International Journal of Infectious Diseases 104 (2021) 1–6

Contents lists available at ScienceDirect

International Journal of Infectious Diseases

journal homepage: www.elsevier.com/locate/ijid

The collective wisdom in the COVID-19 research: Comparison and synthesis of epidemiological parameter estimates in preprints and peer-reviewed articles

Yuejiao Wang[a,b], Zhidong Cao[a,*], Daniel Dajun Zeng[a], Qingpeng Zhang[c], Tianyi Luo[a,b]

[a] Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China
[b] School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China
[c] City University of Hong Kong, Hong Kong 999077, China

ARTICLE INFO

Article history:
Received 16 August 2020
Received in revised form 22 November 2020
Accepted 13 December 2020

Keywords:
COVID-19
Preprint
Peer review
Epidemiology

ABSTRACT

Objectives: We aimed to explore the collective wisdom of preprints related to COVID-19 by comparing and synthesizing them with results of peer-reviewed publications.
Methods: PubMed, Google Scholar, medRxiv, bioRxiv, arXiv, and SSRN were searched for papers regarding the estimation of four epidemiological parameters of COVID-19: the basic reproduction number, incubation period, infectious period, and case-fatality-rate. Distributions of parameters and timeliness of preprints and peer-reviewed papers were compared. Four parameters in two groups were synthesized by bootstrapping, and their validities were evaluated by simulated cumulative cases of the susceptible-exposed-infectious-recovered-dead-cumulative (SEIRDC) model.
Results: A total of 106 papers were included for analysis. The distributions of four parameters in two literature groups were close, and the timeliness of preprints was better. Synthesized estimates of the basic reproduction number (3.18, 95% CI 2.85–3.53), incubation period (5.44 days, 95% CI 4.98–5.99), infectious period (6.25 days, 95% CI 5.09–7.51), and case-fatality-rate (4.51%, 95% CI 3.41%–6.29%) were obtained. Simulated cumulative cases of the SEIRDC model matched well with the onset cases in China.
Conclusions: The validity of the COVID-19 parameter estimations of the preprints was on par with that of peer-reviewed publications, and synthesized results of literatures could reduce the uncertainty and be used for epidemic decision-making.

Background

The outbreak of COVID-19 has posed a significant global threat. In response to the emerging infectious disease, the number of research papers has exploded in both formal publications and preprints (Wellcome Trust, 2020). Many journals used a fast track to publish COVID-19 research and made all COVID-19 work freely accessible to facilitate information sharing. In contrast to previous Zika and Ebola outbreaks, scientists were more enthusiastic about posting articles on preprint archives due to the very high transmissibility of COVID-19 (Fraser et al., 2020; Johansson et al., 2018). Many major results were first posted online as preprints before being formally published in journals. However, there were also voices questioning preprints' authority (Maggio et al., 2018; Soderberg et al., 2020; Tang, 2020), believing that preprints pose the risk of dissemination of unconfirmed results and even rumors as they were not peer reviewed. We identified 12 papers and news articles and found that scientists were skeptical of preprints mainly because rigorous peer reviews are absent, and thus the conclusions of preprints may not be reliable (see Supplementary material—support_background_papers). However, there are few studies on the validity of the results reported in preprints. Scientists are overwhelmed by mixed and sometimes contradictory conclusions (Vuong, 2020), and the scientific community and policymakers face new challenges: how valid are the results of the preprints compared to journal papers, and how do we comprehensively integrate results from massive studies efficiently?

Numerous preprints and peer-reviewed articles have estimated the four epidemiological parameters: the basic reproduction number (R_0) (Chowell et al., 2007), incubation period, infectious period, and case-fatality-rate (CFR). It is critical to accurately estimate these four parameters because they indicate the transmission dynamics and severity of COVID-19. Based on various cases data sets and methods, estimates of preprints and peer-

* Corresponding author at: No. 95, Zhongguancun East Road, Haidian District, Beijing, China.
E-mail address: zhidong.cao@ia.ac.cn (Z. Cao).

https://doi.org/10.1016/j.ijid.2020.12.040

reviewed papers have varied over time. We searched for reviews using Google with the terms ("epidemiology" AND ("meta-analysis" OR "reviews") AND "COVID-19") on May 23, 2020. Nine studies were identified (see Supplementary material—support_-background_papers). Majumder and Mandl (2020) discovered 11 studies related to R_0 estimation on Google Scholar and four preprint servers by February 9, 2020. They used a consensus-based approach to yield average R_0 estimates for preprints and journal papers. Park et al. (2020) searched on PubMed and preprint archives on February 21, 2020, and listed all estimates of R_0, incubation period and CFR. This study did not analyze the differences between preprints and peer-reviewed papers, nor did it propose a reasonable method to synthesize various results. Alqahtani et al. (2020) searched MEDLINE and Google Scholar from inception date to March 16, 2020, and did not include preprints in the formal analysis of severity. Other meta-analyses on epidemiology contained only a small number of preprints (Hu et al., 2020; Siordia, 2020; Xie and Chen, 2020). The scientific value of the preprints was largely overlooked by most of the reviews. However, we argue that the collective wisdom contained in the large number of preprints should not be neglected. Further, the potential of synthesizing preprint results with journal paper results should be explored.

Taking epidemiological parameters as objects, we aimed to quantitatively compare the validity of the preprints with peer-reviewed papers and to synthesize the estimations of the two types of literature to mitigate the impact of uncertainty. This study compared and synthesized results for four parameters estimates (R_0, incubation period, infectious period, and CFR) in two literature groups and two pandemic stages. Further, based on the historical data of COVID-19 in China, we evaluated the effectiveness of the synthesized parameters in predicting the epidemic trend (Chowell et al., 2007). Our findings explored the collective wisdom in an epidemic crisis and indicated the academic value of the preprints.

Methods

Search strategy and selection criteria

As the number of daily new cases of COVID-19 in China has been below 100 since the end of March 2020 and the Chinese mainland has entered a new stage of resuming work and production, we only searched and analyzed papers about the epidemic in China. We searched PubMed, Google Scholar, and four popular preprint servers (i.e. medRxiv, bioRxiv, arXiv, and SSRN) for papers published from 23 January to 20 March 2020, using the following terms: "2019-nCoV", "coronavirus", or "COVID-19". Through screening of the titles, we removed papers focused on clinical treatment or papers whose research scopes were countries other than China. Then, the full-text screening was operated to remove comments, news, or papers that did not contain estimates for any of the following epidemiological parameters: (i) R_0, the average number of secondary cases generated by an index case in the totally susceptible population; (ii) incubation period, the average time from infection to illness; (iii) infectious period, the period of time when an infected person is capable of transmitting the virus to others; and (iv) CFR, the percentage of patients who die from a given disease. Finally, we noted that some of the later published papers directly adopted an earlier estimate. Among papers related to the incubation period, 12 cited the same paper published in New England journal of Medicine on January 29 (Li et al., 2020), which estimated the incubation period to be 5.2 days. Thus, we removed the papers that adopted the same estimate of an earlier paper. For the preprints that were published in a certain journal by 20 March, we only kept the journal version.

Data analysis

The following information was manually extracted from each paper: title, publication date (T_P), manuscript submission date (T_S), publication source, estimates for the corresponding four parameters (R_0, incubation period, infectious period, and CFR), and their uncertainty intervals (if available). The publication delay (T_D) of each paper was calculated by the difference between T_P and T_S. As a few peer-reviewed papers did not provide T_S, the latest date for cases data collection in that paper was approximated as the T_S. Based on publication sources, the literature collection was divided into preprint group and peer-reviewed group.

To compare the parameter estimations and timeliness between the preprints and peer-reviewed papers, the distributions of four parameters estimates and T_D of the two groups were separately plotted using the "seaborn" toolbox in Python 3.7.3. Next, we used

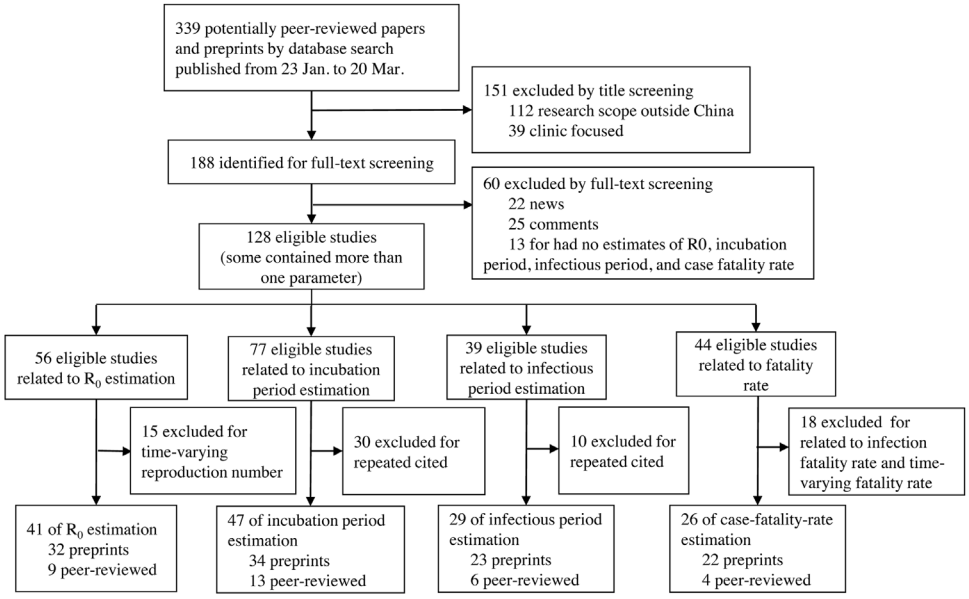

Figure 1. Study selection process.

the bootstrap method to estimate the means and 95% confidence intervals (95% CIs) of the four parameters in the two groups (Efron, 1979). Means of the entire parameter set (not grouped) were obtained as well. The bootstrap method was conducted by the built-in function "bootci" of Matlab R2017a.

With more awareness of the COVID-19 outbreak and more data being accumulated, the parameter estimates changed over time. To integrate results from literatures in the time dimension, the estimations of R_0, incubation period, infectious period, and CFR were ranked in the chronological order based on T_P. Taking February 13 as the demarcation point, the whole period from January 23 to March 20 was divided into two stages. Stage one was from January 23 to February 12, and stage two was from February 13 to March 20. The first stage was the epidemic development period, and the second stage was the epidemic recession period. As the Chinese government isolated and treated 14,840 mild or clinically confirmed cases in Hubei Province on February 12, which further prevented the interpersonal transmission of the virus, the number of daily confirmed cases in China began to decline on February 13. The same bootstrap method was used to calculate the iterative updates of parameter estimates over time and the means of the four parameters in the two stages.

To evaluate the effectiveness of the four synthesized parameters in predicting the epidemic trend, we randomly assembled the means of bootstrap samples of the four parameters of the whole dataset and put them into the SEIRDC model (Supplementary material—SEIRDC model and Supplementary Figure 1) (Chowell et al., 2007). Considering the outbreak of COVID-19 in mainland China, 1.4 billion population in China were set as the susceptible population with an initial infected person. After 1000 Monte Carlo simulations of the SEIRDC model, we obtained the mean curves of simulated cumulative infected cases (C(t)). On the other hand, the cumulative onset infections in China after epidemiological retrospective investigation were obtained from the report of the WHO-China joint mission on Coronavirus Disease 2019. The simulated curve and real curve were aligned and compared based on the date when the cumulative infections reached 100. Accumulated infections in China up to March 20 were also obtained from the official website of the National Health Commission of China.

Results

After selection from the 339 potential literatures (Figure 1), 106 papers were included in the final collection. Fifty-eight (54.7%) papers contained estimates for more than one parameter. There were far more preprints than peer-reviewed papers; preprints accounted for 78.0% (32/41), 72.3% (34/47), 79.3% (23/29), and 84.6% (22/26) of the papers related to R_0, incubation period, infectious period, and CFR, respectively. Papers included and the main characteristics (four parameters, T_P, T_S, and T_D) extracted from them were summarized (see Supplementary material—support_meta_papers; also available at: https://github.com/wyj1996/4-key-epidemiological-parameters-of-COVID19).

The distributions and quantiles of R_0, incubation period, and infectious period in the two groups were close (Figure 2A), but the quartiles of CFR in the preprints group (2.84%, IQRs 1.38%–5.13%, $p < 0.05$, Figure 2A) was much lower than that in the peer-reviewed group (5.6%, IQRs 4.7%–8.1%, $p < 0.05$). Regarding the comparison of timeliness, T_D of the preprints were much lower than the peer-reviewed papers (Figure 2B). The quantiles of T_D of the preprints were 1 (IQRs 0–2, $p < 0.05$), 2 (IQRs 1–3.25, $p < 0.05$), 2 (IQRs 1–2.25, $p < 0.05$), and 2 (IQRs 1–3, $p < 0.05$), corresponding to the four parameters, respectively. Additionally, the review speed of different journals varied greatly: the T_D of peer-reviewed papers were 7 (IQRs 5–16, $p < 0.05$), 7 (IQRs 2–24.5, $p < 0.05$), 5 (IQRs 2–31, $p < 0.05$), and 16 (IQRs 10–16.75, $p < 0.05$), respectively.

Synthesized estimations of the four parameters of each literature group generated by bootstrap method are listed in Table 1. We estimated the mean R_0 in the preprints group to be 3.20 (95% CI 2.92–3.59, $p < 0.05$), which is slightly higher than the mean R_0 of 3.07 (95% CI 2.23–4.17, $p < 0.05$) in the peer-reviewed group. Similarly, the mean incubation period (5.61 days, 95% CI 5.07–6.29, $p < 0.05$) and infectious period (6.54 days, 95% CI 5.24–8.08, $p < 0.05$) in the preprints group were still longer than that in the peer-reviewed group, which were 5.04 (95% CI 4.41–5.72, $p < 0.05$) days and 5.25 (95% CI 3.32–7.25, $p < 0.05$) days. The estimate of CFR in the preprints group was 4.26% (95% CI 3.10%–6.31%, $p < 0.05$), but the mean CFR in the peer-reviewed group (6.10%, 95% CI 4.00%–7.62%, $p < 0.05$) was much higher and had a smaller range of uncertainty.

Regardless of groups, iterative updates of four parameters estimates in time dimension were shown in Figure 3, and the synthesized estimations for the two pandemic stages were shown in Table 1. Aside from the CFR data points in stage one being significantly less than those in stage two, the data of the other three parameters were relatively evenly distributed (Figure 3). The mean R_0 in stage one (3.10, 95% CI 2.64–3.73, $p < 0.05$) was close to R_0 in stage two (3.25, 95% CI 2.89–3.64, $p < 0.05$). The mean incubation period (5.14 days, 95% CI 4.63–5.63, $p < 0.05$) in stage one was slightly shorter than that in stage two (Table 1). However, the mean infectious period declined from 7.19 (95% CI 4.86–9.65, $p < 0.05$) days to 5.76 (95% CI 4.73–7.38, $p < 0.05$) days. Because of

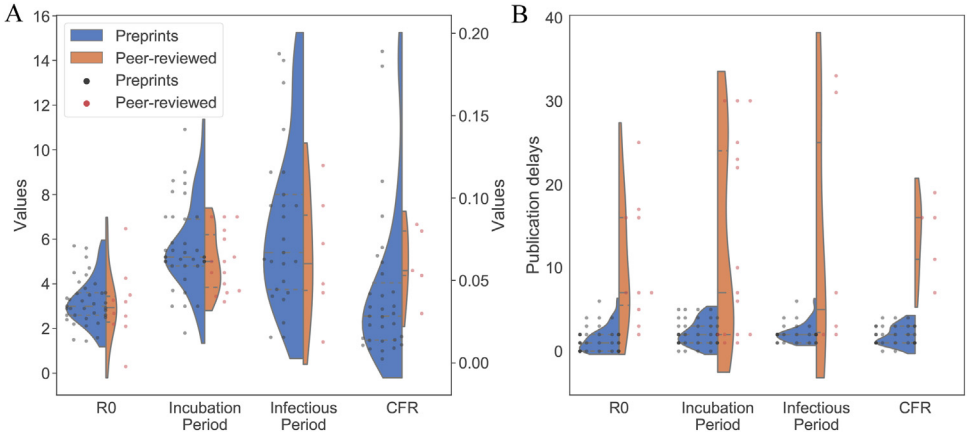

Figure 2. (A) Full distributions of estimates of R_0, incubation period, infectious period, and CFR and (B) their corresponding publication delays in the preprints groups and the peer-reviewed group. CFR corresponds to the right y-axis in Figure 2A; R_0: the basic reproduction number; CFR: case-fatality-rate.

Table 1
Estimates of R_0, incubation period, infectious period, and CFR generated by bootstrap method.

Groups	R_0	Incubation period (days)	Infectious period (days)	CFR (%)
Preprints	3.20 (95% CI 2.92–3.59)	5.61 (95% CI 5.07–6.29)	6.54 (95% CI 5.24–8.08)	4.26 (95% CI 3.10–6.31)
Peer-reviewed	3.07 (95% CI 2.23–4.17)	5.04 (95% CI 4.41–5.72)	5.25 (95% CI 3.32–7.25)	6.10 (95% CI 4.00–7.62)
Stage one	3.10 (95% CI 2.64–3.73)	5.14 (95% CI 4.63–5.63)	7.19 (95% CI 4.86–9.65)	5.93 (95% CI 3.27–11.42)
Stage two	3.25 (95% CI 2.89–3.64)	5.63 (95% CI 5.00–6.50)	5.76 (95% CI 4.73–7.38)	4.09 (95% CI 3.09–6.24)
Overall	3.18 (95% CI 2.85–3.53)	5.44 (95% CI 4.98–5.99)	6.25 (95% CI 5.09–7.51)	4.51 (95% CI 3.41–6.29)

All 95% CIs with $p < 0.05$; R_0: the basic reproduction number; CFR: case-fatality-rate.

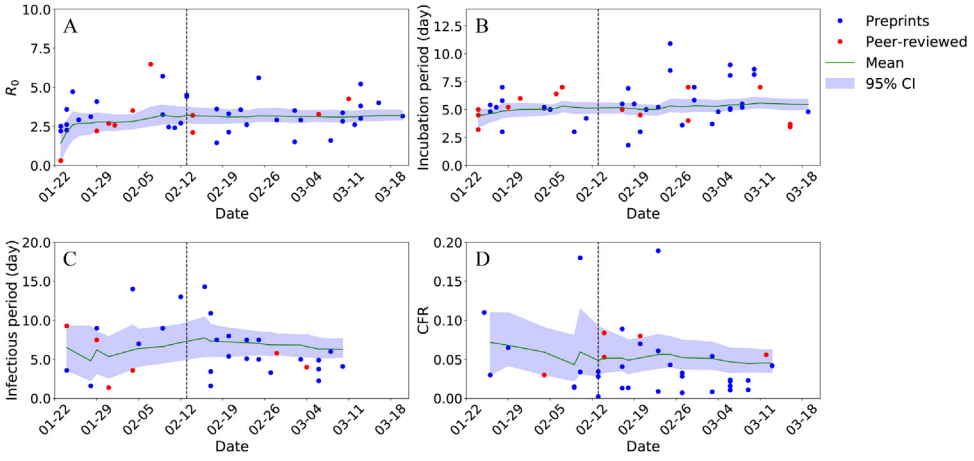

Figure 3. Distributions of (A) R_0, (B) incubation period, (C) infectious period, and (D) CFR ordered by publication time. R_0: the basic reproduction number; CFR: case-fatality-rate.

limited data, mean CFR in stage one was 5.93% with a larger 95% CI between 3.27% and 11.42%, and CFR in stage two was 4.09% (95% CI 3.09%–6.24%, $p < 0.05$). The overall estimations of these four parameters were also given in Table 1. The gaps between the overall estimates and the estimates of the preprint group were smaller because preprints accounted for the majority.

To evaluate the effectiveness of the four synthesized parameters in predicting the epidemic trend, C(t) of mainland China from December 26, 2019, to February 19, 2020, obtained by 1000 Monte Carlo simulation of the SEIRDC model, was shown in Figure 4 (please refer to the method section for the specific simulation method). December 26 was the date when both the cumulative onset infections (the blue curve) and the simulated infections (the red curve) reached 100. By December 31, 2019, the cumulative onset cases almost exactly matched the simulated mean (Figure 4). From January 1 to February 7, the onset infections were slightly above the simulated mean but still within the range of simulations. During this period, China adopted many prevention and control measures, including traffic restriction and makeshift hospitals in Wuhan. After February 7, the epidemic in China was effectively contained by quarantine and treatment measures, and as of February 19, the cumulative onset infections of China had stabilized at about 75,100 (Figure 4). The officially reported cumulative confirmed cases (the green curve) significantly lagged behind the onset infections in the early period, and the two curves were not roughly equal until February 13.

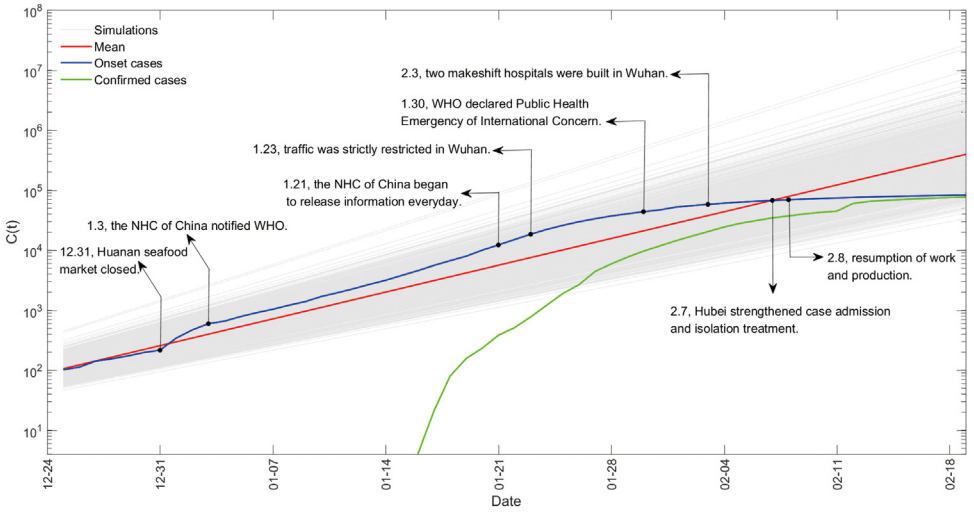

Figure 4. The simulation results of the cumulative number of infections of China.

Discussion

In this study, based on literatures of COVID-19 in mainland China by March 20, 2020, we compared distributions of parameter estimations between the preprints group and the peer-reviewed group and synthesized the estimations according to group or publication date. Results showed that, except for CFR, the distributions and synthesized estimates of R_0, incubation period, and infectious period were similar between the two groups. Estimates of R_0 and incubation period remained stable in two pandemic stages, but the estimates of infectious period and CFR in stage two declined significantly. Further, the SEIRDC simulations of COVID-19 outbreak in China evaluated the applicability and validity of the comprehensive parameters space. The actual cumulative onset infections and the simulation results matched well.

The validity of the overall parameter distribution of the preprints was quantitatively analyzed in this study. The validity of some preprints is controversial. Further, there were far more preprints than peer-reviewed papers because the former are simply reviewed by volunteers on the platform. For the same reason, preprints are also disseminated in a timely manner compared to peer-reviewed publications. Some scientists argued that the conclusions of the preprints may be misleading and should not be widely adopted (Elmore, 2018), while others thought that because the preprint was available to the public, authors would pay more attention to their personal reputations, and the quality of the preprint would thus not be uncontrolled (Berg et al., 2016; Callaway and Powell, 2016). The comparison results of this study showed that the estimated distributions of R_0, incubation period, and infectious period in the preprints group were similar to that of the peer-reviewed group (Figure 2A). Further, the distributions of preprints were more concentrated, and thus, the ranges of 95% CIs of preprints were smaller (Table 1). These suggested that, in the outbreak of COVID-19, although the results of some preprints may be biased, the validity of synthesized parameter estimates of the preprints were at the same level as the peer-reviewed papers, and the synthesized estimations of preprints were even more robust. Therefore, it is not wise to neglect the collective wisdom contained in the large number of preprints.

The iterative estimations of parameters from the time dimension can reflect the trend of the epidemic. Compared with stage one, the corresponding infection period in stage two was shorter and CFR was reduced (Table 1). The possible reason for the changes was that many patients diagnosed in the late period were included for the parameter estimation in studies that were posted in stage two. Due to effective control measures in China, the speed of testing was improved and the cure rate also increased, and thus, the infectious period and CFR decreased. This indicated that integrating epidemiological parameters in the time dimension could also reflect changes in the epidemic situation. Additionally, in this study, the preprints accounted for 72%–85% of the total literature, and so these trends were largely reflected by preprints. Further, as the publication delay of preprint was shorter (Figure 2B), the preprints allowed us to obtain the latest information to assist emergency decision-making in a timely manner.

The results of many papers inevitably have uncertainties. It is unreasonable to make decisions based on the conclusion of a single paper. To mitigate the effects of uncertainty, it is more robust to synthesize the results from multiple papers. In this study, through random sampling and 1000 Monte Carlo simulations, the four parameter estimates of published papers were utilized in a comprehensive way. In Figure 4, in the absence of human intervention, before the closure of the Huanan seafood market, the cumulative onset cases and the simulated mean were almost in line. After January 23, due to strong isolation and control measures, the rate of cumulative onset cases slowed. The cumulative onset cases were always within the uncertain range of the simulations. However, at that time, the official confirmed cases (Figure 4) that scholars could obtain for analysis were far behind the real onset cases, indicating that even if scientific data were lacking and delayed in the early period of COVID-19, the whole parameter space still grasped well the pattern of epidemic spread. This not only reminds us to comprehensively refer to the results of all published papers but also reflects the practical value of the preprints because they are the majority.

This study has some limitations. First, the data officially reported by China did not fully represent all the infections and deaths because many patients died without diagnoses in the early period. Further, with the huge burden on the medical system in Hubei Province, it was impossible to detect and report all cases without omission. We can only prove the validity of our parameter estimations to a certain extent, but we cannot deny the reference value of collective wisdom from literatures. Second, in our study, the validity of the preprints was only compared and evaluated based on the overall distribution, thus demonstrating their academic value on the task of estimating the four epidemiological parameters of COVID-19. However, this does not mean that the estimate of every preprint in the fusion set is accurate; at least in some preprints, the parameter estimations may be biased for various reasons. Further, the method of our study cannot be arbitrarily extended to other fields because this article only performs bootstrap synthesis on four important epidemic parameters; whether synthesizing preprints and peer-reviewed papers is useful for parameter estimation in other fields needs further exploration. Additionally, without rigorous rounds of review, there may be more minor errors in preprints, and thus we should be cautious about their results. When scientists publish their own preprints, they should not lower the standards of preprints because of the simple review process. Further, scientists should further standardize the publication process of preprints and guide the media to scientifically report preprints.

In conclusion, our quantitative analysis shows that the overall validity of the preprints in parameter estimation is not less than that of the peer-reviewed papers. The latest information on the epidemic can be obtained more sensitively through preprints. Furthermore, the simulation of the COVID-19 in China proved that the synthesis of whole parameter space is an effective way to reduce uncertainty and to grasp the pattern of transmission. In response to future public health crises, scientists should be more proactive in promoting the development of preprint platforms and quality monitoring (Johansson et al., 2018), while more automated literature analysis and integration methods should be developed to make collective intelligence more applicable to decision-making.

Funding sources

This work was supported by the National Natural Science Foundation of China [grant numbers 72042018, 91546112, and 71621002] and the Beijing Municipal Natural Science Foundation [grant number L192012].

Conflict of interest statement

All authors declared no conflict of interest.

Ethical approval

The survey was discussed with the Institute of Automation Chinese Academy of Sciences, who reviewed the content and advised that it was exempt from the ethics committee review.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.ijid.2020.12.040.

References

Alqahtani JS, Oyelade T, Aldhahir AM, Alghamdi SM, Almehmadi M, Alqahtani AS, et al. Prevalence, severity and mortality associated with COPD and smoking in patients with COVID-19: a rapid systematic review and meta-analysis. PLoS One 2020;15(5):e0233147, doi:http://dx.doi.org/10.1371/journal.pone.0233147.

Berg JM, Bhalla N, Bourne PE, Chalfie M, Drubin DG, Fraser JS, et al. Preprints for the life sciences. Science 2016;352(6288):899–901, doi:http://dx.doi.org/10.1126/science.aaf9133.

Callaway E, Powell K. "Hug a preprint, biologists!" ASAPbio meeting discusses the ins and outs of posting work online before peer review. Nature 2016;530:265.

Chowell G, Nishiura H, Bettencourt LM. Comparative estimation of the reproduction number for pandemic influenza from daily case notification data. J R Soc Interface 2007;4(12):155–66, doi:http://dx.doi.org/10.1098/rsif.2006.0161.

Elmore SA. Preprints: what role do these have in communicating scientific results?. Toxicol Pathol 2018;46(4):364–5, doi:http://dx.doi.org/10.1177/0192623318767322.

Efron B. Bootstrap methods: another look at the jackknife. Ann Stat 1979;7(1):1–26.

Fraser N, Brierley L, Dey G, Polka JK, Pálfy M, Coates JA. Preprinting a pandemic: the role of preprints in the COVID-19 pandemic. BioRxiv 2020;, doi:http://dx.doi.org/10.1101/2020.05.22.111294 2020.05.22.111294.

Hu Y, Sun J, Dai Z, Deng H, Li X, Huang Q, et al. Prevalence and severity of corona virus disease 2019 (COVID-19): a systematic review and meta-analysis. J Clin Virol 2020;127:104371, doi:http://dx.doi.org/10.1016/j.jcv.2020.104371.

Johansson MA, Reich NG, Meyers LA, Lipsitch M. Preprints: an underutilized mechanism to accelerate outbreak science. PLoS Med 2018;15(4):e1002549, doi:http://dx.doi.org/10.1371/journal.pmed.1002549.

Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus?infected pneumonia. N Engl J Med 2020;382(13):1199–207, doi:http://dx.doi.org/10.1056/NEJMoa2001316.

Majumder MS, Mandl KD. Early in the epidemic: impact of preprints on global discourse about COVID-19 transmissibility. Lancet Glob Health 2020;8(5):e627–30, doi:http://dx.doi.org/10.1016/S2214-109X(20)30113-3.

Maggio LA, Artino [241_TD$DIFF]Jr. AR, Driessen EW. Preprints: facilitating early discovery, access, and feedback. Perspect Med Educ 2018;7(5):287–9, doi:http://dx.doi.org/10.1007/s40037-018-0451-8.

Park M, Cook AR, Lim JT, Sun Y, Dickens BL. A systematic review of COVID-19 epidemiology based on current evidence. J Clin Med 2020;9(4):967, doi:http://dx.doi.org/10.3390/jcm9040967.

Soderberg CK, Errington T, Nosek BA. Credibility of preprints: an interdisciplinary survey of researchers. 2020. . . [Accessed 18 March 2020] https://osf.io/preprints/metaarxiv/kabux/.

Siordia [241_TD$DIFF]Jr. JA. Epidemiology and clinical features of COVID-19: a review of current literature. J Clin Virol 2020;127:104357, doi:http://dx.doi.org/10.1016/j.jcv.2020.104357.

Tang G. Distortion and misreading: rumors from preprints spread in China. , doi:http://dx.doi.org/10.2139/ssrn.3584183 SSRN. 3584183.

Vuong Q-H. The rise of preprints and their value in social sciences and humanities. Sci Ed 2020;7(1):70–2, doi:http://dx.doi.org/10.6087/kcse.193.

Wellcome Trust. Sharing research data and findings relevant to the novel coronavirus (COVID-19) outbreak. 2020. . . [Accessed 18 March 2020] https://wellcome.ac.uk/coronavirus-covid-19/open-data.

Xie M, Chen Q. Insight into 2019 novel coronavirus — an updated interim review and lessons from SARS-CoV and MERS-CoV. Int J Infect Dis 2020;94:119–24, doi:http://dx.doi.org/10.1016/j.ijid.2020.03.071.