



# Unjustified Poisson assumptions lead to overconfident estimates of the effective reproductive number

## Original Paper

**Cite this article:** Němcová B, Goldstein IH, Sebastian J, Minin VM and Bracher J (2026). Unjustified Poisson assumptions lead to overconfident estimates of the effective reproductive number. *Epidemiology and Infection*, **154**, e79, 1–8  
<https://doi.org/10.1017/S0950268826101605>

Received: 02 December 2025  
Revised: 23 April 2026  
Accepted: 05 May 2026

### Keywords:

Infectious disease epidemiology; Effective reproductive number; Uncertainty quantification; Renewal equation; Poisson distribution

### Corresponding author:

Barbora Němcová;  
Email: [barbora.sobolova@kit.edu](mailto:barbora.sobolova@kit.edu)

Barbora Němcová<sup>1,2</sup> , Isaac H. Goldstein<sup>3</sup>, Jessalyn Sebastian<sup>4</sup>,  
Volodymyr M. Minin<sup>4</sup> and Johannes Bracher<sup>1,5</sup>

<sup>1</sup>Institute of Statistics, Karlsruhe Institute of Technology, Germany; <sup>2</sup>Helmholtz Information & Data Science School for Health, Helmholtz Association Helmholtz Information & Data Science Academy, Germany; <sup>3</sup>Department of Statistics, Stanford University, USA; <sup>4</sup>Department of Statistics, University of California Irvine, USA and <sup>5</sup>Heidelberg Institute for Theoretical Studies, Germany

### Abstract

Time-varying effective reproductive numbers of infectious diseases are commonly estimated using renewal equation models. In the widely applied R package EpiEstim and various related tools, this approach is combined with a Poisson distributional assumption. This has been criticized on various occasions, mostly on grounds of general model realism or a desire to estimate dispersion parameters. Here, we argue that an important issue arising from the Poisson assumption is that inference about the effective reproductive number becomes overconfident in the presence of overdispersion. The extent to which standard errors are underestimated follows from theory on generalized linear models in a straightforward manner. We therefore recommend to replace the Poisson assumption by quasi-Poisson or negative binomial extensions, and contrast their respective properties. We illustrate our arguments in detailed simulation studies and three case studies on Ebola, pandemic influenza, and COVID-19.

### Introduction

The effective reproductive number, defined as the expected number of new infections generated by a single infectious individual in a population that is no longer fully susceptible [1], is a key quantity characterizing disease outbreak dynamics. It is commonly estimated using renewal equation models, often combined with Poisson distributional assumptions. This approach was popularized by the R package EpiEstim [2], which due to its broad applicability, ease of use, and thorough documentation has become a standard tool [3]. It has also been integrated in numerous methodological extensions (e.g., [4–6]) and alternative R packages including EpiInvert [7], ern [8], estimateR [9], and most recently Rt.GLM [10]. In this article, we are concerned with the uncertainty surrounding estimated reproductive numbers based on the renewal equation and a Poisson assumption for the observed incidence. We demonstrate that the associated standard errors are underestimated in the presence of overdispersion, a common feature of infectious disease surveillance data. Especially when reported incidence is high, uncertainty intervals based on Poisson assumptions get vanishingly narrow, overstating the certainty of the inference. While this has been pointed out empirically [11–13], we provide a concise summary of relevant statistical properties. We moreover contrast several methods to account for overdispersion, which turn out to impact not only estimated standard errors, but also point estimates.

Transmission heterogeneity, superspreading and overdispersion are widely studied phenomena in the epidemic modelling literature. A number of recent contributions have been made with respect to renewal equation modelling [14–17]. Many of them, most explicitly Zhang et al. and Pakkanen et al. [18, 19], link this to branching process models and the seminal work by Lloyd-Smith [20]. Branching processes accommodate individual-level heterogeneity in a very natural way, and with specific distributional choices can be shown to be equivalent to renewal equation models [18].

Gostic et al. [1, p. 3] argue that appropriate confidence intervals rather than mere point estimates are needed to assess whether the effective reproductive number has crossed the threshold value of one, and discuss various sources of uncertainty. The point of our paper is that even the best understood source of uncertainty – classic statistical estimation uncertainty – is handled insufficiently by many commonly used methods, and discuss the necessary remedies. This matters when estimating and interpreting reproductive numbers is the primary goal, but also when estimates feed into multi-step analyses. Overconfidence may then permeate into the downstream parts (see [21] for a general discussion). In renewal equation-based forecasting, it is therefore common to propagate uncertainty about the reproductive number using a multiple imputation approach [22]. This seems to be less common in regression modelling exercises

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.



linking estimated reproductive numbers to intervention measures, with, for example, Li and colleagues [23] ignoring the associated estimation uncertainty. However, as pointed out, for example, by Lison et al. [24], this neglects part of the uncertainty and may lead to unwarranted conclusions about the effectiveness of intervention measures. The underlying statistical issues have been discussed in detail in the econometric literature on *generated outcome variables* [25].

The remainder of the paper outlines the challenges arising from the presence of overdispersion in the classic Poisson renewal equation model for estimating the effective reproductive number, along with possible model extensions. We illustrate our arguments using simulations and case studies on COVID-19, Ebola, and pandemic influenza.

## Methods

### The Poisson renewal equation model

While the effective reproductive number can be defined in multiple ways [26], we focus on the *instantaneous* effective reproductive number, which is considered particularly suitable for real-time monitoring [1]. It is defined as the expected number of secondary infections occurring at time  $t$  that are generated by individuals infected before  $t$ , with each infective's contribution weighted by their relative infectiousness at time  $t$ . The effective reproductive number is often denoted by  $R_e$  or  $R_t$ , as its variation over time is what makes real-time estimation relevant. However, in EpiEstim and related approaches, this temporal variation is not modelled explicitly. Instead, the reproductive number is assumed to be constant within estimation windows, usually of length  $T = 7$  days. We hence drop the index  $t$  and write simply  $R$ , while indexing the days in the estimation window by  $t = 1, \dots, T$ . It is implicitly understood that the estimation exercise will be repeated for a set of rolling time windows, leading to a trajectory of estimates.

Daily or weekly incidence counts  $X_t, t = 1, \dots, T$  are linked to  $R$  via the renewal equation [27] and a Poisson assumption,

$$X_t | X_{t-1}, \dots, X_0 \sim \text{Pois}(\mu_t), \mu_t = R \times \Lambda_t, \Lambda_t = \sum_{d=1}^D w_d \times X_{t-d}. \quad (1)$$

Here,  $\mu_t$  is the expectation of the Poisson distribution, while the probabilities  $w_1, \dots, w_D$  represent a known discrete-time serial interval distribution. Serial intervals are expressed in the same time steps as incidence counts, and are assumed to have a maximum possible value of  $D$ . We note that the distributional assumption holds conditionally on past values  $X_{t-1}, \dots, X_0$ , on which  $X_t$  depends through  $\Lambda_t$  (see Web Appendix 1 of [2]). The Poisson distribution is chosen as it makes Bayesian inference straightforward without the need for Monte Carlo methods [2].

Throughout the paper, we refer to  $w_1, \dots, w_D$  as the distribution of the serial interval (the time between the onset of symptoms in a pair of consecutive cases) rather than the generation time (the time between two consecutive infections), as the exact times of infection are typically unobserved. It is known that using the serial interval distribution, which has a larger variance than the generation time distribution, introduces some light bias in the estimation of reproductive numbers [28]. However, this has become standard practice [2], and we do not explore this aspect in detail.

Equation (1) can be read as a Poisson regression with an identity link, no intercept and  $\Lambda_t$  as the sole covariate [12, 13]. The properties of the resulting estimates can hence be studied in the

framework of generalized linear regression [29]. This is most straightforward for maximum likelihood (ML) estimates; the results, however, also translate to the Bayesian approach by [2], which behaves similarly unless incidence is very low (Supplement B.1.2). The ML estimator takes the intuitive form

$$\hat{R}_{\text{Po}} = \frac{\sum_{t=1}^T X_t}{\sum_{t=1}^T \Lambda_t}, \quad (2)$$

where the index 'Po' indicates the underlying distributional assumption. The estimated standard error

$$\hat{\text{se}}(\hat{R}_{\text{Po}}) = \frac{\hat{R}_{\text{Po}}}{\sqrt{\sum_{t=1}^T X_t}} \quad (3)$$

depends on  $\hat{R}_{\text{Po}}$  and the total incidence over the estimation window, but not the variability of the time series conditioned on past values. This is a consequence of the equidispersion property of the Poisson distribution, which implies  $\sigma_t^2 = \mu_t$ , with  $\sigma_t^2$  the conditional variance  $\text{Var}(X_t | X_{t-1}, \dots, X_0)$ . Derivations of equations (2) and (3) are provided in Supplement B.1.1, where we also discuss the challenges due to the dynamic nature of model (1).

The estimated standard error (3) can be used to derive Wald-type confidence intervals, which for a confidence level of 95% are given by  $\hat{R}_{\text{Po}} \pm 1.96 \times \hat{\text{se}}(\hat{R}_{\text{Po}})$ . Empirical coverage fractions of such intervals in simulation studies are a straightforward way to assess uncertainty quantification under varying conditions. We will address this in Section 'Simulation study'.

### Overdispersion and how to account for it

#### Impact on the Poisson estimator

In practice, infectious disease incidence data usually display overdispersion relative to the Poisson model. Throughout the paper, we conceive this as conditional variances exceeding conditional means,  $\sigma_t^2 > \mu_t$ . Overdispersion may be due to a variety of factors, including superspreading events [20] and incomplete reporting [30]. The Poisson assumption will then cause our inference for  $R$  to be overconfident; see Supplementary Section A for an intuitive illustration. Well-known theory on generalized linear regression [29, Sec. 5.5] implies that expression (3) underestimates the true standard error of  $\hat{R}_{\text{Po}}$  by a factor of  $\sqrt{\phi}$ , with

$$\phi = \frac{\sum_{t=1}^T \sigma_t^2}{\sum_{t=1}^T \mu_t}. \quad (4)$$

The *variance inflation factor*  $\phi$  depends on the relationship between  $\mu_t$  and  $\sigma_t^2$ . We note that equation (4) holds irrespective of the exact form of this relationship, even though in practice one usually makes more specific assumptions (see the next section). As we shall see in Section 'Illustration for three outbreaks',  $\phi$  may be quite large in practice, meaning that standard errors estimated under the Poisson assumption may be considerably too small.

Translating equation (4) to the width of the resulting confidence interval, a  $(1 - \alpha)$  confidence interval from the Poisson renewal equation model has coverage probability

$$1 - 2 \times \Phi \left[ \Phi^{-1} \left( \frac{\alpha}{2} \right) / \sqrt{\phi} \right], \quad (5)$$

where  $\Phi$  is the cumulative distribution function of the standard normal distribution. This matches the intended level  $(1 - \alpha)$  only if

$\phi = 1$ , and decreases if  $\phi > 1$ . The derivation of the coverage probability (5) is provided in Supplement B.2.2.

In the following, we will outline three options to account for overdispersion, all of which have been applied in the literature. While this is typically motivated by the desire to estimate dispersion parameters [16, 18] and general model realism [31], our focus is on uncertainty quantification.

**Quasi-Poisson regression**

Quasi-Poisson regression, used by [13], serves to correct the uncertainty quantification for the Poisson model by estimating the variance inflation factor  $\phi$  from equation (4). Assuming a linear relationship between  $\mu_t$  and  $\sigma_t^2$ , it can be estimated from the residuals of the Poisson model (see Supplement B.2.1). The corrected standard error is then obtained by multiplying expression (3) with  $\sqrt{\hat{\phi}}$ . We note that quasi-Poisson regression is not a full probability model, making it unsuitable, for example, for likelihood-based model comparisons and for sampling future trajectories.

**Negative binomial regression**

The most common full probability model to generalize the Poisson assumption is the negative binomial distribution. We here parameterize it by its mean and a dispersion parameter, which, for example, in the R software is called the ‘size parameter’. Despite its name, a higher value of the dispersion parameter indicates lower dispersion (see below for formulas), and the Poisson distribution is the limiting case of an infinite dispersion parameter. The dispersion parameter can be handled in two ways ([32], p.73), which we call *NegBin-L* and *NegBin-Q*, as they imply *linear* and *quadratic* mean–variance relationships. Irrespective of which is used, estimates and standard errors for  $R$  are obtained using standard regression methodology and naturally account for overdispersion.

**NegBin-L.** In the NegBin-L renewal equation (see, e.g., [16, 18]),

$$X_t | X_{t-1}, \dots, X_0 \sim \text{NegBin}(\text{mean} = \mu_t, \text{disp} = \zeta \times \mu_t), \quad (6)$$

the dispersion parameter is assumed to be proportional to the mean. This implies a linear mean–variance relationship  $\sigma_t^2 = (1 + 1/\zeta) \times \mu_t$  as in the quasi-Poisson model, and can be motivated by overdispersed individual offspring distributions [18]. While some differences between NegBin-L and quasi-Poisson regression exist [33], in our applications we found similar behaviour, with point estimates almost identical to those from equation (2).

**NegBin-Q.** We therefore focus primarily on NegBin-Q, defined as

$$X_t | X_{t-1}, \dots, X_0 \sim \text{NegBin}(\text{mean} = \mu_t, \text{disp} = \psi), \quad (7)$$

and implying a quadratic mean–variance relationship  $\sigma_t^2 = \mu_t + \mu_t^2/\psi$ . This version, employed, for example, by [12, 18, 31, 34], can be motivated by temporal variation of  $R$  over the estimation window [18]. Despite the mean structure  $\mu_t = R \times \Lambda_t$  mirroring the Poisson version, the ML estimator  $\hat{R}_{\text{NBQ}}$  is not identical to  $\hat{R}_{\text{Po}}$  from equation (2). No closed expression exists, but if overdispersion is sufficiently strong, the approximation

$$\hat{R}_{\text{NBQ}} \approx \frac{1}{T} \times \sum_{t=1}^T \frac{X_t}{\Lambda_t} \quad (8)$$

holds; see derivation in Supplement B.3.1. Unlike in equation (2), the ratios  $X_1/\Lambda_1, \dots, X_T/\Lambda_T$  thus contribute equally to the estimate. This is because they are all equally informative under a quadratic

mean–variance relationship, irrespective of the magnitude of  $\Lambda_t$  and  $X_t$  (see discussion in [35]). This also explains why the approximate standard error

$$\widehat{\text{se}}(\hat{R}_{\text{NBQ}}) \approx \hat{R}_{\text{NBQ}} \times \sqrt{(\hat{\psi} \times T)^{-1}} \quad (9)$$

features the number of observations  $T$  rather than the total incidence as in equation (3). It moreover depends on the estimated dispersion parameter  $\hat{\psi}$ , which reflects the conditional variability of the time series and can be computed from the model residuals (Supplement B.3). It generally holds that  $\widehat{\text{se}}(\hat{R}_{\text{NBQ}}) \geq \widehat{\text{se}}(\hat{R}_{\text{Po}})$ , that is, the estimated standard error is at least equal to its Poisson counterpart. The approximations (8) and (9) are assessed empirically in Supplementary Figure S22.

ML estimation in the negative binomial renewal equation models (6) and (7) can run into convergence issues if the data used for fitting show no or very little overdispersion. This is because the dispersion parameter  $\zeta$  in the NegBin-L model (or  $\psi$  in the NegBin-Q) is usually estimated on a log-transformed scale to ensure positivity. If the highest likelihood is achieved for a very high value of  $\zeta$ , the likelihood as a function of  $\log(\zeta)$  becomes almost flat for large  $\log(\zeta)$ . This causes numerical failure in the computation of point estimates and/or standard errors. Whenever this occurs, we revert to the Poisson model (1), and report the corresponding point estimates and standard errors. In such low-dispersion settings, reverting to the Poisson is numerically convenient and statistically appropriate, as the Poisson distribution can be understood as the limiting form of the negative binomial when overdispersion vanishes.

Quasi-Poisson regression has certain advantages over the negative binomial versions if the mean–variance relationship is misspecified [36]. In our setting, however, both equations (2) and (8) remain unbiased and consistent in this case, and only their efficiency is reduced. How strongly the two estimators differ is an empirical question we will return to in Section ‘Illustration for three outbreaks’.

**Comparing model fits and testing for the presence of conditional overdispersion.** To assess whether and to what degree the negative binomial models improve the fit relative to the Poisson, we compare the AIC values and employ likelihood ratio (LR) tests. The quasi-Poisson model cannot be included in this comparison, as it does not have a proper likelihood function. In the LR test, we test the null hypothesis of no conditional overdispersion. Specifically, we test  $H_0 : 1/\zeta = 0$  (NegBin-L) and  $H_0 : 1/\psi = 0$  (NegBin-Q) against one-sided alternatives. To account for the fact that the null hypotheses are on the border of the parameter space we employ the correction suggested by Molenberghs and Verbeke [37].

**Setup of the simulation study**

To support our claim that Poisson assumptions lead to overconfidence in the estimation of  $R$  we conduct a simulation study. We combine different distributional assumptions for the data-generating process and the model used for  $R$  estimation, and compute empirical coverage fractions of Wald confidence intervals. These are given by the fraction of cases in which confidence intervals with a given nominal level contain the true reproductive number  $R$ .

In the main manuscript, we focus on time series generated from the NegBin-L renewal equation (6). We generate 1,000 trajectories from each of four scenarios, which are defined by a combination of two parameters: two values of the reproductive

number  $R \in \{1.5, 2.5\}$  and two levels of overdispersion  $\xi \in \{2, 0.2\}$ . For all scenarios, we used the same shifted gamma serial interval distribution with a mean of 7.5 days and a standard deviation of 2.1 days, cut off at  $D = 14$  days. All trajectories are initialized by the same sequence of values (2, 7, 6, 4, 3, 7, 7, 6, 6, 6, 4, 9, 3, 7). In [Supplementary Section C.2.1](#), we include additional simulation scenarios, where the trajectories are initialized by higher values around 100 in order to produce different trajectories with overall higher incidence. After this deterministic initialization, we generate random trajectories of length 28 days. The first 14 days serve to avoid dependence on the specific initialization sequence ('burn-in period'), while the last 14 days are used to estimate the value of  $R$ . We emphasize that the true  $R$  is a constant throughout the 28 days, such that in terms of the mean structure the assumptions shared by all considered models are fulfilled and the theoretical arguments from the previous sections apply. While in reality, smooth variation of  $R$  would be expected, our simulation study is concerned with the behaviour of EpiEstim-like methods if all assumptions other than the Poisson distribution are fulfilled. It therefore inherits the constraint of time-constant  $R$  values. An adapted simulation where the reproductive number varies smoothly over time is available in [Supplementary Section C.2.3](#).

The serial interval is motivated by the estimate [38] provide for respiratory syncytial virus (RSV; see also [39]). In [Supplementary Section C.2.2](#), we provide further simulation scenarios with serial intervals that are motivated by influenza (mean 3.7 days, sd 2.1 days [40]) and measles (mean 13.7 days, sd 1.5 days [41]).

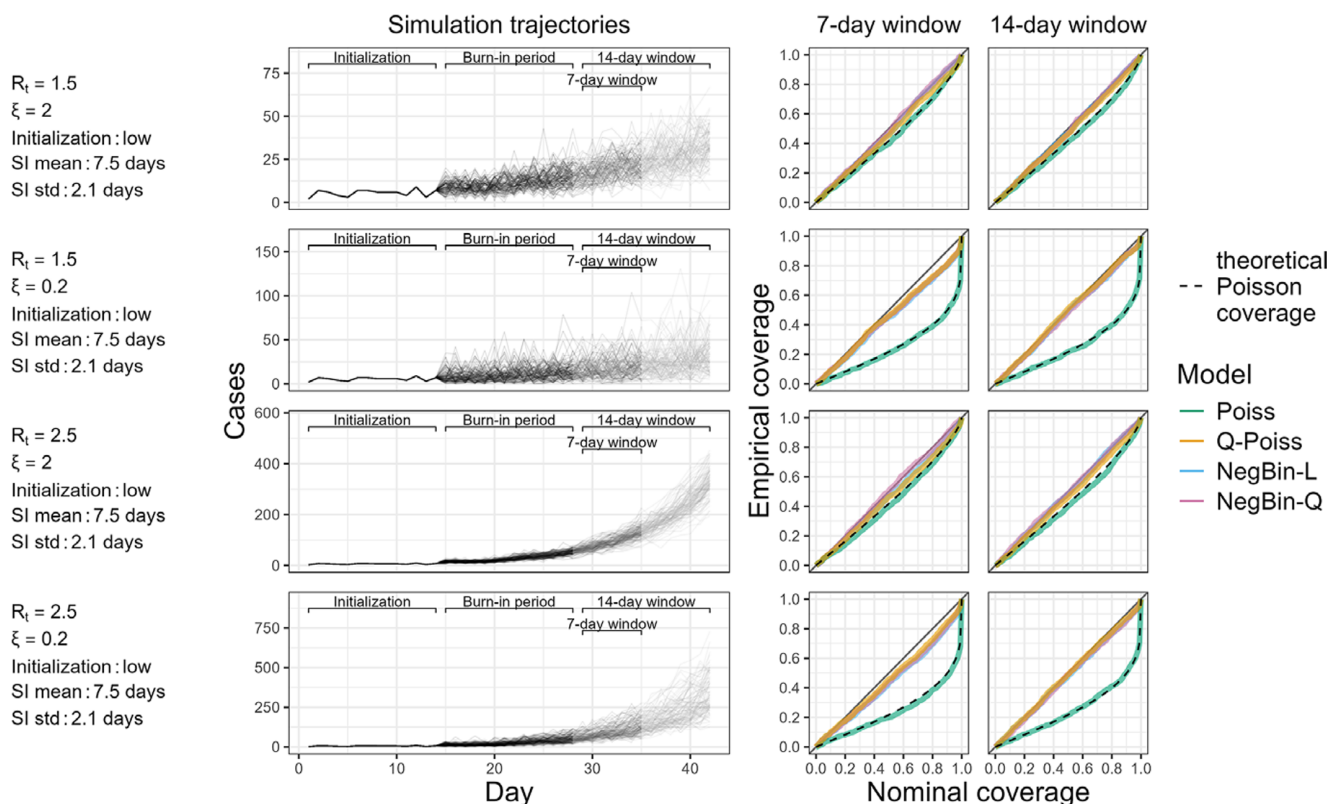
For each simulated trajectory,  $R$  values were estimated under the four distributional assumptions discussed in Section 'Overdispersion and how to account for it'. The Poisson and quasi-Poisson models were fitted using the `stats::glm` function in R [42], while `gamlss::gamlss` [43] was employed for the negative binomial models. To assess the impact of small sample biases, we applied two different estimation window lengths (7 and 14 days) across all scenarios. As described in Section 'Negative binomial regression', in cases of convergence issues due to low degrees of observed overdispersion, we reverted from the negative binomial models to the Poisson version.

To make our assessment comprehensive, we ran corresponding simulation studies using the NegBin-Q and Poisson versions of the model as the data-generating processes. In addition, and similarly to [44], we study a setting where none of the considered models matches the data-generating process. To this end, we simulated time series from an overdispersed branching process model. Details on all of these settings are provided in [Supplementary Section C](#).

## Results

### Simulation study

[Figure 1](#) shows the sampled trajectories in the different NegBin-L scenarios (left column), along with how often confidence intervals produced under different distributional assumptions cover the true value of  $R$  (right column). These coverage proportions are shown for the two window lengths (7 and 14 days) and as a function of the



**Figure 1.** Simulation trajectories generated by the NegBin-L model with a serial interval typical to RSV and empirical coverage of the four models in scenarios initialized by low values (see [Supplementary Figure S3](#) for corresponding results with higher initial values). The left panel shows 1,000 incidence trajectories generated from the renewal equation with the NegBin-L distribution across four scenarios. These are defined by different parameter combinations specified on the left margin of the figure. The right panels display empirical coverage of the true  $R$  for Poisson, quasi-Poisson and negative binomial models across nominal coverage levels, using 7-day (middle column) and 14-day (right column) estimation windows. Dashed black lines indicate our theoretical expectation for the coverage of the Poisson model based on [equation \(5\)](#).

nominal coverage level  $(1 - \alpha)$ . Perfect calibration corresponds to  $(1 - \alpha) \times 100\%$  confidence intervals covering the true value of  $R$  in  $(1 - \alpha) \times 100\%$  of cases for all  $1 \leq \alpha \leq 1$ . The plotted curve should therefore closely follow the diagonal. Values above indicate over-coverage, while values under the diagonal point to undercoverage and hence overconfidence.

From Figure 1, it is clear that the Poisson model consistently undercovers. The undercoverage is relatively mild in scenarios with low overdispersion (e.g., for  $R = 1.5$ ,  $\zeta = 2$ , the 90% confidence interval covers the true value in 82% of the cases), but becomes quite substantial as overdispersion increases. This is especially evident at high nominal coverage levels. For instance, for  $R = 1.5$ ,  $\zeta = 0.2$ , the actual coverage of the Poisson model reaches just 50% for a nominal level of 90%. The observed coverage level of the Poisson model is well-approximated by our theoretical expectation based on formula (5), which is shown as a dashed black line. This underscores the validity of our theoretical arguments.

In contrast, the three models accounting for overdispersion generally exhibit similar coverage behaviour with the empirical coverage close to the nominal levels. It thus does not seem crucial whether the mean–variance relationship is specified correctly. A slight tendency to undercover can be spotted for higher nominal coverage levels and larger dispersion values. This issue is less pronounced when the longer 14-day window is used, suggesting the effect of small sample biases when using the 7-day estimation window.

We note that in 18.7% of simulation runs, the fitting algorithm of the negative binomial models did not converge, making it necessary to revert to the Poisson model. In line with our reasoning from Section ‘Negative binomial regression’, this occurred almost exclusively in the scenarios with low overdispersion.

For completeness, we display the empirical distributions of the estimated reproductive number  $\hat{R}$  and the associated standard errors in Supplementary Figure S2, as well as the empirical distribution of the dispersion parameter  $\hat{\zeta}$ , in Supplementary Figure S17. Results when using higher values for initialization, or different serial intervals are similar to the ones discussed above and included in Supplementary Sections C.2.1 and C.2.2, respectively. Simulation scenarios where the reproductive number varies smoothly over time are discussed in Supplementary Section C.2.3, again without any qualitative differences in results. In Supplementary Section C.2.4, we show results for data generated from the NegBin-Q rather than NegBin-L process. Again, it is not crucial whether the mean–variance relationship is specified correctly, and the NegBin-Q, NegBin-L and quasi-Poisson models all perform reasonably. Incidence time series from the Poisson model and associated estimates are shown in Supplementary Section C.2.5. Especially in the high-incidence settings, the resulting bundles of trajectories are very tight, demonstrating that the Poisson distribution in such settings implies an unrealistically low degree of variability. Confidence intervals from both negative-binomial models become somewhat conservative in this setting, a finding we explain in more detail in Section C.2.5 of the Supplement. Finally, Supplementary Section C.2.6 shows results for data generated from a branching process with an overdispersed

offspring distribution, instead of the renewal equation model. Again, all qualitative patterns remain unchanged.

### Illustration for three outbreaks

To illustrate the impact of the distributional choice in real-world settings, we present three case studies covering different pathogens and locations. These are based on previous applications of EpiEstim summarized in Table 1. We use shifted gamma serial interval distributions with means and standard deviations retrieved from the respective references. For the influenza and COVID-19 examples, we also use the corresponding estimation window length. For the Ebola data, we present a variation of the analysis from [5]. We aggregate the daily data to weekly totals, rescale the serial interval distribution and employ an estimation window of 4 weeks. This illustrates our point in a setting, where the incidence data are in a weekly resolution and the serial interval is long, so that the times between symptom onsets can likewise be considered in a weekly resolution. As in the simulation study, we show Wald confidence intervals in all cases.

Figure 2 shows the three incidence time series (top row) and  $R$  estimates based on the different distributional assumptions. The second and third rows show the Poisson and quasi-Poisson models, respectively. While the point estimates agree by definition, the Poisson version yields much narrower 95% confidence intervals. For influenza and COVID-19, they are visually indiscernible from the point estimate. For Ebola, the intervals are wider, as suggested by equation (3) for lower incidence counts. Nonetheless, they rarely include  $R = 1.0$ , implying high confidence about epidemic growth or decline. The intervals from the quasi-Poisson model (shown in orange), on the other hand, often contain values above and below  $R = 1.0$ .

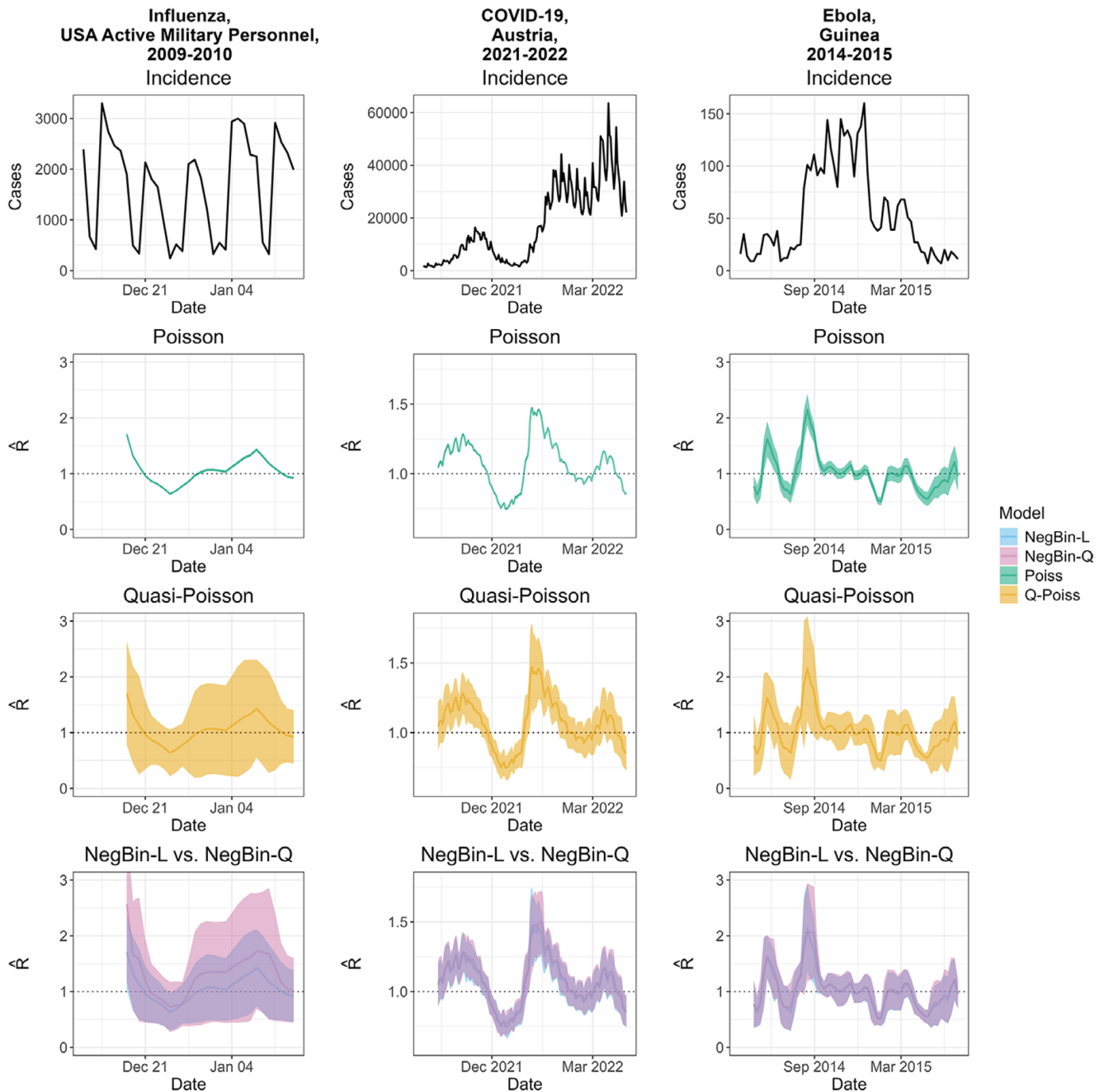
The fourth row overlays results from the NegBin-L model (which behaves similarly to the quasi-Poisson model; Supplementary Figure S21), and the NegBin-Q model. The confidence intervals are comparable for Ebola and COVID-19. For influenza, the NegBin-Q intervals are wider, and there are considerable differences in the point estimates. These are due to the different weighting of the observations in estimators (2) and (8), in conjunction with weekday effects. We discuss this in detail in Supplementary Section A. We note that weekday effects also occur to a lesser degree in the COVID-19 example, but the resulting differences in estimates and confidence intervals are minor.

Ignoring weekday effects can be argued to yield a mis-specified model, and indeed this aspect receives some discussion in [6]. The resulting unexplained variability is absorbed by the dispersion parameters, leading to confidence intervals that may become overly wide. Despite these issues, we kept the influenza example to illustrate the difference between NegBin-L and NegBin-Q, as well as the difficulty of fitting time-homogeneous models to daily data.

In Table 2, we report the average AIC values for the Poisson and negative binomial models across all estimation windows. We additionally calculate the proportion of windows in which the negative

**Table 1.** Summary of the case studies

Reference	Disease	Location	Year	Resolution	Serial interval mean (sd)	Estimation window width $T$
[6], data from [45]	Pandemic influenza	US (military personnel)	2009–2010	Daily	3.6 (1.6) days	7 days
[46]	COVID-19	Austria	2021–2022	Daily	3.37 (1.83) days	13 days
[5]	Ebola	Guinea	2014–2015	Weekly	2.19 (1.33) weeks	4 weeks



**Figure 2.** Incidence and estimated effective reproductive numbers for three outbreaks: influenza among active military personnel in the USA, 2009–2010 (left column); COVID-19 in Austria, 2021–2022 (middle); and Ebola in Guinea, 2014–2015 (right). The first row shows case incidences, which are in daily resolution for influenza and COVID-19, and in weekly resolution for Ebola. The second and third rows show the  $\hat{R}$  estimates from the Poisson and quasi-Poisson models, respectively. All estimates are aligned with the last day of the respective estimation window. The fourth row shows a comparison of the  $\hat{R}$  estimates from the NegBin-Q and NegBin-L models. An additional plot overlaying the quasi-Poisson and NegBin-L versions as well as the estimated dispersion parameters are available in [Supplementary Figures S20 and S21](#).

**Table 2.** Summary of the model comparisons using AIC values and the likelihood ratio test to detect conditional overdispersion

Disease	Poisson		NegBin-L		NegBin-Q		
	Mean AIC	Mean AIC	% lower AIC than Poisson	% LR-test rejections	Mean AIC	% lower AIC than Poisson	% LR-test rejections
Pandemic Influenza	7,517.84	122.74	100.00	100.00	122.06	100.00	100.00
COVID-19	11,926.48	243.77	100.00	100.00	243.98	100.00	100.00
Ebola	44.73	35.44	79.03	75.81	35.53	77.42	75.81

binomial versions have lower AIC than the Poisson model, and how often the LR test rejects the null hypothesis of no conditional overdispersion. For COVID-19 and influenza, the average AIC values are drastically lower for the negative binomial models and the window-wise AIC values are lower in every instance. The LR test likewise rejects the null hypothesis in every single case. For Ebola, the differences in AIC are more moderate due to the lower magnitude of the Ebola incidence and the reduced number of observations per window, yet both the AIC comparison and the LR test favour the negative binomial models in more than 75% of the cases. Boxplots of the AIC values are provided in [Supplementary Figure S19](#).

## Discussion

We argued that unjustified Poisson assumptions lead to overconfident estimation of reproductive numbers and discussed options to account for overdispersion. The quasi-Poisson and NegBin-L versions agreed well in terms of point estimates and uncertainty intervals. The NegBin-Q approach yielded somewhat different results, but in simulation studies, we found that uncertainty quantification was reliable even if the assumed mean–variance relationship was incorrect. In applications to influenza, COVID-19 and Ebola, all three extensions yielded considerably widened uncertainty intervals relative to the Poisson model. The wider intervals included the threshold  $R=1.0$  considerably more often than the Poisson intervals. Differences between estimates from the NegBin-L and quasi-Poisson models on the one hand and the NegBin-Q model on the other hand turned out to be substantial in the case study on influenza. This, however, could be attributed to unaccounted weekday effects, meaning that such discrepancies are indicative of more general model mis-specification issues.

In light of these results, we strongly recommend to account for overdispersion in the estimation of reproductive numbers. While the conjugate prior approach from [2] is difficult to extend, standard regression functions for likelihood-based inference can be adapted. We are aware that an MCMC-based extension of EpiEstim to account for overdispersion in a Bayesian framework is under development, and we recommend it to future users.

Alternative tools like EpiNow2 [47] and EpiLPS [31] already account for overdispersion. EpiNow2 further avoids the need for estimation windows and handles delays between infection events and reporting. This makes estimates more aligned with the actual infection dynamics. Due to its increased complexity, however, EpiNow2 requires more tuning and computational effort, which may present an obstacle to some users. Alternatively, users constrained to renewal equation models with sliding windows can leverage standard regression functionality for quasi-Poisson and negative binomial models. As visible from our accompanying codes (see code availability statement below) and [Supplementary Section E](#), this can be accomplished in a few lines of code.

Our work focuses on correctly quantifying the statistical uncertainty about  $R$ . In statistics, there is sometimes a preference to err on the conservative side, and hence prefer too wide over too narrow uncertainty intervals. It could be argued, though, that in a broader decision-making context there is a price to this, as decision-making may be slowed down by overly stringent criteria. A related discussion on converting  $R$  estimates and their uncertainty into decision rules can be found in [48].

We conclude by discussing some limitations of our arguments and suggested extensions. By using off-the-shelf generalized linear models, we are deliberately agnostic to the mechanisms behind

overdispersion and conflate different sources, such as superspreading, underreporting, and other reporting artefacts. It is thus difficult to assign any mechanistic interpretation to the estimated dispersion parameters, and they cannot serve to derive implications on disease emergence and control [20]. We note that models explicitly separating different sources of overdispersion exist [30], but their identifiability usually hinges on certain quantities being known from external sources (e.g., reporting probabilities).

In the examples of influenza in US military personnel and (to a lesser degree) COVID-19 in Austria, the model appeared to be mis-specified due to weekday patterns. These may inflate the estimated dispersion parameters and standard errors. More generally, estimating the dispersion parameter along with  $R$  from short time windows may lead to instabilities. For negative binomial regression, it is known that small-sample biases can occur [49], with a tendency to underestimate standard errors. Bias-corrected estimators exist and could be used instead of our ML estimators, but we do not expect this to have an impact on our conclusions. Finally, many other layers of complexity like unobserved initial generations in an outbreak [4] and changes in testing volumes [13] could be addressed. However, we consider these outside the scope of the present article.

A promising regression-based approach replacing sliding window approach by splines has recently been proposed by Nouvellet [10], and can remedy some of the aforementioned weaknesses. It estimates the temporal variation in  $R$  flexibly, potentially accounting for covariates like weekday effects. Unfortunately, the new framework currently only supports the Poisson distribution, and is thus likewise affected by the pitfalls we described. Extensions, however, could easily be implemented, potentially even allowing for time-varying dispersion parameters via the `gamlss` framework [43].

**Supplementary material.** The supplementary material for this article can be found at <http://doi.org/10.1017/S0950268826101605>.

**Author contribution.** Conceptualization: B.N., V.M., J.B.; Data curation: B.N., I.G., J.S.; Formal analysis: B.N., I.G., J.S., J.B.; Funding acquisition: V.M., J.B.; Investigation: B.N., I.G., J.S., J.B.; Methodology: B.N., V.M., J.B.; Project administration: B.N., J.B.; Resources: I.G., J.S., V.M., J.B.; Software: B.N., I.G., J.S.; Supervision: V.M., J.B.; Validation: B.N., I.G., J.S., J.B.; Visualization: B.N., I.G., J.S.; Writing - original draft: B.N., J.B.; Writing - review & editing: B.N., I.G., J.S., V.M., J.B.

**Funding statement.** B.N. and J.B. were supported by the German Research Foundation (DFG), project 512483310. B.N. was moreover supported by the Helmholtz Association under the joint research school HIDSS4Health – Helmholtz Information and Data Science School for Health. I.H.G. was supported by a Stanford Center for Computational, Evolutionary and Human Genomics Fellowship. V.M.M. acknowledges support from NIH grant R01AI170204.

**Reproducibility.** The repository [https://github.com/barbora-sobolova/epies\\_tim\\_overdispersion/](https://github.com/barbora-sobolova/epies_tim_overdispersion/) contains code and data to reproduce our analyses.

## References

- [1] Gostic KM, et al. (2020) Practical considerations for measuring the effective reproductive number,  $R_t$ . *PLoS Computational Biology* **16**(12), 1–21.
- [2] Cori, et al. (2013) A new framework and software to estimate time-varying reproduction numbers during epidemics. *American Journal of Epidemiology* **178**(9), 1505–1512.
- [3] Nash RK, Nouvellet P and Cori A (2022) Real-time estimation of the epidemic reproduction number: Scoping review of the applications and challenges. *PLOS Digital Health* **1**(6), 1–17.

- [4] **Brizzi A, O'Driscoll M and Dorigatti I** (2022) Refining reproduction number estimates to account for unobserved generations of infection in emerging epidemics. *Clinical Infectious Diseases* **75**(1), e114–e121.
- [5] **Green WD, Ferguson NM and Cori A** (2022) Inferring the reproduction number using the renewal equation in heterogeneous epidemics. *Journal of the Royal Society Interface* **19**(188), 20210429.
- [6] **Nash RK**, et al. (2023) Estimating the epidemic reproduction number from temporally aggregated incidence data: A statistical modelling approach and software tool. *PLoS Computational Biology* **19**(8), 1–14.
- [7] **Alvarez L**, et al. (2021) Computing the daily reproduction number of COVID-19 by inverting the renewal equation using a variational technique. *Proceedings of the National Academy of Sciences* **118**(50), e2105112118.
- [8] **Champredon D, Papst I and Yusuf W** (2024) ern: An R package to estimate the effective reproduction number using clinical and wastewater surveillance data. *PLoS One* **19**(6), 1–22.
- [9] **Scire J**, et al. (2023) estimateR: an R package to estimate and monitor the effective reproductive number. *BMC Bioinformatics* **24**(1), 310.
- [10] **Nouvellet P** (2025) Rtglm: Unifying estimation of the time-varying reproduction number,  $R_t$ , under the generalised linear and additive models. *Epidemics* **53**, 100857.
- [11] **Huisman JS**, et al. (2022) Estimation and worldwide monitoring of the effective reproductive number of SARS-CoV-2. *eLife* **11**, e71345.
- [12] **Brockhaus EK**, et al. (2023) Why are different estimates of the effective reproductive number so different? A case study on COVID-19 in Germany. *PLoS Computational Biology* **19**(11), 1–27.
- [13] **Goldstein IH, Wakefield J, Minin VM**. (2023) Incorporating testing volume into estimation of effective reproduction number dynamics. *Journal of the Royal Statistical Society Series A: Statistics in Society* **187**(2), 436–453.
- [14] **Schneckenreither G**, et al. (2023) Assessing the heterogeneity in the transmission of infectious diseases from time series of epidemiological data. *PLoS One* **18**(5), 1–26.
- [15] **Johnson KD** et al. (2021) Disease momentum: Estimating the reproduction number in the presence of superspreading. *Infectious Disease Modelling* **6**, 706–728.
- [16] **Ho F**, et al. (2023) Accounting for the potential of overdispersion in estimation of the time-varying reproduction number. *Epidemiology* **34**(2), 201–205.
- [17] **Meagher J and Friel N** (2022) Assessing epidemic curves for evidence of superspreading. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* **185**(4), 2179–2202.
- [18] **Zhang Y, Britton T and Zhou X** (2022) Monitoring real-time transmission heterogeneity from incidence data. *PLoS Computational Biology* **18**(12), 1–24.
- [19] **Pakkanen MS**, et al. (2023) Unifying incidence and prevalence under a time-varying general branching process. *Journal of Mathematical Biology* **87**(2), 35.
- [20] **Lloyd-Smith JO**, et al. (2005) Superspreading and the effect of individual variation on disease emergence. *Nature* **438**, 355–359.
- [21] **Zelner J**, et al. (2021) Accounting for uncertainty during a pandemic. *Patterns* **2**(8), 100310.
- [22] **Nouvellet P**, et al. (2018) A simple approach to measure transmissibility and forecast incidence. *Epidemics* **22**, 29–35.
- [23] **Li Y**, et al. (2021) 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. *The Lancet Infectious Diseases* **21**(2), 193–202.
- [24] **Lison A**, et al. (2023) Effectiveness assessment of non-pharmaceutical interventions: lessons learned from the COVID-19 pandemic. *The Lancet Public Health* **8**(4), e311–e317.
- [25] **Dumont M**, et al. (2005) Correcting standard errors in two-stage estimation procedures with generated regressands. *Oxford Bulletin of Economics and Statistics* **67**(3), 421–433.
- [26] **White LF**, et al. (2020) Statistical estimation of the reproductive number from case notification data. *American Journal of Epidemiology* **190**(4), 611–620.
- [27] **Fraser C** (2007) Estimating individual and household reproduction numbers in an emerging epidemic. *PLoS One* **2**(8), 1–12. <https://doi.org/10.1371/journal.pone.0000758>.
- [28] **Britton T and Scalia Tomba G** (2019) Estimation in emerging epidemics: biases and remedies. *Journal of the Royal Society Interface* **16**(150), 20180670. <https://doi.org/10.1098/rsif.2018.0670>.
- [29] **Fahrmeir L, Kneib T and Lang S** (2009) *Regression*. Heidelberg: Springer.
- [30] **Bracher J and Held L** (2020) A marginal moment matching approach for fitting endemic-epidemic models to underreported disease surveillance counts. *Biometrics* **77**(4), 1202–1214.
- [31] **Gressani O**, et al. (2022) EpiLPS: A fast and flexible Bayesian tool for estimation of the time-varying reproduction number. *PLoS Computational Biology* **18**(10), 1–27.
- [32] **Cameron AC and Trivedi PK** (1998) *Regression Analysis of Count Data*. Cambridge: Cambridge University Press.
- [33] **Hilbe JM** (2011) *Negative Binomial Regression*, 2nd Edn. Cambridge: Cambridge University Press.
- [34] **Held L, Höhle M and Hofmann M** (2005) A statistical framework for the analysis of multivariate infectious disease surveillance counts. *Statistical Modelling* **5**(3), 187–199.
- [35] **Ver Hoef JM and Boveng PL** (2007) Quasi-Poisson vs negative binomial regression: how should we model overdispersed count data? *Ecology* **88**(11), 2766–2772.
- [36] **Wooldridge JM** (2010) *Econometric Analysis of Cross Section and Panel Data*. Cambridge: The MIT Press.
- [37] **Molenberghs G and Verbeke G** (2007) Likelihood ratio, score, and Wald tests in a constrained parameter space. *The American Statistician* **61**(1), 22–27.
- [38] **Crowcroft NS**, et al. (2008) Respiratory syncytial virus infection in infants admitted to paediatric intensive care units in London, and in their families. *European Journal of Pediatrics* **167**(4), 395–399.
- [39] **Vink MA, Bootsma MCJ and Wallinga J** (2014) Serial intervals of respiratory infectious diseases: A systematic review and analysis. *American Journal of Epidemiology* **180**(9), 865–875.
- [40] **Morgan OW**, et al. (2010) Household transmission of pandemic (H1N1) 2009, San Antonio, Texas, USA, April–May 2009. *Emerging Infectious Diseases* **16**(4), 631–637.
- [41] **Fine PEM** (2003) The interval between successive cases of an infectious disease. *American Journal of Epidemiology* **158**(11), 1039–1047.
- [42] **R Core Team** (2013) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.
- [43] **Stasinopoulos DM and Rigby RA** (2007) Generalized additive models for location scale and shape (GAMLSS) in R. *Journal of Statistical Software* **23**(7), 1–46.
- [44] **Challen R and Danon L**. Instantaneous reproduction number estimation from modelled incidence. Preprint, medRxiv, 2025. <https://www.medrxiv.org/content/early/2025/12/15/2025.12.11.25342071> (accessed 26 March 2026).
- [45] **Riley P, Cost AA and Riley S** (2016) Intra-weekly variations of influenza-like illness in military populations. *Military Medicine* **181**(4), 364–368.
- [46] **Richter L** et al. (2022) Epidemiologische Parameter des COVID19 Ausbruchs - Update 01.04.2022, Österreich, 2020–2022. Technical report, Austrian Agency for Health and Food Safety (AGES), 2022. [https://www.ages.at/fileadmin/Corona/Epidemiologische-Parameter/updates/Update\\_Epidemiologische\\_Parameter\\_des\\_COVID19\\_Ausbruchs\\_2022-04-01.pdf](https://www.ages.at/fileadmin/Corona/Epidemiologische-Parameter/updates/Update_Epidemiologische_Parameter_des_COVID19_Ausbruchs_2022-04-01.pdf) (accessed 28 November 2025).
- [47] **Abbott S**, et al. (2020) Estimating the time-varying reproduction number of SARS-CoV-2 using national and subnational case counts. *Wellcome Open Research* **5**(112). <https://doi.org/10.12688/wellcomeopenres.16006.2>.
- [48] **Sebastian JN, Minin VM**. (2025) Gaussian process priors with Markov properties for effective reproduction number inference. Preprint, arXiv, <https://arxiv.org/abs/2511.18797> (accessed 26 March 2026).
- [49] **Kenne Pagui EC, Salvan A and Sartori N** (2022) Improved estimation in negative binomial regression. *Statistics in Medicine* **41**(13), 2403–2416.