Title:

Unraveling the flaws of estimates of the infection fatality rate for COVID-19

Short running title:

Flaws in Infection Fatality Rates Estimates

Authors:

Chen Shen[1]	M.A.	
Derrick VanGennep[1]	Ph.D.	
Alexander F. Siegenfeld[1,2]	Ph.D.	
Yaneer Bar-Yam[1]	Ph.D.	Corresponding Author yaneer@necsi.edu

1 New England Complex Systems Institute

2 Department of Physics, Massachusetts Institute of Technology

Highlight (Teaser):

> The infection fatality rate (IFR) of COVID-19 is of importance for policymaking. We show that there are significant flaws in many studies estimating the IFR and used as references by public health authorities.

Author statements:

> The authors contribute equally to this work.
>
> We have no conflict of interest.
>
> There is no fundiing for this submission.

Unraveling the flaws of estimates of the infection fatality rate for COVID-19

Chen Shen, Derrick Van Gennep, Alexander F. Siegenfeld and Yaneer Bar-Yam
New England Complex Systems Institute

The infection fatality rate (IFR) of COVID-19 is one of the measures of disease impact that can be of importance for policy making. However, we have found that there are significant flaws in many of the studies that attempt to estimate the IFR

which lead to biased results, potentially affecting public discourse and policy actions.

In order to demonstrate some of the flaws, we consider papers cited in a meta-analysis[2] that the CDC had used as the basis for its estimate of IFR for an extended period of time.[1] We have previously written a criticism of one of the 24 reports on used by the meta-analysis[2] based on the Diamond Princess cruise[3,4] Starting from an incorrectly estimated 1.3% IFR for the Diamond Princess—the actual IFR for the Diamond Princess being 2.0%—they attempted to model the IFR in China and reported a value 0.6%; however, the multiple errors in their calculations, including dividing by the same factor twice, render their results meaningless. Here we discuss flaws in five additional studies cited in the meta-analysis. We note that these studies were not selected because of the presence of flaws, but instead they were sequentially chosen in each of the three categories in the pre-print version of the meta-analysis. We discuss two types of problems: flaws in the studies per se, and flaws in the way the meta-analysis used them.

I. *The Rate of Underascertainment of Novel Coronavirus (2019-nCoV) Infection: Estimation Using Japanese Passengers Data on Evacuation Flights*.[5] At the end of January, Japan conducted three repatriation flights of Japanese nationals and their families on Jan 29-31 from Wuhan, China. Upon disembarkation from the planes, tests were performed of the passengers and the test results were used to estimate IFR in Wuhan.

Non-representative sample – Japanese nationals likely differ in ways relevant to disease transmission due to their social networks and geographic areas they frequent, and the 565 passengers are not independent samples due to the presence of family units and correlated social networks and behaviors.

Nonsensical mathematical analysis – the authors write an equation in which they set the cumulative number of infections in Wuhan c(t)/q (the number of detected infections divided by the ascertainment rate) times the detection window T divided by the total population of Wuhan n equal to the fraction of evacuees who tested positive (8/565)—i.e.

$$\frac{8}{565} = \frac{c(t)T}{qn}.$$

However, the left-hand side of the equation is dimensionless while the right-hand side has units of time. This renders the equation meaningless.

No external validity to the ascertainment rate – even if it were correctly derived, however, it has no validity for application to other contexts, as it applies only to a particular snapshot in time when tests in China were just starting to ramp up. Any application of this rate to convert CFR to IFR under other conditions is therefore inappropriate and misleading (see Section II).

Ignoring confidence interval uncertainty – the ascertainment rate comes with a 95% CI range that can half or double the point estimate of q equals 10%, which authors are evidently aware of as they noted q with $(5.0 - 20\%)$ 95% CI. Yet when calculating IFR subsequently, where IFR = CFR × q, they use 10% for q without addressing the error range. Of course, given all of the flaws present in deriving the ascertainment rate, even a correctly calculated error range would be meaningless.

II. *Real-Time Estimation of the Risk of Death from Novel Coronavirus (COVID-19) Infection: Inference Using Exported Cases*.[6] This study modeled the epidemic growth rate in China and estimated a CFR of $5 - 8\%$. and an IFR of $0.5 - 0.8\%$ by multiplying the cCFR of the study by the ascertainment rate from the previously described study in Section I.

Inappropriate duplicate use of studies – there is no independent assessment of the ascertainment rate.[6] Since the ascertainment rate is a key determining factor for IFR, this is not an independent evaluation of the IFR as would be required for correct inclusion in a meta-analysis.

Flawed ascertainment rate – we note again that the ascertainment rate that is being used has multiple invalidating flaws detailed previously.

III. *Estimating the burden of SARS-CoV-2 in France*.[7] By assuming that the French and Diamond Princess cruise ship populations differ only in their age/sex composition and by assuming a particular distribution of infected individuals across age/sex groups, the authors estimate the IFR for each age/sex group in France. They report an overall IFR of 0.5% (95% credible interval: 0.3 to 0.9%).

Mistatement of study's conclusions – In the abstract the authors state that "0.5% of those infected die" but later in their manuscript they note that all COVID-19 deaths occurring outside of hospitals—and in particular all retirement/nursing home populations—are excluded from their analyses. While it is valid to conduct an analysis that applies only to part of a population, it is misleading to report results derived from such analyses as if they apply to the population as a whole as the authors do in the abstract, particularly given that they end up excluding 9604 of the total 25990 COVID-19 deaths reported by May 7 (the final date for data used in this study).[7] If hospitalized deaths are included, it increases their IFR estimates by a factor of approximately 1.59.

Unjustified assumptions regarding relative probabilities of infection – They assume that infections for each age/sex group are proportional to contact rate based on the supposed linearity of their Fig. S17. However, Fig. S17 actually has substantial curvature to it, bringing this assumption into question. Furthermore, the data they use to ascertain the contact rates of each of the age/sex groups is from a study conducted in 2012 rather than during the pandemic period. The authors then make a series of assumptions that are not validated by empirical data as to the effect of the lockdown on these contact rates in each age group.

Non-representative sample (Diamond Princess) used to set the overall IFR – The IFRs by age/sex group that they calculate are essentially derived from those on the Diamond Princess, using the French data to understand the relative effect of age/sex on IFR. Their analysis is accurate only to the extent that the IFR for each age/sex group on the Diamond Princess equals the IFR for that age/sex group in France. This assumption is not justified; individuals on the Diamond Princess could substantially differ from individuals of the same age and sex in France. In particular, cruise ship passengers are, for their age, likely to be healthier and (due to active screening) to have received more rapid medical attention than the general population. Their sensitivity analysis in which the Diamond Princess population has a 25% lower IFR for each age/sex group is inadequate; it would not be surprising if the Diamond Princess IFRs for each particular age/sex group differed from their French counterparts by a greater factor.

No justification of sensitivity analysis for Bayesian priors – Their paper is based on Bayesian analysis, but they provide no justification for their choice of priors, nor is any sensitivity analysis for the priors reported. It is therefore unclear how their results would differ had they chosen a different set of priors.

Duplicate use of Diamond Princess data in meta-analysis: By including both the Diamond Princess study[4] and this study that estimated IFR based on the Diamond Princess data, the meta-analysis effectively includes IFR estimation of the Diamond Princess data twice, a similar meta-analysis error as found in Section II.

IV. *Characteristics of COVID-19 infection in Beijing*.[8] This is an observational study of patients from Beijing, with 262 confirmed cases that were transferred to designated hospitals and 3 deaths. It is used as 1.15% IFR in the meta-analysis.

Non-representative population sample – The patients over-represent travelers (from Wuhan), a group that is typically younger and healthier than the general population.

Meta-analysis flaw – The study reports a CFR but the meta-analysis[2] misstates the CFR as the IFR.

V. *COVID-19 Antibody Seroprevalence in Santa Clara County, California*.[9] This widely criticized study estimated a 2.8% prevalence rate, meaning infections were more than 50 times higher than the confirmed cases in Santa Clara county at the time of the study. The estimated IFR was 0.17%, one of the lowest IFR estimations the meta-analysis cited.

Non-representative sample – The population sample was obtained by on-line sign-up without measures to determine whether or not it was a representative sample.

Incorrect statistical accounting of false positives – There were only 50 positive test results out of a sample of 3330, a number that could be dominated by noise from even a small false positive rate of the antibody tests.

VI. *Final Notes.* Our review points out multiple flaws in a selection of papers used to estimate IFR. In addition to these specific flaws, none of the studies discussed above

attempted to estimate the number of deaths from COVID-19 that were not attributed to COVID-19, leading to systematic underestimates of the IFR, as noted in the meta-analysis.[2] Excess mortality data[10] suggests a significant number of unattributed deaths from COVID-19, although some of the excess mortality may be from indirect effects of the pandemic (which are nevertheless an impact of the pandemic).

It is also important to note that the IFR is not a biological invariant but rather significantly varies with the distribution of age, health and other qualities of those infected as well as the medical care they receive. Thus, the IFR calculated for a particular population of infected individuals is not immediately generalizable to other populations. Meta-analyses that consider IFR as a well-defined quantity and that average values from various populations to produce a single estimate without accounting for such differences are therefore problematic. The lead author has recently posted an additional meta-analysis that investigates age-stratified IFRs;[11] such an analysis is an improvement but still does not take into account other factors such as differences in the underlying health of each age group and differences in access to treatment.

References

1 CDC. COVID-19 Pandemic Planning Scenarios. https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html (accessed November 18, 2020).

2 Meyerowitz-Katz G, Merone L. A systematic review and meta-analysis of published research data on COVID-19 infection-fatality rates. *medRxiv* 2020; DOI: 10.1101/2020.05.03.20089854 (preprint).

3 Russell T, Hellewell J, Jarvis C, et al. Estimating the infection and case fatality ratio for COVID-19 using age-adjusted data from the outbreak on the Diamond Princess cruise ship. *medRxiv* 2020; DOI: 10.1101/2020.03.05.20031773 (preprint).

4 Shen C, Van Gennep D, Bar-Yam Y. The IFR of the Diamond Princess has been misreported, best current value is 2.0%. *New England Complex Systems Institute.* https://necsi.edu/the-ifr-of-the-diamond-princess-has-been-misreported-best-current-value-is-2 (accessed November 18, 2020).

5 Nishiura H, Kobayashi T, Yang Y, 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**, 419. DOI: 10.3390/jcm9020419.

6 Jung S, Akhmetzhanov AR, Hayashi K, 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**, 523. DOI: 10.3390/jcm9020523.

7 Salje H, Tran Kiem C, Lefrancq N, et al. Estimating the burden of SARS-CoV-2 in France. *Science* 2020; **369**:6500, 208-211. DOI: 10.1126/science.abc3517.

8 Tian S, Hu N, Lou J, et al. Characteristics of COVID-19 infection in Beijing. *Journal of Infection* 2020; **80**, 401–406. DOI: 10.1016/j.jinf.2020.02.018.

9 Bendavid E, Mulaney B, Sood N, et al. COVID-19 Antibody Seroprevalence in Santa Clara County, California. *medRxiv* 2020; DOI:10.1101/2020.04.14.20062463 (preprint).

10 Rossen LM, Branum AM, Ahmad FB, Sutton P, Anderson RN. Excess Deaths Associated with COVID-19, by Age and Race and Ethnicity — United States, January 26–October 3, 2020. *MMWR Morb Mortal Wkly Rep* 2020; **69**: 1522–7.

11 Levin AT, Hanage WP, Owusu-Boaitey N, Cochran KB, Walsh SP, Meyerowitz-Katz G. Assessing the Age Specificity of Infection Fatality Rates for COVID-19: Systematic Review, Meta-Analysis, and Public Policy Implications. *medRxiv* 2020; DOI: 10.1101/2020.07.23.20160895 (preprint).