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 R0 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 R0 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 our manuscript to make it approachable for a wider audience.

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 as-

sumptions 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 in the main text:

"Despite a wide range of models considered across Study 1–7, all of them assume that the epidemic grows exponentially in the beginning. The IDEA model (used in Study 7) includes a discount parameter d that allows the model to deviate from the exponential growth when $d \neq 0$ (Fisman et al., 2013), but Study 7 estimates $d = 0$ across all parameters they consider. Even though some studies consider reported cases up to Janaury 26, 2020 — three days after the travel restriction that took place on January 23, 2020 (Tian et al., 2020) — the exponential growth assumption can still describe the number of reported cases reasonably well; given the incubation period of around 5 days (Lauer et al., 2020) as well as reporting delay of around 5 days (Sun et al., 2020), the majority of reported cases during study periods are likely to have been infected prior to the travel ban.

When the epidemic is growing exponentially, the estimated basic reproductive number entirely depends on the exponential growth rate r and the *intrinsic* generation-interval distribution $g(\tau)$, which describes the infection time of secondary cases caused by a primary case in a fully susceptible population, via the Euler-Lotka equation (Wallinga and Lipsitch, 2007):

$$\frac{1}{\mathcal{R}_0} = \int \exp(-r\tau) g(\tau)\, d\tau. \tag{1}$$

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 or credible intervals associated with the exponential growth rates or parameters of the generation-interval distributions will be necessarily narrow. Therefore, it is sufficient to consider the estimates and assumptions about the exponential growth rates and the shapes of the generation-interval distributions to understand disparate estimates of the basic reproductive 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.

Since we are interested in the uncertainty, expressed by probability distributions, the overall probabilities are not affected by the assumptions of gamma or uniform distributions. We have added following sentences in the main text:

"This assumption does not affect our analysis because we are re-estimating the probability distribution of r_i that is consistent with the reported values of $(\mathcal{R}_0)_i$, $\bar{G}_i$, and κ_i; in other words, we still obtain the same estimates and the associated uncertainty of $(\mathcal{R}_0)_i$ if we calculate it from r_i, $\bar{G}_i$, and κ_i."

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. We have also added a table to provide a better overview of the models and data used in each study.

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 (Nishiura et al., 2009; McBryde et al., 2009; Roberts and Nishiura, 2011; Trichereau et al., 2012; Nishiura and Chowell, 2015; Park et al., 2019) 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. We have also added the following text:

"We note that this approach does not account for non-independence between the parameter estimates made by different modelers. In this case, most estimates primarily depend on reported cases from China, particularly from Hubei Province. Differences among estimates are likely to driven by differences in estimation methods and underlying assumptions, rather than epidemiological differences. The pooled estimates can become sharper (i.e., give narrower credible intervals) as we add more models even when the models or the data no longer add more information about the epidemic. Since SARS-CoV-2 had primarily spread in Hubei Province, China during this period, it is not possible to include independent sources of data from other countries. Thus, the pooled estimator should be interpreted with care."

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?

All of these issues raised here directly affect estimates of growth rate r, which in turn affect estimates of the basic reproduction number $\mathcal{R}_0$. Therefore, comparing the growth rate with generation-interval distribution parameters allow us to take these assumptions into account implicitly. We have made this clearer in the main text.

Reviewer #2

We might have inadvertently confused reviewer 2 to some extent. 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 already 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 R0 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).

We have made significant changes to our Methods section.

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.

We have removed the formula from the caption. However, we decided to include the formula in the Results section as we think it provides better flow:

"We further explore how the effects of uncertainties in generation-interval

distributions change when there is less uncertainty associated with the exponential growth rate. This hypothetical example reflects recent scenarios, as increased data availability will allow researchers to estimate r with more certainty. To simulate estimates of the exponential growth rate with stronger confidence, we use $\hat{\mu}_r = (\mu_r + 3 \times \text{median}(\mu_r))/4$ instead of μ_r; then $\hat{\mu}_r$ has the same median as μ_r but the associated 95% CI is 4 times narrower (0.16–0.19 days^{-1})."

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.

We have made significant revisions to the Methods section. In particular, we made it very clear that we are simply averaging across the reported estimates and are not relying on any data or dynamical model. Therefore, we do not think that a directed acyclic graph is necessary with the current text, particularly given that we are simply averaging across reported values.

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 R0.

Done.

"Here, we assume that all 7 studies are equally weighted. During the initial phase of an outbreak of a novel virus such as SARS-CoV-2, uncertainties in parameter estimates are primarily driven by the underlying assumptions; narrow confidence intervals from a study are likely to reflect their strong assumptions rather than its precision of inference. Moreover, we cannot assess which assumptions are more realistic during the initial phase of an outbreak a novel virus due to the scarcity of information. Therefore, it is difficult to come up with a sensible way to weigh different estimates, and we

do not explore weighted estimation."

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. On the other hand, 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 (μ_r, μ_G, and μ_κ). The authors do not give information about the other sources of uncertainty linked with the variation of parameter sigma (σ_r, σ_G, σ_κ).

Variation parameters do not affect the basic reproduction number.

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.

We do not rely on data or fit a model to data. Therefore, it is not possible to compare with the reported cases.

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.

We have changed the figure caption as follows:

"We inferred point estimates (black), uniform distributions (orange) or con-

fidence/credible intervals (purple) for each parameter from each study, and combined them into pooled estimates using a Bayesian multilevel model (red). Points represent medians calculated from the parameter set $(\bar{G}_i, \kappa_i, r_i)$ for each study i (orange and purple). Error bars represent 95% equi-tailed quantiles calculated from the parameter set $(\bar{G}_i, \kappa_i, r_i)$ for each study i. Red violin plots represent distributions of 2000 posterior samples. Open triangle: we assumed $\kappa = 0.5$ for Study 2 which does not report generation-interval dispersion."

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 R0 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.

We have removed most, if not all, interpretation of the results or opinions based on the results from the Results section.

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 R0.

This is a very good point. We have added the following paragraph to the main text:

"Here, we focused on the estimates of $\mathcal{R}_0$ that were published within a very short time frame (January 23–26, 2020). Since the estimates were published as pre-prints, rather than in peer-reviewed journals, the quality of the analyses as well as the resulting estimates have not been finalized. For example, Study 4 initially estimated $\mathcal{R}_0 = 3.8$ (95% CI: 3.6–4.0; Read et al. (2020a)) but revised their estimate on January 28, 2020: $\mathcal{R}_0 = 3.11$ (95% CI: 2.39–4.13; Read et al. (2020b)); we did not include their revised estimates in our analysis in order to focus on available information at the very beginning of the outbreak. Some studies also do not provide detailed description of their methods, data, and/or assumptions. The quality of these analyses adds further uncertainty to their results that are not captured by their uncertainty quantification (e.g., reported credible intervals) or by our analysis. While pre-prints can play an important role in disseminating information, they should be interpreted with care."

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.

Some of the issues raised here (e.g., day of the week effect and limited diagnostic resources) were already mentioned in the Discussion of the previous version of the manuscript. We now mention other points as well:

"There are other important phenomena not covered by our simple framework. Examples that seem relevant to this outbreak include: changing reporting rates, reporting delays (including the effects of weekends and holidays), and changing generation intervals. For emerging pathogens such as SARS-CoV-2, there may be an early period of time when the reporting rate is very low due to limited awareness or diagnostic resources; for example, Zhao et al. (2020) (Study 6) demonstrated that estimates of $\mathcal{R}_0$ can change from 5.47 (95% CI: 4.16–7.10) to 3.30 (95% CI: 2.73–3.96) when they assume 2-fold changes in the reporting rate between January 17, when the official diagnos-

tic guidelines were released (World Health Organization, 2020), and January 20. Delays between key epidemiological timings (e.g., infection, symptom onset, and detection) can also shift the shape of an observed epidemic curve and, therefore, affect parameter estimates as well as predictions of the course of an outbreak (Tariq et al., 2019). Even though a constant delay between infection and detection may not affect the estimate of the growth rate, it can still affect the associated credible intervals. Other factors related to reporting — including changes in case definition, saturation in diagnostic test capacity, transparency of data, and representativeness of samples — will also affect the inference. Finally, generation intervals can become shorter throughout an epidemic as intervention strategies, such as quarantine, can reduce the infectious period (Hethcote et al., 2002); since we are primarily focusing on the outbreak in Wuhan city before confinement, generation intervals are unlikely to vary significantly. All of these factors, including fitting to cumulative curves or ignoring process errors, affect the estimation of the exponential growth rate (as well as the associated uncertainties), which in turn affects the estimation of the basic reproductive number."

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

We explain why we do not consider revised estimates in several places. See above in the discussion about pre-prints.

Page 6, Table 2. Please recheck the parameters of Study 5, which is inconsistent with the 90confidence 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.

We match probability, not necessarily equi-tailed quantiles. A Gamma distribution distribution with mean 2.2 and shape 12 has a 90% probability of containing a value between 1.27 and 3.34. We make this clear in our main text now:

"We used gamma distributions to model values reported with confidence or credible intervals and uniform distributions to model values reported with ranges; when confidence or credible intervals are reported, we parameterize the gamma distribution such that (i) its mean matches the estimated value

and (ii) the probability that a random variable following the specified gamma distribution falls between the lower and upper confidence or credible limits is equal to the reported confidence or credible level. This probability does not need to be equi-tailed. For example, Study 3 estimated $\mathcal{R}_0 = 2.92$ (95% CI: 2.28–3.67); we model this estimate as a gamma distribution with a mean of 2.92 and a shape parameter of 67, which has a 95% probability of containing a value between 2.28 and 3.67 (see Table 2 for a complete description)."

Page 6, Line 29. Are the priors for μ_r, μ_G, and μ_κ weakly informative? Did the authors check the sensibility with new priors?

We believe that these priors are sufficiently wide and therefore did not try different priors. We provide a better justification in the main text:

"These priors are chosen such that their 95% quantiles are sufficiently wider than biologically realistic parameter ranges. Specifically, 95% quantiles for μ_r, μ_G, and μ_κ are 0.02–0.40 days^{-1}, 0.8–19.5 days, and 0.1–1.4, respectively. Parameters that are outside these ranges are biologically unrealistic for SARS-CoV-2 outbreaks. Therefore, we do not expect our results to be sensitive to these priors."

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

Done.

"Weak priors (e.g. half-Cauchy(0,5)) can lead to inefficient sampling and poor convergence."

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?

We now check the sensitivity of the value of kappa of Study 2 on the pooled estimate of the basic reproduction number.

"We note that our estimate of $\mathcal{R}_{\text{pool}}$ is relatively insensitive to our assumption of $\kappa = 0.5$ for Study 2: assuming $\kappa = 0.1$ gives $\mathcal{R}_{\text{pool}} = 3.0$ (95% CI: 2.2–4.7), whereas assuming $\kappa = 0.9$ gives $\mathcal{R}_{\text{pool}} = 2.9$ (95% CI: 2.1–4.4)."

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.

"Even though some studies consider reported cases up to Janaury 26, 2020 — three days after the travel restriction that took place on January 23, 2020 (Tian et al., 2020) — the exponential growth assumption can still describe the number of reported cases reasonably well; given the incubation period of around 5 days (Lauer et al., 2020) as well as reporting delay of around 5 days (Sun et al., 2020), the majority of reported cases during study periods are likely to have been infected prior to the travel ban."

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".

We have changed this explanation (and added Sensitivity Analysis section in the Methods):

"Finally, Figure 3 compares the base estimates (based on r_i, $\bar{G}_i$, and κ_i for each study i) and 21 substitute estimates (3 parameter substitutions $\times$ 7 studies). All but 8 substitute estimates have wider credible intervals compared to their corresponding base estimates — namely, $\bar{G}$-substitute estimates for Study 1 and 7, r-substitute estimates for Study 1 and 2, and κ-substitute estimates for Study 3, 6, and 7."

Reviewer #3

Here, Park and colleagues present a straight-forward, but important analysis of 7 early estimates of R0 for the ongoing covid outbreak. They stress that multiple forms of uncertainty have to be taken into account when calculating R0 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.

References

Fisman, D. N., T. S. Hauck, A. R. Tuite, and A. L. Greer (2013). An IDEA for short term outbreak projection: nearcasting using the basic reproduction number. *PloS One 8*(12).

Hethcote, H., M. Zhien, and L. Shengbing (2002). Effects of quarantine in six endemic models for infectious diseases. *Math Biosci 180*(1-2), 141–160.

Lauer, S. A., K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, A. S. Azman, N. G. Reich, and J. Lessler (2020). The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. *Ann Intern Med*.

McBryde, E., I. Bergeri, C. van Gemert, J. Rotty, E. Headley, K. Simpson, R. Lester, M. Hellard, and J. E. Fielding (2009). Early transmission characteristics of influenza A (H1N1) v in Australia: Victorian state, 16 May–3 June 2009. *Euro Surveill 14*(42), 19363.

Nishiura, H., C. Castillo-Chavez, M. Safan, and G. Chowell (2009). Transmission potential of the new influenza A (H1N1) virus and its age-specificity in Japan. *Euro Surveill 14*(22), 19227.

Nishiura, H. and G. Chowell (2015). Theoretical perspectives on the infectiousness of Ebola virus disease. *Theor Biol Med Model 12*(1), 1.

Park, S. W., D. Champredon, J. S. Weitz, and J. Dushoff (2019). A practical generation-interval-based approach to inferring the strength of epidemics from their speed. *Epidemics 27*, 12–18.

Read, J. M., J. R. Bridgen, D. A. Cummings, A. Ho, and C. P. Jewell (2020a). Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions. https://www.medrxiv.org/content/10.1101/2020.01.23.20018549v1. Accessed 26, January, 2020.

Read, J. M., J. R. Bridgen, D. A. Cummings, A. Ho, and C. P. Jewell (2020b). Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions. https://www.medrxiv.org/content/10.1101/2020.01.23.20018549v2. Accessed 5, February, 2020.

Roberts, M. G. and H. Nishiura (2011). Early estimation of the reproduction number in the presence of imported cases: pandemic influenza H1N1-2009 in New Zealand. *PLoS One 6*(5).

Sun, K., J. Chen, and C. Viboud (2020). Early epidemiological analysis of the coronavirus disease 2019 outbreak based on crowdsourced data: a population-level observational study. *Lancet Dig Health*.

Tariq, A., K. Roosa, K. Mizumoto, and G. Chowell (2019). Assessing reporting delays and the effective reproduction number: The Ebola epidemic in DRC, May 2018–January 2019. *Epidemics 26*, 128–133.

Tian, H., Y. Liu, Y. Li, C.-H. Wu, B. Chen, M. U. G. Kraemer, B. Li, J. Cai, B. Xu, Q. Yang, B. Wang, P. Yang, Y. Cui, Y. Song, P. Zheng, Q. Wang, O. N. Bjornstad, R. Yang, B. T. Grenfell, O. G. Pybus, and C. Dye (2020). An investigation of transmission control measures during the first 50 days of the covid-19 epidemic in china. *Science*.

Trichereau, J., C. Verret, A. Mayet, G. Manet, C. Decam, J.-B. Meynard, X. Deparis, and R. Migliani (2012). Estimation of the reproductive number for A (H1N1) pdm09 influenza among the French armed forces, September 2009–March 2010. *J Infect 64*(6), 628–630.

Wallinga, J. and M. Lipsitch (2007). How generation intervals shape the relationship between growth rates and reproductive numbers. *Proc R Soc Lond B Biol Sci 274*(1609), 599–604.

World Health Organization (2020). Laboratory testing for 2019 novel coronavirus (2019-nCoV) in suspected human cases. https://www.who.int/publications-detail/laboratory-testing-for-2019-novel-coronavirus-in-suspected-human-cases-20200117. Accessed February 4, 2020.

Zhao, S., J. Ran, S. S. Musa, G. Yang, Y. Lou, D. Gao, L. Yang, and D. He (2020). Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. https://www.biorxiv.org/content/10.1101/2020.01.23.916395v1. Accessed 26, January, 2020.