1 Estimating SARS-CoV-2 variant strain infectiousness in Japan as of March 28, 2021

2

3

4 Junko Kurita¹⁾, Tamie Sugawara²⁾, Yasushi Ohkusa²⁾

5

6 1) Department of Nursing, Tokiwa University, Ibaraki, Japan

7 2) National Institute of Infectious Diseases, Tokyo, Japan

8

9

10 Corresponding author: Junko Kurita, kuritaj@tokiwa.ac.jp

11

12

13 Keywords: infectiousness, reproduction number, SARS-CoV-2, variant strain,

14 VOC202012/01

15

16

17

18 Abstract

19 Background: A variant strain of SARS-CoV-2, VOC202012/01, emerged in the UK
20 in September, 2020. Its infectiousness was estimated as higher than that of the original
21 strain.

22 Object: We estimated the infectiousness of the variant strain of SARS-CoV-2 in
23 comparison to that of the original strain under conditions prevailing in Japan.

24 Methods: We estimated infectiousness through a simple susceptible–infected–recovered
25 (SIR) model by strain. The study period was March 1–28, 2020. The information used
26 for the study was available as of April 3, 2020.

27 Results: The estimated reproduction number of the SARS-CoV-2 variant strain was
28 1.799; its 95% CI was [1.642, 1.938]. The onset date of the first case in Japan was
29 estimated as December 4 ([November 16, December 14]), 2020. However,
30 infectiousness of the original strain was estimated as 1.123 ([1.093, 1.166]).

31 Discussion and Conclusion: We demonstrated that infectiousness increased by 0.684 or
32 60%, increasing from the original to the variant. That finding might be comparable to
33 that of a study conducted in the UK. However, the difference must be monitored
34 continuously and carefully.

35

36 **Introduction**

37

38 A variant strain of SARS-COV-2, VOC202012/01, emerged in the UK in
39 September 2020. Its infectiousness and pathogenicity were estimated as 75% higher
40 than those of the original strain, which was previously circulating before emerging
41 variant strain[1].

42 In Japan, the first case was detected on January 6 as an imported case from the UK.
43 Subsequently until the end of March, 2021, community outbreak of the variant strain
44 occurred in Japan. Since the beginning of March, 2021, the outbreak was remarkable in
45 Japan. Actually, the Ministry of Health, Labour and Welfare reported prevalence of the
46 variant strain in the community outbreak as 7, 10, 16, and 20% for each week of March
47 in Japan [2].

48 However, the estimated infectiousness of the variant strain was not evaluated
49 outside Europe. Even for the outbreak caused by the original strain, outbreak situations
50 differed among countries. For example, prevalence of SARS-COV-2, with mainly the
51 original strain, in the USA were 9361 and 6558 cases in the UK per 100 thousand
52 residents, but in Japan it was limited to 372 as of the end of March, 2021 [3].
53 Infectiousness of the variant strain might differ country-by-country, as that of the

original strain has. Therefore, we evaluated the infectiousness of the variant strain compared to that of the original strain under the same conditions using a simple deterministic SIR model in Japan.

Methods

We defined the variant strains as mutated at N501Y and/or E484K together because the numbers of patients were not reported by subtype among variant strains. Therefore, we cannot specifically examine differences among variant strains.

For a simple deterministic SIR model [4–6] applied to the epidemic curve for Japan’s 120 million populations, we used the numbers of symptomatic patients reported by strain from MHLW for March 1–28, 2021, which were published [7] on April 4, 2021. We excluded patients who were presumed to be infected abroad because those patients presumably do not represent cases of community-acquired infection in Japan. We assumed that the two strains had similar properties, except for infectiousness. Because the SIR model is a system reflecting daily differences, we converted weekly data simply to daily data divided by 7.

The incubation period is assumed to conform to the empirical distribution in Japan.

72 Letting the probability of onset on day k after exposure in the empirical distribution be
 73 $p(k)$, then $p(k)/\prod_{m=0}^{k-1}(1-p(m))$ of infected patients showed no symptom until $k-1$ days
 74 from exposure showed onset on day k . That is, $1-p(k)/\prod_{m=0}^{k-1}(1-p(m))$ of them showed
 75 onset on $k+1$ day after exposure, or later.

76 For symptomatic patients with unknown onset date, we estimated the onset date as
 77 follows: Letting $f(k)$ represent the probability of a k day delay in the empirical
 78 distribution of reporting delay from onset and letting N_t denote the number of patients
 79 for whom onset dates were not available for publication at date t , then the number of
 80 patients for whom the onset date was known is $t-1$. The number of patients for whom
 81 onset dates were not available was estimated as $f(1)N_t$. Similarly, the number of patients
 82 with onset date $t-2$ and for whom onset dates were not available were estimated as
 83 $f(2)N_t$. Therefore, the total number of patients for whom the onset date was not available,
 84 given an onset date of s , was estimated as $\sum_{k=1} f(k)N_{s+k}$ for the long duration extending
 85 from s .

86 Moreover, the reporting delay from the Ministry of Health, Labour and Welfare
 87 (MHLW) might be considerable. In other words, if $s+k$ is larger than that in the current
 88 period t , then $s+k$ represents the future for period t . For that reason, N_{s+k} is not
 89 observable. Such a reporting delay engenders underestimation bias of the number of

90 patients. For that reason, it must be adjusted as $\sum_{k=1}^{t-s} f(k) N_{s+k} / \sum_{k=1}^{t-s} f(k)$. Similarly,
 91 patients for whom the onset dates were available are expected to be affected by the
 92 reporting delay. Therefore, $M_{s|t} / \sum_{k=1}^{t-s} f(k)$, where $M_{s|t}$ represents the reported number of
 93 patients for whom onset dates were within period s , extended until the current period t .
 94 We used these arranged data to produce the epidemic curve for SIR modelling.

95 The distribution of infectiousness in symptomatic cases was assumed to be 30% on
 96 the onset day, 20% on the following day, and 10% for the subsequent 5 days [8]. If
 97 designating the distribution of infectiousness as $g(k)$, where k represents days from
 98 onset and the number of symptomatic patients on day t and onset was k days ago as
 99 $I(t,s)$, then the number of newly infected patients on day t was $R(t) S(t) \sum_{k=0} g(k) I(t,k)$
 100 $/ Y(t)$, where $S(t)$ denotes the number of susceptible population and $Y(t)$ represents the
 101 total population on day t .

102 We run the simulation from the initial case under the presumed reproduction
 103 number. Simulation results of the epidemic curve, which was defined as the number of
 104 patients by onset date, was converted to the predicted reporting number of
 105 laboratory-confirmed patients by the empirical distribution of reporting delay from
 106 onset. To that was added a constant proportion of asymptomatic patients, 20%,
 107 assuming no infectiousness among asymptomatic patients [8].

108   If the number of the estimated reported laboratory-confirmed patients on a day in
109   the simulation was higher than the actual data on March 1, then we compared the model
110   prediction and observation for four weeks to seek reproduction number by types to
111   achieve the minimum absolute difference among them through grid search on [1,2] by
112   0.0001. The bootstrapping procedure we used was fully replicated bootstrapping for a
113   fixed number of initial cases. There were N patients in the data, with numbering of the
114   patients from the initial case to the last case. Initially, no patient was on the
115   bootstrapped epidemic curve. If a random variable drawn from a uniform distribution of
116         $(0,1)$ was t included in the interval $[i/(N-1), (i+1)/(N-1)]$, then we added one to the onset
117   date of the $i+1$ th patient to the bootstrapped epidemic curve. We replicated this
118   procedure $N-1$ times. Thereby, we obtain the bootstrapped epidemic curve with $N-1$
119   patients. Finally, we added the initial patient, for whom the onset date was January 14,
120   to the bootstrapped epidemic curve. Consequently, we obtained a bootstrapped
121   epidemic curve with N patients.

122   The study period was March 1–28, 2021 for estimating the infectiousness of the
123   two strains. The information used for the study was available on April 4, 2021. The
124   empirical distribution of the incubation period and reporting delay were based on the
125   outbreak in Japan from January 14, 2020 to the end of February, 2021.

126

127 **Results**

128

129   In the four weeks I March 2021, 7264, 7940, 8960 and 12045 laboratory- confirmed
130 cases including variant strain cases as well as the original strain cases were reported [7].
131 Multiplying proportion of the variant strain described above to these number, the
132 weekly number of reported laboratory-confirmed patients by strain in March, 2021 were
133 shown in Figure 1. One seventh of these data were used for the following analysis as
134 daily data.

135   Figure 2 presents an empirical distribution of the duration of onset to reporting in
136 Japan. The maximum delay was 31 days. Figure 3 depicts the empirical distribution of
137 incubation periods among 125 cases for which the exposed date and onset date were
138 published by MHLW in Japan. The mode was 6 days. The average was 6.6 days.

139   The estimated reproduction number of the variant strain of SARS-CoV-2 was
140 1.799; its 95% CI was [1.642, 1.938]. The onset date of the first case in Japan was
141 estimated as December 4 ([November 16, December 14]), 2020. However, the
142 infectiousness of the original strain was estimated as 1.123 ([1.093, 1.166]). Figures 4
143 and 5 show the fitted line and its 95% CI of the number of laboratory-confirmed cases

144 by the model for the two strains. Although stepwise data were difficult to mimic by the
145 model, bands of CI fairly covered the observation data.

146

147 **Discussion**

148

149 Similar differences of infectiousness were found for the variant strain and the
150 original strain, as described for an earlier study [1]. The infectiousness increased by
151 0.684, indicating a 60% increase from the original to the variant. The rate of increase
152 was slightly lower than that presented in an earlier report [1]. It might reflect the
153 outbreak situation including vaccine coverage.

154 The vaccine coverage in March, 2021 in Japan was less than 1%. Therefore, it
155 probably did not affect the outbreak at all. However, with increasing vaccine coverage
156 in Japan, differences among strains are not expected to be affected if the vaccine
157 effectiveness is equal for all strains. Alternatively, differences among strains will
158 become increasingly evident if vaccine effectiveness against variant strains is less than
159 the effectiveness against the original strain.

160 We estimated the onset date of the initial case of the community outbreak by a

161 variant strain as December 4 ([November 16, December 14]), 2020. Actually, the first
162 case of the variant strain was detected on January 6, 2021 as an imported case from UK.
163 Therefore, we predicted the initial case of the variant strain as earlier than one month
164 from the officially reported imported case. As one might expect, the earlier stage of the
165 outbreak in the variant strain was not recognized in December, 2020 and January, 2021
166 because viral genomic sequencing, which can detect a variant strain, was not applied to
167 the specimen from the community outbreak in Japan. Therefore, we cannot deny the
168 possibility that a community outbreak by the variant strain occurred earlier than
169 December, 2020, in Japan. The earliest case with the variant strain in the UK was
170 reportedly identified on September 20, 2020 [1]. Therefore, no community outbreak was
171 likely to have started in Japan two months later than in the UK.

172 The present study has some limitations. First, because of data limitations, some
173 differences among variant strains were not considered. However, examination of those
174 differences might be important for predicting which subtype of variant strain is
175 dominant in the outbreak.

176 Secondly, the obtained result reflects information up to March, 2021. Changes in
177 climate or mobility might have affected the results. In addition, as described above,
178 increasing vaccine coverage might eventually affect results. Therefore, it is necessary to

179 monitor differences of infectiousness among strains carefully and continuously.

180

181 **Conclusion**

182

183 We demonstrated that infectiousness of the variant was greater by 0.684, or 60%,
184 compared to the original strain of COVID-19. The findings might be comparable to
185 those of a study conducted in the UK [1]. Nevertheless, careful and continuous
186 monitoring of differences among strains is necessary.

187 The present study is based on the authors' opinions. Its results do not reflect any
188 stance or policy of their professionally affiliated bodies.

189

190 **Acknowledgments**

191

192 We acknowledge the great efforts of all staff at public health centers, medical
193 institutions, and other facilities who are fighting the spread and destruction associated
194 with COVID-19.

195

196   **References**

- 197   1.   Leung K, Shum MHH, Leung GM, Lam TTY, Wu JT. Early transmissibility
198   assessment of the N501Y mutant strains of SARS-CoV-2 in the United Kingdom,
199   October to November 2020. Euro Surveill 2021;26:2002106. doi:
200   10.2807/1560-7917.ES.2020.26.1.2002106.
- 201   2.   Ministry of Health, Labour and Welfare, Implementation rate and positive rate of
202   screening test for the variant strain (mechanical estimation) preliminary figures.
203   <https://www.mhlw.go.jp/content/000766381.pdf> (in Japanese) [accessed on April 9,
204   2021]
- 205   3.   World Health Organization. Coronavirus disease (COVID-19) Weekly
206   Epidemiological Update and Weekly Operational Update
207   <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-report>
208   [s](#) [accessed on April4, 2021]
- 209   4.   Kurita J, Sugawara T, Ohkusa Y. Estimated effectiveness of school closure and
210   voluntary event cancellation as COVID-19 countermeasures in Japan. J Infect
211   Chemother 2021, 27:62-4. doi: 10.1016/j.jiac.2020.08.012. Epub 2020 Aug 19.
- 212   5.   Sugishita Y, Kurita J, Sugawara T, Ohkusa Y. Effects of voluntary event
213   cancellation and school closure as countermeasures against COVID-19 outbreak in

- 214 Japan. Plos one 2020. <https://doi.org/10.1371/journal.pone.0239455>
- 215 6. Ohkusa Y, Sugawara T, Taniguchi K, Okabe N. Real-time estimation and
- 216 prediction for pandemic A/H1N1(2009) in Japan. J Infect Chemother. 2011; 17:
- 217 468-472.
- 218 7. Japan Ministry of Health, Labour and Welfare. Press releases. Available from:
- 219 https://www.mhlw.go.jp/stf/newpage_10723.html (in Japanese) [accessed on April
- 220 4, 2021]
- 221 8. Cao S, Gan Y, Wang C, Bachmann M, Wei S, Gong J, Huang Y, Wang T, Li L, Lu
- 222 K, Jiang H, Gong Y, Xu H, Shen X, Tian Q, Lv C, Song F, Yin X, Lu Z.
- 223 Post-lockdown SARS-CoV-2 nucleic acid screening in nearly ten million residents
- 224 of Wuhan, China. Nat Commun 2020;11:5917. doi: 10.1038/s41467-020-19802-w.
- 225

1 Fig. 1. Number of laboratory confirmed cases infected by original and variant strain of
2 SARS-CoV-2 in March, 2021 reported by MHLW, Japan.
3 (Cases)

4
5 (Days in March, 2021)

6 Note: Orange bars represent numbers of laboratory-confirmed cases infected by the
7 original strain weekly. Blue bars represent patients infected by the variant strain.

8 Laboratory-confirmed cases included asymptomatic and presymptomatic cases. Dates
9 are the reported date: not the onset or infected date.

1 Fig. 2. Empirical distribution of durations from onset to report by MHLW, Japan. (%)

2

3 Note: Bars represent the probabilities of duration from onset to report based on data of

4 657 patients in Japan for whom onset dates were available. The horizontal axis shows

5 the duration from exposure to onset in days. The vertical axis shows the probability of

6 the incubation period. Data were obtained from MHLW, Japan.

1 Fig. 3. Empirical distribution of the incubation period published by MHLW, Japan. (%)

2 (Cases)

3

4

(Days)

5 Notes: Bars show the distribution of incubation periods for 125 cases for which the

6 exposure date and onset date were published by MHLW, Japan. The horizontal axis was

7 measured in duration from exposure to onset in days. The vertical axis was the

8 probability of the incubation period. Patients for whom incubation was longer than 14

9 days are included in the bar shown for day 14.

1 Fig. 4. Observed data of the original strain and prediction with 95% CI

2 (Number of cases)

3
4 (Days in March, 2021)

5 Note; Bars represent the reported number of the test confirmed cases. The bold line

6 represents its prediction by the model. Thin lines represent its 95% CI.

1 Fig. 5. Observed data of the variant strain and its prediction with 95% CI.

2 (Number of cases)

4 (Days in March, 2021)

5 Note: Bars represent the reported numbers of the test confirmed cases. Bold line

6 represents its prediction by the model. Thin lines represent 95% CI.