

Dear Editor:

Thank you for the chance to revise and resubmit our manuscript. Below please find our responses to reviewers.

## Reviewer #1

The manuscript presents a statistical framework for comparing different estimates of  $R_0$  decomposed in their building blocks (exponential growth, average generation interval, and dispersion in the generation interval). The analysis allows for pulled estimates of the different quantities, combining information from the different studies. In addition, authors provide a reflection on how the uncertainty in the different components is propagated in the  $R_0$  estimate and its confidence interval.

It seems to me that the paper is well written.

Thank you!

As a disclaimer, I must say that I am not sufficiently expert of the methodology used here to be fully confident in my judgment of the soundness and the completeness of the analysis. Maybe it is related to my limitation, but I find the main focus of the paper a little too technical for the Journal (e.g. see the last point below). Thus, I believe that authors could make an extra effort to convince a more general and interdisciplinary public on the importance of the point they are raising.

We have made significant revisions to the manuscript to address this point.

I am not sure whether authors incorporate well all assumptions of the studies they consider.

During the exponential growth period, the basic reproduction number entirely depends on the exponential growth rate and the generation-interval distributions. Therefore, all model assumptions can be boiled down to assumptions about the exponential growth rate and the shape of the generation-

interval distribution. So we are effectively considering all model assumptions. We have tried to make this point clearer:

“Despite a wide range of models considered across Study 1–7, all of them assume that the epidemic grows exponentially in the beginning. While the IDEA model (used in Study 7) includes a discount parameter  $d$  that allows the model to deviate from the exponential growth, Study 7 estimates  $d = 0$  across all parameters they consider, in which case the IDEA model becomes equivalent to the exponential growth model. When the epidemic is growing exponentially, the estimated basic reproduction number entirely depends on the exponential growth rate  $r$  and the *intrinsic* generation-interval distribution  $g(\tau)$ :

$$1/\mathcal{R}_0 = \int \exp(-r\tau)g(\tau) d\tau. \quad (1)$$

In other words, all model assumptions come down to assumptions about the exponential growth rate  $r$  and the shape of the generation-interval distribution  $g(\tau)$ . For example, if a model relies on strong assumptions about the underlying observation or process model, the estimated confidence intervals of the model parameters will be necessarily narrow. Therefore, it is sufficient to consider the estimates and assumptions about the exponential growth rates and shape of the generation-interval distribution to understand the estimates of the basic reproduction number.”

In addition, it is not clear to me on what extent the assumptions made by authors themselves affect the results - e.g. the assumption of a gamma and uniform distribution for the estimates adopted to compute the growth rate.

???

Other hypotheses underlie these works other than average generation interval, its coefficient of variation and the exponential growth. For instance, what about the offspring distribution (Poisson vs. negative binomial), or the observation model? Authors should provide a synthesis of all assumptions of these studies

See above.

and discuss how well the gamma approximation framework embrace all cases.

The gamma approximation framework assumes that the generation-interval distribution is gamma distributed. Other assumptions mentioned above (e.g., offspring distributions or the observation model) do not affect the generation-interval distributions. Moreover, this approximation has been widely used previously and has been tested across multiple diseases. We have clarified the text and included appropriate citations:

“Here, we use the gamma approximation framework to the generation-interval distribution to (i) characterize the amount of uncertainty present in the exponential growth rates and the shape of the generation-interval distribution and (ii) assess the degree to which these uncertainties affect the estimate of  $\mathcal{R}_0$ . The gamma distribution provides a reasonable approximation for generation-interval distributions of many diseases, including Ebola, influenza, measles, and rabies.”

Authors state that “We note that this approach does not account for non-independence between the parameter estimates made by different modelers.” Authors should discuss this point more in depth. This seems to me an important limitation. With this respect I wonder what data the studies focus on. Are all analyzing the same epidemic (in China) considering same or similar data? Again, more details on these studies (in this case the location, period of the epidemic, and data source) should be provided.

We now provide a table that summarizes this information.

At the beginning of the Discussion section authors list a series of important issues and possible sources of errors that often limit the utility of early outbreak analyses (e.g. fitting the cumulative curve of cases, assuming a fixed generation interval when in fact it varies during the outbreak, changing reporting rate, etc.). Compared with these issues the main focus of the paper (not properly accounting for the uncertainty in the generation time and growth rate parameters) appears a small caveat. Maybe author could give a sense of the importance of the element they are focusing on with the respect to the other issues?

## Reviewer #2

We might have inadvertently confused reviewer 2. Our method does not rely on any data or dynamical models but the reviewer asks us to compare our results with the reported number of cases (this is not possible as we do not make any predictions about the epidemic trajectory). We use a multi-level model to calculate pooled estimates of parameters that have been estimated or assumed from other studies. Then, we use the pooled estimates to calculate  $\mathcal{R}_0$ . We have made significant improvement to our manuscript to make our methods clearer.

The subject is interesting for the ongoing pandemic. The authors propose an estimation method for the basic reproduction number and describe its uncertainty from the estimates made by several early studies. The consensus epidemiological estimate could be beneficial for making control strategies.

Thank you!

The authors propose a statistical method for an epidemiological situation. They combine basic reproduction rate estimates and other estimates in the early stage of the outbreak and their uncertainty to obtain a global estimate with its uncertainty. So, this subject is at the interface of biomathematics and epidemiology of the infectious disease, which is the scope of the journal.

The interesting and primordial subject of the estimation of  $R_0$  and its uncertainty is studied using the Bayesian inference tools. The authors do not develop enough the underlying hypotheses and the method used. Many elements are not described, even in the appendix. This should be described to estimate the validity of these underlying assumptions.

More specific points:

The authors must use the same variable names and their units throughout the documents: e.g., for kappa, which is a key parameter in this study. In the document, it appears in following various terms that make the reader confused: GI variation, squared coefficient of variation, generation-interval dispersion. Please always keep the same terms for the same variables. The same remark is also

applied to the mean generation interval.

Done.

The unit of time changes throughout the document. Choose the same time unit (e.g., days, which is the most used and change the Equation 5, Page 6, which is in weeks).

Done.

As in the Bayesian inference, there are no Confidence Intervals like in frequentist inference, but Credible Interval. The authors should replace Confidence Intervals with Credible Intervals.

Done.

## Methods

The authors should clearly explain the hypotheses they have done in the Methods section (e.g., the source of the uncertainty).

The authors do not describe all the methods used in this article within the Methods section. Some of them are in the figure captions, e.g., Figures 2 and 3. Also, the authors should describe all the methods in this section. Using subheadings could help the reader to better understand the methods, e.g., likelihood, prior distributions, sensitivity analysis, etc.

Done.

Please clearly explain the formula in the caption of Figure 2 (Page 8, line 33) in the Methods section.

Done.

The Bayesian estimation framework is not clearly explained. Readers could easily get confused about which variables serve as observed data, parameters, or hyperparameters. It is also not very clear about the number of observed data, e.g.,  $i = 7$  means that there are only 7 data (1 per study) or there are 700,000

simulated data (100,000 per study) in the model. We strongly suggest that the authors should provide a directed acyclic graph (DAG) representing the entire model, which is a common practice for the Bayesian estimation.

Done.

The authors consider that all the 7 studies have the same weight in the estimation of basic reproduction weight even though their estimations and precision vary a lot. The authors might explain why they make this hypothesis and why they did not explore a weighted estimation based on the confidence of the estimation of  $R_0$ .

Done.

In each chain, the authors use 500,000 sampling steps and remove every 1000th step. Therefore, the posterior distributions of each parameter consist of 2,000 samples (500 samples per chain with 4 independent chains). Is it sufficient for the effective sample size? Is it too much for the burn-in steps of 500,000 and the thinning steps of 1,000?

This may be an overkill but we wanted to make sure that the posterior samples are not autocorrelated. This is definitely sufficient for the effective sample size to reliably calculate the 95% confidence intervals, especially given that we're only averaging across previously estimated parameter values, rather than explicitly fitting a model to data.

## Results

In Figures 2 and 3, the authors replace the reported parameters ( $r$ ,  $\bar{G}$ , and  $\kappa$ ) with the pooled estimates  $\mu$  ( $\mu_r$ ,  $\mu_G$ , and  $\mu_\kappa$ ). The authors do not give information about the other sources of uncertainty linked with the variation of parameter sigma ( $\sigma_r$ ,  $\sigma_G$ ,  $\sigma_\kappa$ ).

Done.

The authors do not provide the validation on their estimates, e.g., a posterior predictive check and comparisons with the reported cases during the study period.

No...

In figure 1, the posterior density distribution for the 3 parameters [exponential growth rate ( $r$ ), Mean generation interval ( $\bar{G}$ ), Generation interval dispersion ( $\kappa$ )] must replace the segment representing the median and the CI. In figure 1, the legend must be self-sufficient. 'See text' it is not possible. A short explanation is necessary.

Done.

In the first paragraph of the Results page 7 the value of the median and the CI of the 3 parameters must be presented.

Done.

The estimation of  $R_0$  median and its credible interval "2.9 (95% CI: 2.1-4.5)" must be presented in page 7 with the figure 2 in the second paragraph of Results and not at the end of the Results part (page 9 line 26).

Done.

The uncertainty related to  $\kappa$  is not described in Figure 2. Even the authors mentioned in Page 7, Lines 44-50, the results should also be visualized in Figure 2 along with other parameters.

Done.

Discussions

In the Results section, the authors occasionally discuss the results, i.e., Page 8, line 50 to Page 9, line 23. The opinions should be in the Discussion section.

Done.

In the Discussion section, the authors should discuss the quality of the reported estimates from 7 studies used in this subject. Some of them do not provide detailed methods and assumptions. Also, most of them are preprint articles. It could be another important source of uncertainty on  $R_0$ .

Although the authors are in the case of the COVID-19 epidemic in Hubei before any confinement, the authors do not mention in the introduction or in the discussion the other sources of uncertainty linked to the data (such as the variations in diagnostic results related to the day of the week, a saturation of diagnostic test capacity, transparency of data, representativeness of samples and improvement of detection capacity as time goes by).

Page 4, Table 1. Please check Study 4, the article has been revised since 27 January 2020, and their estimates also changed.

Page 6, Table 2. Please recheck the parameters of Study 5, which is inconsistent with the 90% confidence interval in Table 1. The 90% confidence interval for a Gamma distribution with mean 2.2 and shape 12 is 1.27 to 3.34, while the interval in Table 1 is 1.4 to 3.8.

Done.

Page 6, Line 29. Are the priors for  $\mu_r$ ,  $\mu_G$ , and  $\mu_K$  weakly informative? Did the authors check the sensibility with new priors?

Done.

Page 6 Lines 44-47. Please clarify the sentence: "Alternative choices of prior ... can lead to poor mixing".

Done.

Page 7, Line 20. Figure 1 - Caption: Did the authors test the sensitivity of the value of kappa of Study 2 (e.g., 0 or 1) on the pooled estimate?

Page 7, Line 38. The authors might give some information about the incubation period of COVID-19. It could explain why using the data of 26 January is still acceptable even if it was the 3rd day of the city lockdown.

Done.

Page 8, Line 50. Please clarify the sentence: "We find that incorporating uncertainties one at a time increases the width of the confidence intervals in all but 7



cases". It is not clear particularly at "in all but 7 cases".

Done.

### **Reviewer #3**

Here, Park and colleagues present a straight-forward, but important analysis of 7 early estimates of  $R_0$  for the ongoing covid outbreak. They stress that multiple forms of uncertainty have to be taken into account when calculating  $R_0$  and its confidence intervals, lest the intervals be too narrow and/or the estimates over-confident. The methods are sound, and the results compelling. I really don't have any major comments, though I wonder if it would be better suited to another journal. I do apologize to the authors for the tardiness of my review.

Thank you for the review. We have made significant revisions to our manuscript to make it approachable for a wider audience.