

R_t : a decent summary statistic

Justin Lessler and C. Jessica E. Metcalf

The effective reproductive number, R_t , is an expression of how many infections are caused by each prevalent infection in the subsequent generation of transmission. It tells us many things, whether the epidemic is growing or declining, what proportion of infections would need to be stopped to “turn the curve” of the epidemic, and, if R_0 is known, the current impact of interventions and immunity. Hence, as a headline indicator, R_t has value for public health policy. However, it has limits, as it smooths over critical impacts of stochasticity, social structure, and the time distribution of infections. Here we discuss perspectives on R_t that highlight challenges in estimation, interpretation, and the search for an equally concise alternative.

Heterogeneity in contacts can profoundly affect epidemic trajectories, leading to significant deviations from standard epidemic theory (Bansal, Grenfell, and Meyers 2007; Anderson and May 1992). Coffeng and de Vlas use Individual Based Models to explore the implications of this idea, focussing on the challenge of adjusting R_0 (and thus R_t) appropriately - early infection of individuals with high contact rates biases naive estimates of transmission upwards; and thus overestimates both the peak of the outbreak and the impact of interventions (Gomes et al. 2020; Britton, Ball, and Trapman 2020).

The challenge comes in practically accounting for this phenomena in ongoing epidemics. As noted by Jewell and Lewnard, the potential impact of heterogeneity drove debate in mid- to late-2020 about the number of infections that we would need to force $R_t < 1$ (often referred to as the “herd immunity threshold”), with some claiming that this level could be as low as 10-20% (Gomes et al. 2020). Later events have shown that this was not the case, exposing the inherent difficulty in accounting for heterogeneity as it happens: we don’t, in fact, know the structure of the relevant contact networks, nor how they are changing over time.

This is one of many factors that Jewell and Lewnard consider as they ask whether the challenges in R_t ’s estimation and interpretation make it of little use in the public health response. They point out that R_t estimates are necessarily delayed as they require mapping past infections to their offspring, and that for many reasons (including those highlighted by Coffeng and de Vlas) its relationship with fundamental drivers of transmission and their observable correlates, such as mobility, susceptibility and underlying incidence is uncertain. Jewell and Lewnard argue that R_t based public health decisions may be fundamentally misguided if they do not take into account other factors. For instance a high R_t is meaningless if no cases are present, and an R_t less than 1 is no reason to roll back control measures if prevalence is high. Accordingly, they suggest a greater emphasis on measures more directly aimed at estimating infection prevalence.

We certainly agree such information is a powerful supplement to more readily available surveillance data. However, measures of the rate of change remain important. If infections are low but increasing fast, drastic measures may be needed. Even if infections are not present, an estimate of how quickly disease would spread if introduced can play a critical role in preparedness. Jewell and Lewnard don't disagree, but argue that the rate of change in infections (or correlates) tells you most of what you need to know, and there may be limited benefit to turning these into real-time estimates of R_t .

Alongside R_t , the field of infectious disease dynamics has long considered instantaneous measures of growth, termed r_t , a fundamental observed quantity from which R_t is derived (Wallinga and Lipsitch 2007). Parag, Thompson and Donnelly ask whether we might be better served by focusing on r_t directly as a concrete measure that includes the times-scale of growth, capturing in a single number what R_t must be combined with the generation time to reflect. They find, however, that the data-smoothing assumptions necessary to estimate r_t are similar to assumptions about the distribution of generation times necessary to estimate R_t . Hence, the inferential challenges in the two tasks are essentially equivalent, and suffer equally, for example, from the bias-variance challenge. Each may have advantages from a communication perspective, with R_t providing more information on the scale of intervention required, while r_t gives a more direct estimate of the speed of outbreak growth; but neither avoids the challenges raised by the other contributors.

These papers highlight that R_t is not a magical quantity that resolves all conundra, but is rather a semi-mechanistic summary statistic that captures the outcome of many underlying processes in a digestible format. Like any summary statistic, its power is only fully realized when used with other statistics or compared across groups (much as a mean is of limited use without a measure of dispersion), and it must be estimated on a scale commensurate with the processes being characterized. It seems unlikely that another statistic summarizing the same information would not share R_t 's shortcomings, as these derive from the nature of the processes being characterized. While there is ample room for innovation, we are confident that R_t will remain a fundamental quantity of interest in future epidemics.

References

- Anderson, Roy M., and Robert M. May. 1992. *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press.
- Bansal, Shweta, Bryan T. Grenfell, and Lauren Ancel Meyers. 2007. "When Individual Behaviour Matters: Homogeneous and Network Models in Epidemiology." *Journal of the Royal Society, Interface / the Royal Society* 4 (16): 879–91.
- Britton, Tom, Frank Ball, and Pieter Trapman. 2020. "A Mathematical Model Reveals the Influence of Population Heterogeneity on Herd Immunity to SARS-CoV-2." *Science* 369 (6505): 846–49.
- Gomes, M. Gabriela M., Rodrigo M. Corder, Jessica G. King, Kate E. Langwig, Caetano Souto-Maior, Jorge Carneiro, Guilherme Goncalves, Carlos Penha-Goncalves, Marcelo U.

Ferreira, and Ricardo Aguas. 2020. "Individual Variation in Susceptibility or Exposure to SARS-CoV-2 Lowers the Herd Immunity Threshold." *medRxiv : The Preprint Server for Health Sciences*, May. <https://doi.org/10.1101/2020.04.27.20081893>.

Wallinga, J., and M. Lipsitch. 2007. "How Generation Intervals Shape the Relationship between Growth Rates and Reproductive Numbers." *Proceedings. Biological Sciences / The Royal Society* 274 (1609): 599–604.

-----GRAVEYARD-----

R is the worst epidemic descriptor, except for all the others.

Perhaps a more interesting question is not when R_t is wrong, but is it even telling us what we want to know when it is right. Jewell and Lewnard examine some of the intricacies of estimating R_t , and, indeed, approximating changes in it via changes in e.g., mobility (often highly context specific), or susceptibility (with many remaining questions around heterogeneity in immunity from different vaccines and in different populations). They argue that policy application of this quantity would benefit from greater scrutiny. For example, is lifting restrictions when R_t falls below a threshold like 1 is sensible, given that this neglects current prevalence, and is also estimated with an inevitable delay, given the time required for cases or deaths to enter registration systems? They suggest that a focus on changes in R_t might be more informative in shaping policy.

Finally, Parag, Thompson and Donnelly ask whether more concise alternatives offer any benefit over R_t ? In particular, they examine the epidemic growth rate rt , defined as the rate of change of log transformed case incidence. This quantity thus

(innovation around data: <https://www.medrxiv.org/content/10.1101/2021.04.29.21255961v1> ? or actual innovation in framing?)