

# Complexity of the Basic Reproduction Number ( $R_0$ )

Paul L. Delamater, Erica J. Street, Timothy F. Leslie, Y. Tony Yang, Kathryn H. Jacobsen

The basic reproduction number ( $R_0$ ), also called the basic reproduction ratio or rate or the basic reproductive rate, is an epidemiologic metric used to describe the contagiousness or transmissibility of infectious agents.  $R_0$  is affected by numerous biological, sociobehavioral, and environmental factors that govern pathogen transmission and, therefore, is usually estimated with various types of complex mathematical models, which make  $R_0$  easily misrepresented, misinterpreted, and misapplied.  $R_0$  is not a biological constant for a pathogen, a rate over time, or a measure of disease severity, and  $R_0$  cannot be modified through vaccination campaigns.  $R_0$  is rarely measured directly, and modeled  $R_0$  values are dependent on model structures and assumptions. Some  $R_0$  values reported in the scientific literature are likely obsolete.  $R_0$  must be estimated, reported, and applied with great caution because this basic metric is far from simple.

The basic reproduction number ( $R_0$ ), pronounced “R naught,” is intended to be an indicator of the contagiousness or transmissibility of infectious and parasitic agents.  $R_0$  is often encountered in the epidemiology and public health literature and can also be found in the popular press (1–6).  $R_0$  has been described as being one of the fundamental and most often used metrics for the study of infectious disease dynamics (7–12). An  $R_0$  for an infectious disease event is generally reported as a single numeric value or low–high range, and the interpretation is typically presented as straightforward; an outbreak is expected to continue if  $R_0$  has a value  $>1$  and to end if  $R_0$  is  $<1$  (13). The potential size of an outbreak or epidemic often is based on the magnitude of the  $R_0$  value for that event (10), and  $R_0$  can be used to estimate the proportion of the population that must be vaccinated to eliminate an infection from that population (14,15).  $R_0$  values have been published for measles, polio, influenza, Ebola virus disease, HIV disease, a diversity of vectorborne infectious diseases, and many other communicable diseases (14,16–18).

Affiliations: University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA (P.L. Delamater); George Mason University, Fairfax, Virginia, USA (E.J. Street, T.F. Leslie, K.H. Jacobsen); George Washington University, Washington, DC, USA (Y.T. Yang)

DOI: <https://doi.org/10.3201/eid2501.171901>

The concept of  $R_0$  was first introduced in the field of demography (9), where this metric was used to count offspring. When  $R_0$  was adopted for use by epidemiologists, the objects being counted were infective cases (19). Numerous definitions for  $R_0$  have been proposed. Although the basic conceptual framework is similar for each, the operational definitions are not always identical. Dietz states that  $R_0$  is “the number of secondary cases one case would produce in a completely susceptible population” (19). Fine supplements this definition with the description “average number of secondary cases” (17). Diekmann and colleagues use the description “expected number of secondary cases” and provide additional specificity to the terminology regarding a single case (13).

In the hands of experts,  $R_0$  can be a valuable concept. However, the process of defining, calculating, interpreting, and applying  $R_0$  is far from straightforward. The simplicity of an  $R_0$  value and its corresponding interpretation in relation to infectious disease dynamics masks the complicated nature of this metric. Although  $R_0$  is a biological reality, this value is usually estimated with complex mathematical models developed using various sets of assumptions. The interpretation of  $R_0$  estimates derived from different models requires an understanding of the models’ structures, inputs, and interactions. Because many researchers using  $R_0$  have not been trained in sophisticated mathematical techniques,  $R_0$  is easily subject to misrepresentation, misinterpretation, and misapplication. Notable examples include incorrectly defining  $R_0$  (1) and misinterpreting the effects of vaccination on  $R_0$  (3). Further, many past lessons regarding this metric appear to have been lost or overlooked over time. Therefore, a review of the concept of  $R_0$  is needed, given the increased attention this metric receives in the academic literature (20). In this article, we address misconceptions about  $R_0$  that have proliferated as this metric has become more frequently used outside of the realm of mathematical biology and theoretic epidemiology, and we recommend that  $R_0$  be applied and discussed with caution.

## Variations in $R_0$

For any given infectious agent, the scientific literature might present numerous different  $R_0$  values. Estimations of the  $R_0$  value are often calculated as a function of 3 primary parameters—the duration of contagiousness after a person becomes infected, the likelihood of infection per contact between a susceptible person and an infectious person or

vector, and the contact rate—along with additional parameters that can be added to describe more complex cycles of transmission (19). Further, the epidemiologic triad (agent, host, and environmental factors) sometimes provides inspiration for adding parameters related to the availability of public health resources, the policy environment, various aspects of the built environment, and other factors that influence transmission dynamics and, thus, are relevant for the estimation of  $R_0$  values (21). Yet, even if the infectiousness of a pathogen (that is, the likelihood of infection occurring after an effective contact event has occurred) and the duration of contagiousness are biological constants,  $R_0$  will fluctuate if the rate of human–human or human–vector interactions varies over time or space. Limited evidence supports the applicability of  $R_0$  outside the region where the value was calculated (20). Any factor having the potential to influence the contact rate, including population density (e.g., rural vs. urban), social organization (e.g., integrated vs. segregated), and seasonality (e.g., wet vs. rainy season for vectorborne infections), will ultimately affect  $R_0$ . Because  $R_0$  is a function of the effective contact rate, the value of  $R_0$  is a function of human social behavior and organization, as well as the innate biological characteristics of particular pathogens. More than 20 different  $R_0$  values (range 5.4–18) were reported for measles in a variety of study areas and periods (22), and a review in 2017 identified feasible measles  $R_0$  values of 3.7–203.3 (23). This wide range highlights the potential variability in the value of  $R_0$  for an infectious disease event on the basis of local sociobehavioral and environmental circumstances.

### Various Names for $R_0$

Inconsistency in the name and definition of  $R_0$  has potentially been a cause for misunderstanding the meaning of  $R_0$ .  $R_0$  was originally called the basic case reproduction rate when George MacDonald introduced the concept into the epidemiology literature in the 1950s (17,19,24,25). Although MacDonald used  $Z_0$  to represent the metric, the current symbolic representation ( $R_0$ ) appears to have remained largely consistent since that time. However, multiple variations of the name for the concept expressed by  $R_0$  have been used in the scientific literature, including the use of basic and case as the first word in the term, reproduction and reproductive for the second word, and number, ratio, and rate for the final part of the term (13). Although the frequent use of the term basic reproduction rate is in line with MacDonald's original terminology (9), some users interpret the use of the word rate as suggesting a quantity having a unit with a per-time dimension (7). If  $R_0$  were a rate involving time, the metric would provide information about how quickly an epidemic will spread through a population. But  $R_0$  does not indicate whether new cases will occur within 24 hours after the initial case or months later, just as  $R_0$  does not indicate whether the

disease produced by the infection is severe. Instead,  $R_0$  is most accurately described in terms of cases per case (7,13). Calling  $R_0$  a rate rather than a number or ratio might create some undue confusion about what the value represents.

### $R_0$ and Vaccination Campaigns

Vaccination campaigns reduce the proportion of a population at risk for infection and have proven to be highly effective in mitigating future outbreaks (26). This conclusion is sometimes used to suggest that an aim of vaccination campaigns is to remove susceptible members of the population to reduce the  $R_0$  for the event to  $<1$ . Although the removal of susceptible members from the population will affect infection transmission by reducing the number of effective contacts between infectious and susceptible persons, this activity will technically not reduce the  $R_0$  value because the definition of  $R_0$  includes the assumption of a completely susceptible population. When examining the effect of vaccination, the more appropriate metric to use is the effective reproduction number ( $R$ ), which is similar to  $R_0$  but does not assume complete susceptibility of the population and, therefore, can be estimated with populations having immune members (16,20,27). Efforts aimed at reducing the number of susceptible persons within a population through vaccination would result in a reduction of the  $R$  value, rather than  $R_0$  value. In this scenario, vaccination could potentially end an epidemic, if  $R$  can be reduced to a value  $<1$  (16,27,28). The effective reproduction number can also be specified at a particular time  $t$ , presented as  $R(t)$  or  $R_t$ , which can be used to trace changes in  $R$  as the number of susceptible members in a population is reduced (29,30). When the goal is to measure the effectiveness of vaccination campaigns or other public health interventions,  $R_0$  is not necessarily the best metric (10,20).

### Measuring and Estimating $R_0$

Counting the number of cases of infection during an epidemic can be extremely difficult, even when public health officials use active surveillance and contact tracing to attempt to locate all infected persons. Although measuring the true  $R_0$  value is possible during an outbreak of a newly emerging infectious pathogen that is spreading through a wholly susceptible population, rarely are there sufficient data collection systems in place to capture the early stages of an outbreak when  $R_0$  might be measured most accurately. As a result,  $R_0$  is nearly always estimated retrospectively from seroepidemiologic data or by using theoretical mathematical models (31). Data-driven approaches include the use of the number of susceptible persons at endemic equilibrium, average age at infection, final size equation, and intrinsic growth rate (10). When mathematical models are used,  $R_0$  values are often estimated by using ordinary differential equations (8–10,19,31), but high-quality data

are rarely available for all components of the model. The estimated values of  $R_0$  generated by mathematical models are dependent on numerous decisions made by the modeler (8,32,33). The population structure of the model, such as the susceptible-infectious-recovered model or susceptible-exposed-infectious-recovered model, which includes compartments for persons who are exposed but not yet infectious, as well as assumptions about demographic dynamics (e.g., births, deaths, and migration over time), are critical model parameters. Population mixing and contact patterns must also be considered; for example, for homogeneous mixing, all population members are equally likely to come into contact with one another, and for heterogeneous mixing, variation in contact patterns are present among age subgroups or geographic regions. Other decisions include whether to use a deterministic (yielding the same outcomes each time the model is run) or stochastic (generating a distribution of likely outcomes on the basis of variations in the inputs) approach and which distributions (e.g., Gaussian or uniform distributions) to use to describe the probable values of parameters, such as effective contact rates and duration of contagiousness. Furthermore, many of the parameters included in the models used to estimate  $R_0$  are merely educated guesses; the true values are often unknown or difficult or impossible to measure directly (31,34,35). This limitation is compounded as models become more complex and, thus, require more input parameters (20,35), such as when using models to estimate the value of  $R_0$  for infectious pathogens with more complex transmission pathways, which can include vectorborne infectious agents or those with environmental or wildlife reservoirs. In summary, although only 1 true  $R_0$  value exists for an infectious disease event occurring in a particular place at a particular time, models that have minor differences in structure and assumptions might produce different estimates of that value, even when using the same epidemiologic data as inputs (20,31,32,36,37).

### Obsolete $R_0$ Values

New estimates of  $R_0$  have been produced for infectious disease events that occurred in recent history, such as the West Africa Ebola outbreak (34,38,39). However, for many vaccine-preventable diseases, the scientific literature reports  $R_0$  values calculated much further back in history. For example, the oft-reported measles  $R_0$  values of 12–18 are based on data acquired during 1912–1928 in the United States ( $R_0$  of 12.5) and 1944–1979 in England and Wales ( $R_0$  of 13.7–18.0) (14), even though more recent estimates of the  $R_0$  for measles highlight a much greater numeric range and variation across settings (23). For pertussis ( $R_0$  of 12–17), the original data sources are 1908–1917 in the United States ( $R_0$  of 12.2) and 1944–1979 in England and Wales ( $R_0$  of 14.3–17.1) (14). The major changes that

have occurred in how humans organize themselves both socially and geographically make these historic values extremely unlikely to match present day epidemiologic realities. Behavioral changes undoubtedly have altered contact rates, which are a key component of  $R_0$  calculations. Yet, these  $R_0$  values have been repeated so often in the literature that newer  $R_0$  values generated by using modern data might be dismissed if they fall outside the range of previous estimates. Given that  $R_0$  is often considered when designing and implementing vaccination strategies and other public health interventions, the use of  $R_0$  values derived from older data is likely inappropriate (23). Decisions about public health practice should be made with contemporaneous  $R_0$  values or  $R$  values instead.

### Conclusions

Although  $R_0$  might appear to be a simple measure that can be used to determine infectious disease transmission dynamics and the threats that new outbreaks pose to the public health, the definition, calculation, and interpretation of  $R_0$  are anything but simple.  $R_0$  remains a valuable epidemiologic concept, but the expanded use of  $R_0$  in both the scientific literature and the popular press appears to have enabled some misunderstandings to propagate.  $R_0$  is an estimate of contagiousness that is a function of human behavior and biological characteristics of pathogens.  $R_0$  is not a measure of the severity of an infectious disease or the rapidity of a pathogen's spread through a population.  $R_0$  values are nearly always estimated from mathematical models, and the estimated values are dependent on numerous decisions made in the modeling process. The contagiousness of different historic, emerging, and reemerging infectious agents cannot be fairly compared without recalculating  $R_0$  with the same modeling assumptions. Some of the  $R_0$  values commonly reported in the literature for past epidemics might not be valid for outbreaks of the same infectious disease today.

$R_0$  can be misrepresented, misinterpreted, and misapplied in a variety of ways that distort the metric's true meaning and value. Because of these various sources of confusion,  $R_0$  must be applied and discussed with caution in research and practice. This epidemiologic construct will only remain valuable and relevant when used and interpreted correctly.

This research was supported through a grant from the George Mason University Provost Multidisciplinary Research Initiative.

### About the Author

Dr. Delamater is an assistant professor in the Department of Geography and a faculty fellow at the Carolina Population Center at the University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA. His research focuses on the geographic aspects of health, disease, and healthcare.

## References

1. Johnson R. The 10 epidemics that almost wiped out mankind. *Business Insider*. 2011 Sep 19 [cited 2017 Nov 17]. <http://www.businessinsider.com/epidemics-pandemics-that-almost-wiped-out-mankind-2011-9>
2. Coy P. A primer on the deadly math of Ebola. *Bloomberg Businessweek*. 2014 Sep 26 [cited 2017 Nov 17]. <http://www.bloomberg.com/news/articles/2014-09-26/ebolas-deadly-math>
3. Doucleff M. No, seriously, how contagious is Ebola? *NPR*. 2014 Oct 2 [cited 2017 Nov 17]. <http://www.npr.org/sections/health-shots/2014/10/02/352983774/no-seriously-how-contagious-is-ebola>
4. Freeman C. Magic formula that will determine whether Ebola is beaten. *Telegraph*. 2014 Nov 6 [cited 2017 Nov 17]. <http://www.telegraph.co.uk/news/worldnews/ebola/11213280/Magic-formula-that-will-determine-whether-Ebola-is-beaten.html>
5. McKenna M. The mathematics of Ebola trigger stark warnings: act now or regret it. *WIRED*. 2014 Sep 14 [cited 2017 Nov 17]. <https://www.wired.com/2014/09/r0-ebola/>
6. Fox M. Ebola epidemic was driven by just a few infected people, study finds. *NBC News*. 2017 Feb 13 [cited 2017 Nov 17]. <https://www.nbcnews.com/storyline/ebola-virus-outbreak/superspreaders-drove-ebola-epidemic-study-finds-n720321>
7. Heesterbeek JAP, Dietz K. The concept of  $R_0$  in epidemic theory. *Stat Neerl*. 1996;50:89–110. <http://dx.doi.org/10.1111/j.1467-9574.1996.tb01482.x>
8. Keeling MJ, Grenfell BT. Individual-based perspectives on  $R_0$ . *J Theor Biol*. 2000;203:51–61. <http://dx.doi.org/10.1006/jtbi.1999.1064>
9. Heesterbeek JAP. A brief history of  $R_0$  and a recipe for its calculation. *Acta Biotheor*. 2002;50:189–204. <http://dx.doi.org/10.1023/A:1016599411804>
10. Heffernan JM, Smith RJ, Wahl LM. Perspectives on the basic reproductive ratio. *J R Soc Interface*. 2005;2:281–93. <http://dx.doi.org/10.1098/rsif.2005.0042>
11. Roberts MG. The pluses and minuses of  $R_0$ . *J R Soc Interface*. 2007;4:949–61. <http://dx.doi.org/10.1098/rsif.2007.1031>
12. Pellis L, Ball F, Trapman P. Reproduction numbers for epidemic models with households and other social structures. I. Definition and calculation of  $R_0$ . *Math Biosci*. 2012;235:85–97. <http://dx.doi.org/10.1016/j.mbs.2011.10.009>
13. Diekmann O, Heesterbeek JAP, Metz JAJ. On the definition and the computation of the basic reproduction ratio  $R_0$  in models for infectious diseases in heterogeneous populations. *J Math Biol*. 1990;28:365–82. <http://dx.doi.org/10.1007/BF00178324>
14. Anderson RM, May RM. Directly transmitted infectious diseases: control by vaccination. *Science*. 1982;215:1053–60. <http://dx.doi.org/10.1126/science.7063839>
15. Anderson RM, May RM. Vaccination and herd immunity to infectious diseases. *Nature*. 1985;318:323–9. <http://dx.doi.org/10.1038/318323a0>
16. Anderson RM, May RM. *Infectious diseases of humans: dynamics and control*. Oxford: Oxford University Press; 1991.
17. Fine PEM. Herd immunity: history, theory, practice. *Epidemiol Rev*. 1993;15:265–302. <http://dx.doi.org/10.1093/oxfordjournals.epirev.a036121>
18. Chowell G, Nishiura H. Transmission dynamics and control of Ebola virus disease (EVD): a review. *BMC Med*. 2014;12:196. <http://dx.doi.org/10.1186/s12916-014-0196-0>
19. Dietz K. The estimation of the basic reproduction number for infectious diseases. *Stat Methods Med Res*. 1993;2:23–41. <http://dx.doi.org/10.1177/096228029300200103>
20. Ridenhour B, Kowalik JM, Shay DK. Unraveling  $R_0$ : considerations for public health applications. *Am J Public Health*. 2014;104:e32–41. <http://dx.doi.org/10.2105/AJPH.2013.301704>
21. Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. Superspreading and the effect of individual variation on disease emergence. *Nature*. 2005;438:355–9. <http://dx.doi.org/10.1038/nature04153>
22. Anderson RM. Directly transmitted viral and bacterial infections of man. In: Anderson RM, editor. *The population dynamics of infectious diseases: theory and applications*. London: Chapman and Hall; 1982. p. 1–37.
23. Guerra FM, Bolotin S, Lim G, Heffernan J, Deeks SL, Li Y, et al. The basic reproduction number ( $R_0$ ) of measles: a systematic review. *Lancet Infect Dis*. 2017;17:e420–8. [http://dx.doi.org/10.1016/S1473-3099\(17\)30307-9](http://dx.doi.org/10.1016/S1473-3099(17)30307-9)
24. MacDonald G. The analysis of equilibrium in malaria. *Trop Dis Bull*. 1952;49:813–29.
25. MacDonald G. *The epidemiology and control of malaria*. London: Oxford University Press; 1957.
26. Bärnighausen T, Bloom DE, Cafiero-Fonseca ET, O'Brien JC. Valuing vaccination. *Proc Natl Acad Sci U S A*. 2014;111:12313–9. <http://dx.doi.org/10.1073/pnas.1400475111>
27. Anderson RM. The concept of herd immunity and the design of community-based immunization programmes. *Vaccine*. 1992;10:928–35. [http://dx.doi.org/10.1016/0264-410X\(92\)90327-G](http://dx.doi.org/10.1016/0264-410X(92)90327-G)
28. Plans Rubió P. Is the basic reproductive number ( $R_0$ ) for measles viruses observed in recent outbreaks lower than in the pre-vaccination era? *Euro Surveill*. 2012;17:22. <http://dx.doi.org/10.2807/ese.17.31.20233-en>
29. Nishiura H, Chowell G. The effective reproduction number as a prelude to statistical estimation of time-dependent epidemic trends. In: Chowell G, Hyman JM, Bettencourt LMA, Castillo-Chavez C, editors. *Mathematical and statistical estimation approaches in epidemiology*. Dordrecht (the Netherlands): Springer Netherlands; 2009. p. 103–21.
30. Mercer GN, Glass K, Becker NG. Effective reproduction numbers are commonly overestimated early in a disease outbreak. *Stat Med*. 2011;30:984–94.
31. Li J, Blakeley D, Smith RJ. The failure of  $R_0$ . *Comput Math Methods Med*. 2011;2011:527610. <http://dx.doi.org/10.1155/2011/527610>
32. Breban R, Vardavas R, Blower S. Theory versus data: how to calculate  $R_0$ ? *PLoS One*. 2007;2:e282. <http://dx.doi.org/10.1371/journal.pone.0000282>
33. Artalejo JR, Lopez-Herrero MJ. On the exact measure of disease spread in stochastic epidemic models. *Bull Math Biol*. 2013;75:1031–50. <http://dx.doi.org/10.1007/s11538-013-9836-3>
34. Khan A, Naveed M, Dur-E-Ahmad M, Imran M. Estimating the basic reproductive ratio for the Ebola outbreak in Liberia and Sierra Leone. *Infect Dis Poverty*. 2015;4:13. <http://dx.doi.org/10.1186/s40249-015-0043-3>
35. Roberts M, Heesterbeek H. Bluff your way in epidemic models. *Trends Microbiol*. 1993;1:343–8. [http://dx.doi.org/10.1016/0966-842X\(93\)90075-3](http://dx.doi.org/10.1016/0966-842X(93)90075-3)
36. Bani-Yaghoob M, Gautam R, Shuai Z, van den Driessche P, Ivanek R. Reproduction numbers for infections with free-living pathogens growing in the environment. *J Biol Dyn*. 2012;6:923–40. <http://dx.doi.org/10.1080/17513758.2012.693206>
37. Nishiura H. Correcting the actual reproduction number: a simple method to estimate  $R_0$  from early epidemic growth data. *Int J Environ Res Public Health*. 2010;7:291–302. <http://dx.doi.org/10.3390/ijerph7010291>
38. Legrand J, Grais RF, Boelle PY, Valleron AJ, Flahault A. Understanding the dynamics of Ebola epidemics. *Epidemiol Infect*. 2007;135:610–21. <http://dx.doi.org/10.1017/S0950268806007217>
39. Fisman D, Khoo E, Tuite A. Early epidemic dynamics of the West African 2014 Ebola outbreak: estimates derived with a simple two-parameter model. *PLoS Curr*. 2014;6:ecurrents.outbreaks.89c0d3783f36958d96ebbae97348d571. <https://doi.org/10.1371/currents.outbreaks.89c0d3783f36958d96ebbae97348d571>

Address for correspondence: Paul L. Delamater, University of North Carolina at Chapel Hill, Campus Box 3220, Chapel Hill, NC 27599-3220, USA; email: pld@email.unc.edu