medRxiv preprint doi: https://doi.org/10.1101/2020.02.12.20022434.this version posted February 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Early epidemiological assessment of the transmission potential and virulence of 2019 Novel Coronavirus in Wuhan City: China, 2019–2020

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: 343 (Max 350)

Main: 2575

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.12.20022434.this version posted February 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Abstract

Background:

Since the first cluster of cases was identified in Wuhan City, China, in December, 2019, 2019–nCoV has rapidly spread across China as well as caused multiple introductions in 25 countries as of February, 2020. 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 emerging disease. The present study aimed to estimates of the transmissibility and virulence of 2019–nCov in Wuhan City, China, by reconstructing the underlying transmission dynamics.

Methods:

We employ statistical methods and publicly available epidemiological datasets to jointly derive estimates of transmissibility and severity associated with the novel coronavirus. For estimation, the daily series of laboratory–confirmed nCov cases and deaths in Wuhan City and epidemiological data of Japanese evacuees from Wuhan City on board government–chartered flights were used.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.12.20022434.this version posted February 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Results:

We found that our posterior estimates of basic reproduction number (*R*) in Wuhan City, China in 2019–2020 is calculated to be as high as 7.05 (95%CrI: 6.11–8.18) and the enhanced public health intervention after January 23rd in 2020 has declined *R* to 3.24 (95%CrI: 3.16–3.32), with the total number of infections (i.e. cumulative infections) estimated at 983006 (95%CrI: 759475– 1296258) in Wuhan City, raising the proportion of infected individuals to 9.8% (95%CrI: 7.6–13.0%). We also found that most recent crude infection fatality ratio (IFR) and time–delay adjusted IFR is estimated to be 0.07% (95% CrI: 0.05%–0.09%) and 0.23% (95%CrI: 0.17–0.30%), which is several orders of magnitude smaller than the crude CFR at 4.06%

Conclusions:

We have estimated key epidemiological parameters of the transmissibility and virulence of 2019–nCov in Wuhan, China, 2019-2020 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; 2019–nCov; China

medRxiv preprint doi: https://doi.org/10.1101/2020.02.12.20022434.this version posted February 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Background

The novel coronavirus (2019–nCoV) 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, 2019–nCoV has rapidly spread across China as well as caused multiple introductions in 25 countries as of February, 2020 [2]. Nevertheless, China is bearing the burden of this emerging infectious disease, especially the city of Wuhan located in Hubei province, where the first cluster of severe pneumonia caused by the novel virus was identified. Meanwhile, the cumulative number of laboratory confirmed cases and deaths in mainland China has reached 28001 and 642, respectively, as of February 5th, 2020 [2].

Because the morbidity and mortality burden associated with the novel coronavirus has disproportionally affected the city of Wuhan, 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 28001 2019-nCov cases reported in China, 11618 cases (37.3%) are from Wuhan City. In terms of the death count, a total of 478 deaths (74.5%) have occurred in Wuhan city out of the 642 deaths reported throughout China. To guide the effectiveness of interventions, it is crucial to gauge the uncertainty relating to key epidemiological parameters relating to 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 [3-4]. Recently more evidence suggests that a substantial fraction of the infected individuals with the novel coronavirus show little if any symptoms, which suggest the

medRxiv preprint doi: https://doi.org/10.1101/2020.02.12.20022434.this version posted February 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

need to reassess the transmission potential of emerging disease [5-6]. 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 [2, 8-9]. As of February 8th, 2020, a total of 14982 confirmed cases including 608 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) [6]. As of February 9th, a total of four flights left Wuhan City. We collected information on the dates when those fights left Wuhan City and the number of passengers with confirmed cases to calibrate our model (Table S1)

Statistical analysis

Using the following integral equation model, we estimate the reproduction number of 2019-nCov. Here, infected and reported cases are denoted by i and c, respectively.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.12.20022434.this version posted February 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

We connected a 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,

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

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

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

where $period_1$ represents the corresponding period from the start of study period to January 23rd 2020, 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 (2019–nCoV) to mitigate transmission while parameter β_1 scales 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

medRxiv preprint doi: https://doi.org/10.1101/2020.02.12.20022434.this version posted February 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

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

$$\delta_2 = \begin{cases} \alpha_1, & \text{if } t = period_2, \\ \alpha_2, & \text{if } t = period_3, \\ 1, & \text{otherwise}, \end{cases}$$

where period$_2$ and period$_3$ represents the corresponding periods from the start of our study period to the Jan 17, and from Jan 18 to Jan 20, 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), motivated by a previous study [12]. 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 expectation E[h]. The likelihood function for the time series of observed cases that we employ to estimate the effective reproduction number and other relevant parameters is given by:

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

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

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

medRxiv preprint doi: https://doi.org/10.1101/2020.02.12.20022434.this version posted February 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

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

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

Serial interval estimates of 2019–nCov were derived from previous studies of nCov, indicating that it follows a gamma distribution with the mean and SD at 7.5 and 3.4 days, respectively, based on ref. [14]. The maximum value of the serial interval was fixed at 28 days as the cumulative probability distribution of the gamma distribution up to 28 days reaches 0.999.

Infection fatality ratio

Crude CFR and crude IFR is defined as the number of cumulative deaths 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].

medRxiv preprint doi: https://doi.org/10.1101/2020.02.12.20022434.this version posted February 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

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

where c_t represents the number of new cases with reported day t, and D_t is the number of new deaths with reported day t_i [2,8-9, 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 distribution of each parameter and based on the samples drawn from the posterior distributions. All statistical analyses were conducted in R version 3.5.2 (R Foundation for Statistical Computing, Vienna, Austria) using the 'rstan' package.

Results

The daily series of 2019–nCoV laboratory–confirmed incidence and

medRxiv preprint doi: https://doi.org/10.1101/2020.02.12.20022434.this version posted February 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

cumulative incidence in Wuhan in 2019–2020 are displayed in Figure 1. Overall, our dynamical models yield a good fit to the temporal dynamics (i.e. incidence, cumulative incidence) including an exponential growth pattern in Wuhan. In incidence data, a few fluctuations are seen, probably indicating surveillance system likely missed many cases during the early transmission phase (Figure 1).

Our posterior estimates of basic reproduction number (*R*) in Wuhan City, China in 2019–2020 was estimated to be as high as 7.05 (95%CrI: 6.11–8.18). The time–dependent scaling factor quantifying the extent of enhanced public health intervention on *R* is 0.46 (95%CrI: 0.39–0.54) and this has declined *R* to 3.24 (95%CrI: 3.16–3.32) after January 23rd, 2020. Other parameter estimates for the probability of occurrence and reporting rate are 0.97 (95% CrI: 0.82–1.00) and 0.015 (95% CrI: 0.012–0.02), respectively. Moreover, the time–dependent scaling factor quantifying the extent of reporting rate, α, is estimated to be 0.08 (95% CrI: 0.03–0.21) before January 17 and to be 0.98 (95% CrI: 0.91–1.00) from January 17 to January 20.

The total number of estimated laboratory–confirmed cases (i.e. cumulative cases) is 14433 (95% CrI: 12339–15104) and respectively, while the actual numbers of reported laboratory–confirmed cases during our study period is 14982. Moreover, we inferred the total number of 2019–nCov infections (Figure S1). Our results indicate that the total number of infections (i.e. cumulative infections) is 983006 (95%CrI: 759475–1296258).

The Observed and posterior estimates of the cumulative number of deaths of the 2019–nCov epidemic in Wuhan are displayed in Figure 2, and model–based posterior estimates of the cumulative number of deaths is 610 (95%CrI: 546–680), while actual number of reported deaths is 608. The estimated temporal variation in the

medRxiv preprint doi: https://doi.org/10.1101/2020.02.12.20022434.this version posted February 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

death risk caused by 2019–nCov in Wuhan, China, 2019–2020 is shown in Figure 3 and Figure S2. Observed and posterior estimated of crude CFR in Wuhan City is presented in Figure 2A, while observed and posterior estimates of time–delay adjusted CFR is shown in Figure 2B. Furthermore, Figure 3A and 3B illustrates time–delay no–adjusted IFR and time–delay adjusted IFR, respectively.

The latest estimate of the crude CFR and time–delay adjusted CFR in Wuhan appeared to be 4.51% (95% CrI: 4.02–5.32%) and 15.93% (95% CrI: 14.60–17.28%), respectively, whereas the latest model–based posterior estimates of time–delay not adjusted IFR and adjusted IFR, presented in Figure 3 C and D, are 0.07%(95% CrI: 0.05%–0.09%) and 0.23% (95%CrI: 0.17–0.30%), respectively, while the observed crude CFR is calculated to be 4.06% (Table 1).

Discussion

In this study we derived estimates of the transmissibility and virulence of 2019–nCov in Wuhan City, China, by reconstructing the underlying transmission dynamics. 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 7.05 (95%CrI: 6.11–8.18). The time–dependent scaling factor quantifying the extent of enhanced public health intervention on *R* is 0.46 (95%CrI: 0.39–0.54) and this has declined *R* to 3.24 (95%CrI: 3.16–3.32) after January 23rd in 2020. These *R* estimates capturing the underlying transmission dynamics modify the impact of 2019–Cov, with the total number of

medRxiv preprint doi: https://doi.org/10.1101/2020.02.12.20022434.this version posted February 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

infections (i.e. cumulative infections) estimated at 983006 (95%CrI: 759475– 1296258) in Wuhan City, raising the proportion of infected individuals to 9.8% (95%CrI: 7.6–13.0%) with a catchment population in Wuhan City of 10 million people. These sustained high R values in Wuhan City even after the lockdown and mobility restrictions suggests that transmission is occurring inside the household or in healthcare settings [19], which is a landmark of past SARS and MERS outbreaks [20-21].Considering the potent transmissibility of 2019-nCov in confined settings, as illustrated by the ongoing 2019-nCov outbreak aboard a cruise ship, the Diamond Princess, where the total number of secondary or tertiary infections has reached 135 as of February 10th, 2020 [22], it is crucial to prevent further hospital-based transmission by strengthening infection control measures.

Our most recent estimates of the crude CFR and time−delay adjusted CFR are at 4.51% (95% CrI: 4.02–5.32%) and 15.93% (95% CrI: 14.60–17.28%), respectively. In contrast, our most recent crude IFR and time–delay adjusted IFR is estimated to be 0.07%(95% CrI: 0.05%–0.09%) and 0.23% (95%CrI: 0.17–0.30%), which is several orders of magnitude smaller than the crude CFR at 4.06%. 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 [23-24]. 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 [25-26].

Public health authorities are interested in quantifying *R* and CFR to measure the transmission potential and virulence of an infectious disease, especially when

medRxiv preprint doi: https://doi.org/10.1101/2020.02.12.20022434.this version posted February 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

emerging/re–emerging epidemics occur in order to decide the intensity of the public health response. Given a substantial portion of unobserved infections due to 2019–nCov, *R* estimates derived from infections and IFR are probably more realistic than *R* solely derived from observed cases and the CFR as an index. [19, 27-28]

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 due to a low reporting rate. A very low reporting probability suggests that it is difficult to diagnose 2019–nCov cases or a breakdown in medical care delivery. Moreover, we also identified a remarkable change in reporting rate, estimated to be 12–fold lower in the 1st period (–Jan 16, 2020) and about the same during the 2nd period (January 17 – 20, 2020), relative to the that estimated after January 21st 2020.

Our results are not free from the limitations. First, our methodology aims to capture the underlying transmission dynamics. 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 local residents in Wuhan, underestimating the reproduction number.

Conclusion

In summary, we have estimated key epidemiological parameters of the transmissibility and virulence of 2019–nCov in Wuhan, China, 2019-2020 using an

medRxiv preprint doi: https://doi.org/10.1101/2020.02.12.20022434.this version posted February 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

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.

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.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.12.20022434.this version posted February 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

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.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.12.20022434.this version posted February 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

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. The State Council, The People's Republic of China. [cited 2020 Feb 7]. http://www.gov.cn/
3. 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
4. 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.
5. Linton NM, Kobayashi T, Yang Y, Hayashi K, Akhmetzhanov AR, Jung SM, et al. **Epidemiological characteristics of novel coronavirus infection: A**

medRxiv preprint doi: https://doi.org/10.1101/2020.02.12.20022434.this version posted February 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

statistical analysis of publicly available case data. *medRxiv* 2020.01.26.20018754; doi: https://doi.org/10.1101/2020.01.26.20018754

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

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

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

10. 2020 Hubei lockdowns, Wikipedia. https://en.wikipedia.org/wiki/2020_Hubei_lockdowns

11. 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

12. Zhao S, Lin Q, Ran J, Musa SS, Yang G, Wang W, et al. **Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the**

medRxiv preprint doi: https://doi.org/10.1101/2020.02.12.20022434.this version posted February 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

early phase of the outbreak. *Int J Infect Dis*. 2020. pii: S1201-9712(20)30053-9. doi: 10.1016/j.ijid.2020.01.050.

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**

medRxiv preprint doi: https://doi.org/10.1101/2020.02.12.20022434.this version posted February 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

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**. Submitted.
19. Gamerman, D. & Lopes, H. F. **Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Interference**. 2nd edn (Chapman & Hall/CRC, 2006).
20. 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.
21. 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.
22. Jessie Yeung, Steve George, James Griffiths, Adam Renton and Amy Woodyatt, CNN. Coronavirus kills 97 people in one day, while cruise ship cases almost double. Updated 9:02 p.m. ET, February 10, 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.12.20022434.this version posted February 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

https://edition.cnn.com/asia/live-news/coronavirus-outbreak-02-10-20-intl-hnk/h_4d69299c0570721079dd1115dceefb73

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

24. 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.

25. 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.

26. 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.

27. J.Y. Wong, P. Wu, H. Nishiura, E. Goldstein, E.H. Lau, L. Yang, et al. **Infection fatality risk of the pandemic A(H1N1)2009 virus in Hong Kong**. *Am J Epidemiol*. 2013;**177 (8)**:pp. 834-840

medRxiv preprint doi: https://doi.org/10.1101/2020.02.12.20022434.this version posted February 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

28. Presanis AM, De Angelis D; New York City Swine Flu Investigation Team, Hagy A, Reed C, Riley S, Cooper BS, et al. **The severity of pandemic H1N1 influenza in the United States, from April to July 2009: a Bayesian analysis**. *PLoS Med.* 2009;**6(12)**:e1000207.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.12.20022434.this version posted February 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

Figures

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

Observed and posterior estimates of the cumulative deaths of the 2019–nCov 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 2019–nCov in Wuhan, China, 2019–2020

medRxiv preprint doi: https://doi.org/10.1101/2020.02.12.20022434.this version posted February 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

(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.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.12.20022434.this version posted February 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

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

Observed daily new cases and posterior estimates of infections of the 2019–nCov 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 2019–nCov 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.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.12.20022434.this version posted February 13, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license .

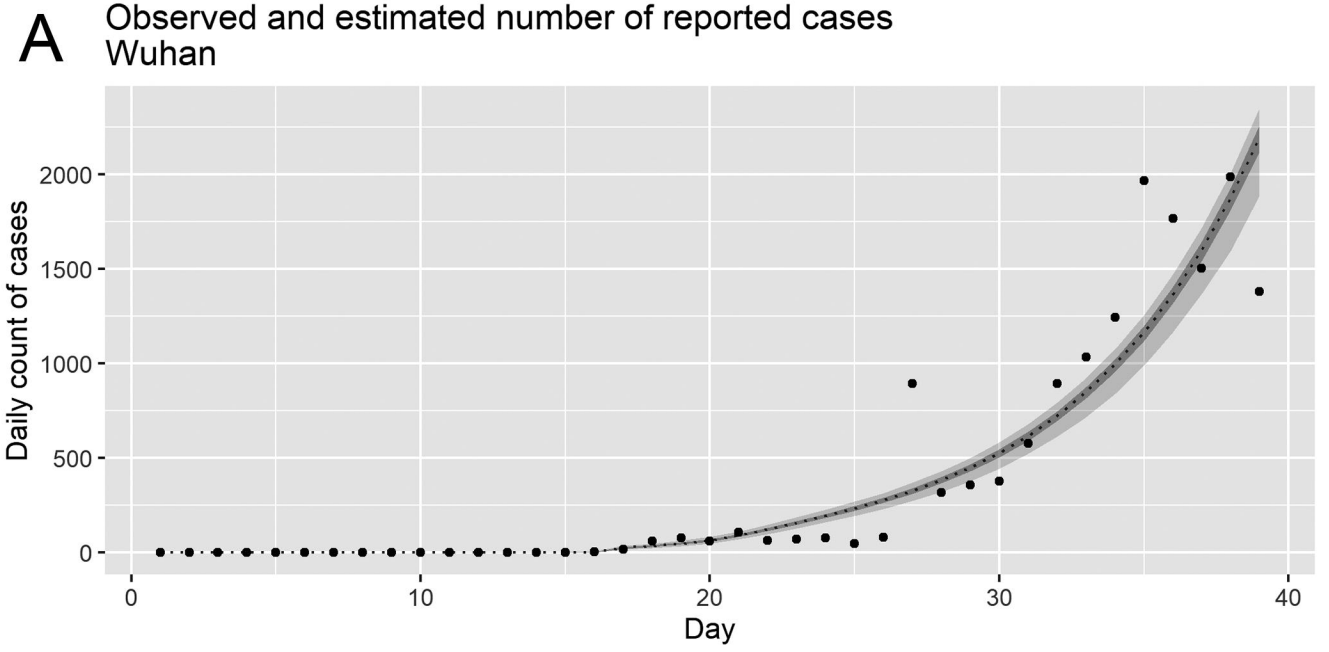
Tables

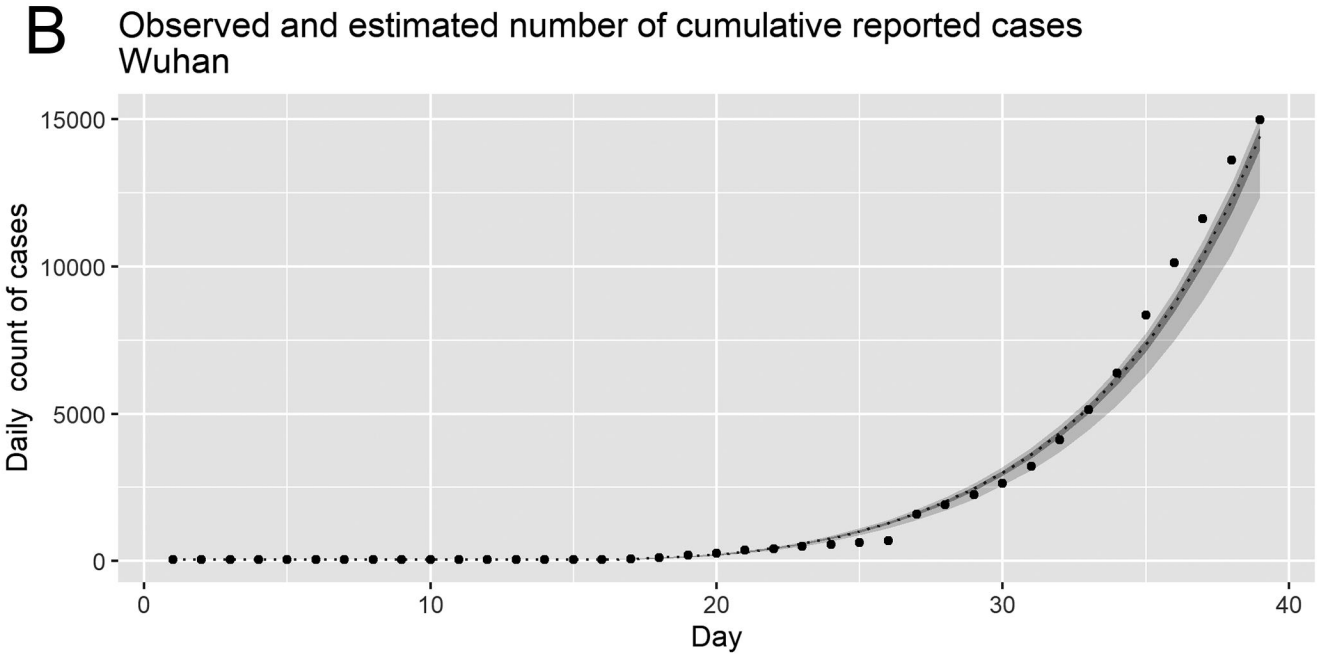
Table 1 – Death risk by 2019–nCov in Wuhan City, China, 2020 (As of February 9, 2020)

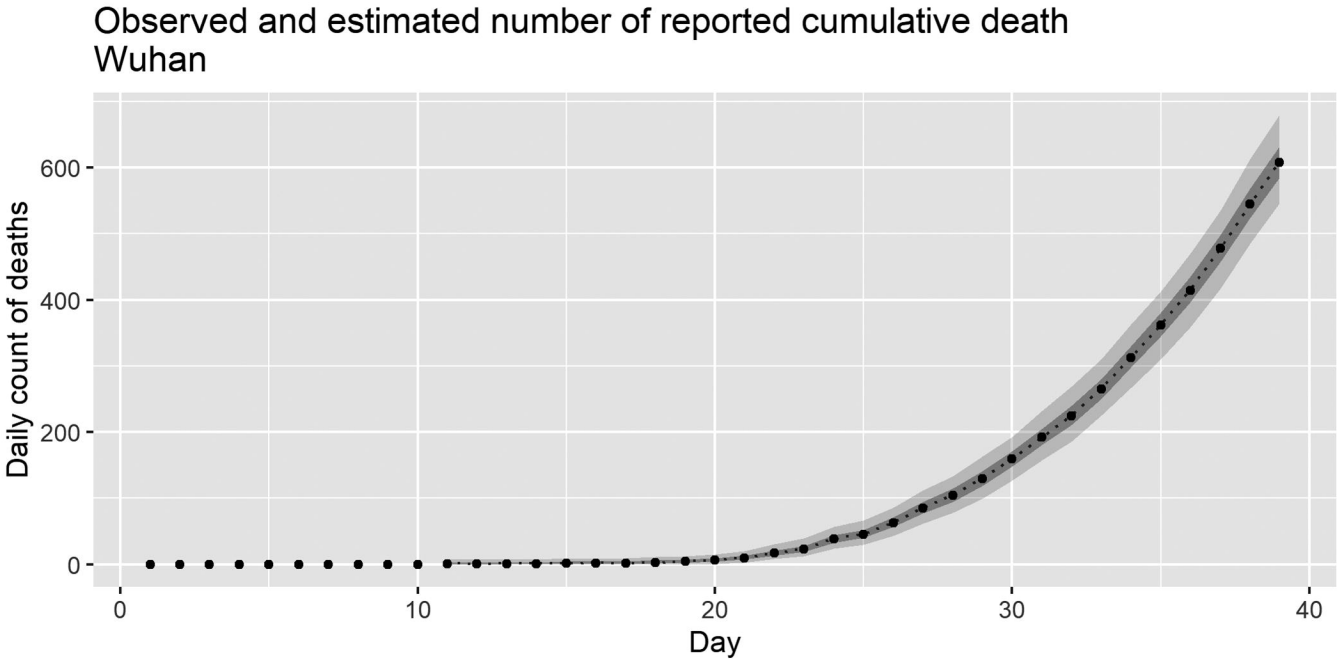
Death Risk	Latest estimate	Range of median estimates
Crude CFR (Observed)	4.06%	2.02 – 9.03%
Crude CFR (Estimated)	4.50% (95%CrI[‡]: 4.02 – 5.31%)	3.00 – 5.80%
Time delay adjusted CFR	15.93% (95%CrI: 14.60 – 17.28%)	4.22 – 34.54%
Crud IFR	0.07% (95%CrI: 0.05 – 0.09%)	0.02 – 0.36%
Time delay adjusted IFR	0.23% (95%CrI: 0.17 – 0.30%)	0.06 – 0.49%

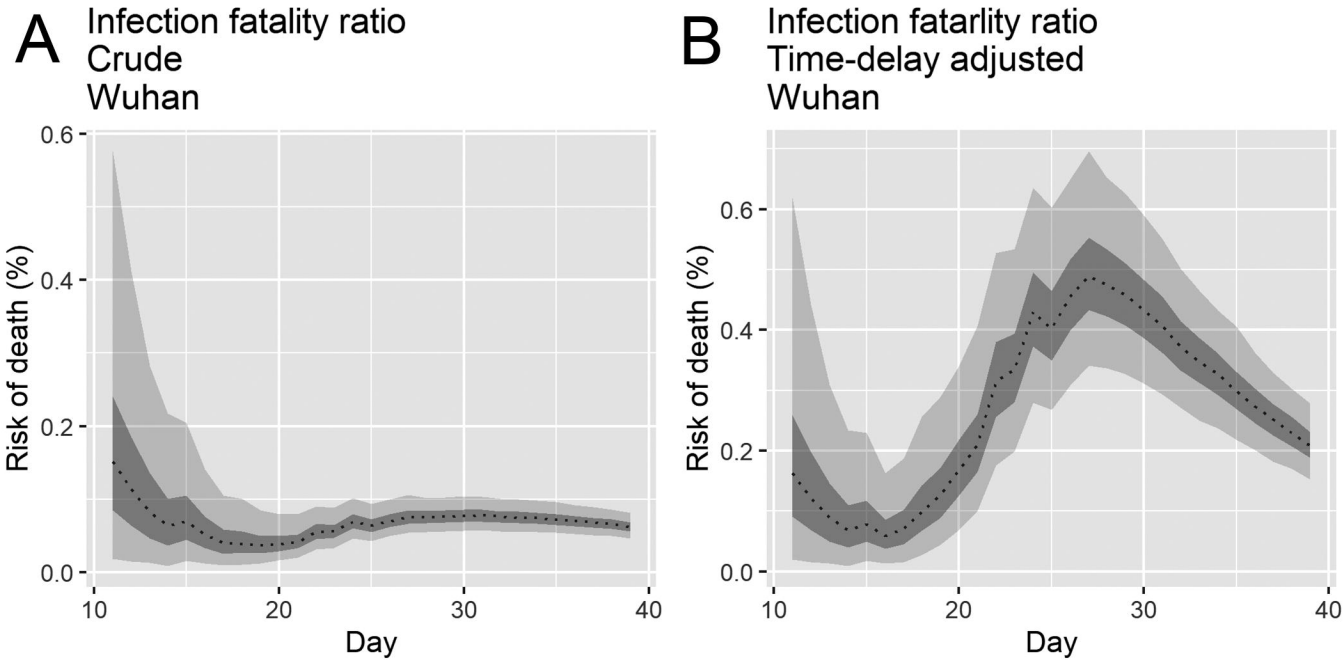
CrI: Credibility intervals, CFR: Case fatality ratio, IFR: Infection fatality ratio

[‡]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
15000
10000
5000
0
0
10
20
30
40
Day


Observed and estimated number of reported cumulative death
Wuhan
Daily count of deaths
600
400
200
0
0
10
20
30
40
Day


A
Infection fatality ratio
Crude
Wuhan
Risk of death (%)
0.6
0.4
0.2
0.0
10
20
30
40
Day
B
Infection fatarlity ratio
Time-delay adjusted
Wuhan
Risk of death (%)
0.6
0.4
0.2
0.0
10
20
30
40
Day