COVID-19 and the difficulty of inferring epidemiological parameters from clinical data

Simon N. Wood[1], Ernst C. Wit[2], Matteo Fasiolo[1] & Peter J. Green[1]

Knowing the infection fatality ratio (IFR) is crucial for epidemic management: for immediate planning; for balancing the life years saved against those lost to the consequences of management; and for evaluating the ethics of paying substantially more to save a life year from the epidemic than from other diseases. Impressively, Verity et al. (2020) rapidly assembled case data and used statistical modelling to infer the IFR for COVID-19. We have attempted an in-depth statistical review of their paper, eschewing statistical nit-picking, but attempting to identify the extent to which the (necessarily compromised) data are more informative about the IFR than the modelling assumptions. First the data.

- Individual level data for outside China appear problematic, because different countries have differing levels of ascertainment and different disease-severity thresholds even for classification as a case. Their use in IFR estimation would require country-specific model ascertainment parameters, about which we have no information. Consequently these data provide no useful information on IFR.
- Repatriation flight data provide the sole information on Wuhan prevalence (excepting the lower bound of confirmed cases). 689 foreign nationals eligible for repatriation are doubtfully representative of the susceptible population of Wuhan. Hence it is hard to see how to usefully incorporate the 6 positive cases from this sample.
- Case-mortality data from China provide an upper bound for IFR, and, with extra assumptions, on the age dependence of IFR. Since prevalence is unknown, they contain no information for estimating the absolute IFR magnitude.
- Because of extensive testing, the Diamond Princess (used only for validation by Verity et al.) supplies data on both infections and symptomatic cases, with fewer ascertainment problems. These data appear directly informative about IFR. Against this, the co-morbidity load on the DP is unlikely to fully represent any population of serious interest (perhaps fewer very severe but more milder co-morbidities).

Secondly, the modelling assumptions: we see two primary problems.

1. Verity et al. correct the Chinese case data by assuming that ascertainment differences across age groups determine case rate differences. Outside Wuhan they replace observed case data by the cases that would have occurred if each age group had the same per-capita observed case rate as the 50-59 group. They assume complete ascertainment for the 50-59s. These are very strong modelling assumptions that will greatly affect the results: but the published uncertainty bounds reflect no uncertainty about them. In Wuhan, the complete ascertainment assumption is relaxed by introducing a parameter, but one for which the data appear uninformative, so the results will be driven by the assumed uncertainty.

2. Generically, Bayesian models describe uncertainty both in the data and in prior beliefs about the studied system. Only when data are informative about the targets of modelling can we be sure that prior beliefs play a small role in what the model tells us about the world. In this case the data are especially uninformative: we suspect results are mostly the consequence of what our prior beliefs were.

Taken together these problems indicate that Verity et al.'s IFRs should be treated very cautiously when planning. While awaiting actual measurements, we would base IFRs on the DP data, with the Chinese case-fatality data informing the dependence of IFR on age: in supplementary material we provide a crude Bayesian model with its IFR estimates by age. Corresponding population IFR estimates and 95% credible intervals are China: 0.43% (.23,.65), UK: 0.55% (.30,.82); India 0.20% (.11,.30). The strong assumptions required, by this approach too, emphasize the need for improved data. We should replace complex models of inadequate clinical data, with simpler models of epidemiological prevalence data from appropriately designed random sampling using antibody or PCR tests.

Verity, R., L. C. Okell, I. Dorigatti, P. Winskill, C. Whittaker, N. Imai, G. Cuomo-Dannenburg, H. Thompson, P. G. Walker, H. Fu, et al. (2020). Estimates of the severity of coronavirus disease 2019: a model-based analysis. *The Lancet Infectious Diseases.*

[1] University of Bristol, UK
[2] Università della Svizzera italiana, Switzerland