

Discussion of “An epidemiological forecast model and software assessing interventions on the COVID-19 epidemic in China”

SHANNON GALLAGHER*¹

¹*Department of Clinical Research, National Institute of Allergy and Infectious Diseases*

Wang et al. provide a useful and important step in making epidemic models accessible to both experts and non-experts alike. Via their R package ‘eSIR’, the authors create a tool that allows for flexible modeling of SIR epidemics via a Bayesian hierarchical model.

During the ongoing COVID-19 pandemic, statistical models are in abundance (see <https://github.com/midas-network/COVID-19>) and eSIR differentiates their contribution from other works through an abundance of accessibility, reproducibility, and transparency in their modeling framework.

With regards to accessibility, the authors complement their ideas with purposeful figures, equations, and verbal explanations, which serve to explain their contributions than any one feature alone. Moreover, the authors release their work along with their data via an R package and a basic Shiny app.

With regards to reproducibility, all code is open-source and is available publicly online. Moreover, the authors provide the code for their calculations along with a vignette, and the package “just works.”

Finally, the authors are generally transparent in stating their assumptions and processing of the data. Data are messy, and this is demonstrated in the sudden change in case counts in China on January 12, 2020. Wang et al. clearly explain their procedure to smooth over or ‘calibrate’ the discrepancy and present their modeling results with and without this calibration.

One line of questioning I would pursue, with regards to transparency, is the origin of the observed number of recovered/removed individuals Y_t^R . New case counts are commonly reported but new recovered/removed are reported far less. Since the number of recovered is associated with estimating γ , the recovery rate and hence R_0 , it is important to analyze how robust estimates are to any pre-processing steps of Y_t^R .

The implementation of quarantine procedures in the eSIR model is both useful and novel, especially with the comparison of the inclusion of the global modifier in the SIR model and the SIR model with an additional quarantine compartment. This shows an interesting view in how two models which attempt to describe the same phenomenon (i.e. quarantining a population) can produce different results, especially with regards to R_0 , the initial reproduction number.

The models described in this paper describe a population that is homogeneous with regards to infection dynamics or at least, on average acts in the same manner. Future work could be dedicated to extending the general SIR model to sub-groups within the population such as to close-quarters, children and adults, or other socio-economic factors. A consequence of adding more and more strata is additional parameters to estimate which may hinder computational tractability of the model. Moreover, simply developing the ODEs which such a model would follow ($f(\cdot)$ in the authors notation) may be unclear. Since individual-level models and agent-based models are used to model groups with multiple strata and local, heterogeneous interaction, it would be interesting to analyze if and when compartment models become supplanted by these individual-level and agent-based models.

*Email: shannon.gallagher@nih.gov