Response to Lourenço et al. 2020

Carson C. Chow
Laboratory of Biological Modeling, NIDDK, NIH
carsonc@nih.gov

March 27, 2020

A recent paper by Lourenço et al. [1] claims that it is possible that the Covid-19 epidemic in the UK and Italy may have started a month before the first reported death and a large fraction of the population may have already been infected and are thus no longer susceptible. The authors employ a standard SIR model but instead of fitting the model to the number of cases directly, which could underestimate the total number, they instead fit to the cumulative number of deaths. They then show that a model with a very low rate of progression from infection to severe illness is consistent with the data and thus the virus may spread already spread to much of the UK population conferring herd immunity.

Here, it is argued that while this scenario is possible it is somewhat unlikely. The problem is that the model is under constrained by death data alone so that a wide range of parameter values can fit the data equally well. Repeating their analysis with broader prior distributions shows that the posterior distributions for the model parameters and in particular the rate of progression of illness as well as the time delay between infection and death are broad and actually peaked at levels for higher rate of progression and lower time delay.

The model used in [1] is

$$\frac{dy}{dt} = \beta y(1-z) - \sigma y \tag{1}$$

$$\frac{dz}{dt} = \beta y(1-z) \tag{2}$$

$$D = N\rho\theta z(t-\psi) \tag{3}$$

where y is the fraction of the infected population, z is the fraction not susceptible, D is the cumulative number of deaths, N is the total population, $\beta = R_0\sigma$ is the transmission coefficient, σ is the loss of infection rate, ρ is the fraction of population at risk of severe infection, θ is the probability of death with severe infection, and ψ is the time delay between infection and death. The model is fit to the UK data for cumulative deaths [2] using a Metropolis-Hastings algorithm as in [1]. However, non-informative priors are used for all parameters except for ψ which is constrained from becoming too large (ψ values greater than 50 are weakly penalized). As noted below, the model parameters are not identifiable and thus priors are necessary to regularize the fit. Note that ρ and θ are also not identifiable so the product $\rho \times \theta$ is fit as one parameter. The model simulation was initiated by setting $z = 0$ and $y = 1/N$ mimicking a single person becoming infected at time 0. The predicted cumulative death number D is compared to the data after a time shift of ψ.

The MCMC results for the posterior distributions are shown in Figure, which are generated by 100000 samples after ensuring equilibrium is reached. They are broad for all parameters and in fact the modes favor the conclusion of high ρ and low ψ indicating that there is no widespread non-symptomatic infection. Models were also fit with $R_0 = 2.25$ and similar results were obtained. These results can be understood by observing that in the linear regime where $z << 1$, which is where the model is fit to the data, z has the approximate exponential form $z \propto e^{(\beta-\sigma)(t)} + C$ and thus $D = Ae^{(\beta-\sigma)(t-\psi)} + C = e^{(\beta-\sigma)(t-\psi)-\log A} + C$, where A includes ρ. This implies that the model prediction is approximately invariant along the curve $(\beta-\sigma)\psi - \log A$, along which individual parameters are not identifiable. In the MCMC samples there was a correlation coefficient of approximately 0.5 between these two quantities.

The basic conclusion as the authors noted is that the model is underconstrained by cumulative death rates alone and thus a wide variety of parameter settings can fit the model. Although widespread infection seems unlikely, their suggestion of widespread testing for anti-bodies to SARS-CoV-2 is most certainly warranted.

All code was written in the Julia language and available at https://github.com/ccc1685/covid-19.git. I would like to thank Kevin Hall for helpful suggestions. This work was supported by the Intramural Research Program of the NIH, NIDDK.

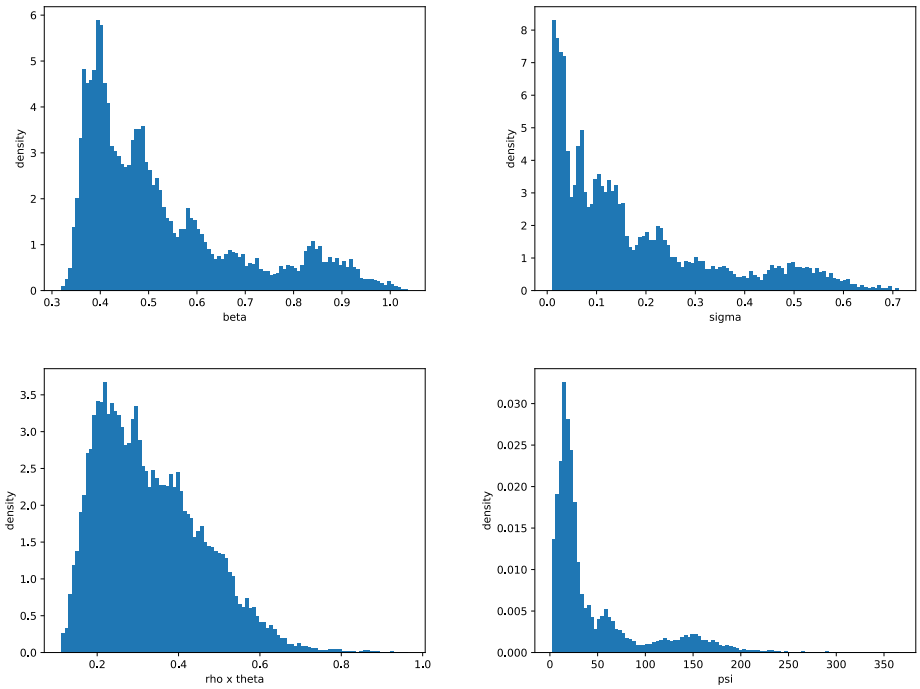

Figure 1: Posteriors for model parameters

References

[1] José Lourenćo, Robert Paton, Mahan Ghafari, Moritz Kraemer, Craig Thompson, Peter Simmonds, Paul Klenerman, and Sunetra Gupta. Fundamental principles of epidemic spread highlight the immediate need for large-scale serological surveys to assess the stage of the SARS-CoV-2 epidemic. Preprint 2020.

[2] Novel Coronavirus (COVID-19) Cases provided by CSSE at Johns Hopkins University: https://github.com/CSSEGISandData/COVID-19.