STATISTICAL MODELS FOR COVID-19 INCIDENCE, CUMULATIVE PREVALENCE, AND R_t

Nicholas P. Jewell, Ph.D.
Department of Medical Statistics
London School of Hygiene & Tropical Medicine

We appreciate the opportunity to discuss the thought-provoking paper by Quick, Dey and Lin (2021), which we shall refer to as QDL in the following for convenience.

There is a lot of material to unpack in the paper, and I fear that I will not do justice to this impressive work by largely ignoring many of the modeling and computational techniques used in the authors' analyses of COVID-19 infection counts and ascertained cases, calibrated by various seropositivity assessments, all data covering both space and time. According to the authors prescription, I will refer to the proposed model through their acronym, MERMAID. I hope that my overlooking key issues will be redressed by other discussants and by future research inspired by the paper.

General Remarks

From a broad perspective, it is useful to begin a discussion by setting expectations for a retrospective statistical model of community transmission. In considering this for MERMAID, the following quote from Chris Whitty (the current Chief Medical officer for England) in the pre-COVID-19 world of 2015 is on point:

"The tendency of some modelers to present them as scientific predictions of the future rather than models does not help. Models are widely used in government, and some models have arguably too much influence. They are generally most useful when they identify impacts of policy decisions which are not predictable by commonsense; the key is usually not that they are 'right', but that they provide an unexpected insight."

While MERMAID is not designed for predictive purposes, I think readers can usefully approach the methodology and analyses as an exercise in reading the results from any modeling exercise: what were the unexpected insights learned?

Characteristics of MERMAID

There has been a plethora of quantitative models that try to both explain patterns of COVID-19 infections and provide insight into the scale of future transmissions and associated hospitalizations. Many of these approaches correctly identify inaccuracies in available counts of infections, compounded by lags in hospitalization or mortality

counts that are often used as more reliable sources. A partial remedy for underreporting (particularly likely for infections with mild or no symptoms) is to pre-process infection counts using external information about ascertainment rates. Here, QDL models infections, ascertainment, lags, and connections to seroprevalence surveys, simultaneously. It would be interesting to elucidate the statistical advantages of this more comprehensive approach as compared to more narrow analyses of adjusted infection counts to the extent that infection patterns are of primary interest.

MERMAID involves four components: (i) a model of the (unobserved) number of infections per unit time, (ii) aallocation of infections to a potential day of confirmation, (iii) a delay model that focuses on *observed* confirmed cases based on the number of infections and an ascertainment probability that depends on testing metrics, and finally (iv) calibration of the number of infections over time as they are reflected in population estimates of cumulative seroprevalence. The models work at a regional level (here states in the US), largely because there is anticipated and clear variation in many of the model parameters from state to state. The exception is the inclusion of regression effects for R_t based on crude measures of non-pharmaceutical interventions (NPIs) where state effects are inextricably linked as discussed further below.

There are two general characteristics of MERMAID that are important to note: (i) the model is *statistical*, or purely empirical, as compared to methods that rely on dynamic transmission models whether these are deterministic or stochastic, and (ii) the model is fundamentally *retrospective* in nature. Let me discuss both of these issues in turn.

MERMAID is purely empirical

To a statistician, this is an attractive feature--so long as the statistical model is compatible with a reasonable understanding of how available data is generated—as it avoids detailed epidemiological understanding of transmission characteristics. At first glance, the initial stage of MERMAID modelling seems to involve an epidemiological model for infection growth, centred on R_t, the reproductive number associated with SARS-Cov-2 infections; however, this does not appear essential to the overall approach as noted below. Loosely speaking, R_t refers to the expected number of infector-infectee pairs, arising from a single infector, at time t. One disadvantage of R_t for public consumption is that does not carry any information on time, that is, how long it takes for an infector to transmit the pathogen to subsequent infectees.

As an aside, estimation of R_t depends on the distribution of the generation interval, and not the serial interval; these two distributions the same mean, but have different variances in general which is important. Here, the usually unobserved generation

interval refers to the time between infections for an infector-infectee pair (always positive), whereas the serial interval describes the time between symptom onset for the same pair, a time that is more likely observable but may be negative. It is known that use of an estimated serial interval distribution in place of that for the generation interval may result in biased estimates of R_t (Gostic *et al.*, 2020). It is unclear whether use of the serial interval is an issue here. A related concern is that is likely that the generation and serial interval distributions vary over time and space, particularly as a pathogen mutates and responds to local environments, whereas MERMAID assumes that the serial distribution is fixed.

The growth rate, r_t (the rate of change of the log-transformed case incidence), presents an alternative to R_t with advantages that (i) it does carry information on the speed at which cases grow, (ii) does not depend on any epidemiological model of transmission, and (iii) is more understandable to lay people (Parag and Donnelly, 2020). There are challenges to estimate either measure from available data: Parag, Thompson, and Donnelly (2021) discuss the relative merits of the two estimates with a focus on guiding public health interventions in real time. The growth rate, therefore, appears to provide an equally effective basis for the first part of the MERMAID model, removing any reliance on a formal transmission model and thereby stressing the empirical nature of the model.

Before going further, it is important to note that empirical models carry their own assumptions, and that not all such models of epidemic growth have been successful in the past (Ellenberg and Morris, 2021; Jewell and Jewell, 2021; Jewell, Lewnard and Jewell, 2020ab). To its great credit, MERMAID avoids the pitfalls of earlier statistical modelling attempts of COVID-19 infection counts, nor is there any attempt to extrapolate the model to the future as we discuss further below. As noted, MERMAID still requires a lot of assumptions to yield valid interpretations, many of them discussed carefully by the authors. Other discussants will further elaborate on other assumptions, and I return to interpretation of the R_t regression below. I also point to the challenges of modelling ascertainment rates, here allowed to potentially depend on the number of tests and/or the fraction of the population being tested, data that are often publicly available. However, in many locations, ascertainment also depended heavily on the strategy for testing, that is who was targeted for testing and at what level, particularly with regard to asymptomatic testing. Such strategic decisions are far less documented and hard to capture quantitatively.

The authors note that identifiability of the ascertainment model depends on the availability of seroprevalence data. This suggests the possibility of purely statistical

correlation between estimates of ascertainment rates and seroprevalence. However, this phenomenon is identified as a primary finding of the data analyses, raising the question of how much of the detected relationship reflects a true effect and how much is related to an identifiability artefact.

MERMAID is necessarily retrospective

The chronological time period covered in the illustrative data analyses is March to December 2020, largely covering the first wave of the SARS-Cov-2 outbreak. The intent is thus to provide insight into what happened in the past, and not what will happen in the future. Undoubtedly, this is of considerable interest but it is less useful for real-time policy decisions that hinge on influencing where communities might go in the future. To be fair, empirical models (and mechanistic models for that matter) have not, in general, successfully *predicted* COVID-19 infection counts more than a month out, and model ensemble estimates only partially address key uncertainties. Epidemic models differ fundamentally from many apparently similar complex prediction efforts (such as national meteorological models) since human behavior and governmental policies are influenced by predictions, necessarily modifying subsequent outcomes.

This issue is particularly important to emphasize for infectious disease outbreaks where interventions are needed *before* exponential growth is established—delay in action contributes a huge fraction of future disease burden (Roberts, 2020). A key public health dilemma when responding to a novel pathogen is that intervention may be required before a substantial amount of data is available. In such cases, evidence-based decisions may not be feasible, and mitigation efforts must be launched on the bases of other scientific assessments including mathematical models. In the presence of potential exponential growth, public health experts recognize that early interventions must be implemented before there is a significant caseload and will be subject to significant criticism *particularly if they are successful.* At this point in the COVID-19 pandemic, the public health community must assess the success off various policy decisions, and ask both what have we learned and how we can improve in the future.

Above, we briefly stated what components MERMAID combines into a comprehensive model of past infection patterns. But what does the model not address? Amongst other issues, MERMAID does not yet allow for (i) natural feedback loops, (ii) day of the week effects, (iii) variants of concern, (iv) vaccine, natural immunity and cross-immunity effects, (v) subgroup (i.e. age) variation. In addition, it does not address the impact of school closures and re-openings on community transmission, nor the impact of other major events in communities (e.g. Euro 2020—see Smith *et al.*, 2021).

An additional challenge to the application of both mathematical and statistical models to the COVID-19 pandemic, is that models generally fail to link different risk communities—for example, a connection between community transmission and risk within long-term care facilities. This linkage of transmission between distinct at-risk communities was ultimately recognized to be very important in understanding HIV transmission dynamics.

The simulations provided by QDL provide useful insight into the robustness of MERMAID to some assumptions. It would be of great interest to consider additional comparisons with alternative empirical models. In such comparisons, statisticians can provide valuable insights into the right metric for comparison. In a related but different point, how to best calibrate an empirical model in the presence of poor-quality data is important. MERMAID uses community seroprevalence surveys that are largely based on convenience samples of one form or another, and are only available at idiosyncratic intervals. Is there any advantage to community testing of sewage samples for SARS-CoV-2 (Larsen and Wigginton, 2020)? Are there better approaches? We return to this below.

Regression Models for R_t

This component is perhaps the aspect of MERMAID that is of most interest to policy makers, and simultaneously the most challenging. As QDL emphasize, their regression models for R_t are associational and not causal. It is necessarily hard to quantify the effects of population interventions on R_t that yield a causal interpretation, in part because R_t is multifactorial and subject to inherent feedback mechanisms. Transmission rates depend on the (i) intensity and nature of contacts (i.e. mixing), (ii) infectivity associated with contacts, and the (iii) duration of infectiousness. Further, policy interventions are varied, directed at each of these factors with different impacts, and constantly change over time.

As such, models as they currently stand, including MERMAID, often cannot answer a key question required for a data-driven policy decision: how much reduction in R_t can be expected through implementation of a defined "amount" of mitigation? As we have argued elsewhere, there has been far less reliance on randomized experiments than might be desirable (Jewell and Jewell, 2021; Fretheim, 2020; McCartney, 2020), other than in trials of vaccine and therapeutic efficacy, perhaps reflecting a form of experiment aversion where social policies are concerned. For many intervention questions associated with reducing the risk of COVID-19 infection, there is clear evidence of equipoise, reducing ethical concerns associated with experimentation.

For identifiability reasons, MERMAID constrains the regression effects on R_t to be constant across states, i.e. the impact of interventions must be consistent in different regions. QDL invoke the assumption that the effects of mitigation policies are constant both over time and regions. But, as we illustrate below, regression effects are likely to vary depending on location (i.e. causal effects from one location have poor transportability to another).

For example, Miller *et al.* (2020) estimate changes in estimated R_t over time in various US states in the early stages of the COVID-19 pandemic. Their Figure 3 (reproduced here, in part, as Figure 1) compares R_t estimates with a mobility measure of Relative Routing Volume (RRV) for four specific states. Note that chronological time in Figure 1 flows from right to left since mobility decreased throughout March 2020. As seen in Figure 1, this association between R_t and reduction in RRV is quite different across the states, with the estimated R_t falling below one at different levels of RRV depending on the state. For Louisiana, R_t was reduced to one when RRV fell to 65% (58-75%) of baseline levels. On the other hand, New York's R_t fell below one only when RRV was reduced to 48% (43-56%) of baseline. Reductions in RRV below 80% of baseline delivered diminishing returns in reducing R_t in Louisiana, while the slope in New York was maximized at RRV around 50% of baseline.

For COVID-19, regression models for R_t have been almost universally cross-sectional in nature. However, observed relationships when mobility decreases may not be a reasonable description of what may happen to R_t "in reverse," that is, when mobility increases. The same is likely true for other mitigation strategies such as mask mandates and transportation restrictions. Specifically, Nouvellet *et al.* (2021) demonstrate quite different regression effects on R_t between an increasing and decreasing mobility covariate. An interesting extension to the MERMAID regression model for R_t might exploit *longitudinal* data to capture effects that permit interactive effects with the covariate "direction".

Clearly, an important goal is to understand the "left to right" relationship—i.e. as mitigations decrease—sufficiently well to develop strategies for easing restrictions that increase mobility, and loosen mask mandates etc., while at the same time minimizing increases in R_t.

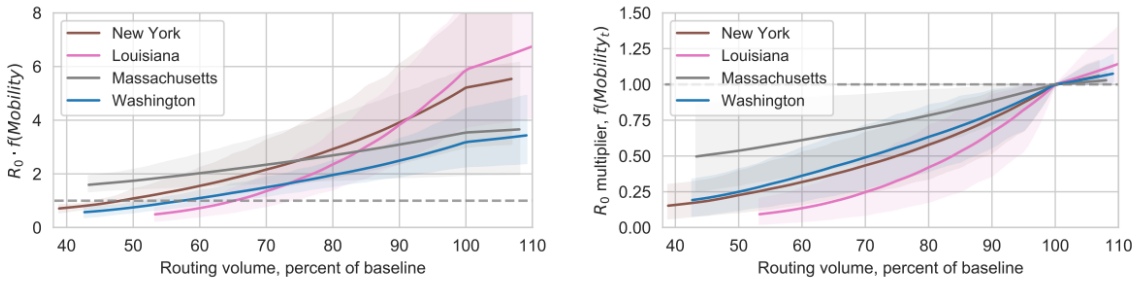

Figure 1. Miller *et al*. (2020). Left: Inferred relationship between the reproduction number and mobility volume change, for four US states. Right: the multiplier effect on initial reproductive number estimates as a function of relative change in mobility from baseline for the same four US states.

As QDL stress, it is important not to ignore cautionary lessons learned regarding the challenges of causal inference from observational, indeed here ecological, associations. The regression effects reported by MERMAID cannot be interpreted causally due to the role of other factors. And, any such regression estimate, and related uncertainty in inference, are sensitive to model selection in a variety of ways.

Better Data not Better Models?

We teach that the science of epidemiology focuses on three questions: Who? When? Where? The cornerstone of any epidemiological description of an emerging outbreak begins with an understanding of how many individuals have been infected by a pathogen, and who they are with regard to basic demographics and risk factors. Crucial to these goals are valid and direct epidemiological measures of community infection and prevalence rates. Routine surveillance allows for targeting of intervention responses and effective mobilization of health care resources. This information is best obtained through adaptable and integrated disease surveillance systems that can capture both new and past infections. There is a lack of national and state hospitalization data due, in part, to the fragmented nature of the US health care system. In the UK, the Office of National Statistics (ONS) has been more proactive with regard to reporting infection and death counts based on probability samples (UK Office for National Statistics, 2021). In addition, the UK REACT studies—also based on probability samples—have provided regular updates on infection rates and community prevalence (Imperial College, London, 2021).

MERMAID is forced to rely on convenience samples which should only be used as a last resort. Unfortunately, the use of population samples has been the exception rather than the norm in the US. Early opportunities to launch seroprevalence surveys in the US

were missed, in part due to a lack of supply of test kits early in the epidemic. A natural option for capturing blood samples from a nationally representative sample would have been the annual National Health and Nutrition Examination Survey (NHANES), but this was suspended on March 6, 2020 due to COVID-19.

In addition, the US has been slow to institute regular and comprehensive sequencing of SARS-Cov-2 positive samples to detect variants. In the UK, the new and fast-spreading SARS-CoV-2 variant (B.1.1.7) was only identified quickly because of the implementation of regular, systematic sequencing of a large sample of positive SARS-CoV-2 tests. There was no such timely systematic attempt in the US, although more sequence data is being obtained now due to the appearance of several additional variants of concern across the world. Further, coordinated contact tracing data — which is extremely useful in assessing transmissibility and factors that affect transmission — has been generally lacking.

It is notable that this was not the case in earlier pandemics when resources, technology and understanding of survey methodology were much less advanced. In the winter of 1918/1919, the US Public Health Service carried out a large door-to-door survey (with a sample size that exceeded 145,000) to measure the morbidity and mortality of the 1918/1919 influenza pandemic (Morabia, 2020).

So, with all due respect to the authors' extraordinary efforts in developing and applying MERMAID, I end my remarks by suggesting that, in the future, it will be better data that we need rather than better models. A national strategy for addressing surveillance data requirements for the next novel pathogen is urgently required.

REFERENCES

Ellenberg, S. S. and Morris, J. S. (2021) AIDS and CCOVID: A tale of two pandemics and the role of statisticians. *Stat. Med.*, 40(11), 2499-2510.

Fretheim, A., Flatø, M, Steens, A., *et al*. (2020) COVID-19: we need randomised trials of school closures. *J. Epidem. Comm. Health*, 74:1078–1079.

Gostic, K. M., McGough, L., Baskerville, E. B., Abbott, S., Joshi, K., Tedijanto, C., *et al.* (2020) Practical considerations for measuring the effective reproductive number, R_t. *PLoS Comput. Biol.* 16(12): e1008409. https://doi.org/10.1371/journal. pcbi.1008409

Imperial College, London. Accessed August 23, 2021. https://www.imperial.ac.uk/medicine/research-and-impact/groups/react-study/

Jewell, B. L. and Jewell, N. P. (2021) On the role of statisticians and modelers in responding to AIDS and COVID-19. *Stat. Med.*, 40(11), 2530-2535.

Jewell, N. P., Lewnard, J. A., and Jewell, B. L. (2020a) Predictive mathematical models of the COVID-19 pandemic. *J. Amer. Med. Assoc.*, 369 doi:10.1001/jama.2020.6585.

Jewell, N. P., Lewnard, J. A., and Jewell, B. L. (2020b) Caution warranted: Using the Institute for Health Metrics and Evaluation model for predicting the course of the COVID-19 pandemic. *Ann. Int. Med.* 2020;173:226-227. doi:10.7326/M20-1565

Larsen, D. A. and Wigginton, K. R. (2020) Tracking COVID-19 with wastewater. *Nature Biotech.,* 38:1151-1153.

McCartney, M. We need better evidence on non-drug interventions for covid-19. (2020) *BMJ*, 370:m3473

Miller, A. C., Foti, N. J., Lewnard, J. A., Jewell, N. P., Guestrin, C., and Fox, E. B. (2020) Mobility trends provide a leading indicator of changes in SARS-CoV-2 transmission. *medRxiv*: DOI:10.1101/2020.05.07.20094441.

Nouvellet *et al.*, Reduction in mobility and COVID-19 transmission. (2021) *Nature Communications*, 12, 1090. https://doi.org/10.1038/s41467-021-21358-2

Parag, K. and Donnelly, C. (2020) Using information theory to optimise epidemic models for real-time prediction and estimation. *PLOS Comput. Biol.*, 16, e1007990.

Parag, K., Thompson, R. N., and Donnelly, C. (2021) Are epidemic growth rates more informative than reproduction numbers? *JRSS*, to appear

Quick, C., Dey, R., and Lin, X. (2021) Regression models for understanding COVID-19 epidemic dynamics with incomplete data. *JASA*, in press

Roberts, S. (2020) The exponential power of now. *New York Times*, March 13, 2020.

Smith, J. A. E., Hopkins, S., Turner, C., Dack, K., Trelfa, A., Peh, J., and Monks, P. S. (2021) Public health impact of mass sporting and cultural events in a rising COVID-19

prevalence in England.
https://khub.net/documents/135939561/338928724/Public+health+impact+of+mass+sporting+and+cultural+events+in+a+rising+COVID-19+prevalence+in+England.pdf/05204895-1576-1ee7-b41e-880d5d6b4f17

UK Office for National Statistics. Accessed August 23, 2021.
https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/datalist