

## **Epidemiological and clinical features of the 2019 novel coronavirus outbreak in China**

Yang Yang<sup>1\*</sup>, Qing-Bin Lu<sup>2</sup>, Ming-Jin Liu<sup>1</sup>, Yi-Xing Wang<sup>3</sup>, An-Ran Zhang<sup>1,3,4</sup>, Neda Jalali<sup>1</sup>, Natalie E. Dean<sup>1</sup>, Ira Longini<sup>1</sup>, M. Elizabeth Halloran<sup>5</sup>, Bo Xu<sup>6</sup>, Xiao-Ai Zhang<sup>3</sup>, Li-Ping Wang<sup>7\*</sup>, Wei Liu<sup>3\*</sup>, Li-Qun Fang<sup>3\*</sup>

<sup>1</sup>Department of Biostatistics, College of Public Health and Health Professions, and Emerging Pathogens Institute, University of Florida, Gainesville, Florida, U. S. A;

<sup>2</sup>Department of Laboratorial Science and Technology, School of Public Health, Peking University, Beijing, P. R. China; <sup>3</sup>State Key Laboratory of Pathogen and Biosecurity, Beijing Institute of Microbiology and Epidemiology, Beijing, P. R. China;

<sup>4</sup> Department of Epidemiology, School of Public Health, Shandong University, Jinan, P. R. China; <sup>5</sup>Vaccine and Infectious Diseases Division, Fred Hutchinson Cancer Research Center, Seattle, Washington, U. S. A.; <sup>6</sup>Army Engineering University of PLA, Nanjing, P.R. China; and <sup>7</sup>Division of Infectious Disease, Key Laboratory of Surveillance and Early-warning on Infectious Disease, Chinese Centre for Disease Control and Prevention.

**\*Correspondence:**

L. Q. Fang or W. Liu, State Key Laboratory of Pathogen and Biosecurity, 20 Dong-Da Street, Beijing Institute of Microbiology and Epidemiology, Fengtai District, Beijing, China 100071 (fang\_lq@163.com or lwbime@163.com); Y. Yang, Department of Biostatistics, College of Public Health and Health Professions, and Emerging Pathogens Institute, University of Florida, Gainesville, FL, USA 32611 (yangyang@ufl.edu); L. P. Wang, Division of Infectious Disease, Key Laboratory of Surveillance and Early-warning on Infectious Disease, Chinese Centre for Disease Control and Prevention, Beijing, China 102206 (wanglp@chinacdc.cn)

## Abstract

**Background** The ongoing outbreak of the 2019 novel coronavirus (2019-nCoV) in China has led to the declaration of Public Health Emergency of International Concern by the World Health Organization.

**Methods** All 2019-nCoV infected patients reported to Chinese Center for Disease Control and Prevention up to 26 January 2020 were included for analysis. Disease and death incidence were compared between demographic groups and baseline conditions. Case fatality rates (CFRs) and the basic reproductive number  $R_0$  was estimated with a transmission model.

**Results:** As of 26 January 2020, a total of 8866 patients including 4021 (45.35%) laboratory confirmed patients were reported from 30 provinces. Nearly half of the patients were aged 50 years or older (47.7%). There was a clear gender difference in incidence with 0.31 (male) vs. 0.27 (female) per 100,000 people ( $P < 0.001$ ). The median incubation period was 4.75 (interquartile range: 3.0-7.2) days. About 25.5%, 69.9% and 4.5% patients were diagnosed with severe pneumonia, mild pneumonia, and non-pneumonia, respectively. The overall CFR was estimated be 3.06% (95% CI 2.02-4.59%), but male patients,  $\geq 60$  years old, baseline diagnosis of severe pneumonia and delay in diagnosis were associated with substantially elevated CFR. The  $R_0$  was estimated to be 3.77 (95% CI 3.51-4.05), ranging 2.23-4.82 in sensitivity analyses varying the incubation and infectious periods.

**Conclusions** Compared with SARS-CoV, 2019-nCoV had comparable transmissibility and lower CFR. Our findings based on individual-level surveillance data emphasize the importance of early detection of elderly patients, particularly males, before symptoms progress to severe pneumonia.

## **Introduction**

The fast-growing outbreak of the 2019 novel coronavirus (2019-nCoV), which originated from Wuhan in central China at the beginning of December 2019, reached multiple continents in merely a month. On 30 January 2020, the World Health Organization (WHO) declared it to be a Public Health Emergency of International Concern.<sup>1,2</sup> As of 31 January 2020, the disease was confirmed in more than 9,000 patients in all 31 provinces of mainland China and 138 patients in 23 other countries and regions (including Hong Kong, Macau and Taiwan, China), with a total of 213 deaths in mainland China and 1 in Hong Kong.<sup>3</sup> Despite the lockdown of the source city since 23 January 2020, the massive human movement during the Chinese traditional new year period may have fueled the spread of the disease. To date, there is no antiviral treatment or vaccine specifically designed for this virus with field-proven effectiveness. Nonpharmaceutical interventions such as shutdown of public gathering places, wearing of facial masks and social distancing could slow the spread of the disease; however, when used alone, these measures may not succeed in fully containing an outbreak of a novel and highly transmissible pathogen.<sup>4</sup>

Despite the difficulty in accurate real-time assessment of the epidemic because of the insufficient supply of testing kits and delayed reporting by overwhelmed healthcare facilities, a few studies have made progress to understand the molecular, clinical and epidemiological features of the 2019-nCoV.<sup>2,5-11</sup> For example, the 2019-nCoV more severely affect older patients with comorbidities.<sup>12</sup> Analysis on family clusters and an assessment of the basic reproductive number based on early investigation data have shown efficient person-to-person transmissibility of the virus.<sup>7,10</sup> Nevertheless, there is an urgent need to verify and update these early findings as the number of patients accumulates. Here we summarize our findings on clinical and epidemiological characteristics of the 2019-nCoV based on the surveillance data of confirmed and suspected 2019-nCoV patients in China up to 26 January 2020.

## Methods

### Data sources

Soon after 2019-nCoV was identified as the etiological pathogen of the pneumonia outbreak, the disease was classified as Class B infectious disease and managed as Class A.<sup>13</sup> Confirmed and suspected patients are required to be reported within 24 hours to the National Notifiable Infectious Disease Surveillance System, according to the standard protocol issued by National Health Commission of the People's Republic of China (NHCC).<sup>14</sup> Publicly accessible cumulative numbers of confirmed and suspected cases (see Supplementary Methods for definitions), which are updated daily by the NHCC, were collected for preliminary analyses.<sup>15</sup> The data on confirmed and suspected patients reported by 26 January 2020 were obtained from the Chinese Public Health Science Data Center. Case data included basic demographic information, case classification, date of symptom onset, date of diagnosis, date of hospitalization, date of discharge or date of death, etc. Population data at the prefecture (city) level in the year of 2017 were obtained from the National Bureau of Statistics of the People's Republic of China.

### Statistical analysis

Incidence (per 100,000 people) is defined as the number of patients as of 26 January 2020 divided by the population size in a given prefecture. Observed case fatality rate (CFR) is defined as the observed number of deaths divided by the number of patients. The patient's location was mapped at the prefecture level using a geographic information system (ArcGIS, Environmental Systems Research Institute, Redlands, California). Statistical analyses such as between-sample comparisons were performed using R 3.6.1 (R core team, 2017) and Stata, version 14.0 (Stata Corp LP, College Station, TX, USA).

To account for the fact that the final clinical outcome has not resolved for the majority of patients, especially the newly identified cases, we estimate the CFR by restricting the analysis to patients with symptom onset at least  $d$  days earlier than 26

January 2020,  $d$  varying from 10 to 14 days. For each value of  $d$ , a CFR is calculated as the proportion of fatal cases among all patients meeting this restriction. The average of these estimates serves as the final estimate of CFR, which is referred to as adjusted CFR. We estimated adjusted CFR by age group, gender, baseline severity level and delay in diagnosis for confirmed patients only.

To estimate the basic reproductive number,  $R_0$ , defined as the average number of secondary infections a patient can generate in a fully susceptible population, we used a chain binomial model.<sup>16,17</sup> Another important epidemiological quantity is the effective reproductive number,  $R_t$ , that measures the variation of a pathogen's transmissibility over time in response to, e.g., climate change or intervention programs. The estimation of  $R_t$  is achieved with a sliding time window of 5 days to estimate within-window transmissibility. Due to lack of individual level contact data, we assume homogeneous mixing and restricted the analysis to the transmission dynamics in Wuhan. Sensitivity analyses were performed by varying settings of the natural history of disease, epidemic growth phase, and for confirmed only vs. confirmed and suspected patients. For comparison, we also estimated  $R_t$  using an R package EpiEstim.<sup>18</sup>

Data collection and analysis were considered as part of a continuing public health outbreak investigation and exempt from institutional review board approval.

## **Results**

### **Epidemiological description**

The earliest symptom onset of confirmed patients can be traced back to 7 December 2019. As of 26 January 2020, a total of 8866 probable patients were reported and 4021 (45.35%) were laboratory-confirmed, in 30 provinces of China (Table S1). The mean ( $\pm$ standard deviation) age among confirmed patients was  $49\pm 16$  years, comparable with suspected patients. However, the proportion of females differs slightly, 45% among confirmed and 49% among suspected patients. Fourteen children

<10 years old were confirmed to be infected with 2019-nCoV. The median time from disease onset to diagnosis among confirmed patients was 5 (IQR: 2-9) days (Fig. S1A). The median delay from symptom onset to diagnosis decreased dramatically from before 14 January 2020 to after 22 January 2020 (Fig. S1B-S1D). About 41.38% (1664/4176) of confirmed patients occurred in Wuhan city (Table S1).

The disease incidence initially remained low and sporadic until January 1, 2020, when an abrupt jump was seen, followed by an exponential growth until 23 January 2020 (Fig. S2A). This increasing phase overlaps with the population movement period before the spring festival that officially started on 10 January 2020. The declining trend during the last three days of the study period was likely due to the delay in diagnosis or reporting, as the spring festival occurred on January 25. January 17 seems to be a tipping point of the epidemic, when the daily number of confirmed patients outside Wuhan begin to surpass that in Wuhan (Fig. S2B). Suspected patients had the similar pattern as that of the confirmed patients, however with a 2–3 days lag (Fig. S2C). The fast increase in the number of patients in Wuhan during the middle of January was followed by the spread to other provinces during the second half of the month, particularly to the neighboring provinces such as Henan, Sichuan and Hunan (Fig. S2D–E). Zhejiang, and Guangdong, the provinces popular for migrant workers, also saw sharp increases. When the number of affected towns was analyzed, a similar trend was observed, except that a remarkable increase in affected towns was observed outside Hubei province (Fig. S3).

Geographic clustering of patients is clearly seen at the township level, mostly in Hubei province (65.35% of all patients), with incidences ranging from 0.75 per 100,000 to 15.81 per 100,000 at the prefecture level. Other clusters are notable in neighboring provinces as well as in Beijing, the Yantze River delta near Shanghai, southeast of Zhejiang, and the Pearl River delta near Guangzhou and Hong Kong. The earlier the reporting of confirmed patients, the higher the incidence at the prefecture level (Fig. 1A). The reporting time of the first case was negatively correlated with the size of outflow population from Wuhan to each affected prefecture during 1–26

January 2020 (Spearman correlation,  $r=-0.568$ ,  $P<0.001$ ) (Fig. 1B).

Patients 30–65 years old dominated confirmed patients (2873, 71.45%), with the highest case number of 139 at 56 years old. A similar age distribution was observed in the suspected patients (Fig. 2A-C). Highest incidence was observed among adults  $\geq 50$  years, with the lowest incidence in the age group younger than 20 years (Fig. 2D-F and Table S2). Overall, males experienced a higher incidence than females (0.31 per 100,000 vs. 0.27 per 100,000,  $P<0.001$ ). However, age-gender pattern differs between Wuhan, the source of the outbreak, and other areas. The high incidence subpopulation outside Wuhan tended to be younger than that in Wuhan. Significant gender difference was only found outside Wuhan, 0.19 per 100,000 among males vs. 0.15 per 100,000 among females ( $P<0.001$ ).

Among the confirmed patients, 935 (25.5%) and 2563 (69.9%) were diagnosed with severe and mild pneumonia, respectively, and 167 (4.5%) had no evidence of pneumonia (Table S3). Patients diagnosed with severe pneumonia were significantly older (Mean $\pm$ SD, 55 $\pm$ 15 years old) and had a higher proportion of males (61.5%), in comparison to those with mild pneumonia (45 $\pm$ 15 years old and 52.71% male) and non-pneumonia (42 $\pm$ 16 years old and 52.1% male). For 13 pediatric patients  $\leq 10$  years old with severity information, 10 had mild pneumonia and 3 had non-pneumonia. The median (IQR) interval from disease onset to diagnosis was longer for severe pneumonia, 8 (4-12) days, than for mild pneumonia, 4 (2-7) days, and non-pneumonia, 3 (1-5) days. The proportion of non-pneumonia increased from earlier disease onset to late disease onset. Higher frequency of mild pneumonia and non-pneumonia was diagnosed from Wuhan than from the regions outside Wuhan.

### **Case fatality of 2019-nCoV infection**

Fatal outcome developed in 58 confirmed patients and 18 suspected patients, the former leading to an observed CFR of 1.44% (95% CI 1.10%-1.86%). Confirmed patients with severe pneumonia experienced a much higher observed CFR, 5.88%, than those with mild pneumonia (0.12%) and non-pneumonia (0%). Additional results on observed CFRs can be found in Supplementary Appendix, Table S4 and Fig. S3.



We estimated the overall adjusted CFRs among confirmed patients to be 3.06% (95% CI 2.02-4.59%) (Table 1). The adjusted CFR in male patients more than tripled that in female patients, 4.45% (95% CI 2.81-6.93%) vs. 1.25% (95% CI 0.43-3.29%). Patients 60 years or older were also subject to a much more excessive adjusted CFR of 5.30% (95% CI 3.25-8.46%), compared to the younger patients, 1.43% (95% CI 0.61-3.15%). Diagnosis of severe pneumonia at baseline is another leading risk factor for death, associated with an adjusted CFR of 6.23% (95% CI 3.87-9.79%). The adjusted CFR among patients with mild or no pneumonia at baseline was relatively low, 0.68-1.16%, depending on whether patients with unknown baseline severity were classified as mild or not (Supplementary Methods). A delay from onset to diagnosis >5 days also doubled the adjusted CFR from 1.34% (95% CI 0.35-5.12%) to 3.07% (95% CI 2.02-4.60%). Further stratified analyses by more than one baseline variables found that patients meeting any two of the three characteristics, male patient, 60 years or older and severe pneumonia at diagnosis, had further increased CFRs. In particular, the adjusted CFR reached as high as 9.47% (5.34-15.99%) among older male patients diagnosed with severe pneumonia (Table S5).

### **Person to person transmissibility of 2019-nCoV**

As a necessary input for estimating  $R_0$ , the distribution of the incubation period was estimated from 125 patients with clearly defined exposure periods (Supplementary Methods). We estimated a median (IQR) duration of 4.8 (3.0, 7.2) days for the incubation period (Fig. S5), based on which we constructed several sensitivity analysis settings for the incubation and infectious periods to assess person-to-person transmissibility of the 2019-nCoV (Table S6). Using a mean incubation period of 5 days and a mean infectious period of 7 days, we estimated the  $R_0$  to be 3.77 (95% CI 3.51-4.05), assuming 100% initial reporting rate before 1 January 2020 (Table 2). The various settings of the natural history of disease and reporting rate yielded  $R_0$  estimates ranging from 2.23 to 4.82. Higher estimates are associated with longer incubation and infectious periods and a higher initial reporting rate. Using all confirmed and suspected patients for estimation gives slightly higher

estimates, ranging from 2.31 to 5.18 (Table S7). Moving the estimation window from the period of 25 December 2019 – 12 January 2020 to the period of 28 December 2019 – 15 January 2020 lowered the estimate to 3.00 (95% CI 2.81-3.20) based on confirmed patients (Table S8) and 3.28 (95% CI 3.14-3.43) based on all patients (Table S9), under the median incubation and infectious periods and 100% initial reporting. The effective reproductive number,  $R_t$ , started to cross the critical threshold of 1 near 25 December 2019, and peaked at 8-15 around 3 January 2020 depending on the setting of the natural history of disease, and quickly descended to below the threshold near 16 January 2020 (Fig. 3). The declining trend was most likely due to delayed reporting rather than decreasing transmissibility. The shape of the  $R_t$  curve is moderately sensitive to the incubation period but not to the infectious period. Similar to the results for  $R_0$ , a lower initial reporting rate was associated with a lower  $R_t$  curve (Fig. S6-S7). Using all patients led to high  $R_t$  during the early phase of the epidemics but peak values are similar (Figs. S8-S10). The  $R_t$  curve obtained using the EpiEstim package shows a bimodal shape, with the two modes near 22 December 2019 and the middle of January in 2020 (Fig. S11). Assuming a mean serial interval of 9 days, the average  $R_t$  during 25 December 2019 – 12 January 2020 is close to 3, comparable to our estimate for  $R_0$ . A longer mean serial interval is associated with a higher  $R_t$  curve, consistent with our method.

## Discussion

Similar to the severe acute respiratory syndrome coronavirus (SARS-CoV) and the Middle East Respiratory Syndrome coronavirus (MERS-CoV), the 2019-nCoV adversely affected the elderly male subpopulation more than any other subpopulations, in terms of both the highest incidence of confirmed patients and the highest CFR.<sup>10</sup> The higher prevalence of ACE2 receptors, to which the virus binds, in the lungs of Asian males could have contributed to this gender difference.<sup>19</sup> The age effect is also obvious in female patients, although to a less extent. The shift to a younger age profile among patients who were identified outside Wuhan could be due to the fact that these patients who had travelled to or from Wuhan tend to represent a younger population.

While more data are needed to exclude the possibility of viral adaptation to younger hosts, the observed CFR is similar between Wuhan and non-Wuhan patients after adjusting for age and gender.

We estimated the overall adjusted CFR to be 3.06% for the 2019-nCoV, which is lower than those of SARS-CoV (9.2%) and MERS-CoV (34.4%).<sup>20</sup> The limited number of discharges so far has impeded the use of any advanced method such as the competing risk model for estimating CFR.<sup>21</sup> We restricted our analyses to patients with symptom onset at least 10 days before 26 January 2020 to reduce bias from unresolved final clinical outcomes. However, another serious source of bias is under-detection and under-reporting of mild cases especially those without pneumonia (4.5% in our data), which in turn could have substantially inflated the overall CFR estimate. Therefore, our estimate should be treated as an upper bound. The shortened delay from symptom onset to diagnosis over the epidemic course might have helped reducing CFR. Early diagnosis of elderly patients, especially males, with fever or respiratory symptoms before they progress to severe pneumonia is thus an important target for preventing fatal outcomes.

In addition to atypical non-pneumonia patients, a couple of asymptomatic infections have also been reported for both adults and children.<sup>7</sup> An asymptomatic young woman who came back from Wuhan was suspected to be the source infector of three of her relatives in her hometown in Henan Province who had no travel history to Wuhan.<sup>22</sup> It remains necessary to evaluate the transmissibility of mild or asymptomatic infections and their contribution to the overall epidemic, using both field contact tracing data and modeling approaches.

Our estimate for  $R_0$ , 3.77, is higher than recently published estimates. An estimate of  $R_0=2.0$  based on 425 early reported patients is likely an underestimation given the serious delay in case confirmation during the early phase.<sup>11</sup> Another estimate of  $R_0=2.7$  was also based on surveillance data, but the methodology was different.<sup>23</sup> Our method considers right censoring of infections who might not have developed symptoms by the end of the study period, which is important given the ongoing

increasing trend of the epidemic. In addition, our analysis distinguished patients living and diagnosed within Wuhan from those Wuhan residents who were diagnosed outside Wuhan, a major source of exporting the disease to other parts of China. The transmissibility of 2019-nCoV is comparable to that of SARS-CoV in the range of 2.9-3.3<sup>24</sup> and much higher than that of MERS-CoV. Both 2019-nCoV and SARS-CoV resulted from occasional spill-over from non-specific intermediate host mammals, whereas MERS-CoV has a clear and constant animal reservoir.

Our findings are constrained to a limited time frame and might have missed time-changing disease features, given the fast evolution of the epidemic. The most recent curve of daily new numbers of confirmed patients shows that the exponential growth extended beyond 25 January 2020 with a steeper slope,<sup>25</sup> indicating the possibility of underestimation in our analysis. In particular, the declining trend of  $R_t$  after the beginning of January was most likely due to delayed reporting rather than reduced transmissibility, as many cases with symptom onsets during the last week of the study period had not been reported to China CDC when this analysis was performed. However, further evaluation of transmissibility will need to address the complexity raised by implementation of nonpharmaceutical interventions such as travel ban.

It remains a challenging task to contain an outbreak of a novel pathogen capable of efficient person to person transmission in this highly mobile world, in particular when treatment and prevention options are limited. Existing antiviral treatments such as lopinavir/ritonavir and remdesivir have been evaluated for treating SARS-CoV and MERS-CoV infections,<sup>26-28</sup> and some have been used or are being considered for treating 2019-nCoV infections.<sup>29</sup> Clinical trials need to be carefully designed and implemented to assess their efficacies, which could be challenging given the overwhelmed healthcare resources in China. Equally important is improvement in case detection and management in the most vulnerable elderly population.

## **Acknowledgement**

This work was financially supported by grants from the China Mega-Project on Infectious Disease Prevention (No. 2018ZX10713001, 2018ZX10713002, 2017ZX10103004 and 2018ZX10101003-002), the National Natural Science Funds (No. 81825019), and the U.S. National Institute of Health (R01 AI139761 and R01 AI116770). We thank the staff members at the China CDC for their assistance in data preparation.

## References

1. WHO. Statement on the second meeting of the International Health Regulations (2005) Emergency Committee regarding the outbreak of novel coronavirus (2019-nCoV). 2020. at [https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-\(2005\)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-\(2019-ncov\).](https://www.who.int/news-room/detail/30-01-2020-statement-on-the-second-meeting-of-the-international-health-regulations-(2005)-emergency-committee-regarding-the-outbreak-of-novel-coronavirus-(2019-ncov).))
2. Zhu N, Zhang D, Wang W, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* 2020.
3. National Health Commission of the People's Republic of China. Update on the novel coronavirus pneumonia outbreak (Jan 31, 2020). 2020. at <http://www.nhc.gov.cn/xcs/yqtb/202002/84faf71e096446fdb1ae44939ba5c528.shtml>.)
4. Peak CM, Childs LM, Grad YH, Buckee CO. Comparing nonpharmaceutical interventions for containing emerging epidemics. *Proc Natl Acad Sci U S A* 2017;114:4023-8.
5. Benvenuto D, Giovannetti M, Ciccozzi A, Spoto S, Angeletti S, Ciccozzi M. The 2019-new coronavirus epidemic: evidence for virus evolution. *J Med Virol* 2020.
6. Chan JF, Kok KH, Zhu Z, et al. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. *Emerg Microbes Infect* 2020;9:221-36.
7. Chan JF-W, Yuan S, Kok K-H, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet* 2020.
8. Chen N, Zhou M, Dong X, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* 2020.
9. Hon KL, Leung CW, Cheng WT, et al. Clinical presentations and outcome of severe acute respiratory syndrome in children. *Lancet* 2003;361:1701-3.
10. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020.
11. Li Q, Guan X, Wu P, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N Engl J Med* 2020.
12. Zhang N, Wang L, Deng X, et al. Recent advances in the detection of respiratory virus infection in humans. *J Med Virol* 2020.
13. National Health Commission of the People's Republic of China. Announcement on novel coronavirus. 2020. at <http://www.nhc.gov.cn/xcs/zhengcwj/202001/44a3b8245e8049d2837a4f27529cd386.shtml>.)
14. National Health Commission of the People's Republic of China. Diagnosis and treatment of

- pneumonia infected with novel coronavirus. 2020. at <http://www.nhc.gov.cn/yzygj/s7653p/202001/f492c9153ea9437bb587ce2ffcbee1fa.shtml>.)
15. National Health Commission of the People's Republic of China. The latest news on Health Emergency Office. 2020. at [http://www.nhc.gov.cn/yjb/pqt/new\\_list.shtml](http://www.nhc.gov.cn/yjb/pqt/new_list.shtml).)
  16. Yang Y, Longini IM, Jr., Halloran ME. Design and Evaluation of Prophylactic Interventions Using Infectious Disease Incidence Data from Close Contact Groups. *J R Stat Soc Ser C Appl Stat* 2006;55:317-30.
  17. Yang Y, Sugimoto JD, Halloran ME, et al. The transmissibility and control of pandemic influenza A (H1N1) virus. *Science* 2009;326:729-33.
  18. Cori A, Ferguson NM, Fraser C, Cauchemez S. A new framework and software to estimate time-varying reproduction numbers during epidemics. *Am J Epidemiol* 2013;178:1505-12.
  19. Zhao Y, Zhao Z, Wang Y, Zhou Y, Ma Y, Zuo W. Single-cell RNA expression profiling of ACE2, the putative receptor of Wuhan 2019-nCoV. 2020.
  20. Wang C, Horby PW, Hayden FG, Gao GF. A novel coronavirus outbreak of global health concern. *Lancet* 2020.
  21. Kalbfleisch JD, Prentice RL. The statistical analysis of failure time data. *J Am Stat Assoc* 2003;98:1084-5.
  22. The Mammoth News. Five of the woman's relatives were diagnosed with no symptoms after she returned to Anyang from Wuhan. 2020. at <https://news.sina.com.cn/s/2020-01-28/doc-iihnzakh6756480.shtml>.)
  23. 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. *The Lancet* 2020.
  24. WHO. Consensus document on the epidemiology of severe acute respiratory syndrome (SARS). . 2003. at <https://www.who.int/csr/sars/en/WHOconsensus.pdf>.)
  25. China CDC. Distribution of pneumonic outbreaks of novel coronavirus infection. 2020.
  26. Chu CM, Cheng VC, Hung IF, et al. Role of lopinavir/ritonavir in the treatment of SARS: initial virological and clinical findings. *Thorax* 2004;59:252-6.
  27. Can an anti-HIV combination or other existing drugs outwit the new coronavirus? 2020. at <https://www.sciencemag.org/news/2020/01/can-anti-hiv-combination-or-other-existing-drugs-outwit-new-coronavirus>.)
  28. Sheahan TP, Sims AC, Leist SR, et al. Comparative therapeutic efficacy of remdesivir and combination lopinavir, ritonavir, and interferon beta against MERS-CoV. *Nat Commun* 2020; 11:222.
  29. Holshue ML, DeBolt C, Lindquist S, et al. First Case of 2019 Novel Coronavirus in the United States. *N Engl J Med* 2020.

### **Figure Legends:**

#### **Fig. 1. The geographic location of 4021 patients with confirmed 2019-nCoV infection (Panel A) and the spatial dispersion of confirmed patients from Wuhan city to the other 253 cities in China (Panel B), as of 26 January 2020.**

Panel A, the illness onset day of each patient (blue) and the incidence rate of each city (red) are color-differentiated.

Panel B, the time interval (days) from the first case-reporting date of each affected city to 7 December 2019 (the first patient reported in China) is color-differentiated in blue. The location of each dot shows the township of the first patient in each affected city.

The proportion of Wuhan-originated outflow population to each affected city during 1-26 January 2020 is color-differentiated in red as background. Transportation networks are shown in green for railways and light brown for highways and freeways.

#### **Fig. 2. Age and sex patterns of the patients infected with 2019-nCoV.**

Age distribution stratified by sex is shown for all patients (Panel A), for confirmed patients (Panel B) and for suspected patients (Panel C). Disease incidence stratified by sex is shown for confirmed patients in mainland China (Panel D), for confirmed patients reported in Wuhan city (Panel E) and for confirmed patients reported outside Wuhan city (Panel F).

#### **Fig. 3. Estimates (solid red) and the 95% pointwise confidence band (dashed red) for real time effective reproductive numbers from 24 December 2019 to 18 January 2020 based on confirmed cases who are residents of Wuhan. Results are stratified by assumption settings about the incubation (inc.) and infectious (inf.) periods. Reporting rate before 1 January 2020 is assumed to be 100%. Numbers of confirmed cases are shown as the histogram (gray). Model-predicted case numbers (solid green) and 95% confidence band (green dashed) are also shown.**

<b>Table 1. Estimates for adjusted case fatality rate (CFR) among 2019-nCoV patients, stratified by case type and baseline characteristics.</b>				
Case type	Missing severity classified as mild	No. of cases	No. of deaths (Observed CFR, %)	CFR, % (95% CI)
<b>Confirmed patients</b>				
<b>Gender</b>				
Male		2213	44 (1.99)	4.45 (2.81, 6.93)
Female		1808	14 (0.77)	1.25 (0.43, 3.29)
<b>Age Group</b>				
<60 y		2969	15 (0.51)	1.43 (0.61, 3.15)
≥60 y		1052	43 (4.09)	5.30 (3.25, 8.46)
<b>Severity</b>				
Severe		927	47 (5.07)	6.23 (3.87, 9.79)
Mild	Yes	3094	11 (0.36)	1.16 (0.47, 2.69)
	No	2727	3 (0.11)	0.68 (0.13, 3.07)
<b>Time from onset to diagnosis</b>				
≤5 days <sup>§</sup>		2054	18 (0.88)	1.34 (0.35, 5.12)
>5 days		1967	40 (2.03)	3.07 (2.02, 4.60)
<b>Overall</b>		<b>4021</b>	<b>58 (1.44)</b>	<b>3.06 (2.02, 4.59)</b>

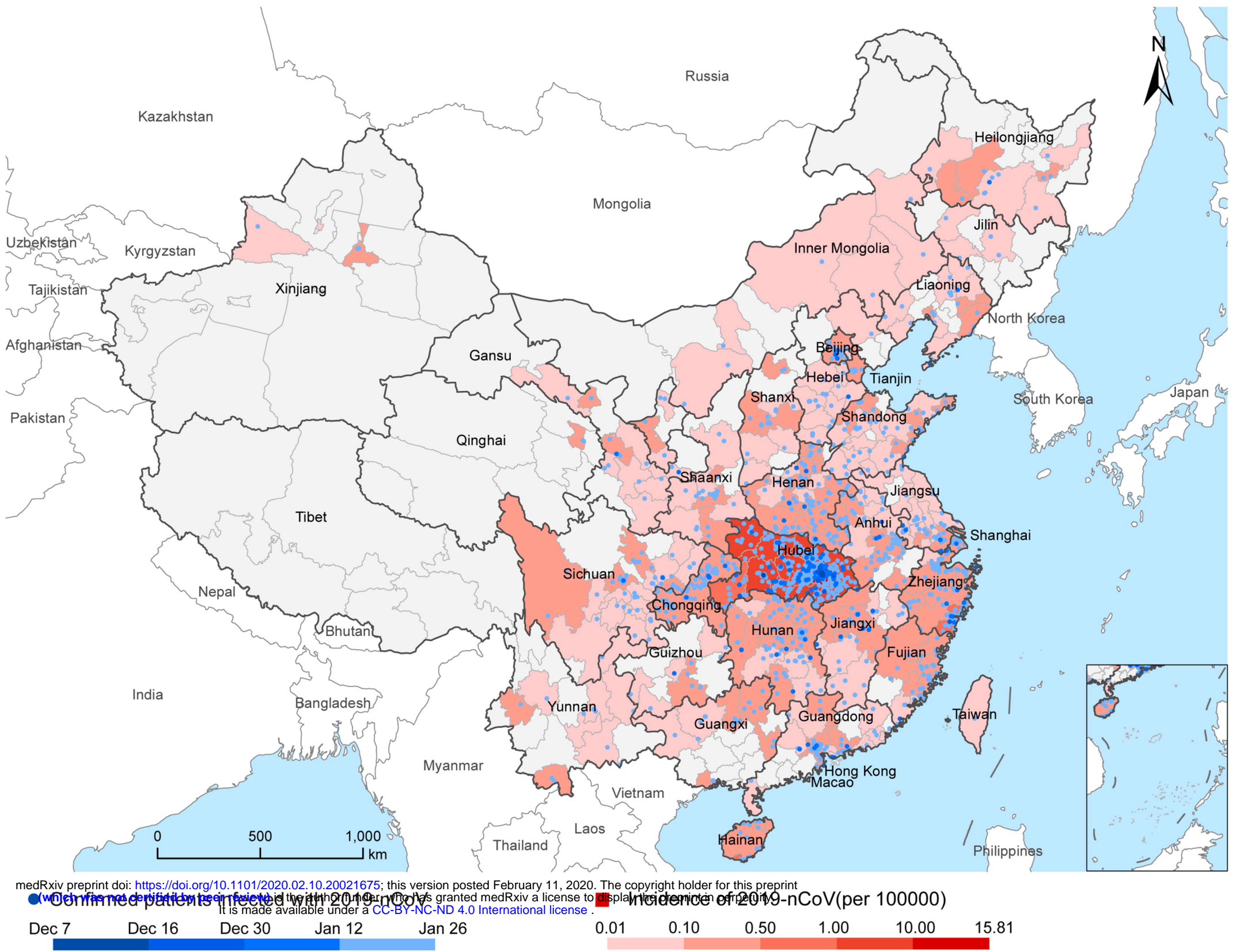
§ These cases all had symptom onset dates < 14 days before 26 Jan. 2020. CFR is calculated for this group by varying  $d$  from 5 to 8 days, rather than from 10 to 14 days (see methods).



**Table 2. Estimates of  $R_0$  for different settings of reporting rate before 1 January 2020, incubation period and infectious period. Only confirmed patients are used for estimation. The time window used for estimation is from 25 December 2010 to 12 January 2020**

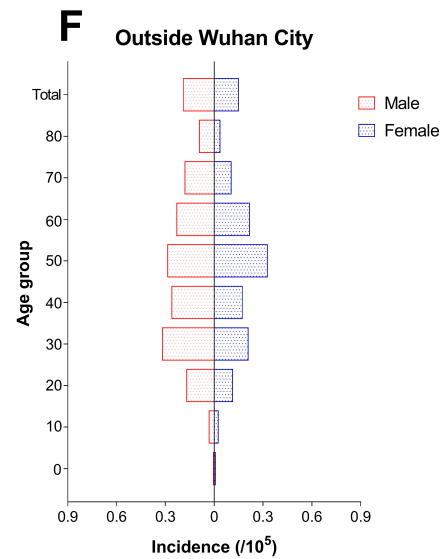
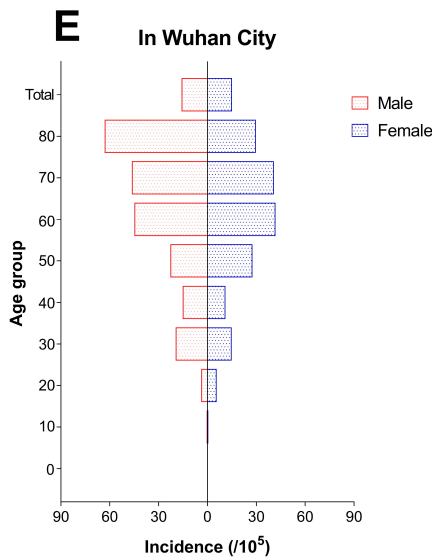
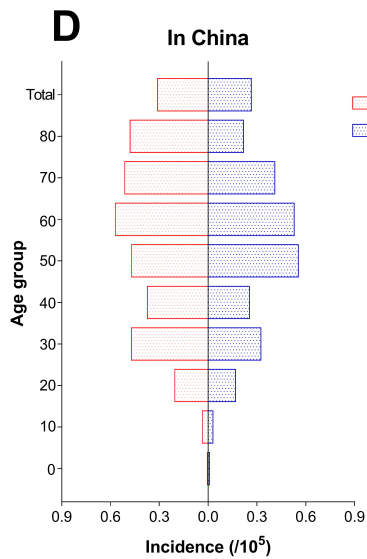
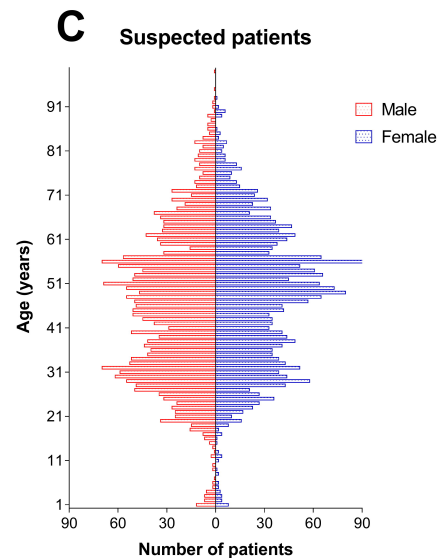
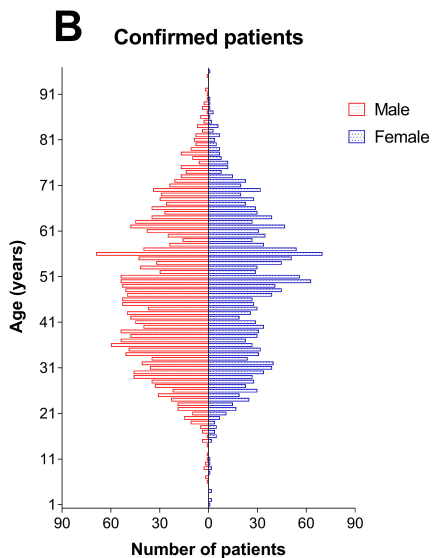
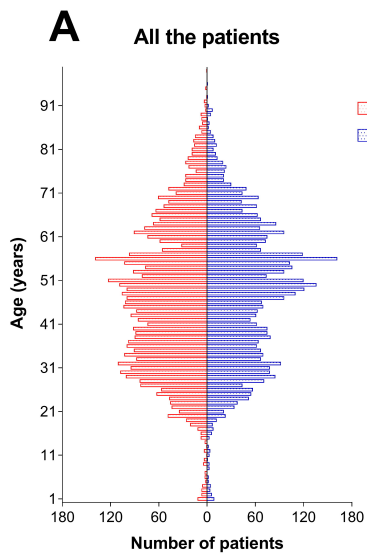
Reporting rate	Incubation period	Infectious period	$R_0$	Standard error	95% CI
100%	Short	Short	2.805	0.111	(2.595, 3.031)
		Median	3.204	0.126	(2.966, 3.461)
		Long	3.643	0.143	(3.373, 3.934)
	Median	Short	3.308	0.122	(3.077, 3.557)
		Median	3.772	0.139	(3.510, 4.054)
		Long	4.281	0.156	(3.985, 4.598)
	Long	Short	3.737	0.134	(3.483, 4.010)
		Median	4.255	0.152	(3.967, 4.563)
		Long	4.819	0.171	(4.495, 5.167)
50%	Short	Short	2.511	0.100	(2.324, 2.714)
		Median	2.790	0.110	(2.582, 3.014)
		Long	3.081	0.121	(2.852, 3.328)
	Median	Short	2.930	0.109	(2.725, 3.152)
		Median	3.249	0.120	(3.022, 3.494)
		Long	3.581	0.132	(3.331, 3.849)
	Long	Short	3.274	0.119	(3.049, 3.515)
		Median	3.621	0.131	(3.373, 3.887)
		Long	3.980	0.144	(3.708, 4.272)
30%	Short	Short	2.225	0.088	(2.059, 2.405)
		Median	2.408	0.095	(2.228, 2.601)
		Long	2.590	0.102	(2.397, 2.798)
	Median	Short	2.560	0.096	(2.378, 2.755)
		Median	2.763	0.103	(2.568, 2.973)
		Long	2.966	0.111	(2.757, 3.191)
	Long	Short	2.822	0.104	(2.626, 3.034)
		Median	3.037	0.112	(2.826, 3.264)
		Long	3.251	0.119	(3.026, 3.494)

A

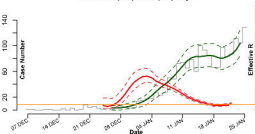


B

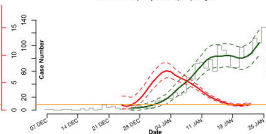




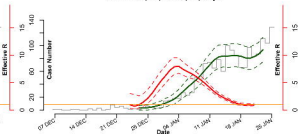
Mean Inc. (Inf.) = 3.8 (5.2) days



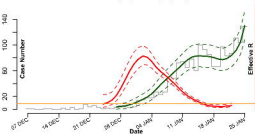
Mean Inc. (Inf.) = 3.8 (7.2) days



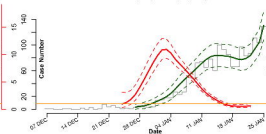
Mean Inc. (Inf.) = 3.8 (9.2) days



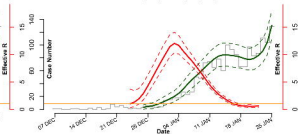
Mean Inc. (Inf.) = 5.2 (5.2) days



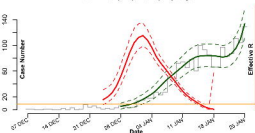
Mean Inc. (Inf.) = 5.2 (7.2) days



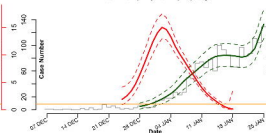
Mean Inc. (Inf.) = 5.2 (9.2) days



Mean Inc. (Inf.) = 6.4 (5.2) days



Mean Inc. (Inf.) = 6.4 (7.2) days



Mean Inc. (Inf.) = 6.4 (9.2) days

