ORIGINAL ARTICLE

Journal Section

On the prior modelling for the basic reproductive ratio

Luiz Max Carvalho[1,3*] | **Marcio Marciel Bastos**[2*] | **Daniel A. Villela**[3] | **Leonardo S. Bastos**[3] | **Flávio Codeço Coelho**[2]

[1]School of Biological Sciences, University of Edinburgh, Edinburgh, UK.

[2]School of Applied Mathematics, Getùlio Vargas Foundation, Rio de Janeiro, Brazil.

[3]Scientific Computing Programme (PROCC), Oswaldo Cruz Foundation, Rio de Janeiro, Brazil.

Correspondence
Luiz Max Carvalho, Scientific Computing Programme (PROCC), Oswaldo Cruz Foundation, Rio de Janeiro, Brazil.
Email: lmax.procc@gmail.com

Funding information
Funder One, Funder One Department, Grant/Award Number: 123456, 123457 and 123458; Funder Two, Funder Two Department, Grant/Award Number: 123459

The basic reproductive number, $\mathcal{R}_0$, is a central quantity in theoretical epidemiology, defining the threshold between disease-free equilibria and epidemics. $\mathcal{R}_0$ can usually be written as a ratio between the rate of creation of infected individuals and the rate of removal of infected individuals. Bayesian inference for epidemic models usually proceeds by assigning Gamma or log-normal priors to these rates, obtaining posterior distributions and often computing their ratio to recover a posterior distribution for $\mathcal{R}_0$. In this paper we show that these modelling choices lead to induced distributions on the quantity of interest ($\mathcal{R}_0$) that have poor statistical properties. We propose new classes of priors and also extend the usual approach by correcting the distribution of $\mathcal{R}_0$ to explicitly consider the population size N. Our findings show that care is needed when constructing priors for inference of epidemic models to ensure valid conclusions about the quantities of interest.

KEYWORDS
Epidemic models, basic reproductive number, prior modelling, Bayesian inference.

Abbreviations: ABC, a black cat; DEF, doesn't ever fret; GHI, goes home immediately.

*Equally contributing authors.

- Collect the models that give the gamma ratio for the $\mathcal{R}_0$;
- select studies and data sets to evaluate;
- study moment-matching priors with better tail properties
- Question: does restricting the support of $f_{\mathcal{R}_0}(r)$ to $[0, N]$ lead to improved inference?
- Derive $E[\mathcal{R}_0]$, Var($\mathcal{R}_0$), etc, with the support correction.

| EPIDEMIC MODELS

What are epidemic models?

Why they are useful?

$\mathcal{R}_0$ is represents the threshold between epidemic and disease-free equilibria. Useful summary; often reported

SIR model

$$\begin{aligned} \frac{dS}{dt} &= -\beta SI \\ \frac{dI}{dt} &= \beta SI - \gamma I \\ \frac{dR}{dt} &= \gamma I \end{aligned}$$

where $S(t) + I(t) + R(t) = N \; \forall t$, β is the transmission (infection) rate and γ is the recovery rate. The basic reproductive number is

$$\mathcal{R}_0 = \frac{\beta N}{\gamma}. \tag{1}$$

| BAYESIAN ESTIMATION OF TRANSMISSION AND RECOVERY RATES

(Coelho et al., 2011)

LITERATURE REVIEW

$$p(\boldsymbol{\theta}|\boldsymbol{Y}) \propto \mathcal{L}(\boldsymbol{Y}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}) \tag{2}$$

| The induced prior on $\mathcal{R}_0$

When constructing prior distributions it is important to consider the effects of these prior distributions on quantities of interest (q.o.i.) in the model, particularly when these are non-linear transforms of the parameters (Seaman III et al., 2012). Even if a prior distribution seems uninformative on the scale of the parameter it was designed to model, the **induced** distribution on the q.o.i. might be informative.

In this section we expand the discussion in Section 2.4 of Clancy et al. (2008) about the induced distribution on $\mathcal{R}_0$ when using Gamma priors on (γ, β) and also consider the case where log-normal priors are used. For simplicity we will

assume that $\pi(\boldsymbol{\theta}) = f_\gamma(\gamma) f_\beta(\beta)$.

Gamma priors

Suppose the *a priori* uncertainty about parameters can be represented by Gamma distributions:

$$\begin{aligned} f_\beta(b \mid k_\beta, \theta_\beta) &= \frac{1}{\Gamma(k_\beta)\theta_\beta^{k_\beta}} b^{k_\beta - 1} \exp\left(-\frac{b}{\theta_\beta}\right), \\ f_\gamma(g \mid k_\gamma, \theta_\gamma) &= \frac{1}{\Gamma(k_\gamma)\theta_\gamma^{k_\gamma}} g^{k_\gamma - 1} \exp\left(-\frac{g}{\theta_\gamma}\right), \end{aligned}$$

then the probability distribution function (pdf) of $\mathcal{R}_0$ is given by

$$f_{\mathcal{R}_0}(r \mid k_\beta, \theta_\beta, k_\gamma, \theta_\gamma, N) = \frac{(N\theta_\beta\theta_\gamma)^{k1+k2}}{\mathcal{B}(k_\beta, k_\gamma)(N\theta_\beta)^{k_\beta}\theta_\gamma^{k_\gamma}} r^{k_\beta - 1}(\theta_\gamma r + N\theta_\beta)^{-(k_\beta + k_\gamma)}, \tag{3}$$

where $\mathcal{B}(a, b) = \Gamma(a + b)/\Gamma(a)\Gamma(b)$ is the Beta function. We give a derivation of this result in the Appendix. Clancy et al. (2008) call this distribution a scaled F distribution. We will instead henceforth refer to it as the Gamma ratio distribution. The (cumulative) distribution function of the Gamma ratio is

$$F_{\mathcal{R}_0}(x) = \frac{(N\theta_\beta\theta_\gamma)^{k1+k2}}{\mathcal{B}(k_\beta, k_\gamma)(N\theta_\beta)^{k_\beta}\theta_\gamma^{k_\gamma}} \frac{x^{k_\beta}\left(\frac{\theta_\gamma x}{\theta_\beta N} + 1\right)^{(k_\beta + k_\gamma)} {}_2F_1\left(k_\beta, k_\beta + k_\gamma, k_\beta + 1, -\frac{\theta_\gamma x}{\theta_\beta N}\right)}{k_\beta\left(\theta_\gamma x + \theta_\beta N\right)^{(k_\beta + k_\gamma)}}, \tag{4}$$

where ${}_2F_1(a, b, c, z)$ is the Gaussian hypergeometric function. The expectation of the Gamma ratio distribution is then

$$\begin{aligned} E[\mathcal{R}_0] &= \int_0^\infty r f_{R0}(r) dr, \\ &= \frac{N\theta_\beta}{\theta_\gamma}\frac{k_\beta}{(k_\gamma - 1)}, \end{aligned} \tag{5}$$

which is defined only for $k_\gamma > 1$. The variance can be computed as

$$\begin{aligned} \mathrm{Var}(\mathcal{R}_0) &= E[\mathcal{R}_0^2] - E[\mathcal{R}_0]^2, \\ &= \left(\frac{N\theta_\beta}{\theta_\gamma}\right)^2 \frac{(k_\beta + k_\gamma - 1)k_\beta}{(k_\gamma - 2)(k_\gamma - 1)^2}, \end{aligned} \tag{6}$$

and only exists for $k_\gamma > 2$. The mode is

$$\frac{N\theta_\beta}{\theta_\gamma}\frac{k_\beta - 1}{(k_\gamma + 1)}. \tag{7}$$

Please notice that the formulae given here might differ from those given in Clancy et al. (2008) because we use a shape/scale parametrisation whilst those authors assume the prior distributions on β and γ are parametrised in terms of shape and rate. As discussed by Clancy et al. (2008) (section 3.2), the induced prior on $\mathcal{R}_0$ displays some undesirable statistical properties. For instance, for fairly standard choices of the hyperparameters, namely $k_\beta = k_\gamma = 1$ and $\theta_\beta = \theta_\gamma = \epsilon$ where ϵ is a large positive number, one ends up with a distribution on $\mathcal{R}_0$ for which $E[\mathcal{R}_0] > 1$, the variance

is undefined and $\Pr(\mathcal{R}_0 > 1) = 1/2$, which assigns equal probability to epidemic and disease-free equilibria. We also note that $f_{\mathcal{R}_0}(r)$ is concave only for

$$r' < \frac{N\theta_\beta\sqrt{\left(k_\beta - 1\right)k_\gamma^2 + \left(k_\beta^2 + k_\beta - 2\right)k_\gamma + 2k_\beta^2 - 2k_\beta} + \left(k_\beta - 1\right)N\theta_\beta k_\gamma + \left(2k_\beta - 2\right)N\theta_\beta}{\theta_\gamma k_\gamma^2 + 3\theta_\gamma k_\gamma + 2\theta_\gamma}. \tag{8}$$

The (log) concavity property is important for instance when computing *maximum a posteriori* (MAP) estimates or employing most approximate inference methods that require unimodality. Of course, one might argue that if the model is parametrised in terms of β and γ, so long as the individual priors for these parameters are (log) concave, there would be no problems. We however believe that in a setting where the induced distribution on a transformation of the parameters is also of interest, it is reasonable to expect such properties from the induced distribution. In other words, it would be desirable to construct the priors on the natural parameters of the model so as to induce a well-behaved probability distribution on the q.o.i., $\mathcal{R}_0$.

Log-normal priors

Another popular class of probability distributions to model strictly positive quantities is the log-normal. In this case, suppose the uncertainty about β and γ can be represented by

$$\begin{aligned} h_\beta(b \mid \mu_\beta, \sigma_\beta) &= \frac{1}{b\sigma_\beta\sqrt{2\pi}}\exp\left(-\frac{\left(\ln b - \mu_\beta\right)^2}{2\sigma_\beta^2}\right), \\ h_\gamma(g \mid \mu_\gamma, \sigma_\gamma) &= \frac{1}{g\sigma_\gamma\sqrt{2\pi}}\exp\left(-\frac{\left(\ln g - \mu_\gamma\right)^2}{2\sigma_\gamma^2}\right). \end{aligned}$$

It is straightforward to show that the induced distribution on $\mathcal{R}_0$ is

$$h_{\mathcal{R}_0}(r \mid \mu_\beta, \sigma_\beta, \mu_\gamma, \sigma_\gamma) = \frac{1}{r\sqrt{2\pi\left(\sigma_\beta^2 + \sigma_\gamma^2\right)}}\exp\left(-\frac{\left(\ln r - \ln N - \mu_\beta + \mu_\gamma\right)^2}{2\left(\sigma_\beta^2 + \sigma_\gamma^2\right)}\right), \tag{9}$$

i.e. a log-normal distribution with parameters $\mu_{\mathcal{R}_0} = \ln N + \mu_\beta - \mu_\gamma$ and $\sigma_{\mathcal{R}_0} = \sigma_\beta^2 + \sigma_\gamma^2$. Under the justification of employing a non-informative prior, researchers might be tempted to choose $\mu_\beta = \mu_\gamma = 0$ and $\sigma_\beta = \sigma_\gamma = 100$, say – see e.g. Ho et al. (2018), section 5.1. This apparently non-informative choice of hyperparameters leads to a prior on $\mathcal{R}_0$ for which $E[\mathcal{R}_0] = \exp(10^4)$ and $\Pr(\mathcal{R}_0 > 100) = 0.49$, which are not reasonable.

Informative priors on β and γ

So far we have discussed the setting where so-called non-informative priors are employed. We now study the induced distribution of $\mathcal{R}_0$ in the case where informative priors are used. Consider modelling [DISEASE] , where the transmission rate has a mean of [XX] and the average recovery rate is [YY]. To construct the Gamma priors we picked $k_\beta = 2$, $\theta_\beta = 1/4000$, $k_\gamma = 40$ and $\theta_\gamma = 1/200$. For the log-normal priors we matched the first two moments of the Gamma priors,

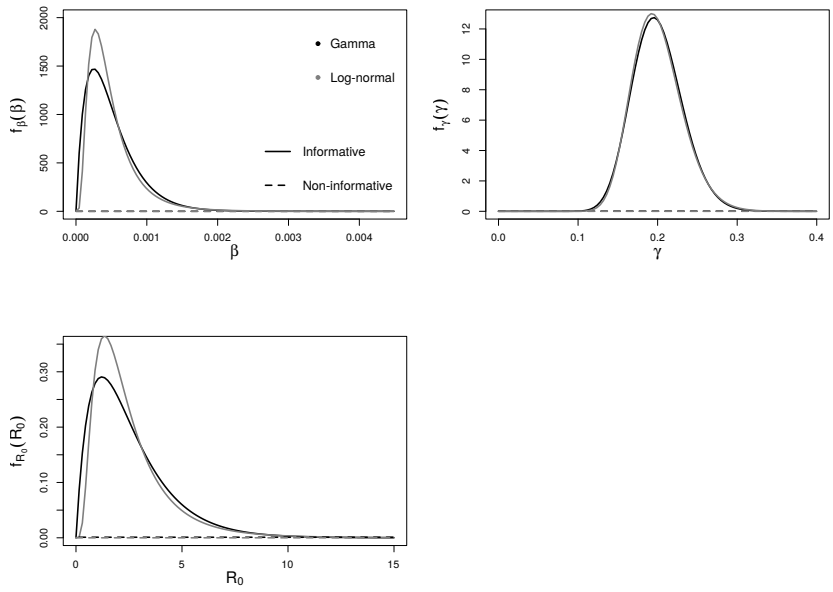

FIGURE 1 Informative and uninformative priors for the transmission (β) and recovery (γ) rates and the induced density on the basic reproductive number, $\mathcal{R}_0$. Solid lines show the informative prior, whilst dashed lines show the uninformative formulation (see text). Black lines represent the Gamma priors (and Gamma ratio for $\mathcal{R}_0$) and grey lines pertain to the moment-matching log-normal densities.

i.e., we set

$$\mu_\beta = \ln\left(\frac{k_\beta\theta_\beta}{\sqrt{1+\frac{1}{k_\beta}}}\right) \text{ and } \sigma_\beta = \ln\left(1+\frac{1}{k_\beta}\right),$$
$$\mu_\gamma = \ln\left(\frac{k_\gamma\theta_\gamma}{\sqrt{1+\frac{1}{k_\gamma}}}\right) \text{ and } \sigma_\gamma = \ln\left(1+\frac{1}{k_\gamma}\right).$$

Figure 1 shows the resulting densities, along with uninformative priors obtained by setting $k_\beta = k_\gamma = 1$ and $\theta_\gamma = \theta_\beta = 100$ – log-normal priors again constructed via moment-matching. For these computations we used $N = 1000$.

Accounting for Biology

The physical interpretation of $\mathcal{R}_0$ is the number of secondary infections caused by a single infectious individual in a susceptible population. This means that $\mathcal{R}_0$ cannot exceed the population size N, which for simplicity we will assume is fixed through time. One can obtain a distribution on $\mathcal{R}_0$ with the correct bounds by correcting the distribution to have

all of its probability density in the interval $[0, N]$:

$$f'_{\mathcal{R}_0}(r \mid k_\beta, \theta_\beta, k_\gamma, \theta_\gamma, N) = \begin{cases} \frac{1}{F_{\mathcal{R}_0}(N)} f_{\mathcal{R}_0}(r \mid k_\beta, \theta_\beta, k_\gamma, \theta_\gamma, N), \ 0 < r < N, \\ 0, \ \text{otherwise} \end{cases} \tag{10}$$

DISCUSSION

(Weidemann et al., 2014) present skew normal priors as a more flexible approach to Gamma and Beta priors, since the SN has three parameters. This makes it easier to construct priors from a point estimate of central tendency and confidence/credibility intervals.

ACKNOWLEDGEMENTS

Acknowledgements should include contributions from anyone who does not meet the criteria for authorship (for example, to recognize contributions from people who provided technical help, collation of data, writing assistance, acquisition of funding, or a department chairperson who provided general support), as well as any funding or other support information.

CONFLICT OF INTEREST

You may be asked to provide a conflict of interest statement during the submission process. Please check the journal's author guidelines for details on what to include in this section. Please ensure you liaise with all co-authors to confirm agreement with the final statement.

| APPENDIX

To derive the induced distribution, we begin by noting that for $N > 1$, the distribution of $\beta^* = \beta N$ is a Gamma distribution with parameters k_β and $N\theta_\beta$. Under the assumption of independence $\pi(\beta^*, \gamma) = f_{\beta^*}(\beta^*) f_\gamma(\gamma)$, we can write

$$\mathcal{R}_0 = \beta^*/\gamma, \tag{11}$$

$$f_{\mathcal{R}_0}(r) = A \int_0^\infty \gamma(\gamma r)^{k_\beta - 1} e^{-\frac{\gamma r}{N\theta_\beta}} \gamma^{k_\gamma - 1} e^{-\frac{\gamma}{\theta_\gamma}} d\gamma, \tag{12}$$

$$A := \frac{1}{\Gamma(k_\beta)(N\theta_\beta)^{k_\beta} \Gamma(k_\gamma) \theta_\gamma^{k_\gamma}}.$$

Rearranging yields

$$f_{\mathcal{R}_0}(r) = A r^{k_\beta - 1} \int_0^\infty \gamma^{k_\beta + k_\gamma - 1} e^{-B\gamma} d\gamma, \tag{13}$$

$$B := \frac{\theta_\gamma r + N\theta_\beta}{N\theta_\beta \theta_\gamma}.$$

Noticing the integral in (13) is the kernel of a Gamma pdf gives the result in (3). For a slightly different derivation, based on generalised Gamma distributions, see Coelho and Mexia (2007).

REFERENCES

Clancy, D., O'Neill, P. D. et al. (2008) Bayesian estimation of the basic reproduction number in stochastic epidemic models. *Bayesian Analysis*, **3**, 737–757.

Coelho, C. A. and Mexia, J. T. (2007) On the distribution of the product and ratio of independent generalized gamma-ratio random variables. *Sankhyā: The Indian Journal of Statistics*, 221–255.

Coelho, F. C., Codeço, C. T. and Gomes, M. G. M. (2011) A bayesian framework for parameter estimation in dynamical models. *PloS one*, **6**, e19616.

Ho, L. S. T., Crawford, F. W., Suchard, M. A. et al. (2018) Direct likelihood-based inference for discretely observed stochastic compartmental models of infectious disease. *The Annals of Applied Statistics*, **12**, 1993–2021.

Seaman III, J. W., Seaman Jr, J. W. and Stamey, J. D. (2012) Hidden dangers of specifying noninformative priors. *The American Statistician*, **66**, 77–84.

Weidemann, F., Dehnert, M., Koch, J., Wichmann, O. and Höhle, M. (2014) Bayesian parameter inference for dynamic infectious disease modelling: rotavirus in germany. *Statistics in medicine*, **33**, 1580–1599.