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

Ernst C. Wit, Università della Svizzera italiana

Knowing the infection fatality ratio (IFR) is of crucial importance for evidence-based epidemic management: for immediate planning; for balancing the life years saved against the life years lost due the consequences of management; and for evaluating the ethical issues associated with the tacit willingness to pay substantially more for life years lost to the epidemic, than for those to other diseases. Against this background, in an impressive paper, Verity et al. (2020) have rapidly assembled case data and used statistical modelling to infer the IFR for COVID-19.

Given the importance of the issues, the necessarily compromised nature of the data and the consequent heavy reliance on modelling assumptions, my collaborators and I present an in-depth statistical review of what has been done. We have attempted this, conscious that the circumstances require setting aside the usual standards of statistical nit-picking. Facilitated by Verity et al. (2020)'s exemplary provision of their code and data, we have attempted to identify to what extent the data are sufficiently informative about the IFR to play a greater role than the modelling assumptions, and have tried to identify those assumptions that appear to play a key role.

After having identified some of the weakness in the analysis, we propose a crude alternative Bayesian model to estimate the IFR, which results in lower values. Nevertheless, we do not believe that it is possible to model our way out of the deficiencies in the clinical data in order to estimate crucial epidemiological parameters. There is an urgent need to replace complex models of inadequate clinical data, with simpler models using adequate epidemiological prevalence data based on appropriately designed, random sampling.

This is joint work with Simon N. Wood, Matteo Fasiolo and Peter J. Green