Estimation of the onset ratio and the number of asymptomatic patients of COVID-19 from the proportion of untraceable patients

Takashi Odagaki[1,2*]

[1]Kyushu University, Fukuoka, 819-0395, Japan
[2]Research Institute for Science Education, Inc., Kyoto, 603-8346, Japan.
*Corresponding author. Email: t.odagaki@kb4.so-net.ne.jp

Abstract

A simple method is devised to estimate the onset ratio of COVID-19 patients from the proportion of untraceable patients tested positive, which allows us to obtain the number of asymptomatic patients, the number of infectious patients and the effective reproduction number. The recent data in Tokyo indicate that there are about six to ten times as many infectious patients in the city as the daily confirmed new cases. It is shown that, besides social distancing and use of effective masks, a quarantine measure on non-symptomatic patients is critically important in controlling the pandemic.

Introduction

The pandemic COVID-19 is still prevalent in most parts of the world despite of continuous efforts by governments to control it. Difficulty of the control lies in the fact that patients of COVID-19 take a route different from common epidemics. In common epidemics, infected individuals show symptoms and become infectious after an incubation period. Then they are treated and recover from the disease. In COVID-19, infected individuals become infectious before symptom-onset and asymptomatic patients who do not show any symptoms before their recovery are infectious. In order to formulate proper strategies in controlling COVID-19, it is important to know the proportion of asymptomatic and pre-symptomatic patients [1-3], who will be called non-symptomatic patients collectively, in addition to the number of symptomatic patients. However, it is unpractical to identify all non-symptomatic patients in the entire population by PCR tests. Therefore, it is an important problem to devise a method for estimating the number of non-symptomatic infectious patients from data reported daily such as the confirmed new cases and the proportion of untraceable patients tested positive. Here, symptomatic patients are divided into two groups; one is traceable patients who know from whom they have gotten infected and the other is untraceable patients who have had no contact with pre-symptomatic and symptomatic patients.

In this paper, I propose a simple method by which the onset ratio x of COVID-19 patients can be estimated from the proportion f of untraceable patients tested positive and show that the proportion of the infectious patients can be obtained from the proportion of untraceable patients. I first analyze the infection process on the basis of the SIQR (Susceptible-Infectious-Quarantined-Removed) model [4] and find a relation between x and f. Then, I argue that the number of infectious patients and the number of new patients on a given day can be related to the proportion f. I also

discuss the effective reproduction number which depends on a quarantine rate of non-symptomatic patients and show that the quarantine of non-symptomatic patients is critically important in controlling COVID-19. I analyze the situation in Tokyo and discuss why COVID-19 does not converge in Tokyo.

Classification of infectious patients

I classify infected individuals on the basis of the SIQR model [4] and the mean field approach as follows. I first assume that patients of COVID-19 follow the same disease progression day by day at the same pace on average since their infection. Namely, I assume that the timeline of infection follows infected, infectious, symptom/quarantined or asymptomatic and that the time between these events is deterministic. On a particular day which I call day zero, there are many infected individuals who can be classified by the number of days since their infection. The number of patients, some of whom are identified as patients by PCR test and quarantined on the day zero, is denoted by n_0. I denote the average onset ratio of COVID-19 by x, and then the number of infected symptomatic individuals in the group of patients is given by xn_0 and the number of asymptomatic individuals in the group of patients is given by $(1-x)n_0$. I denote by $n_{-1}, n_{-2}, \cdots, n_{-k}$ the number of infected individuals who got infected one, two,$\cdots$, k days later than the day when the n_0 individuals got infected. Similarly, I denote by $n_1, n_2, \cdots, n_{\ell-1}$the number of

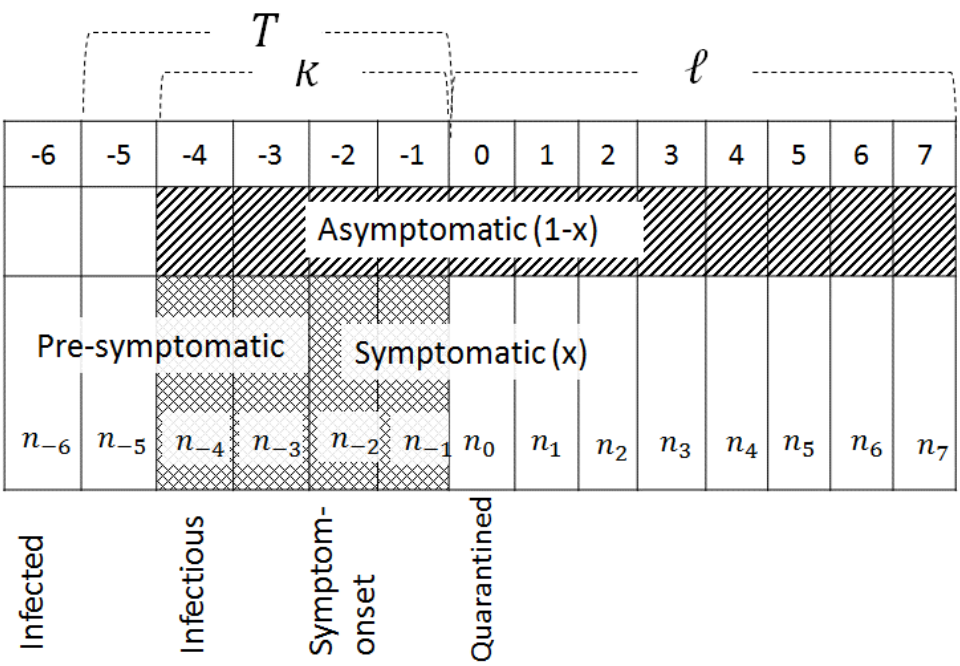

Fig.1. Break down of COVID-19 patients into asymptomatic, pre-symptomatic and symptomatic patients. Patients in the shaded area are infectious. The hatched region represents infectious asymptomatic patients who produce untraceable patients and the cross-hatched region corresponds to infectious pre-symptomatic and symptomatic patients who produce traceable patients. T denotes the effective period of PCR test before quarantine.

individuals who got infected one, two, $\cdots, \ell - 1$ days earlier than the day when the $n_0$individuals got infected [5]. Here, k is the sum of the infectious period before symptom-onset and the period between the onset date and the quarantined date, and ℓ denotes the infectious period of asymptomatic patients after symptomatic patients in the same group are quarantined. I assume that the proportion $1 - x$ of asymptomatic patients is common in every group of patients. Figure 1 shows schematically the breakdown of patients into these three groups, where patients in the shaded area are infectious. Note that in the present analysis, the latent and incubation periods do not play any roles.

I assume that all symptomatic patients will be tested intentionally some days after the symptom-onset and quarantined. I also assume that PCR tests, which is effective since T days before day zero, are conducted on the general population and denote the quarantine rate of patients by q. If the quarantine measure on infected individuals is taken effectively, these infected patients decrease by a factor $(1 - q)$ every day. Therefore, the number of infectious patients $\mathfrak{I}$ on day zero is given by

$$\mathfrak{I} = \sum_{i=1}^{\mathrm{k}} n_{-i}(1-q)^{T-i} + (1-x)\sum_{i=0}^{\ell-1} n_i(1-q)^{T+i}, \tag{1}$$

and the new patients ΔI infected on day zero who will be identified some days later are given by

$$\Delta I = (1-\mu)(1-a)\left\{\sum_{i=1}^{\mathrm{k}} \beta_{-i}\, n_{-i}(1-q)^{T-i} + (1-x)\sum_{i=0}^{\ell-1} \beta_i\, n_i(1-q)^{T+i}\right\}, \tag{2}$$

where β_j is the infection coefficient of patient group n_j, μ is the fraction of immunized individuals who are no longer susceptible, and a represents the reduction ratio of social contacts among people due to lockdown measures. Here, I assume $T \geq k$ for simplicity.

In order to make the following description transparent, I define an average of n_j and a weighted average of β_j as follows:

average of $n_{-1}, n_{-2}, \cdots, n_{-k}$

$$\langle n \rangle_k = \frac{1}{k}\sum_{i=1}^{k} n_{-i}\,(1-q)^{T-i}, \tag{3}$$

average of $n_0, n_1, n_2, \cdots, n_{\ell-1}$

$$\langle n \rangle_\ell = \frac{1}{\ell}\sum_{i=0}^{\ell-1} n_i(1-q)^{T+i}, \tag{4}$$

weighted average of $\beta_{-1}, \beta_{-2}, \cdots, \beta_{-k}$

$$\langle \beta \rangle_k = \sum_{i=1}^{k} \beta_{-i} n_{-i}\,(1-q)^{T-i} \Big/ \sum_{i=1}^{k} n_{-i}\,(1-q)^{T-i}, \tag{5}$$

weighted average of $\beta_0, \beta_1, \beta_2, \cdots, \beta_{\ell-1}$

$$\langle \beta \rangle_\ell = \sum_{i=0}^{\ell-1} \beta_i n_i\,(1-q)^{T+i} \Big/ \sum_{i=0}^{\ell-1} n_i(1-q)^{T+i}, \tag{6}$$

and write Eqs. (1) and (2) as

$$\mathfrak{I} = k\langle n \rangle_k + (1-x)\ell\langle n \rangle_\ell \tag{7}$$

and

$$\Delta I = (1-\mu)(1-a)\{k\langle\beta\rangle_k\langle n\rangle_k + (1-x)\ell\langle\beta\rangle_\ell\langle n\rangle_\ell\}. \tag{8}$$

Relation between the proportion of untraceable patients and the onset ratio

In this section, I argue that the traceability of patients can be related to infection from asymptomatic patients and derive a relation between the onset ratio and the proportion of untraceable patients. Out of the newly infected individuals ΔI, $\Delta Q \equiv x\Delta I$ will show symptoms some days later and $(1-x)\Delta I$ will not show any symptoms. Patients ΔQ showing symptoms will be identified by PCR tests and be listed as daily confirmed new cases. I assume that they are classified into two groups ΔQ_t and ΔQ_{unt}, where ΔQ_t are traceable patients who are infected from symptomatic and pre-symptomatic patients and ΔQ_{unt} are untraceable patients who are infected from asymptomatic patients. Namely, ΔQ_t and ΔQ_{unt} are given by

$$\Delta Q_t = (1-\mu)(1-a)x^2 k\langle\beta\rangle_k\langle n\rangle_k, \tag{9}$$

$$\Delta Q_{unt} = (1-\mu)(1-a)x(1-x)\{k\langle\beta\rangle_k\langle n\rangle_k + \ell\langle\beta\rangle_\ell\langle n\rangle_\ell\}. \tag{10}$$

Therefore, the proportion of untraceable cases in the daily confirmed new cases is given by

$$f \equiv \frac{\Delta Q_{unt}}{\Delta Q} = \frac{(1-x)[k\langle\beta\rangle_k\langle n\rangle_k + \ell\langle\beta\rangle_\ell\langle n\rangle_\ell]}{k\langle\beta\rangle_k\langle n\rangle_k + (1-x)\ell\langle\beta\rangle_\ell\langle n\rangle_\ell}, \tag{11}$$

which does not depend on μ and a explicitly. This relation can be inverted to get

$$x = \frac{1-f}{1-\frac{\ell\langle\beta\rangle_\ell\langle n\rangle_\ell}{k\langle\beta\rangle_k\langle n\rangle_k + \ell\langle\beta\rangle_\ell\langle n\rangle_\ell}f}, \tag{12}$$

which indicates that the onset ratio x can be obtained from the proportion of untraceable patients f once other parameters are known.

In order to incorporate the trend of infection status [6] into the present analysis, I first define trend parameters τ_k and τ_ℓ by $\tau_k = \langle n\rangle_k/(1-q)^T n_0$ and $\tau_\ell = \langle n\rangle_\ell/(1-q)^T n_0$. It is apparent that the infection status after day zero will be increasing, stationary and decreasing when $\tau_k > 1$, $\tau_k = 1$ and $\tau_k < 1$, respectively. Similarly, $\tau_\ell > 1$, $\tau_\ell = 1$ and $\tau_\ell < 1$ indicate that the infection status before day zero has been decreasing, stationary and increasing, respectively. It is straight forward to express x in terms of τ_0 and τ_1 as

$$x = \frac{1-f}{1-\frac{\ell\langle\beta\rangle_\ell\tau_\ell}{k\langle\beta\rangle_k\tau_k + \ell\langle\beta\rangle_\ell\tau_\ell}f}. \tag{13}$$

Therefore, ΔI is written as

$$\Delta I = (1-\mu)(1-a)\left\{k\langle\beta\rangle_k\tau_k + \frac{\langle\beta\rangle_k\langle\beta\rangle_\ell \mathrm{k}\tau_k\ell\tau_{\mathcal{L}}}{k\langle\beta\rangle_k\tau_k + \ell\langle\beta\rangle_\ell\tau_\ell(1-f)}f\right\}(1-q)^T n_0. \tag{14}$$

Since the confirmed new cases on day zero $(\Delta Q)_0$ is given by $(\Delta Q)_0 = x(1-q)^T n_0$, $\mathfrak{I}$ can be expressed in terms of τ_k, τ_ℓ and $(\Delta Q)_0$ as

$$\mathfrak{I} = \frac{1}{1-f}\left[k\tau_k + \frac{(\langle\beta\rangle_k - \langle\beta\rangle_\ell)k\tau_k\ell\tau_\ell}{k\langle\beta\rangle_k\tau_k + \ell\langle\beta\rangle_\ell\tau_\ell}f\right](\Delta Q)_0. \tag{15}$$

Number of infectious patients and asymptomatic patients in Tokyo

As an example of analysis, I apply the present analysis to the infection status in Tokyo in the middle of June, 2021. In Tokyo, PCR tests have not conducted on non-symptomatic patients and I can set $q = 0$. I assume that the infection status was stationary in the middle of June, 2021, and set $\tau_k = \tau_\ell = 1$ and $\langle\beta\rangle_k = \langle\beta\rangle_\ell$. Then, Eq. (13) reduces to

$$x = \frac{1-f}{1-\frac{\ell}{k+\ell}f}. \tag{16}$$

According to Ministry of Health, Labour and Welfare of Japan (MHLW), transmission occurs in the 2 days before symptom-onset and in the 7∼10 days post symptom-onset [7]. These estimations are consistent with those reported in Ontario, Canada: Ontario Agency for Health Protection and Promotion reported that the transmission occurs in the 3∼5 days before onset and in the 8∼10 days post onset [5]. Assuming on average a patient will be tested and quarantined next day of symptom-onset, I set as a model case $k = 3days$ and $\ell = 7days$. Then, Eq. (16) is reducible to

$$x = \frac{1-f}{1-\frac{7}{10}f}, \tag{17}$$

which is shown in Fig.2. In Tokyo, the proportion of untraceable patients is $f = 50\sim60\%$ [8]. Using the lower value f = 50 %, I find $x = 77\%$. This value becomes $x = 69\%$ if $f = 60\%$ is used. These values are consistent with observations of the onset ratio $x = 76\%$ [9] or 75% [10].

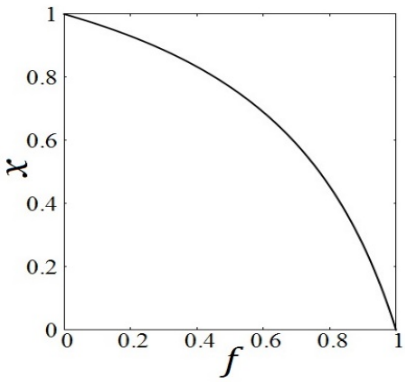

Fig. 2. The onset ratio x is plotted as a function of the proportion of untraceable patients f when $k = 3days$ and $\ell = 7days$.

On the same conditions, the number of infectious patients Eq. (15) can be written as

$$\frac{\Im}{(\Delta Q)_0} = \frac{k}{1-f}. \tag{18}$$

If there were no untraceable patients, all infected individuals would be symptomatic and, therefore, the number of infectious patients will be given by the number of new cases times days during which they are infectious, i.e.

$\Im/(\Delta Q)_0 = k$ when $f = 0$ as Eq. (18) indicates. Equation (18) shows $\Im/(\Delta Q)_0 = 2k$ and $2.5k$ when $f = 50{\sim}60\%$ respectively. Figure 3 shows f dependence of $\Im/(\Delta Q)_0$ when $k = 3days$ and $\ell = 7days$. It is important to note that in Tokyo $f = 50{\sim}60\%$ and $k = 3days$, and thus there are 6~7.5 times more infectious patients than the daily

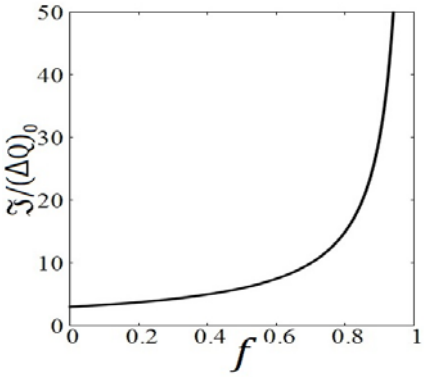

Fig. 3. The ratio of the number of infectious patients in the city $\Im$ and the daily confirmed new cases $(\Delta Q)_0$ is plotted as a function of the proportion of untraceable patients f when $k = 3days$.

confirmed new cases. It should also be mentioned that the number of asymptomatic infectious patients $\Im_{as}$ excluding pre-symptomatic patients is given by

$$\frac{\Im_{as}}{(\Delta Q)_0} = \frac{fk}{1-f}, \tag{19}$$

and thus $\Im_{as}/(\Delta Q)_0 = 3$ when $k = 3days$, $\ell = 7days$ and $f = 50\%$.

In January 2022, the number of patients in Tokyo was increasing due to the omicron variant. I can assume that the difference between $\langle\beta\rangle_k$ and $\langle\beta\rangle_\ell$ is small and q is negligible in Eq. (15). From the infection curve increasing exponentially with rate 0.2, τ_k is estimated to be $\tau_k = 1.24$. In this period, the proportion of untraceable patients was 65.5% [8], and thus, setting $k = 3days$, I get $\Im/(\Delta Q)_0 \approx 11$.

Around the end of February 2022, the sixth wave in Tokyo [8] seems to be in the stationary state with $f = 61\%$, and thus I can set $\tau_k = 1$. Therefore, $\Im/(\Delta Q)_0$ is estimated to be $\Im/(\Delta Q)_0 \approx 7.7$.

Effective reproduction number and assessment of policies

I first define an effective infectious period d_{eff} and an effective infection coefficient β_{eff} as follows:

$$d_{eff} = (1-q)^T\{k\tau_\ell + (1-x)\ell\tau_\ell\}, \tag{20}$$

$$\beta_{eff} = \frac{k\langle\beta\rangle_k\tau_k + (1-x)\ell\langle\beta\rangle_\ell\tau_\ell}{k\tau_\ell + \ell\tau_\ell}. \tag{21}$$

Equation (8) can now be expressed as

$$\Delta I = (1-\mu)(1-a)\beta_{eff}d_{eff}n_0, \tag{22}$$

and the effective reproduction number R_{eff} on day zero is given by

$$R_{eff} \equiv \frac{\Delta I}{n_0} = (1-\mu)(1-a)\beta_{eff} d_{eff}. \tag{23}$$

It should be remarked that when d_{eff} is equal to the recovery time or the inverse of the recovery rate γ and β_{eff} is a constant, the expression for R_{eff} is identical to that defined in the standard SIR model. It is important to note that, in contrast to the standard SIR model, the effective reproduction number R_{eff} defined by Eq. (23) depends on the quarantine rate of non-symptomatic patients through d_{eff}, which has a significant effect in reducing the effective reproduction number.

The basic strategy against COVID-19 is to bring the effective reproduction number smaller than unity so that the number of patients decreases. Equation (23) indicates that there are three tactics: (1) increase immunized people by vaccination, if it is effective in preventing transmission of SARS-CoV-2, or by natural immunity to make $(1-\mu)$ smaller, (2) enforce social distancing by various lockdown measures to make $(1-a)$ smaller, and (3) quarantine asymptomatic patients by PCR tests to make d_{eff} smaller. In countries like Taiwan, Australia and New Zeeland who have succeeded in controlling COVID-19 before the vaccination was started, PCR tests have been conducted more than 100 times per positive patient in a well-designed manner.

In Tokyo, the PCR test has been used only to confirm the infection of SARS-CoV-2 for people who show some symptoms, which means $q \approx 0$, and thus it has not been contributing to the battle against COVID-19. Furthermore, lockdown measures have been very sloppy. Instead of enforcing social distancing among the entire population it has been applied only specific targets, like the night life district on 2020 and restaurants serving alcoholic beverages in 2021 and 2022. The policy targeting on certain shops and opening hours has only limited effects on social distancing since people gather together in parks or on streets. Furthermore, the policy has been enforced and lifted every one or two months, which has caused the wavy infection curve [11,12]. It could be possible to increase a by, for example, promoting Telework, limiting working days, reducing crowd in commuter trains and banning gatherings.

Discussion

In the present analysis, traceable and untraceable patients tested positive are related to infection from pre-symptomatic and symptomatic patients and from asymptomatic patients, respectively. Although this assumption may not be rigorous, it is a good assumption if people cooperate in the investigation of infection route by the health department of local government.

When the infection status is increasing or decreasing [6], the relation Eq. (15) between the number of infectious patients and the number of the newly confirmed new cases reduces to

$$\frac{\Im}{(\Delta Q)_0} = \frac{k\tau_k}{1-f} \tag{24}$$

if the condition $\langle\beta\rangle_k = \langle\beta\rangle_\ell$ is satisfied. This ratio becomes large in the increasing status and small in the decreasing status compared to Eq. (18). In this case, the effective reproduction number depends on the infection status through d_{eff}. In general, d_{eff} is an increasing function of τ_k and τ_ℓ. Since the increasing state corresponds to $\tau_k > 1 > \tau_\ell$ and the decreasing state to $\tau_k < 1 < \tau_\ell$, d_{eff} can increase or decrease depending on the relation between τ_k and τ_ℓ.

The infection coefficient of SARS-CoV-2 depends on variants. When a new variant with stronger infection coefficient emerges, $\langle\beta\rangle_k$ becomes larger than $\langle\beta\rangle_\ell$. The number of infectious patients will increases for a stronger variant when other parameters are the same.

In 2021, vaccination has been progressed in many countries and the infection status seems to be improving at least in the reduction of serious cases. The effect of vaccination appears through μ in Eq. (2) which depends on variants of the virus. Therefore policies for controlling COVID-19 should not rely only on the vaccination, and a proper combination of policies on vaccination, social distancing and quarantine of non-symptomatic patients must be designed in each country around the world [13].

Conclusion

It has been shown that the onset ratio and the number of infectious patients can be estimated from the proportion of untraceable patients tested positive. The effective reproduction number depends not only on the vaccination rate and effects of social distancing but also on d_{eff} which is controlled by quarantine measures on pre-symptomatic and asymptomatic patients. The effect of the quarantine measure appears as reduction of the effective reproduction number by $(1-q)^T$ which could be significantly large compared to the effects of the lockdown measures and the vaccination.

References

1. D. P. Oran and E. J, Topol, Prevalence of Asymptomatic SARS-CoV-2 Infection. Ann. Intern. Med. **174**, 286-287 (2021). https://doi.org/10.7326/L20-1285.
2. R. Subramanian, Q. He, and M. Pascual, Quantifying asymptomatic infection and transmission of COVID-19 in New York City using observed cases, serology, and testing capacity, Proc. Nat. Acad. Sci., **118**, e2019716118 (2021). https://doi.org/10.1073/pnas.2019716118
3. B. T. Montague, M. F. Wipperman, A. T. Hooper, S. C. Hamon, R. Crow, F. Elemo, L. Hersh, S. Langdon, J. D. Hamilton, M. P. O'Brien, E. A. F. Simões. Anti-SARS-CoV-2 IgA Identifies Asymptomatic Infection in First Responders. J. Infect. Dis. **225**, 578-586 (2022). https://doi.org/10.1093/infdis/jiab524
4. T. Odagaki, Analysis of the outbreak of COVID-19 in Japan by SIQR model, Infectious Disease Modelling, **5**, 691-698 (2020). https://doi.org/10.1016/j.idm.2020.08.013
5. Ontario Agency for Health Protection and Promotion (Public Health Ontario). COVID-19 overview of the period of communicability – what we know so far. Toronto, ON: Queen's Printer for Ontario; 2021. https://www.publichealthontario.ca/-/media/documents/ncov/covid-wwksf/2021/03/wwksf-period-of-communicability-overview.pdf?la=en
6. T. Odagaki and R. Suda, Classification of the infection status of COVID-19 in 190 countries, J. Clinical Trials **11,** 472 (2021). https://doi.org/10.1101/2020.12.17.20248445
7. Ministry of Labour, Health and Welfare, Japan, "Guidance for treatment of COVID-19" ver5.1, July, 5, (2021). https://www.mhlw.go.jp/content/000801626.pdf

8. Tokyo Metropolitan Government, “Updates on COVID-19 in Tokyo”, July 20 (2021) https://www.fukushihoken.metro.tokyo.lg.jp/iryo/kansen/corona_portal/info/monitoring.images/0715graph3-1.JPG

9. M. A. Johansson, T. M. Quandelacy, S. Kada, P. V. Prasad, M. Steele *et al*, SARS-CoV-2 Transmission From People Without COVID-19 Symptoms, JAMA Network Open. **4**(1) e2035057 (2021). http://doi:10.1001/jamanetworkopen.2020.35057

10. M. Alene, L. Yismaw, M. A. Assemie. D. B. Ketema, B. Mengist *et al*, Magnitude of asymptomatic COVID-19 cases throughout the course of infection: A systematic review and meta-analysis, PLoS ONE **16**(3), e0249090 (2021). https://doi.org/10.1371/journal.pone.0249090

11. T. Odagaki, Self-organized wavy infection curve of COVID-19, Scientific Reports, **11**, 1936 (2021). https://doi.org/10.1038/s41598-021-81521-z

12. T. Odagaki, Self-organization of oscillation in an epidemic model for COVID-19, Physica A **573**, 125925 (2021). https://doi.org/10.1016/j.physa.2021.125925

13. A. Singh, M. Arquam, Epidemiological modeling for COVID-19 spread in India with the effect of testing, Physica A: Statistical Mechanics and its Applications **592**, 126774 (2022). https://doi.org/10.1016/j.physa.2021.126774

Acknowledgments

This work was supported in part by JSPS KAKENHI Grant Number 18K03573.