Estimating the COVID-19 *R* number: a bargain with the devil?

Chris T. Bauch
Department of Applied Mathematics
University of Waterloo
cbauch@uwaterloo.ca

The deeper understanding Faust sought
Could not from the Devil be bought.
But now we are told
By theorists bold
All we need know is R_0.[1]

-- Robert May, 1936-2020

Bob May's limerick alludes to both the promises and dangers of characterizing epidemic control by a single number. The basic reproduction number R_0 is the average number of infections produced by a single infectious person in a population with no immunity. R_0 has a close relative named the effective reproduction number *R*: the average number of infections produced by a single infected person in a population with partial immunity. In *The Lancet Infectious Diseases*[2], Li and colleagues estimate how the imposition and lifting of non-pharmaceutical interventions (NPIs) changed the *R* number for SARS-CoV-2 in 131 countries in the first half of 2020.

If $R<1$, an epidemic eventually dies out because each infected person generates less than one new infection. (This may take a long time if there are currently many infections, like the proverbial small rudder on a big ship.) However, when $R>1$, the epidemic may continue growing. *R* can also change over time: NPIs such as closing schools, physical distancing, and mask use can reduce *R*. Hence *R* is often used to gauge whether pandemic mitigation is working.

Li and colleagues compared daily estimates of *R* at the country level against a database describing which NPIs each country applied, and when. They found that imposing NPIs significantly reduced *R*, but lifting them later on increased *R*. School closure, a public events ban, and internal movement limits--both when being imposed and when lifted--had the biggest individual effect, changing *R* between 7% and 25%.

NPIs in combination are even more effective. The combined effect of school and workplace closure, a ban on public events and gatherings of 10 or more persons, internal movement limits, and a stay at home requirement reduced *R* by a whopping 52% (CI 29%-68%). The R_0 for SARS-CoV-2 lies somewhere between 2 and 3[3]. Hence, early pandemic interventions must reduce *R* by between 50% and 67% in order to bring it below one. The authors' estimate does not include the effects of contact tracing and isolation. Despite this, the estimate suggests that it might have been exceedingly difficult to flatten the curve in Spring 2020, had the R_0 for SARS-CoV-2 been a little higher.

But *R* is not without shortcomings. Just as our body mass index does not tell us everything about our state of health, a single number cannot provide a complete picture of the state of a pandemic. National-level estimates can hide local heterogeneity. Seasonal differences in contact patterns from Spring to Autumn are not captured by the short time windows used in many epidemiological studies. Reporting delays, stochastic effects and superspreading[4] can also bias *R*. Moreover, *R* does not tell us what proportion of infections are caused by an infected individual before symptom onset. This crucial distinction for infection control may explain why SARS-CoV-1 did not cause a pandemic while SARS-CoV-2 did, despite their comparable R_0 values[5-7].

Li and colleagues discuss some of these limitations and also raise the issue of "behavioural inertia". Timelines of decision-making lend the perception that governments can turn NPIs on and off like a switch. But in fact, populations can take weeks to adjust their mobility patterns in response to imposition of NPIs[2,8]. This effect probably contributes to the authors' finding that NPIs did not exhibit their maximal effect on *R* until up to 4 weeks later.

R promises crystal clarity in a time when there are no crystal balls. Hence, the allusion to R_0 as a bargain with the devil. Statistician George Box has been widely paraphrased as writing "All models are wrong, but some are useful"[9]. I like to re-paraphrase this as: some models are useful precisely because they are wrong. A model including all the real-world details of a study system would no longer be a model, because it would be the system itself.

Despite *R*'s imperfections, the findings of Li and colleagues tell us that NPIs work, and they tell us which ones work best. This is crucial, given that some NPIs have massive socio-economic impacts. In a similar vein, transmission models that project COVID-19 cases and deaths under different NPI scenarios could be highly valuable for optimizing a country's portfolio of NPIs[10-13]. Moreover, I think *R* provides a social utility that epidemiologists easily overlook. The success of large-scale NPIs requires population adherence. *R* can stimulate populations to act, and gives them useful feedback on the fruits of their labour. Perhaps this is one reason that *R* has entered our vernacular in 2020.

References

1. Fine PEM, Hassell MP, Levin BR, et al. *Population Biology of Infectious Diseases: Report of the Dahlem Workshop on Population Biology of Infectious Disease Agents Berlin 1982, March 14–19*. Vol. 25. Springer Science & Business Media, 2012.
2. Li Y, Campbell H, Kulkarni D, et al. The temporal association of introducing and lifting non-pharmaceutical interventions with the time-varying reproduction number (R) of SARS-CoV-2: a modelling study across 131 countries.
3. Hilton J, Keeling MJ. Estimation of country-level basic reproductive ratios for novel Coronavirus (SARS-CoV-2/COVID-19) using synthetic contact matrices. *PLoS computational biology* 2020; 16: e1008031.

4. Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. Superspreading and the effect of individual variation on disease emergence. *Nature* 2005; 438: 355-359.
5. Bauch CT, Lloyd-Smith JO, Coffee MP, Galvani AP. Dynamically modeling SARS and other newly emerging respiratory illnesses: past, present, and future. *Epidemiology* 2005; 16: 791-801.
6. Fraser C, Riley S, Anderson RM, Ferguson NM. Factors that make an infectious disease outbreak controllable. *Proceedings of the National Academy of Sciences* 2004; 101: 6146-6151.
7. Moghadas SM, Fitzpatrick MC, Sah P, et al. The implications of silent transmission for the control of COVID-19 outbreaks. *Proceedings of the National Academy of Sciences* 2020; 117: 17513-17515.
8. COVID-19 Community Mobility Reports. https://www.google.com/covid19/mobility/. Accessed 5 September 2020.
9. Box GEP. Science and statistics. *Journal of the American Statistical Association* 1976; 71: 791-799.
10. Karatayev VA, Anand M, Bauch CT. Local lockdowns outperform global lockdown on the far side of the COVID-19 epidemic curve. *Proceedings of the National Academy of Sciences* 2020; 117: 24575-24580.
11. Tuite AR., Fisman DN, Greer AL. Mathematical modelling of COVID-19 transmission and mitigation strategies in the population of Ontario, Canada. *CMAJ* 2020; 192: E497-E505.
12. Chinazzi M, Davis JT, Ajelli M, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* 2020; 368: 395-400.
13. Firth JA., Hellewell J, Klepac P, et al. Using a real-world network to model localized COVID-19 control strategies. *Nature medicine* 2020: 1-7. https://doi.org/10.1038/s41591-020-1036-8

~~*~*~*~*~*

Old material:

Are their projected impacts of NPIs conservative on account of delays, saturation?

R has entered the vernacular during the COVID-19 pandemic in much the same way as other mathematical terms have been popularized in recent years, such as “tipping” points”. (And in fact, R=1 is also a tipping point).

If imposition of NPIs is also a reaction to population movement that begins before the NPIs become available, then impact of NPIs is under-estimate, if only because the NPIs are a signifier of undercurrents in the broader population that allow--or demand--application of NPIs.

Superspreading events can not only bias *R* estimates. More

On the network analog of a population where superspreading is possible (power law network), there are no more thresholds.[5]

Second reason, pandemic are social-epidemiological phenomena. Coupled behaviour-diseaese models. The idea that these interventions are being put in place absent of changes in transmission are not considered. For example, if there is an increase in transmission then some authorities may choose to reimplement interventions. This would clearly impact the utility and interpretation of these results. Second reason the estimates are conservative.

For instance, due to stochasticity, the sizes of outbreaks tend to vary on a continuum in the vicinity of R=1 and many outbreaks can die out due to stochastic effects even when R<1 [4] For example, heterogeneity--superspreading.