## ABSTRACT

The basic reproduction number, $R_0$, and its real-time analogue, $R_t$, are summary measures that reflect the ability of an infectious disease to spread through a population. Estimation methods for $R_t$ have a long history, have been widely developed, and now enhanced by application to the COVID-19 pandemic. While retrospective analyses of $R_t$ have provided insight into epidemic dynamics and the effects of control strategies in prior outbreaks, misconceptions around the interpretation of $R_t$ have arisen with broader recognition and near real-time monitoring of this parameter alongside reported case data during the COVID-19 pandemic. Here we discuss some widespread misunderstandings regarding the use of $R_t$ as a barometer for population risk and its related use as an "on/off" switch for policy decisions regarding relaxation of non-pharmaceutical interventions. Computation of $R_t$ from downstream data (e.g. hospitalizations) when infection counts are unreliable exacerbates lags between when transmission happens and when events are recorded., We also discuss analyses that have shown various relationships between $R_t$ and measures of mobility, vaccination coverage, and a test-trace-isolation intervention in different settings.

## Section 1. INTRODUCTION

Quantification of local transmission patterns is a necessary tool in assessing spread and determining appropriate intervention strategies during an infectious disease outbreak. During the current COVID-19 pandemic, reproduction number ($R$) estimates have been widely used to summarize transmission dynamics in a community. Generally speaking, reproduction numbers aim to define the expected number of new infections generated from a single infected individual under various conditions or assumptions. The basic reproduction number ($R_0$)—likely the most familiar to scientific and popular audiences alike—is typically taken to indicate the expected number of infections resulting from an infectious individual encountering a fully-susceptible population; this definition is sometimes further extended to represent conditions at the beginning of an outbreak before specifically enacted interventions are in play.

However, various other reproductive number parameters are also of value for describing transmission. In some instances, $R_0$ is defined as a mean of individual-specific reproductive numbers $R_i$ for members of a population, the distribution of which has importance to transmission dynamics and the ease of control but is rarely characterized.[1–3] A value which has gained particular interest is the time-varying effective reproductive number ($R_t$),[4] or instantaneous reproductive number, which tracks changes over time in the number of secondary infections caused by each case owing to buildup of immunity in the population, changes in behavior, and implementation of new interventions.

Real time estimates of $R_t$ have been widely reported and often used to compare regions within countries with regard to progress in stopping or slowing growth in infections. In various settings, such values have also been prioritized as a criterion for the lifting and reintroduction of nonpharmaceutical interventions. Here we briefly review the interpretation and estimation of various reproductive number parameters, and highlight experience with the use of such measures during the COVID-19 pandemic to inform public health decision-making.

## Section 2. REPRODUCTION NUMBERS, THEIR INTERPRETATION, AND FACTORS INFLUENCING THEM

Reproduction numbers depend on several factors, most prominently (i) the intensity of contacts between susceptible and infectious persons in the population, together with mixing patterns, (ii) the infectivity associated with a contact (that is, the probability of transmission per contact between a susceptible and infectious persons), and (iii) the duration of the infectiousness period. In fact, these multiple components underly the varied strategies used to reduce transmission (and thus $R_t$) in a community. For example, (i) shelter-in-place policies reduce contact rates, (ii) vaccination, masks, and avoidance of close or physical contact reduce infectivity, and (iii) contact tracing, quarantining and isolation all potentially reduce the time of exposure between infectious and susceptible individuals.

Thus, $R$ (defined as either $R_0$ or $R_t$) summarizes many social and biological effects in a single measure and thus—while useful—may be subject to overly simplistic interpretations and limited in its value in driving policy decisions. The simplest such decision for which $R$ estimates may provide insight is whether current levels of public health intervention are sufficient to prevent uncontrolled growth of an epidemic, which may be expected under any scenario resulting in $R > 1$. This assessment may be of greatest importance at early stages of an outbreak, especially when the agent involved is a novel pathogen with unknown transmissibility.

Second, quantitative estimates of $R$ provide an indication of the level—and possibly, by extensions, pathways—of intervention required to bring transmission under control by achieving $R < 1$. During the COVID-19 pandemic, analyses have made assessments of the impacts of various non-pharmaceutical interventions aiming to reduce interactions in particular settings, both through analyses of the next-generation matrix[5,6] and empirically, based on ecological studies.[7] Estimates of the proportion of the population that must be protected from infection to achieve $R < 1$, through natural immunity in response to infection or vaccination, are among the most valuable insights of this nature. Issues relating to such “herd immunity thresholds” have become a flashpoint of controversy during the COVID-19 pandemic, initially focused on levels of natural immunity but now including population vaccination coverage targets. Whereas early modeling work demonstrated the potential for severe and protracted outbreaks in the absence of effective vaccination,[8,9] subsequent theoretical studies suggested that heterogeneity in individual risk of acquiring or transmitting infection could lower such thresholds,[10] including to implausibly low values in the range of 10-20%.[11] The validity of these conclusions that transmission could be brought under control with low population immunity have ultimately been brought into question by the persistence of severe epidemics with $R > 1$ even amid high population seroprevalence and aggressive vaccine rollout.[12–14] Related debate has surrounded the interpretation of early $R_0$ estimates in light of the possibility for immunological cross-recognition of SARS-CoV-2 among individuals with recent exposure to endemic coronaviruses.[15] However, it is crucial to note that empirical estimates of reproduction numbers in the setting of SARS-CoV-2-naïve populations implicitly account for the impact of such immunity on transmission dynamics. This definitional ambiguity is a limitation of $R_0$, and underscores the need for authors to provide clear descriptions of the interpretation of reproduction numbers they report.

Among the most closely monitored parameters of SARS-CoV-2 spread has been $R_t$. The need to determine the impact and effectiveness of nonpharmaceutical interventions has resulted in great interest in monitoring changes over time in $R_t$ in various settings. Some jurisdictions, including the United Kingdom, have gone so far as to suggest changes in $R_t$ as a basis for lifting or reimplementation of interventions.[16] However, this approach conveys undue confidence in the validity of real-time $R_t$ estimates, for which reliable public health data collection and quantitative estimation strategies present continuing challenges[17] (see Section 3), and belies a more fundamental misunderstanding in the epidemiologic meaning of this parameter. Inherent in its

definition, $R_t$ conveys information about growth in the number of infected and does not therefore carry information about the current level or prevalence of infection. With that in mind, estimates of $R_t$ are not sufficient as an "on/off" switch regarding relaxation of social distancing and other measures to restrict contacts such as shuttering certain kinds of businesses and events. For example, policy makers may be keenly interested in the impact of indoor restaurant dining depending on epidemiological attributes of what is going on at that time in the community. What will happen if restrictions are lifted to some degree and what degree should that be? Naturally, the most important issue at hand is the prevalence of infection in the community (as that would determine whether going to a restaurant results in a lower or higher risk of exposure), and not $R_t$. The social conditions in place in Taiwan or New Zealand, where infection prevalence is low, could possibly support an $R_t \approx 3$ as these populations are mixing relatively freely; this matters little, however, when there is little or no infection to circulate.

To be fair, the reliance on $R_t$ as a trigger for policy interventions may be a result of overly simplistic interpretation of what is often complex and nuanced decision making. For example, Mahase[16] stated that the UK Prime Minister Boris Johnson "told the public on 10 May [2020] that easing lockdown in England would depend on whether the reproduction number could be kept down." While $R_t$ was emphasized in the transcript of his remarks, he specifically stated that "the Covert Alert Level will be determined primarily by R *and the number of coronavirus cases* (emphasis added), and that "we will be monitoring the R *and the number of new infections and the progress we are making*" (emphasis added).[18]

This raises a second issue regarding over-reliance on $R_t$ as a summary measure, namely that it is necessarily estimated in a delayed fashion by all available methods. We discuss this briefly below and, as noted above, the complexity that transmission depends differently on the distinct factors that influence $R_t$ . The need to implement appropriate policies with considerable speed (to the extent possible), and often before many cases have been detected by surveillance systems, is counter-intuitive to many policymakers and the media, but critical to stop spread during an exponential growth phase. Using a mathematical transmission model, Pei *et al*.[19] estimate that 56% (95% CI: 44-64%) of reported deaths in the United States (US) as of May 3, 2020 could have been avoided had observed control measures been implemented one week earlier, reinforcing the earlier and simpler calculations of Jewell and Jewell[20] that estimated an approximately 60% reduction in US COVID-19 deaths associated with a similar one week advance in mitigation measures in early spring 2020. This significantly limits the value of a necessarily time-delayed estimate of $R_t$ as a trigger for intervention even when supplemented by estimates of population prevalence of active infection. The combination of exponential growth and considerable levels of asymptomatic transmission allowed the rapid growth of SARS-CoV-2 infections before action was taken even on second or subsequent waves of infection; a similar problem occurred with the rapid asymptomatic spread of HIV infections in the 1980s albeit on a quite different time scale.[21]

Given this lag in estimation of $R_t$ and difficulties associated with obtaining data of sufficient quality to support accurate and timely calculations, it is natural to seek precursors of $R_t$ using more available and reliable information, "predictive correlates," or "surrogates," for $R_t$ if you will. We discuss some examples briefly in Section 3, again with cautions about interpretation.

## Section 3. ESTIMATION OF $R_t$

We do not discuss here detailed estimation strategies for reproduction numbers, and the necessary underlying data requirements and assumptions, but refer to recent reviews by Gostic *et al.*[17] and O'Driscoll *et* al.[22] Both analyses consider various statistical approaches to estimation of $R_t$ and discuss comparative performance on simulated epidemics with due attention paid to practical considerations for implementing differing estimation strategies.

It goes without saying that estimates of $R$ critically depend on the nature of available data on infection counts and their sequelae. As is widely acknowledged, a critical issue at the beginning of an epidemic is the paucity of high-quality information that allows precise statistical estimation under any approach. An intuitive approach, direct counting of secondary cases linked to each index infected, is especially prone to undercounting in such circumstances, especially before investigations can be informed by detailed understanding of transmission pathways and the clinical spectrum of infection. Ecological methods for analysis of time series of case numbers are therefore a mainstay of estimation approaches.[4,17,23,24] While incomplete ascertainment will not introduce bias if the proportion of infections ascertained (and the clinical stage at which they are ascertained) remains constant from one generation of infection to the next,[25] this circumstance is unlikely to be met as enhanced clinical awareness, testing effort, and public health surveillance capacity contribute to improvements in ascertainment over time.[26]

Using observed infection numbers may be highly misleading as a basis for estimating $R_t$ when infection testing has been so variable not only over time but also geographically and demographically. Given the inadequacy of surveillance for monitoring population prevalence of infection, it is attractive to exploit data that may be less subject to inaccurate and unreliable reporting, such as deaths or hospitalizations due to COVID-19. Unfortunately, deaths are also subject to inaccuracies with considerable evidence of both underreporting and variable reporting delays;[27] while clinical criteria for hospitalization may change as hospitals approach capacity limits or as clinical management strategies improve in ambulatory and other care settings. Further, the use of COVID-19 deaths or hospitalizations inevitably introduces a longer lag after infection so that "real time estimation" reflects transmission patterns as long as a month previously. Differentiating genuine reductions in transmission intensity from biases similar to censoring, due to the delayed presentation of cases that were recently infected, poses further difficulty for which statistical deconvolution approaches remain

underdeveloped.[28] As discussed below, this challenge limits the value of transmission intensity estimates in informing real-time public health policy changes.

Gostic *et al.*[17] provide a detailed discussion of additional sources of bias in estimation of $R_t$. Reconstructing 'true' infection counts over time requires an accurate description of the delay distribution between infection and detection and an allowance for (right) truncation at time $t$, Further, $R_t$ fundamentally depends on the generation interval distribution that describes the time between infection of an index case and a subsequent transmission event to a susceptible. This is often approximated by the serial distribution, which measures the time between onset of symptoms of the index case and infected susceptible. The latter is observable, in principle, whereas the former is usually not directly calculable. While these two distributions have the same mean, their variance (and form) is different, and this misspecification introduces bias in estimation of $R_t$ (typically away from the null value of $R_t = 1$, and increasing with more extreme values of $R_t$). Smoothing of observed infection counts is usually needed to accommodate stochasticity in surveillance patterns and this introduces further potential for bias. We emphasize that the best estimation strategy in the world cannot overcome inadequate infection surveillance systems whose sampling and delay characteristics vary over space and time.

## Section 4. CORRELATES FOR CHANGES IN $R_t$

A major complexity in understanding the association between population characteristics/policies and changes in $R_t$ is the multifactorial nature of the measure, as noted in Section 1. Quite different "correlates" may exist that reflect their impact on (i) mixing patterns, (ii) susceptibility measures, and (iii) the duration of infectiousness, in some instances in near-real time. We consider each of these potential associations here. Note that there has been significant work done on each of these questions by exploiting various dynamic mathematical models of disease spread. Here, however, we focus on the effect of correlate changes in the field, where it is much harder to obtain high quality evidence, and interpret what is observed causally. There is insufficient space to comment on each of these illustrations in detail and each has their individual strengths and weaknesses. We make some brief remarks regarding inference associated with each approach.

The first association is illustrated by analyses that exploit Apple Maps routing-based mobility data to show quite varied relationships between $R_t$ and mobility measures in different locales. As one example, Miller *et al*.[29] examined the association of mobility with estimates of $R_t$ across various states in the US during the early stages of the pandemic in early 2020. Here, mobility was captured through a measure of Relative Routing Volume (RRV) that captures relative changes in requests for directions in Apple Maps as compared to a baseline date of January 13, 2020, prior to the onset of the pandemic in the US.

Miller *et al.*'s Figure 2[29] is a version of a ubiquitous plot--seen in many sources--of changes in estimated $R_t$ over time in various US states in the early stages of the pandemic. Of more interest, their Figure 3 (reproduced here, in part, as Figure 1) compares these $R_t$ estimates with RRV for four specific states. The association between estimated $R_t$ and reduction in RRV varies considerably across the states, with the estimated $R_t$ falling below one at different levels of RRV depending on the state. For Louisiana, $R_t$ is reduced to one when RRV falls to 65% (58-75%) of baseline levels. On the other hand, New York's $R_t$ falls below one only when RRV is 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.

In chronological time, Figure 1 should be read from right to left as March 2020 was a period of decreasing mobility over that month. The observed relationships may not be a reasonable description of what may happen to $R_t$ "in reverse," that is, in periods where mobility increases as social restrictions are lifted. Clearly, an important goal is to understand the "left to right" relationship sufficiently well to develop strategies for easing restrictions that increase mobility while at the same time minimizing increases in $R_t$.

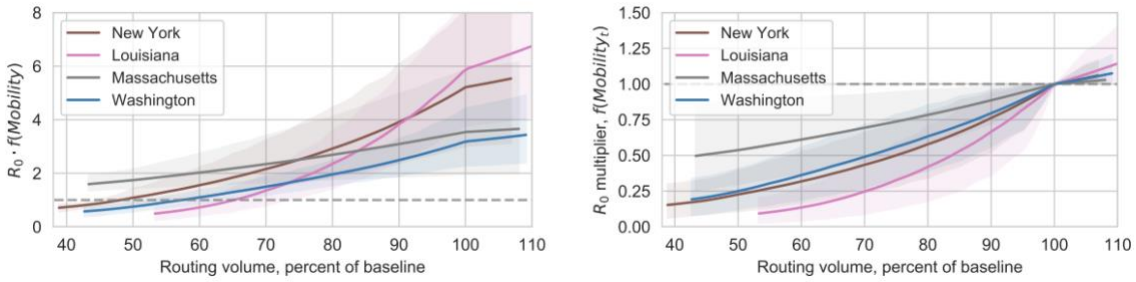


Figure 1. Miller *et al*. (2020). Left: Inferred relationship between the reproduction number and mobility volume change, for four US states. Right: the multiplier effect on initial reproductive number estimates as a function of relative change in mobility from baseline for the same four US states.

It goes without saying that Figure 1 represents associations rather than establishing a formal causal relationship. It is important not to ignore cautionary lessons learned regarding the difficulty of causal inference from observational, indeed here ecological, associations. There have been other analogous efforts to establish links between various population mixing factors associated with transmission and estimates of $R_t$. For example, Unwin *et al*.[30] model a relationship between infection transmission and mobility measures in the US from Google's COVID-19 Mobility Report using both COVID-19 infection and mortality counts. Here, mobility is similarly measured as a percentage change (from an appropriate baseline) in the number of visits to various venues including grocery markets, retailers, parks, transit stations, work sites, etc. Both mobility, infection and mortality data were aggregated at the US state-level. The authors ultimately claim, for example, that "if mobility stopped entirely (100% reduction in

average mobility then $R_t$ would be reduced by 55.1% [26.5%-77.0%].” However, such an estimate cannot be interpreted causally due to the role of other factors. Any such estimate, and related uncertainty in inference, are sensitive to model selection in a variety of ways.

Several researchers have provided more sophisticated analyses of this kind that use similar proxies for population mobility. Specifically, Brooks-Pollock *et al*. use data from the 2010 UK Social Contact Survey, coupled with Google community mobility reports, to estimate $R_t$, exploiting UK death data in March 2020 for calibration.[31] This approach is then used to quantify the impact of various social distancing policies by allowing the latter to modify the rate of social contacts. These kinds of analyses are essentially retrospective in nature and remain subject to causal inference challenges.

The second kind of correlate relationship is illustrated by examining the impact of vaccination programs on transmission. Figure 2 reproduces a graphic of Segal[32] that compares the trajectory of $R_t$ for forty days in two distinct periods in Israel. The second wave reflects the changes in $R_t$ after October 2020 when shelter-in-place restrictions were lifted, a period that saw $R_t$ increase from about 0.7 to 1.2 in about six weeks. The third wave curves shows analogous changes in $R_t$ after exiting “lockdown” in early Spring 2021. A critical difference between these two periods is that vaccinations were being administered around and after the peak of the third wave whereas there were no available vaccines in October and November 2020. Of course, vaccine administration did not occur overnight to the same extent as easing restrictions did; other data (from the Israel Ministry of Health website, not shown here) indicates that about 5% of the Israeli population had received one vaccination dose by the beginning of January 2021, rising to between 50-60% by mid-March (for two doses the analogous figures are close to 0% at the beginning of 2021, rising to about 45% by mid-March), with both increases relatively linear in time; during this period about 65% of the population was eligible for vaccination due to age and prior SARS-CoV-2 restrictions. This period essentially covers the 40-day period associated with the third wave in Figure 3. In principle, one can attempt to quantify the pattern of reduction in $R_t$ corresponding to increases in population vaccination percentages, in a manner analogous to Figure 1 for a mobility measure. However, such a strategy is fraught with difficulty, at least when applied to a single example: (i) the levels of natural immunity were naturally higher as the third wave progressed as compared to the second wave, (ii) policy and population responses to lockdown easing may have differed across the two waves (particularly given the experience of the second wave leading to large outbreaks), and (iii) the possibility that mixing grew faster after the third wave as compared to the second because of awareness of increasing vaccine administration. Nevertheless, the impact of increasing vaccination coverage, essentially by reducing susceptibles, appears dramatic and encouraging, but quantifying the “vaccination effect” on $R_t$ remains challenging.

There have been fewer field investigations of the third influence on $R_t$, namely interventions that shorten the effective period of infectiousness. One such approach is evidenced by attempts to capture the impact of imposing contact tracing, community

testing and case isolation on transmission. For example, Kendall *et al.*[33] examine the effect of the roll out of the National Health Service Test and Trace programme on the

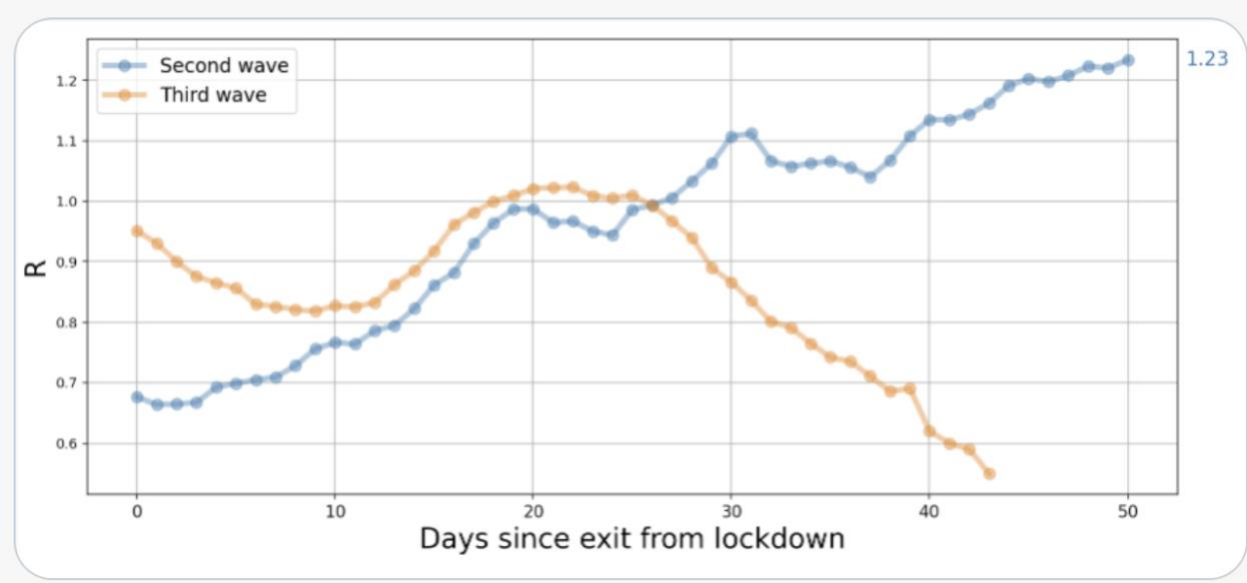


Figure 2. Israel reproductive number estimates per Segal (2021) for two periods where lockdown restrictions were eased: October 2020 forward (second wave) and mid-January 2021 forward (third wave).

Isle of Wight in May 2020. Their analysis estimated a greater than 50% reduction in $R_t$ which may be attributable, in part at least, to the intervention effort. However, this differs from the two previous examples in that the intervention, or putative correlate for changes in $R_t$, is binary—it is a comparison of a Test and Trace programme versus the status quo. There is no continuous 'dial' of how levels of contact tracing and testing modulate the value of $R_t$ in a community quantitatively, and thus no information on what might be important targets for community test and trace efforts.

The Isle of Wight example further reflects significant challenges to attempts to ascertain markers for $R_t$ and changes thereof. First, quantification of the association between community measures and estimates of $R_t$ is rarely simple, as reflected by both the mobility and vaccination examples. Relationships are unlikely to be linear and are potentially influenced by other local characteristics, making it difficult to isolate the impact of a single putative marker for disease transmission. Second, as noted above, associations are often made ecologically in large groups, thereby making causal inferences much harder to support, and transportability of effects likely less successful. Effect modification of relationships between a community marker and $R_t$ may be even harder to quantify but are no less important for policy decisions.

## Section 5. SUMMARY

A primary lesson of the pandemic has been the need to move extremely rapidly in times of early exponential growth—even a few days can make a substantial difference in averting infections. This raises a concern as to whether calculation of $R_t$, which at best reflects infection patterns some weeks in the past, is of value to policymakers who may find themselves always chasing the curve. In addition, outbreaks are often geographically variable and the locality of an $R_t$ estimate may not match the area where interventions are targeted. Perhaps, it is the detection and quantification of local changes in $R_t$ instead of, or at least in addition to, absolute levels, that may be more useful in influencing policy decisions. Many comparisons of locations with apparently similar levels of estimated $R_t$ exhibit different transmission patterns subsequently. In many cases, attempts to infer signals from such comparisons often do not allow for the role of chance in epidemic spread.

Despite valiant attempts, estimation of $R_t$ based on observed case numbers is always likely to be subject to substantial error, particularly in situations where asymptomatic transmission plays a major role. To be fair, the reservations expressed in these paragraphs apply similarly to the use of other quantitative measures of transmission--based on inadequate, and likely inaccurate, epidemiologic data—to respond to rapidly changing conditions. A fundamental lesson from almost all infectious disease outbreaks is that public health responses must be mounted *before* any significant evidence of transmission is evident. It is unclear whether policymakers, and the public, are willing to accept such actions in the absence of immediate evidence of overt disease spread, but the consequences of inaction have been demonstrated time and again.

With the strong emphasis on $R_t$ among policymakers and the public, valid and direct epidemiologic measures of community SARS-CoV-2 infection rates received less emphasis than might be ideal, at least early in the epidemic. Estimates of community transmission and seroprevalence were often initially based on convenience sampling rather than population-based strategies. It is challenging to interpret estimates of infection intensity when the latter necessarily were based on testing data in circumstances where testing strategies varied significantly over time and locale depending on a host of factors, not the least the availability of tests in the early stages of the pandemic. Similar considerations affected ad hoc approaches to seroprevalence. The use of population sampling methods has been the exception rather than the norm. It is remarkable that this was not necessarily the case in earlier pandemics when resources, technology and understanding of survey methodology were much less advanced, and the nature of the infectious agent was less well characterized. For example, 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.[34] With the exception of the UK REACT study,[35] few countries have launched comprehensive, systematic surveillance of SARS-CoV-2 active infection (or seroprevalence) to obtain an unbiased view of transmission intensity, and thus the need for non-pharmaceutical interventions to mitigate risk. Of note, provided study procedures achieve a fast turnaround in processing specimens, changes in infection prevalence can be detected in near real-time, affording an earlier view into transmission dynamics than $R_t$ estimates that must

inherently be delayed by at least one generation of infection; downward changes in prevalence of active infection necessarily indicate $R < 1$, whereas increases necessarily indicate $R > 1$. Information on seroprevalence is also of great importance to other aspects of epidemiologic studies and public health response, for instance providing a denominator to enable estimation of infection-to-hospitalization and infection-to-fatality ratios.[36] A fundamental approach to future outbreak responses must surely stress the enormous value of basic—but high quality—epidemiological surveillance data that effectively captures "how, when, and why" infections are occurring, rather than remaining "blind" and overly dependent on predictive models.

## REFERENCES


1 Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. Superspreading and the effect of individual variation on disease emergence. *Nature* 2005; 438: 355–9.

2 Laxminarayan R, Wahl B, Dudala SR, *et al.* Epidemiology and transmission dynamics of COVID-19 in two Indian states. *Science (80- )* 2020. DOI:10.1126/science.abd7672.

3 Sun K, Wang W, Gao L, *et al.* Transmission heterogeneities, kinetics, and controllability of SARS-CoV-2. *Science (80- )* 2021. DOI:10.1126/science.abe2424.

4 Wallinga J, Teunis P. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *Am J Epidemiol* 2004. DOI:10.1093/aje/kwh255.

5 Jarvis CI, Van Zandvoort K, Gimma A, *et al.* Quantifying the impact of physical distance measures on the transmission of COVID-19 in the UK. *BMC Med* 2020. DOI:10.1186/s12916-020-01597-8.

6 Van Den Driessche P, Watmough J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Math. Biosci. 2002; **180**: 29–48.

7 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. *Lancet Infect Dis* 2020. DOI:10.1016/S1473-3099(20)30785-4.

8 Kissler SM, Tedijanto C, Goldstein E, Grad YH, Lipsitch M. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science (80- )* 2020. DOI:10.1126/science.abb5793.

9 Ferguson NM, Laydon D, Nedjati-Gilani G, *et al.* Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. *imperial.ac.uk* 2020. DOI:10.25561/77482.

10 Britton T, Ball F, Trapman P. A mathematical model reveals the influence of population heterogeneity on herd immunity to SARS-CoV-2. *Science (80- )* 2020. DOI:10.1126/science.abc6810.

11 Aguas R, Corder RM, King JG, Gonçalves G, Ferreira MU, Gomes MGM. Herd immunity thresholds for SARS-CoV-2 estimated from unfolding epidemics. medRxiv. 2020. DOI:10.1101/2020.07.23.20160762.

12 Buss LF, Prete CA, Abrahim CMM, *et al.* Three-quarters attack rate of SARS-CoV-2 in the Brazilian Amazon during a largely unmitigated epidemic. *Science (80- )* 2021. DOI:10.1126/science.abe9728.
13 Anand S, Montez-Rath M, Boyd SD, Garcia P, Parsonnet J, Chertow GM. SARS-COV-2 antibody prevalence in patients on dialysis in the US in January 2021. *medRxiv* 2021.
14 Malani A, Shah D, Kang G, *et al.* Seroprevalence of SARS-CoV-2 in slums versus non-slums in Mumbai, India. Lancet Glob. Heal. 2021. DOI:10.1016/S2214-109X(20)30467-8.
15 Lourenço J, Pinotti F, Thompson C, Gupta S. The impact of host resistance on cumulative mortality and the threshold of herd immunity for SARS-CoV-2. medRxiv. 2020. DOI:10.1101/2020.07.15.20154294.
16 Mahase E. Covid-19: What is the R number? *BMJ* 2020. DOI:10.1136/bmj.m1891.
17 Gostic KM, McGough L, Baskerville EB, *et al.* Practical considerations for measuring the effective reproductive number, Rt. PLoS Comput. Biol. 2020. DOI:10.1371/journal.pcbi.1008409.
18 Prime Minister's statement on coronavirus (COVID-19)" 10 May 2020. https://www.gov.uk/government/speeches/pm-address-to-the-nation-on-coronavirus-10-may-2020
19 Pei S, Kandula S, Shaman J. Differential effects of intervention timing on COVID-19 spread in the United States. *Sci Adv* 2020. DOI:10.1126/sciadv.abd6370.
20 Jewell B, Jewell N. The Huge Cost of Waiting to Contain the Pandemic. New York Times. 2020. https://www.nytimes.com/2020/04/14/opinion/covid-social-distancing.html.
21 Jewell BL, Jewell NP. On the role of statisticians and modelers in responding to AIDS and COVID-19. *Stat Med* 2021; 40(11): 2530-2535. https://doi.org/10.1002/sim.8943
22 O'Driscoll M, Harry C, Donnelly CA, Cori A, Dorigatti I. A comparative analysis of statistical methods to estimate the reproduction number in emerging epidemics with implications for the current COVID-19 pandemic. *Clin Infect Dis* 2020. DOI:10.1093/cid/ciaa1599.
23 Dietz K. The estimation of the basic reproduction number for infectious diseases. *Stat Methods Med Res* 1993. DOI:10.1177/096228029300200103.
24 Cori A, Ferguson NM, Fraser C, Cauchemez S. A new framework and software to estimate time-varying reproduction numbers during epidemics. *Am J Epidemiol* 2013. DOI:10.1093/aje/kwt133.
25 Lipsitch M, Cohen T, Cooper B, *et al.* Transmission dynamics and control of severe acute respiratory syndrome. *Science* 2003; **300**: 1966–70.
26 Pitzer V, Chitwood M, Havumaki J, *et al.* The impact of changes in diagnostic testing practices on estimates of COVID-19 transmission in the United States. *medRxiv Prepr Serv Heal Sci* 2020. DOI:10.1101/2020.04.20.20073338.
27 Jewell NP, Lewnard JA, Jewell BL. Caution Warranted: Using the Institute for Health Metrics and Evaluation Model for Predicting the Course of the COVID-19 Pandemic. *Ann Intern Med* 2020. DOI:10.7326/m20-1565.
28 Miller AC, Hannah L, Futoma J, *et al.* Statistical deconvolution for inference of

infection time series. medRxiv. 2020. DOI:10.1101/2020.10.16.20212753.

29 Miller AC, Foti NJ, Lewnard JA, Jewell NP, Guestrin C, Fox EB. Mobility trends provide a leading indicator of changes in SARS-CoV-2 transmission. medRxiv. 2020. DOI:10.1101/2020.05.07.20094441.

30 Unwin HJT, Mishra S, Bradley VC, *et al.* State-level tracking of COVID-19 in the United States. *Nat Commun* 2020. DOI:10.1038/s41467-020-19652-6.

31 Brooks-Pollock E, Read JM, McLean AR, *et al*. Mapping social distancing measures to the reproduction number for COVID-19. Phil. Trans. R. Soc. B 2021; 376: 20200276. https://doi.org/10.1098/rstb.2020.0276.

32 Segal E. [@segal_eran]. (2021, March 25) *Israel: Continuing to crush the curve*[Tweet], Twitter, https://twitter.com/segal_eran/status/137499592842851532

33 Kendall M, Milsom L, Abeler-Dörner L, Wymant C, Ferretti L, Briers M, Holmes C, Bonsall D, Abeler J, Fraser C. Epidemiological changes on the Isle of Wight after the launch of the NHS Test and Trace programme: a preliminary analysis. *Lancet Digital Health.* 0 (2020), doi:10.1016/S2589-7500(20)30241-7.

34 Morabia A. The US Public Health Service house-to-house canvass survey of the morbidity and mortality of the 1918 influenza pandemic. *Amer. J. Pub. Health* 2020 0, e1_e8, https://doi.org/10.2105/AJPH.2020.306025.

35 Riley S, Atchison C, Ashby D, *et al.* REal-time Assessment of Community Transmission (REACT) of SARS-CoV-2 virus: Study protocol [version 1;peer review: awaiting peer review]. *Wellcome Open Res* 2020.

36 O'Driscoll M, Dos Santos GR, Wang L, *et al.* Age-specific mortality and immunity patterns of SARS-CoV-2. *Nature* 2020. DOI:10.1038/s41586-020-2918-0.