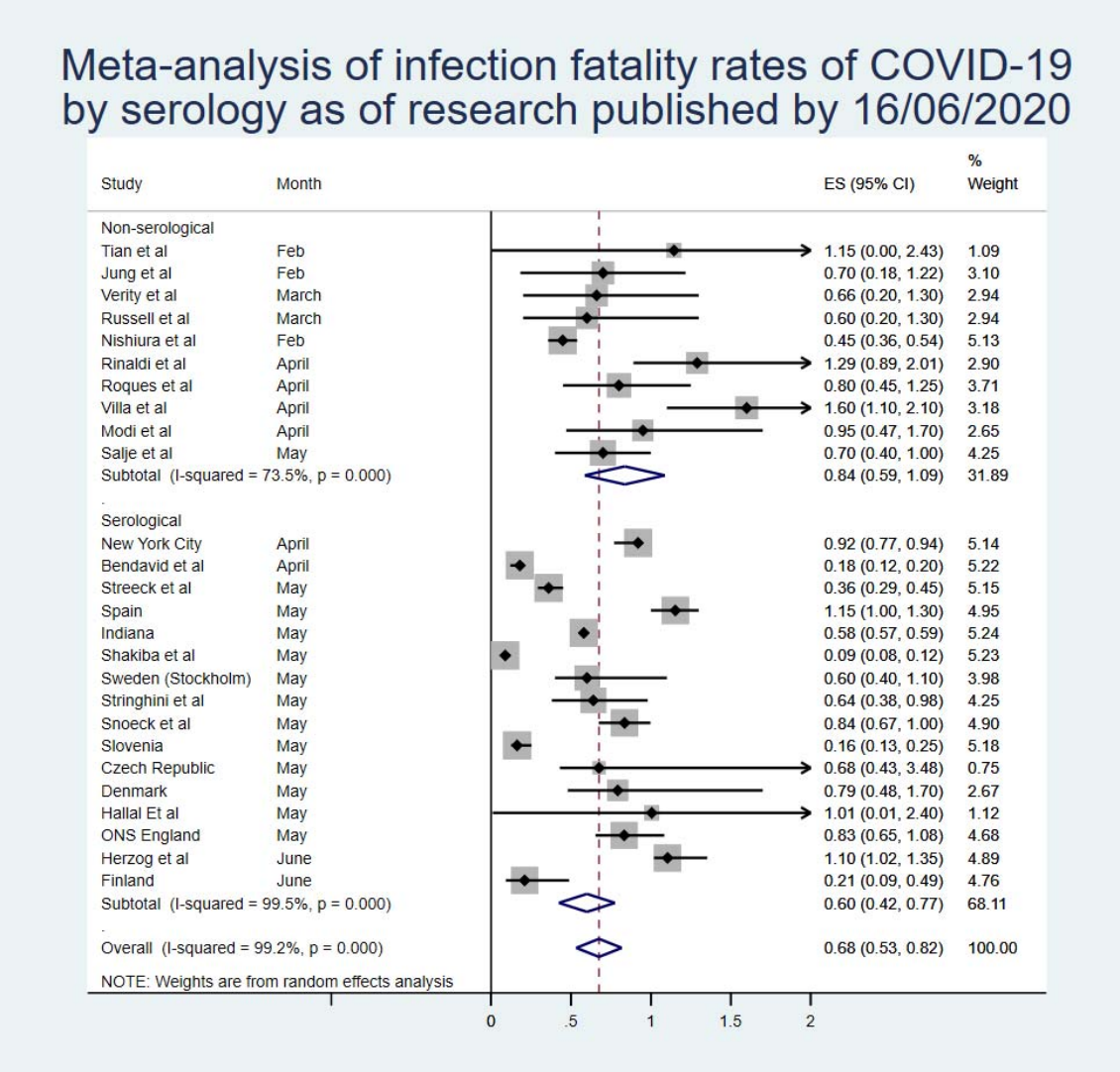



**Figure 4**

There were not sufficient data in the included research to perform a meta-analysis of IFR by age. However, qualitatively synthesizing the data that was presented indicates that the expected IFR below the age of 60 years is likely to be reduced by a large factor. This is supported by studies examining the CFR which were not included in the quantitative synthesis, as well as studies examining IFR in selected populations younger than 70 years of age, that demonstrate a strong age-related gradient to the death rate from COVID-19.

Plotting the studies using a funnel plot produced some visual indication of publication bias, with more high estimates than would be expected, however the Egger's regression was not significant (p=0.74).

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

**Risk of Bias**

As previously noted, all estimates obtained from modelling studies are considered to be at a high risk of bias due to the heterogeneity and difficulty in rating these studies for accuracy. After using the rating "risk of bias" tool for prevalence studies, 6 studies were considered to be at a low risk of bias, 4 studies at a moderate risk of bias, and the remaining 5 estimates at a high risk of bias. This is summarized in the table below (full scoring in supplementary materials):

Table 2 – risk of bias in included serosurveys

| Study | Overall Risk of Bias |
|---|---|
| New York City | Moderate |
| Bendavid et al | High |
| Streeck et al | Low |
| Spain | Low |
| Indiana | High |
| Shakiba et al | Moderate |
| Sweden (Stockholm) | Moderate |
| Stringhini et al | Low |
| Wu et al | High |
| Snoeck et al | Low |
| Slovenia | High |
| Czech Republic | Low |
| Denmark | Moderate |
| Hallal Et al | Low |
| Herzog et al | Moderate |
| Finland | Moderate |
| ONS England | Moderate |

In general, the primary reason for down-rating studies was non-response bias, the lack of representativeness of the population sample, and a lack of information across all fields. Some reports were published with minimal information, which substantially increased the uncertainty and thus the risk of bias in these estimates.

The sensitivity analysis by study quality results are below. Broadly, study quality was correlated with a higher inferred IFR, with lower-quality serosurveys reporting higher estimates of population

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

prevalence than randomly-sampled population-wide prevalence estimates. Restricting the analysis to only those studies at a low risk of bias resulted in modestly reduced heterogeneity and an increased IFR of 0.76% (0.37-1.15%).

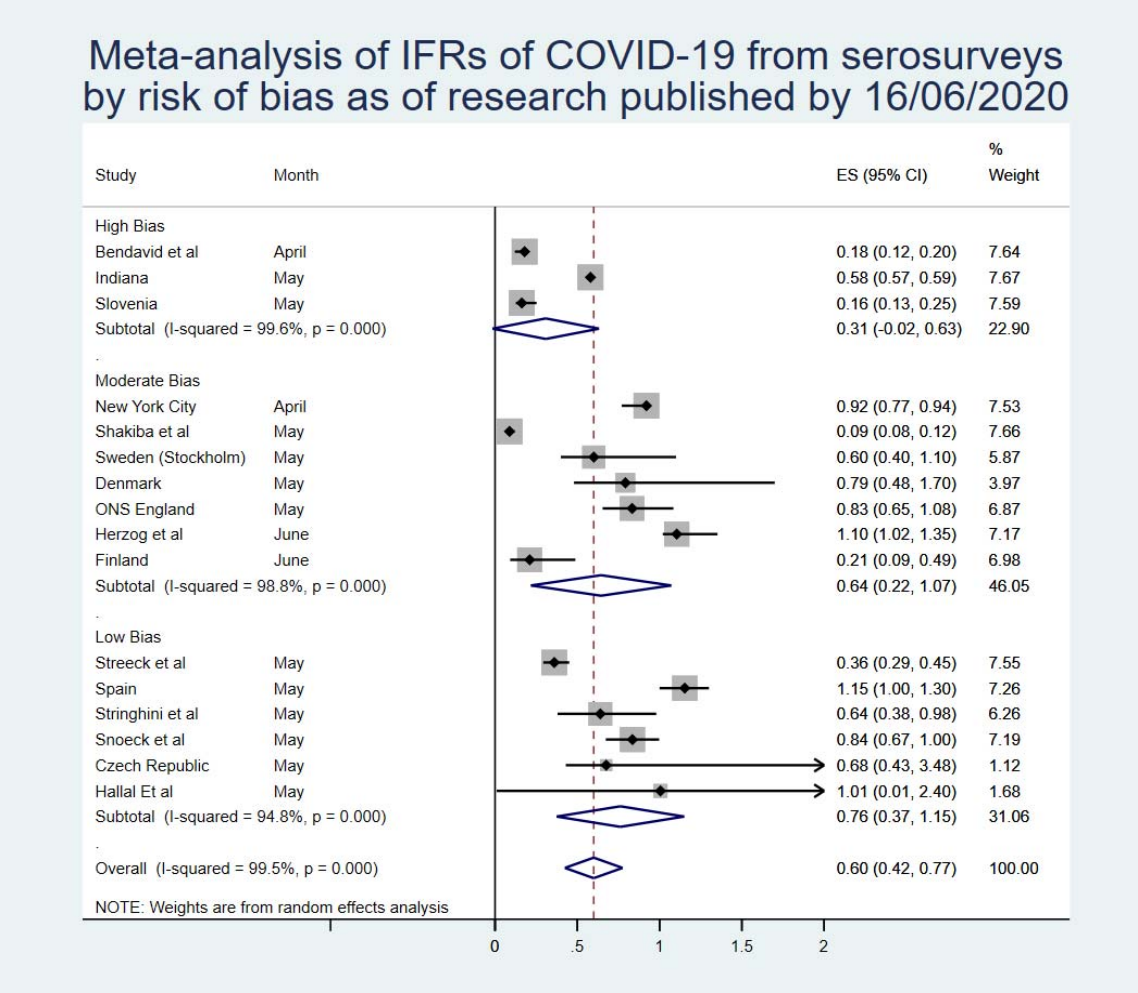



**Figure 5**

**Other Estimates of IFR**

Several estimates of IFR were identified but not included in the meta-analysis as they did not meet the inclusion criteria. The aggregated best estimate from the Centre for Evidence-Based Medicine at Oxford University of 0.1-0.41% (52), and the preprint estimate reported by Grewelle and Leo of 1.04% (0.77-1.38%) (53) were both pertinent but could not be included due to collinearity.

Similarly, the estimate of symptomatic IFR produced by Basu of 1.3% (0.6-2.1%) (54) was excluded due to the exclusion of asymptomatic cases. Using reported estimates of asymptomatic cases, this

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

estimate would likely match the meta-analytic IFR, however this correction could not be applied for the estimates in this study as it could easily introduce bias into the results.

## Discussion

As pandemic COVID-19 progresses, it is useful to use the IFR when reporting figures, particularly as some countries begin to engage in enhanced screening and surveillance, and observe an increase in positive cases who are asymptomatic and/or mild enough that they have so far avoided testing (55). It has been acknowledged that there is significant asymptomatic carriage – potentially up to 50% of all patients - and that asymptomatic transmission may also be possible with COVID-19 (29, 56) and use of IFR would aid the capture of these individuals in mortality figures. IFR modelling, calculation and figures, however, are inconsistent.

The main finding of this research is that there is very high heterogeneity among estimates of IFR for COVID-19 and therefore it is difficult to draw a single conclusion regarding the number. Aggregating the results together provides a point-estimate of 0.68% (0.53-0.82%), but there remains considerable uncertainty about whether this is a reasonable figure or simply a best guess. It appears likely, however, that the true population IFR in most places from COVID-19 will lie somewhere between the lower bound and upper bounds of this estimate.

One reason for the very high heterogeneity is likely that different countries and regions will experience different death rates due to the disease. One factor that may impact this is government response, with more prepared countries suffering lower death rates than those that have sufficient resources to combat a large outbreak (57). Moreover, it is very likely, given the evidence around age-related fatality, that a country with a significantly younger population would see fewer deaths on average than one with a far older population, given similar levels of healthcare provision between the

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

two. For example, Israel, with a median age of 30 years, would expect a lower IFR than Italy, with a much higher median age (45.4 years).

Some included studies (2, 37) compared fatality during COVID-19 pandemic with previous years' average fatality, determining that mortality has been higher during pandemic and whilst correlation doesn't necessarily equate to causation, it is reasonable to link the events as causal given the high CFR observed across countries. It is highly likely from the data analysed that IFR increases with age-group, with those aged over 60 years old potentially experiencing the highest IFR, in one case close to 15% (37). Given the elderly are the most vulnerable in society to illness and likely to carry a higher disease burden owing to increased susceptibility and comorbidity (58, 59), the lower IFRs observed in the younger populations may skew the figure somewhat. There are some reasonable estimates of fatality in younger age-groups that were not included in the population estimates (11, 20, 23), which imply a substantially lower rate of death in the population under 70. While these studies were not considered applicable for quantitative synthesis, they imply that the IFR for <70 year olds is likely lower than 0.1%, and may be less than 10x the rate of death in over-70s. Another recently published estimate stratified infection-fatality by age and found a very low risk for under 50s that increased exponentially with age from 0.0016% <50 years to 0.14% for 50-64 year olds and up to 5.6% for those 65 years and older (60).

While not included in the quantitative synthesis, one paper did examine the extreme lower bound of IFR of COVID-19 in situations where the healthcare system has been overwhelmed. This is likely to be higher than the IFR in a less problematic situation but demonstrates that the absolute minimum in such a situation cannot be lower than 0.2%, and is likely much higher than this figure in most scenarios involving overburdened hospitals.

Of note, there appears to be a divergence between estimates based on serosurveys and those that are modelled or inferred from other forms of testing, with the IFR based purely on serosurveillance being 0.60% (0.43-0.77%). Some have argued that serological surveys are the only proper way to

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

estimate IFR, which would lead to the acceptance of this slightly lower IFR as the most likely estimate (61). However, even these estimates are very heterogeneous in quality, with some extremely robust data such as that reported from the Spanish and Swedish health agencies (40, 62), and some that have clear and worrying flaws such as a study from Iran where death estimates are reportedly substantially lower than the true figure (45). However, when taking quality into account, and only analysing those serosurveys that had a low risk of bias, it is interesting to note that the inferred IFR rises substantially to 0.76% (0.37-1.15%). This may be due to the bias in lower-quality serosurveys being towards a higher prevalence (27), which in turn lowers the IFR substantially.

Another key issue is accounting for deaths. While official death counts were used for all serosurvey estimates, and included in all modelled estimates, these counts are increasingly being recognized as undercounts of the true death figure (37). Published research is already estimating that, even in many wealthy countries with excellent death reporting systems, more than 50% of COVID-19 deaths are likely being missed (63, 64). It is not unlikely that, after correcting for excess mortality not captured in official death reporting systems, the IFR of COVID-19 in most populations would be higher than 1%.

Conversely, there is evidence that the tests used in these serosurveys have drawbacks despite their high specificity and sensitivity. For example, in asymptomatic/mild cases, the tests may have reduced sensitivity, leading to a biased overestimation of the IFR (65). A recent systematic review and meta-analysis of serological tests for COVID-19 found that even the better serology tests would likely overestimate prevalence in an area with few cases and underestimate prevalence when many people had already been infected (9). In areas with a prevalence of 1-2%, for example, the systematic review implies that a study employing an enzyme-linked immunosorbent assay to examine antibodies would produce an estimated infection rate almost double the true prevalence. This would then cause the IFR be underestimated by the same fraction.

There are a number of limitations to this research. Importantly, the heterogeneity in the meta-analysis was very high. This may mean that the point-estimates are less reliable than would be

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

expected. It is also notable that any meta-analysis is only as reliable as the data contained within – this research included a very broad range of studies that address slightly different questions with a very wide range of methodological rigor, and thus cannot represent certainty of any kind. While modelling studies were not formally graded, at least one has already been critiqued for simple mathematical errors, and given that many were pre-prints it is hard to ascertain if they have provided accurate representations of the data. Serology studies were at variable risk of bias, and analysing by only the highest quality serosurveys produced a higher estimate than relying on lower quality studies.

Moreover, the quality of included serosurvey estimates was often questionable. Many countries have a clear political motivation to present lower estimates, making it challenging to ascertain whether these may have biased the reporting of results, particularly for those places that have only presented results as press releases thus far. Some have also been criticized for sampling issues that would likely lead to a biased overestimate of population infection rates (10).

Accounting for right-censoring in these estimates was also a challenge. Using a 10-day cutoff for deaths is far too crude a method to create a reliable estimate. In some cases, this could be an overestimate, due to the seroconversion process taking almost as much time as the median time until death. Conversely, there is a long tail for COVID-19 deaths (15), and therefore it is almost certain that some proportion of the 'true' number of deaths will be missed by using a 10-day cutoff, biasing the estimated IFRs down. This may be why serosurvey estimates at first appear to result in somewhat lower IFRs than modelled and observational data suggests.

It is also important to recognize that this is a living estimate. With new data being published every single day during this pandemic, in a wide variety of languages and in innumerable formats, it is impossible to collate every single piece of information into one document no matter how rigorous. Moreover, this aggregated estimate is only as correct as the most recent search – the point estimate has not shifted substantially due to the inclusion of new research, but the confidence interval has changed. It is almost certain that, over the course of coming months and years, the IFR will be revised

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

a number of times. In particular, it is vital that future research stratifies this estimate by age, as this appears to be the most significant factor in risk of death from COVID-19.

This research has a range of very important implications. Some countries have announced the aim of pursuing herd immunity with regards to COVID-19 in the absence of a vaccination. The aggregated IFR would suggest that, at a minimum, you would expect 0.45-0.53% of a population to die before the herd immunity threshold of the disease (based on R0 of 2.5-3 (34)) was reached (66). As an example, in the United States this would imply more than 1 million deaths at the lower end of the scale.

This also has implications for future planning. Governments looking to exit lockdowns should be prepared to see a relatively high IFR within the population who are infected, if COVID-19 re-emerges. This should inform the decision to relax restrictions, given that the IFR for people infected with COVID-19 appears to be not insignificant even in places with very robust healthcare systems.

## Conclusions

Based on a systematic review and meta-analysis of published evidence on COVID-19 until May, 2020, the IFR of the disease across populations is 0.68% (0.53-0.82%). However, due to very high heterogeneity in the meta-analysis, it is difficult to know if this represents the 'true' point estimate. In particular, higher-quality serosurveys with lower risk of bias appeared to generate higher IFRs. It is likely that, due to age and perhaps underlying comorbidities in the population, different places will experience different IFRs due to the disease. Given the issues with mortality recording, it is also likely that this represents an underestimate of the true IFR figure. More research looking at age-stratified IFR is urgently needed to inform policy-making on this front.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

Author declarations

The Authors declare no conflicts of interest. No funding was received for this study. A preprint version can be found here: https://www.medrxiv.org/content/10.1101/2020.05.03.20089854v1

This research was not funded.

1. Kahathuduwa CN, Dhanasekara CS, Chin S-H. Case fatality rate in COVID-19: a systematic review and meta-analysis. medRxiv. 2020:2020.04.01.20050476.
2. Rinaldi G, Paradisi M. An empirical estimate of the infection fatality rate of COVID-19 from the first Italian outbreak. medRxiv. 2020:2020.04.18.20070912.
3. Wong JY, Wu P, Nishiura H, Goldstein E, Lau EHY, Yang L, et al. Infection Fatality Risk of the Pandemic A(H1N1)2009 Virus in Hong Kong. American Journal of Epidemiology. 2013;177(8):834-40.
4. Thompson WW, Moore MR, Weintraub E, Cheng P-Y, Jin X, Bridges CB, et al. Estimating Influenza-Associated Deaths in the United States. American Journal of Public Health. 2009;99(S2):S225-S30.
5. Spychalski P, Błażyńska-Spychalska A, Kobiela J. Estimating case fatality rates of COVID-19. The Lancet Infectious Diseases.
6. Variation in False-Negative Rate of Reverse Transcriptase Polymerase Chain Reaction–Based SARS-CoV-2 Tests by Time Since Exposure. Annals of Internal Medicine.0(0):null.
7. Xiao AT, Tong YX, Zhang S. False-negative of RT-PCR and prolonged nucleic acid conversion in COVID-19: Rather than recurrence. Journal of Medical Virology.n/a(n/a).
8. Winter AK, Hegde ST. The important role of serology for COVID-19 control. The Lancet Infectious Diseases.
9. Lisboa Bastos M, Tavaziva G, Abidi SK, Campbell JR, Haraoui L-P, Johnston JC, et al. Diagnostic accuracy of serological tests for covid-19: systematic review and meta-analysis. BMJ. 2020;370:m2516.
10. Bendavid E, Mulaney B, Sood N, Shah S, Ling E, Bromley-Dulfano R, et al. COVID-19 Antibody Seroprevalence in Santa Clara County, California. medRxiv. 2020:2020.04.14.20062463.
11. Erikstrup C, Hother CE, Pedersen OBV, Mølbak K, Skov RL, Holm DK, et al. Estimation of SARS-CoV-2 infection fatality rate by real-time antibody screening of blood donors. medRxiv. 2020:2020.04.24.20075291.
12. Simon P. Robust Estimation of Infection Fatality Rates during the Early Phase of a Pandemic. medRxiv. 2020:2020.04.08.20057729.
13. Hoy D, Brooks P, Woolf A, Blyth F, March L, Bain C, et al. Assessing risk of bias in prevalence studies: modification of an existing tool and evidence of interrater agreement. Journal of clinical epidemiology. 2012;65(9):934-9.
14. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. The Lancet Infectious Diseases. 2020;20(5):533-4.
15. Giorgi Rossi P, Emilia-Romagna C-wg, Broccoli S, Angelini P. Case fatality rate in patients with COVID-19 infection and its relationship with length of follow up. J Clin Virol. 2020;128:104415-.
16. Gudbjartsson DF, Helgason A, Jonsson H, Magnusson OT, Melsted P, Norddahl GL, et al. Spread of SARS-CoV-2 in the Icelandic Population. N Engl J Med. 2020.
17. Doi A, Iwata K, Kuroda H, Hasuike T, Nasu S, Kanda A, et al. Estimation of seroprevalence of novel coronavirus disease (COVID-19) using preserved serum at an outpatient setting in Kobe, Japan: A cross-sectional study. medRxiv. 2020:2020.04.26.20079822.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

18. Takita M, Matsumura T, Yamamoto K, Yamashita E, Hosoda K, Hamaki T, et al. Preliminary Results of Seroprevalence of SARS-CoV-2 at Community Clinics in Tokyo. medRxiv. 2020:2020.04.29.20085449.
19. 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. 2020:2020.05.11.20095158.
20. Valenti L, Bergna A, Pelusi S, Facciotti F, Lai A, Tarkowski M, et al. SARS-CoV-2 seroprevalence trends in healthy blood donors during the COVID-19 Milan outbreak. medRxiv. 2020:2020.05.11.20098442.
21. 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. medRxiv. 2020:2020.04.27.20082289.
22. 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. 2020:2020.04.18.20071134.
23. Thompson C, Grayson N, Paton R, Lourenço J, Penman B, Lee LN, et al. Neutralising antibodies to SARS coronavirus 2 in Scottish blood donors - a pilot study of the value of serology to determine population exposure. medRxiv. 2020:2020.04.13.20060467.
24. Ed Slot BMH, Chantal B.E.M. Reusken et al. Herd immunity is not a realistic exit strategy during a COVID-19 outbreak, 28 April 2020, PREPRINT (Version 1) available at Research Square [+https://doi.org/10.21203/rs.3.rs-25862/v1+]. Herd immunity is not a realistic exit strategy during a COVID-19 outbreak. In Review. 2020.
25. Silveira M, Barros A, Horta B, Pellanda L, Victora G, Dellagostin O, et al. Repeated population-based surveys of antibodies against SARS-CoV-2 in Southern Brazil. medRxiv. 2020:2020.05.01.20087205.
26. 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 in Boise, Idaho. Journal of Clinical Microbiology. 2020:JCM.00941-20.
27. Sood N, Simon P, Ebner P, Eichner D, Reynolds J, Bendavid E, et al. Seroprevalence of SARS-CoV-2–Specific Antibodies Among Adults in Los Angeles County, California, on April 10-11, 2020. JAMA. 2020.
28. Wilson L. SARS-CoV-2, COVID-19, Infection Fatality Rate (IFR) Implied by the Serology, Antibody, Testing in New York City. SSRN. 2020.
29. Nishiura H, Kobayashi T, Yang Y, Hayashi K, Miyama T, Kinoshita R, et al. The Rate of Underascertainment of Novel Coronavirus (2019-nCoV) Infection: Estimation Using Japanese Passengers Data on Evacuation Flights. Journal of clinical medicine. 2020;9(2).
30. Jung SM, Akhmetzhanov AR, Hayashi K, Linton NM, Yang Y, Yuan B, et al. Real-Time Estimation of the Risk of Death from Novel Coronavirus (COVID-19) Infection: Inference Using Exported Cases. Journal of clinical medicine. 2020;9(2).
31. Salje H, Tran Kiem C, Lefrancq N, Courtejoie N, Bosetti P, Paireau J, et al. Estimating the burden of SARS-CoV-2 in France. Science. 2020:eabc3517.
32. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. The Lancet Infectious diseases. 2020.
33. Tian S, Hu N, Lou J, Chen K, Kang X, Xiang Z, et al. Characteristics of COVID-19 infection in Beijing. The Journal of infection. 2020;80(4):401-6.
34. Russell TW, Hellewell J, Jarvis CI, van Zandvoort K, Abbott S, Ratnayake R, et al. Estimating the infection and case fatality ratio for coronavirus disease (COVID-19) using age-adjusted data from the outbreak on the Diamond Princess cruise ship, February 2020. Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin. 2020;25(12).
35. Roques L, Klein E, Papaix J, Sar A, Soubeyrand S. Using early data to estimate the actual infection fatality ratio from COVID-19 in France. medRxiv. 2020:2020.03.22.20040915.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

36. Villa M, Myers JF, Turkheimer F. COVID-19: Recovering estimates of the infected fatality rate during an ongoing pandemic through partial data. medRxiv. 2020:2020.04.10.20060764.
37. Modi C, Boehm V, Ferraro S, Stein G, Seljak U. Total COVID-19 Mortality in Italy: Excess Mortality and Age Dependence through Time-Series Analysis. medRxiv. 2020:2020.04.15.20067074.
38. Streeck H, Schulte B, Kuemmerer B, Richter E, Hoeller T, Fuhrmann C, et al. Infection fatality rate of SARS-CoV-2 infection in a German community with a super-spreading event. medRxiv. 2020:2020.05.04.20090076.
39. Bassett BA. Strict Lower Bound on the COVID-19 Fatality Rate in Overwhelmed Healthcare Systems. medRxiv. 2020:2020.04.22.20076026.
40. ESTUDIO NACIONAL DE SERO-EPIDEMIOLOGÍA DE LA INFECCIÓN POR SARS-COV-2 EN ESPAÑA. Instituto de Salud Carlos III; 2020.
41. IU, ISDH release preliminary findings about impact of COVID-19 in Indiana. Indiana: Indiana University.
42. 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. 2020:2020.05.11.20092916.
43. First study carried out on herd immunity of the population in the whole territory of Slovenia. In: Slovenia Ro, editor. Slovenia: Republic of Slovenia.
44. Collective immunity study SARS-COV-2-CZ-Preval: preliminary results. Czech Republic; 2020.
45. Shakiba M, Hashemi Nazari SS, Mehrabian F, Rezvani SM, Ghasempour Z, Heidarzadeh A. Seroprevalence of COVID-19 virus infection in Guilan province, Iran. medRxiv. 2020:2020.04.26.20079244.
46. Coronavirus (COVID-19) Infection Survey pilot: 5 June 2020. In: Statistics OFN, editor. England: Office For National Statistics; 2020.
47. Hallal P, Hartwig F, Horta B, Victora GD, Silveira M, Struchiner C, et al. Remarkable variability in SARS-CoV-2 antibodies across Brazilian regions: nationwide serological household survey in 27 states. medRxiv. 2020:2020.05.30.20117531.
48. Notat: Foreløbige resultater fra den repræsentative seroprævalensundersøgelse af COVID-19. In: Institut SS, editor. Denmark: Statens Serum Institut; 2020.
49. Folkhälsomyndigheten. The infection fatality rate of COVID-19 in Stockholm – Technical report. Sweden: Public Health Agency of Sweden; 2020.
50. Weekly Report of the Population Serology Survey of the Corona Epidemic. Finland: Finnish Government; 2020.
51. 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. The Lancet.
52. Jason Oke CH. Global Covid-19 Case Fatality Rates. Centre for Evidence-Based Medicine: University of Oxford; 2020.
53. Grewelle R, De Leo G. Estimating the Global Infection Fatality Rate of COVID-19. medRxiv. 2020:2020.05.11.20098780.
54. Basu A. Estimating The Infection Fatality Rate Among Symptomatic COVID-19 Cases In The United States. Health affairs (Project Hope). 2020:101377hlthaff202000455.
55. Sutton D, Fuchs K, D'Alton M, Goffman D. Universal Screening for SARS-CoV-2 in Women Admitted for Delivery. N Engl J Med. 2020.
56. Bai Y, Yao L, Wei T, Tian F, Jin D-Y, Chen L, et al. Presumed Asymptomatic Carrier Transmission of COVID-19. JAMA. 2020;323(14):1406-7.
57. Scally G, Jacobson B, Abbasi K. The UK's public health response to covid-19. BMJ. 2020;369:m1932.
58. Liu K, Chen Y, Lin R, Han K. Clinical features of COVID-19 in elderly patients: A comparison with young and middle-aged patients. The Journal of infection. 2020:S0163-4453(20)30116-X.

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

59. Rothan HA, Byrareddy SN. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. Journal of Autoimmunity. 2020;109:102433.
60. Perez-Saez J, Lauer SA, Kaiser L, Regard S, Delaporte E, Guessous I, et al. Serology-informed estimates of SARS-COV-2 infection fatality risk in Geneva, Switzerland. medRxiv. 2020:2020.06.10.20127423.
61. Ioannidis J. The infection fatality rate of COVID-19 inferred from seroprevalence data. medRxiv. 2020:2020.05.13.20101253.
62. Första resultaten från pågående undersökning av antikroppar för covid-19-virus. In: Folkhälsomyndigheten, editor. Sweden: Folkhälsomyndigheten; 2020.
63. Modi C, Boehm V, Ferraro S, Stein G, Seljak U. How deadly is COVID-19? A rigorous analysis of excess mortality and age-dependent fatality rates in Italy. medRxiv. 2020:2020.04.15.20067074.
64. Modig K, Ebeling M. EXCESS MORTALITY FROM COVID-19. WEEKLY EXCESS DEATH RATES BY AGE AND SEX FOR SWEDEN. medRxiv. 2020:2020.05.10.20096909.
65. Takahashi S, Bryan Greenhouse, and Isabel Rodríguez-Barraquer. Are Sars-cov-2 Seroprevalence Estimates Biased? OSF Preprints. 2020.
66. Mahase E. Covid-19: UK starts social distancing after new model points to 260 000 potential deaths. BMJ. 2020;368(1089).

**Table 1: number: Results of systematic review of published research data on COVID-19 infection-fatality rates**

| Study | Location | Study period | Method and sample size | Results |
|---|---|---|---|---|
| **Basset *et al* 2020** | New York (NYC) (USA), Madrid, Lombardy | Until 22nd April 2020 (commence date not provided) | Utilised $R_0$ of 2.4 to calculate a predicted infection rate of 81% (UK and USA). | Over the 3 regions, the IFR (using predicted total infection rate of 81%) was calculated at 0.17%, for each region specifically, using the same predicted infection rate: NYC 0.22%, Lombardy 0.15%, Madrid 0.14%. |
| **Bendavid *et al* 2020** | Santa-Clara Country | 2 days | Serological testing of 3,300 local adults and children. Volunteer sampling. Bootstrap procedure used for weighted and unweighted prevalence estimates. | Crude prevalence rate 1.5% (95%CI 1.1-2.0%), unweighted population prevalence 1.2% (bootstrap 95%CI 0.7-1.8%), weighted population prevalence 2.8% (95%CI 1.3-4.7%). Number of infections estimated |

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

| | | | | |
|---|---|---|---|---|
| | | | | to be greater than number of recorded cases. IFR 0.17%. |
| **Government of the Czech Republic** | Czech Republic | Unspecified start date, concluded May 1st 2020 | Tested 26,549 people for antibodies (serology). | Uncovered 107 new cases. |
| **Government of Denmark** | Denmark | Reported on 20th May 2020 | Tested 1,071 people out of a total sample of 2,600 for antibodies. | An estimated 1.1% seroprevalence, with a confidence interval ranging from 0.5-1.8%. |
| **Government of England (Office for National Statistics)** | England, UK | As of May 24th 2020 | Serology samples randomly gathered from 885 people up until 24/05/2020 | 6.78% (5.21-8.64%) tested positive on serological testing regirmen. |
| **Government (State) of Indiana, USA** | Indiana, USA | 7 days | Tested >4,600 using viral PCR and serum for antibodies; 3600 randomly selected individuals and 900 volunteers. | 1.7% tested positive for COVID-19 on PCR plus an additional 1.1% who tested positive for antibodies. Estimated IFR 0.1%. 45% of positive cases report no symptoms. |
| **Government of Finland** | Finland | Week 22 | Random weekly sampling of the Finnish population, week 22 included 178 samples | 5 positives from 178 lead to a 2.81% positive rate, ranging from 1.21-6.41%. |
| **Government of Slovenia** | Slovenia | Not specified | 1367 swabs and 1367 blood samples collected from a representative sample of the population. | 41 people (3.1%) tested positive for COVID-19 antibodies. |

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

| | | | | |
|---|---|---|---|---|
| **Government of Spain** | All provinces of Spain | 27/04/2020-11/05/2020 | 60,983 participants invited, of which so far 37,992 (62.3%) have responded | 5% positive on serology, with death rates varying by region. Calculated IFR between 1-1.3% |
| **Government of Sweden** | Stockholm county | 27/04/2020-3/05/2020 (week 18) | 1,200 weekly samples. Initial analysis was reported from 1,104 samples. | 7.3% tested positive on serology in Stockholm county. Official government report estimates IFR at 0.6% (0.4-1.1%) based on modelling and serological testing. |
| **Hallal *et al* 2020** | Brazil | From 14th May to 21st May 2020 | 46,011 attempts lead to a total of 25,025 samples across every region of Brazil. | An overall seroprevalence of 1.39%, with the authors reporting a calculated IFR of 1%, although it was impossible to ascertain whether this accounted for right-censoring. |
| **Herzog *et al* 2020** | Belgium | Two time periods in April, with the estimate used in this paper from the 20-26th of April 2020 | Total of 7,307 samples from locations around Belgium | 193 out of 3,397 samples tested positive, with a weighted overall seroprevalence of 3.1%. Combined with death estimates, this produced an IFR of 1.1% overall |
| **Jung *et al* 2020** | Cases exported from China and diagnosed outside China | 16 days | A total of 51 cases diagnosed between 24/09/2020 and 09/02/2020. Data collected from government websites or media quoting government announcements. | Mean time from illness onset to death was 20.2 days. Estimated incidence in China on 24/01/2020 was 4718 (95%CI 3328-6278) and CFR 5.3% (95%CI 3.5-7.6%). IFR 0.5-0.8%. |

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

| | | | | |
|---|---|---|---|---|
| **Modi *et al* 2020** | Italy (1688 towns) | Used data from 01/01/2015-28/03/2020 | Utilised data from the Italian Institute of Statistics. Compared death rates during the COVID-19 pandemic to previous death rates by age and region. | Clear increase in deaths was noted for early 2020. IFR increases with age. Range 0.02% (40-49 years old) to 15.1% (>90 years old). |
| **Nishiura *et al* 2020** | Japanese "evacuees" returning to Japan from Wuhan | 3 days | A total of 565 individuals screened for symptoms and tested for COVID-19 (PCR). | A total of 8 passengers tested PCR positive for COVID-19 (1.4%). Estimated ascertainment rate of 9.2%. Estimated IFR 0.3-0.6%. |
| **Rinaldi *et al* 2020** | Northern Italy (10 municipalities in Lombardy) | Utilised 5-year death data until April 2020 | Collected data from the Italian Institute of Statistics. The total population of the included municipalities was 50563. Bayesian model used to estimate IFR. | Deaths between February and April 2020 were 5-fold the 2015-2019 average (341 versus 70). IFR 1.29% (95%CI 0.89-2.01), increasing to 4.25% for those >60 years old (95%CI 3.01-6.39%) |
| **Roques *et al* 2020** | France | 54 days | Obtained data on positive cases and deaths from Johns Hopkins University Centre for Systems Science and Engineering and data on tests performed from Santé Publique France, deaths from nursing homes were added to the official count. | Calculated IFR 0.5% (95%CI 0.3-0.8), when nursing home residents were adjusted for estimated IFR 0.8% (95%CI 0.45-1.25). Estimated ratio between those actually infected and those observed was 8 (95%CI 5-12). |
| **Rosenberg *et al* 2020** | New York State, USA | 9 days | Cross-sectional Seroprevalence | 12.5% of specimens were reactive. |

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

| | | | | |
|---|---|---|---|---|
| | | | study of 15,101 adults. Used IgG immunoassay approved for COVID-19 | Cumulative incidence was noted to be higher in Hispanic people, African American people and non-Hispanic Asian people. |
| **Russel *et al* 2020** | Diamond Princess Cruise Ship | 14-17 days | A total of 3711 passengers and staff were tested (PCR) whilst in quarantine. Utilised data from the World Health Organisation situational reports. | There were 619 confirmed cases (17%), 318 of whom were asymptomatic (51%). Corrected CFR was 2.6% (95%CI 0.89-6.7%). Corrected IFR was 1.3% (95%CI 0.38-3.6%). CFR increased with age (3.6% for those aged 60-69, 95%CI 3.2-4.0) and 14.8% for those >80 years, 95%CI 13.0-16.7). |
| **Salje *et al* 2020** | France (hospital data) and Diamond Princess Cruise Ship | Data available up to 7th May 2020 | Modelling analysis of COVID-19 transmission in France. This included 95,210 hospitalisations and 719 infections from the Diamond Princess Cruise Ship. | 3.6% of infected individuals were hospitalised (95% CI 2.1-5.6%) and this percentage increased with age and gender (0.2% females <20 years; 45.9% males >80 years). IFR 0.7% (95%CI 0.4-1.0%) with a range 0.001-10.1%, increasing with age. |
| **Shakiba *et al* 2020** | Iran | 1 month | Cluster randomised sampling utilised to obtain 551 rapid antibody tests for COVID-19 | 22% antibody seropositivity. 18% (65 subjects) were asymptomatic. IFR 0.08-0.12%. |
| **Snoeck *et al* 2020** | Luxembourg | Data collection commenced | Recruited voluntary residents of | Low prevalence of carriers (0.3%). Seroprevalence of |

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

| | | | | |
|---|---|---|---|---|
| | | April 15th 2020 | Luxembourg. Performed PCR for COVID-19 in 1842 participants and serology testing in 1820 participants. | IgA was 11%, and 2% for IgG. Of 1842 PCR tests, only 6 were inconclusive (0.3%). Time prevalence of COVID-19 was 0.32 (95%CI 0.02-0.63). |
| **Streek *et al* 2020** | Germany | 7 days | A sero-epidemiological CGP and GEP-compliant study in a town exposed to a super-spreading event. Utilised a questionnaire and PCR/serology testing 6 weeks after outbreak. A sample of 919 individuals had evaluable infection status. | Infection rate or 15.5% (95% CI 12.3-19.0%); this was 5-fold reported cases in the community (3.1%). Estimated IFR 0.36% (95% CI 0.29-0.45). |
| **Strighini *et al* 2020** | Switzerland | Three serosurveys over several time periods, with the final results reported on June 2nd | Longitudinal serological survey, with an accompanying paper estimating infection-fatality rate as well. | Infection-fatality rate estimated by authors was 0.64 (0.38-0.98%) after correcting and accounting for demography. |
| **Tian *et al* 2020** | Beijing, China | 21 days | 262 cases retrospectively enrolled and characteristics compared between severe, mild and asymptomatic patients using Mann-Whitney U tests and Wilcoxon tests. | Five patients died and 46 were classified as severe. IFR in Beijing was lower than nationally; 0.9% versus 2.4% ($p$<0.001). |
| **Verity *et al* 2020** | Mainland China and 37 countries | 56 days | Age-stratified CFR estimates on 1334 cases outside | Mean time from illness onset to death 17.8 days |

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .

| | | | | |
|---|---|---|---|---|
| | outside of mainland China | | mainland China. Used prevalence data from PCR-confirmed cases in international residents repatriated from China to determine IFR. | (95%CI 16.9-19.2). CFR in China 1.38% (95%CI 1.23-1.53), increasing with age to 6.8% in those aged >65 years (95%CI 5.7-7.2%) and 13.4% in those aged >80 years (95%CI 11.2-15.9%). IFR 0.66% (95%CI 0.39-1.33%). |
| **Villa *et al* 2020** | Italy | 32 days | Collected data from Italy's Civil Protection Agency from each of Italy's 20 regions. | Estimated an IFR of 1.1% (95%CI 0.2-2.1%) and a CFR of 12.7%. |

medRxiv preprint doi: https://doi.org/10.1101/2020.05.03.20089854; this version posted July 7, 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 4.0 International license .