Epidemic doubling time of the 2019 novel coronavirus outbreak by province in mainland China

Kamalich Muniz-Rodriguez, MPH[1]; Gerardo Chowell, PhD[1]; Chi-Hin Cheung, MS; Dongyu Jia, PhD; Po-Ying Lai, MS; Yiseul Lee, MPH, Manyun Liu, MS; Sylvia K. Ofori, MPH; Kimberlyn M. Roosa, MPH; Lone Simonsen, PhD; Cecile G. Viboud, PhD; Isaac Chun-Hai Fung, PhD

[1] These authors contribute equally as co-first authors

Author affiliations: Georgia Southern University, Georgia, USA (K. Muniz-Rodriguez, D. Jia, M. Liu, S. K. Ofori, I. C.-H. Fung); Georgia State University (G. Chowell, Y. Lee, K. M. Roosa); Independent researcher (C.-H. Cheung); Boston University (P.-Y. Lai), The George Washington University (L. Simonsen), The National Institutes of Health (G. Chowell, C. Viboud)

Email addresses: km11200@georgiasouthern.edu (K. Muniz-Rodriguez); gchowell@gsu.edu (G. Chowell) westerpants@gmail.com (C.-H. Cheung); djia@georgiasouthern.edu (D. Jia); pylai@bu.edu (P.-Y. Lai); ylee97@student.gsu.edu (Y. Lee); ml16842@georgiasouthern.edu (M. Liu); so01935@georgiasouthern.edu (S. K. Ofori); kroosa1@student.gsu.edu (K. M. Roosa); lone@gwu.edu (L. Simonsen); viboudc@mail.nih.gov (C. G. Viboud); cfung@georgiasouthern.edu (I. C.-H. Fung)

Please address correspondence to Isaac Chun-Hai Fung, PhD, Department of Biostatistics, Epidemiology and Environmental Health Sciences, Jiann-Ping Hsu College of Public Health, Georgia Southern University, Statesboro, GA 30460, USA. +1 912 478 5079. Email: cfung@georgiasouthern.edu

Genre: Research Letter (Max. 800 words)

Current word count ~ 645.

Abstract: n=50; Word limit = 50

We analyzed the epidemic doubling time of the 2019-nCoV outbreak by province in mainland China. Mean doubling time ranged from 1.0 to 3.3 days, being 2.4 days for Hubei (January 20-February 2, 2020). Trajectory of increasing doubling time by province indicated social distancing measures slowed the epidemic with some success.

To the editor:

Our ability to estimate basic reproduction numbers for novel infectious diseases is hindered by the dearth of information about their epidemiological characteristics and transmission mechanisms (1). More informative metrics could synthesize real-time information about the extent to which the epidemic is expanding over time. Such metrics would be particularly useful if they rely on minimal data of the outbreak's trajectory (2).

Epidemic doubling times characterize the sequence of times at which the cumulative incidence doubled (3). Here we analyze the evolution of the doubling times and the number of times the cumulative incidence doubles, associated with the novel coronavirus (2019-nCoV) outbreak by province in mainland China (4), from January 20 (when provinces outside Hubei started reporting cases) through February 2, 2020. See Technical Appendix for a sensitivity analysis applied to data from December 31, 2019 through February 2, 2020. If an epidemic is growing exponentially with a constant growth rate r, the doubling time should remain constant, where doubling time = (ln 2) / r. An increase in doubling time could mean the epidemic has slowed down, assuming that the underlying reporting rate remained unchanged (see Technical Appendix and Figure S1).

Cumulative incidence data from December 31, 2019 through February 2, 2020 were retrieved from official webpages of provincial health commissions, and that of the National Health Commission of China (5). They were double-checked against the reported numbers of the provinces according to Centre for Health Protection, Hong Kong, if available (6). Whenever discrepancies arose, the respective provincial government sources were deemed authoritative. Tibet was excluded from further analysis because there was only one case as of February 2, 2020 and thus doubling time could not be calculated. All data analyzed are publicly available.

From January 20 through February 2, 2020, the mean doubling time of the cumulative incidence ranged from 1.0 day (Hunan and Henan) to 3.3 days (Hainan) (Figure 1A). In Hubei, it was estimated as

2.4 days. The cumulative incidence of Hubei doubled 5 times (Figure 1B). Provinces with the cumulative incidence doubled ≥5 times, and mean doubling time <2d included Chongqing, Fujian, Heilongjiang, Henan, Hunan, Jiangxi, Shandong, Shanghai, Shanxi, Sichuan, Yunnan, and Zhejiang. These provinces experienced a faster and consistent epidemic growth (Figures 1 and S2).

The aggregate cumulative incidence of all non-Hubei provinces increased over time (Figure S3) and therefore suggested a sub-exponential growth of the epidemic outside Hubei. The gradual piece-meal increase in doubling time could be explained by the practice of self-quarantine since the Chinese New Year and the different levels of intra-and-inter-provincial travel restrictions imposed across China since the travel quarantine of Wuhan (imposed on Jan 23, 2020) (7).

The limitations of our study included the incompleteness of the cumulative incidence data as reported by mainland Chinese authorities. One potential reason for underreporting is underdiagnosis, due to the lack of diagnostic tests, healthcare workers and other resources. Differential underreporting across provinces could have biased the data. However, as long as the rate of reporting remains constant over time within the same province, the calculation of doubling times remains reliable. However, increased awareness and increased availability of diagnostic tests might have improved the reporting rate over time. This might artificially shorten the doubling time. Nevertheless, apart for Hubei, for the majority of mainland China, cases were only reported since January 20, 2020. It was when the Chinese authorities openly acknowledged the seriousness of the outbreak. Therefore, the bias due to increased awareness might be small to negligible.

Conclusions

We analyzed the epidemic doubling time of the 2019 novel coronavirus outbreak by province in mainland China. The mean doubling time of cumulative incidence in Hubei was 2.4 days (January 20 through February 2, 2020) but the mean doubling time of Henan, Hunan, and Shandong were the lowest.

Trajectory of increasing doubling time by province indicated social distancing measures adopted in China slowed the epidemic with some success.

First author(s) biography

Kamalich Muniz-Rodriguez, MPH, is a doctoral student at the Jiann-Ping Hsu College of Public Health, Georgia Southern University. Her research interests include infectious disease epidemiology, digital epidemiology and disaster epidemiology.

Gerardo Chowell, PhD, is Professor of Epidemiology and Biostatistics, and Chair of the Department of Population Health Sciences at Georgia State University School of Public Health. As a mathematical epidemiologist, Prof Chowell studies the transmission dynamics of emerging infectious diseases, such as Ebola, MERS and SARS.

Acknowledgement

GC acknowledges support from NSF grant 1414374 as part of the joint NSF-NIH-USDA Ecology and Evolution of Infectious Diseases program. ICHF acknowledges salary support from the National Center for Emerging and Zoonotic Infectious Diseases, Centers for Disease Control and Prevention (19IPA1908208). This article is not part of ICHF's CDC-sponsored projects.

Disclaimer

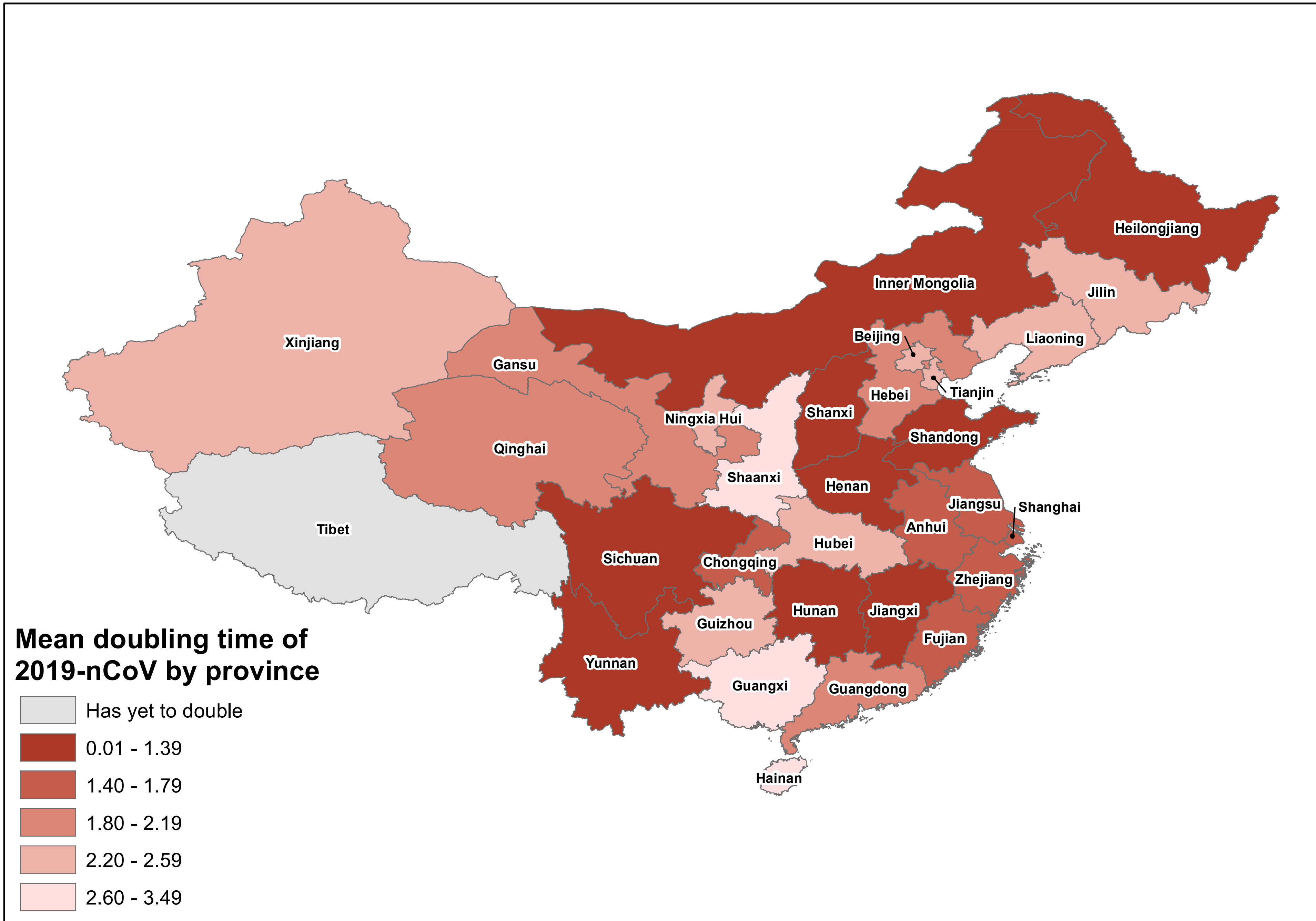
This article does not represent the official positions of the Centers for Disease Control and Prevention, the National Institutes of Health, or the United States Government.

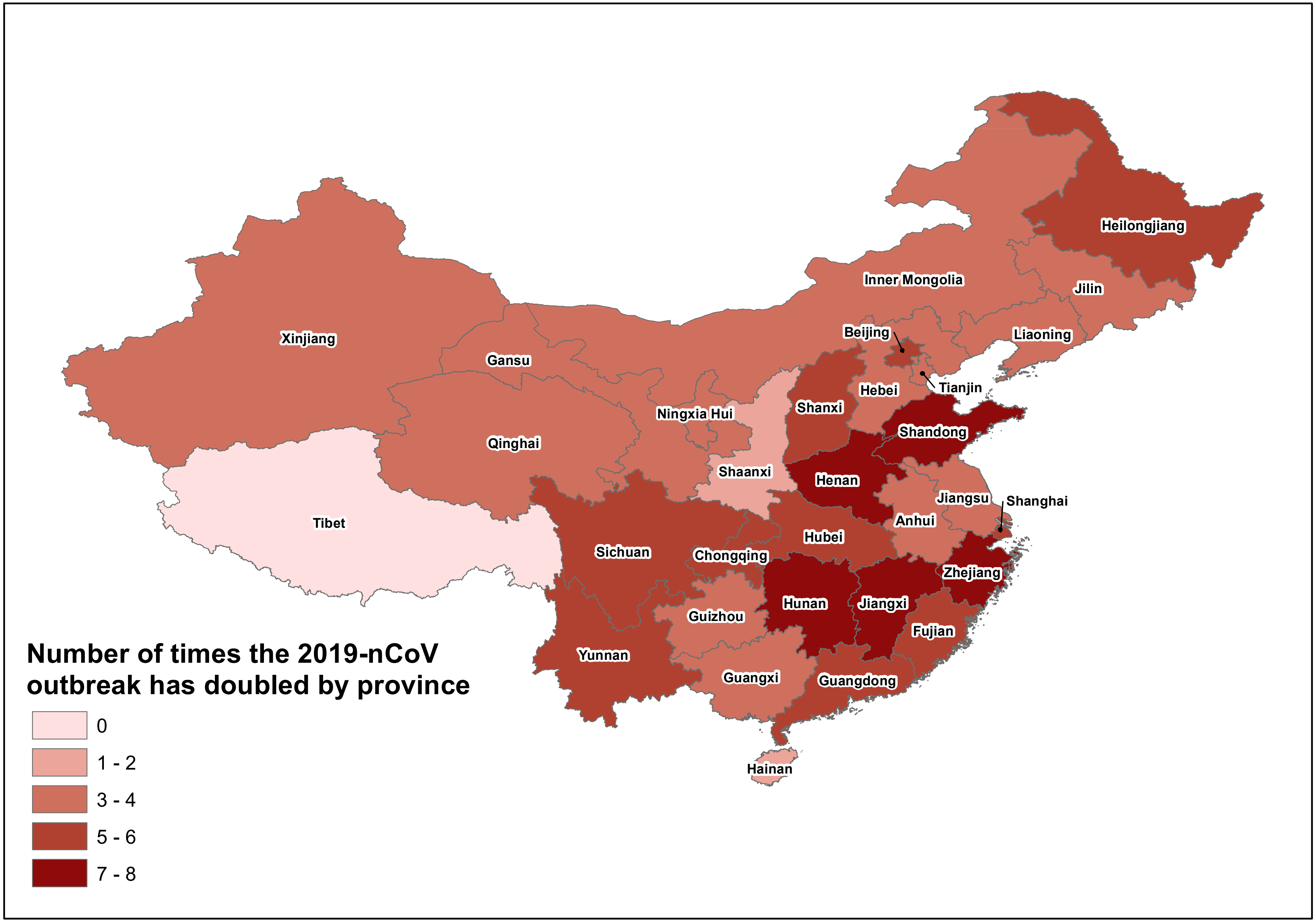
References

1. Anderson RM, May RM. Infectious diseases of humans. Oxford: Oxford University Press; 1991.
2. Drake JM, Bakach I, Just MR, O'Regan SM, Gambhir M, Fung IC-H. Transmission Models of Historical Ebola Outbreaks. Emerging Infectious Disease journal. 2015;21(8):1447.
3. Vynnycky E, White RG. An Introduction to Infectious Disease Modelling. Oxford: Oxford University Press; 2010.
4. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. New England Journal of Medicine. 2020.
5. National Health Commission of the People's Republic of China. 2020 [cited Feb 2, 2020]; Available from: http://www.nhc.gov.cn/
6. Centre for Health Protection, Department of Health, The Government for the Hong Kong Special Administrative Region. 2020 [cited Feb 2, 2020]; Available from: https://www.chp.gov.hk/en/index.html
7. Du Z, Wang L, Cauchemez S, Xu X, Wang X, Cowling BJ, et al. Risk of 2019 novel coronavirus importations throughout China prior to the Wuhan quarantine. medRxiv. 2020:2020.01.28.20019299.

Figure legend

Figure 1. The mean doubling time (Panel A) and the number of times the 2019-nCoV outbreak cumulative incidence has doubled (Panel B) by province in mainland China, from January 20 through February 2, 2020.


Heilongjiang
Inner Mongolia
Jilin
Liaoning
Beijing
Tianjin
Hebei
Xinjiang
Gansu
Ningxia Hui
Shanxi
Shandong
Qinghai
Shaanxi
Henan
Jiangsu
Shanghai
Anhui
Tibet
Sichuan
Chongqing
Hubei
Zhejiang
Hunan
Jiangxi
Guizhou
Fujian
Yunnan
Guangxi
Guangdong
Hainan
Mean doubling time of 2019-nCoV by province
Has yet to double
0.01 - 1.39
1.40 - 1.79
1.80 - 2.19
2.20 - 2.59
2.60 - 3.49


Number of times the 2019-nCoV outbreak has doubled by province
0
1 - 2
3 - 4
5 - 6
7 - 8
Heilongjiang
Inner Mongolia
Jilin
Liaoning
Beijing
Tianjin
Hebei
Xinjiang
Gansu
Ningxia Hui
Shanxi
Shandong
Qinghai
Shaanxi
Henan
Jiangsu
Shanghai
Anhui
Tibet
Hubei
Sichuan
Chongqing
Zhejiang
Hunan
Jiangxi
Guizhou
Fujian
Yunnan
Guangxi
Guangdong
Hainan

Epidemic doubling time of the 2019 novel coronavirus outbreak by province in mainland China

Technical Appendix

Kamalich Muniz-Rodriguez, MPH[1]; Gerardo Chowell, PhD[1]; Chi-Hin Cheung, MS; Dongyu Jia, PhD; Po-Ying Lai, MS; Yiseul Lee, MPH, Manyun Liu, MS; Sylvia K. Ofori, MPH; Kimberlyn M. Roosa, MPH; Lone Simonsen, PhD; Cecile G. Viboud, PhD; Isaac Chun-Hai Fung, PhD

[1] These authors contribute equally as co-first authors

Author affiliations: Georgia Southern University, Georgia, USA (K. Muniz-Rodriguez, D. Jia, M. Liu, S. K. Ofori, I. C.-H. Fung); Georgia State University (G. Chowell, Y. Lee, K. M. Roosa); Independent researcher (C.-H. Cheung); Boston University (P.-Y. Lai), The George Washington University (L. Simonsen), The National Institutes of Health (G. Chowell, C. G. Viboud)

Please address correspondence to Isaac Chun-Hai Fung, PhD, Department of Biostatistics, Epidemiology and Environmental Health Sciences, Jiann-Ping Hsu College of Public Health, Georgia Southern University, Statesboro, GA 30460, USA. +1 912 478 5079. Email: cfung@georgiasouthern.edu

Technical appendix

Additional information on our motivation, scope and methods

Motivation. R_0 is a widely used indicator of transmission potential in a totally susceptible population and is driven by the average contact rate and the mean infectious period of the disease (1). Yet, it only characterizes transmission potential at the onset of the epidemic and varies geographically for a given infectious disease according to local healthcare provision, outbreak response, as well as socioeconomic and cultural factors. Furthermore, estimating R_0 requires information about the natural history of the infectious disease. Thus, our ability to estimate reproduction numbers for novel infectious diseases is hindered by the dearth of information about their epidemiological characteristics and transmission mechanisms. More informative metrics could synthesize real-time information about the extent to which the epidemic is expanding over time. Such metrics would be particularly useful if they rely on minimal data of the outbreak's trajectory.

Scope and definition. We restricted our analysis to mainland China in this paper. A 'province' herein encompasses three different types of political sub-divisions of mainland China, namely, a province, a directly administered municipality (Beijing, Chongqing, Shanghai, and Tianjin) and an autonomous region (Guangxi, Inner Mongolia, Ningxia, Tibet, and Xinjiang). Our analysis does not include Hong Kong Special Administrative Region and Macau Special Administrative Region, which are under effective rule of the People's Republic of China through the so-called 'One Country, Two Systems' political arrangements. Likewise, our analysis does not include Taiwan, which is *de facto* governed by a different government (the Republic of China).

Data apart from epidemic data. Provincial demographic, transportation and socioeconomic data were obtained from the National Bureau of Statistics of China (2) and other sources (see Table S2).

Doubling time calculation. As the epidemic grows, the times at which cumulative incidence doubles are given by t_{d_i} such that $2C(t_{d_i}) = C(t_{d_{i+1}})$ where $t_{d_o} = 0$, $C(t_{d_o}) = C_0$, and $i = 0,1,2,3, \ldots, n_d$ where n_d is the total number of times cumulative incidence doubles (Figure S1). The actual sequence of "doubling times" are defined as follows (Figure S1):

$$d_j = \Delta t_{d_j} = t_{d_j} - t_{d_{j-1}} \text{ where } j = 1,2,3, \ldots, n_d.$$

Doubling time calculation was conducted using MATLAB R2019b (Mathworks, Natick, MA). Multiple linear regression analyses were conducted using R version 3.6.2 (R Core Team). Significance level was a priori decided to be $\alpha = 0.05$.

Additional information on our results and discussion

Demographic, transportation and socioeconomic factors. We performed multiple linear regression models with the latest doubling time, mean doubling time and the slope of the doubling time over the number of times the cumulative incidence doubles as the dependent variables, respectively. We included population density, average temperature in January, average household size, subnational Human Development Index in all models. We included passenger traffic and provincial capital's distance from Wuhan, for railway (models group A) and highway (models group B) respectively. However, none of the independent variables were found statistically significantly ($p > 0.05$) associated with any of the dependent variables (Table S2).

Sensitivity analysis

We performed sensitivity analysis by expanding our data analysis to the data since January 31, 2019, when Hubei first reported a cluster of pneumonia cases with unexplained etiology that turned out to be 2019-nCoV. The only difference between the sensitivity analysis and the main analysis is the inclusion of Hubei data from January 31, 2019 through January 19, 2020, because all other provinces started to report cases on January 20, 2020. The only differences in results were found for Hubei, with the mean doubling time being 3.85 (Figures S4, S6), and the cumulative incidence in Hubei doubled 8 times from January 31, 2019 through February 2, 2020 (Figures S5, S6). The first doubling time of Hubei (Figure S5) was high, reflecting that real-time data was unavailable before mid-January. It was only by January 17, 2020 onwards when data reporting become increasingly transparent and timely.

In our sensitivity analysis, we performed the same multiple regression models previously described, with the mean doubling time, and the slope of the doubling time over the number of times the cumulative incidence doubles as dependent variables. We included population density, average temperature in January, average household size, subnational Human Development Index in all models. We included passenger traffic and provincial capital's distance from Wuhan, for railway and highway respectively. However, none of the independent variables were found statistically significantly ($p > 0.05$) associated with the three dependent variables (results not shown).

Table S1. List of provinces, directly administered municipalities and autonomous regions in mainland China, as displayed in Figures S2, S3, S4 and S5.

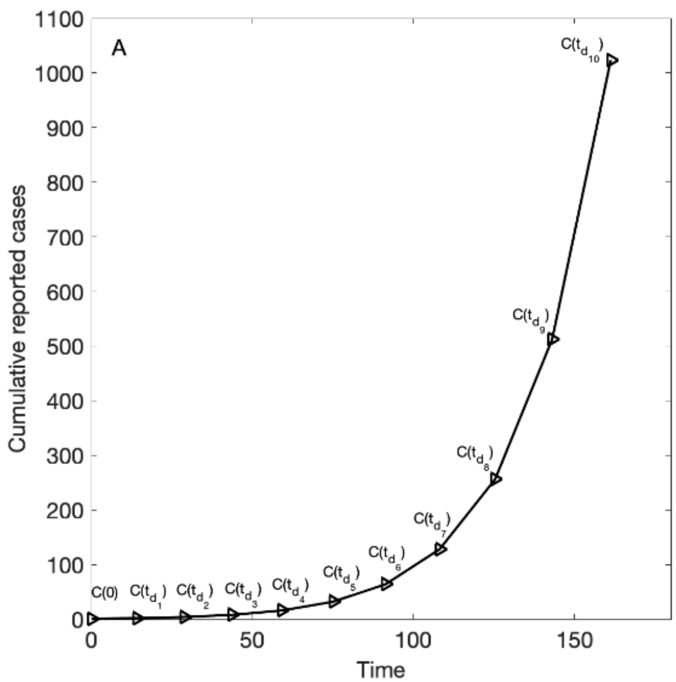
Numbering in Figures 2 and 3	Name
1	Hubei
2	Aggregate of the entire mainland China, except Hubei
3	Anhui
4	Beijing
5	Chongqing
6	Fujian
7	Gansu
8	Guangdong
9	Guangxi
10	Guizhou
11	Hainan
12	Heilongjiang
13	Henan
14	Hebei
15	Hunan
16	Inner Mongolia
17	Jiangsu
18	Jiangxi
19	Jilin
20	Liaoning
21	Ningxia
22	Qinghai
23	Shaanxi
24	Shanxi
25	Shandong
26	Shanghai
27	Sichuan
28	Tianjin
29	Tibet
30	Xinjiang
31	Yunnan
32	Zhejiang

Table S2. The demographic and transportation variables that had been considered for linear regression models:

	Regression models		
Variable	**Group A**	**Group B**	**Reference**
Railway Passenger Traffic (10000 persons)	x		(2)
Highway Passenger Traffic (10000 persons)		x	(2)
Provincial Capital's Distance from Wuhan (Railway) (km)	x		(2)
Provincial Capital's Distance from Wuhan (Highway) (km)		x	(2)
Population Density (10000 person per km^2)	x	x	(3)
Average Temperature for January (°C)	x	x	(4)
Average Household size (2018)	x	x	(5)
(Subnational) Human development index (HDI)	x	x	(6)

Notes: The following data have been studied in preliminary studies but not in our models presented herein: Resident population (year-end) (10000 persons), Urban population (year-end) (10000 persons), Rural population (year-end) (10000 persons), Passenger Traffic (10000 persons), Passenger-Kilometers (100 million passenger-km), Passenger-Kilometers of Railways (100 million passenger-km), Passenger-Kilometers of Highway (100 million passenger-km) (2).

Figure S1. Illustration of the concept of doubling time using a hypothetical data set. Panel A presents the exponential increase of the cumulative reported cases over time and its number at each time when the case number doubled. Panel B presents the doubling time at each time when the cumulative incidence doubles.



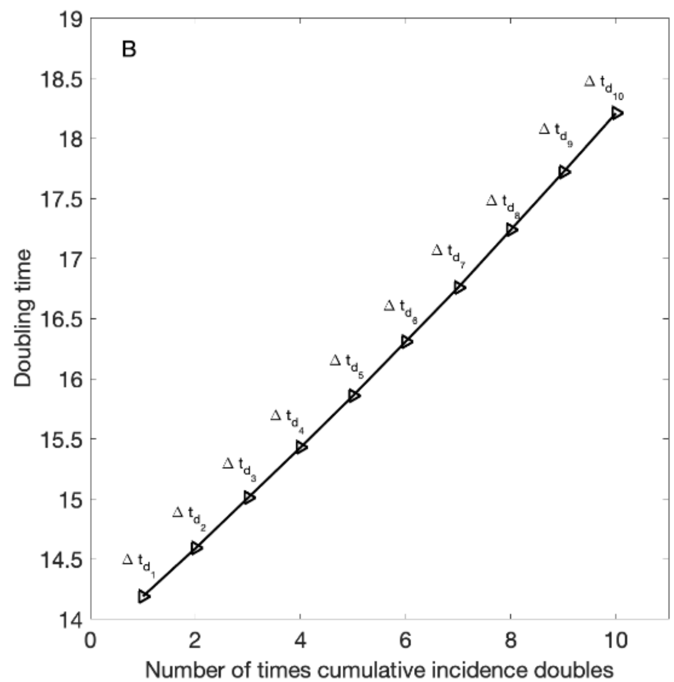


Figure S2. The mean doubling time (days) and the number of times the reported cumulative incidence doubles by province within mainland China from January 20, 2020 through February 2, 2020. Each point represents a province except for Point 2 that is the aggregate of all other provinces in mainland China. Point 1 is Hubei. Point 28 (Tibet) is not available, because there is only 1 confirmed case in Tibet as of February 2, 2020. For others, please refer to Table S1 in Supplementary Materials.

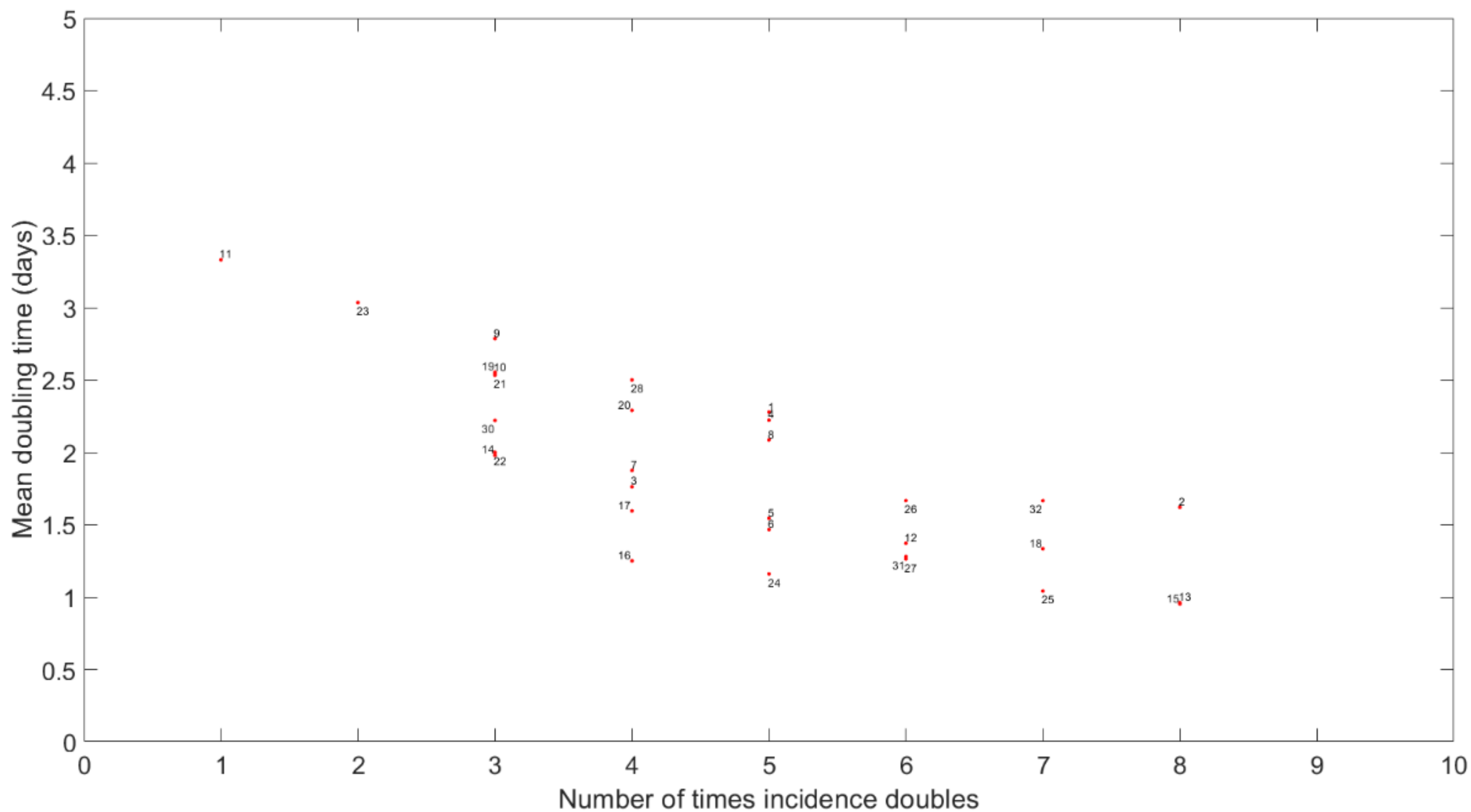


Figure S3. The doubling time (days) each time the reported cumulative incidence doubles and the number of times the reported cumulative incidence doubles by province within mainland China from January 20, 2020 through February 2, 2020. Each point represents a province except for Point 2 that is the aggregate of all other provinces in mainland China. Point 1 is Hubei. Point 28 (Tibet) is not available, because there is only 1 confirmed case in Tibet as of February 2, 2020. For others, please refer to Table S1 in Supplementary Materials.

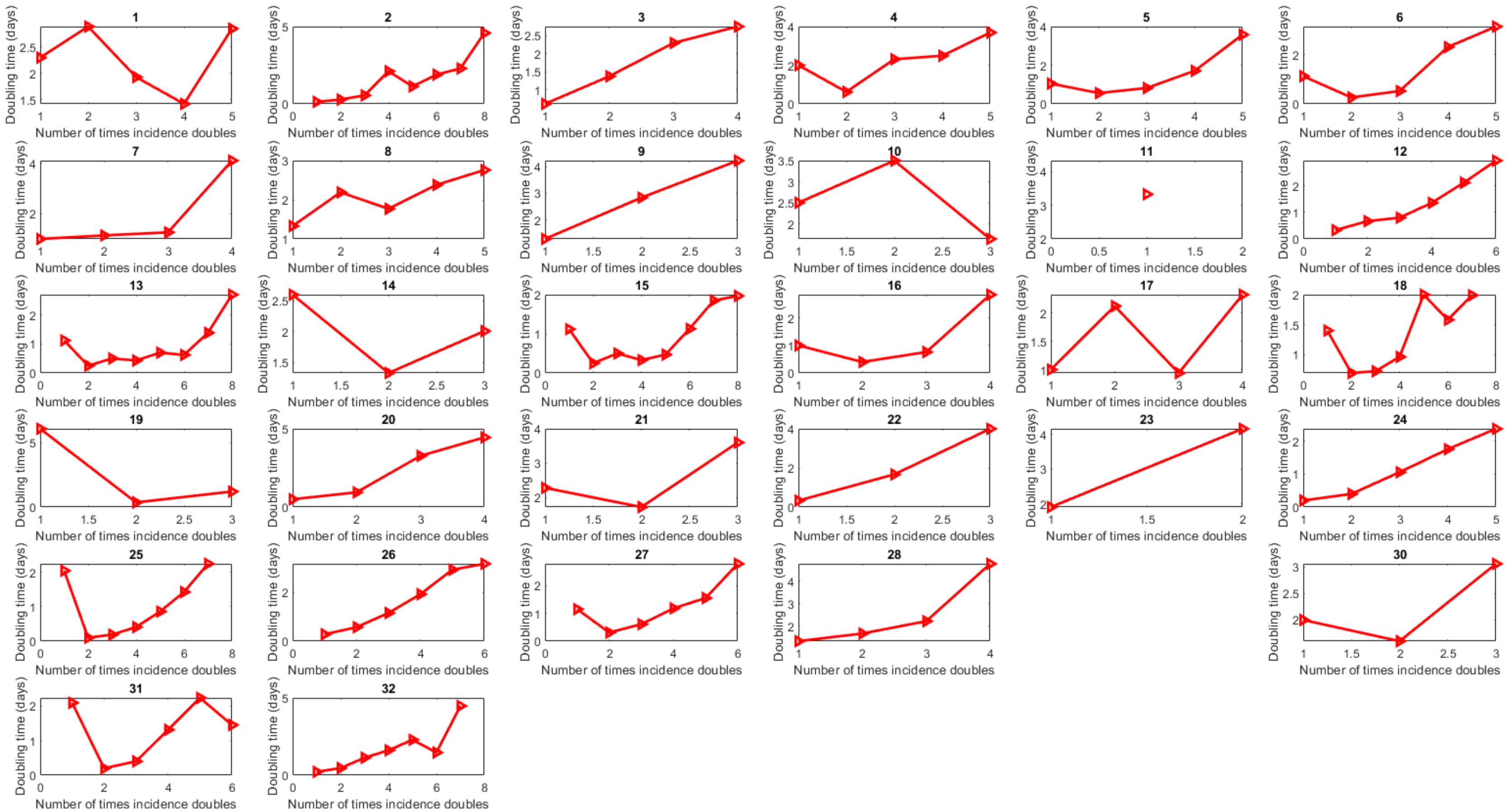


Figure S4. Sensitivity analysis: The mean doubling time (days) and the number of times the reported cumulative incidence doubles by province within mainland China from December 31, 2019 through February 2, 2020. Each point represents a province except for Point 2 that is the aggregate of all other provinces in mainland China. Point 1 is Hubei. Point 28 (Tibet) is not available, because there is only 1 confirmed case in Tibet as of February 2, 2020. For others, please refer to Table S1 in Supplementary Materials.

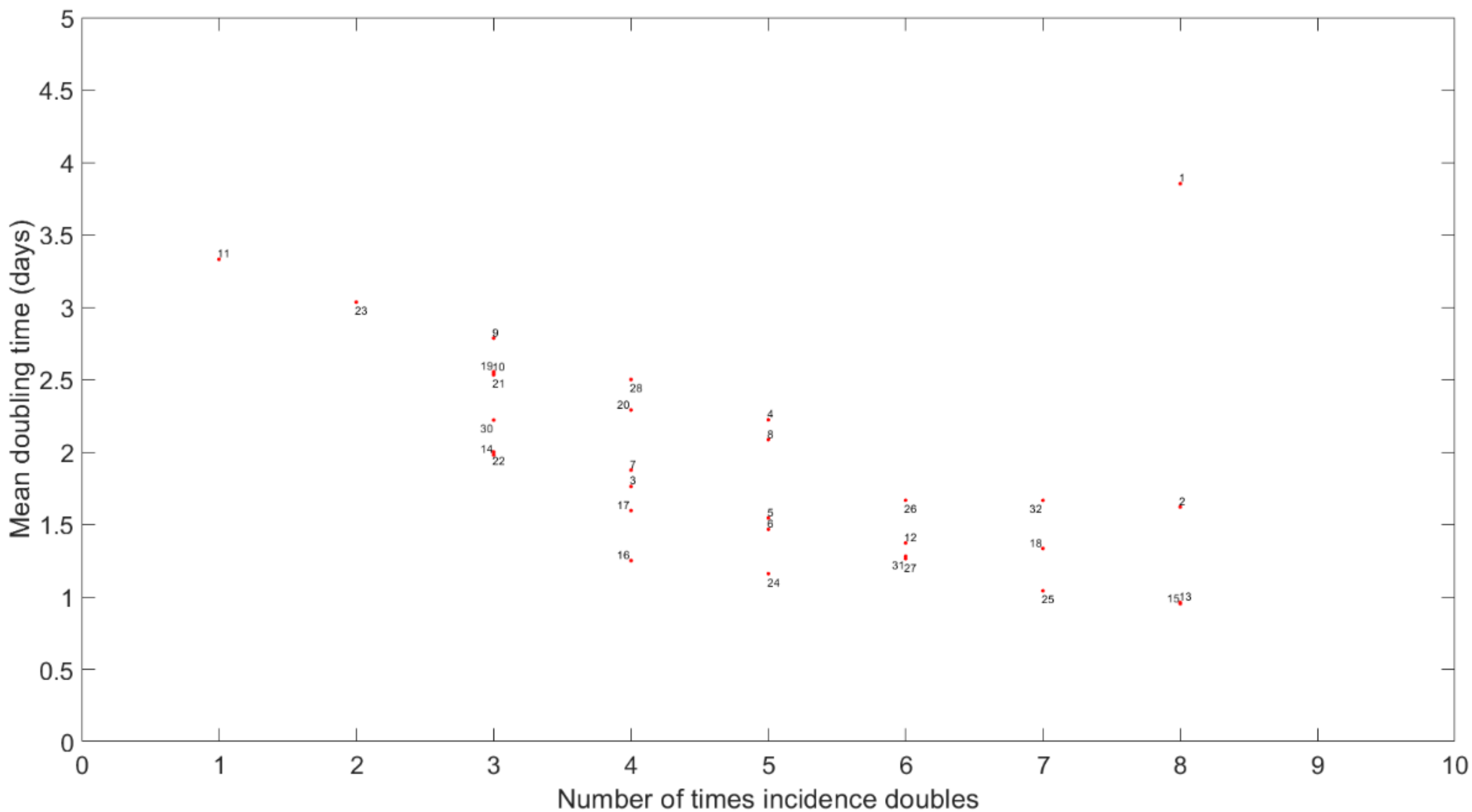


Figure S5. Sensitivity analysis: The doubling time (days) each time the reported cumulative incidence doubles and the number of times the reported cumulative incidence doubles by province within mainland China from December 31, 2019 through February 2, 2020. Each point represents a province except for Point 2 that is the aggregate of all other provinces in mainland China. Point 1 is Hubei. Point 28 (Tibet) is not available, because there is only 1 confirmed case in Tibet as of February 2, 2020. For others, please refer to Table S1 in Supplementary Materials.

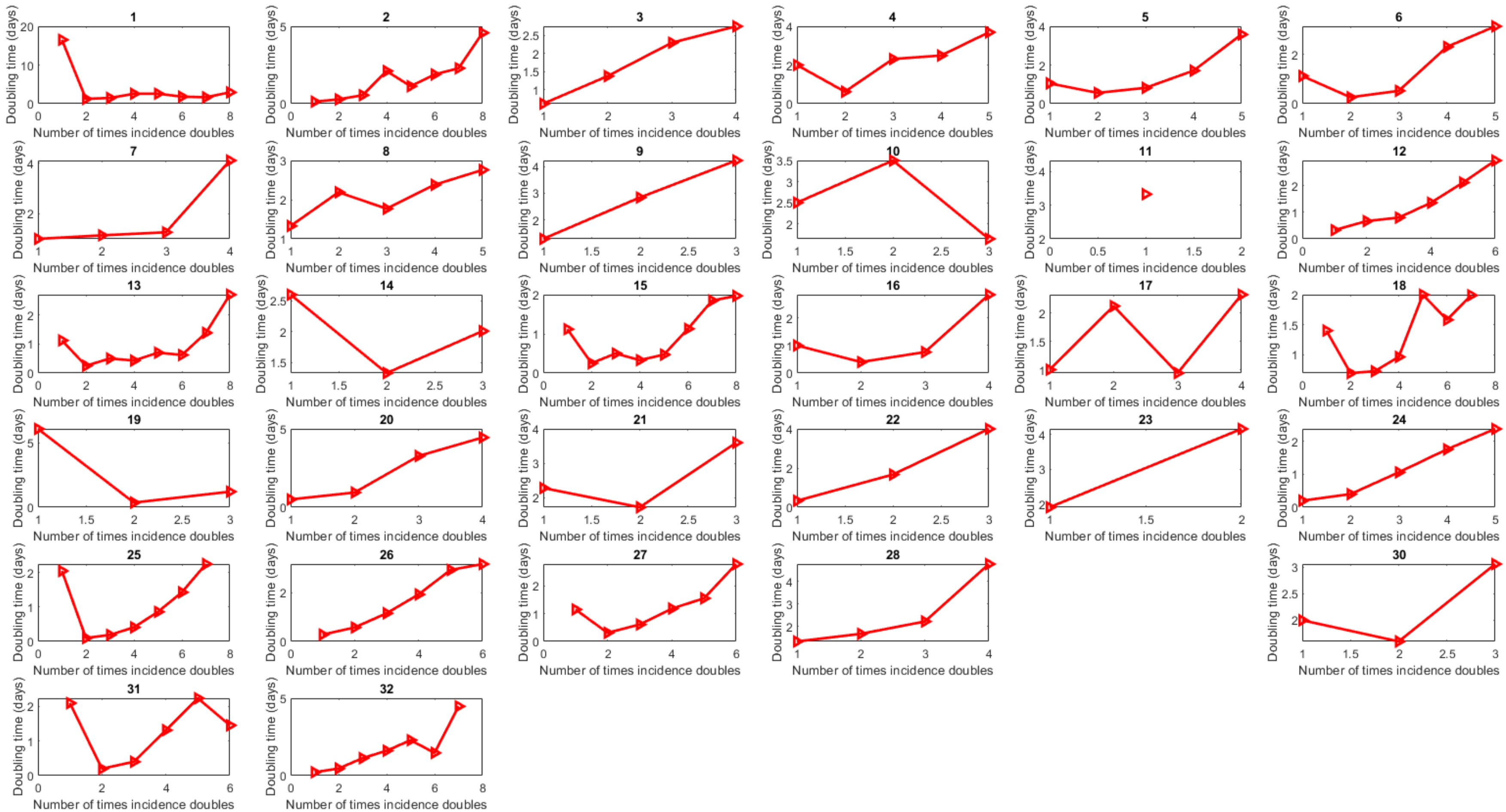


Figure S6. Sensitivity analysis: The mean doubling time of 2019-nCoV by province in mainland China, from December 31, 2019 through February 2, 2020.

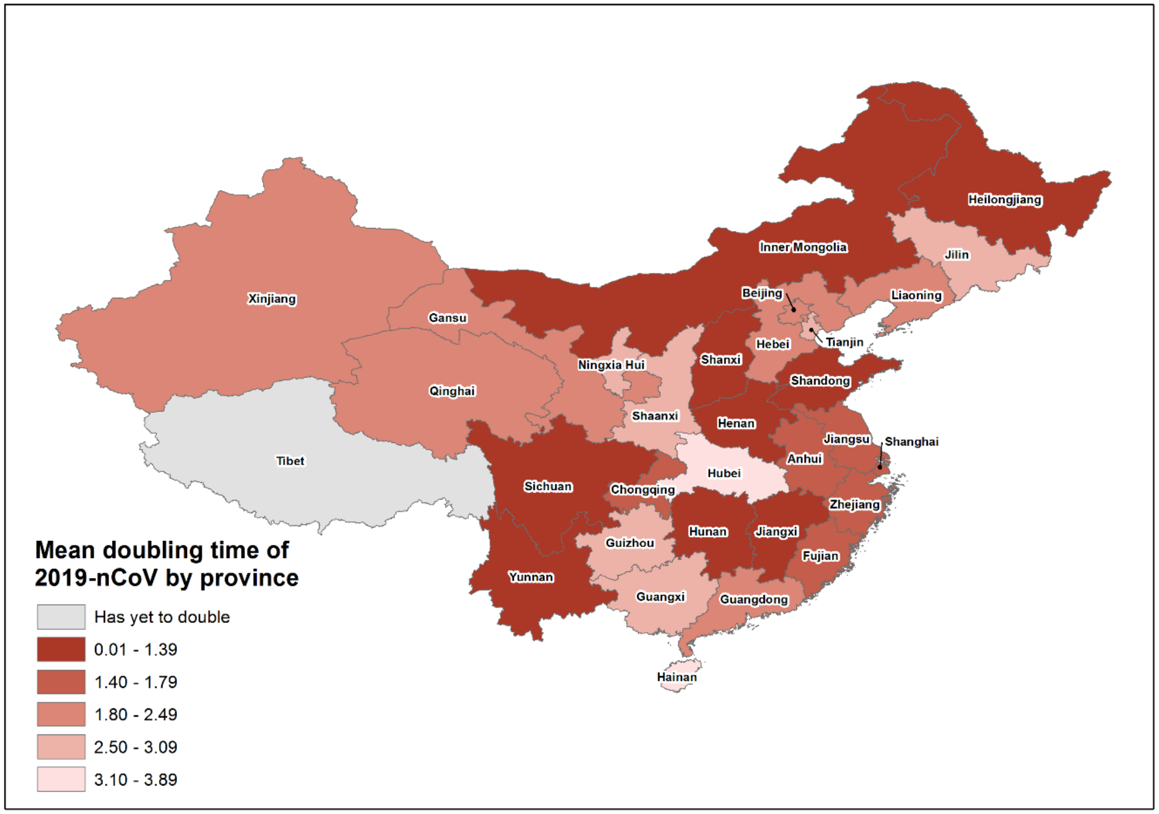
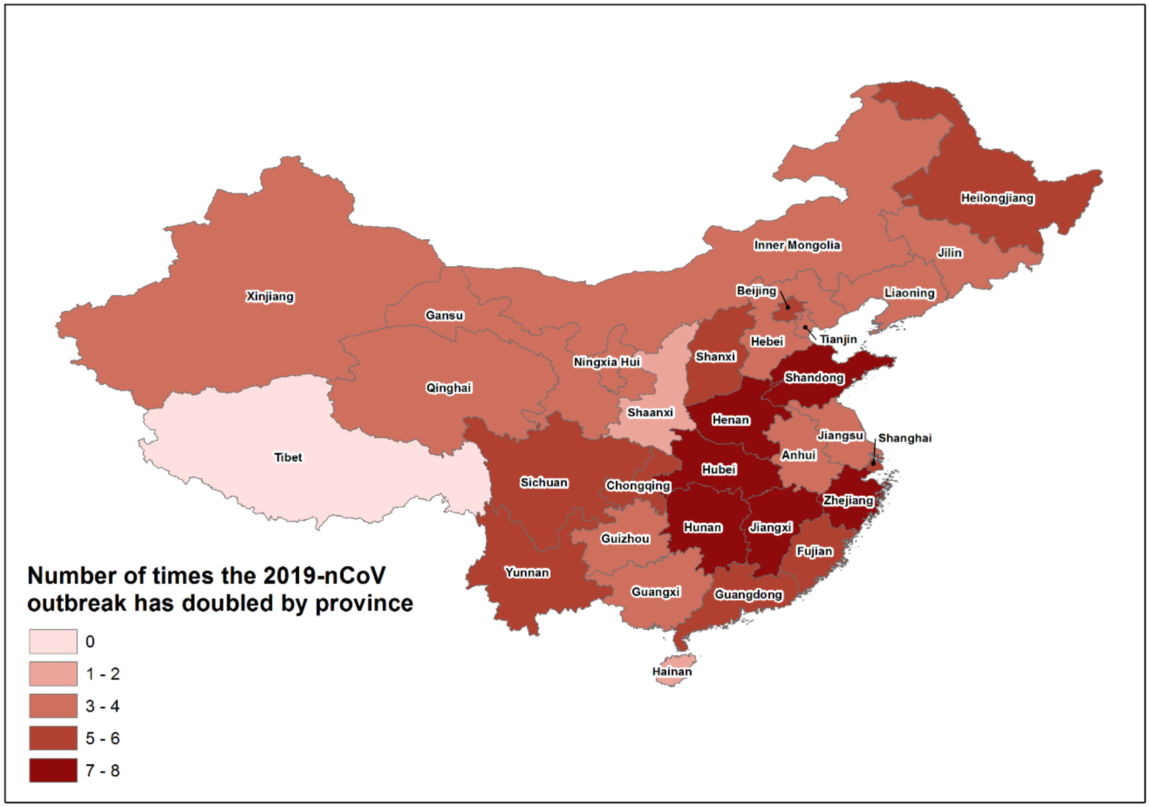


Figure S7. Sensitivity analysis: The number of times the 2019-nCoV outbreak has doubled by province in mainland China, from December 31, 2019 through February 2, 2020.



Authors' contributions

Project management: Dr. Gerardo Chowell, Dr. Isaac Chun-Hai Fung and Ms. Kamalich Muniz-Rodriguez

Manuscript writing: Dr. Isaac Chun-Hai Fung and Dr. Gerardo Chowell

Manuscript editing and data interpretation: Ms. Kamalich Muniz-Rodriguez, Dr. Gerardo Chowell, Dr. Isaac Chun-Hai Fung, Dr. Lone Simonsen, and Dr. Cecile Viboud

MATLAB code and Figure S1: Dr. Gerardo Chowell

Doubling time calculation using MATLAB and Figures S2, S3, S4 and S5: Ms. Kamalich Muniz-Rodriguez, Dr. Gerardo Chowell and Dr. Isaac Chun-Hai Fung

Statistical analysis in R: Dr. Isaac Chun-Hai Fung

Data management and quality check of epidemic data entry: Ms. Kamalich Muniz-Rodriguez

Entry of epidemic data for countries and territories outside mainland China (including Hong Kong, Macao and Taiwan): Ms. Kamalich Muniz-Rodriguez and Ms. Sylvia K. Ofori

Entry of epidemic data for provinces in mainland China: Ms. Manyun Liu (from the early reports, up to Jan 24, 2020 data), Ms. Po-Ying Lai (since Jan 25, 2020 data to today), Mr. Chi-Hin Cheung (since Jan 27, 2020 data to today), and Ms. Kamalich Muniz-Rodriguez and Dr. Isaac Chun-Hai Fung (whenever there is a back-log).

Retrieval of epidemic data from official websites (downloading and archiving of China's national and provincial authorities' press releases): Ms. Manyun Liu and Dr. Dongyu Jia

Retrieval of statistical data from the official website of National Bureau of Statistics of the People's Republic of China: Mr. Chi-Hin Cheung

Retrieval of publicly available statistical data from various sources: Ms. Yiseul Lee, Dr. Isaac Chun-Hai Fung

Map creation (Figures 1, S6 and S7): Ms. Kimberlyn M. Roosa

References

1. Anderson RM, May RM. Infectious diseases of humans. Oxford: Oxford University Press; 1991.
2. Statistical data, National Bureau of Statistics of the People's Republic of China. 2020 [cited Feb 2, 2020]; Available from: http://www.stats.gov.cn/tjsj/
3. Population.City, China, Administrative Units. 2020 [cited Feb 3, 2020]; Available from: http://population.city/china/adm/
4. World Weather Online. 2020 [cited Feb 3, 2020]; Available from: https://www.worldweatheronline.com/
5. CEIC. China Population: No of Person per Household. 2020 [cited Feb 3, 2020]; Available from: https://www.ceicdata.com/en/china/population-no-of-person-per-household
6. Smits J, Permanyer I. The Subnational Human Development Database. figshare. Collection. 2019 [cited Feb 3, 2020]; Available from: https://springernature.figshare.com/collections/The_Subnational_Human_Development_Database/4353632