Journal of Data Science 18 (3), 536–549
July 2020
DOI: 10.6339/JDS.202007_18(3).0019

Meta-Analysis of Several Epidemic Characteristics of COVID-19

Panpan Zhang*[1], Tiandong Wang[2], and Sharon X. Xie[1]

[1]*Department of Biostatistics, Epidemiology and Informatics, University of Pennsylvania, Philadelphia, Pennsylvania, USA*

[2]*Department of Statistics, Texas A&M University, College Station, Texas, USA*

Abstract

As the COVID-19 pandemic has strongly disrupted people's daily work and life, a great amount of scientific research has been conducted to understand the key characteristics of this new epidemic. In this manuscript, we focus on four crucial epidemic metrics with regard to the COVID-19, namely the basic reproduction number, the incubation period, the serial interval and the epidemic doubling time. We collect relevant studies based on the COVID-19 data in China and conduct a meta-analysis to obtain pooled estimates on the four metrics. From the summary results, we conclude that the COVID-19 has stronger transmissibility than SARS, implying that stringent public health strategies are necessary.

Keywords *basic reproduction number, epidemic doubling time, incubation period, sensitivity analysis, serial interval.*

1 Introduction

Coronavirus disease (COVID-19) is an infectious disease caused by *severe acute respiratory syndrome coronavirus 2* (SARS-CoV-2), a newly discovered coronavirus, which leads to respiratory illness and can be transmitted from person to person. Ever since December 2019, when the first case of COVID-19 in Wuhan, P. R. China (or mainland China) was reported, the novel coronavirus has hit most of the countries in the world, with the United States (U.S.) being the one having the largest number of confirmed cases (Worldometers, 2020). On March 11, 2020, the World Health Organization (WHO) declared COVID-19 a global pandemic. As of May 9, 2020, the WHO has reported a total of 3,855,788 confirmed cases all over the world and the total number of deaths has reached 265,862 (https://covid19.who.int/).

The COVID-19 pandemic has significant negative impacts on both the global health and the economy. In the U.S., for example, the unemployment rate has jumped up to 14.7% in April, 2020, reaching the highest rate and the largest monthly increase since January 1948 (U.S. Bureau of Labor Statistics, https://www.bls.gov/bls/newsrels.htm). Continuous efforts have been made by every country in the world to slow the spread of the disease and mitigate the associated negative impacts on various aspects of the society.

So far, a great amount of scientific research has been conducted on COVID-19, ranging from ongoing clinical trials that evaluate potential treatments to statistical analyses on the characteristics of this infectious disease. With more and more COVID-19 data and studies available, it is vitally important to aggregate the information and pool the statistical findings. Hence, here we conduct a meta-analysis based on published studies of the COVID-19 outbreak in mainland China.

Our meta-analysis that accounts for between-study heterogeneity concentrates on four common epidemic metrics that characterize the transmission of the COVID-19:

*Corresponding author. Email: panpan.zhang@pennmedicine.upenn.edu.

Received May, 2020; Accepted July, 2020

(i) *Basic reproduction number*: Often referred to as R_0, the basic reproduction number measures the contagiousness or transmissibility of infectious agents and is interpreted as the expected number of infections caused directly by one case in a completely susceptible population.

(ii) *Incubation period*: This metric is defined as the number of days between when an individual was actually infected and when this person starts to show symptoms. Understanding the incubation period of the COVID-19 is crucial as it provides guidelines on deciding a reasonable length of the quarantine period.

(iii) *Serial interval*: Defined as the time between the start of symptoms in the primary patient (infector) and onset of symptoms in the patient being infected by the infector (the infectee), the serial interval is critical in the calculation of R_0.

(iv) *Epidemic doubling time*: This metric measures the period of time needed for the total number of cases in the epidemic to double, and is an important factor that reflects the speed at which the COVID-19 is spreading.

The rest of this manuscript is organized as follows. We first provide details on the data collection process in Section 2. Then in Section 3, we introduce the key ingredients for a meta-analysis, where the modeling and estimation methods are outlined. Section 4 lists a detailed summary for the four epidemic metrics as given above. Statistical results from the meta-analysis are reported in Section 5, and sensitivity analysis is given in Section 6. Concluding remarks are then disclosed in Section 7.

2 Study Selection

We conduct a comprehensive literature screening for the articles published in scientific journals (including early versions) or preprint platforms, e.g., *medRxiv* and *bioRxiv*. The key words of our searching are COVID-19 (SARS-CoV-2 or 2019-nCOV), the epidemic characteristics of interest (basic reproduction number, incubation period, serial interval and epidemic doubling time), and China, where the selection criteria is relaxed for regional labels. Specifically, we include the studies in the whole country of China, China except for Hubei province (where the city of Wuhan is located), a list of provinces and municipalities, a collection of cities or even a specific region, in order to increase the sample size and the potential statistical power. The variability and heterogeneity across the selected studies are accounted when adopting appropriate methodologies in the meta-analysis.

In addition, we only include studies reporting unambiguous estimates and the associated 95% confidence intervals or standard deviations. If the standard deviation of an estimator is given, we calculate a 95% confidence interval under the assumption of normal approximation provided that the fitted model is valid. This selection criterion is needed to ensure the consistency of analysis results. Besides, we do not include the articles, where the inference about the individual metrics are made based upon aggregate statistics. The inclusion of these studies in meta-analysis usually cause *ecological bias* (also known as aggregation bias or ecological fallacy in the literature).

Some studies included here provide more than one estimates that are obtained from various methods or models. For those studies, we select one estimate and its associated confidence interval in our analysis. We illustrate our choices and the corresponding reasons in details in Section 5.

3 Meta-analysis

Meta-analysis is a statistical procedure aiming to combine scientific results from multiple comparable studies or trials. It is one of the most popular analytical tools in the statistical analysis, which derives a pooled estimate by aggregating relevant information, thus increasing the statistical power. In particular, when comparing different studies addressing the same question, the crux is to measure the standardized difference across various results, i.e. the *effect size*. Meta-analysis is a useful tool for estimating a pooled effect size by combining multiple studies, and potentially improves the precision of estimation.

In general, there are two methods to pool effect sizes from multiple studies: the *fixed-effects model* (FEM) and the *random-effects model* (REM). A FEM assumes all included studies come from the same population, whereas a REM is constructed under the assumption that data are collected from different populations. Although we restrict our data search to mainland China only, there may still exist potential variations in the populations across the studies. Hence, we choose the REM in our analysis.

Let Y_k be the estimator in the k-th study, for $k = 1, 2, \ldots, K$, where K is the total number of collected studies. A REM assumes a *normal-normal hierarchical* model:

$$\begin{aligned} Y_k|\theta_k &\sim \mathcal{N}(\theta_k, s_k^2), \\ \theta_k &\sim \mathcal{N}(\mu, \tau^2). \end{aligned}$$

Here θ_k is the parameter of interest, and s_k^2 is the variance of Y_k. The hyper-parameters μ and τ^2 correspond to the mean effect and the across-study variance that reflects heterogeneity of the population, respectively.

The inference for REM is conducted sequentially as follows: we first estimate the heterogeneity variance τ^2 (denoted as $\hat{\tau}^2$); then given $\hat{\tau}^2$, we obtain an estimate of the effect μ. There are a variety of estimation methods for τ^2; see Veroniki et al. (2016) for a concise survey. We adopt the *Hartung-Knapp-Sidik-Jonkman* (HKSJ) method (Hartung and Knapp, 2001; Sidik and Jonkman, 2002), which is shown to be more robust when the number of studies is small and there is a substantial heterogeneity in the population.

Given $\hat{\tau}^2$, the conditional maximum likelihood estimator of the effect μ becomes

$$\hat{\mu} = \frac{\sum_k w_k y_k}{\sum_k w_k},$$

where the weights are specified as

$$w_k = \frac{1}{s_k^2 + \hat{\tau}^2}, \qquad k = 1, 2, \ldots, K.$$

The weight w_k can be interpreted as the inverse of the variance of the estimate of the k-th study. Then we use the normal approximation to construct confidence intervals, where the standard deviation of $\hat{\mu}$ is given by

$$\hat{\sigma}_\mu = \sqrt{\frac{1}{\sum_k w_k}}.$$

Later in Section 5, we also provide prediction intervals, which are equally significant components in a meta-analysis. In the presence of a substantial heterogeneity index, prediction intervals are recommended to be reported alongside with confidence intervals as they not only quantitatively

provide a range for the effect size of a new study, but also measure the uncertainty of the estimate in a way that acknowledges the heterogeneity (Higgins et al., 2009; Guddat et al., 2012; Nagashima et al., 2019).

4 Epidemic Characteristics

In this section, we briefly introduce the epidemic characteristics that are investigated in this manuscript.

Basic Reproduction Number In epidemiology, the *basic reproduction number*, denoted R_0, is the expected number of cases caused by one case in a completely susceptible population. It is a critical metric to describe the contagiousness or transmissibility of infectious diseases (Delamater et al., 2019). The estimation of R_0 is primarily based on compartmental models, where the *susceptible-infectious-recovered* (SIR) model and its extensions are the most commonly used. A variety of methods have been developed to estimate R_0, such as maximum likelihood methods and Bayesian approaches; we refer the interested readers to Dietz (1993) for a concise survey and to Nikbakht et al. (2019) for a practical comparison of the methods. For the users of statistical software `R`, the package `R0` includes most standard methods for the estimation of R_0.

Since the outbreak of COVID-19, R_0 has been one of the most critical metrics receiving substantial interest in the community, as it provides a basic benchmark (with threshold 1) to define a pandemic. Besides, R_0 helps indicate the potential severity of an epidemic outbreak. It is evident that R_0 is closely related to the fraction of a susceptible population who will eventually be infected (Holme and Masuda, 2015). Furthermore, R_0 is an indispensable component when estimating the effective reproduction numbers (denoted as R_t), which are often used to assess the effectiveness of the intervention procedure to mitigate the spread of an epidemic.

Incubation Period The second characteristic we investigate here is the incubation period. The incubation period of an epidemic is the period from the time of the contact of a transmission source (susceptible or confirmed infector) and the time of symptom onset. The incubation period provides important information during an outbreak, as it helps to determine when the infected individuals who are symptomatic are most likely to spread the disease, and signal necessary public health activities such as monitoring, surveillance and intervention.

From a statistical perspective, the incubation period is a crucial factor to model the current and future trends of an epidemic, as well as evaluate intervention strategies. One of the biggest challenges of estimating the incubation period is that data are often coarsely observed, in addition to several other issues such as censoring and selection bias. The estimation of incubation period is generally based on the methods developed in Reich et al. (2009) and the extensions.

Serial Interval The *serial interval* is defined as the time duration between the onset of symptoms in the primary patient and the onset of symptoms in the secondary patient who receives the disease from that primary one (Lipsitch et al., 2003). It is the sum of the latent period and the duration of infectiousness. Being another crucial factor for constructing epidemiological models, the serial interval is one of the fundamentals for computing and estimating R_0; see Fine (2003) for a summary of the importance of serial interval in epidemiological studies.

The estimation of serial interval is based on the generation time distribution, which is closely related to the infection rate of a epidemic. Standard estimation procedures are given

in Diekmann et al. (2013, Chapter 13). In practice, the `R` package `R0` (Boelle and Obadia, 2015) collects functions that can be used to estimate serial intervals via maximum likelihood methods (White et al., 2009), and the estimation of serial interval in package `EpiEstim` (Cori, 2020) is based on the method developed in Cori et al. (2013).

Epidemic Doubling Time The *epidemic doubling time* (or simply doubling time) is another important index in epidemic studies, as it measures the length of the period during which the number of confirmed cases is doubled. It is evident that the doubling time is inversely related to another epidemic parameters of interest: case-fatality rate. Hence, learning doubling time helps epidemiologists not only understand the transmissibility of an infectious disease but also evaluate its severity. In the course of pathology, the doubling time is useful for analyzing the growth rate of the virus. Doubling time can also be used to assess the effectiveness of public health interventions and protocols, as an increase in the doubling time usually indicates a slowdown in epidemic transmission.

The estimation of the epidemic doubling time is generally model based, where exponential growth models are the most frequently adopted (Galvani et al., 2003; Du et al., 2020c). In this manuscript, we conduct a meta-analysis on the doubling time to study the growth rate of COVID-19.

5 Results

In this section, we apply the methods introduced in Section 3 to estimate the epidemic characteristics of interest listed in Section 4. The analysis is primarily done in `R`, where several standard packages for meta-analysis are utilized: `meta` (Balduzzi et al., 2019) and `dmetar` (Harrer et al., 2019).

Basic Reproduction Number Relevant studies used for our meta-analysis are summarized in Table S1 (in the supplementary document), where a total of 12 research articles are included (Cao et al., 2020; Kucharski et al., 2020; Li et al., 2020; Liu et al., 2020a; Imai et al., 2020; Read et al., 2020; Read, 2020; Shen et al., 2020; Sun et al., 2020; Tian et al., 2020; Wu et al., 2020b; Zhu and Chen, 2020). Note that the number of collected studies is not large, so we depict a *funnel plot* in Figure S1 (in the supplementary document) to see whether there is any bias owing to the small sample size. In a funnel plot, a significantly asymmetry pattern indicates the existence of the publication bias. A more rigorous method is to exploit the *Egger's test* (Egger et al., 1997), for which a small p-value suggests rejecting the null hypothesis (i.e., some bias caused by the small sample size does exist). Recently, an alternative approach has been proposed by Lin and Chu (2018), but is not considered in the present article. The p-value of the Egger's test here is 0.47, showing that no further procedure is needed for correcting the effect size for the present meta-analysis.

Using the estimation procedure demonstrated in Section 3, we see that the estimate of R_0 is 3.19 with a 95% confidence interval $[2.63, 3.74]$ and a 95% prediction interval $[1.28, 5.09]$. The associated *forest plot* is presented in Figure 1. Specifically, *Hedges' g* is a standard measure of effect size proposed by Hedges (1981). The index of heterogeneity, I^2, which ranges from 0 to 1, is used to quantify the dispersion of effect sizes. Here we have $I^2 = 97\%$, indicating a substantial heterogeneity in the population. Hence, the REM is indeed more appropriate than the FEM for our meta-analysis.

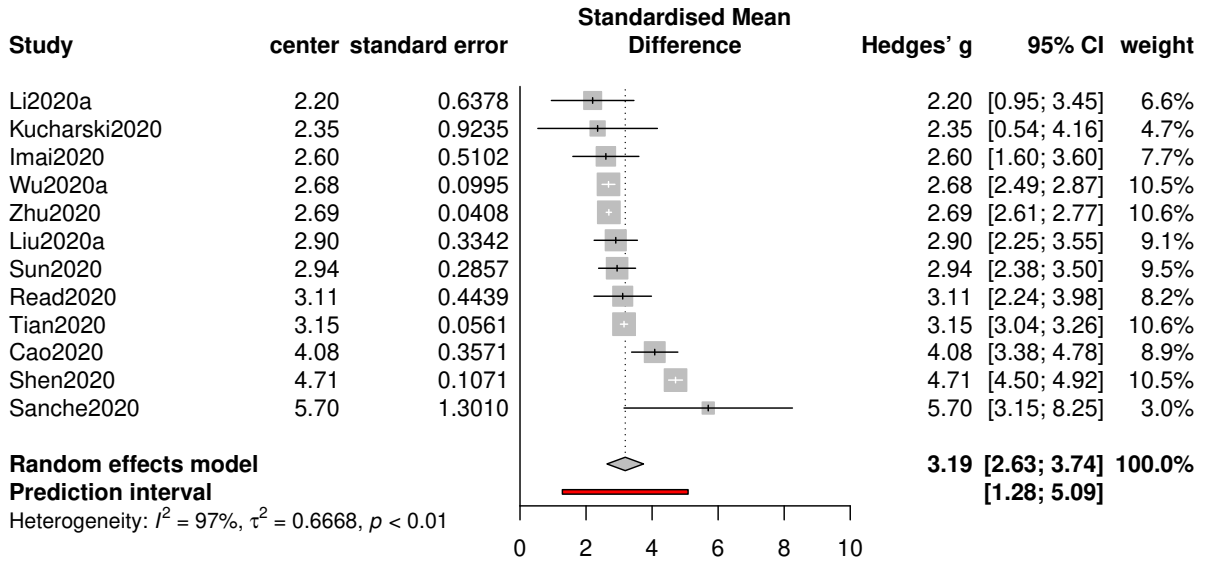

Figure 1: Forest plot of meta-analysis for R_0.

With high heterogeneity in the underlying population, prediction intervals that incorporate heterogeneity are more informative than confidence intervals that focus only on summary estimates. Prediction intervals become useful for predicting the potential effect of an individual study that may be considerably different from the average effect (Riley et al., 2011). Hence, it is recommended to routinely report prediction intervals to allow more informative inferences in REMs. In the present meta-analysis, we report both of the confidence interval and the prediction interval when inferring epidemic characteristics (including R_0).

Based on the results from our meta-analysis, we conclude that the basic reproduction number (R_0) of COVID-19 appears to be greater than that of SARS (point estimate around 3), as reported by WHO in "Consensus document on the epidemiology of severe acute respiratory syndrome (SARS)". However, it is not as large as an average-based estimate (3.28) for COVID-19 reported in Liu et al. (2020b). Compared with the results from other two meta-analyses on COVID-19, our estimate is slightly greater than 3.15 reported in He et al. (2020a), and moderately greater than 3.05 reported in Dong et al. (2020). Nonetheless, we do not observe statistically significant difference in either case, but the high value of R_0 alerts the community that COVID-19 is a highly contagious disease.

Incubation Period The articles collected for analyzing incubation period are: Backer et al. (2020); He et al. (2020b); Lauer et al. (2020); Leung (2020); Linton et al. (2020); Liu et al. (2020a); Han (2020); Qin et al. (2020); Read (2020); Xia et al. (2020). Analogous to the previous part, we give the funnel plot in Figure S2 (in the supplementary document), from which a (roughly) symmetric pattern is observed. This is consistent with the p-value (0.84) from the Egger's test. Hence, again, publication bias is not present here.

The estimate of the mean of the incubation period of COVID-19 based on our meta-analysis is 5.34 (days), with a 95% confidence interval $[4.29, 6.40]$ and a 95% prediction interval $[1.97, 8.73]$; see Figure 2. Our result is greater than the median (4 with *interquartile range* (IQR) from 2 to 7) of the incubation period estimated by Guan et al. (2020), which is not included in our

meta-analysis as no 95% confidence interval is provided therein. One possible reason is that the study period of Guan et al. (2020) is between December 11, 2019 and January 29, 2020, which is considered as a relatively early stage of the COVID-19 outbreak in mainland China. Besides, our estimate is larger than that (5.08) from another meta-analysis in He et al. (2020a).

Our meta-analysis also suggests that the incubation period of COVID-19 is a bit longer than that of SARS (commonly ranging from 3 to 5 according to WHO), and our finding is close to the estimate of the incubation period of COVID-19 by the U.S. Centers for Disease Control and Prevention (CDC). It is evident that the distribution of incubation period is usually heavy-tailed (Backer et al., 2020; Lauer et al., 2020). Our result supports the 14-day monitoring and quarantine periods in implementation.

Serial Interval For this part, we collect 10 studies to proceed along our meta-analysis (Bi et al., 2020; Du et al., 2020c,b; He et al., 2020b; Li et al., 2020; Zhao et al., 2020; Tindale et al., 2020; Wu et al., 2020a; Zhang et al., 2020). Albeit the funnel plot in Figure S3 (in the supplementary document) displaying an asymmetric pattern, the p-value from the Egger's test is 0.26, suggesting that it is not required to implement a correction procedure. The estimate of the serial interval is 5.40 with a 95% confidence interval $[4.53, 6.26]$ and a 95% prediction interval $[2.89, 7.90]$. The associated forest plot is given in Figure 3.

The estimate of the serial interval of COVID-19 based on our meta-analysis is close to that of SARS (5–6 days according to WHO's report in "Consensus document on the epidemiology of severe acute respiratory syndrome (SARS)"). A shorter serial interval of COVID-19, together with a shorter mean incubation period, suggests higher possibility that a transmission is completed before the onset of symptoms. Therefore, reducing the source of transmission (by hospitalizing infected individuals or implementing "stay-at-home" protocols to susceptible individuals) and reasonably extending the quarantine period are extensively helpful to slow the progression of COVID-19.

Epidemic Doubling Time The number of the collected studies for epidemic doubling time is 8 (less than 10); they are: Du et al. (2020a); Kraemer et al. (2020); Li et al. (2020); Muniz-Rodriguez et al. (2020); Volz et al. (2020); Wu et al. (2020b,a). As a small sample size is likely to cause bias in a meta-analysis (Lin, 2018), the visualization of the funnel plot in Figure S4 (in the supplementary document) exhibits asymmetry.

We consider the *trim-and-fill procedure* (Duval and Tweedie, 2000) based on the motivation of reducing the bias owing to a small sample size, and the estimation results before and after implementing the trim-and-fill procedure are given in Figure 4. The estimate without correction is 4.86 (with a 95% confidence interval $[3.26, 6.45]$), which is larger than the trimmed estimate of value 3.48 (with a 95% confidence interval $[1.60, 5.35]$). Hence, even though the p-value of the Egger's test is not significant, a remedy is still necessary, as an extensively small sample size usually jeopardizes the statistical power of the Egger's test. In contrast, when we apply the trim-and-fill procedure to the previous three epidemic metrics, no significant difference has been detected.

Although there seems to be no official report on the epidemic doubling time of SARS from WHO or CDC, we find that an estimate of the doubling time of SARS from a published article (Galvani et al., 2003) is 16.3 through a literature search. The estimate of the doubling time of SARS is three times more than that of COVID-19 based on our analysis, suggesting that COVID-19 is a more contagious disease.

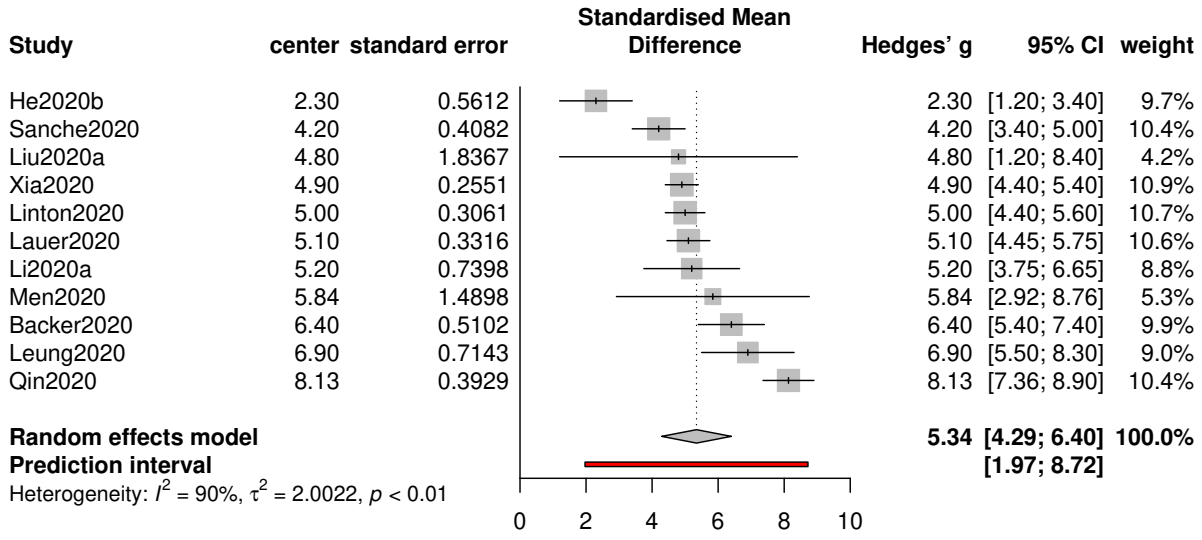

Figure 2: Forest plot of meta-analysis for incubation period.

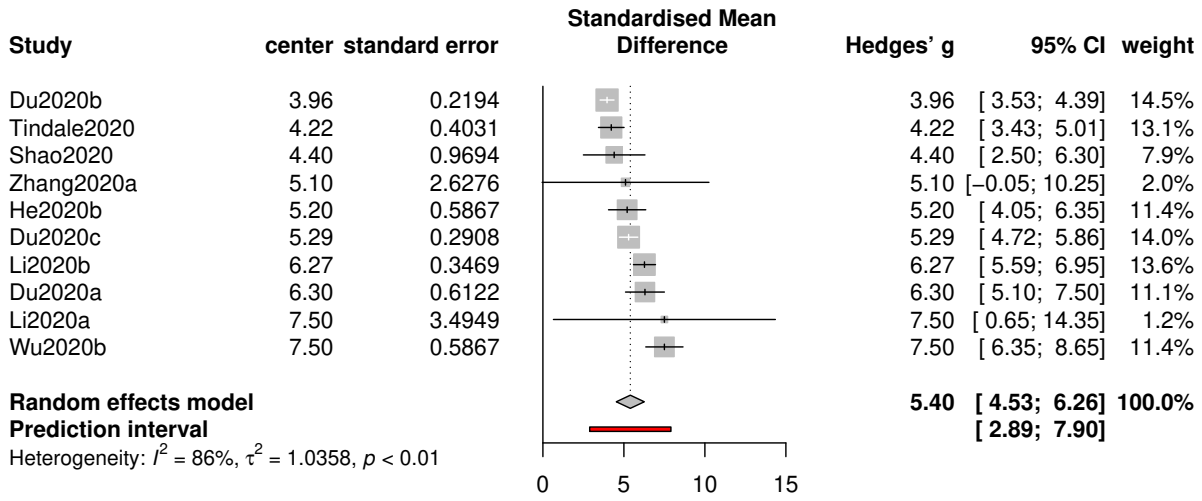

Figure 3: Forest plot of meta-analysis for serial interval.

6 Sensitivity Analysis

Meta-analysis and sensitivity analysis usually go hand in hand. While the meta-analysis focuses on the summary of a systematic review of relevant studies, the sensitivity analysis is used to assess the robustness of the results from the meta-analysis. In practice, the sensitivity analysis is a repeat of the meta-analysis, substituting alternative studies or the results from unclear studies. In other words, the goal of sensitivity analysis is to explore the impact of the meta-analysis by including or excluding studies in the meta-analysis based on some criteria as in Higgins et al. (2019, Section 9.7). In this section, we present a couple of sensitivity analyses for the basic reproduction number R_0. Analogous studies can be carried out for other epidemic characteristics mutatis mutandis. For the sake of brevity, we present the details on the sensitivity analysis of R_0 only in this section.

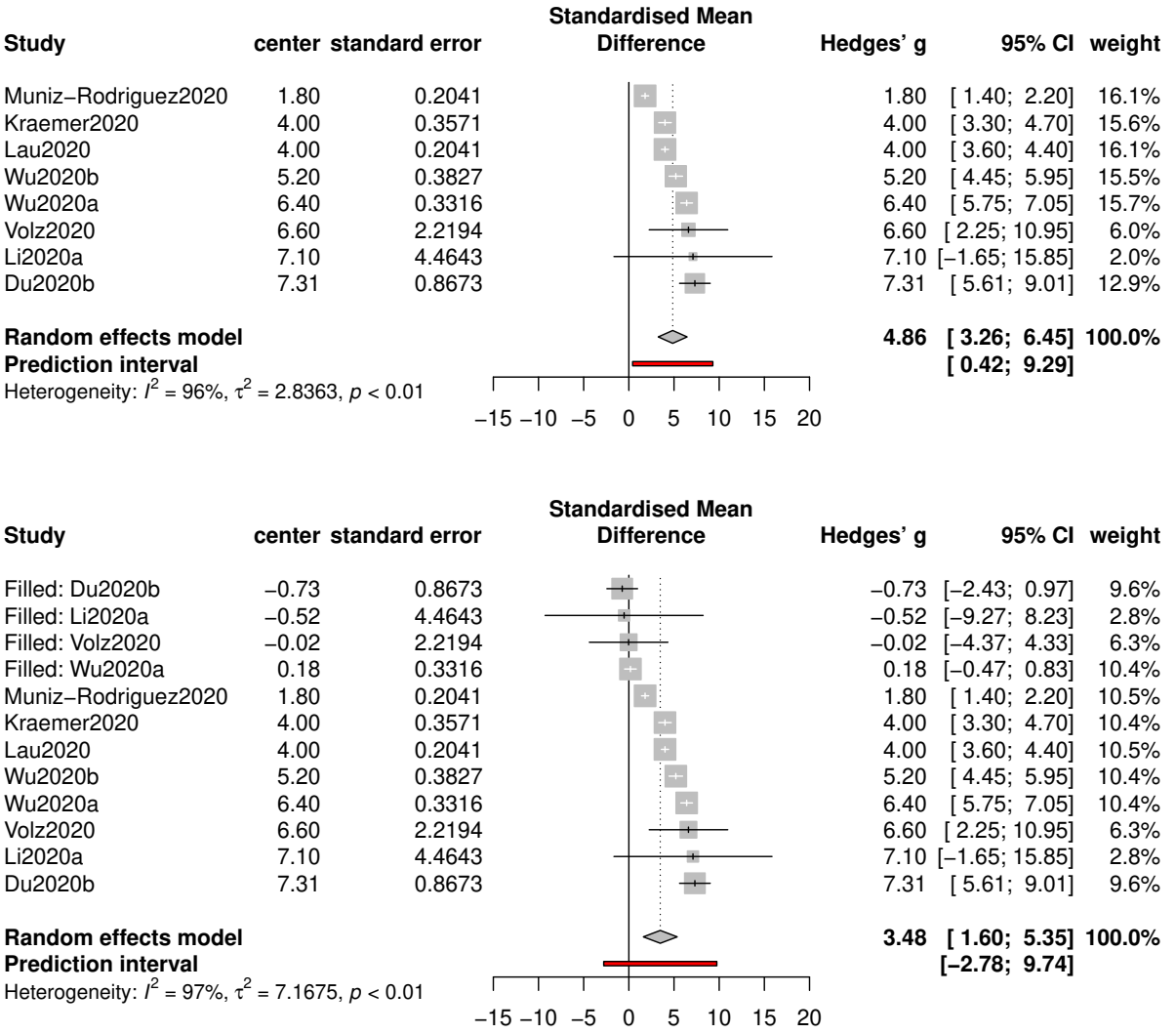

Figure 4: Forest plots of meta-analysis for epidemic doubling time before (top) and after (bottom) implementing the trim-and-fill procedure.

In the first sensitivity analysis, we only include the published articles (a total number of 7 left) in the new analysis, by leaving out preprints that have not yet been through the peer review process. The new estimate of R_0 is 2.85 with a 95% confidence interval $[2.09, 3.60]$. Having observed a large value of heterogeneity index ($I^2 = 89\%$), we accordingly report a 95% prediction interval $[0.58, 5.11]$ as well.

Now instead of leaving out ambiguous results, we expand the scope of our study to the East Asia. We add several studies from Japan, Korea and the Diamond Princess Cruise, which are listed in Table S5 (in the supplementary document). The estimate of R_0 for this analysis is 2.82 with a 95% confidence interval $[2.35, 3.28]$ and a prediction interval $[0.93, 4.71]$.

Table 1: Summary of our numerical findings with respect to R_0 or the basic reproduction number, the incubation period, the serial interval and the doubling time of the epidemic.

Metric	Est.	95% CI	95% PI	$\geq$ or $\leq$ SARS (Est.)
Basic Reproduction Number	3.19	$[2.63, 3.74]$	$[1.28, 5.09]$	$\geq$ (3)
Incubation Period	5.34	$[4.29, 6.40]$	$[1.97, 8.72]$	$\geq$ $(3 \sim 5)$
Serial Interval	5.40	$[4.53, 6.26]$	$[2.89, 7.90]$	$\approx$ $(5 \sim 6)$
Epidemic Doubling Time	3.48	$[1.60, 5.35]$	$[-2.78, 9.74]$	$\leq$ (16.3)

7 Concluding Remarks

In this meta-analysis, we study a collection of recent studies on the COVID-19 pandemic, focusing mainly on four epidemic characteristics: R_0 or the basic reproduction number, the incubation period, the serial interval, and the doubling time of the epidemic. We summarize our numerical findings in Table 1 and include corresponding comparisons with SARS, which is also a viral respiratory illness caused by a coronavirus in 2003. From Table 1, we see that compared to SARS, the COVID-19 has a larger R_0, a longer incubation period and much shorter doubling time, thus suggesting this novel coronavirus be more contagious and stringent public health strategies be necessary.

The numerical results also provide insights on further studies. For example, with pooled estimates of R_0 and the doubling time available, we can then compare them with the *effective reproduction number* (or R_t) or the doubling time after the nationwide lockdown protocol has been implemented in China to see the effectiveness of the public health strategies. The estimates of the incubation period and the serial interval reassure the necessity of reinforcing a 14-day quarantine period to prevent the spread of the disease.

It is also worthwhile noting the potential limitations of the current meta-analysis. To ensure the accuracy of the estimates collected, we do not include publications that report estimates with a 90% confidence interval or estimates with quartiles. This may lead to some loss of precision or publication bias for our estimates in meta-analysis. Meanwhile, the study on the COVID-19 may not be only restricted to the four characteristics listed in the manuscript. Other metrics, e.g., the case-to fatality rate and the testing capacity, and factors, e.g., significant clusters and patients' underlying chronic medical conditions, are also of great importance in the future endeavors to mitigate the negative impacts of the COVID-19 pandemic.

Supplementary Materials

The Supplementary Materials, including the document mentioned in the text and the data/code used in the analyses, can be found on the *Journal of Data Science* website.

Acknowledgments

The authors would like to thank the AE and two anonymous referees for their insightful comments. Dr. Zhang and Dr. Xie received support from the National Institute of Health (NIH) grant R01-NS102324.

References

Backer JA, Klinkenberg D, Wallinga J (2020). Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020. *Eurosurveillance*, 25(5): 2000062.

Balduzzi S, Rücker G, Schwarzer G (2019). How to perform a meta-analysis with r: A practical tutorial. *Evidence-Based Mental Health*, 22(4): 153–160.

Bi Q, Wu Y, Mei S, Ye C, Zou X, Zhang Z, et al. (2020). Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: A retrospective cohort study. *The Lancet Infectious Diseases*. Forthcoming, https://doi.org/10.1016/S1473-3099(20)30287-5.

Boelle PY, Obadia T (2015). *R0: Estimation of R0 and Real-Time Reproduction Number from Epidemics*. R package version 1.2-6.

Cao Z, Zhang Q, Lu X, Pfeiffer D, Jia Z, Song H, et al. (2020). Estimating the effective reproduction number of the 2019-nCoV in China. MedRxiv preprint: https://doi.org/10.1101/2020.01.27.20018952.

Cori A (2020). *EpiEstim: Estimate Time Varying Reproduction Numbers from Epidemic Curves*. R package version 2.2-3.

Cori A, Ferguson NM, Fraser C, Cauchemez S (2013). A new framework and software to estimate time-varying reproduction numbers during epidemics. *American Journal of Epidemiology*, 178(9): 1505–1512.

Delamater PL, Street EJ, Leslie TF, Yang YT, Jacobsen KH (2019). Complexity of the basic reproduction number (R0). *Emerging Infectious Diseases*, 25(1): 1–4.

Diekmann O, Heesterbeek H, Britton T (2013). *Mathematical Tools for Understanding Infectious Disease Dynamics*, volume 7. Princeton University Press.

Dietz K (1993). The estimation of the basic reproduction number for infectious diseases. *Statistical Methods in Medical Research*, 2(1): 23–41.

Dong J, Zhou Y, Zhang Y, Franz D (2020). A validation study for the successful isolation policy in China: A meta-analysis in COVID-19. MedRxiv preprint: https://doi.org/10.1101/2020.04.15.20065102.

Du Z, Wang L, Cauchemez S, Xu X, Wang X, Cowling BJ, et al. (2020a). Risk for transportation of coronavirus disease from Wuhan to other cities in China. *Emerging Infectious Diseases*, 26(5): 1049–1052.

Du Z, Wu Y, Wang L, Cowling BJ, Meyers LA, et al. (2020b). COVID-19 serial interval estimates based on confirmed cases in public reports from 86 chinese cities. MedRxiv preprint: https://doi.org/10.1101/2020.04.23.20075796.

Du Z, Xu X, Wu Y, Wang L, Cowling BJ, Meyers LA (2020c). Serial interval of COVID-19 among publicly reported confirmed cases. *Emerging Infectious Diseases*, 26(6): 1341.

Duval S, Tweedie R (2000). Trim and fill: A simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, 56(2): 455–463.

Egger M, Smith GD, Schneider M, Minder C (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, 315(7109): 629–634.

Fine PE (2003). The interval between successive cases of an infectious disease. *American Journal of Epidemiology*, 158(11): 1039–1047.

Galvani AP, Lei X, Jewell NP (2003). Severe acute respiratory syndrome: Temporal stability and geographic variation in death rates and doubling times. *Emerging Infectious Diseases*, 9(8): 991–994.

Guan Wj, Ni Zy, Hu Y, Liang Wh, Ou Cq, He Jx, et al. (2020). Clinical characteristics of coronavirus disease 2019 in China. *New England Journal of Medicine*, 382(18): 1708–1720.

Guddat C, Grouven U, Bender R, Skipka G (2012). A note on the graphical presentation of prediction intervals in random-effects meta-analyses. *Systematic Reviews*, 1: 34.

Han H (2020). Estimate the incubation period of coronavirus 2019 (COVID-19). MedRxiv preprint: https://doi.org/10.1101/2020.02.24.20027474.

Harrer M, Cuijpers P, Furukawa T, Ebert DD (2019). *dmetar: Companion R Package For The Guide 'Doing Meta-Analysis in R'*. R package version 0.0.9000.

Hartung J, Knapp G (2001). On tests of the overall treatment effect in meta-analysis with normally distributed responses. *Statistics in Medicine*, 20(12): 1771–1782.

He W, Yi GY, Zhu Y (2020a). Estimation of the basic reproduction number, average incubation time, asymptomatic infection rate, and case fatality rate for COVID-19: Meta-analysis and sensitivity analysis. *Journal of Medical Virology*. Forthcoming, https://doi.org/10.1002/jmv.26041.

He X, Lau EH, Wu P, Deng X, Wang J, Hao X, et al. (2020b). Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nature Medicine*, 26(5): 672–675.

Hedges LV (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, 6(2): 107–128.

Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al. (2019). *Cochrane Handbook for Systematic Reviews of Interventions*. John Wiley & Sons, Chichester, UK.

Higgins JP, Thompson SG, Spiegelhalter DJ (2009). A re-evaluation of random-effects meta-analysis. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 172(1): 137–159.

Holme P, Masuda N (2015). The basic reproduction number as a predictor for epidemic outbreaks in temporal networks. *PLoS ONE*, 10(3): e0120567.

Imai N, Cori A, Dorigatti I, Baguelin M, Donnelly C, Riley S, et al. (2020). Report 3: Transmissibility of 2019-nCoV. http://hdl.handle.net/10044/1/77148.

Kraemer MU, Yang CH, Gutierrez B, Wu CH, Klein B, Pigott DM, et al. (2020). The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science*, 368(6490): 493–497.

Kucharski AJ, Russell TW, Diamond C, Liu Y, Edmunds J, Funk S, et al. (2020). Early dynamics of transmission and control of COVID-19: A mathematical modelling study. *The Lancet Infectious Diseases*, 22: 553–558.

Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, et al. (2020). The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: Estimation and application. *Annals of Internal Medicine*, 172(9): 577–582.

Leung C (2020). Estimating the distribution of the incubation period of 2019 novel coronavirus (COVID-19) infection between travelers to Hubei, China and non-travelers. MedRxiv preprint: https://doi.org/10.1101/2020.02.13.20022822.

Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. (2020). Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *New England Journal of Medicine*, 382: 1199–1207.

Lin L (2018). Bias caused by sampling error in meta-analysis with small sample sizes. *PLoS ONE*, 13(9): e0204056.

Lin L, Chu H (2018). Quantifying publication bias in meta-analysis. *Biometrics*, 74(3): 785–794.

Linton NM, Kobayashi T, Yang Y, Hayashi K, Akhmetzhanov AR, Jung Sm, et al. (2020).

Incubation period and other epidemiological characteristics of 2019 novel coronavirus infections with right truncation: A statistical analysis of publicly available case data. *Journal of Clinical Medicine*, 9(2): 538.

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

Liu T, Hu J, Kang M, Lin L, Zhong H, Xiao J, et al. (2020a). Transmission dynamics of 2019 novel coronavirus (2019-nCoV). BioRxiv preprint: https://doi.org/10.1101/2020.01.25.919787.

Liu Y, Gayle AA, Wilder-Smith A, Rocklöv J (2020b). The reproductive number of COVID-19 is higher compared to SARS coronavirus. *Journal of Travel Medicine*, 27(2): 1–4.

Muniz-Rodriguez K, Chowell G, Cheung CH, Jia D, Lai PY, Lee Y, et al. (2020). Doubling time of the COVID-19 epidemic by Chinese province. MedRxiv preprint: https://doi.org/10.1101/2020.02.05.20020750.

Nagashima K, Noma H, Furukawa TA (2019). Prediction intervals for random-effects meta-analysis: A confidence distribution approach. *Statistical Methods in Medical Research*, 28(6): 1689–1702.

Nikbakht R, Baneshi MR, Bahrampour A, Hosseinnataj A (2019). Comparison of methods to estimate basic reproduction number (R0) of influenza, using canada 2009 and 2017-18 a (H1N1) data. *Journal of Research in Medical Sciences*, 24: 67.

Qin J, You C, Lin Q, Hu T, Yu S, Zhou XH (2020). Estimation of incubation period distribution of COVID-19 using disease onset forward time: A novel cross-sectional and forward follow-up study. MedRxiv preprint: https://doi.org/10.1101/2020.03.06.20032417.

Read JM, Bridgen JR, Cummings DA, Ho A, Jewell CP (2020). Novel coronavirus 2019-nCoV: Early estimation of epidemiological parameters and epidemic predictions. MedRxiv preprint: https://doi.org/10.1101/2020.01.23.20018549.

Read MC (2020). EID: High contagiousness and rapid spread of severe acute respiratory syndrome coronavirus 2. *Emerging Infectious Diseases*, 26(7): 1470–1477.

Reich NG, Lessler J, Cummings DA, Brookmeyer R (2009). Estimating incubation period distributions with coarse data. *Statistics in Medicine*, 28(22): 2769–2784.

Riley RD, Higgins JPT, Deeks JJ (2011). Interpretation of random effects meta-analyses. *BMJ*, 342: d549.

Shen M, Peng Z, Xiao Y, Zhang L (2020). Modelling the epidemic trend of the 2019 novel coronavirus outbreak in China. BioRxiv preprint: https://doi.org/10.1101/2020.01.23.916726.

Sidik K, Jonkman JN (2002). A simple confidence interval for meta-analysis. *Statistics in Medicine*, 21(21): 3153–3159.

Sun H, Qiu Y, Yan H, Huang Y, Zhu Y, Chen SX (2020). Tracking and predicting COVID-19 epidemic in China mainland. MedRxiv preprint: https://doi.org/10.1101/2020.02.17.20024257.

Tian H, Liu Y, Li Y, Wu CH, Chen B, Kraemer MU, et al. (2020). An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science*, 368(6491): 638–642.

Tindale L, Coombe M, Stockdale JE, Garlock E, Lau WYV, Saraswat M, et al. (2020). Transmission interval estimates suggest pre-symptomatic spread of COVID-19. MedRxiv preprint: https://doi.org/10.1101/2020.03.03.20029983.

Veroniki AA, Jackson D, Viechtbauer W, Bender R, Bowden J, Knapp G, et al. (2016). Meth-

ods to estimate the between-study variance and its uncertainty in meta-analysis. *Research Synthesis Methods*, 7(1): 55–79.

Volz E, Baguelin M, Bhatia S, Boonyasiri A, Cori Z Aand Cucunubá, Cuomo-Dannenburg G, et al. (2020). Report 5: Phylogenetic analysis of SARS-CoV-2. https://spiral.imperial.ac.uk/handle/10044/1/77169.

White LF, Wallinga J, Finelli L, Reed C, Riley S, Lipsitch M, et al. (2009). Estimation of the reproductive number and the serial interval in early phase of the 2009 influenza A/H1N1 pandemic in the usa. *Influenza and Other Respiratory Viruses*, 3(6): 267–276.

Worldometers (2020). Reported cases and deaths by country, territory, or conveyance. https://www.worldometers.info/coronavirus/. Accessed: 2020-05-09.

Wu JT, Leung K, Bushman M, Kishore N, Niehus R, de Salazar PM, et al. (2020a). Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. *Nature Medicine*, 26(4): 506–510.

Wu JT, Leung K, Leung GM (2020b). Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study. *The Lancet*, 395(10225): 689–697.

Xia W, Liao J, Li C, Li Y, Qian X, Sun X, et al. (2020). Transmission of corona virus disease 2019 during the incubation period may lead to a quarantine loophole. MedRxiv preprint: https://doi.org/10.1101/2020.03.06.20031955.

Zhang J, Litvinova M, Wang W, Wang Y, Deng X, Chen X, et al. (2020). Evolving epidemiology and transmission dynamics of coronavirus disease 2019 outside hubei province, China: A descriptive and modelling study. *The Lancet Infectious Diseases*, 20(7): 793–802.

Zhao S, Gao D, Zhuang Z, Chong MK, Cai Y, Ran J, et al. (2020). Estimating the serial interval of the novel coronavirus disease (COVID-19): A statistical analysis using the public data in hong kong from january 16 to february 15, 2020. MedRxiv preprint: https://doi.org/10.1101/2020.02.21.20026559.

Zhu Y, Chen YQ (2020). On a statistical transmission model in analysis of the early phase of COVID-19 outbreak. *Statistics in Biosciences*. Forthcoming, https://doi.org/10.1007/s12561-020-09277-0.