Reconciling Diverse Estimates of COVID-19 Infection Rates

Authors: Sen Pei, Marta Galanti, Teresa Yamana, Jeffrey Shaman

Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University

Quantification of COVID-19 total infection rates is critical for understanding the penetration of the virus within communities, the epidemiological characteristics of the causative agent, SARS-CoV-2, and proper delineation of the infection fatality rate. The first quantity enables estimation of residual susceptibility within communities and associated risk of (re-)emergent outbreaks. The second set of measures provide insight into the underlying transmission dynamics of the virus and means for controlling its spread. The last metric, the rate of death per infection, is critical for estimating the potential loss of life associated with continued spread.

Reported COVID-19 cases only make up a portion of overall SARS-CoV-2 infections, as many individuals, particularly those experiencing mild or no symptoms, are never tested for the virus. Multiple lines of evidence, including anecdotal reports of superspreader events, cross-sectional studies from cruise ships, and mathematical and statistical analyses, indicate that undocumented infections, though unrecognized, transmit the virus and are important for both its spread within communities and its geographic translocation. By quantifying these otherwise undocumented infections, associated metrics including the ascertainment rate (AR)—the fraction of infections that are reported—the overall infection rate, and the infection fatality rate (IFR)—deaths per infection—can be determined. These measures allow quantification of the depletion of susceptible individuals, the risk of rebound outbreaks once social distancing measures are relaxed, and potential future deaths. However, at present, estimates of undocumented infections and associated metrics are quite divergent. This variance is due to differences in reporting from location-to-location that arise from varying rates of seeking clinical care, differing testing capacities, variable public health data compilation and sharing practices, and differing quantitative approaches.

In recent weeks, more direct measurements of undocumented infections have begun to appear. These measurements include qPCR testing of the broader community for acute infection with the virus and, more recently, seroassays capable of detecting SARS-CoV-2-induced antibodies in samples from blood bank donations and the general population. Seroassay field studies, in particular, provide the opportunity for large-scale quantification of cumulative infection rates—both documented and undocumented—within populations. While these seroprevalence studies are attractive in that they provide direct measurement of indicators of infection, the tests are prone to a number of biases.

One potential bias arises from variability in antibody response to COVID-19 infection. Indeed, preliminary evidence indicates that antibody response is heterogeneous among recovered patients with a substantial percentage (~30%) manifesting low or undetectable antibody levels[1]. Such under-detection would yield low-biased estimates of prevalence and high-biased estimates of AR and IFR. For the purposes of estimating population susceptibility and the chance of future infection, this misidentification might not be as consequential, as individuals without strong antibody titers may already be prone to reinfection. However, it is not known whether low or undetectable antibody levels

post-infection is indicative of susceptibility to repeat infection; indeed, the presence of other factors, such as memory T cells, even in the absence of detectable antibodies might still confer protection.

A second potential bias arises from the sensitivity and specificity of the seroassay employed. Minimizing false negatives and, in particular, false positives, is critical for obtaining unbiased estimates of antibody prevalence, especially if cumulative infection rates are in truth low. Even small increases in the false positive rate (e.g. perhaps due to cross reaction with existing antibodies against other pathogens) could lead to large over-estimation of cumulative infection rates. If test specificity is reliably quantified the effects of these errors can be factored into calculations; however, if specificity is lower than claimed[2], the false discovery rate will unknowingly climb (Figure 1A). For example, consider a county of population 1.6 million with 1,000 confirmed cases and thus a true AR of 10% and true infection rate of 0.63%. A random sample of 10,000 individuals using a seroassay with 95% specificity and 100% sensitivity would yield an AR of 1.1% and an unadjusted infection rate of 5.6%. If specificity drops to 90%, the biases increase, giving an AR of 0.59% and an infection rate of 10.6% (Figure 1A). This example ignores further complications possibly introduced by variable antibody response or waning immunity, which would bias total infection rates low and AR high.

A third potential bias may arise from a failure to appropriately account for immunological and epidemiological lags inherent to COVID-19 disease progression. Immunological lags produce delays from infection acquisition to antibody production, which may vary from individual to individual and depend on the type of antibody measured, e.g. IgM v. IgG. IgM SARS-CoV-2 antibodies don't begin to appear until symptom onset or later and don't persist[3]; symptom onset lags infection acquisition; and IgG SARS-CoV-2 antibodies develop still later and persist longer. As a consequence of these lags, measurements of SARS-CoV-2 antibody rates only provide an estimate of earlier cumulative infection rates (7 days earlier or more), which, if not accounted for, will bias total infection and AR estimates low. Similarly, the lag from infection acquisition to case confirmation is roughly 10-16 days, and the lag from infection to death is even longer. These delays must all be reconciled in order to appropriately calculate AR and IFR.

Early modeling and epidemiological analyses of the COVID-19 pandemic estimated ARs of 6-66% depending on the country and time period. For example, Li et al.[4] estimated a 14% AR in China during mid-January prior to travel restrictions and control efforts. However, a number of more recent studies using serological surveys or alternate analytic approaches have estimated substantially lower ARs ranging from 0.1% to 2%. Recently, Bendavid et al.[5] employed a serosurvey in Santa Clara county California and estimated a 1.2-2% AR. Our own modeling of COVID-19 transmission in China prior to travel restrictions and control measures (January 10-23, 2020) provides a means for exploring whether ARs that low are really feasible. This period of time is during a phase of relatively unconstrained epidemic growth and expansion from the epicenter in Wuhan. While recognizing that our model system could be structurally biased in ways we do not yet recognize, we find that an ascertainment rate of 2% is not feasible without a much higher rate of seeding of latent infections at the January 10 start of model inference and simulation, which requires assumption of either an earlier spillover emergence of the virus or a faster doubling time. These assumptions are at odds with findings based on phylogenetic reconstructions and estimates of transmission rates prior to January 10. However, even with these assumptions, and a fixing of a 2% ascertainment rate, the fit of the model to observations is much weaker (Figure 1B-C).

Accurate estimation of total infections, the AR and the IFR are critical for understanding the progression and future risk from COVID-19. Multiple estimates derived from independent approaches are needed to provide unbiased estimates of these quantities. Seroassays with high sensitivity and specificity can

ultimately support this work; however, the potential biases in the data derived from these tests must be thoroughly vetted. Reliable and widespread antibody testing is needed, as are further determination of variable antibody generation in response to infection and quantification of rates of antibody waning over time. Such information is critical for informing estimation of residual population susceptibility, vaccine needs and back to work strategies.

Acknowledgements

This work was supported by US NIH grant GM110748 and US NSF RAPID grant DMS2027369. JS and Columbia University disclose partial ownership of SK Analytics. JS also discloses consulting for BNI.

References

1. 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;doi:10.1101/2020.03.30.20047365.
2. Lassauniere R, Frische A, Harboe ZB et al. Evaluation of nine commercial SARS-CoV-2 immunoassays. medRxvi. 2020;doi:10.1101/2020.04.09.20056325.
3. Guo L, Ren L, Yang S, et al. Profiling early humoral response to diagnose novel coronavirus disease (COVID-19). Clinical Infectious Diseases. 2020; ciaa310, doi:10.1093/cid/ciaa310.
4. Li R, Pei S, Chen B, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). Science. 2020;doi: 10.1126/science.abb3221
5. Bendavid E, Mulaney B, Sood N, et al. COVID-19 antibody seroprevalence in Santa Clara County, California. medRxiv. 2020; https://www.medrxiv.org/content/10.1101/2020.04.14.20062463v1

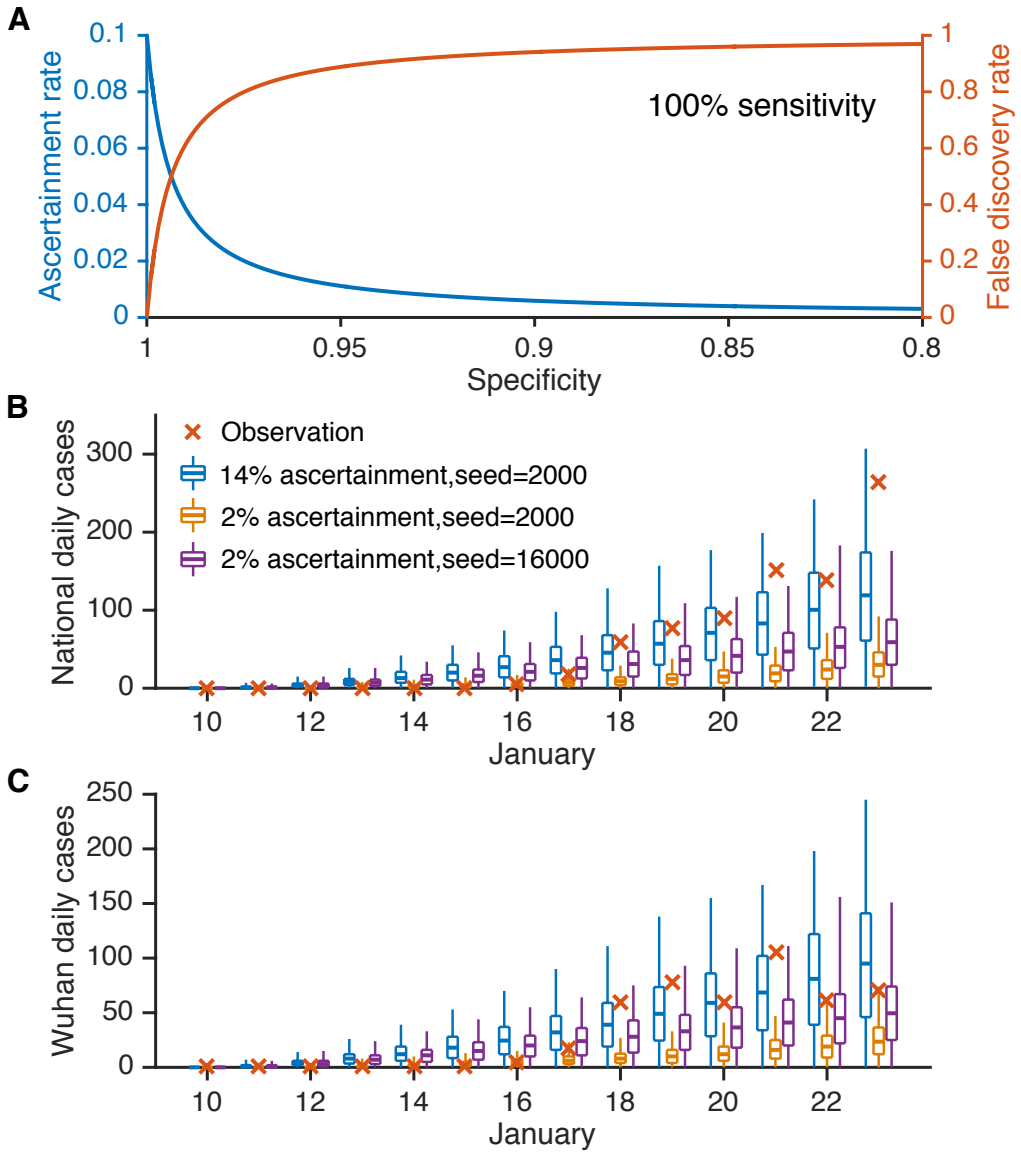

Figure 1. Impact of test specificity and model fittings to confirmed cases in China. Assuming 100% test sensitivity, the ascertainment rate and false discovery rate (i.e., false positive/(false positive + true positive)) for a range of test specificity are shown (Panel A). Model fittings to national daily cases (Panel B) and Wuhan daily cases (Panel C) in China from January 10 2020 to January 23 2020 using different ascertainment rates and seeding scenarios are compared (Model as in Li et al.[4]). Fittings were generated by free simulations using the estimated parameters in each scenario. Boxes and whiskers show the median, interquartile and 95% CI of fittings. The log-likelihoods for the fittings are -217.6 (14% ascertainment, seed=2000), -521.3 (2% ascertainment, seed=2000) and -342.7 (2% ascertainment, seed=16000), respectively.