

**Title:**

Prediction of severe COVID-19 cases requiring a ventilator in Tokyo, Japan

**Authors:**

Junko Kurita<sup>1</sup>, Tamie Sugawara<sup>2</sup>, Yasushi Ohkusa<sup>2</sup>

**Affiliation:**

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

2) National Institute of Infectious Diseases, Tokyo, Japan

**Corresponding author:** Junko Kurita      kuritaj@tokiwa.ac.jp

**Keywords:** COVID-19, medical exhaustion, prediction, severe case, ventilator

## Summary

**Background:** To avoid exhaustion of medical resources by COVID-19, policy-makers must predict care needs, specifically the proportion of severe cases likely to need ventilator care.

**Objective:** This study was designed to use a statistical model to elucidate dynamics of severe cases in Tokyo and to discuss the timing of effective policy activation.

**Methods:** The study extended from April 27 through October 18, 2020 in Japan's Tokyo Metropolitan area. Medical exhaustion was defined as use of more than half of the ventilator capacity available before the COVID-19 outbreak. We regressed the number of severe cases on the newly onset patients of more than 14 days prior. As earlier research indicated, the COVID-19 severity changed at the end of May. Therefore, we added dummy variables to reflect changing severity from June and its product with newly onset patients as the explanatory variable. Then we calculated the threshold using  $R(t)$ :  $R(t)=0.99$  for the number of patients 14 days prior was used as a threshold at which strong countermeasures should be activated to keep to avoid medical exhaustion.

**Results:** The critical number signaling medical exhaustion in Tokyo was defined as 655 cases. We selected 15, 30, 60 and 90 days prior as explanatory variables for explaining the number of severe cases. A coefficient of determination larger than 0.95 was inferred

as indicating good fit. The threshold was estimated as more than 4500 cases for

$R(t)=1.1$  and monotonically decreasing by  $R(t)$  to be 600 cases for  $R(t)=2.5$ .

Discussion and Conclusion: Results showed that newly onset patients reported more than 14 days prior can explain the number of severe cases very well. We also confirmed the threshold number of patients 14 days prior by  $R(t)$  for which strong countermeasures should be activated to avoid medical exhaustion with  $R(t)=0.99$ . This method is expected to be useful for countermeasure activation policies for Tokyo.

1 Introduction

2 As of October 28, 2020, the World Health Organization reported more than 42  
3 million confirmed COVID-19 cases and about 1.15 million deaths worldwide (1). In  
4 Japan, the outbreak reached its peak at the end of July. By October 29, 2020,  
5 approximately 100 thousand cases and 1.7 thousand deaths had been reported (1).

6 For the 29,601 cases reported from February 1 through August 5, 2020, the  
7 case–fatality rate (CFR) was low for people 59 years old and younger. Nevertheless,  
8 among patients 60 years old and older, CFR was found to increase with age. Similarly,  
9 probabilities of ICU admission and ventilator use increased rapidly with age (2).  
10 Patients with underlying diseases such as hypertension or diabetes constituted high-risk  
11 groups for severe cases of COVID-19 (3,4).

12 The Japanese government set guidelines for countermeasures against COVID-19 on  
13 August 28, emphasizing caution for elderly people and high-risk individuals with  
14 underlying diseases (5). Based on those guidelines, medical resource allocations  
15 thereafter emphasized high-risk groups, including elderly people, rather than patients  
16 with mild symptoms or asymptomatic patients.

17 Given these circumstances, understanding and predicting the proportion of severe  
18 cases requiring ventilator treatment are necessary for policy makers if a strong policy

19 requiring self-restriction of travel is to be made, or if customer-service businesses must  
20 be closed. Therefore, this study was conducted to elucidate the dynamics of severe cases  
21 in Tokyo, with discussion of the timing of activating countermeasure policies.

22       Although we applied mathematical models to outbreaks of COVID-19 in Japan,  
23 such models are unable to explain why the outbreak reached a peak at the end of July  
24 under conditions of no herd immunity or lock-down. Actually, at that time in late July,  
25 the Japanese government had encouraged domestic travel and thereby promoted the  
26 spread of the outbreak. Nevertheless, the outbreak was contained in August. Therefore,  
27 to date, no mathematical model has been able to explain the outbreak in Japan. Instead,  
28 we examined a statistical model rather than a mathematical model to explain the course  
29 of the outbreak. We propose some thresholds that are useful to signal activation of  
30 strong countermeasures.

31

## 32 Methods

33       Severe cases of COVID-19 were defined as requiring hospitalization of patients  
34 with required use of a ventilator. The study period extended from April 27, when severe  
35 cases under the current definition began to be reported in Tokyo, to October 18, 2020.  
36 The Tokyo Metropolitan area of Japan was the study area. Medical exhaustion was

37 defined as patients using more than half of the ventilator capacity available before the  
38 COVID-19 outbreak.

39 We regressed the number of severe cases on the newly onset patients reported more  
40 than 14 days prior. As earlier studies have indicated, the severity of COVID-19 changed  
41 at the end of May. Therefore, we added dummy variables to reflect changing severity  
42 from June and its product with newly onset patients as an explanatory variable. We  
43 selected a formula maximizing the adjusted coefficient of determination.

44 Then, we calculated the threshold by assuming constant  $R(t)$  after October 19, 2020.  
45 The timing of  $R(t)=0.99$  was assessed for activation of strong countermeasures to avoid  
46 medical facility exhaustion, with activation occurring when the number of patients 14  
47 days prior reached this criterion level. Medical facility exhaustion was defined as  
48 patients requiring ventilator use exceeding half of the ventilator capacity available  
49 before the COVID-19 outbreak.

50 The study period for the epidemic curve of symptomatic patients was January 16,  
51 2020 through October 18, 2020. For severe cases, it was April 27, 2020, when severe  
52 cases were defined as patients who needed ventilators, through October 18, 2020.  
53 Because incubation periods and delays of reporting varied after onset of symptoms, data  
54 of the prior month were expected to be underreported. Therefore, we did not use the

55 final month of data available to us. However, collection of data of patients using a  
56 ventilator started April 27 in Tokyo. Therefore, data of ventilator usage on or before  
57 April 26 were not used for this study. The study area was metropolitan Tokyo, with 14  
58 million residents. The epidemic curve and information about severe cases in Tokyo  
59 were provided by the Tokyo Metropolitan Government (6).

60 We also performed out of sample prediction assuming that  $R(t)$  will be constant  
61 from October 19, 2020. We calculated the threshold to activate strong policy measures  
62 for reducing  $R(t)$  to be 0.99 for the number of newly onset patients in a day in terms of  
63  $R(t)$ . Such measures might include some requirements of voluntary restrictions against  
64 going out of the home or shutting down customer-service businesses. Under such  
65 policies, one can expect to avoid exhaustion of medical resources and keep the need for  
66 ventilators lower than 655.

67 Finally, to evaluate the model's predictive power, we performed out of sample  
68 forecasting prospectively for 21 days from December 3, 2020. The parameters for  
69 forecasting were fixed. The estimation result used data obtained through October 18,  
70 2020.

71

72 Ethical considerations

73 All information used for this study was collected under the Law of Infection  
74 Control, Japan. Published data were used. Therefore no ethical issue is posed by the  
75 procedures used to conduct this study.

76

## 77 Results

78 In Tokyo, we observed 3,072 patient days of ventilator use from April 27, 2020  
79 through the end of July. From January 16, 2020 through the end of July, 11,619 patients  
80 with symptoms were reported.

81 Because Tokyo had 2,456 ventilators before the COVID-19 outbreak and because  
82 1,310 ventilators were unused by patients (7), if patients needing ventilators were to  
83 become more numerous than 655, the situation would be designated as medical facility  
84 exhaustion.

85 Figure 1 depicts severe cases and newly onset patients. Severe cases fluctuate with  
86 a longer lag than the onset patients. The estimation results are presented in Table 1.  
87 Because 120 days before April 27 was earlier than January 16, when the initial case in  
88 Tokyo showed onset, the specification with onset patients 120 days prior includes fewer  
89 observations than other specifications. Based on adjusted coefficients of determination,  
90 we selected the model with up to 90 days prior. Its adjusted coefficient of determination



91 was higher than 0.95. Figure 2 depicts the observed and fitted numbers of severe cases.

92 It is apparent that these were almost equal to one another.

93 The estimated thresholds of the number of patients 14 days prior by  $R(t)$  in which

94 the timing of activation of strong countermeasures to be  $R(t)=0.99$  should be activated

95 to avoid medical exhaustion are shown at Figure 3. The threshold was estimated as

96 more than 4500 for  $R(t)=1.1$ , decreasing monotonically by  $R(t)$  to be 600 for  $R(t)=2.5$ .

97 Figure 4 depicts out of sample forecasting prospectively for 21 days from

98 December 3, 2020 through February 23, 2021. The observed peak was January 20, 2021,

99 but the peak in forecasting was January 29, 2021. Before the peak date, forecasting was

100 significantly lower than the observed figure, except for early December. However, after

101 the peak date, no significant difference was found among them.

102

103 Discussion

104 Many mathematical models have been applied to the outbreak of COVID-19

105 worldwide (8,9). Unfortunately, these have not succeeded in explaining the course of

106 COVID-19 outbreaks.

107 Especially, in Japan, two peaks were experienced through October. The first peak

108 was that of earlier April; the second one was at the end of July. The first peak might

109 explain the mathematical model with restriction of going out (10) because, since  
110 mid-March, the rate of going out declined, as shown in Apple data. However, the  
111 second peak is inexplicable by the mathematical model. Herd immunity, of course, was  
112 not achieved. No lock-down was activated in Japan at all. Probably, the “new normal”  
113 lifestyle with mask-wearing, social distancing, and restrictions against numerous dining  
114 or night-life establishments became common.

115       Actually, at that time, the government encouraged domestic travel. It promoted the  
116 spread of the outbreak. However, the outbreak was contained in August. Therefore, to  
117 date, no mathematical model has been sufficient to explain the outbreak in Japan. The  
118 policy for severe cases cannot be discussed using mathematical models.

119       Instead, we used a statistical model. It might always show some association of  
120 variables for which there is no causality. We showed clear association among severe  
121 cases and patients who had experienced onset 15 days prior. Moreover, we showed the  
122 threshold of the number of patients 14 days prior by  $R(t)$ , for which the timing of  
123 activation of strong countermeasures to be  $R(t)=0.99$  should be activated to avoid  
124 medical exhaustion. Therefore, the statistical model was probably more useful than the  
125 mathematical model for countermeasures against COVID-19.

126       Regarding the increase phase in June and July, the number of newly onset patients

127 increased from about 10 to 350 during 60 days. If one assumes an average incubation  
128 period to be about 6 (10), then the average  $R(t)$  during this period was  
129  $1.43=(350/10)^{(6/60)}$ . Figure 3 depicts the threshold as around 2200 at  $R(t)=1.43$ . The  
130 peak at the end of July was 350. Therefore, the peak at the end of July accounted for  
131 only 16% of the threshold.

132 We used no information before  $t-14$ . Reporting was usually delayed from onset up  
133 to 31 days (10). Therefore, the number of newly onset patients at the same day was that  
134 which was usually developing dynamically for at least two weeks. Information of the  
135 final two weeks prior to the estimate is not reliable. Therefore, we ignored information  
136 during the final two weeks of the study period. Moreover, because the epidemic curve  
137 changes gradually, information related to day  $t$  is closely correlated with information  
138 related to day  $t-1$ . Therefore, we used no information related to consecutive days. We  
139 used the number of newly onset patients with 15–30 day intervals as explanatory  
140 variables.

141 Prospective out of sample forecasting in Figure 4 shows good fit. Especially in the  
142 declining phase, it fits very well. Therefore, results demonstrated that the estimated  
143 model can predict outcomes well for 21 days.

144 Some limitations might have affected this study. First, we used dummy variables

145 after June to reflect the disease's decreasing severity. However, we do not understand  
146 the reason why. It might be attributable to mutations of the COVID-19 virus.  
147 Alternatively, treatment for patients might have improved since June. Or younger and  
148 therefore milder patients accounted for a larger proportion of the patients than before  
149 June. Moreover, the severity of the disease would decrease if a new drug were  
150 developed. It might alternatively increase if some mutation were to strengthen its  
151 pathogenicity. Severity should be monitored to improve the fit of the statistical model.

152 Secondly, definitions of severe cases differed among prefectures, even in Japan.  
153 Moreover, they probably varied among countries. It remains unknown whether the  
154 approach used for the present study is applicable to other prefectures than Tokyo or  
155 other countries with other definitions of severity. We should check the robustness of  
156 usefulness of the statistical model in other areas in Japan or other countries.

157 Thirdly, we might specifically examine only elderly patients or patients older than  
158 50 as explanatory variables. If the reason for lower severity after June was the larger  
159 proportion of younger patients, then it might be controlled by specifically examining  
160 elderly patients as a high risk group. Exploration of that point remains as a challenge for  
161 future research efforts.

162

163 Conclusion

164 Results demonstrated that newly onset patients more than 14 days prior can  
165 explain the number of severe cases very well. The prospective out of sample forecasting  
166 used for this study can predict observations well for 21 days. Results also show a  
167 threshold of the number of patients 14 days prior by  $R(t)$ . Timing of the activation of  
168 strong countermeasures should be  $R(t)=0.99$  to avoid medical facility exhaustion. That  
169 signal might be useful for countermeasure activation in Tokyo. Careful real-time  
170 monitoring of  $R(t)$  and newly onset patients are expected to be extremely important for  
171 timing of the activation of strong policy measures.

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

174

175 Acknowledgments

176 We acknowledge the great efforts of all staff at public health centers, medical  
177 institutions, and other facilities who are fighting the spread and destruction associated  
178 with COVID-19. This study represents the authors' opinion. It does not reflect any  
179 stance of our affiliation.

180

181 Competing Interest

182 No author has any conflict of interest, financial or otherwise, to declare in relation  
183 to this study.

184

185 Reference

186 1) World Health Organization. [WHO Coronavirus Disease \(COVID-19\) Dashboard](#)

187 <https://covid19.who.int/> [accessed on October 29, 2020]

188 2) National institute of infectious diseases. Descriptive epidemiology of 516 confirmed

189 cases of novel coronavirus infection reported by the national epidemiological

190 surveillance of infectious diseases (NESID) system and active epidemiological

191 surveillance (as of March 23, 2020).

192 [https://www.niid.go.jp/niid/en/2019-ncov-e/2484-idsc/9555-covid19-14-200323-e.h](https://www.niid.go.jp/niid/en/2019-ncov-e/2484-idsc/9555-covid19-14-200323-e.html)

193 [tml](#). [accessed on September 19, 2020]

194 3) Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult

195 inpatients with COVID-19 in Wuhan, China: a retrospective cohort study [published

196 correction appears in Lancet. 2020 Mar 28;395(10229):1038] [published correction

197 appears in Lancet. 2020 Mar 28;395(10229):1038]. Lancet.

198 2020;395(10229):1054-1062. doi:10.1016/S0140-6736(20)30566-3

- 199 4) Wu C, Chen X, Cai Y, et al. Risk Factors Associated With Acute Respiratory  
200 Distress Syndrome and Death in Patients With Coronavirus Disease 2019  
201 Pneumonia in Wuhan, China. *JAMA Intern Med.* 2020;180(7):934-943.  
202 doi:10.1001/jamainternmed.2020.0994
- 203 5) New measures for the coronavirus. August 28, 2020. Novel Coronavirus Taskforce  
204 decision. (in Japanese) <https://www.mhlw.go.jp/content/10900000/000664804.pdf>
- 205 6) Tokyo Metropolitan Government. The Latest Outbreak Situation in Tokyo  
206 <https://stopcovid19.metro.tokyo.lg.jp/>. (in Japanese) [accessed on October 29, 2020]
- 207 7) Cabinet Secretariat. COVID-19 Information and Resources.  
208 <https://corona.go.jp/dashboard>. (in Japanese) [accessed on September 19, 2020]
- 209 8) Kucharski AJ, Russell TW, Diamond C, Liu Y, Edmunds J, Funk S, Eggo RM;  
210 Centre for Mathematical Modelling of Infectious Diseases COVID-19 working  
211 group. Early dynamics of transmission and control of COVID-19: a mathematical  
212 modelling study. *Lancet Infect Dis.* 2020 ;20:553-8.
- 213 9) Kurita J, Sugawara T, Ohkusa Y. Estimated effectiveness of school closure and  
214 voluntary event cancellation as COVID-19 countermeasures in Japan. *J Infect*  
215 *Chemother* 2020 Aug 19;S1341-321X(20)30292-0. doi:  
216 10.1016/j.jiac.2020.08.012.

217 10) Kurita J, Sugawara T, Ohkusa Y. Mobility data can explain the entire COVID?19

218 outbreak course in Japan. medRxiv 2020.04.26.20081315; doi:

219 <https://doi.org/10.1101/2020.04.26.20081315>

220

221

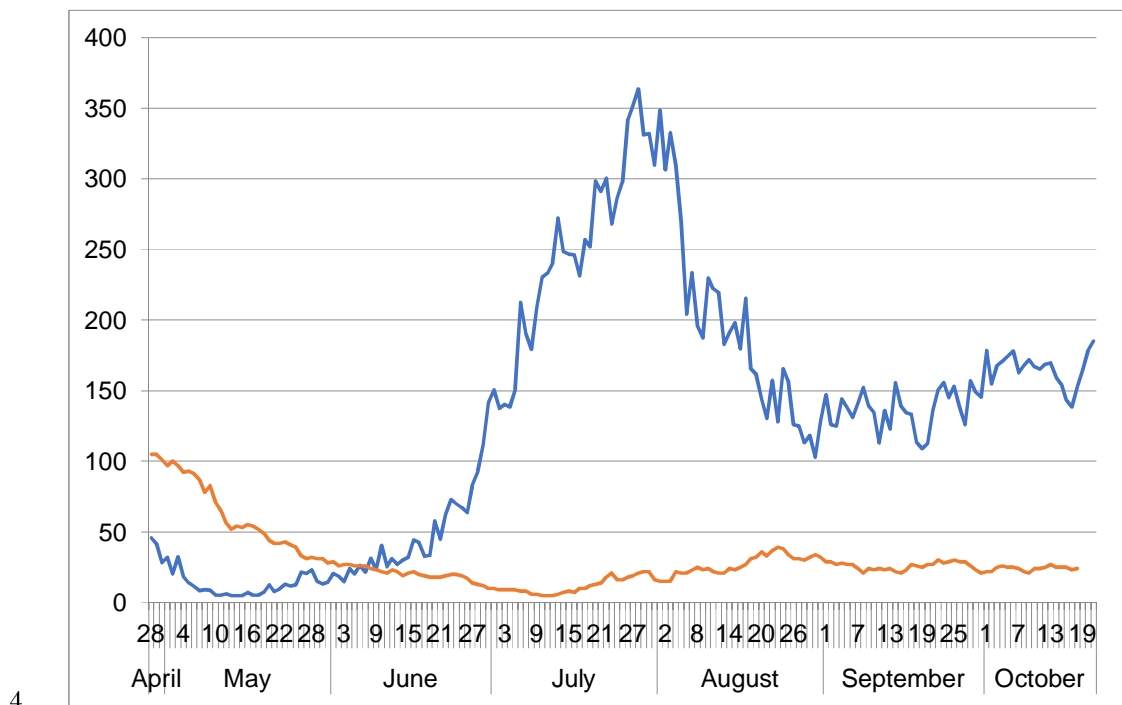


1 Table 1: Estimation result for the number of sever cases

Explanatory variable	Estimated coefficients	<i>p</i> -value	Estimated coefficients	<i>p</i> -value	Estimated coefficients	<i>p</i> -value	Estimated coefficients	<i>p</i> -value
constant	3.899	0.004	3.821	0.025	5.743	0.000	11.549	0.000
t-15	0.011	0.206	0.024	0.002	0.021	0.017	-0.007	0.377
t-30	0.064	0.000	0.059	0.000	0.054	0.000	0.073	0.000
t-60	0.030	0.000	0.029	0.000	0.031	0.000		
t-90	0.013	0.020	0.012	0.084				
t-120	0.020	0.079						
dummy variable for after June	19.256	0.000	15.617	0.000	1.272	0.625	3.097	0.083
Cross term with dummy variable for after June								
t-15	0.059	0.545	0.302	0.000	0.229	0.000	0.175	0.000
t-30	0.268	0.000	0.218	0.000	0.329	0.000	0.300	0.000
t-60	-0.047	0.008	-0.031	0.041	0.030	0.058		
t-90	-0.101	0.000	-0.135	0.000				
t-120	-0.477	0.240						
Number of observations =	157		175		175		175	
Adjusted coefficients of determinant	0.8829		0.9613		0.9443		0.9265	

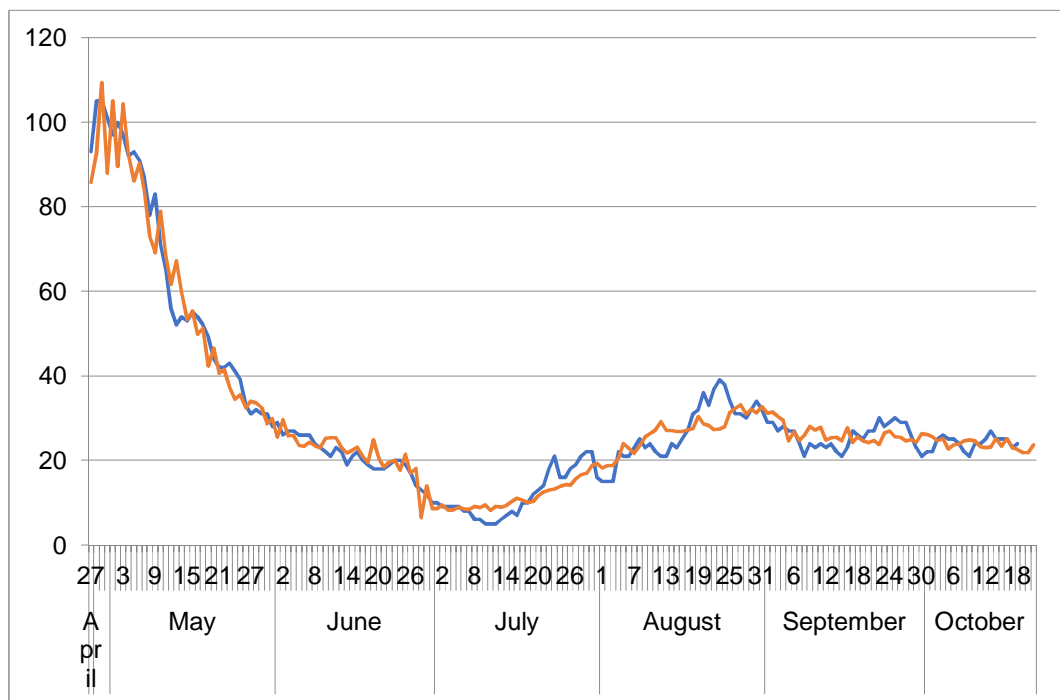
2 Note: t-x (x=15,30,60,90, and 120) indicates the number of newly onset patients x days  
3 before. Panel lower than “Cross term with dummy variable for after June” shows cross  
4 term with dummy variable for after June and he number of newly onset patients x days  
5 before. The most left specification with t-120 has fewer observations than other  
6 specification because 120 day before from April 27 was earlier than January 16 when  
7 the initial case in Tokyo was onset. Based on adjusted coefficients of determinant, we  
8 selected the model with up to 90 days before

- 1 Figure 1: Epidemic curve of the symptomatic confirmed cases and number of patients
- 2 requiring a ventilator in Tokyo, April 28 – October 18, 2020.
- 3 (number of patients)



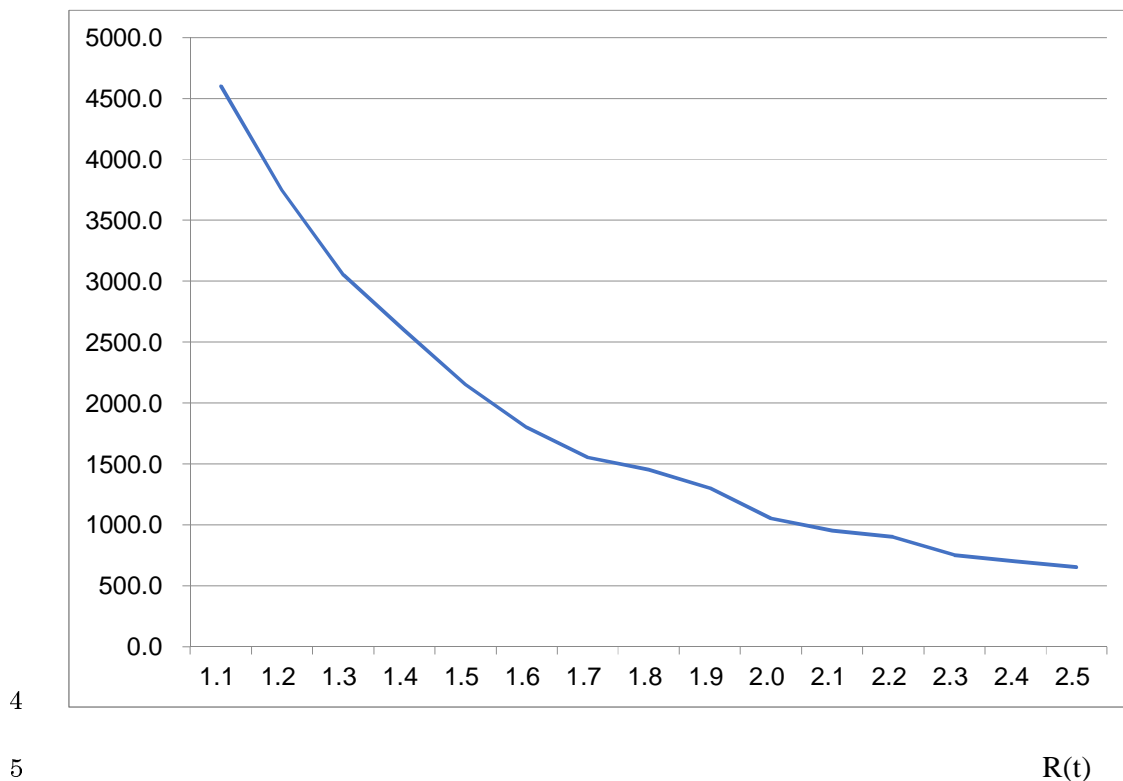
- 5 Notes: The blue line shows the epidemic curve of the symptomatic confirmed cases on
- 6 the onset date. The red line represents the number of patients with a ventilator.

- 1 Figure 2: Estimated and observed number of patients requiring a ventilator in Tokyo,
- 2 April 28 – October 18, 2020.
- 3 (number of patients)



- 4
- 5 Note: Blue The blue line shows the observed number of patients requiring a ventilator.
- 6 The red line represents the estimated number of patients requiring a ventilator.

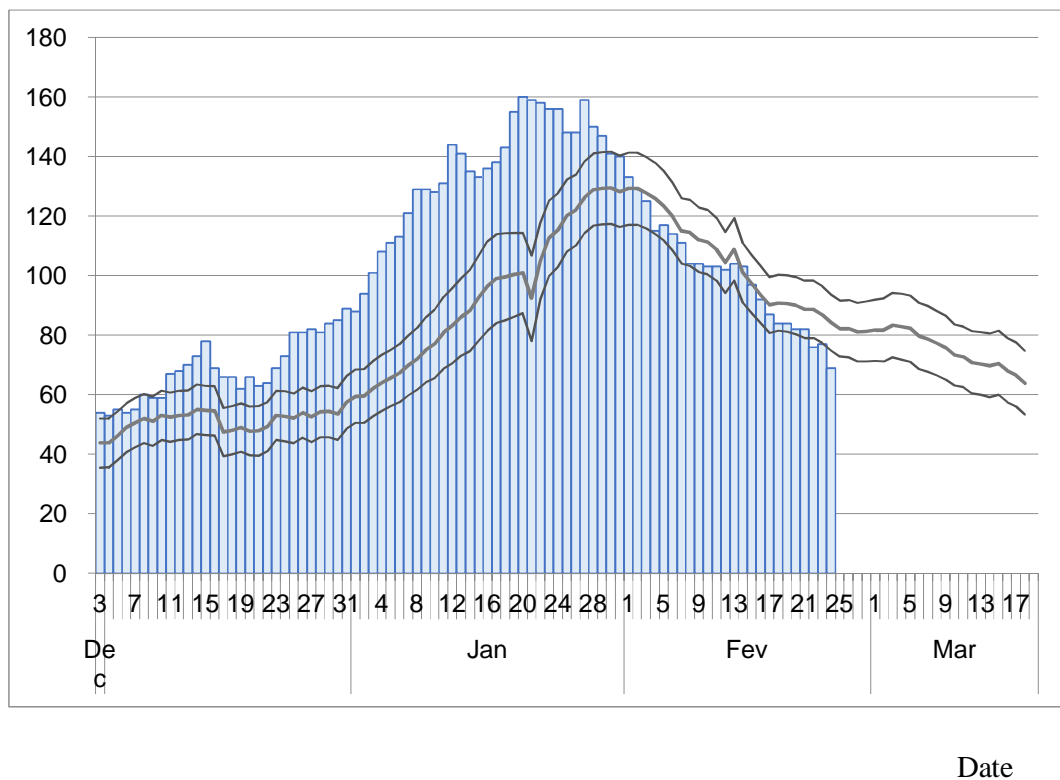
- 1 Figure 3: Estimated threshold of the number of newly onset cases 14 days prior to avoid
- 2 medical facility exhaustion by  $R(t)$ .
- 3 (number of newly onset patients 14 days before)



- 4
- 5
- 6 Notes: The line represents the threshold of the number of newly onset patients on 14
- 7 days prior, by  $R(t)$ . Medical facility exhaustion will occur if newly onset patients during
- 8 14 days prior exceed the threshold.

1 Figure 4: Observed data of severe cases and its forecasting.

2 (number of patients)



5 Notes: Bar represents data of severe cases observed in Tokyo from December 3, 2020.

6 The bold line represents a forecast based the estimation results presented in Table 1.

7 Thin lines represent its 95% CI.