Journal of General and Family Medicine WILEY

A mathematical model for COVID-19 pandemic—SIIR model: Effects of asymptomatic individuals

Masaki Tomochi PhD, Professor[1] (iD) | **Mitsuo Kono DrSci, Professor Emeritus[2]**

[1]Department of Economics, Okinawa International University, Okinawa, Japan

[2]Faculty of Policy Studies, Chuo University, Tokyo, Japan

Correspondence
Masaki Tomochi, Department of Economics, Okinawa International University, 2-6-1 Ginowan, Ginowan, Okinawa 901-2701 Japan.
Email: mtomochi@okiu.ac.jp

Abstract

A new mathematical model called SIIR model is constructed to describe the spread of infection by taking account of the characteristics of COVID-19 and is verified by the data from Japan. The following features of COVID-19: (a) there exist presymptomatic individuals who have infectivity even during the incubation period, (b) there exist asymptomatic individuals who can freely move around and play crucial roles in the spread of infection, and (c) the duration of immunity may be finite, are incorporated into the SIIR model. The SIIR model has the advantage of being able to explicitly handle asymptomatic individuals who are delayed in discovery or are extremely difficult to be discovered in the real world. It is shown that the conditions for herd immunity in the SIIR model become more severe than those in the SIR model; that is, the presence of asymptomatic individuals increases herd immunity threshold (HIT).

KEYWORDS

asymptomatic individuals, COVID-19, finite antibody duration, Health Policy, infectious diseases, SIIR model

1 | INTRODUCTION

A history of mankind is a history of fighting against infectious diseases, such as smallpox, polio, plague, flu, AIDS, SARS, and MARS. Because of the spread of COVID-19 in these days, human beings continue to be suffered by pandemics, that is, danger of collapse of medical, social, and economic system.

Mathematical models have been used to assess the spread of infectious diseases for long time. As infectious diseases change and become more complex, mathematical models have also evolved. Among these models, SIR (susceptible–infectious–recovered [removed]) model, which has three transitioning stages: susceptibility, infectious, and recovered (and removed from the infectious network), has been widely used for discussing the process of infectious diseases spreading. In addition, SEIR model, which divides the

infection stage in the SIR model into two stages, noninfectious (exposed) and infectious, and investigates the four transitioning stages, has also been widely used.[1–9]

Since the mechanism of the spread of COVID-19 consists of direct and indirect contacts between infected and susceptible individuals, it seems that the conventional models sufficiently explain the phenomena if COVID-19 is regarded to be the same as the conventional infectious diseases. However, unique features of COVID-19 should be taken into account. These features are (a) there exist presymptomatic individuals who are infectious even during the incubation period, (b) there exist asymptomatic individuals who can move around freely and spread infection, making it difficult to control the situation, and (c) the duration of immunity may be finite and susceptible population is reproduced; that is, it may take long time to achieve herd immunity and several peaks of the spread of infection may occur.

In the next section, SIIR model is introduced to explain the mechanism of COVID-19 infection spread which cannot be fully explained by the conventional mathematical models such as SIR model and SEIR model. In Section 3, SIIR model is verified by data on Japan to show that the model is suitable. Discussions are given in the last section.
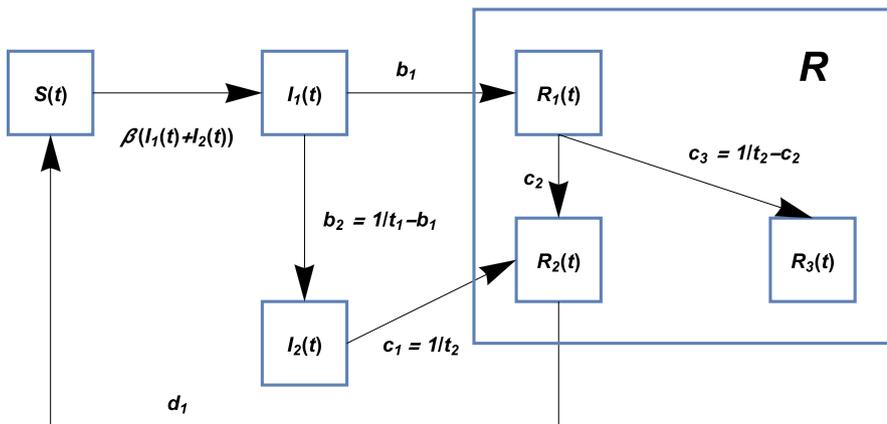
2 | SIIR MODEL

2.1 | Model equations

As one of the features of COVID-19, there is incubation period after infection and presymptomatic carriers can spread the infection during the incubation period. Further, even though the symptomatic carriers or those who are confirmed to be infected are quarantined, there are also asymptomatic individuals who do not develop after the end of the incubation period. This makes the spread of infection difficult to be controlled because asymptomatic individuals are left unchecked without isolation. COVID-19 is characterized by two different infectious states (presymptomatic and asymptomatic) and is pointed out that antibody duration is not so long.[10] These are not considered in SIR or SEIR models and, therefore, a model that takes such characteristics of COVID-19 into account is needed.

Here, parties concerned in SIIR model at time t are $S(t)$ as susceptible population, $I_1(t)$ as presymptomatic population (infectious), $I_2(t)$ as asymptomatic population (infectious), $R_1(t)$ as symptomatic population (not infectious by quarantine), $R_2(t)$ as recovered population (with antibody and not infectious), and $R_3(t)$ as fatalities because of COVID-19 (not infectious). Then, the interrelationship among the above variables is shown in Figure 1 and is described by the following coupled differential equations:

$$\frac{dS(t)}{dt} = -\beta S(t)(I_1(t) + I_2(t)) + d_1 R_2(t), \tag{1}$$

$$\frac{dI_1(t)}{dt} = \beta S(t)(I_1(t) + I_2(t)) - (b_1 + b_2)I_1(t), \tag{2}$$

$$\frac{dI_2(t)}{dt} = b_2 I_1(t) - c_1 I_2(t), \tag{3}$$

$$\frac{dR_1(t)}{dt} = b_1 I_1(t) - (c_2 + c_3)R_1(t), \tag{4}$$

$$\frac{dR_2(t)}{dt} = c_1 I_2(t) + c_2 R_1(t) - d_1 R_2(t), \tag{5}$$

$$\frac{dR_3(t)}{dt} = c_3 R_1(t), \tag{6}$$

which is called SIIR model since there are two states in the infection stage, that is, I_1 and I_2.

In the SIIR model, presymptomatic and asymptomatic individuals are at different stages but they are both not quarantined and are free to move around. Therefore, in order to clearly show that they are infectious, they are denoted as I_1 and I_2, respectively. On the other hand, asymptomatic individuals, recovered individuals with antibody, and those who passed away are not infectious, so they are represented by R, that is, R_1, R_2, and R_3, respectively. The SIIR model is used to distinguish it from the SIR model. The feature of the model is that infectious and noninfectious states have internal structures.

The coefficient β in the SIIR model corresponds to the probability, with which susceptible individuals (S) will be infected by contacting with those who are infectious ($I_1 + I_2$) at a different stage of infectivity, and has a similar role to the one in the SIR model. Presymptomatic individuals (I_1) bifurcate into symptomatic individuals (R_1) who are quarantined and asymptomatic individuals (I_2) who are not quarantined with the coefficients b_1 and b_2, respectively. The meanings of these coefficients are understood in the following way. Assuming presymptomatic population are evenly distributed during the incubation period t_1, out of presymptomatic population $I_1(t)$, those at the end of incubation period t_1 denoted by $n(t_1)$ may be expressed by $n(t_1) = I_1(t)/t_1$. At the end of the incubation period, the i-th presymptomatic individual bifurcates into the symptomatic state with a probability b_{1i} and asymptomatic state with a probability b_{2i} where $b_{1i} + b_{2i} = 1$. On the other

FIGURE 1 Structure of the SIIR model: $S(t)$ as susceptible population, $I_1(t)$ as presymptomatic population (infectious), $I_2(t)$ as asymptomatic population (infectious), $R_1(t)$ as symptomatic population (quarantined and not infectious), $R_2(t)$ as recovered population (recovered with antibody and not infectious), and $R_3(t)$ as fatalities because of COVID-19 (not infectious)

FIGURE 2 Single peak (upper row) and multiple peak (lower row) solution. Upper and lower left: the susceptible population $S(t)$ (blue) and recovered population with antibody $R_2(t)$ (cyan). Upper and lower right: The presymptomatic population $I_1(t)$ (red), asymptomatic population $I_2(t)$ (green), symptomatic and quarantined population $R_1(t)$ (yellow), and fatalities $R_3(t)$ (black)
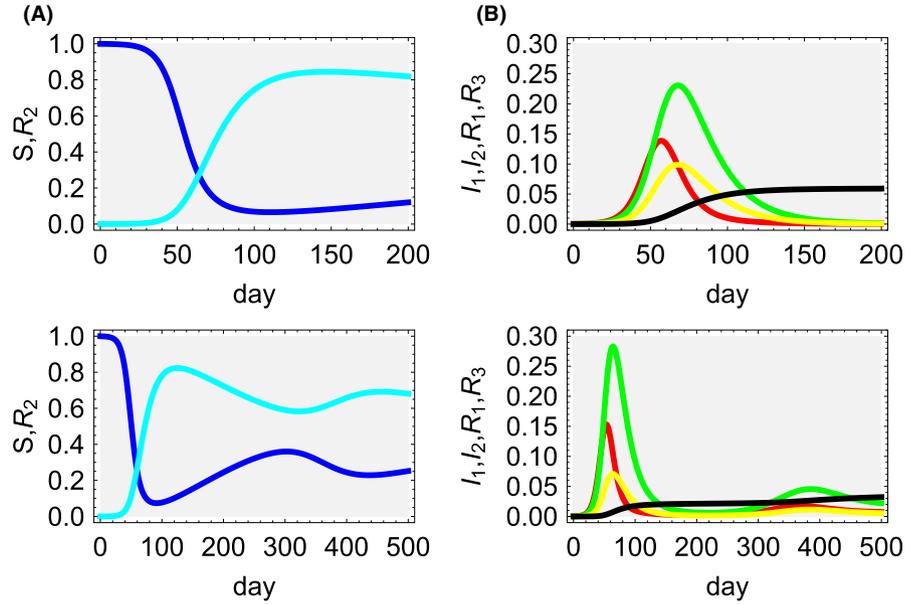
hand, assuming that symptomatic and asymptomatic individuals are equally distributed in the period of the onset, they are given by $R_1(t)/t_2$ and $I_2(t)/t_2$, respectively, at the time of transition from incubation to onset. Thus, we have,

$$\sum_{i=1}^{n(t_1)} b_{1i} = \frac{R_1(t)}{t_2} = \Delta R_1(t), \quad \sum_{i=1}^{n(t_1)} b_{2i} = \frac{I_2(t)}{t_2} = \Delta I_2(t),$$

$$\sum_{i=1}^{n(t_1)} (b_{1i} + b_{2i}) = n(t_1) = \frac{I_1(t)}{t_1} = \frac{1}{t_2}(R_1(t) + I_2(t))$$

Since $I_1(t)$ bifurcates into $\Delta R_1(t)$ with b_1 and $\Delta I_2(t)$ with b_2, we have,

$$\Delta R_1(t) = b_1 I_1(t), \quad \Delta I_2(t) = b_2 I_1(t), \quad \rightarrow \quad b_1 + b_2 = \frac{1}{t_1}. \quad (7)$$

Factors that determine whether those who are infected become symptomatic or asymptomatic have not yet been identified. However, if health factors are not considered, $1/t_1$ is apportioned between b_1 and b_2.

The coefficients c_1, c_2, and c_3 are also related to the onset period. At the end of the onset period t_2, the following,

$$\frac{1}{t_2} R_1(t) = \Delta_1 R_2(t) + \Delta R_3(t),$$

holds at the transition from symptomatic to recovery with antibody or to death. For the transition from asymptomatic to recovery with antibody, the following relation holds:

$$\frac{1}{t_2} I_2(t) = \Delta_2 R_2(t),$$

as long as there is no transition from asymptomatic to death. Therefore, we have,

$$\Delta_1 R_2(t) = c_2 R_1(t), \quad \Delta_2 R_2(t) = c_1 I_2(t), \quad \Delta R_3 = c_3 R_1(t),$$

which leads to,

$$c_1 = c_2 + c_3 = \frac{1}{t_2}.$$

That is, c_1 is the inverse of the onset period, and $1/t_2$ is apportioned between c_2 and c_3.

If recovered individuals R_2 lose their antibody at time t, they become susceptible S again, and the coefficient d_1 represents the inverse of the antibody duration.

In this way, the SIIR model, in spite of a continuous model, is explicitly to capture the duration of each stage. In addition, the SIIR model is a five-variable system because of the following conserved quantity given by,

$$\frac{d}{dt}(S(t) + I_1(t) + I_2(t) + R_1(t) + R_2(t) + R_3(t)) = 0$$
$$\rightarrow N = S(t) + I_1(t) + I_2(t) + R_1(t) + R_2(t) + R_3(t).$$

If the variables are normalized by the conserved quantity N, it is convenient to see the characteristic behaviors common to the system with different N. In verifying the model with data, the parameters and scale conversion with respect to N are required when the theoretical results best-fit the data. For example, for the spread of infection in Japan, it does not make sense to set the question of how the entire population of Japan of one hundred twenty-six million susceptible individuals is infected. This is because the infection cannot evenly spread throughout Japan. There are many high-risk

areas such as densely populated areas and workplaces scattered throughout Japan. Since the spread of infection in each place is similar to each other, by combining the results of the normalized model as a whole of local infections such as cluster infections, community-acquired infections, and domestic infections, the substance of the spread of global infection may become clear. This is the power of mathematical models. So, we normalize the variables in the SIIR model by N and make the following replacements:

$$\left\{ \frac{S(t)}{N}, \frac{I_i(t)}{N}, \frac{R_i(t)}{N} \right\} \;\rightarrow\; \{S(t), I_i(t), R_i(t)\}, \quad \beta N \;\rightarrow\; \beta.$$

By these replacements, Equations from (1) to (6) representing the SIIR model are isomorphic even after normalization.

2.2 | Numerical solutions of SIIR model

Now, we look for numerical solutions of the SIIR model. Figure 2 (single peak for the upper row and multiple peaks for the lower row) shows the time evolution of the variables of the SIIR model where the parameters and initial values are set as follows:

$$N = 1, \quad t_1 = 5, \quad t_2 = 17, \quad \beta = 0.2,$$

$$b_1 = \frac{0.3}{t_1}, \quad b_2 = \frac{1}{t_1} - b_1, \quad c_1 = \frac{1}{t_2}, \quad c_2 = \frac{0.8}{t_2}, \quad c_3 = \frac{1}{t_2} - c_1,$$

$d_1 = 0.001$ (for Figure 2 upper row) and $d_1 = 0.03$ (for Figure 2 lower row) and

$$S(0) = N - I_1(0) = S_0, I_1(0) = 0.0004,$$

$$I_2(0) = R_1(0) = R_2(0) = R_3(0) = 0.$$

As is expected, the susceptible population $S(t)$ (blue) decreases and the presymptomatic population $I_1(t)$ (red) increases, which is followed by the growth of both the asymptomatic population $I_2(t)$ (green) and the symptomatic and quarantined population $R_1(t)$ (yellow). Then, the recovered population with antibody $R_2(t)$ (cyan) and the fatalities $R_3(t)$ (black) increase to certain saturation levels. Since we set $b_2 > b_1$, the asymptomatic population (green) becomes larger than the symptomatic and quarantined population (yellow). In this way, the SIIR model has a definite advantage of handling asymptomatic population which is extremely difficult to be discovered in societies and plays crucial roles in infection spreading. Moreover, because $d_1 \neq 0$, for $t \gg 1/d_1$, the recovered population with antibody (cyan) gradually decreases and the susceptible population (blue) gradually increases. For large t, the susceptible population remains finite, showing realization of herd immunity which is discussed in Section 2.4.

During the intermediate transition state of the entire process of infection spreading and termination, the presymptomatic population, symptomatic and quarantined population, and asymptomatic population have their own peaks in the short term. It is β and t_1 that constitute the basic reproduction number and determine rising speed of the peak of the presymptomatic population. If β is fixed, the larger t_1, the faster presymptomatic population rises, the peak position shifts to the smaller time axis, and the larger the value of the peak. This is because if t_1 is larger, presymptomatic individuals stay in incubation state for a longer time, and conversely, if t_1 is smaller, they immediately move to the next stage. Therefore, the peak width of the presymptomatic population depends on the magnitude of t_1. The presymptomatic individuals bifurcate to the symptomatic individuals with the ratio of b_1 and asymptomatic individuals with $b_2 = 1/t_1 - b_1$. Therefore, when b_1 is smaller, asymptomatic population increases, leading to a larger feedback to presymptomatic population which in turn grows faster with higher peak and narrower width. On the other hand, when b_1 becomes larger, the symptomatic and quarantined population increases, the feedback to presymptomatic population becomes smaller, resulting in slower growth, smaller peak, and wider width. Thus, during the intermediate transition states, b_1 determines position, growth rate, and width of the peak of infectious population.

The t_2 determines how symptomatic and asymptomatic population decline from the peak. The reason is clear from the terms $c_2 = 1/t_2$ and $c_2 + c_3 = 1/t_2$ of Equations (3) and (4). If t_2 becomes larger, the attenuation becomes gentler and the peak width becomes wider.

In the SIIR model, all recovered individuals have antibodies, but since the antibody duration is set as $1/d_1$, they become susceptible again over time, that is, small peaks will be repeated until complete termination of the infection.

2.3 | Basic reproduction number and effective reproduction number for SIIR model

Basic reproduction number of the SIIR model is very much different from that of the SIR model. In the SIR model, $\mathscr{R}_0 > 1$ is a condition for infection to be spread. However, in Figure 2 for the SIIR model, even though we have,

$$\mathscr{R}_0 = \frac{\beta S_0}{b_1 + b_2} = \beta S_0 t_1 = 0.9996 (< 1),$$

spread of infection has still occurred. This is because the condition on $\mathscr{R}_0$ in the SIIR model is relaxed because of the presence of asymptomatic individuals, which is one of the characteristics of the SIIR model. This is understood from the initial behavior of the SIIR model. Since $S(t) \simeq S_0$ can be used near $t \simeq 0$, Equations (2) and (3) are expressed as,

$$\frac{d}{dt} \begin{pmatrix} I_1(t) \\ I_2(t) \end{pmatrix} = A \begin{pmatrix} I_1(t) \\ I_2(t) \end{pmatrix}, \quad A = \begin{pmatrix} \beta S_0 - (b_1 + b_2) & \beta S_0 \\ b_2 & -c_1 \end{pmatrix}.$$

The eigenvalue λ of A is calculated as,

$$\lambda = \frac{1}{2}\{-\{(b_1+b_2)(1-\mathscr{R}_0)+c_1\} \pm \sqrt{\{(b_1+b_2)(1-\mathscr{R}_0)+c_1\}^2 - 4(b_1+b_2)\{c_1-(c_1+b_2)\mathscr{R}_0\}}\},$$

from which the condition $\lambda > 0$ gives,

$$\mathscr{R}_0 > \frac{c_1}{c_1+b_2}. \tag{8}$$

In Equation (8), the condition for the spread of infection in SIR model is recovered for $b_2 = 0$. The right-hand side of Equation (8) is smaller than 1 for $b_2 > 0$. Thus, the presence of asymptomatic individuals reduces the threshold of $\mathscr{R}_0$ and the spread of infection is taken place even for $\mathscr{R}_0 < 1$.

The effective reproduction number $\mathscr{R}_e(t)$ for the SIIR model is calculated from Equation (2) as follows:

$$\frac{dI_1(t)}{dt} = (b_1+b_2)(\mathscr{R}_e(t)-1)I_1(t), \quad \mathscr{R}_e(t) = \frac{\beta}{b_1+b_2}S(t)\frac{I_1(t)+I_2(t)}{I_1(t)}.$$

Since the basic reproduction number $\mathscr{R}_0$ for the SIIR model is specific to the infectious disease and is determined based on the initial value as,

$$\mathscr{R}_0 = \frac{\beta}{b_1+b_2}S_0, \quad S_0 = S(0), \tag{9}$$

the effective reproduction number is represented as,

$$\mathscr{R}_e(t) = \mathscr{R}_0 \frac{S(t)}{S_0}\left(1+\frac{I_2(t)}{I_1(t)}\right). \tag{10}$$

The SIIR model cannot be solved analytically but can be solved numerically, and numerical solutions are used to evaluate Equation (10).

Equation (8) is also obtained based on Equation (10). Using a resolvent operator:

$$\mathscr{L}(c_1) = \frac{d}{dt}+c_1,$$

Equation (3) has the following formal solution:

$$I_2(t) = b_2\mathscr{L}^{-1}(c_1)I_1(t).$$

Near $t=0$, we have,

$$\mathscr{L}^{-1}(c_1) \sim \frac{1}{c_1},$$

leading to

$$\mathscr{R}_e = \mathscr{R}_0\left(1+\frac{b_2}{c_1}\right).$$

The condition $\mathscr{R}_e > 1$ gives Equation (8). Thus, the effective reproduction number depends on both the incubation period $t_1 = 1/(b_1+b_2)$ and the onset period $t_2 = 1/c_1$.

2.4 | Condition for herd immunity

In order for the spread of infection to reach an end, we have $\mathscr{R}_e(t) < 1$, which is equivalent to the condition for the onset of herd immunity:

$$1-\frac{S(t)}{S_0} > 1-\frac{1}{\mathscr{R}_0}\frac{I_1(t)}{I_1(t)+I_2(t)} > 1-\frac{1}{\mathscr{R}_0}. \tag{11}$$

If $S(t)$ is a monotonically decreasing function which is the case for $d_1 = 0$, the condition for the herd immunity of the SIIR model is shown to be more severe than that of the SIR model, and the presence of asymptomatic individuals increases herd immunity threshold (HIT). In addition, since those who lose antibody after recovery are transferred to be susceptible again, $S(t)$ may not be a monotonically decreasing function, for which $\mathscr{R}_e(t)$ oscillates around 1 and the infection peaks will appear multiple times in a period of $1/d_1$, the antibody duration.

2.5 | Cumulative fatalities of SIIR model

It is very much important to examine fatalities caused by infectious diseases. Here, we evaluate cumulative fatalities in the SIIR model. As we see earlier, we have,

$$c_1 = c_2+c_3,$$

and, from Equations (3) and (4), we have,

$$\frac{d}{dt}+c_1 \left[I_2(t)-\frac{b_2}{b_1}R_1(t)\right] = 0, \rightarrow I_2(t)-\frac{b_2}{b_1}R_1(t) = \left\{I_2(0)-\frac{b_2}{b_1}R_1(0)\right\}e^{-c_1 t}.$$

Therefore, we obtain,

$$I_2(t)-\frac{b_2}{b_1}R_1(t) \rightarrow 0 \quad \text{for} \quad t \gg \frac{1}{c_1} = t_2,$$

and, since we set $I_2(0) = R_1(0) = 0$ for the initial values, we have,

$$I_2(t) = \frac{b_2}{b_1}R_1(t). \tag{12}$$

With Equation (12), the SIIR model becomes four-variable system. Now, we rewrite the basic equations and have,

$$\frac{dS(t)}{dt} = -\beta S(t)(I_1(t) + I_2(t)) + d_1 R_2(t), \tag{13}$$

$$\frac{dI_1(t)}{dt} = \beta S(t)(I_1(t) + I_2(t)) - (b_1 + b_2)I_1(t), \tag{14}$$

$$\frac{dI_2(t)}{dt} = b_2 I_1(t) - c_1 I_2(t), \tag{15}$$

$$\frac{dR_2(t)}{dt} = \left(c_1 + c_2 \frac{b_1}{b_2}\right) I_2(t) - d_1 R_2(t). \tag{16}$$

Then, the conservation quantity becomes,

$$1 = S(t) + I_1(t) + \left(1 + \frac{b_1}{b_2}\right) I_2(t) + R_2(t) + R_3(t). \tag{17}$$

For simplicity, we set $d_1 = 0$ for now. And with Equation (12), Equations (6) and (16) are rewritten as,

$$\frac{dR_3(t)}{dt} = c_3 R_1(t) = \frac{b_1 c_3}{b_2} I_2(t), \quad \frac{dR_2(t)}{dt} = \frac{b_1 c_2 + b_2 c_1}{b_2} I_2(t).$$

By subtracting one from the other, we have,

$$\frac{d}{dt}\left\{R_3(t) - \frac{b_1 c_3}{b_1 c_2 + b_2 c_1} R_2(t)\right\} = 0 \rightarrow R_3(t) - \frac{b_1 c_3}{b_1 c_2 + b_2 c_1} R_2(t)$$
$$= R_3(0) - \frac{b_1 c_3}{b_1 c_2 + b_2 c_1} R_2(0) = 0.$$

where the initial values, $R_2(0) = R_3(0) = 0$, are used. Thus,

$$R_2(t) + R_3(t) = \frac{c_1}{b_1 t_1 c_3} R_3(t),$$

is obtained and the conserved quantity becomes,

$$1 = S(t) + I_1(t) + \left(1 + \frac{b_1}{b_2}\right) I_2(t) + \frac{c_1}{b_1 t_1 c_3} R_3(t). \tag{18}$$

Here, if $t \rightarrow \infty$, from Equations (13), (14), (15), and (16), we have,

$$-\beta S_*(I_{1*} + I_{2*}) = 0,$$

$$\beta S_*(I_{1*} + I_{2*}) - (b_1 + b_2)I_{1*} = 0,$$

$$b_2 I_{1*} - c_1 I_{2*} = 0,$$

$$\left(c_1 + c_2 \frac{b_1}{b_2}\right) I_{2*} = 0.$$

where $I_i(\infty) = I_{i*}$ and so on. The solutions are given by,

$$I_{1*} = I_{2*} = 0, \quad S_* = \frac{(b_1 + b_2) c_1}{\beta (c_1 + b_2)} = \frac{c_1}{\beta t_1 (c_1 + b_2)}, \tag{19}$$

and

$$R_{3*} = \frac{c_3 t_1}{c_1} b_1 \left(1 - \frac{c_1}{\beta t_1 (c_1 + b_2)}\right) = \frac{c_3 t_1}{c_1} b_1 \left(1 - \frac{c_1}{\mathcal{R}_0 (c_1 + b_2)}\right) \tag{20}$$

where,

$$\mathcal{R}_0 = \beta S_0 t_1 \simeq \beta t_1$$

is used.

In the case where herd immunity is realized in the SIIR model, what are left finite at the end of the infection process are susceptible population, recovered population with antibody, and fatalities. Because cumulative recovered population density can be expressed through cumulative fatalities density, we have,

Cumulative fatalities $\propto (1 - $ susceptible population$)$

$=$ Cumulative density of the infected population.

The cumulative fatalities will have a maximum value as a function of b_1. The coefficient b_1 is the contribution from $I_2(t)$ and $R_1(t)$ through $R_2(t)$, and the contribution from $R_1(t)$ through $R_3(t)$. The coefficient b_1 in the cumulative density of the infected population is a contribution from $I_2(t)$.

The b_{1*} maximizing R_{3*} is found as follows. Differentiating R_{3*} with respect to b_1 and putting the resultant equation to 0, we have,

$$\frac{dR_{3*}}{db_1} = \frac{c_3 t_1}{c_1}\left\{1 - \frac{c_1}{\mathcal{R}_0}\left(\frac{1}{c_1 + b_2} + \frac{b_1}{(c_1 + b_2)^2}\right)\right\} = 0,$$

which is solved to give,

$$b_{1*} = \frac{1}{t_1} - b_{2*} = \frac{1}{t_1} + \frac{1}{t_2}\left\{1 \pm \sqrt{\frac{1}{\mathcal{R}_0}\left(1 + \frac{t_2}{t_1}\right)}\right\}.$$

From the condition $b_1 < 1/t_1$, b_{1*} is obtained as follows:

$$b_{1*} = \frac{1}{t_1} - \frac{1}{t_2}\left(\sqrt{\frac{t_1 + t_2}{\mathcal{R}_0 t_1}} - 1\right). \tag{21}$$

It is confirmed by Equation (20) that the cumulative fatality density R_{3*} in the final state is an increasing function of the fatality rate c_3. Also, if the onset period $(1/c_1)$ becomes shorter, the recovery rate increases, and R_{3*} is a decreasing function of c_1.

2.6 | Comparison between SIIR model and SIR model

A major feature of the SIIR model is that it considers the existence of asymptomatic individuals, namely, $I_2(t)$, and certainly not for the conventional models such as SIR[1] that is defined as follow:

$$\frac{dS(t)}{dt} = -\beta S(t)I(t), \tag{22}$$

$$\frac{dI(t)}{dt} = \beta S(t)I(t) - \gamma I(t) = (\beta S(t) - \gamma)I(t), \tag{23}$$

$$\frac{dR(t)}{dt} = \gamma I(t). \tag{24}$$

SIR model analyzes three transitioning stages (susceptibility (susceptible population $S(t)$), infectious (infectious population $I(t)$), and recovered (and removed) (recovered population $R(t)$)) where infection and recovery rates are denoted as β and γ, respectively. Comparison between SIIR model and SIR model is shown in Table 1.

If a route which connects presymptomatic individuals with asymptomatic individuals is closed and $b_2 = 0$ is set, $I_2(t) = I_2(0)e^{-c1t}$ in the SIIR model disappears. At the same time, if the antibody duration of the recovered individuals is set as infinity ($d_1 = 0$) and $R_1 + R_2 + R_3 = R$ is set, SIIR model is reduced to be SIR model. Thus, it can be said that SIR model treats those who are removed from the infection networks equally and ignores internal structures. Given a finite antibody duration, these who recovered and gained antibody will be included in the group of susceptible individuals over time, but then, this will be excluded from the SIR treatment.

If we take the ratio of the basic reproduction numbers of SIR and SIIR models, we have,

$$\frac{\mathscr{R}_0(SIIR)}{\mathscr{R}_0(SIR)} = \gamma t_1.$$

Since two models may give different basic reproduction numbers for the same infectious disease, policymakers should pay attention to the model they adopt.

TABLE 1 Comparison table for SIIR model and SIR model

Items	SIIR model	SIR model
Variables	6	3
Incubation period	t_1	None
Onset period	t_2	$1/\gamma$
Asymptomatic individuals	Included	Not included
Basic reproduction number $\mathscr{R}_0$	$\beta S_0 t_1$	$\beta S_0/\gamma$
Effective reproduction number $\mathscr{R}_e(t)$	$\mathscr{R}_0 (S(t)/S_0)(1 + I_2(t)/I_1(t))$	$\mathscr{R}_0 S(t)/S_0$
Herd immunity threshold (HIT)	$1 - (1/\mathscr{R}_0)(1 + I_1(t)/I_2(t))$	$1 - 1/\mathscr{R}_0$
Antibody acquisition	All recovered individuals	All recovered individuals
Antibody duration	$1/d_1$	∞
Infection peak	Single and multiple peaks	Single peak

3 | VERIFICATION ON SIIR MODEL BY DATA

Data used here are collected from the daily number of confirmed cases of COVID-19 determined by PCR or antigen tests and announced daily by the prefectural health centers.[11] The variable to be verified by the data is the daily symptomatic and quarantined population in the SIIR model which is denoted by $\Delta R_1(t)$. This is because, among the infected population (presymptomatic, symptomatic, and asymptomatic population), presymptomatic and asymptomatic population have been rarely tested in Japan. Since the observed daily data are under uncontrolled daily fluctuations, $\Delta R_1(t) = b_1 I_1(t)$ calculated from the SIIR model is deviated a bit from the observed daily data. Therefore, we also compare the cumulative confirmed cases of COVID-19 with the cumulative symptomatic and quarantined population $\sum \Delta R_1(t)$ since the cumulative data are almost free from daily fluctuations which are smoothed out by coarse graining.

Below, the results of the SIIR model are compared with the observed data from Japan. Since the population size of the SIIR model is normalized, the population size of the targeted system must be determined from the data. Now, denoting the number of daily confirmed cases at date k by $g(k)$ and assuming N to be the population size of the targeted system. We look for N that minimizes the following equation:

$$\sum_{k=1}^{K_d} (Nb_1 I_1(k) - g(k))^2 \tag{25}$$

where K_d is the total number of data. Although $I_1(k)$ depends on β, t_1, t_2, b_1, and $I_1(0)$, t_1 and t_2 can be fixed as epidemiological parameters.[12,13] From the basic reproduction number $\mathscr{R}_0$, which is obtained from the initial growth curve of the data, β can be determined as $\beta = \mathscr{R}_0/t_1$. Also, b_1 can be estimated from the position of the peak of the data. In addition, the initial value $I_1(0)$ determines the date when the amplitude rises. Thus, N minimizing Equation (25) is calculated as follows:

$$N = \frac{\sum_{k=1}^{K_d} I_1(k)g(k)}{\sum_{k=1}^{K_d} b_1 I_1(k)^2}. \tag{26}$$

Here, note that, even if $b_1 I_1(k)$ and $g(k)$ are similar in profiles, Equation (25) is not necessarily minimized, as long as the date when the amplitude rises are different. Therefore, Equation (26) is used to calculate the population size of the system, and $I_1(0)$ minimizing Equation (25) for given N is numerically calculated.

There are many places of close contacts almost everywhere in Japan. Infections in a group of susceptible population associated with such places could "fire up" separately and simultaneously, forming a peak of infection. If, after the end of the outbreak prevention policy, the infection is not completely terminated, the spread of infection can start again from new close contact between remaining infected individuals and susceptible individuals. Then, following a similar process to the formation of the first peak, peaks of newly infected individuals appear repeatedly with the period of policy. If

policies like self-restraint requests, closures of schools and restraints on conducting business are made nation-widely in a uniform manner without considering the locality of the infection spread, a larger infection spread caused by "fade out of caution" is likely to happen after the policies are lifted.

Extending the SIIR model to describe this chain of spreads is not really difficult. We can build a scenario where remaining presymptomatic and asymptomatic carriers connect several SIIR systems. In fact, it seems that those carriers are moving to discover new areas of spread. It is possible to create n chained SIIR systems and consider that after the first system has ignited and terminated, the remaining presymptomatic and asymptomatic carriers will drive the spread as the initial value for the second system.

Regarding the characteristic values of infectious disease, according to the WHO site, in April 2, 2020, the incubation period of COVID-19 is 5-6 days on average, mainly based on data from Wuhan, and it is infectious during the incubation period.[14] Lauer et al also reported that the median incubation period of COVID-19 was 5.1 days (95% CI, 4.5-5.8 days) and that the infected individuals will become symptomatic within 11.5 days with 97.5% (95%CI, 8.2-15.6 days).[12] It seems that no definitive data analysis has been conducted regarding the onset period, however, according to a survey report by the National Institute of Infectious Diseases (March 23, 2020) in the guidelines of the Ministry of Health, Labor and Welfare in Japan in April 22, 2020, the average length of hospitalization is 16.6 days.[13] So, the following values,

$$t_1 = 5, \quad t_2 = 17,$$

are used in the numerical calculations of the SIIR model. The mortality rate c_3 is set according to the fatalities in the data, and d_1, which is the inverse of the antibody duration, is set as $d_1 = 0$ since there are no evidence data at this moment though several cases are reported from several countries.

Figure 3 (upper row) shows the number of confirmed cases and the cumulative number of confirmed cases in Japan (red) obtained from data and the corresponding number of new symptomatic and quarantined individuals in the SIIR model (yellow).

Here, two normalized SIIR systems $\mathscr{A}_1$ and $\mathscr{A}_2$ are connected at the day $K = 140$ to describe the spread and termination of infection observed in Japan:

$$\mathscr{A}_i = \{ S^{(i)}(t), I_1^{(i)}(t), I_2^{(i)}(t), R_1^{(i)}(t), R_2^{(i)}(t), R_3^{(i)}(t) \}, \quad i = 1, 2,$$

$$1 = S^{(i)}(t) + I_1^{(i)}(t) + I_2^{(i)}(t) + R_1^{(i)}(t) + R_2^{(i)}(t) + R_3^{(i)}(t).$$

The parameters of $\mathscr{A}_1$ are set as follows:

$$t_1 = 5, \quad t_2 = 17, \quad \beta = 0.16, \tag{27}$$

$$b_1 = \frac{0.06}{t_1}, b_2 = \frac{1}{t_1} - b_1, c_1 = \frac{1}{t_2},$$
$$c_2 = \frac{0.942}{t_2}, c_3 = \frac{1}{t_2} - c_2, d_1 = 0. \tag{28}$$

The initial values of $\mathscr{A}_1$ are as follows:

$$S^{(1)}(0) = 1 - I_1^{(1)}(0), \quad I_2^{(1)}(0) = R_1^{(1)}(0) = R_2^{(1)}(0) = R_3^{(1)}(0) = 0. \tag{29}$$

Then, the following values,

$$N_1 = 316067, \quad I_1(0) = 0.00004,$$

are obtained from Equations (25) and (26). The total number of observation data is set as $K = 140$ days.

After the spread of infection in system $\mathscr{A}_1$ met the condition of herd immunity and settled down, some remaining infectious individuals in $\mathscr{A}_1$ contacted susceptible individuals in system $\mathscr{A}_2$ and caused spread of infection in $\mathscr{A}_2$. This relay of infections from $\mathscr{A}_1$ to $\mathscr{A}_2$ happened at $t = K$ (=140). The parameters for $\mathscr{A}_2$ are taken to be the
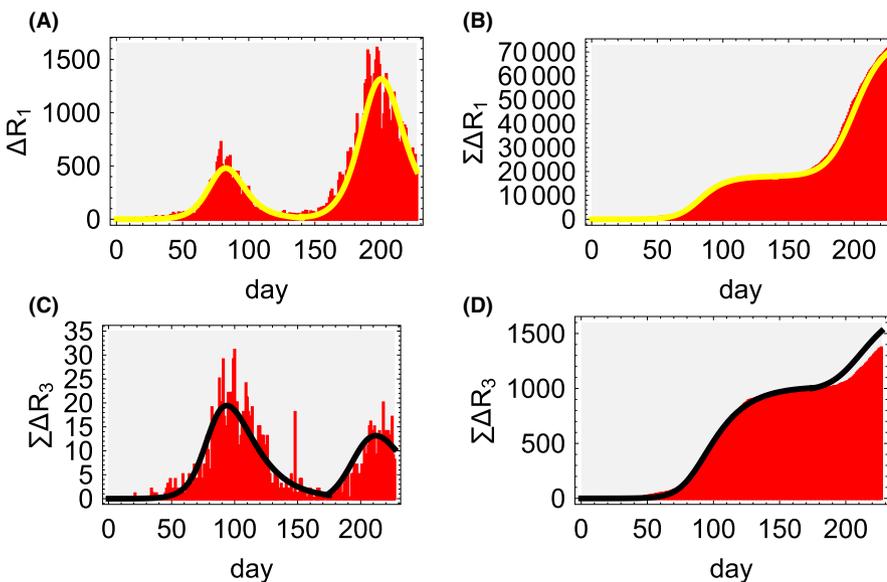
FIGURE 3 Upper left: the number of confirmed cases in Japan (data from January 24 to September 6, 2020: red) and that of new symptomatic and quarantined individuals $b_1 I_1(t)$ (SIIR model: yellow). Upper right: the cumulative number of confirmed cases in Japan (red) and that of new symptomatic and quarantined individuals $\sum_t b_1 I_1(t)$ (SIIR model: yellow). Lower left: daily fatalities by COVID-19 in Japan (red) and $c_3 R_1(t)$ (SIIR model: black). Lower right: cumulative fatalities by COVID-19 in Japan (red) and $\sum_t c_3 R_1(t)$ (SIIR model: black)

same as Equations (27) and (28) in $\mathscr{A}_1$. The initial values at $t = K$ are set as follows:

$$S^{(2)}(K) = 1 - I_1^{(2)}(K), \quad I_2^{(2)}(K) = R_1^{(2)}(K) = R_2^{(2)}(K) = R_3^{(2)}(K) = 0, \quad (30)$$

where the population size of $\mathscr{A}_2$ is calculated from Equations (25) and (26) as follows:

$$N_2 = 1060790, \quad I_1^{(2)}(K) = 0.002.$$

Note that the basic reproduction number here is $\mathscr{R}_0 = 0.79968$ (<1), and the infection does not spread in the SIR model; however, it spreads in the SIIR model as shown in Equation (8).

The asymptomatic population can be obtained from the SIIR models and data. From Equation (12), it can be seen that the asymptomatic population $I_2(t)$ is proportional to the symptomatic and quarantined population $R_1(t)$ and with the given parameters the ratio is given by,

$$\frac{I_2(t)}{R_1(t)} = \frac{b_2}{b_1} = 15.67.$$

In other words, the asymptomatic population in Japan can be approximately 16 times greater than symptomatic and quarantine population. Data also show that asymptomatic individuals who had refrained from going out during the state of emergency declaration went out again after it was lifted, made new contacts, and spread of infection. From this, a major factor leading to the spread of infection could be understood because of the fact that the main target of the PCR test in Japan has been limited to people with subjective symptoms on COVID-19. It is said that the reason for limiting the PCR test in Japan was to prevent medical collapse. However, it is necessary to verify whether the policies taken with the supplementary budget should have been directed to prevent medical collapse.

Now, we take a close look at fatalities caused by COVID-19 in Japan, which can be reproduced quite well with the SIIR model (Figure 3 (lower row). Since R_1 and R_3 are on different stages, there is a difference between the timings when the peaks occur. Therefore, in Figure 3 (lower row), the time when $\mathscr{A}_1$ connects to $\mathscr{A}_2$ is set as $K = 170$, which is different from $K = 140$ used in Figure 3 (upper row), and the fatality rates are set as follows:

$$c_3 = \begin{cases} \dfrac{0.058}{t_2} & \text{for} \quad t \leq K, \\ \dfrac{0.014}{t_2} & \text{for} \quad t \rangle K. \end{cases}$$

In $\mathscr{A}_2$, unlike $\mathscr{A}_1$, more of the infected are young, fewer elderly people with preexisting medical conditions are infected, and, in addition, healthcare professionals have gained experience in $\mathscr{A}_1$ so that they can respond appropriately to treatment. However, in the future, if the infection spreads from the young to the elderly, and if the "GO TO TRAVEL" campaign in Japan boosts the spread of infection to more people including those with preexisting medical conditions and the elderly, then fatalities may also increase. Finding and isolating asymptomatic individuals can help control the spread of the infection.

4 | DISCUSSIONS

In this paper, the SIIR model, dealing with the features of COVID-19 such as presymptomatic individuals who have infectivity even during the incubation period, asymptomatic individuals who can freely move around and play crucial roles in the spread of infection, and limited duration of immunity, has been introduced to show that it can replicate data such as the number of confirmed cases for COVID-19 and that of those who passed away because of COVID-19 in Japan.

Since infectious diseases spread depending on individual's health conditions and behavioral characteristics, the SIIR model cannot provide detailed information for policymaking, as far as our discussion is limited within median or mean. This is the limit of the SIIR model. In the future, we will address the following topics in order to discuss policy issues for the early termination of COVID-19.

- Verification of the SIIR model based on the spread of COVID-19 infection in various countries.
- Formulation of agent-based modeling (ABM) that incorporates diversity such as individual health conditions and behavioral characteristics and verification of ABM by actual data.
- Examination of detailed policies corresponds toward the convergence of infectious diseases.

CONFLICT OF INTEREST
The authors have stated explicitly that there are no conflicts of interest in connection with this article.

AUTHOR CONTRIBUTIONS
All authors had access to the data and a role in writing the manuscript.

INFORMED CONSENT
There is no need to obtain any consent for this article.

ORCID
Masaki Tomochi https://orcid.org/0000-0003-2952-8637

REFERENCES
1. Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics. Proc Roy Soci A. Math Phys Eng Sci. 1927;1:815–21.
2. May RM, Anderson B. Infectious diseases of humans: Dynamics and control. Oxford, England: Oxford University Press; 1992.
3. Iannelli M. Mathematical theory of age structured population dynamics. Pisa: Giardini Editori e Stampatori; 1995.
4. Blower S, Bernoulli D. An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it. Rev Med Virol. 2004;14:275–88.
5. Grimm V, Railsback SF. Individual-based modeling and ecology. Princeton, NJ: Princeton Univ. Press; 2005.

6. Inaba H. Mathematical model for infectious disease. Tokyo, Japan: Baihukan; 2008.

7. Chen Y, Lu P, Chang C, Liu T.A Time-dependent SIR model for COVID-19 with undetectable infected persons. Feb 2020. Available from https://arxiv.org/abs/2003.00122

8. Sameni R.Mathematical modeling of epidemic diseases; A case study of the COVID-19 coronavirus. Mar 2020. Available from https://arxiv.org/abs/2003.11371

9. Jewell NP, Lewnard JA, Jewell B. Predictive mathematical models of the COVID-19 pandemic. J Am Med Assoc. 2020;323(19):1893–94.

10. Heidt A. Studies report rapid loss of COVID-19 antibodies. The Scientist. June 2020. Available from https://www.the-scientist.com/news-opinion/studies-report-rapid-loss-of-covid-19-antibodies67650

11. Japan Broadcasting Corporation. www.nhk.or.jp [internet]. Tokyo: Available from https://www3.nhk.or.jp/news/special/coronavirus/data-all/

12. Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, et al. The incubation Period of coronavirus disease 2019 from publicly reported confirmed cases: estimation and application. Ann Intern Med. 2020;172(9):577–82.

13. National Institute of Infectious Diseases. March 2020. Available from https://www.mhlw.go.jp/content/000624099.pdf

14. WHO. Coronavirus disease 2019 (COVID-19) Situation Report – 73. Available from https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200402-sitrep-73-covid-19.pdf?sfvrsn=5ae25bc76