

## **Supplemental Appendix**

### **High transmissibility of COVID-19 near symptom onset**

Hao-Yuan Cheng, M.D., M.Sc., Shu-Wan Jian, D.V.M., M.PH., Ding-Ping Liu, Ph.D.,  
Ta-Chou Ng, B.Sc., Wan-Ting Huang, M.D. for CDC outbreak investigation team,  
Hsien-Ho Lin M.D., Sc.D.

#### **Affiliations**

Epidemic Intelligence center, Taiwan Centers for Disease Control (HYC, SWC, DPL)  
Office of Preventive Medicine, Taiwan Centers for Disease Control (WTH)  
Institute of Epidemiology and Preventive Medicine, College of Public Health, National  
Taiwan University (TCN, HHL)  
Global Health Program, College of Public Health, National Taiwan University (HHL)

#### **Corresponding author**

Dr. Hsien-Ho Lin, M.D., Sc.D.  
17 Xuzhou Road, Taipei, Taiwan  
hsienho@ntu.edu.tw

## **Content**

<b>1. Member list of Taiwan COVID-19 outbreak investigation team</b>	<b>3</b>
<b>2. Relative risk of secondary clinical transmission by exposure period after symptom onset</b>	<b>4</b>
<b>3. Density function of transmissibility by exposure period after symptom onset</b>	<b>5</b>
<b>4. 77</b>	
<b>5. 1111</b>	

## **1. Member list of Taiwan COVID-19 outbreak investigation team**

### **Taiwan CDC**

#### **Office of Preventive Medicine and Taiwan Field Epidemiology Training Program**

Wan-Ting Huang  
Wan-Chin Chen  
Angela Song-En Huang  
Chia-Ping Su  
Pin-Hui Lee  
Pei-Chun Chan  
Hao-Hsin Wu  
Shih-Tse Huang  
Tsung-Pei Tsou  
Ying-Shih Su  
Yang Li

### **Taipei Regional Control Center**

Hsin-Yi Wei  
Meng-Yu Chen  
Shiao-Ping Tung  
Yu-Fang Tsai  
Xiang-Ting Huang  
Chien-Yu Chou

### **Changhua County Public Health Bureau**

Yen-Po Yeh

### **North Regional Control Center**

Pei-Yuan Wu  
Fang-Tzu Chang  
Chia-Ying Yen  
Hsueh-Mei Chiang  
Ju-Huei Lin

### **Central Regional Control Center**

Kung-Chin Wang  
Ching-Fen Ko  
Pei-Fang Lai  
Du-Ling Lin  
Min-Tsung Lin  
Zhi-Jie Ding  
Hsiao-Chi Wang

### **South Regional Control Center**

Huai-Te Tsai  
Ping-Jung Liu  
Pei-Yi Lin

### **Kao-Ping Regional Control Center**

Hsin-Chun Lee  
Ching-Li Lin  
Chi-Mei Lai  
Hsiao-Mei Liu

## 2. Relative risk of secondary clinical transmission by exposure period after symptom onset

The relative risk of secondary clinical transmission were estimated by fitting a logistic regression model with the piecewise exposure periods as predictors, considering the whole exposure window period. The contact tracing dataset was described in detail in the main text. We implemented the model in a Bayesian framework as below. For each contact  $i$ , the outcome ( $Y_i$ ) is determined the infection probability ( $p_i$ ) which is modeled by a linear predictor of the exposure dose.

$$Y_i \sim \text{Bernoulli}(p_i)$$

$$\text{logit}(p_i) = b_0 + b_1 Z_i^{0-3} + b_2 Z_i^{4-5} + b_3 Z_i^{6-7} + b_4 Z_i^{8-9} + b_5 Z_i^{>9}$$

where  $Z_i^{t_1-t_2}$  denotes the exposure dose of individual  $i$  during the period  $[t_1, t_2]$ , relative to the symptom onset time of the known infector of  $i$ . The exposure dose was defined in two ways.

- (i) Indicator approach:  $Z_i^{t_1-t_2} = 1$  when individual  $i$  was exposed to his/her infector at any time during the period  $[t_1, t_2]$ , otherwise,  $Z_i^{t_1-t_2} = 0$ . The exponentiated coefficient  $\exp(b_m)$  was interpreted as the odds ratio of infection comparing a contact to a non-contact during the  $m$ th exposure period, holding the exposure dose in other periods constant.
- (ii) Person-time approach, where  $Z_i^{t_1-t_2}$  was defined as the contact/exposed duration (days) during the period  $[t_1, t_2]$ . The exponentiated coefficient  $\exp(b_m)$  was interpreted as the odds ratio of infection for an additional day of contact during the  $m$ th exposure period, holding the exposure dose in other periods constant.

Non-informative priors for the regression coefficients were used. For each exposure period  $m = 1, \dots, 5$ ,

$$b_m \sim \text{Normal}(0, 1000)$$

We ran 4 Markov Chain Monte Carlo (MCMC) chains using the No-U-Turn Sampler (NUTS) algorithm in *Stan*<sup>1</sup>. Each chain contains 1000 warmup iterations and 500 samples, rendering to a total 2000 samples.

### 3. Density function of transmissibility by exposure period after symptom onset

The density function of transmissibility represents the time-varying transmissibility of the index case. We fit a gamma distribution to the contact tracing data to estimate the transmissibility since the date of symptomatic onset of the index case. For each contact  $i$  with the exposure period  $[t_1, t_2]$ , we modeled his/her infection probability  $p_i$  to be proportional to the index case's cumulative transmissibility during  $[t_1, t_2]$ , defined as the area under the transmissibility function during the exposure period  $[t_1, t_2]$ .

$$Y_i \sim \text{Bernoulli}(p_i)$$

$$p_i = w \int_{t_1}^{t_2} f(t) dt = w \int_{t_1}^{t_2} \frac{1}{\Gamma(\alpha)\beta^\alpha} t^{\alpha-1} e^{-t/\beta} dt$$

$w$  is a scalar ranging between 0 to 1, and represents the maximum infection probability when the individual is always contacting with the primary case.  $f(t)$  is the gamma-distributed density function of transmissibility, with shape ( $\alpha$ ) and scale ( $\beta$ ). The transmissibility function was then truncated at  $t = 30$  days which is about the maximum observed exposure time in the data, and we can only reliably estimate the transmissibility within this bound. Non-informative priors for the distribution parameters were used.

$$w \sim \text{Uniform}(0, 1)$$

$$\alpha \sim \text{Exponential}(0.001)$$

$$\beta \sim \text{Exponential}(0.001)$$

The estimated posterior distribution of parameters were shown in Table S1. The estimated density function of transmissibility was shown in Figure 2D, and Table S2.

**Table S1.** The estimated distribution parameters of the infectiousness density function with 95% credible interval (CrI).

Parameter	Estimates (95% CrI)
w	0.13 (0.06, 0.23)
shape	0.14 (0.04, 0.37)
scale	1043.72 (6.73, 3768.89)

**Table S2.** The infection probability given the estimated infectiousness density function with 95% credible interval (CrI).

Exposure period	Cumulative infectiousness (95% CrI)
0 ~ 1d	66.94% (94.83%, 28.65%)
0 ~ 3d	76.28% (97.73%, 42.93%)
0 ~ 5d	81.05% (98.75%, 51.8%)
0 ~ 7d	84.34% (99.25%, 58.62%)
0 ~ 9d	86.89% (99.53%, 64.3%)

## 4. Serial interval distribution

### 4.1 Data

We collated the symptom onset dates of COVID-19 patients with confirmed transmission links in published papers, news reports and press releases from local governments. In total, there are 48 transmission pairs where both infectors and infectees developed symptoms after the infection. The exposure settings were categorized into (1) household, (2) community and, (3) others (including workplace, friend, family, and unclear setting). Countries were categorized in to (1) Taiwan, (2) Mainland China, and (3) Others. The full dataset is available upon request.

**Table S3.** The frequencies of the collated transmission pairs by countries and settings.

Country	Setting		
	Household	Community	Others
Mainland China	11	0	3
Germany	0	0	2
Hong Kong	2	3	1
Macau	0	0	1
Singapore	4	2	5
South Korea	2	0	0
Taiwan	5	0	4
Vietnam	3	0	0

### 4.2 Methods

We adopted the statistical method by Backer, et al<sup>2</sup> to deal with uncertain interval data. A gamma distribution was fitted to the observed serial intervals ( $SI$ ) in a Bayesian framework. For each stratum  $k = 1, 2, 3$ ,

$$\begin{aligned} SI^k &\sim \text{Gamma}(\alpha^k, \beta^k) \\ \alpha^k &\sim \text{Exponential}(1/\theta^{shape}) \\ \beta^k &\sim \text{Exponential}(1/\theta^{scale}) \end{aligned}$$

where each stratum ( $k$ ) has its own serial interval distribution. The stratum-specific distribution parameters, shape ( $\alpha^k$ ) and scale ( $\beta^k$ ) are related by an exponential hyper-parameter distribution, with mean  $\theta^{shape}$ , and  $\theta^{scale}$  respectively. Non-informative priors for the hyper-parameters were used.

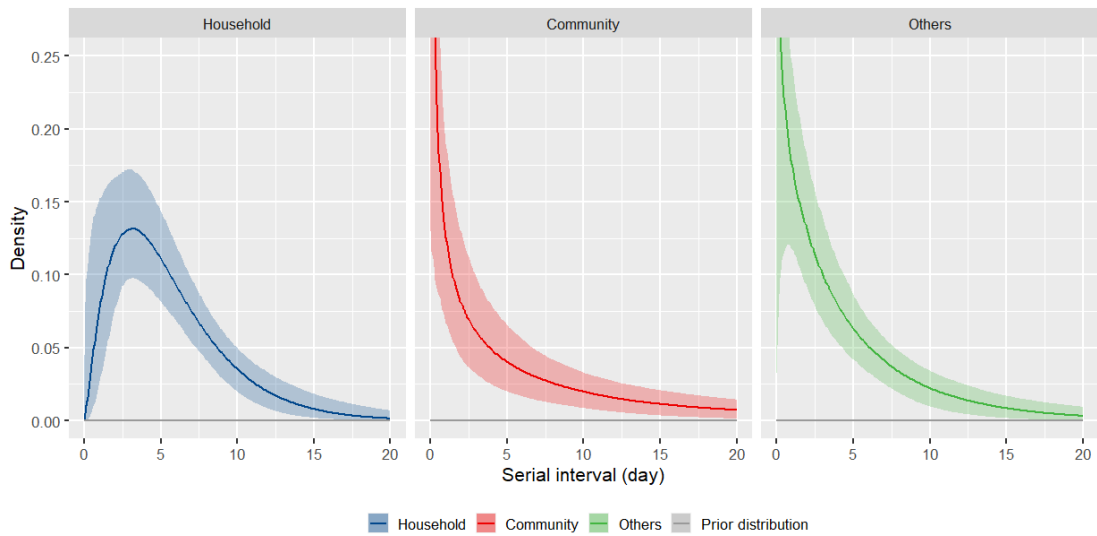
$$\begin{aligned} \theta^{shape} &\sim \text{Half - normal}(0, 1000) \\ \theta^{scale} &\sim \text{Half - normal}(0, 1000) \end{aligned}$$

Since the reported symptom onset dates can be intervals, we set uniform priors for the onset intervals, and estimated the symptom onset dates. For each individual  $i = 1, \dots, n$

$$\begin{aligned}
X_i^0 &\sim \text{Uniform}(XL_i^0, XR_i^0) \\
X_i^1 &\sim \text{Uniform}(XL_i^1, XR_i^1) \\
SI_i &= X_i^1 - X_i^0
\end{aligned}$$

where  $[XL_i^0, XR_i^0]$  is the symptom onset interval for the infector in pair  $i$ , and  $[XL_i^1, XR_i^1]$  is the symptom onset interval for the infectee. We ran 4 MCMC chains using the NUTS algorithm in *Stan*. Each chain contains 1000 warmup iterations and 500 samples, resulting in a total 2000 samples.

### 4.3 Serial interval by exposure setting



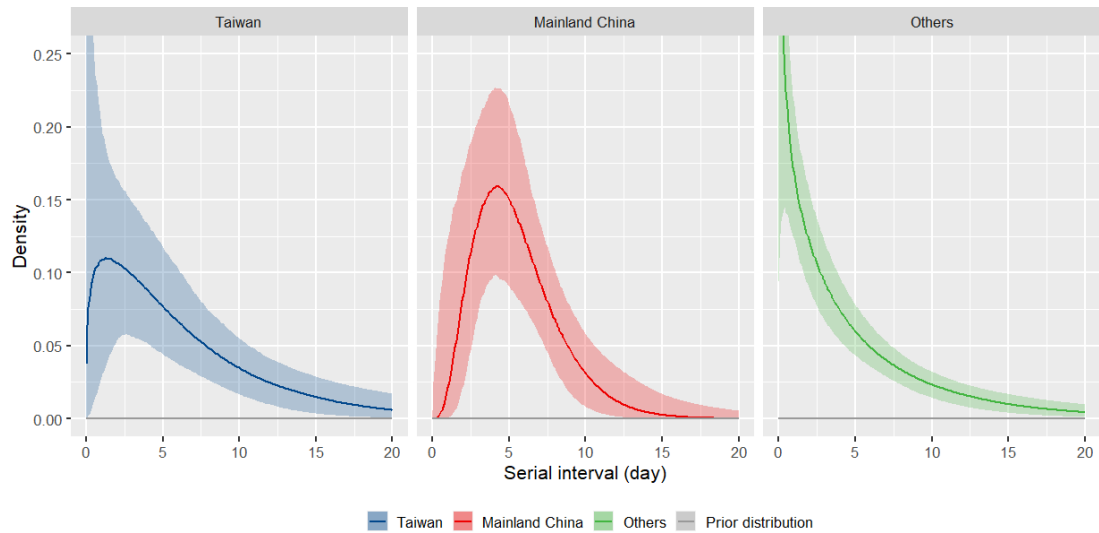
**Figure S1.** The estimated serial interval distributions for each exposure setting (household, community, and others), in solid lines. The shaded areas represent 95% credible interval (CrI).

**Table S4.** The estimated distribution parameters of the serial interval for each exposure setting (household, community, and others), with 95% credible interval (CrI). (S.D. = standard deviation)

Setting	Mean (95% CrI)	S.D. (95% CrI)	Shape (95% CrI)	Scale (95% CrI)
Community	9.91 (2.35,35.77)	15.6 (3.81,59.33)	0.47 (0.16,0.97)	25.86 (5.13,106.36)
Household	5.69 (4.36,7.47)	3.97 (2.76,5.74)	2.15 (1.23,3.38)	2.82 (1.55,4.89)
Others	4.99 (2.75,9.38)	5.82 (3,11.8)	0.81 (0.4,1.37)	6.98 (2.82,16.49)

### 4.4 Serial interval by country



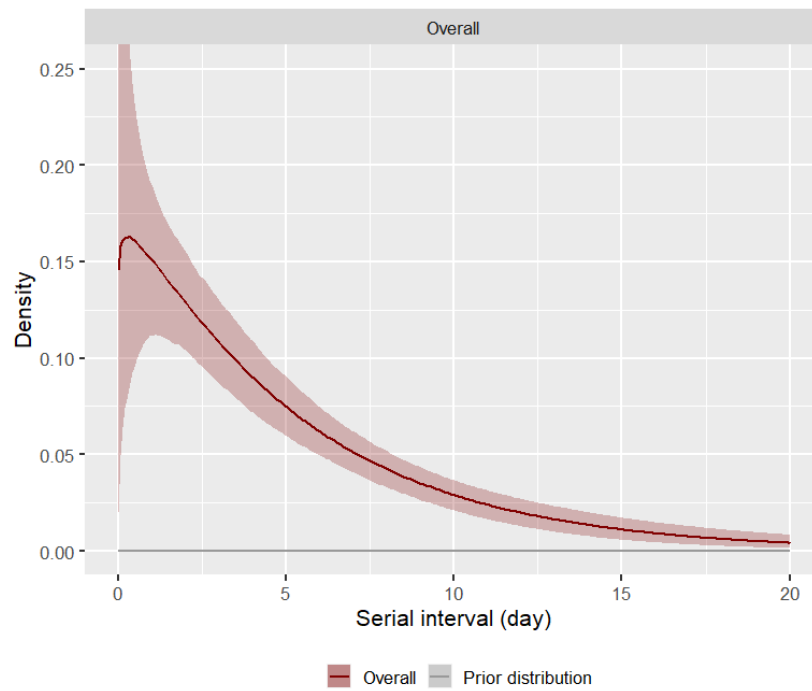


**Figure S2.** The estimated serial interval distributions for each region (Taiwan, Mainland China, and others), in solid lines. The shaded areas represent 95% credible interval (CrI).

**Table S5.** The estimated distribution parameters of the serial interval for each region (Taiwan, Mainland China, and others), with 95% credible interval (CrI). (S.D. =standard deviation)

Region	Mean (95% CrI)	S.D. (95% CrI)	Shape (95% CrI)	Scale (95% CrI)
Mainland China	5.57 (4.2,7.51)	2.86 (1.82,4.97)	4.28 (1.7,8.2)	1.54 (0.66,3.59)
Others	5.09 (3.16,8.09)	6.02 (3.61,10.43)	0.75 (0.44,1.15)	7.23 (3.67,14.89)
Taiwan	7.00 (3.69,13.18)	6.41 (3.19,14.09)	1.37 (0.51,2.73)	6.14 (2.16,17.63)

## 4.5 Overall serial interval



**Figure S3.** The estimated serial interval distribution (overall countries and settings), in solid lines. The shaded areas represent 95% credible interval (CrI).

**Table S6.** The estimated distribution parameters of the serial interval (overall countries and settings), with 95% credible interval (CrI). (S.D. =standard deviation)

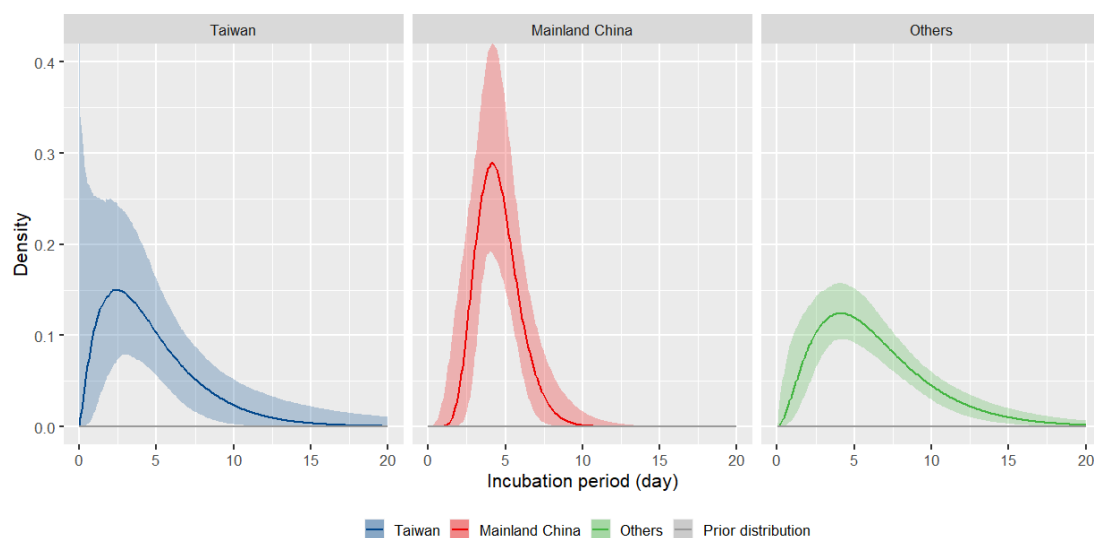
	Mean (95% CrI)	S.D. (95% CrI)	Shape (95% CrI)	Scale (95% CrI)
Overall	5.44 (4.11,7.15)	5.33 (3.82,7.53)	1.07 (0.72,1.47)	5.27 (3.34,8.23)

## 5. Incubation period distribution

### 5.1 Data and methods

Apart from the 9 symptomatic cases with known exposure periods in Taiwan (Figure 1C in the main text), we from publicly available, crowd-sourced dataset provided by Sun, et al<sup>3</sup>. Only cases with known exposure windows (both beginning and end) and the dates of symptom onset were included, because the results were sensitive to ad hoc beginning dates of exposure. In total, 65 cases were included in the analysis of the incubation period. We applied same statistical methods as in the estimation of the serial interval distribution, by substituting  $XL_i^0$  and  $XR_i^0$  by the lower and upper bounds of the exposure period  $EL_i^0$  and  $ER_i^0$  in the previous model.

### 5.2 Incubation period by country

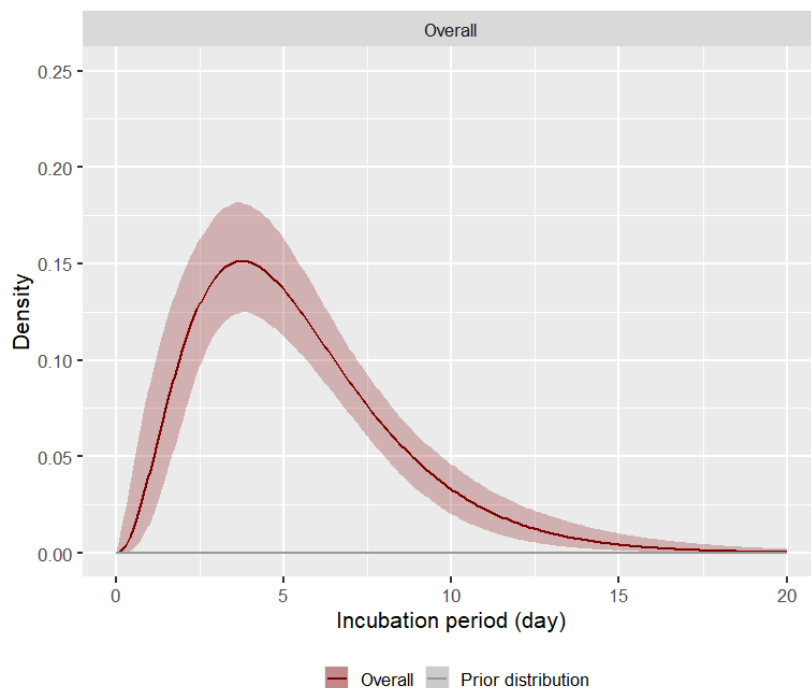


**Figure S4.** The estimated incubation period distributions for each region (Taiwan, Mainland China, and others), in solid lines. The shaded areas represent 95% credible interval (CrI).

**Table S7.** The estimated distribution parameters of the incubation period for each region (Taiwan, Mainland China, and others), with 95% credible interval (CrI). (S.D. =standard deviation)

Region	Mean (95% CrI)	S.D. (95% CrI)	Shape (95% CrI)	Scale (95% CrI)
Mainland China	4.54 (3.9,5.26)	1.43 (0.94,2.18)	11.44 (4.48,23.11)	0.47 (0.2,1.02)
Others	6.35 (5.19,7.75)	3.81 (2.82,5.21)	2.90 (1.72,4.48)	2.32 (1.37,3.86)
Taiwan	4.91 (2.72,8.43)	3.51 (1.76,7.4)	2.30 (0.85,4.94)	2.63 (0.89,6.98)

### 5.3 Overall incubation period



**Figure S5.** The estimated incubation period distribution (overall countries and settings), in solid lines. The shaded areas represent 95% credible interval (CrI).

**Table S8.** The estimated distribution parameters of the incubation period (overall countries and settings), with 95% credible interval (CrI). (S.D. =standard deviation)

	Mean (95% CrI)	S.D. (95% CrI)	Shape (95% CrI)	Scale (95% CrI)
Overall	5.45 (4.71,6.31)	3.07 (2.44,3.9)	3.24 (2.22,4.48)	1.74 (1.17,2.55)

## References

- 1 Carpenter, B. *et al.* Stan: A Probabilistic Programming Language. 2017 **76**, Journal of Statistical Software, doi:10.18637/jss.v076.i01 (2017).
- 2 Backer, J. A., Klinkenberg, D. & Wallinga, J. Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020. *Eurosurveillance* **25** (2020).
- 3 Sun, K., Chen, J. & Viboud, C. Early epidemiological analysis of the coronavirus disease 2019 outbreak based on crowdsourced data: a population-level observational study. *The Lancet Digital Health*, doi:[https://doi.org/10.1016/S2589-7500\(20\)30026-1](https://doi.org/10.1016/S2589-7500(20)30026-1) (2020).