Early epidemiological assessment of the transmission potential and virulence of coronavirus disease 2019 (COVID-19) in Wuhan City: China, January-February, 2020

Authors: Kenji Mizumoto[1,2,3] §, Katsushi Kagaya[2,4], Gerardo Chowell [3]

Affiliations:

[1] Graduate School of Advanced Integrated Studies in Human Survivability, Kyoto University Yoshida–Nakaadachi–cho, Sakyo–ku, Kyoto, Japan
[2] Hakubi Center for Advanced Research, Kyoto University, Yoshidahonmachi, Sakyo–ku,Kyoto, Japan;
[3] Department of Population Health Sciences, School of Public Health, Georgia State University, Atlanta, Georgia, USA
[4] Seto Marine Biological Laboratory, Field Science, Education and Reseach Center, Kyoto University, Shirahama–cho, Nishimuro–gun, Wakayama 649–2211 Japan

§Corresponding author
Email addresses:
KM: mizumotokenji@gmail.com, KK: kagaya.katsushi.8e@kyoto-u.ac.jp, GC: gchowell@gsu.edu

Article type:

Original Research

Word count:

Abstract: 350 (Max 350)
Main: 2768

Abstract

Background:

Since the first cluster of cases was identified in Wuhan City, China, in December, 2019, coronavirus disease 2019 (COVID-19) has rapidly spread across China, causing multiple introductions in 109 countries/territories/areas as of March 10th. Despite the scarcity of publicly available data, scientists around the world have made strides in estimating the magnitude of the epidemic, the basic reproduction number, and transmission patterns. Recently more evidence suggests that a substantial fraction of the infected individuals with the novel coronavirus show little if any symptoms, which suggest the need to reassess the transmission potential of this emerging disease. In this study, we derive estimates of the transmissibility and virulence of COVID-19 in Wuhan City, China, by reconstructing the underlying transmission dynamics using multiple data sources.

Methods:

We employ statistical methods and publicly available epidemiological datasets to jointly derive estimates of transmissibility and severity associated with the novel coronavirus. For this purpose, the daily series of laboratory–confirmed COVID-19 cases and deaths in Wuhan City and epidemiological data of Japanese evacuees from Wuhan City on

board government–chartered flights were integrated into our analysis.

Results:

Our posterior estimates of basic reproduction number (*R*) in Wuhan City, China in 2019–2020 reached values as high as 5.20 (95%CrI: 5.04–5.47) and the enhanced public health intervention after January 23rd in 2020 was associated with a declined *R* at 0.58 (95%CrI: 0.51–0.64), with the total number of infections (i.e. cumulative infections) estimated at 1905526 (95%CrI: 1350283– 2655936) in Wuhan City, raising the proportion of infected individuals to 19.1% (95%CrI: 13.5–26.6%). We also found that most recent crude infection fatality ratio (IFR) and time–delay adjusted IFR is estimated to be 0.04% (95% CrI: 0.03%–0.06%) and 0.12% (95%CrI: 0.08–0.17%), which is several orders of magnitude smaller than the crude CFR estimated at 4.19%

Conclusions:

We have estimated key epidemiological parameters of the transmissibility and virulence of COVID-19 in Wuhan, China during January-February, 2020 using an ecological modelling approach. The power of our approach lies in the ability to infer epidemiological parameters with quantified uncertainty from partial observations collected by surveillance systems.

Keywords: epidemic; transmissibility; mathematical model; COVID-19; China

Background

The novel coronavirus (COVID-19) emerging from China is a deadly respiratory pathogen that belongs to the same family as the coronavirus responsible for the 2002-2003 Severe Acute Respiratory Syndrome (SARS) outbreaks [1]. Since the first cluster of cases was identified in Wuhan City, China, in December, 2019, COVID-19 has rapidly spread across China as well as caused multiple introductions in 109 countries/territories/areas as of March 10th, 2020 [2]. Nevertheless, China has been hit hard by this emerging infectious disease, especially the city of Wuhan in Hubei Province, where the first cluster of severe pneumonia caused by the novel virus was identified. Meanwhile, the cumulative number of laboratory and clinically confirmed cases and deaths in mainland China has reached 80778 and 3158, respectively, as of March 10th, 2020 [3].

Because the morbidity and mortality burden associated with the novel coronavirus has disproportionally affected the city of Wuhan, the center of the epidemic in China, the central government of the People's Republic of China imposed a lockdown and social distancing measures in this city and surrounding areas starting on January 23rd 2020. Indeed, out of the 80778 COVID-19 cases reported in China, 49978 cases (61.9%) are from Wuhan City. In terms of the death count, a total of 2423 deaths (76.7%) have been recorded in Wuhan city out of the 3158 deaths reported throughout China. To guide the effectiveness of interventions, it is crucial to gauge the uncertainty relating to key epidemiological parameters characterizing the transmissibility and the severity of the disease. Despite the scarcity of publicly available data, scientists around the world have made strides in estimating the magnitude of the epidemic, the basic reproduction number, and transmission patterns [4-5]. Moreover, accumulating evidence

suggests that a substantial fraction of the infected individuals with the novel coronavirus show little if any symptoms, which suggest the need to reassess the transmission potential of this emerging disease [6]. For this purpose, in this study we employ statistical methods and publicly available epidemiological datasets to jointly derive estimates of transmissibility and severity associated with the novel coronavirus.

Methods

Epidemiological data

We linked our model to two different datasets. First, the daily series of laboratory–confirmed nCov cases and deaths in Wuhan City were extracted according to date of symptoms onset or reporting date from several sources [3, 7-8]. Our analysis relies on epidemiological data reported prior to February 11th, 2020 because of the change in case definition that was announced on February 12th, 2020 [9]. As of February 11th, 2020, a total of 19559 confirmed cases including 820 deaths were reported in Wuhan City. Second, epidemiological data of Japanese evacuees from Wuhan City on board government–chartered flights were obtained from the Japanese government. After arriving in Japan, all of the Japanese evacuees were kept in isolation for about 14 days and examined for infection using polymerase chain reaction (PCR) tests [7]. As of February 11th, a total of four flights with the Japanese evacuees left Wuhan City. We collected information on the timing of the evacuee fights that left Wuhan City as well as the number of passengers that tested positive for COVID-19 in order to calibrate our model (Table S1).

Statistical analysis

Using the following integral equation model, we estimate the reproduction number of COVID-19. Here, infected and reported cases are denoted by *i* and *c*, respectively.

We connected the daily incidence series with a discrete–time integral equation to describe the epidemic dynamics. Let g_s denote the probability mass function of the serial interval, e.g., the time from illness onset in a primary case to illness onset in the secondary case, of length *s* days, which is given by

$$g_s = G(s) - G(s-1)) \ ,$$

For $s > 0$ where $G(.)$ represents the cumulative distribution function of the gamma distribution. Mathematically, we describe the expected number of new cases with day *t*, $E[c(t)]$ as follows,

$$\mathrm{E}[c(t)] = \sum_{s=1}^{\infty} E[c(t-s)]R,$$

where $E[c(t)]$ represents the expected number of new cases with onset day *t*, where *R* represents the average number of secondary cases per case.

Subsequently, we also employed the time–dependent variation in *R* to take into account the impact of enhanced interventions on the transmission potential. This time dependence was modelled by introducing a parameter δ_1, which is given by

$$\delta_1 = \begin{cases} 1 & otherwise \\ \beta_1 & \text{if } t = period_1 \\ \beta_2 & \text{if } t = period_1 \end{cases}$$

where $period_1$ and period2 represent the corresponding period from January 23rd to February 2nd 2020 and from February 3rd to February 11th, 2020, respectively. January 23rd 2020 is the date when the central government of the People's Republic of

China imposed a lockdown in Wuhan and other cities in Hubei in an effort to quarantine the epicentre of the coronavirus (COVID-19) to mitigate transmission. Furthermore, we evenly divide the interval into two periods to incorporate the time-dependent effects on R using the parameters β_1 and β2 which scale the extent of the intervention, taking values smaller than 1[10].

To account for the probability of occurrence, θ [11] ,we assume that the number of observed cases on day t, $h(t)$, occurred according to a Bernoulli sampling process, with the expected values $E(c_t;H_{t-1})$, where $E(c_t; H_{t-1})$ denotes the conditional expected incidence on day t, given the history of observed data from day 1 to day (t–1), denoted by H_{t-1}. Thus, the number of expected newly observed cases is written as follows:

$$\mathrm{E}[h(t); H_{t-1}] = \begin{cases} (1-\theta) + \theta \mathrm{E}[c(t); H_{t-1}], & \text{if } h = 0, \\ \theta \mathrm{E}[c; H_{t-1}], & \text{otherwise,} \end{cases}$$

Further, we model the time–dependent variation in the reporting probability. This time dependence was modelled by introducing a parameter δ_2, which is given by

$$\delta_2 = \begin{cases} \alpha_1, & \text{if } t = period_3, \\ \alpha_2, & \text{if } t = period_4, \\ 1, & \text{otherwise,} \end{cases}$$

where $period_3$ and $period_4$ represent the corresponding periods from the start of our study period to Jan 16 and from Jan 17 to Jan 22, respectively, while α_1 *and* α_2 scale the extent of the reporting probability (where α_1 and α_2 is expected to be smaller than 1). We evenly divide the duration before the lockdown was put in place into two to incorporate the time dependency of the reporting probability. The number of expected newly observed cases should be updated as

$$\mathrm{E}[h(t); H_{t-1}] = \begin{cases} (1-\theta) + q\delta\theta \mathrm{E}[c(t); H_{t-1}], & \text{if } h_a = 0, \\ q\delta\theta \mathrm{E}[c(t); H_{t-1}], & \text{otherwise,} \end{cases}$$

We assume the incidence, h(t) is the result of the Binomial sampling process with the

147 expectation E[h]. The likelihood function for the time series of observed cases that we
148 employ to estimate the effective reproduction number and other relevant parameters is
149 given by:

$$L_1(U;c) = \prod_{t=1}^{T} \binom{E(h(t);H(t-1))}{c(t)} q^{c(t)} (1-q)^{E(h(t);H(t-1)-c(t))} ,$$

150 where U indicates parameter sets that are estimated from this likelihood.

151 Subsequently, the conditional probability of non–infection given residents in
152 Wuhan City at the time point of t_i, p_{ti}, was assumed to follow a binomial distribution,
153 and the likelihood function is given by:

$$L_2\left(p_{t_i}; M_{t_i}, m_{t_i}\right) = \binom{M_{t_i}}{m_{t_i}} {p_{t_i}}^{m_{t_i}} \left(1-p_{t_i}\right)^{M_{t_i}-m_{t_i}} ,$$

154 Where M_{ti} and m_{ti} is the number of government charted flight passengers and
155 non–infected passengers at the date of t_i, respecitively, and p_{ti} is the proportion of the
156 estimated non–infected population in Wuhan at the date of t_i, calculated from the h(t)
157 and catchment population in Wuhan City [3,13].

158 Serial interval estimates of COVID-19 were derived from previous studies of
159 nCov, indicating that it follows a gamma distribution with the mean and SD at 7.5 and
160 3.4 days, respectively, based on ref. [14]. The maximum value of the serial interval was
161 fixed at 28 days as the cumulative probability distribution of the gamma distribution up
162 to 28 days reaches 0.999.

163

164 **Infection fatality ratio**

165 Crude CFR and crude IFR is defined as the number of cumulative deaths
166 divided by the number of cumulative cases or infections at a specific point in time

without adjusting the time delay from illness onset or hospitalization to death. Next, we employed an integral equation model in order to estimate the real–time IFR. First, we estimated the real–time CFR as described elsewhere [15-17]. For the estimation, we employ the delay from hospitalization to death, f_s, which is assumed to be given by $f_s = F(s) - F(s-1)$ for $s>0$ where $H(s)$ follows a gamma distribution with mean 10.1 days and SD 5.4 days, obtained from the available observed data [18].

$$L_3(\pi; c_t, \theta) = \prod_{t_i} \left(\sum_{t=1}^{t_i} c_t \right) , \left(\pi \frac{\sum_{t=2}^{t_i} \sum_{s=1}^{t-1} c_{t-s} f_s}{\sum_{t=1}^{t_i} c_t} \right)^{D_{t_i}} \left(1 - \pi \frac{\sum_{t=2}^{t_i} \sum_{s=1}^{t-1} c_{t-s} f_s}{\sum_{t=1}^{t_i} c_t} \right)^{\sum_{t=1}^{t_i} c_t - D_{t_i}}$$

where c_t represents the number of new cases with reported day t, and D_{ti} is the number of new deaths with reported day t_i [16-18]. We assume that the cumulative number of observed deaths, D_t is the result of the binomial sampling process with probability π. Subsequently, crude IFR and time–delay adjusted IFR are calculated using the estimated π and h_t.

The total likelihood is calculated as $L=L_1L_2L_3$ and model parameters were estimated using a Monte Carlo Markov Chain (MCMC) method in a Bayesian framework. Posterior distributions of the model parameters were estimated based on sampling from the three Markov chains. For each chain, we drew 100,000 samples from the posterior distribution after a burn–in of 20,000 iterations. Convergence of MCMC chains were evaluated using the potential scale reduction statistic [19-20]. Estimates and 95% credibility intervals for these estimates are based on the posterior probability

185 distribution of each parameter and based on the samples drawn from the posterior
186 distributions. All statistical analyses were conducted in R version 3.5.2 (R Foundation
187 for Statistical Computing, Vienna, Austria) using the 'rstan' package.
188

189 **Results**

190 The daily series of COVID-19 laboratory–confirmed incidence and cumulative
191 incidence in Wuhan in 2019–2020 are displayed in Figure 1. Overall, our dynamical
192 models yield a good fit to the temporal dynamics (i.e. incidence, cumulative incidence)
193 including an early exponential growth pattern in Wuhan. In incidence data, a few
194 fluctuations are seen, probably indicating that the surveillance system likely missed
195 many cases during the early transmission phase (Figure 1).

196 Our posterior estimates of basic reproduction number (R) in Wuhan City, China
197 in 2019–2020 was estimated to be as high as 5.20 (95%CrI: 5.04–5.47). The
198 time–dependent scaling factors quantifying the extent of enhanced public health
199 intervention on R is 0.99 (95%CrI: 0.94–1.00), declining R to 5.12 (95%CrI: 4.98–5.26)
200 from January 23rd to February 1st and 0.11 (95%CrI: 0.10–0.13), declining R to 0.58
201 (95%CrI: 0.51–0.64) from February 2nd to February 11th, 2020. Other parameter
202 estimates for the probability of occurrence and reporting rate are 0.97 (95% CrI:
203 0.84–1.00) and 0.010 (95% CrI: 0.007–0.014), respectively. Moreover, the
204 time–dependent scaling factor quantifying the extent of reporting rate, α, is estimated to
205 be 0.08 (95% CrI: 0.03–0.19) before January 16th and to be 0.99 (95% CrI: 0.96–1.00)
206 from January 17th to January 22nd.

207 The total number of estimated laboratory–confirmed cases (i.e. cumulative
208 cases) is 18913 (95% CrI: 16444–19705) while the actual numbers of reported

209 laboratory–confirmed cases during our study period is 19559 as of February 11th, 2020.
210 Moreover, we inferred the total number of COVID-19 infections (Figure S1). Our
211 results indicate that the total number of infections (i.e. cumulative infections) is
212 1905526 (95%CrI: 1350283– 2655936).

213 The Observed and posterior estimates of the cumulative number of deaths from
214 COVID-19 in Wuhan are displayed in Figure 2, and model–based posterior estimates of
215 the cumulative number of deaths is 820 (95%CrI: 744–900), while actual number of
216 reported deaths is 820. The estimated temporal variation in the death risk caused by
217 COVID-19 in Wuhan, China, 2019–2020 is shown in Figure 3 and Figure S2. Observed
218 and posterior estimated of crude CFR in Wuhan City is presented in Figure 2A, while
219 observed and posterior estimates of time–delay adjusted CFR is shown in Figure 2B.
220 Furthermore, Figure 3A and 3B illustrates time–delay no–adjusted IFR and time–delay
221 adjusted IFR, respectively.

222 The latest estimate of the crude CFR and time–delay adjusted CFR in Wuhan
223 appeared to be 4.2% (95% CrI: 3.9–4.9%) and 12.2% (95% CrI: 11.4–13.0%),
224 respectively, whereas the latest model–based posterior estimates of time–delay not
225 adjusted IFR and adjusted IFR, presented in Figure 3 C and D, are 0.04%(95% CrI:
226 0.03%–0.06%) and 0.12% (95%CrI: 0.08–0.17%), respectively, while the observed
227 crude CFR is calculated to be 4.19% (Table 1).
228

229 Discussion

230 In this study we derived estimates of the transmissibility and virulence of
231 COVID-19 in Wuhan City, China, by reconstructing the underlying transmission

dynamics using multiple data sources. Applying dynamic modeling, the reproduction number and death risks as well as probabilities of occurrence and reporting rate were estimated.

Our posterior estimates of basic reproduction number (*R*) in Wuhan City, China in 2019–2020 is calculated to be as high as 5.20 (95%CrI: 5.04–5.47). The time–dependent scaling factor quantifying the extent of enhanced public health intervention on *R* is 0.99 (95%CrI: 0.94–1.00), declining R to 5.12 (95%CrI: 4.98–5.26) from January 23rd to February 1st and 0.11 (95%CrI: 0.10–0.13), declining R to 0.58 (95%CrI: 0.51–0.64) for February 2nd to February 11th, 2020. These *R* estimates capturing the underlying transmission dynamics modify the impact of COVID-19, with the total number of infections (i.e. cumulative infections) estimated at 1905526 (95%CrI: 1350283– 2655936) in Wuhan City, raising the proportion of infected individuals to 19.1% (95%CrI: 13.5–26.6%) with a catchment population in Wuhan City of 10 million people. Our estimates of mean reproduction number reached values as high as 5.20, an estimate that is slightly higher than previous mean estimates in the range 2.2-3.8 derived by fitting epidemic models to the initial growth phase of the observed case incidence [21-23]. By comparison, the *R* estimate for the Diamond Princess cruise ship in Japan reached values as high as ~11 [24]. Further, these estimates are higher than recent mean *R* estimates derived from the growth rates of the COVID-19 outbreaks in Singapore (R~1.1) [25] and Korea (R~1.5) [26].

The sustained high *R* values in Wuhan City even after the lockdown and mobility restrictions suggests that transmission is occurring inside the household or amplified in healthcare settings [18], which is a landmark of past SARS and MERS outbreaks

[27-28]. Considering the potent transmissibility of COVID-19 in confined settings, as illustrated by COVID-19 outbreaks aboard cruise ships, including the Diamond Princess cruise ship, where the total number of secondary or tertiary infections reached 705 among more than 3,700 passengers as of February 28th, 2020 and also by the COVID-19 outbreak tied to the Shincheonji religious sect in South Korea where church members appear to have infected from seven to 10 people. [29-31], it is crucial to prevent transmission in confined settings including hospital-based transmission by strengthening infection control measures as well as transmission stemming from large social gatherings.

Our most recent estimates of the crude CFR and time−delay adjusted CFR are at 4.2% (95% CrI: 3.9–4.9%) and 12.23% (95% CrI: 11.4–13.0%), respectively. In contrast, our most recent crude IFR and time–delay adjusted IFR is estimated to be 0.04% (95% CrI: 0.03%–0.06%) and 0.12% (95%CrI: 0.08–0.17%), which is several orders of magnitude smaller than the crude CFR at 4.19%. These findings indicate that the death risk in Wuhan is estimated to be much higher than those in other areas, which is likely explained by hospital-based transmission [32]. Indeed, past nosocomial outbreaks have been reported to elevate the CFR associated with MERS and SARS outbreaks, where inpatients affected by underlying disease or seniors infected in the hospital setting have raised the CFR to values as high as 20% for a MERS outbreak [33-34].

Public health authorities are interested in quantifying *R* and CFR to measure the transmission potential and virulence of an infectious disease, especially when emerging/re–emerging epidemics occur in order to decide the intensity of the public health response. In the context of a substantial fraction of unobserved infections due to

COVID-19, *R* estimates derived from the trajectory of infections and the IFR are more realistic indicators compared to estimates derived from observed cases alone [18, 35-36].

Our analysis also revealed a high probability of occurrence and quite low reporting probabilities in Wuhan City. High probability of occurrence in the above equation suggests that zero observed cases at some point is not due to the absence of those infected, but more likely due to a low reporting rate. A very low reporting probability suggests that it is difficult to diagnose COVID-19 cases or a breakdown in medical care delivery. Moreover, we also identified a remarkable change in the reporting rate, estimated to be 12–fold lower in the 1st period (–Jan 16, 2020) and about the same during the 2nd period (January 17th – 22nd , 2020), relative to the that estimated after January 23rd 2020.

Our results are not free from the limitations. First, our methodology aims to capture the underlying transmission dynamics using multiple data sources. By implementing mass screening in certain populations is a useful approach to ascertain the real proportion of those infected and a way of adding credibility to the estimated values. Second, it is worth noting that the data of Japanese evacuee employed in our analysis is not a random sample from the Wuhan catchment population. Indeed, it also plausible that their risk of infection in this sample is not as high as that of local residents in Wuhan, underestimating the fatality risk.

Conclusion

In summary, we have estimated key epidemiological parameters of the

transmissibility and virulence of COVID-19 in Wuhan, China, January-February, 2020 using an ecological modelling approach and several epidemiological datasets. The power of our approach lies in the ability to infer epidemiological parameters with quantified uncertainty from partial observations collected by surveillance systems.

List of abbreviations

CFR: Case fatality ratio, IFR: Infection Fatality ratio, SARS: Severe Acute Respiratory Syndrome, MERS: Middle East Respiratory Syndrome

Additional files

Additional file 1:

Appendix. Table S1. Information related to Japanese evacuees from Wuhan City on board government–chartered flights

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The present study relies on published data and access information to essential components of the data are available from the corresponding author.

Competing interests

The authors declare that they have no competing interests.

Funding

KM acknowledges support from the Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Number 18K17368 and from the Leading Initiative for Excellent Young Researchers from the Ministry of Education, Culture, Sport, Science & Technology of Japan. KK acknowledges support from the JSPS KAKENHI Grant Number 18K19336 and 19H05330. GC acknowledges support from NSF grant 1414374 as part of the joint NSF–NIH–USDA Ecology and Evolution of Infectious Diseases program.

Authors' contributions

KM and GC conceived the early study idea. KM and KK built the model. KM implemented statistical analysis and wrote the first full draft. GC advised on and helped shape the research. All authors contributed to the interpretation of the results and edited and commented on several earlier versions of the manuscript.

Acknowledgements

Not applicable.

REFERENCES

1. Jon Cohen. **Mining coronavirus genomes for clues to the outbreak's origins. Science**. Jan 31, 2020. https://www.sciencemag.org/news/2020/01/mining-coronavirus-genomes-clues-outbreak-s-origins.Accessed, 3 Feb, 2020
2. World Health Organization (WHO). Novel Coronavirus (2019-nCoV) situation reports. Available from: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports Accessed, [cited 2020 Mar 10]

3. The State Council, The People's Republic of China. [cited 2020 Mar 10]. http://www.gov.cn/

4. Nishiura H, Jung SM, Linton NM, Kinoshita R, Yang Y, Hayashi K, et al. **The Extent of Transmission of Novel Coronavirus in Wuhan, China, 2020**. *J Clin Med*. 2020; 9(2); 330

5. Wu JT, Leung K, Leung GM. **Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study.** *Lancet.* 2020. pii: S0140-6736(20)30260-9. doi: 10.1016/S0140-6736(20)30260-9.

6. Linton NM, Kobayashi T, Yang Y, Hayashi K, Akhmetzhanov AR, Jung SM, et al. **Epidemiological characteristics of novel coronavirus infection: A statistical analysis of publicly available case data.** *medRxiv* 2020.01.26.20018754; doi: https://doi.org/10.1101/2020.01.26.20018754

7. Health Commission of Hubei Province, China. [cited 2020 Feb 7]. http://wjw.hubei.gov.cn/

8. Health Commission of Wuhan City, Hubei Province, China [cited 2020 Feb 7] http://wjw.hubei.gov.cn/

9. Clinical guideline for COVID-19, version 5. The State Council, The People's Republic of China. Available from http://www.gov.cn/zhengce/zhengceku/2020-02/05/5474791/files/de44557832ad4be1929091dcbcfca891.pdf [Accessed Feb/29, 2020][in Chinese]

10. Ministry of Health, Labour and Welfare, Japan. https://www.mhlw.go.jp/index.html [in Japanese]

11. 2020 Hubei lockdowns, Wikipedia. https://en.wikipedia.org/wiki/2020_Hubei_lockdowns

12. Li R,a Weiskittel AR,a Kershaw Jr, JA. **Modeling annualized occurrence, frequency, and composition of ingrowth using mixed-effects zero-inflated models and permanent plots in the Acadian Forest Region of North America.** *Can J For Res*. 2011; 41:2077–2089

13. Northeastern University. Laboratory for the Modeling of Biological and Socio‐Technical Systems, 2020. Available online: https://www.mobs‐lab.org/2019ncov.html (accessed on 22 January 2020).

14. Huang C, Wang Y, Li X, et al. **Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China.** *Lancet*. 2020. pii: S0140-6736(20)30183-5.

15. Ghani AC, Donnelly CA, Cox DR, Griffin JT, Fraser C, Lam TH, et al. **Methods for estimating the case fatality ratio for a novel, emerging infectious disease.** *Am J Epidemiol*. 2005; 162: 479-486

16. Nishiura H, Klinkenberg D, Roberts M, Heesterbeek JA. **Early epidemiological assessment of the virulence of emerging infectious diseases: a case study of an influenza pandemic**. *PLoS One*. 2009;4(8):e6852. doi: 10.1371/journal.pone.0006852.

17.Tsuzuki S, Lee H, Miura F, Chan YH, Jung SM, Akhmetzhanov AR, Nishiura H. **Dynamics of the pneumonic plague epidemic in Madagascar, August to October 2017**. *Euro Surveill*. 2017;22(46). doi: 10.2807/1560-7917.ES.2017.22.46.17-00710.

18. Mizumoto K, Chowell G. **Estimating the risk of 2019 Novel Coronavirus death during the course of the outbreak in China, 2020**. *Emerg Infect Dis*. 2020. Accepted.

19. Gamerman, D. & Lopes, H. F. **Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Interference**. 2nd edn (Chapman & Hall/CRC, 2006).

20. Gelman, A. & Rubin, D. B. **Inference from iterative simulation using multiple sequences**. *Stat Sci* **7:**457-472, doi:10.1214/ss/1177011136 (1992).

21. Li Q, Guan X, et. al. **Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia**. *N Engl J Med*. 2020 Jan 29. DOI: 10.1056/NEJMoa2001316 Available at: https://www.nejm.org/doi/full/10.1056/NEJMoa2001316

22. Read JM, Bridgen JR, Cummings DA, Ho A,Jewell CP. **Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions**. *medRxiv*. doi: https://doi.org/10.1101/2020.01.23.20018549

23. Imai N, Cori A, Dorigatti I, Baguelin M, Donnelly CA, Riley S, Ferguson NM. **Report 3: Transmissibility of 2019-nCoV**. https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-2019-nCoV-transmissibility.pdf

24. Mizumoto K, Chowell G. **Transmission potential of the novel coronavirus (COVID-19) onboard the Diamond Princess Cruises Ship, 2020.** *Infect Dis Model*. 2020. 264-270

25. Tariq A, Lee Y, Roosa K, Blumberg S, Yan P, Ma S, Chowell G. **Real-time monitoring the transmission potential of COVID-19 in Singapore, February 2020**. *medRxiv*. doi: https://doi.org/10.1101/2020.02.21.20026435

26. Shim E, Tariq A, Choi W, Lee Y, Chowell G. **Transmission potential of COVID-19 in South Korea.** *medRxiv*. doi:https://doi.org/10.1101/2020.02.27.20028829

27. Chowell G, Abdirizak F, Lee S, Lee J, Jung E, Nishiura H, Viboud C. **Transmission characteristics of MERS and SARS in the healthcare setting: a comparative study.** *BMC Med*. 2015;**13**:210. doi: 10.1186/s12916-015-0450-0.

28. Abdirizak F, Lewis R, Chowell G. **Evaluating the potential impact of targeted vaccination strategies against severe acute respiratory syndrome coronavirus (SARS-CoV) and Middle East respiratory syndrome coronavirus (MERS-CoV) outbreaks in the healthcare setting.** *Theor Biol Med Model*. 2019;**16(1)**:16. doi: 10.1186/s12976-019-0112-6.

29. Blake Essig, Brent Swails, Yoko Wakatsuki and Ben Westcott, CNN. **Top Japanese government adviser says Diamond Princess quarantine was flawed**. Updated 0708 GMT (1508 HKT) February 27, 2020. https://edition.cnn.com/2020/02/27/asia/japan-diamond-princess-quarantine-crew-intl-hnk/index.html

30. Da-hae P, Dam-eun S, Jae-gu K. HANKYOREH. **The reasons why transmission is so prevalent among Shincheonji members**. Mar 2, 2020. http://english.hani.co.kr/arti/english_edition/e_national/930749.html [Accessed Mar/10, 2020]

31. Mizumoto K, Kagaya K, Zarebski A, Chowell G. **Estimating the Asymptomatic Proportion of 2019 Novel Coronavirus onboard the Princess Cruises Ship, 2020**. *Euro Surveill*. 2020. Accepted.

32.Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, et al. **Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China**. *JAMA*. 2020. doi: 10.1001/jama.2020.1585.

33. Mizumoto K, Endo A, Chowell G, Miyamatsu Y, Saitoh M, Nishiura H. **Real-time characterization of risks of death associated with the Middle East respiratory syndrome (MERS) in the Republic of Korea, 2015.** *BMC Med.* 2015;**13**:228. doi: 10.1186/s12916-015-0468-3.

34. Mizumoto K, Saitoh M, Chowell G, Miyamatsu Y, Nishiura H. **Estimating the risk of Middle East respiratory syndrome (MERS) death during the course of the outbreak in the Republic of Korea, 2015**. *Int J Infect Dis.* 2015;**39**:7-9. doi: 10.1016/j.ijid.2015.08.005.

459 35. J.Y. Wong, P. Wu, H. Nishiura, E. Goldstein, E.H. Lau, L. Yang, et al. **Infection**

460 **fatality risk of the pandemic A(H1N1)2009 virus in Hong Kong**. *Am J*

461 *Epidemiol*. 2013;**177 (8)**:pp. 834-840

462 36. Presanis AM, De Angelis D; New York City Swine Flu Investigation Team,

463 Hagy A, Reed C, Riley S, Cooper BS, et al. **The severity of pandemic H1N1**

464 **influenza in the United States, from April to July 2009: a Bayesian analysis**.

465 *PLoS Med.* 2009;**6(12)**:e1000207.

466

Figures

Figure 1. Observed and posterior estimates of the daily new cases and cumulative cases of the COVID-19 cases in Wuhan, China, 2019–2020

Observed and posterior estimates of laboratory–confirmed reported cases (A) and cumulative reported cases (B) are presented.
Observed data are presented in the dot, while dashed line indicates 50 percentile, and areas surrounded by light grey and deep grey indicates 95% and 50% credible intervals (CrI) for posterior estimates, respectively. Epidemic day 1 corresponds to the day that starts at January 1st, 2020.

Figure 2. Observed and posterior estimates of the cumulative deaths of the COVID-19 in Wuhan, China, 2019–2020

Observed and posterior estimates of the cumulative deaths of the COVID-19 in Wuhan is presented. Observed data are presented in the dot, while dashed line indicates 50 percentile, and areas surrounded by light grey and deep grey indicates 95% and 50% credible intervals (CrI) for posterior estimates, respectively. Epidemic day 1 corresponds to the day that starts at January 1st, 2020.

Figure 3.Temporal variation of the infection fatality risks caused by COVID-19 in Wuhan, China, 2019–2020

(A) Posterior estimates of crude infection fatality ratio in Wuhan City. (B) Posterior estimates of time–delay adjusted infection fatality ratio in Wuhan City.

Black dots shows observed data, and light and dark indicates 95% and 50% credible intervals for posterior estimates, respectively. Epidemic day 1 corresponds to the day that starts at January 1st, 2020.

Figure S1. Observed daily new cases and posterior estimates of the daily new infections of the COVID-19 in Wuhan, China, 2019–2020

Observed daily new cases and posterior estimates of infections of the COVID-19 are presented.

Observed data are presented in the dot, while dashed line indicates 50 percentile, and areas surrounded by light grey and deep grey indicates 95% and 50% credible intervals (CrI) for posterior estimates, respectively. Epidemic day 1 corresponds to the day that starts at January 1st, 2020.

Figure S2.Temporal variation of the case fatality risks caused by COVID-19 in Wuhan, China, 2019–2020

(A) Observed and posterior estimates of crude case fatality ratio in Wuhan City, (B) Observed crude case fatality ratio and posterior estimates of time–delay adjusted CFR in Wuhan City.

This figure is submitted to the ref [18]. The purpose of the study is to compare the case fatality ration (CFR. Not IFR) in three different areas (Wuhan City, in Hubei Province excluding Wuhan City and in China excluding Hubei Province) to interpret the current severity of the epidemic in China, and the purpose is different from this study.

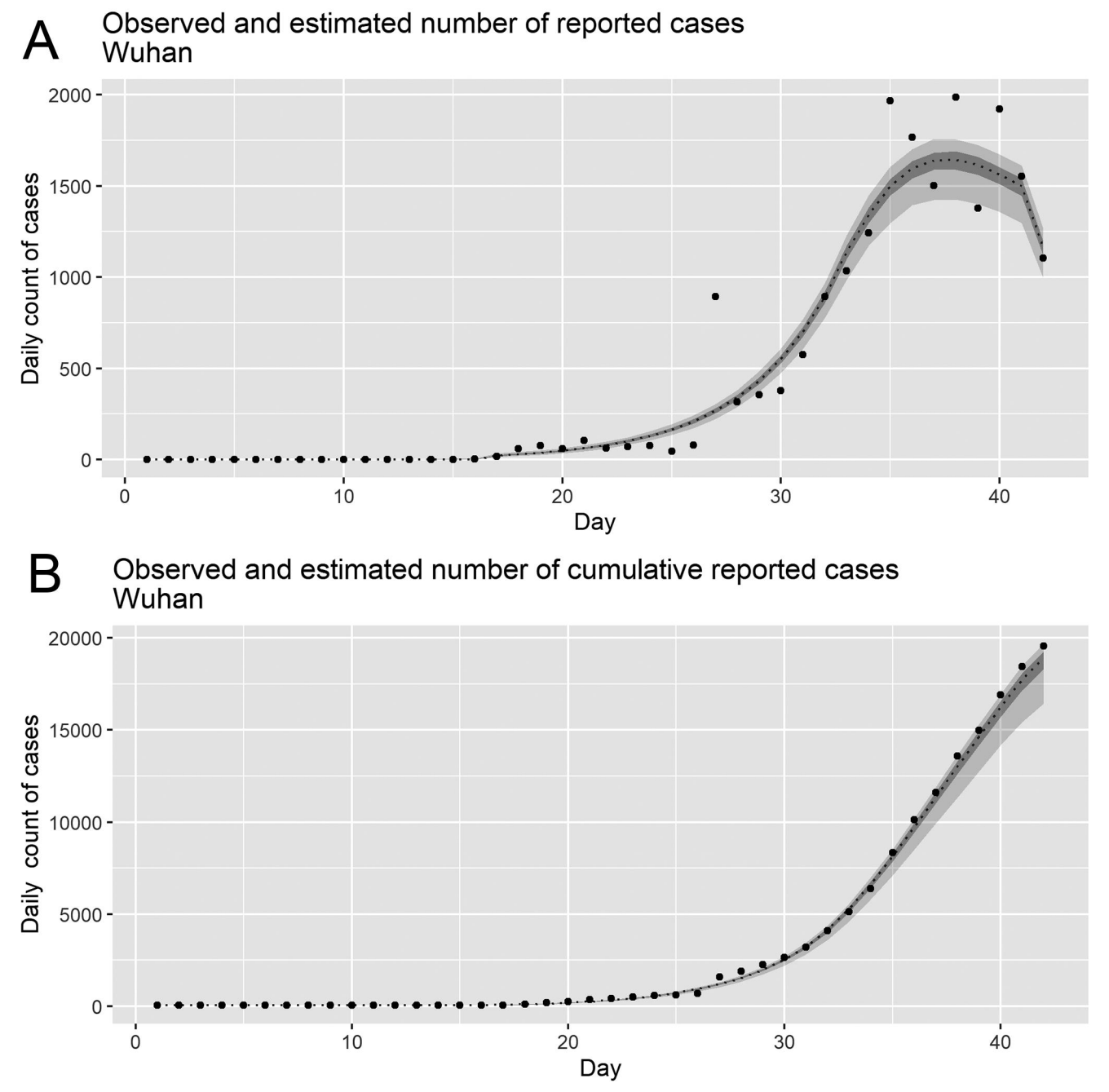
Tables

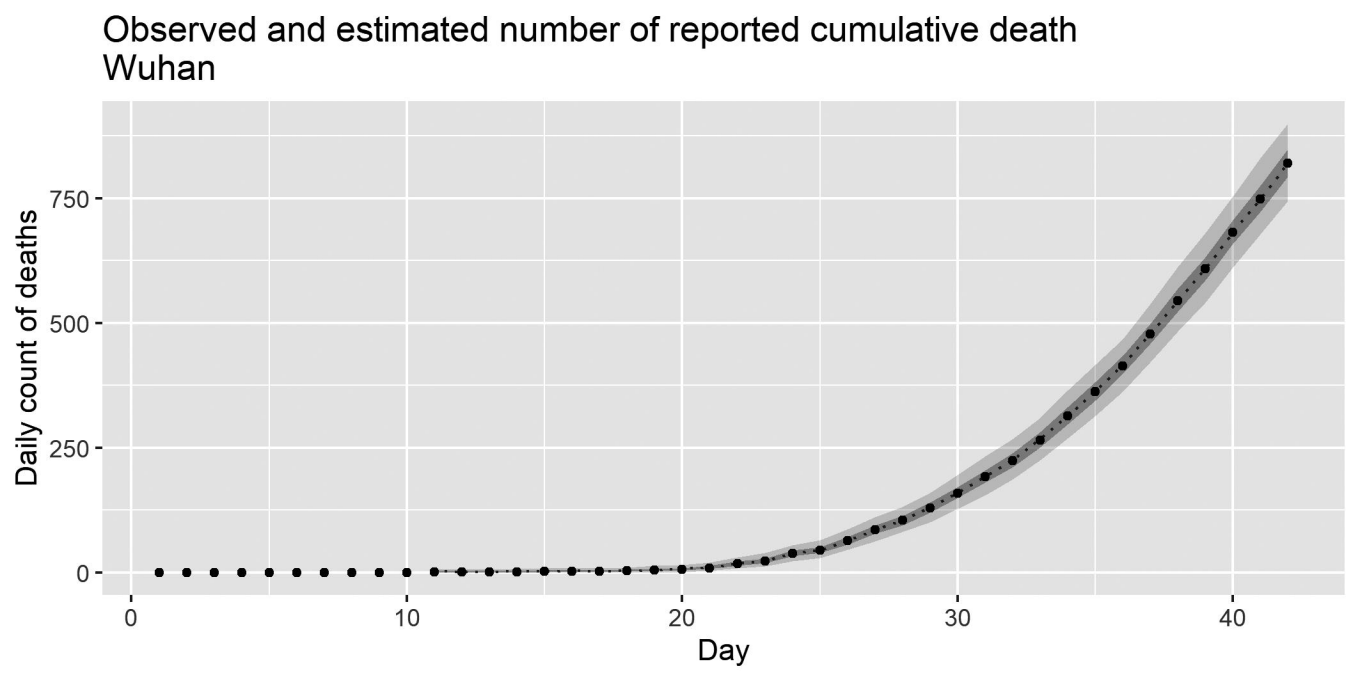
Table 1 – Death risk by COVID-19 in Wuhan City, China, 2020 (As of February 12, 2020)

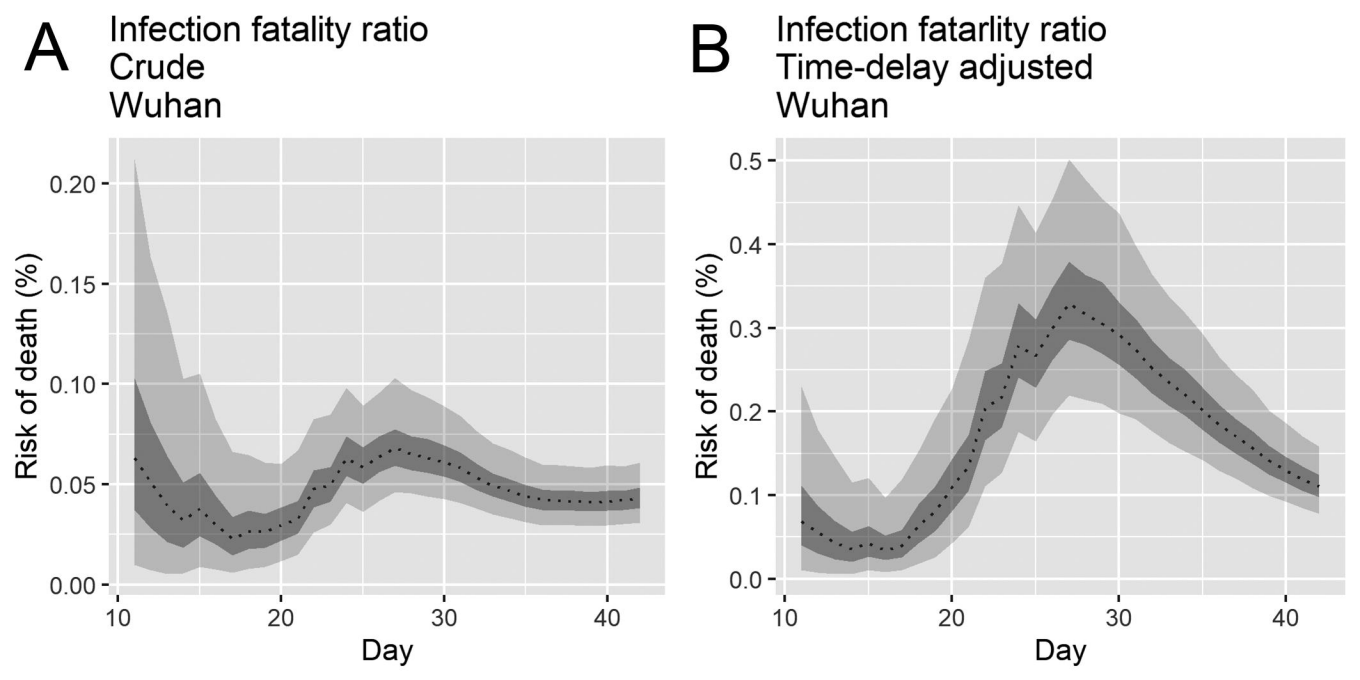
Death Risk	Latest estimate	Range of median estimates
Crude CFR (Observed)	4.19%	2.0 – 9.0%
Crude CFR (Estimated)	4.2% (95%CrI[‡]: 3.9 – 4.9%)	3.4 –7.2%
Time delay adjusted CFR	12.2% (95%CrI: 11.4 – 13.0%)	4.1 – 34.8%
Crud IFR	0.04% (95%CrI: 0.03 – 0.06%)	0.02 – 0.07%
Time delay adjusted IFR	0.12% (95%CrI: 0.08 – 0.17%)	0.04 – 0.33%

CrI: Credibility intervals, CFR: Case fatality ratio, IFR: Infection fatality ratio

[‡]Upper and lower 95% credibility interval


A
Observed and estimated number of reported cases
Wuhan
Daily count of cases
2000
1500
1000
500
0
0
10
20
30
40
Day
B
Observed and estimated number of cumulative reported cases
Wuhan
Daily count of cases
20000
15000
10000
5000
0
0
10
20
30
40
Day


Observed and estimated number of reported cumulative death
Wuhan
Daily count of deaths
750
500
250
0
0
10
20
30
40
Day


A
Infection fatality ratio
Crude
Wuhan
Risk of death (%)
0.00
0.05
0.10
0.15
0.20
10
20
30
40
Day
B
Infection fatarlity ratio
Time-delay adjusted
Wuhan
Risk of death (%)
0.0
0.1
0.2
0.3
0.4
0.5
10
20
30
40
Day