of an outbreak of 2019 novel coronavirus diseases (COVID-19)—China, 2020. China CDC Weekly 2020 [cited 2020 Feb 29]. http://weekly.chinacdc.cn/en/article/id/e53946e2-c6c4-41e9-9a9b-fea8db1a8f51

Address for correspondence: Nick Wilson, Department of Public Health, University of Otago Wellington, Mein St, Newtown, Wellington 6005, New Zealand; email: nick.wilson@otago.ac.nz

Serial Interval of COVID-19 among Publicly Reported Confirmed Cases

Zhanwei Du,[1] Xiaoke Xu,[1] Ye Wu,[1] Lin Wang, Benjamin J. Cowling, Lauren Ancel Meyers

Author affiliations: University of Texas at Austin, Austin, Texas, USA (Z. Du, L.A. Meyers); Dalian Minzu University, Dalian, China (X. Xu); Beijing Normal University Computational Communication Research Center, Zhuhai, China (Y. Wu); Beijing Normal University School of Journalism and Communication, Beijing, China (Y. Wu); Institut Pasteur, Paris, France (L. Wang); University of Hong Kong, Hong Kong, China (B.J. Cowling); Santa Fe Institute, Santa Fe, New Mexico, USA (L.A. Meyers)

We estimate the distribution of serial intervals for 468 confirmed cases of coronavirus disease reported in China as of February 8, 2020. The mean interval was 3.96 days (95% CI 3.53–4.39 days), SD 4.75 days (95% CI 4.46–5.07 days); 12.6% of case reports indicated presymptomatic transmission.

DOI: https://doi.org/10.3201/eid2606.200357

Key aspects of the transmission dynamics of coronavirus disease (COVID-19) remain unclear (*1*). The serial interval of COVID-19 is defined as the time duration between a primary case-patient (infector) having symptom onset and a secondary case-patient (infectee) having symptom onset (*2*). The distribution of COVID-19 serial intervals is a critical input for determining the basic reproduction number (R_0) and the extent of interventions required to control an epidemic (*3*).

To obtain reliable estimates of the serial interval, we obtained data on 468 COVID-19 transmission events reported in mainland China outside of Hubei Province during January 21–February 8, 2020. Each report consists of a probable date of symptom onset for both the infector and infectee, as well as the probable locations of infection for both case-patients. The data include only confirmed cases compiled from online reports from 18 provincial centers for disease control and prevention (https://github.com/MeyersLabUTexas/COVID-19).

Fifty-nine of the 468 reports indicate that the infectee had symptoms earlier than the infector. Thus, presymptomatic transmission might be occurring. Given these negative-valued serial intervals, COVID-19 serial intervals seem to resemble a normal distribution more than the commonly assumed gamma or Weibull distributions (*4*,*5*), which are limited to positive values (Appendix, https://wwwnc.cdc.gov/EID/article/26/7/20-0357-App1.pdf). We estimate a mean serial interval for COVID-19 of 3.96 (95% CI 3.53–4.39) days, with an SD of 4.75 (95% CI 4.46–5.07) days (Figure), which is considerably lower than reported mean serial intervals of 8.4 days for severe acute respiratory syndrome (*5*) to 14.6 days (*6*) for Middle East respiratory syndrome. The mean serial interval is slightly but not significantly longer when the index case is imported (4.06 [95% CI 3.55–4.57] days) versus locally infected (3.66 [95% CI 2.84–4.47] days), but slightly shorter when the secondary transmission occurs within the household (4.03 [95% CI 3.12–4.94] days) versus outside the household (4.56 [95% CI 3.85–5.27] days). Combining these findings with published estimates for the early exponential growth rate COVID-19 in Wuhan (*7*), we estimate an R_0 of 1.32 (95% CI 1.16–1.48) (*5*), which is lower than published estimates that assume a mean serial interval exceeding 7 days (*7*,*8*).

These estimates reflect reported symptom onset dates for 752 case-patients from 93 cities in China, who range in age from 1 to 90 years (mean 45.2 years, SD 17.21 years). Recent analyses of putative COVID-19 infector–infectee pairs from several countries have indicated average serial intervals of 4.0 days (95% CI 3.1–4.9 days; n = 28; unpub. data, H. Nishiura et al., unpub. data, https://doi.org/10.1101/2020.02.03.20019497), 4.4 days (95% CI 2.9–6.7 days, n = 21; S. Zhao et al., unpub. data, https://doi.org/10.1101/2020.02.21.20026559], and 7.5 days (95% CI 5.3–19, n = 6; *8*). Whereas none of these studies report negative serial intervals in which the infectee had symptoms before the infector, 12.6% of the serial intervals in our sample were negative.

[1]These first authors contributed equally to this article.

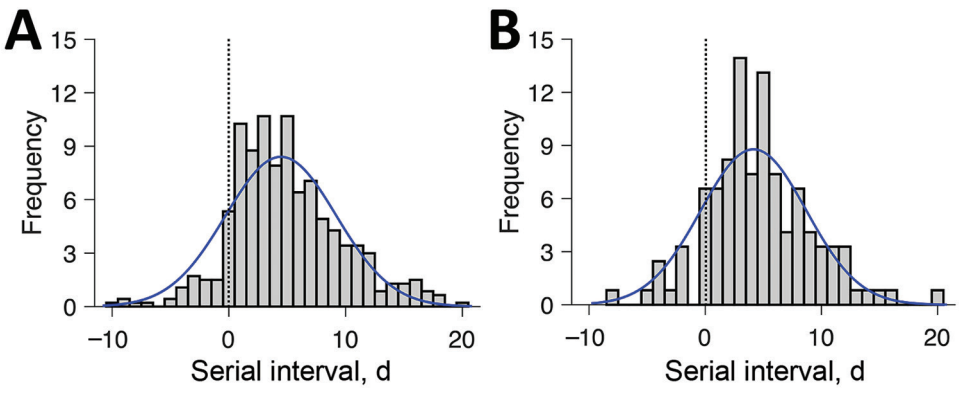


Figure. Estimated serial interval distribution for coronavirus disease (COVID-19) based on 468 reported transmission events, China, January 21–February 8, 2020. A) All infection events (N = 468) reported across 93 cities of mainland China as of February 8, 2020; B) the subset infection events (n = 122) in which both the infector and infectee were infected in the reporting city (i.e., the index patient's case was not an importation from another city). Gray bars indicate the number of infection events with specified serial interval, and blue lines indicate fitted normal distributions. Negative serial intervals (left of the vertical dotted lines) suggest the possibility of COVID-19 transmission from asymptomatic or mildly symptomatic case-patients.

We note 4 potential sources of bias. First, the data are restricted to online reports of confirmed cases and therefore might be biased toward more severe cases in areas with a high-functioning healthcare and public health infrastructure. The rapid isolation of such case-patients might have prevented longer serial intervals, potentially shifting our estimate downward compared with serial intervals that might be observed in an uncontrolled epidemic. Second, the distribution of serial intervals varies throughout an epidemic; the time between successive cases contracts around the epidemic peak (*9*). A susceptible person is likely to become infected more quickly if they are surrounded by 2 infected persons instead of 1. Because our estimates are based primarily on transmission events reported during the early stages of outbreaks, we do not explicitly account for such compression and interpret the estimates as basic serial intervals at the outset of an epidemic. However, if some of the reported infections occurred amid growing clusters of cases, then our estimates might reflect effective (compressed) serial intervals that would be expected during a period of epidemic growth. Third, the identity of each infector and the timing of symptom onset were presumably based on individual recollection of past events. If recall accuracy is impeded by time or trauma, case-patients might be more likely to attribute infection to recent encounters (short serial intervals) over past encounters (longer serial intervals). In contrast, the reported serial intervals might be biased upward by travel-related delays in transmission from primary case-patients that were infected in Wuhan or another city before returning home. If their infectious period started during travel, then we might be unlikely to observe early transmission events with shorter serial intervals. The mean serial interval is slightly higher for the 218 of 301 unique infectors reported to have imported cases.

Given the heterogeneity in type and reliability of these sources, we caution that our findings should be interpreted as working hypotheses regarding the infectiousness of COVID-19, requiring further validation. The potential implications for COVID-19 control are mixed. Although our lower estimates for R_0 suggest easier containment, the large number of reported asymptomatic transmission events is concerning.

We acknowledge the financial support from the US National Institutes of Health (grant no. U01 GM087719) and the National Natural Science Foundation of China (grant no. 61773091).

About the Author

Dr. Du is a postdoctoral researcher in the Department of Integrative Biology at the University of Texas at Austin. He develops mathematical models to elucidate the transmission dynamics, surveillance, and control of infectious diseases.

References

1. Cowling BJ, Leung GM. Epidemiological research priorities for public health control of the ongoing global novel coronavirus (2019-nCoV) outbreak. Euro Surveill. 2020;25. https://doi.org/10.2807/1560-7917.ES.2020.25.6.2000110
2. Svensson A. A note on generation times in epidemic models. Math Biosci. 2007;208:300–11. https://doi.org/10.1016/j.mbs.2006.10.010
3. Wallinga J, Lipsitch M. How generation intervals shape the relationship between growth rates and reproductive numbers. Proc Biol Sci. 2007;274:599–604. https://doi.org/10.1098/rspb.2006.3754
4. Kuk AYC, Ma S. The estimation of SARS incubation distribution from serial interval data using a convolution likelihood. Stat Med. 2005;24:2525–37. https://doi.org/10.1002/sim.2123
5. Lipsitch M, Cohen T, Cooper B, Robins JM, Ma S, James L, et al. Transmission dynamics and control of severe acute

respiratory syndrome. Science. 2003;300:1966–70. https://doi.org/10.1126/science.1086616
6. Park SH, Kim Y-S, Jung Y, Choi SY, Cho N-H, Jeong HW, et al. Outbreaks of Middle East respiratory syndrome in two hospitals initiated by a single patient in Daejeon, South Korea. Infect Chemother. 2016;48:99–107. https://doi.org/10.3947/ic.2016.48.2.99
7. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. N Engl J Med. 2020 Jan 29 [Epub ahead of print]. https://doi.org/10.1056/NEJMoa2001316
8. 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;395:689–97. https://doi.org/10.1016/S0140-6736(20)30260-9
9. Kenah E, Lipsitch M, Robins JM. Generation interval contraction and epidemic data analysis. Math Biosci. 2008;213:71–9. https://doi.org/10.1016/j.mbs.2008.02.007

Address for correspondence: Lauren Ancel Meyers, J.T. Patterson Labs Bldg, University of Texas at Austin, 2415 Speedway, Austin, TX 78712, USA; email: laurenmeyers@austin.utexas.edu

Indirect Virus Transmission in Cluster of COVID-19 Cases, Wenzhou, China, 2020

Jing Cai,[1] Wenjie Sun,[1] Jianping Huang,[1] Michelle Gamber, Jing Wu, Guiqing He

Author affiliations: Wenzhou Sixth People's Hospital, Wenzhou Central Hospital Medical Group, Wenzhou, China (J. Cai, J. Huang, G. He); The Second Affiliated Hospital of Fujian Traditional Chinese Medical University, Fuzhou, China (W. Sun); Robert Stempel College of Public Health and Social Work, Florida International University, Miami, Florida, USA (W. Sun); Shenandoah University, Winchester, Virginia, USA (M. Gamber); Huashan Hospital, Fudan University, Shanghai, China (J. Wu)

DOI: https://doi.org/10.3201/eid2606.200412

To determine possible modes of virus transmission, we investigated a cluster of coronavirus disease cases associated with a shopping mall in Wenzhou, China. Data indicated that indirect transmission of the causative virus occurred, perhaps resulting from virus contamination of common objects, virus aerosolization in a confined space, or spread from asymptomatic infected persons.

[1]These authors contributed equally to this article.

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the causative agent of coronavirus disease (COVID-19), is presumed to spread primarily via respiratory droplets and close contact. However, these transmission modes do not explain all cases. To determine how the virus may have spread among a cluster of COVID-19 cases associated with a shopping mall in Wenzhou (a city with 8 million residents), China, we monitored and traced close contacts and hypothesized possible transmission modes. We analyzed clinical and laboratory data for cases by using real-time reverse transcription PCR (*1*). The study was approved with written consent from the Ethics Committee of Wenzhou Central Hospital and written informed consent from all case-patients.

On January 20, 2020, a 23-year-old man (patient E) sought care at a hospital after 11 days of fever and headache. On January 21, COVID-19 was confirmed for patient E and his co-worker, patient G. The Wenzhou Center for Disease Control and Prevention traced and tested their contacts, and by January 28, COVID-19 was confirmed for 7 persons (patients A–G) from the same office (on floor 7).

Patient A, a 30-year-old woman, the only case-patient who indicated that she had been in Wuhan, China, returned from Wuhan on December 18, 2019. On January 15–16, 2020, she had a fever, but symptoms resolved without treatment. Despite symptom resolution, on January 30 she was confirmed to have SARS-CoV-2 infection. If patient A is the index patient, infected in Wuhan, her incubation period would have been 28 days, which would be extremely long, according to updated information (W.J. Guan et al., unpub. data, https://www.medrxiv.org/content/10.1101/2020.02.06.20020974v1). Asymptomatic carrier transmission has been reported for SARS-CoV-2 (*2*); hence, patient A could have been screened as a close contact during her incubation period and then hospitalized on the basis of a positive test (PCR) result only. However, her clinical symptoms did not appear until after hospitalization. Because persons with asymptomatic COVID-19 can spread the virus, patient A also could have been an asymptomatic carrier with a persistent infection (*3*).

On January 22, the mall was shut down. During January 19–February 9, COVID-19 was diagnosed for 7 mall staff from floors B1–3 and for 10 mall customers. Close contacts associated with the mall were traced, and COVID-19 was confirmed for 11 persons. Sixteen patients had had direct contact with other patients or had gone shopping in the mall. The average incubation period was 7.3 (range 1–17) days.

The mall has 8 floors above ground and several basement levels; floors B1 to 6 are commercial

Article DOI: https://doi.org/10.3201/eid2606.200357

Serial Interval of COVID-19 among Publicly Reported Confirmed Cases

Appendix

Data

We collected publicly available data on 6,903 confirmed cases from 271 cities of mainland China, that were available online as of February 8, 2020. The data were extracted in Chinese from the Web sites of provincial public health departments and translated to English (https://github.com/MeyersLabUTexas/COVID-19). We then filtered the data for clearly indicated transmission events consisting of: (i) a known *infector* and *infectee*, (ii) reported locations of infection for both cases, and (iii) reported dates and locations of symptom onset for both cases. We thereby obtained 468 infector–infectee pairs identified via contact tracing in 93 Chinese cities between January 21, 2020 and February 8, 2020 (Appendix Figure 1). The index cases (infectors) for each pair are reported as either importations from the city of Wuhan (N = 239), importations from cities other than Wuhan (N = 106) or local infections (N = 122). The cases included 752 unique individuals, with 98 index cases who infected multiple people and 17 individuals that appear as both infector and infectee. They range in age from 1 to 90 years and include 386 females, 363 males and 3 cases of unreported sex.

Inference Methods

Estimating Serial Interval Distribution

For each pair, we calculated the number of days between the reported symptom onset date for the infector and the reported symptom onset date for the infectee. Negative values indicate that the infectee developed symptoms before the infector. We then used the fitdist function in Matlab (*1*) to fit a normal distribution to all 468 observations. It finds unbiased estimates of the mean and standard deviation, with 95% confidence intervals. We applied the

same procedure to estimate the means and standard deviations with the data stratified by whether the index case was imported or infected locally.

Estimating the Basic Reproduction Number (R_0)

Given an epidemic growth rate *r* and *normally distributed generation times* with mean (μ) and standard deviation (σ), the basic reproduction number is given by

$R_0 = e^{r\mu - (1/2)r^2\sigma^2}$ (*2*).

Since we do not know the COVID-19 generation time distribution, we use our estimates of the COVID-19 serial interval distribution as an approximation (Appendix Table 1), noting that the serial interval distribution tends to be more variable than the generation time distribution. We assume that the COVID-19 growth rate (*r*) is 0.10 per day [95% CI 0.050–0.16] based on a recent analysis of COVID-19 incidence in Wuhan, China (*3*). To estimate R_0, we take 100,000 Monte Carlo samples of the growth rate ($r \sim N(.1, 0.028)$) and the mean and standard deviation of the serial interval ($(\mu, \sigma) \sim N(\mathbf{M}, \Sigma)$ where $\mathbf{M} = (3.96, 4.75)$ and $\Sigma = \begin{bmatrix} 0.048 & 2.4 \cdot 10^{-18} \\ 2.4 \cdot 10^{-18} & 0.24 \end{bmatrix}$). We thereby estimate an R_0 of 1.32 [95% CI 1.16–1.48].

Model Comparison

We used maximum likelihood fitting and the Akaike information criterion (AIC) to evaluate four candidate models for the COVID-19 serial interval distributions: normal, lognormal, Weibull and gamma. Since our serial interval data includes a substantial number of non-positive values, we fit the four distributions both to truncated data in which all non-positive values are removed and to shifted data in which 12 days are added to each observation (Appendix Figure 1, Appendix Tables 2, 3). The lognormal distribution provides the best fit for the truncated data (followed closely by the gamma and Weibull). However, we do not believe there is cause for excluding the non-positive data and would caution against making assessments and projections based on the truncated data. The normal distribution provides the best fit for the full dataset (shifted or not) and thus is the distribution we recommend for future epidemiologic assessments and planning.

Supplementary Analysis

To facilitate interpretation and future analyses, we summarize key characteristics of the COVID-2019 infection report dataset.

Age Distribution

Of the 737 unique cases in the dataset, 1.7%, 3.5%, 54.1%, 26.1% and 14.5% were ages 0–4, 5–17, 18–49, 50–64, and over 65 years, respectively. Across all transmission events, approximately one third occurred between adults ages 18 to 49, ~92% had an adult infector (over 18), and over 99% had an adult infectee (over 18) (Appendix Table 4).

Secondary Case Distribution

Across the 468 transmission events, there were 301 unique infectors. The mean number of transmission events per infector is 1.55 (Appendix Figure 2) with a maximum of 16 secondary infections reported from a 40 year old male in Liaocheng city of Shandong Province.

Geographic Distribution

The 468 transmission events were reported from 93 Chinese cities in 17 Chinese provinces and Tianjin (Appendix Figure 3). There are 22 cities with at least five infection events and 71 cities with fewer than five infection events in the sample. The maximum number of reports from a city is 72 for Shenzhen, which reported 339 cumulative cases as of February 8, 2020.

References

19. MathWorks. Fit probability distribution object to data—MATLAB fitdist [cited 2020 Feb 19]. https://www.mathworks.com/help/stats/fitdist.html

20. Wallinga J, Lipsitch M. How generation intervals shape the relationship between growth rates and reproductive numbers. Proc Biol Sci. 2007;274:599–604. PubMed https://doi.org/10.1098/rspb.2006.3754

21. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. N Engl J Med. 2020;NEJMoa2001316; Epub ahead of print. PubMed https://doi.org/10.1056/NEJMoa2001316

Appendix Table 1. Model comparison for COVID-19 serial intervals based on 122 reported transmission events in China between January 21, 2020 and February 8, 2020 in which both the infector and infectee were infected in the reporting city (i.e., the index case was not an importation from another city)

Data	Distribution	Mean/Shape (95% CI)	SD/Scale (95% CI)	AIC
Original data	Normal (Mean, SD)	3.66 (2.84– 4.47)	4.54 (4.03–5.20)	718.45
Truncated (>0)	Normal (Mean, SD)	5.21 (4.47–5.95)	3.66 (3.20–4.26)	524.34
	Lognormal (Shape, Scale)	2.14 (1.64–2.78)	2.44 (1.81–3.28)	486.81
	Gamma (Shape, Scale)	1.40 (1.25–1.55)	0.75 (0.65–0.87)	487.68
	Weibull (Shape, Scale)	5.81 (5.05–6.68)	1.52 (1.30–1.77)	489.59
Shifted (+12d)	Normal (Mean, SD)	13.66 (12.84– 14.47)	4.54 (4.03– 5.20)	718.45
	Lognormal (Shape, Scale)	7.52 (5.88–9.61)	1.82 (1.41–2.34)	730.75
	Gamma(Shape, Scale)	2.55 (2.47–2.62)	0.41 (0.37–0.47)	755.39
	Weibull(Shape, Scale)	15.19 (14.32–16.11)	3.18 (2.79–3.63)	722.22

Appendix Table 2. Estimated serial interval distributions based on the location of index infection (imported versus local) and the secondary infection (household versus nonhousehold)*

Group	Mean (95 CI%)	SD (95 CI%)	Proportion of serial intervals <0
All (N = 468)	3.96 (3.53–4.39)	4.75 (4.46–5.07)	12.61% (N = 59)
Locally infected index case (n = 122)	3.66 (2.84–4.47)	4.54 (4.03–5.20)	14.75% (N = 18)
Imported index case (n = 346)	4.06 (3.55–4.57)	4.82 (4.48–5.21)	11.85% (N = 41)
Household secondary infection (n = 104)	4.03 (3.12–4.94)	4.69 (4.12–5.43)	16.35% (N = 17)
Nonhousehold secondary infection (n = 180)	4.56 (3.85–5.27)	4.80 (4.35–5.36)	11.11% (N = 20)

*We assume that the serial intervals follow normal distributions and report the estimated means and standard deviations for all 468 infector–infectee pairs reported from 93 cities in mainland China by February 8, 2020, 122 pairs in which the index case was infected locally, 346 pairs in which the index case was an importation from another city, 104 pairs in which the secondary transmission event occurred within a household, and 180 pairs in which the secondary transmission event was reported as non-household. The rightmost column provides the proportion of infection events in which the secondary case developed symptoms before the index case.

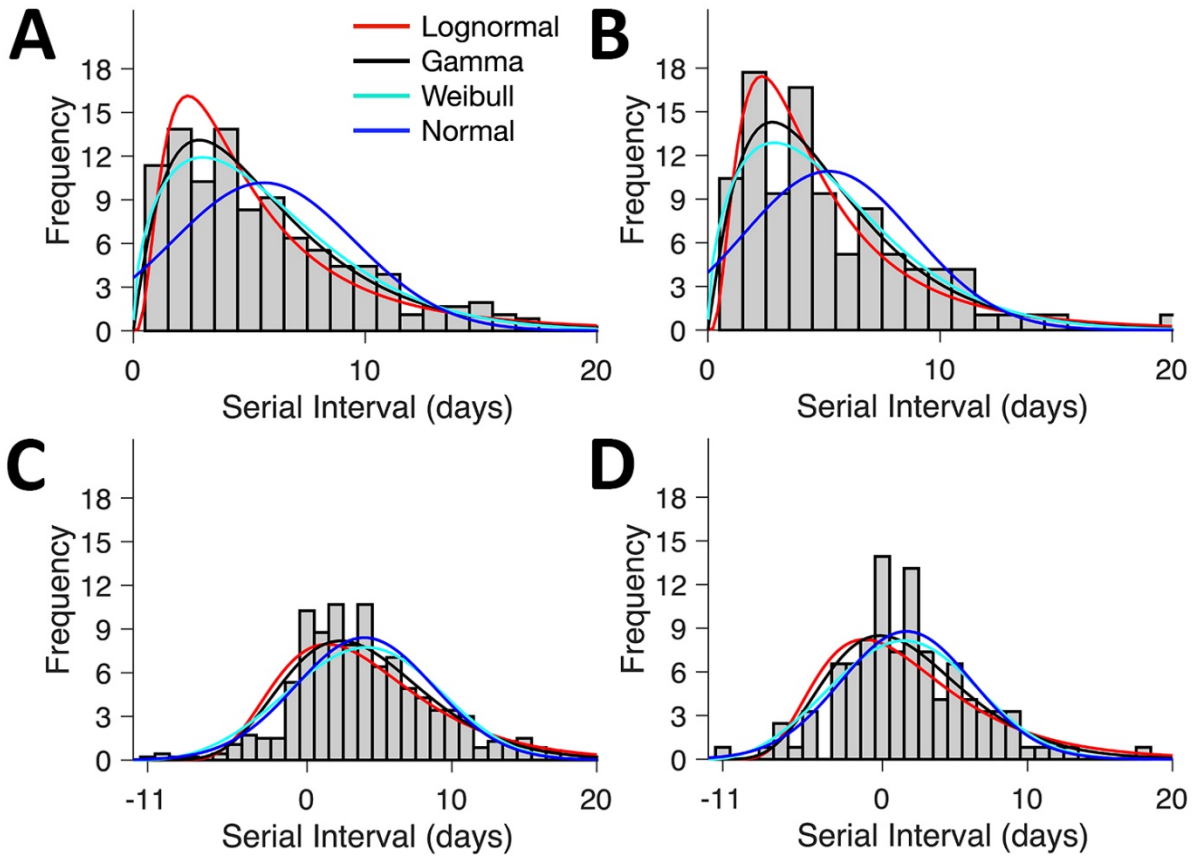
Appendix Table 3. Model comparison for COVID-19 serial intervals based on all 468 reported transmission events in China between January 21, 2020 and February 8, 2020.

Data	Distribution	Mean/Shape (95% CI)	SD/Scale (95% CI)	AIC
Original data	Normal (Mean, SD)	3.96 (3.53– 4.39)	4.75 (4.46– 5.07)	2,788.77
Truncated (>0)	Normal (Mean, SD)	5.62 (5.21–6.03)	3.92 (3.66–4.23)	2,014.12
	Lognormal (Shape, Scale)	2.02 (1.76–2.31)	2.78 (2.39–3.25)	1,886.73
	Gamma (Shape, Scale)	1.46 (1.38–1.54)	0.78 (0.72–0.84)	1,898.63
	Weibull (Shape, Scale)	6.25 (5.81–6.72)	1.50 (1.38–1.62)	1,892.04
Shifted (+12d)	Normal (Mean, SD)	15.96 (15.53–16.39)	4.75 (4.46–5.07)	2,788.77
	Lognormal (Shape, Scale)	9.83 (8.66–11.15)	1.62 (1.43–1.85)	2,822.71
	Gamma(Shape, Scale)	2.72 (2.69–2.75)	0.35 (0.33–0.38)	2,898.24
	Weibull(Shape, Scale)	17.66 (17.19–18.14)	3.56 (3.32–3.80)	2,806.21

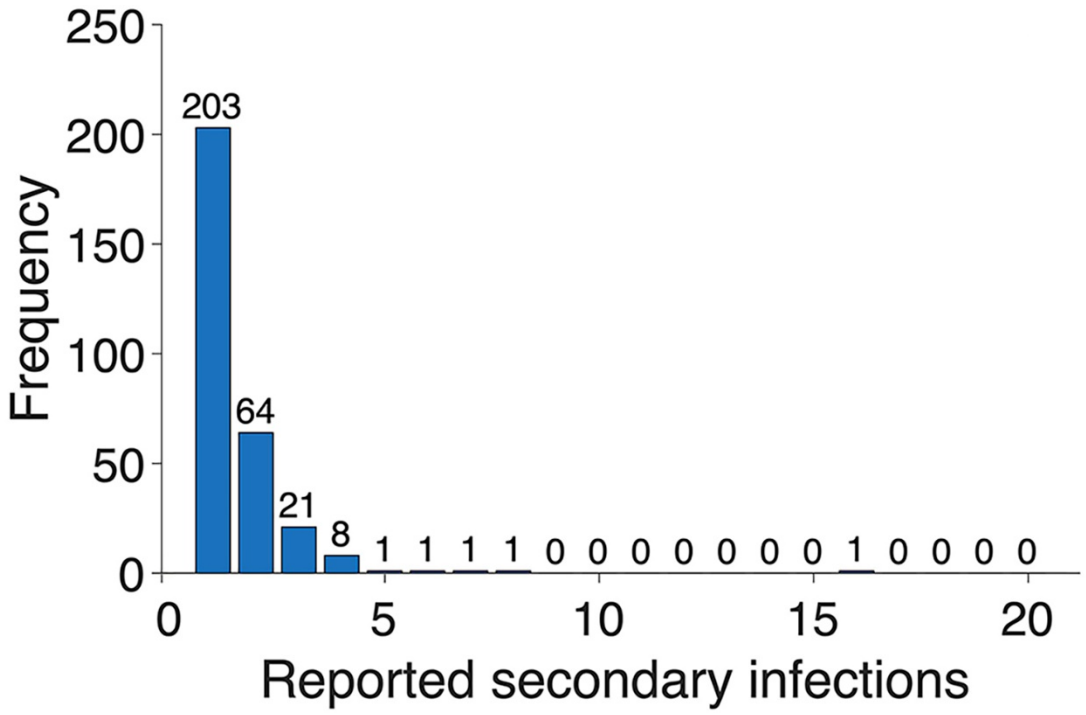
Appendix Table 4. Age distribution for the 457 of 468 infector–infectee pairs*

	Infectee age group, y					
Infector age group, y	0–4	5–17	18–49	50–46	≥65	Total
0–4	0	0	0	0	0	0
5–17	0	0	1	0	1	2
18–49	12	18	154	60	44	288
50–46	1	5	47	49	13	115
≥65	0	1	22	10	19	52
Total	13	24	224	119	77	457

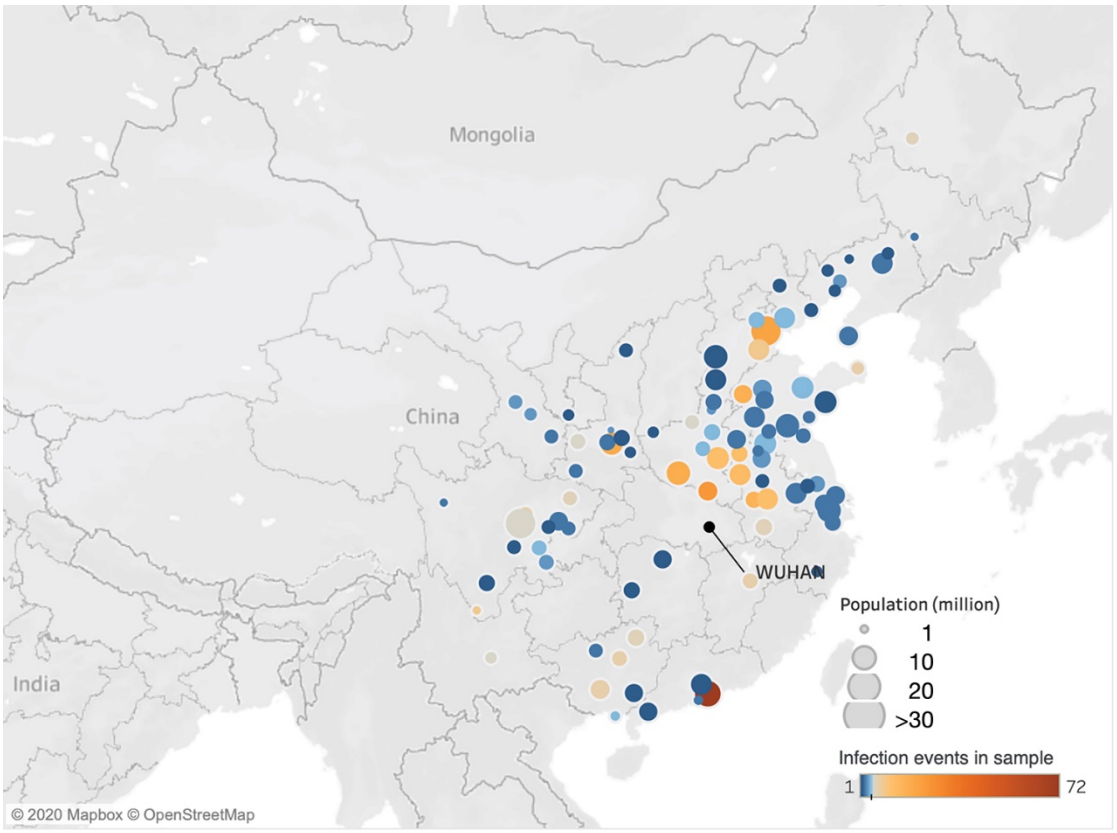
*Each value denotes the number of infector–infectee pairs in the specified age combination. Age was not reported for the remaining 11 pairs.



Appendix Figure 1. Maximum likelihood distributions fit to transformed COVID-19 serial intervals (468 reported transmission events across 93 cities in Mainland China between January 21, 2020 and February 8, 2020). To evaluate several positive-valued distributions (lognormal, gamma and Weibull), we took two approaches to addressing the negative-valued data. First, we left truncated the data (i.e., removed all non-positive values) for (a) all 468 infection events and (b) the subset of infection events (N = 122) in which both the infector and infectee were infected in the reporting city (i.e., the index case was not an importation from another city). Second, we shifted the data by adding 12 days to each reported serial interval for (c) all infection events and (d) the subset of infection events in which both the infector and infectee were locally infected. Bars indicate the number of infection events with the specified serial interval and colored lines indicate the fitted distributions. Parameter estimates and AIC values are provided in Appendix Table 3.



Appendix Figure 2. Number of infections per unique index case in the infection report dataset. There are 301 unique infectors across the 468 infector-infectee pairs. The number of transmission events reported per infector ranges from 1 to 16, with ~55% having only one.



Appendix Figure 3. Geographic composition of the infection report dataset. The data consist of 468 infector-infectee pairs reported by February 8, 2020 across 93 cities in mainland China. Colors represent the number of reported events per city, which range from 1 to 72, with an average of 5.03 (SD 8.54) infection events. The 71 cities with fewer than five events are colored in blues; the 22 cities with at least five events are colored in shades of orange.