Practical considerations for measuring the effective reproductive number, R_t

Katelyn M. Gostic[1*], Lauren McGough[1], Edward B. Baskerville[1], Sam Abbott[2], Keya Joshi[3], Christine Tedijanto[3], Rebecca Kahn[3], Rene Niehus[3], James Hay[3], Pablo M. De Salazar[3], Joel Hellewell[2], Sophie Meakin[2], James Munday[2], Nikos I. Bosse[2], Katharine Sherrat[2], Robin N. Thompson[2,4], Laura F. White[5], Jana S. Huisman[6,7], Jérémie Scire[7,8], Sebastian Bonhoeffer[6], Tanja Stadler[7,8], Jacco Wallinga[9,10], Sebastian Funk[2], Marc Lipsitch[3], Sarah Cobey[1]

1 Department of Ecology and Evolution, University of Chicago, Chicago, IL, USA
2 Centre for Mathematical Modelling of Infectious Diseases, Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, London, UK
3 Center for Communicable Disease Dynamics, Department of Epidemiology, T.H. Chan School of Public Health, Harvard University, Cambridge, MA, USA
4 Mathematical Institute, University of Oxford, Oxford, UK
5 Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA
6 Department of Environmental Systems Science, ETH Zürich, Zürich, Switzerland
7 Department of Biosystems Science and Engineering, ETH Zürich, Switzerland
8 Swiss Institute of Bioinformatics, Basel, Switzerland
9 Centre for Infectious Disease Control, National Institute for Public Health and the Environment, Bilthoven, The Netherlands
10 Department of Biomedical Data Sciences, Leiden University Medical Centre, Leiden, The Netherlands

* kgostic@uchicago.edu

Abstract

Estimation of the effective reproductive number, R_t, is important for detecting changes in disease transmission over time. During the COVID-19 pandemic, policymakers and public health officials are using R_t to assess the effectiveness of interventions and to inform policy. However, estimation of R_t from available data presents several challenges, with critical implications for the interpretation of the course of the pandemic. The purpose of this document is to summarize these challenges, illustrate them with examples from synthetic data, and, where possible, make recommendations. For near real-time estimation of R_t, we recommend the approach of Cori et al. (2013), which uses data from before time t and empirical estimates of the distribution of time between infections. Methods that require data from after time t, such as Wallinga and Teunis (2004), are conceptually and methodologically less suited for near real-time estimation, but may be appropriate for retrospective analyses of how individuals infected at different time points contributed to spread. We advise against using methods derived from Bettencourt and Ribeiro (2008), as the resulting R_t estimates may be biased if the underlying structural assumptions are not met. Two key challenges common to all approaches are accurate specification of the generation interval and reconstruction of the time series of new infections from observations occurring long after the moment of

transmission. Naive approaches for dealing with observation delays, such as subtracting delays sampled from a distribution, can introduce bias. We provide suggestions for how to mitigate this and other technical challenges and highlight open problems in R_t estimation.

Author summary

The effective reproductive number, R_t, is a key epidemic parameter used to assess whether an epidemic is growing, shrinking or holding steady. R_t estimates can be used as a near real-time indicator of epidemic growth or to assess the effectiveness of interventions. But due to delays between infection and case observation, estimating R_t in near real-time, and correctly inferring the timing of changes in R_t is challenging. Here, we provide an overview of challenges and best practices for accurate, timely R_t estimation.

Introduction

The effective reproductive number, denoted R_e or R_t, is the expected number of new infections caused by an infectious individual in a population where some individuals may no longer be susceptible. Estimates of R_t are used to assess how changes in policy, population immunity, and other factors have affected transmission at specific points in time [1–5]. The effective reproductive number can also be used to monitor near real-time changes in transmission [6–10]. For both purposes, estimates need to be accurate and correctly represent uncertainty, and for near real-time monitoring, they also need to be timely.

We consider two potential forms of bias in R_t estimates, systematic over- or underestimation and temporal inaccuracy. Misspecification of the generation interval is a large potential source of over- or underestimation, and we find that R_t estimates are most prone to this kind of bias when the true value is substantially greater or less than one. This situation might arise at the beginning of the COVID-19 pandemic (when R_t is relatively high) or after particularly effective interventions (when it might be low). Over or underestimation would have particularly strong practical consequences near the control threshold of $R_t = 1$, but the biases we observe are smallest in absolute terms in this range.

Another challenge is that depending on the methods used, R_t estimates may be leading or lagging indicators of the true value [4, 11], even measuring transmission events that occurred days or weeks ago if the data are not properly adjusted. Temporal inaccuracy in R_t estimation is particularly concerning when trying to infer how changes in behavior have affected transmission [1–5]. Temporal inaccuracy, for instance in the estimation of the date on which R_t falls below one, is a focus of this Perspective. We find that it has several possible causes and can be difficult to avoid.

This perspective focuses on the three main empirical methods to estimate R_t [12–15]. As an alternative to the methods reviewed here, it is possible to infer changes in transmission using a dynamical compartment model (e.g. [3, 16–18]). The accuracy and timeliness of R_t estimates obtained in this way should be assessed on a case-by-case basis, given sensitivity to model structure and data availability.

We use synthetic data to compare the accuracy of three common empirical methods to estimate R_t, first under ideal conditions, in the absence of parametric uncertainty and with all infections observed at the moment they occur. This idealized analysis is intended to illustrate the inputs needed to estimate R_t accurately, to highlight the intrinsic differences between the methods, and to examine specific causes of bias and

temporal inaccuracy one by one. However, we emphasize that our idealized analyses overestimate the potential accuracy of R_t estimates obtained from real-world data, even if best practice are followed. The results show that the method of Cori et al. [14] is best for near real-time estimation of R_t. For retrospective analysis, the methods of Cori et al. or of Wallinga and Teunis may be appropriate, depending on the aims.

Later we add realism and address practical considerations for working with imperfect data. These analyses emphasize potential errors introduced by uncertainty in the intrinsic generation interval and imperfect case observation, and also the need to adjust for delays in case observation, right truncation, and the need to choose an appropriate smoothing window given the sample size. Finally, we emphasize that most off-the-shelf tools leave it up to the user to account for these five sources of uncertainty when calculating confidence intervals. Failure to propagate uncertainty in R_t estimates can lead to over-interpretation of the results and could falsely imply that confidence or credible intervals have crossed the critical threshold.

Synthetic data

We used synthetic data to compare three common R_t estimation methods. Synthetic data were generated from a deterministic or stochastic SEIR model in which the transmission rate drops and spikes abruptly, representing the adoption and lifting of public health interventions. Results were similar whether data were generated using a deterministic or stochastic model. For simplicity, we show deterministic outputs throughout the document, except in the section on smoothing windows, where stochasticity is a conceptual focus.

In our model, all infections are locally transmitted, but all three of the methods we test can incorporate cases arising from importations or zoonotic spillover [12, 13, 15]. Estimates of R_t are likely to be inaccurate if a large proportion of cases involve transmission outside the population. This situation could arise when transmission is low (e.g., at the beginning or end of an epidemic) or when R_t is defined for a population that is connected to others via migration.

A synthetic time series of new infections (observed at the $S \to E$ transition) was input into the R_t estimation methods of Wallinga and Teunis, Cori et al., and Bettencourt and Ribeiro [12–14]. Following the published methods, we also tested the Wallinga and Teunis estimator using a synthetic time series of symptom onset events, extracted daily from the $E \to I$ transition. In the synthetic data, the generation interval followed a gamma distribution with shape 2 and rate $\frac{1}{4}$, which is the sum of exponentially distributed residence times in compartments E and I, each with mean of 4 days [19]. The methods of Cori et al. and of Wallinga and Teunis can accommodate any positive, discrete generation interval distribution [12, 14, 20]. We chose a discretized, gamma-distributed generation interval for our simulations, which is similar in shape to the interval of COVID-19 [21, 22]. However, the method of Bettencourt and Ribeiro implicitly assumes the generation interval follows an exponential distribution, as in an SIR model with no latent period [13, 19], and thus could not match the assumptions of the synthetic data.

In the synthetic data, R_0 was set to 2.0 initially, then to 0.8 and 1.15, to simulate the adoption and later the partial lifting of public health interventions. To mimic estimation in real-time, we truncated the time series at $t = 150$, before the end of the epidemic. Estimates from the methods of Wallinga and Teunis and Cori et al. were obtained using the R package EpiEstim [20]. Estimates based on the method of Bettencourt and Ribeiro were obtained by translating code from [6, 23] to the Stan language [24]. We initially assumed all infections were observed. Unless otherwise noted, the smoothing window was set to 1 day (effectively, estimates were not smoothed). To

mimic the timescale of observations, we used daily time steps when generating synthetic data and performing analyses.

Comparison of common methods

The effective reproductive number at time t can be defined in two ways: as the instantaneous reproductive number or as the case reproductive number [14, 25]. The instantaneous reproductive number measures transmission at a specific point in time, whereas the case reproductive number measures transmission by a specific cohort of individuals (Fig. 1). (A cohort is a group of individuals with the same date of infection or the same date of symptom onset.) The case reproductive number is useful for retrospective analyses of how individuals infected at different time points contributed to spread. It is a more natural choice for analyses that consider heterogeneity among individuals. For example, the case reproductive number of Wallinga and Teunis can be adapted to incorporate data on observed transmission chains [12, 21, 26] or to produce age-structured R_t estimates, given an age-structured contact matrix [27]. The instantaneous reproductive number is more appropriate for analyses estimating the reproductive number of the infected population on specific dates, especially when aiming to study how interventions or other extrinsic factors have affected transmission at a given point in time.

More formally, the **instantaneous reproductive number** is defined as the expected number of secondary infections occurring at time t, divided by the number of infected individuals, each scaled by their relative infectiousness at time t (an individual's relative infectiousness is based on the generation interval, and time since infection) [14, 25]. The instantaneous reproductive number can be calculated exactly for a compartment model (SIR or SEIR) as follows, where $\beta(t)$ is the time-varying transmission rate, $S(t)$ the fraction of the population that is susceptible, and D the mean duration of infectiousness:

$$R_t^{\text{inst}} = \beta(t)S(t)D. \tag{1}$$

The methods of Cori et al. [14, 15] and methods adapted from Bettencourt and Ribeiro [6, 13, 23] estimate the instantaneous reproductive number from observations. We tested their accuracy under idealized conditions, assuming perfect knowledge of the generation interval and delay distributions used to generate the synthetic data (Fig. 2 and S2 Fig).

The method of Cori et al. estimates R_t as

$$R_t = \frac{I_t}{\sum_{s=1}^{t} I_{t-s} w_s}, \tag{2}$$

where I_t is the number of infections incident on day t and w_s is generation interval, or the probability that s days separate the moment of infection in an index case and a daughter case [14]. Conceptually, this estimator describes the number of new infections incident on day t relative to the number (I_{t-s}) and current infectiousness (w_s) of individuals who became infected s days in the past, and who many now be shedding virus.

The only parametric assumption required by this method is the form of the generation interval. The standard assumption is that w_s follows a discretized gamma distribution [14, 20], but the estimator accepts any parametric or empirical discrete distribution with support on positive values (the same is true of the Wallinga and Teunis method [12]). Thus, when testing the method on COVID-19-like epidemic processes, we could specify the gamma-distributed generation interval of the synthetic data perfectly. When tested under idealized conditions, we found that the method of

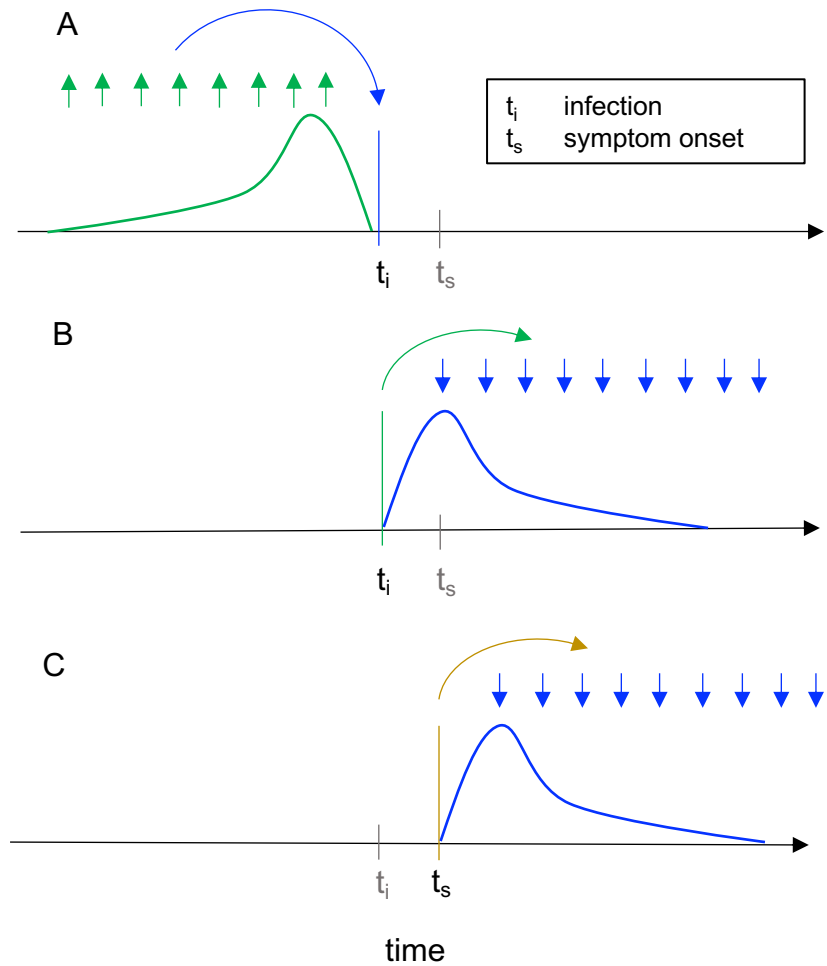


Fig 1. Instantaneous reproductive number as estimated by the method of Cori et al. versus cohort reproductive number estimated by Wallinga and Teunis. For each definition of R_t, arrows show the times at which infectors (upwards), and their infectees (downwards) appear in the data. Curves show the generation interval distribution (A,B), or serial interval distribution (C). A: The instantaneous reproductive number quantifies the number of new infections incident at a single point in time (t_i, blue arrow), relative to the number of infections incident in the previous generation (green arrows), and their current infectiousness (green curve). This figure illustrates the method of Cori et al. B & C: The case reproductive number of Wallinga and Teunis is the average number of new infections that an individual who becomes (B) infected on day t_i (green arrow) or (C) symptomatic on day t_s (yellow arrow) will eventually go on to cause (blue downward arrows show timing of daughter cases). The first definition applies when estimating the case reproductive number using inferred times of infection, and the second applies when using data on times of symptom onset.

Cori et al. accurately estimated R_t, even tracking abrupt changes (Fig. 2). This is the method we recommend for estimation of the instantaneous reproductive number.

Bettencourt and Ribeiro [13] derive an approximate relationship between the R_t and the exponential growth rate of the epidemic, where g is the mean generation time:

$$I_{t+1} = I_t e^{\frac{R_t - 1}{g}}. \tag{3}$$

Under the assumption that R_t evolves through time as a Gaussian process, equation 3 facilitates efficient Bayesian R_t estimation [23]. The key disadvantage of this method is that equation 3 is derived from an SIR model, and therefore implicitly assumes the generation interval follows an exponential distribution [19], whereas empirically, generation intervals can be heavy or light tailed, including for COVID-19 [21, 22, 28]. Because equation 3 underestimates variability in the gamma-distributed generation interval of our SEIR-type synthetic data, we find that this method produces biased R_t estimates, especially when R_t is substantially higher than one (Fig. 2A). These biases are consistent with established theory, in which the coefficient of variation of the generation interval modulates the relationship between exponential growth rate and the reproductive number of an epidemic [19].

In its current form, we do not recommend using the method of Bettencourt and Ribeiro given that unrealistic structural assumptions lead to bias. However, a generalized version capable of accommodating more realistic generation intervals, which implicitly involves different assumptions about the underlying epidemic process [19, 29], could provide several advantages. When implemented as a Gaussian process [23], we found that the Bettencourt and Ribeiro method was computationally efficient. Also, because it penalizes large jumps in R_t across consecutive time steps, it returns smoother estimates than the method of Cori et al., which is advantageous if unmodeled reporting effects, rather than bursts in transmission, are the dominant cause of variability in daily observations.

Finally, the **case or cohort reproductive number** is the expected number of secondary infections that an individual who becomes infected at time t will eventually cause as they progress through their infection [14, 19, 25] (Fig. 1B,C). The case reproductive number R_t^{case} can be calculated exactly at time t within the synthetic data as the convolution of the generation interval distribution $w(\cdot)$ and the instantaneous reproductive number, R_t^{inst}, described in Equation 1 [19],

$$R_t^{\text{case}} = \int_{u=t}^{\infty} R_u^{\text{inst}} w(u - t)\, du. \tag{4}$$

The method of Wallinga and Teunis [12] estimates the case reproductive number from observations. The first step is to estimate the likelihood that case j (infected at time t_j) infected case i, relative to the likelihood that any other case in the data infected case i:

$$p_{ij} = \frac{w(t_i - t_j)}{\sum_{i \neq k} w(t_i - t_k)}. \tag{5}$$

Then, the individual reproductive number of case j is defined as

$$R_j = \sum_i p_{ij}. \tag{6}$$

The case reproductive number at time t is defined as the expected value of R_j for all individuals infected at time t. Like the method of Cori et al., the only parametric assumption required by the method of Wallinga and Teunis is the discrete distribution of the generation interval. An assumption common to the three tested methods is that all infections are observed. Below we discuss the consequences of partial observation, including that confidence or credible intervals around R_t estimates do not account for uncertainty from partial observation.

Practically speaking, there are several important differences between the case reproductive number (estimated by Wallinga and Teunis) and the instantaneous

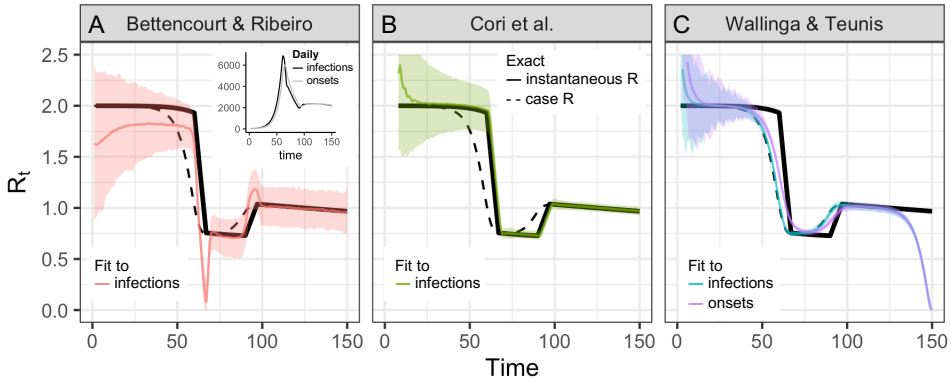


Fig 2. Accuracy of R_t estimation methods given ideal, synthetic data. Solid and dashed black line shows the instantaneous and case reproductive numbers, respectively, calculated from synthetic data. Colored lines show estimates and confidence or credible intervals. To mimic an epidemic progressing in real-time, the time series of infections or symptom onset events up to $t = 150$ was input into each estimation method (inset). Terminating the time series while R_t is falling or rising produces similar results S2 Fig. **(A)** By assuming a SIR model (rather than SEIR, the source of the synthetic data), the method of Bettencourt and Ribeiro systematically underestimates R_t when the true value is substantially higher than one. The method is also biased as transmission shifts. **(B)** The Cori method accurately measures the instantaneous reproductive number. **(C)** The Wallinga and Teunis method estimates the cohort reproductive number, which incorporates future changes in transmission. Thus, the method produces R_t estimates that lead the instantaneous effective reproductive number and becomes unreliable for real-time estimation at the end of the observed time series without adjustment for right truncation [4, 30]. In (A,B) the colored line shows the posterior mean and the shaded region the 95% credible interval. In (C) the colored line shows the maximum likelihood estimate and the shaded region the 95% confidence interval.

reproductive number (estimated by Cori et al. or Bettencourt and Ribeiro). First, estimators of the instantaneous reproductive number were partly developed for near real-time estimation and only use data from before time t (Fig. 1A). Under ideal conditions without observation delays and a window size of one day, neither method is affected by the termination of the synthetic time series at $t = 150$ (Figs. 2 A&B). These methods are similarly robust if the time series ends while R_t is rapidly falling (S2 FigA) or rising (S2 FigB). (Below we discuss more realistic conditions, e.g., in which data at the end of a right-truncated time series are incomplete due to observation delays.)

Unlike the instantaneous reproductive number, the case reproductive number is inherently forward-looking (Fig. 1B,C): near the end of a right-truncated time series, it relies on data that have not yet been observed. Extensions of the Wallinga and Teunis method can be used to adjust for these missing data and to obtain accurate R_t estimates to the end of a truncated time series [4, 30]. But as shown in Fig. 2C, without these adjustments the method will always underestimate R_t at the end of the time series, even in the absence of reporting delays. Mathematically, this underestimation occurs because calculating the case reproductive number involves a weighted sum across transmission events observed after time t. Time points not yet observed become missing terms in the weighted sum. Similarly, any infections that occurred before the first observed date are missing terms in the denominator of the Cori et al. estimator, and so the method of Cori et al. often overestimates R_t early in the time series.

Effectively, the case reproductive number is shifted forward in time relative to the instantaneous reproductive number of Cori et al. (S3 Fig). This temporal shift occurs whether or not a smoothing window is used. The case reproductive number produces leading estimates of changes in the instantaneous reproductive number (Fig. 2, S3 Fig) because it uses data from time points after t, whereas the instantaneous reproductive number uses data from time points before t (Fig. 1). Shifting the case reproductive number back in time by the mean generation interval usually provides a good approximation of the instantaneous reproductive number [2], because the case reproductive number is essentially a convolution of the instantaneous reproductive number and the generation interval (Equation 4) [19]. The case reproductive number generally changes more smoothly than the instantaneous reproductive number [25] (Fig. 2B,C), but if a smoothing window is used, the estimates become more similar in shape and smoothness. Overall, for real-time analyses aiming to quantify the reproductive number at a particular moment in time or to infer the impact of changes in policy, behavior, or other extrinsic factors on transmission, the instantaneous reproductive number will provide more temporally accurate estimates and is most appropriate. The case reproductive number of Wallinga and Teunis considers the reproductive number of specific individuals, and therefore is more appropriate for analyses aiming to incorporate individual-level covariates such as age [27] or transmission cluster membership [12] when estimating R_t.

Summary

- The Cori method most accurately estimates the instantaneous reproductive number in real-time. It uses only past data and minimal parametric assumptions.
- The method of Wallinga and Teunis estimates a slightly different quantity, the case or cohort reproductive number. The case reproductive number is conceptually less appropriate for real-time estimation but may be useful in retrospective analyses, especially those involving individual-level covariates.
- Methods adapted from Bettencourt and Ribeiro [6, 13] can lead to biased R_t estimates if the underlying structural assumptions are not met.

Generation interval misspecification

When estimating R_t from observed data, misspecification of the generation interval is a large potential source of bias. Regardless of the method used, R_t estimates are not only sensitive to the mean generation time but also to the variance and form of the generation interval distribution [19].

The renewal equation is a cornerstone of demographic theory and forms the mathematical backbone of the R_t estimators described above [19]. Within the renewal equation, the generation interval mechanistically links the reproductive number R to observables such as the epidemic growth rate r or the number of new infections per day [19]. Wallinga and Lipsitch [19] describe how the exponential growth rate of the Bettencourt and Ribeiro estimator (their equation 3.1) and continuous-time equivalents of the Cori et al. and Wallinga and Teunis estimators (their equations 4.1 and 4.2) can be derived from the renewal equation model.

Originally developed in the context of population biology, the renewal equation is usually expressed as $b(t) = \int_{a=0}^{\infty} b(t-a)n(a)da$, where $b(t)$ is the number of births at time t and $n(a)$ is fecundity at age a, scaled by the probability of surviving to age a. When used to describe epidemic dynamics, the renewal equation model is expressed in

terms of $I(t)$, the number of infections incident at time t; $S(t)$, the susceptible fraction; and $w(\cdot)$, the generation interval distribution [29, 31, 32]. Note that $R_0 S(t) = R_t$.

$$I(t) = R_0 S(t) \int_0^\infty I(t-s) w(s) ds \tag{7}$$

The difficulty is that the "intrinsic" generation interval of the renewal equation, which is the interval needed for accurate R_t estimation, is conceptually and quantitatively different from the generation intervals observed in practice [31–33]. In the renewal equation, the generation interval describes the time distribution of all infectious contact events, whether or not the contacted individual is susceptible. (Note that a different factor in the equation, $S(t)$, scales the probability of contacting a susceptible individual and causing a new infection.) The demographic analogy is that $w(\cdot)$ describes changes over time in the fecundity (infectiousness) of an index case, while $S(t)$ determines whether offspring of that index case are viable. In practice, we only observe generation intervals between viable pairs. Thus, unlike the intrinsic generation interval, observed intervals are sensitive to changes in the rate of susceptible depletion and can be biased estimators of the intrinsic interval, and of R_t [31–33]. A related challenge is that interventions such as contact tracing and self-isolation can limit transmission late in the course of infectiousness, shortening observed generation and serial intervals [32, 34, 35]. Methods for accurate estimation of the generation interval from contact tracing data involve adjusting for right truncation, and accounting for population susceptibility at the times transmission pairs are observed [32, 33].

The serial interval, defined as the time between symptom onset in an infector-infectee pair, is more easily observed than the generation interval and often used in its place. Although the serial and generation interval are often conflated, failure to understand the differences between these related quantities can bias R_t estimates [21, 32]. The serial interval and the generation interval have the same mean, but usually different variance [36, 37], and the serial interval can be negative (e.g., for COVID-19 [38–40]), whereas the generation interval cannot [21, 32]. The intrinsic generation interval can be estimated from contact tracing data (i.e., estimates of the serial interval) [32, 33].

Fig. 3A illustrates the consequences of misspecifying the mean generation interval in the method of Cori et al. If the mean generation interval is set too high, R_t values will typically be further from 1 than the true value—too high when $R_t > 1$ and too low when $R_t < 1$. If the mean is set too low, R_t values will typically be closer to 1 than the true value. These biases are relatively small when R_t is near the critical threshold of one but increase as R_t takes substantially higher or lower values (Fig. 3). Therefore, biases from misspecification of the generation interval may be greatest early in an epidemic, when the sensitivity of high R_t values to bias may be compounded by limited data and highly uncertain generation interval estimates.

The consequences of misspecifying the form and variance of the generation interval distribution are illustrated in Fig. 2A and Fig. 3B. As explained above, biases in the Bettencourt and Ribeiro estimator arise entirely from misspecification of the form of the generation interval, even if the mean is correctly specified. Similarly, the accuracy of the Cori et al. and Wallinga and Teunis estimators in Fig. 2 B,C are contingent on perfect specification of the generation interval; in practice, if the mean, variance or form of a pathogen's true generation interval is uncertain, R_t estimates obtained using these methods can be biased.

EpiEstim [20] allows users to account for uncertainty in the mean and standard deviation of the generation interval by resampling over a range of plausible values [14, 20]. Similarly, Bayesian methods such as EpiNow2 [41] and the rt.live adaptation [6] of the Bettencourt and Ribeiro method allow users to specify the prior variance of the mean and standard deviation. Uncertainty around an incorrect value can

widen the resulting 95% interval but will not shift the assumed central value toward the truth, and will not correct bias in the central R_t estimates.

Joint estimation of both R_t and the serial interval is possible, depending on data quality and magnitude of R_t [12, 42, 43], and the EpiEstim [15, 20] package provides an off-the-shelf option for joint estimation. However, these off-the-shelf methods should be used with caution, as they estimate the observed serial interval, not the intrinsic generation interval, and do not account for changes over time in behavior or susceptibility.

Summary

- The intrinsic generation interval is required to correctly define the relationship between R_t and incident infections.
- The intrinsic generation interval is rarely observable, and care must be taken to estimate it from proxies such as the serial interval.

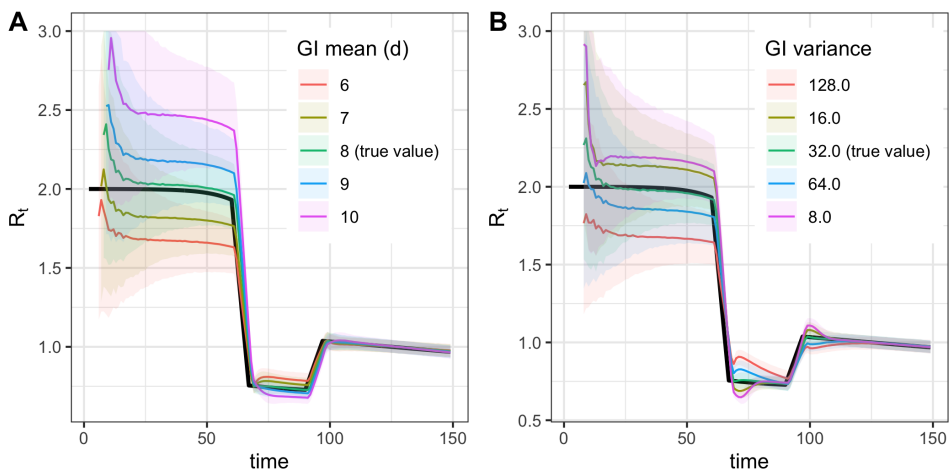


Fig 3. Biases from misspecification of the generation interval mean (A) or variance (B). Demonstrated using the method of Cori et al.

Adjusting for delays

Estimating R_t requires data on the daily number of new infections (i.e., transmission events). Due to lags in the development of detectable viral loads, symptom onset, seeking care, and reporting, these numbers are not readily available. All observations reflect transmission events from some time in the past. In other words, if d is the delay from infection to observation, then observations at time t inform R_{t-d}, not R_t (Fig. 4). Obtaining temporally accurate R_t estimates thus requires assumptions about lags from infection to observation. If the distribution of delays can be estimated, then R_t can be estimated in two steps: first by inferring the incidence time series from observations and then by inputting the inferred time series into an R_t estimation method. Alternatively, the unobserved time series could be inferred simultaneously with R_t or treated as a latent state. Such methods are now under development and available in a development version of the R package EpiNow2 [41, 44].

Two simple but mathematically incorrect methods for inference of unobserved times of infection have been applied to COVID-19: convolution and temporal shifts. The

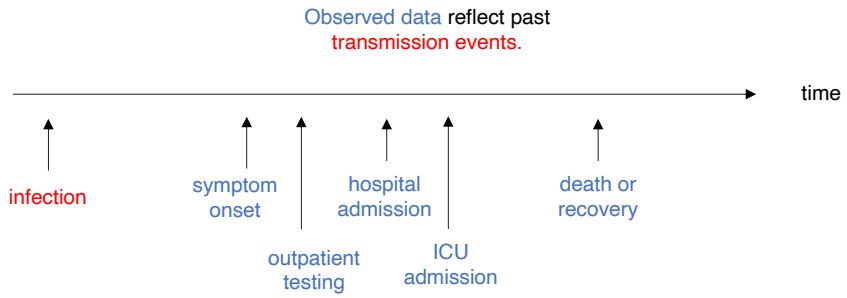


Fig 4. R_t is a measure of transmission at time t. Observations after time t must be adjusted.

errors introduced by these methods may be tolerable if delays to observation are relatively short and not highly variable, and if R_t is not rapidly changing. But when dealing with longer or more variable observation delays, or when aiming to infer the timing of changes in R_t accurately, these methods may not be sufficient.

One method infers each individual's time of infection by subtracting a sample from the delay distribution from each observation time. This is mathematically equivalent to convolving the observation time series with the reversed (backward) delay distribution, but convolution does not accurately infer the underlying time series of infections from observations [45–47]. The forward delay distribution has the effect of spreading out infections incident on a particular day across many days of observation. This blurring into the future is biologically realistic and reflects individual variation in disease progression and care seeking. To recover the original time series of infections from observations requires deblurring. Instead, as illustrated in (S1 Fig), backward convolution unrealistically spreads them out further. An unintended consequence of added blurring from backward convolution is that it can help smooth over weekend effects and other observation noise. But a crucial pitfall is that this blurring also smooths over true variation in R_t: peaks, valleys, and changes in slope of the latent time series of infection events. Convolution and other approaches that blur or oversmooth can therefore prevent or delay detection of changes in R_t and can impede accurate inference of the timing of these changes (Fig. 5C).

The second simple-but-incorrect method to adjust for lags is to shift either the raw inputs (the observed time series) or outputs of R_t estimation on the time axis. R_t estimates obtained by applying the methods of Cori et al. or Bettencourt and Ribeiro to unadjusted data will lag the true instantaneous R_t by roughly the mean delay from infection to observation. Because the case reproductive number leads the instantaneous reproductive number by roughly one generation interval, the unadjusted estimates obtained from the Wallinga and Teunis method will lag the true instantaneous reproductive number by roughly the mean delay to observation minus the mean generation interval. Unlike backward convolution, temporal shifting does not further blur the observed time series. Thus, if the mean delay is known accurately, this method is preferable to subtracting samples from the delay distribution (Fig. 5 A,B). However, shifting the input time series does not undo the blurring effect of the original delay, which, like backward convolution can impede accurate inference of changes in R_t. Shifting inputs or R_t estimates by a fixed amount also fails to account for realistic uncertainty in the true mean delay, which will not be known exactly and might change over time.

More reliable methods to reconstruct the incidence time series are now under

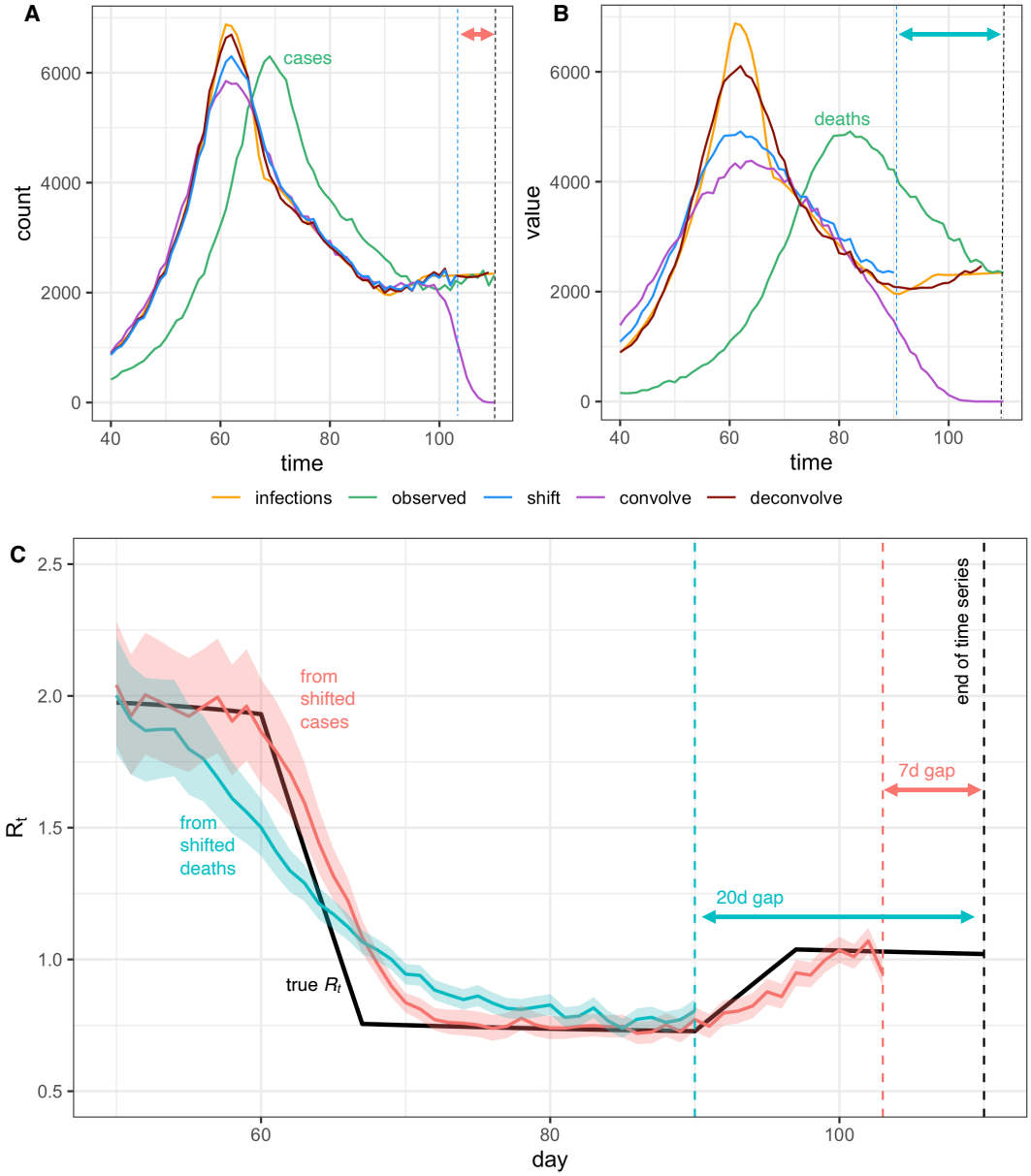


Fig 5. Pitfalls of simple methods to adjust for delays to observation when estimating R_t. Infections back calculated from (A) observed cases or (B) observed deaths either by shifting the observed curve back in time by the mean observation delay (shift), by subtracting a random sample from the delay distribution from each individual time of observation (convolve), or by deconvolution (deconvolve), without adjustment for right truncation. Neither back-calculation strategy accurately recovers peaks or valleys in the true infection curve. The inferred infection curve is less accurate when the variance of the delay distribution is greater (B vs. A). (C) Posterior mean and credible interval of R_t estimates from the Cori et al. method. Inaccuracies in the inferred incidence curves affect R_t estimates, especially when R_t is changing (here R_t was estimated using shifted values from A and B). Finally we note that shifting the observed curves back in time without adjustment for right truncation leads to a gap between the last date in the inferred time series of infection and the last date in the observed data, as shown by the dashed lines and horizontal arrows in A-C.

development. Given a known delay distribution, one potential solution is to infer the unlagged signal using maximum-likelihood deconvolution. This method was applied to AIDS cases, which feature long delays from infection to observation [47], and in the reconstruction of incidence from mortality times series for the 2009 H1N1 pandemic [45]. It is now being applied to COVID-19 [8, 48]. Fig. 5 shows an example of deconvolution applied following the methods in [45]. The method of [47] is implemented in the backprojNP function within the surveillance R package [49, 50]. In principle, deconvolution can more accurately estimate the latent time series than temporal shifting or backward convolution, but the method is sensitive to misspecification of the mean, variance or form of the delay distribution, and the stringency of the stopping condition of the deconvolution algorithm. It can also be difficult to quantify uncertainty in the deconvolved time series [41] and to implement deconvolution while adjusting for right truncation.

A potential alternative to deconvolution is R_t estimation models that include forward delays to observation in the inference process or that treat the time series of infections as latent states. Such methods are in development within the R package EpiNow2 [41]. An additional advantage of inferring the time series of infections jointly with R_t is seamless integration of various sources of uncertainty, e.g., in R_t and reporting. By comparison, the two-step approach of first transforming the observed time series and then calculating R_t requires users to propagate uncertainty from the back-calculation step into the R_t estimation step. A final advantage of latent state methods is that they could in theory facilitate inference from multiple data streams simultaneously. For example, by assuming that cases, hospitalizations, and deaths all arise from a common infection process, these methods might be able to infer the incident time series of infections more accurately and precisely, potentially while also estimating delays and changes in ascertainment for specific data sources (e.g., outpatient cases).

Deconvolution or R_t estimation methods that include a forward observation process are particularly useful when delays to observation are relatively long and variable, and in analyses that require accurate inference of the timing and speed of changes in R_t. If delays to observation are relatively short, or if R_t is not substantially changing, then deconvolution may not be necessary. For example, when working with synthetic case data in which the mean delays to observation are short and known accurately, the underlying infection curve (Fig. 5A) and underlying R_t values (Fig. 5C) can be recovered with reasonable accuracy simply by shifting the observed time series. But longer and more variable delays to observation worsen inference of the underlying incidence curve (Fig. 5B). In turn, this makes it more difficult to infer the speed and timing of abrupt changes in R_t and to relate those changes to policies, behaviors or other extrinsic epidemic drivers at specific points in time. For example, simply shifting a times series of observed deaths by the mean delay does not accurately recover the underlying curves of infections or R_t (Fig. 5B,C). The marginal value of deconvolution or other methods that infer the unobserved time series of infections is greater when delays to observation are longer and more variable, and when R_t is changing.

Another advantage of working with observations nearer the time of infection, such as times of symptom onset among newly symptomatic individuals, is that they provide more information about recent transmission events and therefore allow R_t to be estimated in closer to real-time (Fig. 5C) [46]. Of course, this advantage could be offset by sampling biases and reporting delays. Users will need to balance data quality with the length of the observation delay when selecting inputs.

Further investigation is needed to determine the best methods for inferring infections from observations if the underlying delay distribution is uncertain. If the delay distribution is severely misspecified, all three approaches (deconvolution, shifting by the mean delay, or convolution) will incorrectly infer the timing of changes in incidence. In

this case, methods such as deconvolution or shifting by the mean delay might more accurately estimate the magnitude of changes in R_t but at the cost of spurious precision in the inferred timing of those changes. Ideally, the delay distribution could be inferred jointly with the underlying times of infection or estimated as the sum of the incubation period distribution and the distribution of delays from symptom onset to observation (e.g. from line-list data).

Summary

- Estimating the instantaneous reproductive number requires data on the number of new infections (i.e., transmission events) over time. These inputs must be inferred from observations using assumptions about delays between infection and observation.
- Inferring the unlagged time series of infections using deconvolution, or within an R_t estimation model that includes forward delays, can improve accuracy.
- A less accurate but simpler approach is to shift the observed time series by the mean delay to observation. If the delay to observation is not highly variable, and if the mean delay is known exactly, the error of this approach may be tolerable. A key disadvantage is that shifting by a fixed amount of time does not account for uncertainty or individual variation in delay times.
- Sampling from the delay distribution to impute individual times of infection from times of observation accounts for uncertainty but blurs peaks and valleys in the underlying incidence curve, which in turn compromises the ability to rapidly detect changes in R_t.

Adjusting for right truncation

Near real-time estimation requires not only inferring times of infection from the observed data but also adjusting for missing observations of recent infections. The absence of recent infections is known as "right truncation". Without adjustment for right truncation, the number of recent infections will appear artificially low because they have not yet been reported [4, 30, 51–55]. Thus, adjusting for right truncation is particularly important in analyses with the goal of near real-time R_t estimation.

Figure 5 illustrates the consequences of failure to adjust for right truncation when inferring times of infection from observations. Subtracting the mean observation delay m from times of observation ("shift" method in Fig. 5A,B) leaves a gap of m days between the last date in the inferred infection time series and the last date in the observed data. This hampers recent R_t estimation (Fig. 5C). Inferring the underlying times of infection by subtracting samples from the delay distribution ("convolve" method in Fig. 5A,B) dramatically underestimates the number of infections occurring in the last few days of the time series.

The simplest approach is to drop estimates on the last few dates, or to flag them as unreliable [8]. But many methods are available to adjust for right truncation, which can improve real-time analysis [41, 51–56]. These methods infer based on past reporting delays the total number of infections, observed and not-yet-observed, at the end of the time series.

In short, accurate near real-time R_t estimation requires both inferring the infection time series from recent observations and adjusting for right truncation. Errors in either step could amplify errors in the other. Joint inference approaches for near real-time R_t estimation, which simultaneously infer times of infection and adjust for right truncation, are now in development [41].

Summary

- At the end of a truncated time series, some infections will not yet have been observed. Infer the missing data to obtain accurate recent R_t estimates.

Accounting for incomplete observation

The effect of incomplete case observation on R_t estimation depends on the observation process. If the fraction of infections observed is constant over time, R_t point estimates will remain accurate and unbiased despite incomplete observation [12, 14, 43, 57, 58]. Data obtained from carefully designed surveillance programs might meet these criteria. But even in this best-case scenario, because the estimation methods reviewed here assume all infections are observed, confidence or credible intervals obtained using these methods will not include uncertainty from incomplete observation. Without these statistical adjustments, practitioners and policy makers should beware false precision in reported R_t estimates.

Sampling biases will also bias R_t estimates [57]. COVID-19 test availability, testing criteria, interest in testing, and even the fraction of deaths reported [59] have all changed over time. If these biases are well understood, it might be possible to adjust for them when estimating R_t. Observed hospitalizations and deaths may be less sensitive to changes in test availability and testing effort, but may be more sensitive to other factors, such as hospital bed availability. Hospitalizations and deaths will also vary in their representativeness of mean transmission rates, depending on which age groups are being infected. No matter what data is used, one potential solution is to flag R_t estimates as potentially biased in the few weeks following known changes in data collection or reporting. At a minimum, practitioners and policy makers should understand how the data underlying R_t estimates were generated and whether they were collected under a standardized testing protocol.

Summary

- R_t point estimates will remain accurate given imperfect observation of cases if the fraction of cases observed is time-independent and representative of a defined population. But even in this best-case scenario, confidence or credible intervals will not accurately measure uncertainty from imperfect observation.

- Changes over time in the type or fraction of infections observed can bias R_t estimates. Structured surveillance with fixed testing protocols can reduce or eliminate this problem.

Smoothing windows

Because R_t estimators incorrectly assume all infections are observed, day-of-week reporting effects and stochasticity in the number of observations per day can cause spurious variability in R_t estimates, especially if the number of observations per day is low [14]. The Cori method incorporates a sliding window to smooth R_t estimates, but the size of the smoothing window can affect both the temporal and quantitative accuracy of estimates. Larger windows effectively increase sample size by drawing information from multiple time points but blur what may be biologically meaningful changes in R_t. Some smoothing approaches can also cause R_t estimates to lead or lag the true value (Fig. 6).

Lags can be particularly severe when using the conventions suggested by Cori et al., in which R_t is reported on the last date in a given window, rather than on the middle date. Although this convention returns R_t estimates to the last date in the time series, which is convenient for real-time estimation, R_t estimates reported at the end of a window are based entirely on data from the past and therefore lag the instantaneous R_t (Fig 6B). Instead, we recommend reporting R_t at the midpoint of the smoothing window (Fig 6A), which produces estimates that are more accurately oriented in time. An apparent disadvantage is that a centered window precludes R_t estimation on the last $w/2$ time units, where w is the width of the window. However, we argue that temporal accuracy is preferable, and that if daily counts are low, failure to produce estimates on the last few days in the time series realistically acknowledges uncertainty about recent trends. Thus, for SARS-CoV-2 and other pathogens with short timescales of infection, near real-time R_t estimation requires large enough daily counts to permit a small window (e.g., a few days).

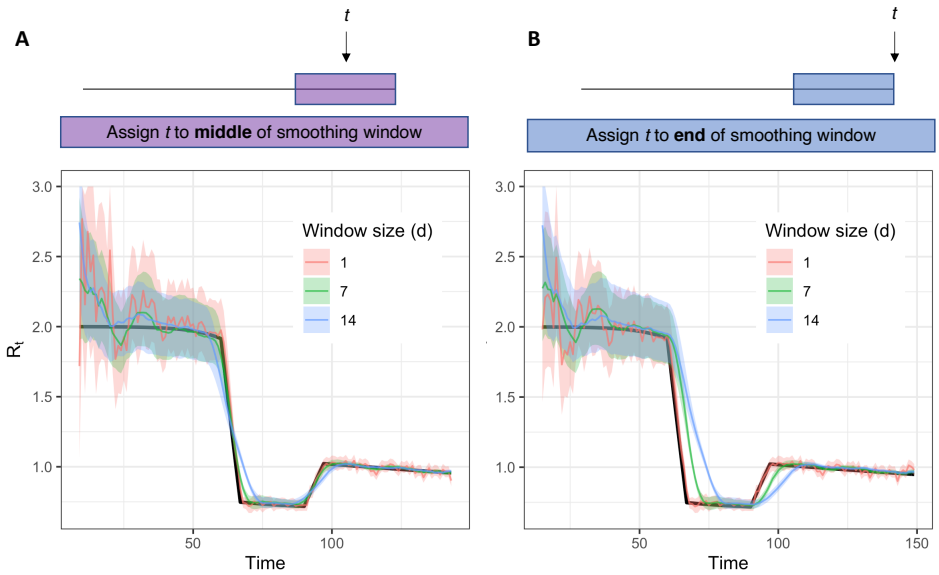


Fig 6. Accuracy of R_t estimates given smoothing window width and location of t within the smoothing window. Estimates were obtained using synthetic data drawn from the $S \to E$ transition of a stochastic SEIR model (inset) as an input to the method of Cori et al. Colored estimates show the posterior mean and 95% credible interval. Black line shows the exact instantaneous R_t calculated from synthetic data.

Although the sliding window increases statistical power to infer R_t, it does not by itself accurately calculate confidence intervals. Thus, underfitting and overfitting are possible. The risk of overfitting in the Cori method is determined by the length of the time window that is chosen. In other words, there is a trade-off in window length between picking up noise with very short windows and over-smoothing with very long ones. To avoid this, one can choose the window size based on short-term predictive accuracy, for example using leave-future-out validation to minimize the one-step-ahead log score [60]. Proper scoring rules such as the Ranked Probability Score can be used in the same way, and a time-varying window size can be chosen adaptively [41].

Summary

- If R_t appears to vary abruptly due to underreporting, a wide smoothing window can help resolve R_t. However, wider windows can also lead to lagged or inaccurate R_t estimates.
- If a wide smoothing window is needed, consider reporting R_t for t corresponding to the middle of the window.
- To avoid overfitting, choose a smoothing window based on short-term predictive accuracy [60] or use an adaptive window [41].

Conclusion

We tested the accuracy of several methods for R_t estimation in near real-time and recommend the methods of Cori et al. [14], which are currently implemented in the R package EpiEstim [20]. The Cori et al. method estimates the instantaneous rather than the case reproductive number and is conceptually appropriate for near real-time estimation. The method uses minimal parametric assumptions about the underlying epidemic process and can accurately estimate abrupt changes in the instantaneous reproductive number using ideal, synthetic data.

Most epidemiological data are not ideal, and statistical adjustments are needed to obtain accurate and timely R_t estimates. First, to obtain timely and temporally accurate R_t estimates, considerable pre-processing is needed to infer the underlying time series of infections (i.e., transmission events) from delayed observations and to adjust for right truncation. Best practices for this inference are still under investigation, especially if the delay distribution is uncertain. The smoothing window must also be chosen carefully, potentially adaptively, and daily counts must be sufficiently high for changes in R_t to be resolved on short timescales. To avoid biases in R_t estimates, the generation interval distribution must be estimated and specified accurately. Finally, to avoid false precision in R_t, uncertainty arising from delays to observation, from adjustment for right truncation, and from imperfect observation must be propagated. The functions provided in the EpiEstim package quantify uncertainty arising from the R_t estimation model but currently not from uncertainty arising from imperfect observation or delays.

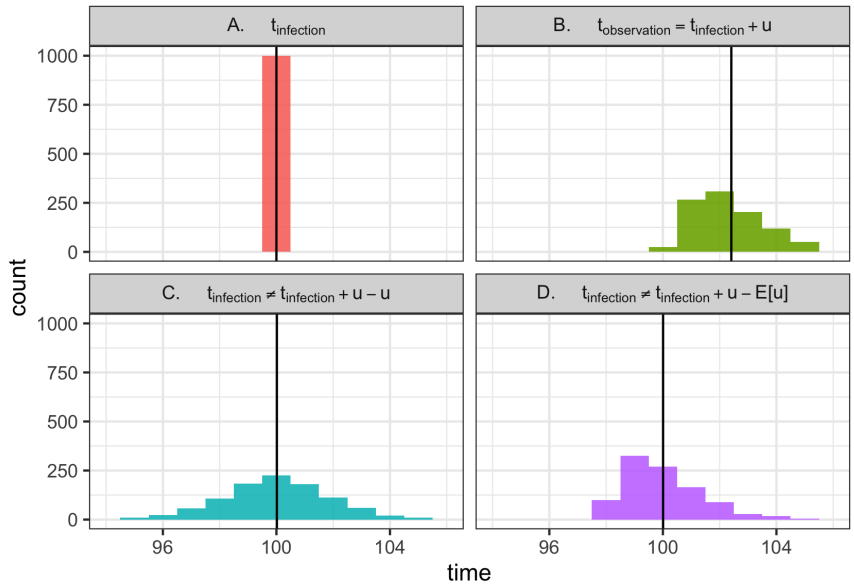
Work is ongoing to determine how best to infer infections from observations and to account for all relevant forms of uncertainty when estimating R_t. Some useful extensions of the methods provided in EpiEstim have already been implemented in the R package EpiNow2 [41, 44], and further updates to this package are planned as new best practices become established.

But even the most powerful inferential methods, extant and proposed, will fail to estimate R_t accurately if changes in sampling are not known and accounted for. If testing shifts from more to less infected subpopulations, or if test availability shifts over time, the resulting changes in case numbers will be ascribed to changes in R_t. Thus, structured surveillance also belongs at the foundation of accurate R_t estimation. This is an urgent problem for near real-time estimation of R_t for COVID-19, as case counts in many regions derive from clinical testing outside any formal surveillance program. Deaths, which are more reliably sampled, are lagged by 2-3 weeks and still subject to biases in underreporting. The establishment of sentinel populations (e.g., outpatient visits with recent symptom onset) for R_t estimation could thus help rapidly identify the effectiveness of different interventions and recent trends in transmission.

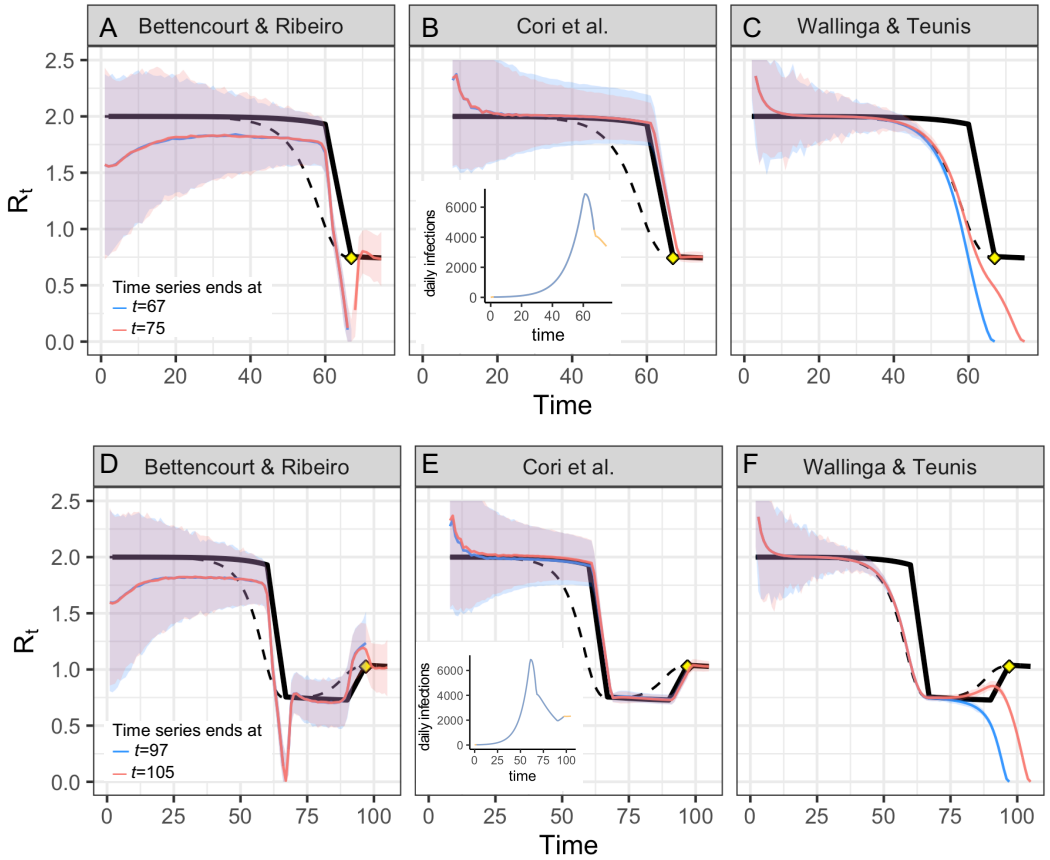
Code availability

All code for analysis and figure generation is available at https://github.com/cobeylab/Rt_estimation.

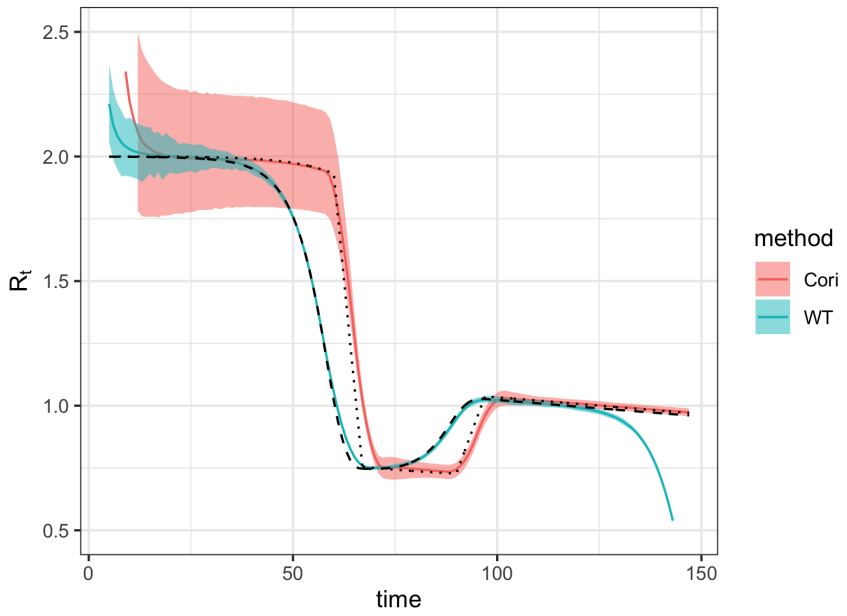
Supporting information



S1 Fig. Why is deconvolution needed to recover latent times of infection? (A) Consider 1000 individuals, all infected at time 100. (Vertical line shows the mean). (B) Now consider the times at which these individuals are observed. Logically, $t_{\text{observation}} = t_{\text{infected}} + u$, where u is a random variable describing the delay between infection and observation. Mathematically, this is a convolution of the infection time and the delay distribution. Because u has non-zero variance, observation times are not only shifted into the future but also are blurred across many dates. This blurring is biologically realistic; due to variability in disease progression and care seeking, individuals with the same date of infection will not necessarily be observed at the same time. (C) Using the observations in B, we aim to recover the latent times of infection shown in A. Doing so would require not only shifting into the past but also removing the variance introduced by the observation process, which can be achieved by deconvolution. Instead, as demonstrated here, a common strategy is to subtract u from the times of observation, effectively repeating the convolution shown in B, but this time moving backward in time rather than forward. This is not the correct inverse operation. It fails to remove variance introduced by the observation process (the forward convolution) and adds new, biologically unrealistic variance, further blurring the inferred times of infection. (D) Shifting the times of observation by the mean delay $E[u]$ is also incorrect, as it does not remove the variance from the forward convolution in B. But if the mean delay time is known exactly, this approach is preferable to C, as it avoids adding even more variance. Ultimately, deconvolution methods would be needed to recover A from the observations in B while properly accounting for uncertainty.



S2 Fig. real-time accuracy when R_t is rising or falling. (A-C) Alternate version of Fig. 2 in which the time series ends on the day R_t first hits its minimum value after falling abruptly (time 67, yellow point), or eight days after the changepoint (time 75). (D-F) The time series ends on the day R_t stops rising (time 97, yellow point), or eight days later (time 105). Estimates of the instantaneous reproductive number (A,B,D,E) remain accurate to the end of the time series, and estimates do not change as new observations become available in the 8 days following the changepoint. As in the main text, estimates of the unadjusted case reproductive number (C,F) depend on data from not-yet-observed time points. These estimates become more accurate as new observations are added to the end of the time series (orange vs. blue). Methods to infer the number of not-yet-observed infections can help make estimates of the case reproductive number more accurate in real-time [4, 30]. All panels show fits to the time series of new infections, and assume all infections are observed instantaneously. Solid black line shows the instantaneous reproductive number, and dashed black line shows the case reproductive number. Colored lines and confidence region show posterior mean and 95% credible interval (A,B,D,E) or maximum likelihood estimate and 95% confidence interval (C,F).



S3 Fig. Smoothed estimates of Cori et al. and Wallinga and Teunis. Both were estimated using a 7-day smoothing window on a synthetic time series of new infections, observed without delay. The estimates of Cori et al. and Wallinga and Teunis are similar in shape when smoothed, but the estimate of Wallinga and Teunis (the case reproductive number) leads that of Cori et al. (the instantaneous reproductive number) by roughly 8 days, or the mean generation interval. Solid colored lines and confidence regions show the posterior mean and 95% credible interval (Cori et al.) or maximum likelihood estimate and 95% confidence interval (Wallinga and Teunis). Dotted and dashed lines show the exact instantaneous reproductive number and case reproductive number, respectively.

Acknowledgements

We are grateful to Michael Höhle for helpful comments. KG was supported by the James S. McDonnell Foundation. LM was supported by the National Institute Of General Medical Sciences of the National Institutes of Health under Award Number F32GM134721. ML acknowledges support from the Morris-Singer Fund and from Models of Infectious Disease Agent Study (MIDAS) cooperative agreement U54GM088558 from the National Institute Of General Medical Sciences. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute Of General Medical Sciences or the National Institutes of Health. LFW acknowledges support from the National Institutes of Health (R01 GM122876). SA, JH, SM, JM, NIB, KS, RNT, SF acknowledge funding from the Wellcome Trust (210758/Z/18/Z). Thanks to Christ Church (Oxford) for funding via a Junior Research Fellowship (RNT). This project has been funded in whole or in part with Federal funds from the National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services, under CEIRS Contract No. HHSN272201400005C (SC).

References

1. Pan A, Liu L, Wang C, Guo H, Hao X, Wang Q, et al. Association of Public Health Interventions With the Epidemiology of the COVID-19 Outbreak in Wuhan, China. JAMA. 2020;323(19):1915. doi:10.1001/jama.2020.6130.

2. Sciré J, Nadeau SA, Vaughan TG, Gavin B, Fuchs S, Sommer J, et al. Reproductive number of the COVID-19 epidemic in Switzerland with a focus on the Cantons of Basel-Stadt and Basel-Landschaft. Swiss Medical Weekly. 2020;150(19-20):w20271.

3. Kucharski AJ, Russell TW, Diamond C, Liu Y, Edmunds J, Funk S, et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. The lancet infectious diseases. 2020;.

4. Cauchemez S, Boëlle PY, Donnelly CA, Ferguson NM, Guy T, Leung GM, et al. Real-time Estimates in Early Detection of SARS. Emerging Infectious Diseases. 2006;12(1):110–113. doi:10.3201/eid1201.050593.

5. Flaxman S, Mishra S, Gandy A, et al. Estimating the number of infections and the impact of nonpharmaceutical interventions on COVID-19 in 11 European countries. Imperial College London; 2020. Available from: `https://doi.org/10.25561/77731`.

6. rt.live.; 2020. Available from: `http://rt.live` [cited 3-June-2020.].

7. covidactnow.; 2020. Available from: `https://covidactnow.org/?s=39636` [cited 3-June-2020.].

8. Effective reproductive number.; 2020. Available from: `https://ibz-shiny.ethz.ch/covid-19-re/` [cited 20-Aug-2020.].

9. Coronavirus disease 2019 Real-time dashboard.;. Available from: `https://covid19.sph.hku.hk/` [cited 3-June-2020].

10. Modeling Covid-19.;. Available from: `https://modelingcovid.com/` [cited 3-June-2020].

11. Lipsitch M, Joshi K, Cobey SE. Comment on Pan A, Liu L, Wang C, et al. Association of Public Health Interventions With the Epidemiology of the COVID-19 Outbreak in Wuhan, China. JAMA. 2020;doi:10.1001/jama.2020.6130.

12. Wallinga J, Teunis P. Different Epidemic Curves for Severe Acute Respiratory Syndrome Reveal Similar Impacts of Control Measures. American Journal of Epidemiology. 2004;160(6):509–516. doi:10.1093/aje/kwh255.

13. Bettencourt LMA, Ribeiro RM. Real Time Bayesian Estimation of the Epidemic Potential of Emerging Infectious Diseases. PLoS ONE. 2008;3(5):e2185. doi:10.1371/journal.pone.0002185.

14. Cori A, Ferguson NM, Fraser C, Cauchemez S. A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics. American Journal of Epidemiology. 2013;178(9):1505–1512. doi:10.1093/aje/kwt133.

15. Thompson RN, Stockwin JE, van Gaalen RD, Polonsky JA, Kamvar ZN, Demarsh PA, et al. Improved inference of time-varying reproduction numbers during infectious disease outbreaks. Epidemics. 2019;29:100356. doi:https://doi.org/10.1016/j.epidem.2019.100356.

16. Dehning J, Zierenberg J, Spitzner FP, Wibral M, Neto JP, Wilczek M, et al. Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. Science. 2020;doi:10.1126/science.abb9789.

17. Lemaitre JC, Perez-Saez J, Azman AS, Rinaldo A, Fellay J. Assessing the impact of non-pharmaceutical interventions on SARS-CoV-2 transmission in Switzerland. Swiss medical weekly. 2020;150(2122):w20295. doi:10.4414/smw.2020.20295.

18. Camacho A, Kucharski AJ, Funk S, Breman J, Piot P, Edmunds WJ. Potential for large outbreaks of Ebola virus disease. Epidemics. 2014;9:70–8. doi:10.1016/j.epidem.2014.09.003.

19. Wallinga J, Lipsitch M. How generation intervals shape the relationship between growth rates and reproductive numbers. Proceedings of the Royal Society B: Biological Sciences. 2007;274(1609):599–604.

20. Cori A. EpiEstim: Estimate Time Varying Reproduction Numbers from Epidemic Curves; 2020.

21. Ganyani T, Kremer C, Chen D, Torneri A, Faes C, Wallinga J, et al. Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data, March 2020. Eurosurveillance. 2020;25(17). doi:https://doi.org/10.2807/1560-7917.ES.2020.25.17.2000257.

22. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. Science. 2020;368(6491). doi:10.1126/science.abb6936.

23. Systrom K. The Metric We Need to Manage COVID-19.; 2020. Available from: `http://systrom.com/blog/the-metric-we-need-to-manage-covid-19/` [cited 3-June-2020].

24. Stan Development Team. RStan: the R interface to Stan; 2020. Available from: `http://mc-stan.org/`.

25. Fraser C. Estimating Individual and Household Reproduction Numbers in an Emerging Epidemic. PLoS ONE. 2007;2(8):e758. doi:10.1371/journal.pone.0000758.

26. Hens N, Calatayud L, Kurkela S, Tamme T, Wallinga J. Robust reconstruction and analysis of outbreak data: influenza A(H1N1)v transmission in a school-based population. American journal of epidemiology. 2012;176(3):196–203. doi:10.1093/aje/kws006.

27. White LF, Archer B, Pagano M. Determining the dynamics of influenza transmission by age. Emerging themes in epidemiology. 2014;11(1):4. doi:10.1186/1742-7622-11-4.

28. Park SW, Champredon D, Weitz JS, Dushoff J. A practical generation-interval-based approach to inferring the strength of epidemics from their speed. Epidemics. 2019;27:12 – 18. doi:https://doi.org/10.1016/j.epidem.2018.12.002.

29. Champredon D, Dushoff J, Earn DJD. Equivalence of the Erlang-Distributed SEIR Epidemic Model and the Renewal Equation. SIAM Journal on Applied Mathematics. 2018;78(6):3258–3278. doi:10.1137/18M1186411.

30. Cauchemez S, Boëlle PY, Thomas G, Valleron AJ. Estimating in Real Time the Efficacy of Measures to Control Emerging Communicable Diseases. American Journal of Epidemiology. 2006;164(6):591–597. doi:10.1093/aje/kwj274.

31. Champredon D, Dushoff J. Intrinsic and realized generation intervals in infectious-disease transmission. Proceedings of the Royal Society B: Biological Sciences. 2015;282(1821):20152026. doi:10.1098/rspb.2015.2026.

32. Park SW, Sun K, Champredon D, Li M, Bolker BM, Earn DJD, et al. Cohort-based approach to understanding the roles of generation and serial intervals in shaping epidemiological dynamics. medRxiv. 2020;doi:10.1101/2020.06.04.20122713.

33. Park SW, Champredon D, Dushoff J. Inferring generation-interval distributions from contact-tracing data. Journal of the Royal Society, Interface. 2020;17(167):20190719. doi:10.1098/rsif.2019.0719.

34. Ali ST, Wang L, Lau EHY, Xu XK, Du Z, Wu Y, et al. Serial interval of SARS-CoV-2 was shortened over time by nonpharmaceutical interventions. Science. 2020;doi:10.1126/science.abc9004.

35. Lipsitch M, Cohen T, Cooper B, Robins JM, Ma S, James L, et al. Transmission dynamics and control of severe acute respiratory syndrome. Science. 2003;300(5627):1966–1970.

36. Britton T, Scalia Tomba G. Estimation in emerging epidemics: biases and remedies. Journal of The Royal Society Interface. 2019;16(150):20180670. doi:10.1098/rsif.2018.0670.

37. Svensson [U+FFFD] A note on generation times in epidemic models. Mathematical Biosciences. 2007;208(1):300–311. doi:10.1016/j.mbs.2006.10.010.

38. Du Z, Xu X, Wu Y, Wang L, Cowling BJ, Meyers LA. Serial Interval of COVID-19 among Publicly Reported Confirmed Cases. Emerging infectious diseases. 2020;26(6). doi:10.3201/eid2606.200357.

39. Tindale L, Coombe M, Stockdale JE, Garlock E, Lau WYV, Saraswat M, et al. Transmission interval estimates suggest pre-symptomatic spread of COVID-19. medRxiv. 2020; p. 2020.03.03.20029983. doi:10.1101/2020.03.03.20029983.

40. Nishiura H, Linton NM, Akhmetzhanov AR. Serial interval of novel coronavirus (COVID-19) infections. International Journal of Infectious Diseases. 2020;doi:10.1016/j.ijid.2020.02.060.

41. Abbott S, Hellewell J, Thompson R, Sherratt K, Gibbs H, Bosse N, et al. Estimating the time-varying reproduction number of SARS-CoV-2 using national and subnational case counts [version 1; peer review: awaiting peer review]. Wellcome Open Research. 2020;5(112). doi:10.12688/wellcomeopenres.16006.1.

42. Moser CB, Gupta M, Archer BN, White LF. The impact of prior information on estimates of disease transmissibility using Bayesian tools. PloS one. 2015;10(3).

43. White LF, Wallinga J, Finelli L, Reed C, Riley S, Lipsitch M, et al. Estimation of the reproductive number and the serial interval in early phase of the 2009 influenza A/H1N1 pandemic in the USA. Influenza and other respiratory viruses. 2009;3(6):267–76. doi:10.1111/j.1750-2659.2009.00106.x.

44. Sam Abbott, Hellewell J, Munday J, Thompson R, Funk S. EpiNow: Estimate Realtime Case Counts and Time-varying Epidemiological Parameters; 2020. Available from: `https://github.com/epiforecasts/EpiNow`.

45. Goldstein E, Dushoff J, Ma J, Plotkin JB, Earn DJD, Lipsitch M. Reconstructing influenza incidence by deconvolution of daily mortality time series. Proceedings of the National Academy of Sciences of the United States of America. 2009;106(51):21825–9. doi:10.1073/pnas.0902958106.

46. wyler d, petermann m. A pitfall in estimating the effective reproductive number Rt for COVID-19. medRxiv. 2020;doi:10.1101/2020.05.12.20099366.

47. Becker NG, Watson LF, Carlin JB. A method of non-parametric back-projection and its application to AIDS data. Statistics in Medicine. 1991;10(10):1527–1542. doi:10.1002/sim.4780101005.

48. Marschner IC. Back-projection of COVID-19 diagnosis counts to assess infection incidence and control measures: analysis of Australian data. Epidemiology and Infection. 2020;148:e97. doi:10.1017/S0950268820001065.

49. Meyer S, Held L, Höhle M. Spatio-Temporal Analysis of Epidemic Phenomena Using the R Package surveillance. Journal of Statistical Software. 2017;77(11):1–55. doi:10.18637/jss.v077.i11.

50. Salmon M, Schumacher D, Höhle M. Monitoring Count Time Series in R: Aberration Detection in Public Health Surveillance. Journal of Statistical Software. 2016;70(10):1–35. doi:10.18637/jss.v070.i10.

51. Lawless JF. Adjustments for reporting delays and the prediction of occurred but not reported events. Canadian Journal of Statistics. 1994;22(1):15–31. doi:10.2307/3315826.n1.

52. McGough SF, Johansson MA, Lipsitch M, Menzies NA. Nowcasting by Bayesian Smoothing: A flexible, generalizable model for real-time epidemic tracking. PLOS Computational Biology. 2020;16(4):e1007735. doi:10.1371/journal.pcbi.1007735.

53. Kalbfleisch JD, Lawless JF. Inference Based on Retrospective Ascertainment: An Analysis of the Data on Transfusion-Related AIDS. Journal of the American Statistical Association. 1989;84(406):360–372. doi:10.1080/01621459.1989.10478780.

54. Höhle M, Heiden Mad. Bayesian nowcasting during the STEC O104:H4 outbreak in Germany, 2011. Biometrics. 2014;70(4):993–1002. doi:10.1111/biom.12194.

55. Kassteele Jvd, Eilers PHC, Wallinga J. Nowcasting the Number of New Symptomatic Cases During Infectious Disease Outbreaks Using Constrained P-spline Smoothing. Epidemiology. 2019;30(5):737–745. doi:10.1097/ede.0000000000001050.

56. Guenther F, Bender A, Katz K, Kuechenhoff H, Hoehle M. Nowcasting the COVID-19 Pandemic in Bavaria. medRxiv. 2020;doi:10.1101/2020.06.26.20140210.

57. Pitzer VE, Chitwood M, Havumaki J, Menzies NA, Perniciaro S, Warren JL, et al. The impact of changes in diagnostic testing practices on estimates of COVID-19 transmission in the United States. medRxiv. 2020;doi:10.1101/2020.04.20.20073338.

58. White LF, Pagano M. Reporting errors in infectious disease outbreaks, with an application to Pandemic Influenza A/H1N1. Epidemiologic perspectives & innovations : EP+I. 2010;7(1):12. doi:10.1186/1742-5573-7-12.

59. Weinberger D, Cohen T, Crawford F, Mostashari F, Olson D, Pitzer VE, et al. Estimating the early death toll of COVID-19 in the United States. medRxiv. 2020;doi:10.1101/2020.04.15.20066431.

60. Parag K, Donnelly C. Optimising Renewal Models for Real-Time Epidemic Prediction and Estimation. bioRxiv. 2019;doi:10.1101/835181.