medRxiv preprint doi: https://doi.org/10.1101/2020.06.09.20127043; this version posted June 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Serial Interval Distribution of SARS-CoV-2 Infection in Brazil

Carlos A. Prete Jr[1], Lewis Buss[2], Victor Bertollo Porto[3], Darlan da Silva Candido[4], Fábio Ghilardi[2], Amy Dighe[5], Christl A. Donnelly[5,6], Oliver G. Pybus[4], Oliver J. Brady[7], Wanderson K. de Oliveira[3], Júlio H. R. Croda[3], Nuno Rodrigues Faria [2,4,5], Ester C. Sabino[2], Vítor H. Nascimento[1].

1. Department of Electronic Systems Engineering, University of São Paulo, Brazil
2. Instituto de Medicina Tropical, University of São Paulo, Brazil
3. Secretaria de Vigilância em Saúde, Ministério da Saúde, Brazil
4. Department of Zoology, University of Oxford, United Kingdom
5. MRC Centre for Global Infectious Disease Analysis, Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, United Kingdom
6. Department of Statistics, University of Oxford, United Kingdom
7. London School of Hygiene and Tropical Medicine, London, United Kingdom

Abstract

Using 65 transmission pairs of SARS-CoV-2 reported to the Brazilian Ministry of Health we estimate the mean and standard deviation for the serial interval to be 2.97 and 3.29 days respectively. We also present a model for the serial interval probability distribution using only two parameters.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.09.20127043; this version posted June 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Text

The severe acute respiratory syndrome novel coronavirus 2 (SARS-CoV-2) first emerged in China in December 2019 [1] and was declared a pandemic by the World Health Organization on 11th March 2020. The transmission dynamics, such as the serial interval – the time between symptom onset in a primary case and their secondary cases in a chain of transmission [2] – are still being established. The epidemic growth rate and the serial interval jointly determine the basic reproduction number [3], and the existence of pre-symptomatic transmission can be inferred when the serial interval is shorter than the incubation period.

Data for this analysis were provided by the Brazilian Ministry of Health following ethical approval (CONEP protocol number 30127020.0.0000.0068). All confirmed cases of SARS-CoV-2 infection notified on the REDCap system between 25th February and 19th March were analyzed. We received the linked pairs data (suspected and primary case) by the Brazilian Ministry of Health. We measured the serial intervals by computing the difference of dates of symptom onset.

In our dataset there were four negative serial intervals (symptom onset in the infectee preceded the infector) and seven zero-valued serial intervals (symptom onset on the same day for the infector-infectee pair). The mean (standard deviation) was 2.97 (3.29) days with a median of 3 days.

It is common to censor observed serial intervals to include only positive values; however, there is no strong theoretical justification for this. As such, we fit a serial interval probability distribution that allows for negative values by modelling the serial interval as

$$\text{Serial Interval} = \lfloor t_{\text{inf},1} + \Delta t_{\text{lat},1} + \Delta t_{\text{lat},2} + \Delta t_{\text{int},2} \rfloor - \lfloor t_{\text{inf},1} + \Delta t_{\text{lat},1} + \Delta t_{\text{int},1} \rfloor,$$

medRxiv preprint doi: https://doi.org/10.1101/2020.06.09.20127043; this version posted June 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

where $\lfloor \cdot \rfloor$ is the floor operator, $t_{\mathrm{inf},1}$ is a random variable uniformly distributed in the interval [0,1], $(\Delta t_{\mathrm{lat},1},\ \Delta t_{\mathrm{lat},2})$ are independent and identically distributed (i.i.d.) chi-squared random variables of mean 2.79 days and $(\Delta t_{\mathrm{int},1}, \Delta t_{\mathrm{int},2})$ are i.i.d. chi-squared random variables of mean 1.31 days. See the **Supplemental Material** for theoretical motivation and method.

Figure 1 shows the comparison between the measured and modelled serial intervals as well as means $\pm$ standard deviations for a number of serial interval probability distributions presented in the literature [4–9]. We conclude that our proposed model for the serial interval probability distribution approximates well the serial intervals measured from data.

In addition, we provide two other analyses of the data in the Supplemental Material in which we i.) fit a normal distribution, and ii.) censor the data to include only positive values, best fit by a lognormal distribution.

The mean serial interval in Brazil of 2.97 days is the shortest reported to our knowledge, but we emphasize that the serial interval is not usually concentrated around the mean due to its high standard deviation (3.29 days) and non-symmetry. Du et al. [4] took an alternative approach fitting a normal distribution to 468 publicly reported serial intervals in which 12% had negative values – see **Supplemental Material** for comparison with our raw data. 8.38% of the measured serial intervals are negative, which is allowed by our model. Negative serial intervals are described in the literature [2] and can occur with pre-symptomatic transmission, and a wide range for the incubation period – the classical example being HIV.

In two transmissions clusters in Tianjin and Singapore the mean incubation periods were 9.0 and 7.1 days, respectively [5]. Our serial interval estimate is shorter than the reported incubation period, implying that the infectious period for SARS-CoV-2 begins before symptom onset. This

medRxiv preprint doi: https://doi.org/10.1101/2020.06.09.20127043; this version posted June 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

is in line with recent observations of a large proportion of undocumented infections, complicating containment of the virus [10].

Our estimates are based on national case notification data, and this is a strength compared to datasets complied solely from media reports. However, there are at least two important sources of bias. Firstly, there is a tendency for secondary cases to recall more recent contacts – recency or recall bias – resulting in a shorter estimate of the serial interval. Similarly, contact tracing is likely to be most effective for recent contacts. Secondly, self-isolation following symptom onset will remove longer serial intervals that would have occurred due to transmission during the symptomatic phase. This second point highlights the role of contextual factors – isolation practices, population density, location of transmission, proportion of the population infected etc. – in determining the serial interval, which is not simply a biological constant of the virus. Given this consideration, it is important generate multiple estimates in different locations and stages of the epidemic. Our report is valuable as the first estimate from Brazil, and Latin America more generally.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.09.20127043; this version posted June 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

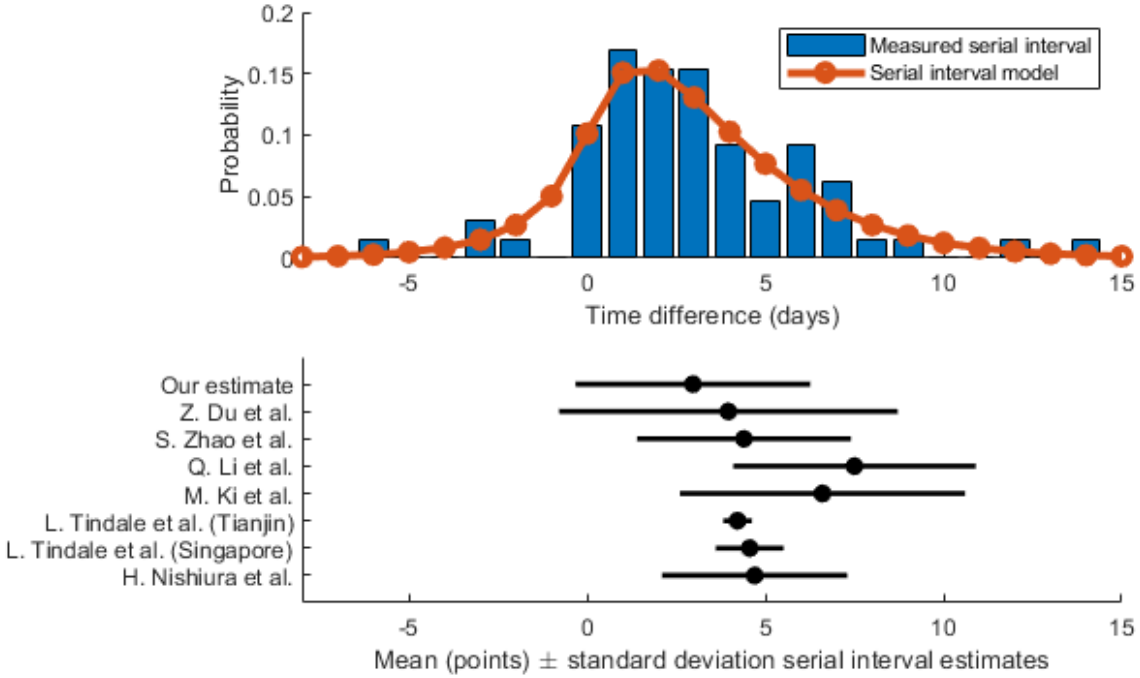

Figure 1. Upper panel – Serial interval estimation from Brazilian SARS-CoV-2 confirmed cases compared to the modelled distribution. Lower panel – mean ± standard deviation serial interval estimates from the current literature.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.09.20127043; this version posted June 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Acknowledgements

This work was supported by a Medical Research Council (MR/S0195/1) and FAPESP (2018/14389-0) CADDE partnership award and a John Fell Research Fund (grant 005166). NRF is supported by a Wellcome Trust and Royal Society Sir Henry Dale Fellowship (204311/Z/16/Z). DDSC is supported by the Clarendon Fund and by the Oxford University Zoology Department. CAPJ and VHN are supported in part by FAPESP grant 2018/12579-7. CAD is thankful for Centre funding from the UK Medical Research Council (MRC) and the UK Department for International Development (DFID) under the MRC/DFID Concordat agreement and is also part of the EDCTP2 programme supported by the European Union (MR/R015600/1). AD is also supported by the UK Medical Research Council (MRC).

Conflict of Interest

The authors declare no conflict of interest.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.09.20127043; this version posted June 12, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

References

1. WHO - Pneumonia of unknown cause – China. Available at: https://www.who.int/csr/don/05-january-2020-pneumonia-of-unkown-cause-china/en/.

2. Fine PEM. The Interval between Successive Cases of an Infectious Disease. American Journal of Epidemiology **2003**; 158:1039–1047.

3. Wallinga J, Lipsitch M. How generation intervals shape the relationship between growth rates and reproductive numbers. Proc R Soc B **2007**; 274:599–604.

4. Du Z, Xu X, Wu Y, Wang L, Cowling BJ, Meyers LA. The serial interval of COVID-19 from publicly reported confirmed cases. Epidemiology, 2020. Available at: http://medrxiv.org/lookup/doi/10.1101/2020.02.19.20025452. Accessed 24 March 2020.

5. Tindale L, Coombe M, Stockdale JE, et al. Transmission interval estimates suggest pre-symptomatic spread of COVID-19. Epidemiology, 2020. Available at: http://medrxiv.org/lookup/doi/10.1101/2020.03.03.20029983. Accessed 24 March 2020.

6. Nishiura H, Linton NM, Akhmetzhanov AR. Serial interval of novel coronavirus (2019-nCoV) infections. Infectious Diseases (except HIV/AIDS), 2020. Available at: http://medrxiv.org/lookup/doi/10.1101/2020.02.03.20019497. Accessed 24 March 2020.

7. Li Q, Guan X, Wu P, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. N Engl J Med **2020**; :NEJMoa2001316.

8. Ki M, Task Force for 2019-nCoV2. Epidemiologic characteristics of early cases with 2019 novel coronavirus (2019-nCoV) disease in Republic of Korea. Epidemiol Health **2020**; Available at: http://e-epih.org/journal/view.php?doi=10.4178/epih.e2020007. Accessed 13 April 2020.

9. Zhao S, Gao D, Zhuang Z, 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. Epidemiology, 2020. Available at: http://medrxiv.org/lookup/doi/10.1101/2020.02.21.20026559. Accessed 24 March 2020.

10. Li R, Pei S, Chen B, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). Science **2020**; :eabb3221.