Title: The time-varying serial interval of the coronavirus disease (COVID-19) and its gender-specific difference: A data-driven analysis using public surveillance data in Hong Kong and Shenzhen, China from January 10 to February 15, 2020

Running title: Time-varying serial interval of COVID-19 and its gender-specific difference

Shi Zhao[1,2,*], Peihua Cao[3], Marc KC Chong[1,2], Daozhou Gao[4], Yijun Lou[5], Jinjun Ran[6], Kai Wang[7], Weiming Wang[8,*], Lin Yang[9], Daihai He[5,*], and Maggie H Wang[1,2]

1 JC School of Public Health and Primary Care, Chinese University of Hong Kong, Hong Kong, China
2 Shenzhen Research Institute of Chinese University of Hong Kong, Shenzhen, China
3 Clinical Research Centre, Zhujiang Hospital, Southern Medical University, Guangzhou, China
4 Department of Mathematics, Shanghai Normal University, Shanghai, China
5 Department of Applied Mathematics, Hong Kong Polytechnic University, Hong Kong, China
6 School of Public Health, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong, China
7 Department of Medical Engineering and Technology, Xinjiang Medical University, Urumqi, 830011, China
8 School of Mathematics and Statistics, Huaiyin Normal University, Huaian, China
9 School of Nursing, Hong Kong Polytechnic University, Hong Kong, China

* Correspondence to: zhaoshi.cmsa@gmail.com (SZ), weimingwang2003@163.com (WW), and daihai.he@polyu.edu.hk (DH)

Word count 564

Email address of all authors
SZ: zhaoshi.cmsa@gmail.com
PC: cphcc@smu.edu.cn
MKCC: marc@cuhk.edu.hk
DG: dzgao@shnu.edu.cn
YL: yijun.lou@polyu.edu.hk
JR: jimran@connect.hku.hk
KW: wangkaimath@sina.com
WW: weimingwang2003@163.com
LY: l.yang@polyu.edu.hk
DH: daihai.he@polyu.edu.hk
MHW: maggiew@cuhk.edu.hk

Main text

To the Editor

An outbreak of coronavirus disease (COVID-19), started in Wuhan, China in the end of 2019 [1], has now reached 40 countries and poses huge threat to global public health and economic [2]. Given the risk of human-to-human transmission, the serial interval (SI), which refers to the time interval from symptom onset of a primary case (i.e., infector) to that of a secondary case (i.e., infectee) [3], is the other essential quantity besides the basic reproduction number to drive the spreading speed.

We examine the publicly available materials and collect the records of COVID-19 transmission events in two neighboring large cities, Hong Kong [4] and Shenzhen [5], in south China from January 10 to February 15, 2020 and extract the SI data. We identify 48 transmission events including 21 in Hong Kong and 27 in Shenzhen, among which 40 events contain the gender information of the primary cases. The last onset date of the primary cases among all collected transmission events is February 2, 2020.

To explore the temporal patterns and the gender-specific difference of SI, we adopt two regression models as follows.

- Model (**1**): log-linear form for the percentage change, $\mathbf{E}[\ln(\mathrm{SI}_{i,t})] = \alpha_1 G_i + \alpha_2 t + \alpha_0$; and

- Model (**2**): linear form for the unit change, $\mathbf{E}[\mathrm{SI}_{i,t}] = \beta_1 G_i + \beta_2 t + \beta_0$.

The $\mathbf{E}[\cdot]$ is the expectation, and the α and β are the regression coefficients. The $\mathrm{SI}_{i,t}$ represents the SI of the i-th primary case whose onset date is the t-th day. The G_i denotes the gender of the i-th primary case. Hence, the $[\exp(\alpha_2) - 1] \times 100\%$ quantifies the percentage change, and β_2 quantifies the unit change (day) in the SI, namely change per day in the calendar date. The gender-specific difference can be interpreted similarly. We fit both models via the standard least square approach.

In Fig 1, we find that SI had been decreasing by 0.4 (95%CI: 0.1−0.7) per day, or 6.2% (95%CI: 0.4−11.6%) in percentage, from January 10 to February 2 in Hong Kong and Shenzhen. The Pearson correlation coefficient between the SI and calendar date is estimated at −0.37 with p-value < 0.01. The SI of male primary cases is 3.5 days (95%CI: 1.2−5.7) shorter than that of a female primary case, or 49.7% (95%CI: 15.3−70.1%) less in percentage. To verify, we additionally conduct Cox proportional hazard modelling analysis using similar formula as in models (1) and (2), and calculate the hazard ratio estimates. We find the

association between SI and calendar date as well as gender-specific difference hold consistently and significantly.

The shortening in SI over time is likely due to the strengthening of the public health control measures. The contact tracing and timely isolation of confirmed COVID-19 infections could lead to shorter observed SI due to right censoring 'bias' [6, 7]. As such, we call the observed SI under the effects of control measures the effective SI, which has a mean 5.2 days from our dataset. This appears slightly but not significantly shorter than the previous estimated `intrinsic' SI with a mean 7.5 days [1]. The mechanism behind the gender difference remains uncovered, but may be partly due to that male cases are more severe than female cases ("*officials recorded a 2.8% fatality rate for male patients versus 1.7% for women*" [8]). In this work, we report two findings in the SI of COVID-19 in Hong Kong and Shenzhen, and their implication warrants further investigation.

Declarations

The funding agencies had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; or decision to submit the manuscript for publication.

Conflict of Interests

The authors declared no competing interests.

Authors' Contributions

SZ and DH conceived the study, carried out the analysis, and drafted the first manuscript. All authors discussed the results, critically read and revised the manuscript, and gave final approval for publication.

References

1.   Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, Ren R, Leung KSM, Lau EHY, Wong JY *et al*: **Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia**. *New England Journal of Medicine* 2020.

2.   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.

3.   Fine PEM: **The Interval between Successive Cases of an Infectious Disease**. *American Journal of Epidemiology* 2003, **158**(11):1039-1047.

4.   **The collection of Press Releases by the Centre for Health Protection (CHP) of Hong Kong.** [https://www.chp.gov.hk/en/media/116/index.html]

5.   **The collection of outbreak situation reports of COVID-19 in shenzhen, Shenzhen Municiple health Commission.** [http://wjw.sz.gov.cn/yqxx/]

6.   Zhao S, Gao D, Zhuang Z, Chong M, Cai Y, Ran J, Cao P, Wang K, Lou Y, Wang W *et al*: **Estimating the serial interval of the novel coronavirus disease (COVID-19): A statistical analysis using the public data in Hong Kong from January 16 to February 15, 2020**. *medRxiv* 2020:2020.2002.2021.20026559.

7.   Nishiura H, Linton NM, Akhmetzhanov AR: **Serial interval of novel coronavirus (2019-nCoV) infections**. *medRxiv* 2020:2020.2002.2003.20019497.

8.   Novel Coronavirus Pneumonia Emergency Response Epidemiology Team: **[The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in China]**. *Zhonghua liu xing bing xue za zhi = Zhonghua liuxingbingxue zazhi* 2020, **41**(2):145-151.

Figure
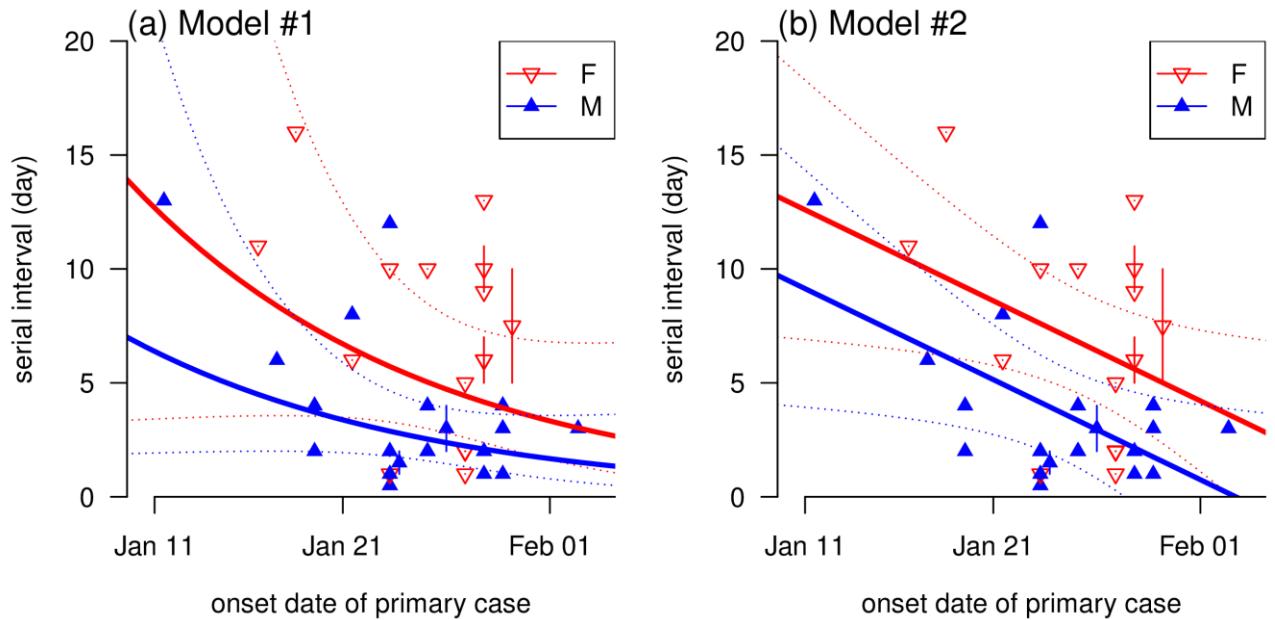
Figure 1.

The observed (dots and bars) and fitted (curves) serial interval (SI) of COVID-19. The results of model (1) are shown in panel (a), and those of model (2) are shown in panel (b). In both panels, the red represents the female primary cases, and blue represents the male primary cases. The dots are the observed (or median) SI, and the bars are the ranges of SI for multiple primary cases. The bold curves are the fitting results, and the dashed curves are the 95% CIs.

Supplementary Material

S1 Key R code for analysis

```
#
require(readxl)
library(survival)
#

SI.data = read_excel(path = 'comb_data.xlsx', sheet = 'comb_data', na = 'NA')
SI.data = as.data.frame(SI.data)

SI.data = SI.data[,1:9]

SI.data$Infector.date.lwr <- as.Date(SI.data$Infector.date.lwr,format ='%m/%d/%Y')

SI.data$Infector.date.upr <- as.Date(SI.data$Infector.date.upr,format ='%m/%d/%Y')

SI.data$Infectee.date <- as.Date(SI.data$Infectee.date,format ='%m/%d/%Y')

SI.data$mid.Infectee.date.index = (as.numeric(SI.data$Infector.date.lwr -as.Date('2019-12-31')) +

as.numeric(SI.data$Infector.date.upr -as.Date('2019-12-31'))) /2


lwr.time.array = NULL

upr.time.array = NULL

surv.code.array = NULL

for (i in 1:nrow(SI.data)) {#      i = 1

 temp.data = SI.data[i,]

 temp.lwr = as.numeric(temp.data$Infectee.date - temp.data$Infector.date.upr)

 temp.upr = as.numeric(temp.data$Infectee.date - temp.data$Infector.date.lwr)

 temp.lwr = ifelse(temp.lwr <=0, 0.5, temp.lwr)

 temp.upr = ifelse(temp.upr <temp.lwr, temp.lwr, temp.upr)

 temp.code = ifelse(temp.upr > temp.lwr, 3, 1)

 lwr.time.array = c(lwr.time.array, temp.lwr)
```

```
  upr.time.array = c(upr.time.array, temp.upr)

  surv.code.array = c(surv.code.array, temp.code)

}

SI.data$lwr.SI = lwr.time.array

SI.data$upr.SI = upr.time.array

SI.data$mid.SI = (SI.data$lwr.SI + SI.data$upr.SI) /2

SI.data$surv.code = surv.code.array


#

sel.data = SI.data

#

cor.test(sel.data$mid.Infectee.date.index, sel.data$mid.SI, method = 'p')

#

simple.lm = lm(log(mid.SI) ~ mid.Infectee.date.index +Infector.gender, data = sel.data)

simple.lm = lm(c(mid.SI) ~ mid.Infectee.date.index +Infector.gender, data = sel.data)

#

surv.obj = Surv(time = sel.data$mid.SI, event = rep(1, nrow(sel.data)))

cx.mod = coxph(surv.obj ~ mid.Infectee.date.index +Infector.gender, data = sel.data)
```

item.index	place	Infector.date.lwr	Infector.date.upr	Infector.gender	Infector.age	Infectee.date	Infectee.gender	Infectee.age	Source
1	HongKong	1/22/2020	1/22/2020	M	60	1/30/2020	F	28	https://www.info.gov.hk/gia/general/202002/05/P2020020500690.htm
2	HongKong	1/22/2020	1/30/2020	NA	NA	2/4/2020	F	56	https://www.info.gov.hk/gia/general/202002/05/P2020020500690.htm
3	HongKong	1/28/2020	1/28/2020	F	57	1/29/2020	M	68	https://www.info.gov.hk/gia/general/202002/09/P2020020900704.htm
4	HongKong	1/28/2020	1/29/2020	NA	NA	1/30/2020	M	24	https://www.info.gov.hk/gia/general/202002/09/P2020020900704.htm
5	HongKong	1/28/2020	1/30/2020	NA	NA	2/1/2020	F	91	https://www.info.gov.hk/gia/general/202002/09/P2020020900704.htm
6	HongKong	1/28/2020	1/28/2020	F	57	1/30/2020	F	25	https://www.info.gov.hk/gia/general/202002/09/P2020020900704.htm
7	HongKong	1/28/2020	1/30/2020	F	NA	2/4/2020	M	22	https://www.info.gov.hk/gia/general/202002/09/P2020020900704.htm
8	HongKong	1/28/2020	1/30/2020	F	NA	2/4/2020	F	51	https://www.info.gov.hk/gia/general/202002/09/P2020020900704.htm
9	HongKong	1/28/2020	1/28/2020	F	57	2/2/2020	F	50	https://www.info.gov.hk/gia/general/202002/09/P2020020900704.htm
10	HongKong	1/28/2020	2/2/2020	F	NA	2/7/2020	M	23	https://www.info.gov.hk/gia/general/202002/09/P2020020900704.htm
11	HongKong	1/28/2020	1/28/2020	F	57	1/30/2020	F	55	https://www.info.gov.hk/gia/general/202002/11/P2020021100018.htm
12	HongKong	1/28/2020	1/30/2020	F	NA	2/8/2020	M	52	https://www.info.gov.hk/gia/general/202002/11/P2020021100018.htm
13	HongKong	1/30/2020	1/30/2020	M	37	2/2/2020	F	37	https://www.info.gov.hk/gia/general/202002/11/P2020021100773.htm
14	HongKong	1/30/2020	1/30/2020	M	37	2/3/2020	F	62	https://www.info.gov.hk/gia/general/202002/11/P2020021100773.htm
15	HongKong	1/30/2020	2/2/2020	NA	NA	2/3/2020	M	75	https://www.info.gov.hk/gia/general/202002/11/P2020021100773.htm
16	HongKong	1/30/2020	1/30/2020	M	37	1/31/2020	F	67	https://www.info.gov.hk/gia/general/202002/13/P2020021300700.htm
17	HongKong	1/30/2020	1/31/2020	NA	NA	2/8/2020	M	37	https://www.info.gov.hk/gia/general/202002/13/P2020021300700.htm
18	HongKong	1/30/2020	1/30/2020	M	37	2/3/2020	F	41	https://www.info.gov.hk/gia/general/202002/14/P2020021400711.htm
19	HongKong	1/28/2020	1/28/2020	F	86	1/30/2020	M	69	https://www.info.gov.hk/gia/general/202002/11/P2020021100018.htm
20	HongKong	1/28/2020	1/28/2020	F	86	1/30/2020	F	63	https://www.info.gov.hk/gia/general/202002/11/P2020021100018.htm
21	HongKong	1/29/2020	1/29/2020	M	43	1/31/2020	M	66	https://www.info.gov.hk/gia/general/202002/13/P2020021300700.htm
22	Shenzhen	1/12/2020	1/15/2020	NA	NA	1/22/2020	F	38	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022872.htm
23	Shenzhen	1/20/2020	1/20/2020	M	73	1/24/2020	M	49	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022873.htm
24	Shenzhen	1/20/2020	1/20/2020	M	73	1/22/2020	F	49	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022874.htm
25	Shenzhen	1/24/2020	1/24/2020	M	64	1/24/2020	M	40	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022875.htm
26	Shenzhen	1/22/2020	1/22/2020	F	30	1/28/2020	F	38	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022876.htm
27	Shenzhen	1/24/2020	1/24/2020	M	61	1/26/2020	F	35	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022877.htm
28	Shenzhen	1/24/2020	1/24/2020	M	61	1/26/2020	M	39	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022878.htm
29	Shenzhen	1/18/2020	1/18/2020	M	64	1/24/2020	M	37	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022879.htm
30	Shenzhen	1/12/2020	1/12/2020	M	69	1/25/2020	F	25	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022880.htm
31	Shenzhen	1/24/2020	1/24/2020	F	59	1/25/2020	F	31	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022881.htm
32	Shenzhen	1/24/2020	1/24/2020	M	51	1/25/2020	M	48	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022882.htm
33	Shenzhen	1/24/2020	1/25/2020	M	NA	1/26/2020	M	34	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022883.htm
34	Shenzhen	1/26/2020	1/26/2020	M	34	1/28/2020	F	34	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022884.htm
35	Shenzhen	1/26/2020	1/26/2020	M	34	1/28/2020	F	4	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022885.htm
36	Shenzhen	1/26/2020	1/26/2020	M	34	1/30/2020	F	50	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022886.htm
37	Shenzhen	1/17/2020	1/17/2020	F	64	1/28/2020	F	37	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022887.htm
38	Shenzhen	1/26/2020	1/26/2020	M	78	1/30/2020	F	78	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022888.htm
39	Shenzhen	1/24/2020	1/24/2020	F	44	2/3/2020	M	13	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022889.htm
40	Shenzhen	1/19/2020	1/19/2020	F	62	2/4/2020	F	33	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022890.htm
41	Shenzhen	1/29/2020	1/30/2020	NA	NA	2/5/2020	F	60	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022891.htm
42	Shenzhen	1/24/2020	1/24/2020	M	61	2/5/2020	M	2	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022892.htm
43	Shenzhen	1/26/2020	1/26/2020	F	50	2/5/2020	F	37	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022893.htm
44	Shenzhen	2/3/2020	2/3/2020	M	54	2/6/2020	M	22	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022894.htm
45	Shenzhen	1/29/2020	1/29/2020	F	58	2/11/2020	M	25	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022895.htm
46	Shenzhen	1/29/2020	1/29/2020	F	58	2/7/2020	F	58	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022896.htm
47	Shenzhen	1/23/2020	1/24/2020	NA	NA	2/1/2020	M	56	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022897.htm
48	Shenzhen	1/29/2020	1/29/2020	M	66	1/30/2020	F	38	http://wjw.sz.gov.cn/yqxx/202002/t20200222_19022898.htm