medRxiv preprint doi: https://doi.org/10.1101/2020.06.28.20142133; this version posted July 24, 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 The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. All rights reserved. No reuse allowed without permission.

Superspreading in Early Transmissions of COVID-19 in Indonesia

Agus Hasan[1,*], Hadi Susanto[2,3], Muhammad Firmansyah Kasim[4], Nuning Nuraini[5], Bony Lestari[6], Dessy Triany[7], and Widyastuti Widyastuti[8]

[1]Mærsk McKinney Møller Institute, University of Southern Denmark, Denmark
[2]Department of Mathematics, Khalifa University, United Arab Emirates
[3]Department of Mathematical Sciences, University of Essex, United Kingdom
[4]Department of Physics, University of Oxford, United Kingdom
[5]Department of Mathematics, Institut Teknologi Bandung, Indonesia
[6]Department of Epidemiology and Biostatistics, Universitas Padjadjaran, Indonesia
[7]Public Health Office of Batam City, Indonesia
[8]Public Health Office of Jakarta Province, Indonesia
[*]agha@mmmi.sdu.dk

ABSTRACT

We estimate the basic reproduction number $\mathscr{R}_0$ and the overdispersion parameter $\mathscr{K}$ at two regions in Indonesia: Jakarta-Depok and Batam. Based on the first 1288 confirmed cases in both regions, we find a high degree of individual-level variation in the transmission. The basic reproduction number $\mathscr{R}_0$ is estimated at 6.79 and 2.47, while the overdispersion parameter $\mathscr{K}$ of a negative-binomial distribution is estimated at 0.06 and 0.2 for Jakarta-Depok and Batam, respectively. This suggests that superspreading events played a key role in the early stage of the outbreak, i.e., a small number of infected individuals are responsible for large amounts of COVID-19 transmission.

Introduction

The first two confirmed cases of COVID-19 in Indonesia was announced in March 2, 2020 by the president himself. Jakarta and its buffer zones including Depok have become the epicenter of the outbreak. In its early transmission, about 50% of the cases were from the cities[1]. The virus quickly spread to all 34 provinces, with more than 93,000 confirmed cases and 4,500 deaths as of July 23, 2020. As a country with the highest death toll in South East Asia and one with the lowest global testing rate, the first government effort to control the disease spread was by establishing a COVID-19 response acceleration task force led by the head of the National Disaster Management Agency. The government further introduced Large-Scale Social Restrictions, which included measures such as closing public places, restricting public transport, and limiting travel to and from restricted regions. On May 19, the first milestone of 10,000 PCR tests per day was reached. In June 2020, WHO announced that only Jakarta meets the minimum requirement of 1 test per 1,000 population per week for a reliable positivity rate calculation[2].

As the Indonesian government still struggles to mitigate the COVID-19 epidemic in the country, it becomes critical to understand its infection spreads. Here, we present a first quantitative analysis of early transmission dynamics of the disease and evidence of superspreading events in Indonesia based on contact tracing of secondary cases. Infection clusters are difficult to predict and thus difficult to prevent. Nevertheless, understanding their environmental and behavioral drivers has been recently recognized to be extremely important to inform strategies for prevention and control[3] as the reproduction number only is no longer enough[4]. In our analysis of Jakarta, we combine data from Depok that is a main buffer city of the capital. Outside the epicentre, we present data from Batam that has received thousand test kits and protective equipment from Singapore as a comparison. There were 218 cases from 4,092 PCR tests, as of June 26[5].

Results

Based on 1288 confirmed cases in March and April 2020 in the regions Jakarta-Depok[1] (population 12.6 millions) and Batam[5] (population 1.4 millions), we find a high degree of individual-level variation in the transmission. Early in the outbreak, superspreaders transmit the majority of coronavirus cases, as can be seen from Figure 1. Defining a superspreader as an individual with at least eight transmissions of the disease, there were 44 clusters in Jakarta-Depok region (2 - 31 March, 2020) and 3 clusters in Batam city (19 March to 7 April, 2020), as can be seen from Table 1.

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/10.1101/2020.06.28.20142133; this version posted July 24, 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. All rights reserved. No reuse allowed without permission.

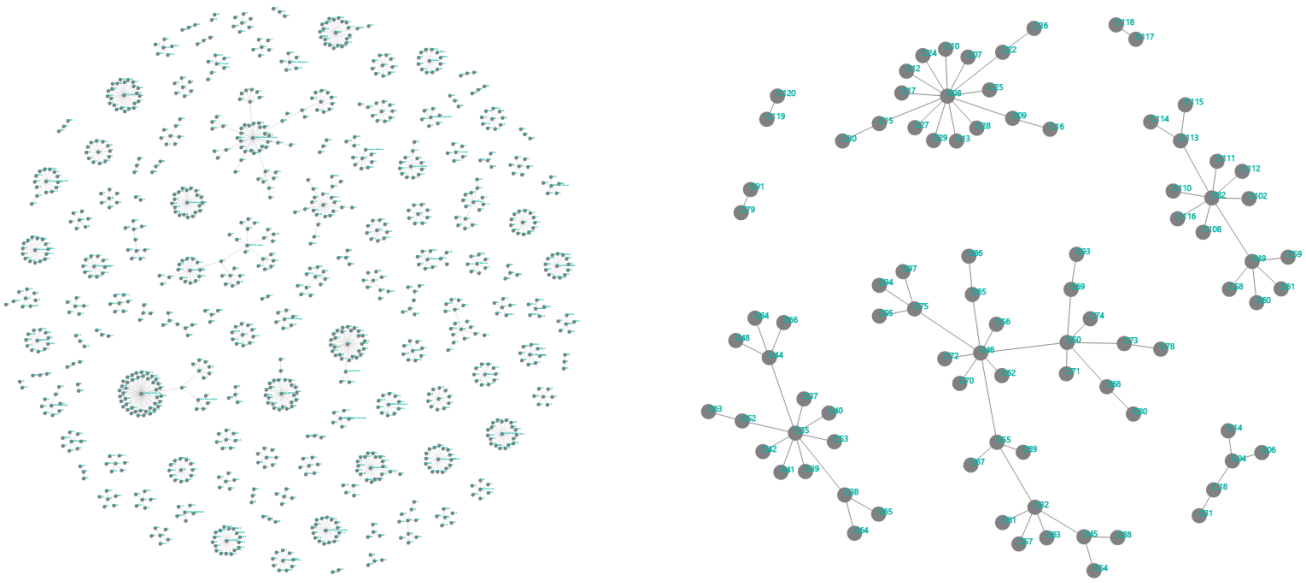

Figure 1. Networks of secondary cases based on local transmissions in Jakarta-Depok (left) and Batam (right) regions.

Jakarta-Depok		Batam	
Secondary cases	Frequency	Secondary cases	Frequency
0	1017	0	63
1	31	1	12
2	36	2	3
3,6	17	3	4
4,13	4	4,7-9,13	1
5	22	5	2
7	11		
8,11,12	3		
9	8		
10	7		
14,17,20,21,26,27,36,43	1		
15,16,19,23	2		
Total	1199	Total	89

Table 1. Secondary cases data of the two regions based on local transmissions.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.28.20142133; this version posted July 24, 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. All rights reserved. No reuse allowed without permission.

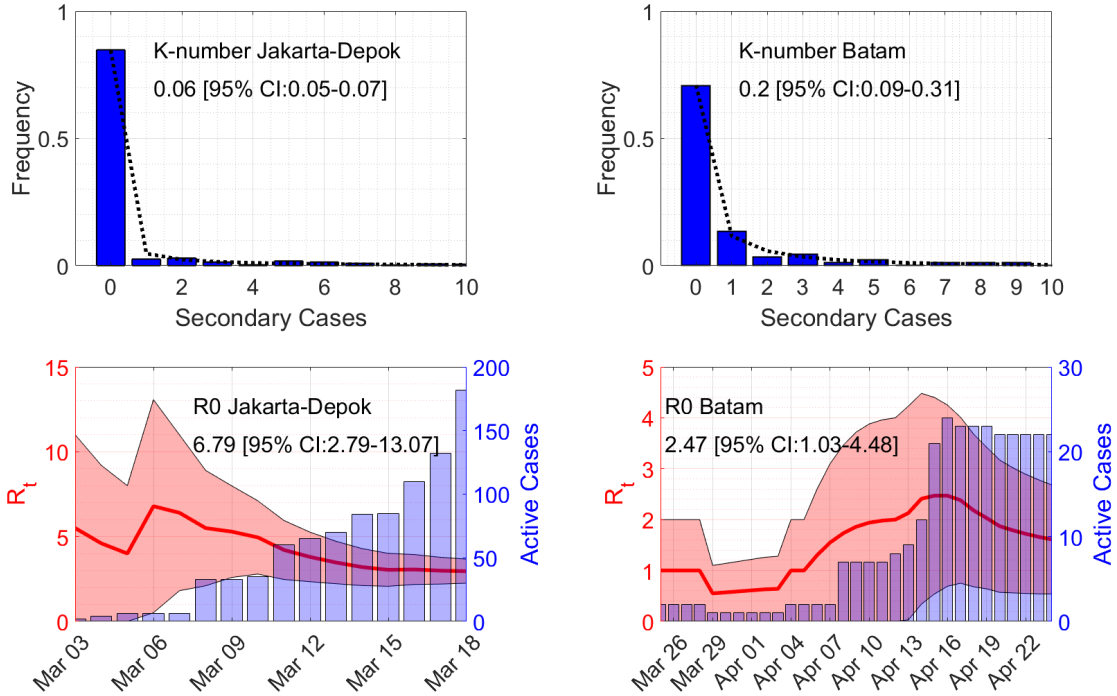

Figure 2. Estimates of the overdispersion parameter $\mathscr{K}$ and the basic reproduction number $\mathscr{R}_0$. Lines in the top figures show maximum-likelihood fits for the negative binomial distribution.

We estimated the reproduction number $\mathscr{R}_0$ to be at 6.79 and 2.47, while the overdispersion parameter $\mathscr{K}$ of a negative-binomial distribution is obtained at 0.06 and 0.2 for Jakarta-Depok and Batam, respectively (see Figure 2), suggesting that a small number of infected individuals are responsible for large amounts of the disease transmission. Indeed, between 10-15% of all infections were responsible for 80% of onward transmission events.

The fact that Indonesia has a moderate number of positive cases, compared to other countries with roughly the same population size such as Russia, USA, and Brazil, can be caused by three reasons: a low number of tests, the success of Large-Scale Social Restrictions, and overdispersion of the COVID-19 transmission. Our result clearly indicates that the transmission is overdispersed, even though it does not exclude the other possibilities. Therefore, close-interaction activities such as traditional markets, religious gathering, and wedding parties need to be adapted if not restricted as they can become transmission hot spots. Effective response factors to reduce the transmission include aggressive implementation of non-pharmaceutical interventions such as rapid identification and isolation of cases. Furthermore, timeliness is critical to prevent or limit their extent since delay of diagnosis is the most common cause of superspreading clusters[6].

Methods

We use a sequential Bayesian method based on the discrete SIR model to estimate the basic reproduction number $\mathscr{R}_0$[7]. Time series of the active cases at day n is assumed to follow the model $I_{n+1} = (\gamma(\mathscr{R}_t - 1) + 1) I_n$, where $\mathscr{R}_t$ is the instantaneous reproduction number and $1/\gamma$ is the infectious time taken to be 9 days herein[8]. Assuming a Poisson distribution for the arrival of active cases and using Bayes' theorem, the posterior probability of $\mathscr{R}_t$ given a confirmed case I_{n+1} is computed through

$$P\left((\gamma(\mathscr{R}_t - 1) + 1) I_n | I_{n+1}\right) \propto \prod_{m=1}^{n} P\left(I_{m+1} | (\gamma(\mathscr{R}_t - 1) + 1) I_m\right).$$

The formula comes from the assumption that the prior distribution of $\mathscr{R}_t$ at day n is taken from the posterior of day $(n-1)$ with a uniform prior at day 1. Estimated $\mathscr{R}_0$ is then the maximum of $\mathscr{R}_t$, see Figure 2.

The parameter $\mathscr{K}$ is then defined as the overdispersion parameter in the negative binomial distribution[9], which can be found by fitting the distribution to the secondary cases data. The number of secondary cases by one person x can be expressed

medRxiv preprint doi: https://doi.org/10.1101/2020.06.28.20142133; this version posted July 24, 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. All rights reserved. No reuse allowed without permission.

in terms of the overdispersion parameter $\mathscr{K}$ and the probability p,

$$P(x|\mathscr{K},p) = \frac{\Gamma(x+\mathscr{K})}{x!\Gamma(\mathscr{K})}(1-p)^{\mathscr{K}}p^x.$$

We use a MATLAB® function *nbinfit* to estimate the overdispersion parameter $\mathscr{K}$. The function uses derivative-free method to find the maximum of the likelihood function.

Data and Code Availability

Secondary cases data (both in Excel and Power BI) and a MATLAB® code to produce Figure 1 and Figure 2 are available at: https://github.com/agusisma/covidindonesia.

References

1. Jakarta City Government. Data covid-19. *Web Reports* (2020).
2. WHO Indonesia Situation Report – 11. Coronavirus disease 2019 (covid-19) situation report 11. *WHO, 10 June 2020* 8 (2020).
3. Frieden, T. & Lee, C. Identifying and interrupting superspreading events–implications for control of severe acute respiratory syndrome coronavirus 2. *Emerg Infect Dis.* **26**, 1059–1066, DOI: 10.3201/eid2606.200495 (2020).
4. Adam, D. A guide to r – the pandemic's misunderstood metric. *Nature* **583**, 346–348, DOI: 10.1038/d41586-020-02009-w (2020).
5. Batam City Government. Data covid-19. *Web Reports* (2020).
6. Lloyd-Smith, J., Schreiber, S., Kopp, P. & Getz, W. M. Superspreading and the effect of individual variation on disease emergence. *Nature* **438**, 355–359, DOI: 10.1038/nature04153 (2005).
7. Susanto, H. *et al.* How many can you infect? simple (and naive) methods of estimating the reproduction number. *Commun. Biomath. Sci.* **3**, 28–36 (2020).
8. To, K. K.-W. *et al.* Temporal profiles of viral load in posterior oropharyngeal saliva samples and serum antibody responses during infection by sars-cov-2: an observational cohort study. *The Lancet Infect. Dis.* **20**, 565–74, DOI: 10.1016/S1473-3099(20)30196-1 (2020).
9. Lloyd-Smith, J. Maximum likelihood estimation of the negative binomial dispersion parameter for highly overdispersed data, with applications to infectious diseases. *PLoS ONE* **2**, e180, DOI: 10.1371/journal.pone.0000180 (2007).

Acknowledgements

We would like to thank Public Health Office of Jakarta Province, Communication, Information and Statistic Office of Jakarta Province, Jakarta Smart City Unit, Human Settlements and Spatial Planning Office of Jakarta Province, Public Health Offices of Batam, and Depok Task Force for COVID-19 Control for providing the data, feedback, and discussions.

Author contributions statement

A.H., H.S., and M.F.K. analysed the data, calculated $\mathscr{R}_0$ and $\mathscr{K}$, and discussed the results, N.N., B.L., D.T., and W.W. collected the data. All authors reviewed the manuscript.