On the effectiveness of the search and find method to suppress spread of SARS-CoV-2

By Hirotaka SUGAWARA*1,†

(Edited by Toshimitsu YAMAZAKI, M.J.A.)

Abstract: Search and find methods*) such as cluster tracing1)–6) or large-scale PCR testing**) of those who exhibit no symptoms or only mild symptoms of COVID-19 is shown by data analysis to be a powerful means to suppress the spread of COVID-19 instead of, or in addition to, lockdown of the entire population. Here we investigate this issue by analyzing the data from some cities and countries and we establish that search and find method is as powerful as lockdown of a city or a country. Moreover, in contrast to lockdown, it neither causes inconvenience to citizens nor does it disrupt the economy. Generally speaking, it is advisable that both social distancing and increased test numbers be employed to suppress spread of the virus. The product of the total test number with the rate of positive cases is the crucial index.

Keywords: SARS-CoV-2, COVID-19, cluster, PCR test, SIR model

1. Introduction—illustrative example to show the method—

We consider first the case of Tokyo to illustrate our method of analysis. In this introductory section, we focus on the number of those who are infected, but the origin of whose infections cannot be traced to a “cluster”: The Japanese government has been initially concentrating on “cluster tracing”, in which they assumed that the major part of the infections originate from a “cluster”. But the subsequent rapid increase in infections without clear origin indicates a weakness of concentrating only on clusters. In later sections we do not distinguish between infections with or without the knowledge of their sources. Both cases are taken into account by a single parameter \( \sigma \) to be defined later. “Cluster tracing” increases \( \sigma \) by increasing the positivity rate of tests, and large-scale PCR testing increases \( \sigma \) by increasing the number of tests.

In Tokyo, the daily number of those infected by the virus whose origin of infection cannot be traced, between March 25 and April 4 (https://stopcovid19.metro.tokyo.lg.jp/en/), is:

\[
\begin{align*}
13, 22, 20, 22, 24, 7, 50, 40, 38, 59, 80
\end{align*}
\]

This is depicted in Fig. 1. We fit the data with a function of the form \( \exp[5x] \) where \( 5 \) is a constant and \( x \) is the date variable (Fig. 2). We note the following:

1. The fitting using Mathematica (least squares method) gives a curve: \( y = 11.0 \exp[0.15x] \).

2. This result shows that it takes 5 days for the infection number with unknown sources to double. \( 5 > 0 \) during this period between March 25 and April 4.

3. To decrease the number of the infected, we must have \( \sigma < 0 \).

How can we achieve it? Obviously, fitting by a linear polynomial is not sufficient.

2. Model construction

We develop a model on the following naive considerations and justify it later in Appendix 1

---

*) The Japanese government adopted this policy at the onset of COVID-19 infection in Japan (https://www.mhlw.go.jp/content/109060000000599837.pdf (in Japanese)). This policy is particularly effective when there are “superspreaders”, for which there is some evidence as observed in Refs. 1–6.

**) See Appendix 6, and the references cited there, on the sensitivity of the test. The importance of PCR testing for suppression of the spread of SARS-CoV-2 is stressed, for example, in a message from OECD (https://www.oecd.org/coronavirus/policy-responses/testing-for-covid-19-a-way-to-lift-confinement-restrictions-89756248/). Some scientific analyses include the research by Takashi Odagaki (see Ref. 7).
based on the standard model of epidemiology, i.e., the SIR model.8),9)

We use the following notation and derive a simple formula taking the Tokyo case as an example:

1. \(N_T\) the number of infections in Tokyo on the date \(T\);
2. \(n_T\) those who were officially recognized among \(N_T\) and therefore isolated;
3. \(U_T = N_T - n_T\) those who stay in town unrecognized as infected;
4. \(\kappa U_T\) total increase of number of infected per day (negative \(\kappa\) means the decrease). Note that those who are officially recognized as infected do not contribute to this number.

5. Suppose we put \(T = \text{March 25}\), then the number of those infected after the \(m\)-th day becomes,

\[
\prod_{n=1}^{m} (1 + \kappa(n)) U_{25} = \exp \left[ \sum_{n=1}^{m} \log(1 + \kappa(n)) \right] U_{25}
\]

\[ \simeq \exp \left[ \sum_{n=1}^{m} \kappa(n) \right] U_{25}. \quad [2] \]

6. This can be written in the continuous limit as,

\[ \exp \left[ \int_0^x \kappa(t)dt \right] U_{25}. \quad [3] \]

And it is a solution to,

\[ dU/dx = \kappa(x)U. \quad [4] \]

Here we use the continuous variable \(x\) instead of \(m\).

In the main body of this paper we use \(x\) as the time variable henceforth but we use \(t\) to signify the time variable in the appendices.

7. The above data shown in Eq. [1] and in Fig. 1 are for \(dn/dx\) and not for \(U\). How are they related?

Before answering this question, we point out that \(\kappa\) is composed of three components, as derived in Appendix 1:

\[ \kappa = K - \rho - \sigma, \quad [5] \]

where \(K\) stands for the contribution of the infection rate, \(\sigma\) will be defined later in Eq. [6], and \(\rho\) stands for the recovery rate. We also note that \(R_t = K/\rho + \sigma = \kappa/\rho + \sigma + 1\), which is usually called the “effective reproduction number”. It depends on time through the time dependence of \(\kappa\) and \(\sigma\).

The parameters \(K\), \(\rho\) and \(\sigma\) in Eq. [5] are explained as follows:

\(K\): the infection rate, which can be written as \(K = \lambda f^2\) with \(\lambda\) the infection rate at normal “individual contact rate". If we reduce the contact rate by a factor \(f < 1\) the rate becomes \(K = \lambda f^2\). The contact rate of general public can be reduced only by lockdown of the entire city or by an official order to citizens to stay home and/or wear masks and maintain social distancing. We may call \(f^2\) the “herd contact rate”.

Another way to explain this is the following:

The “herd contact rate" is proportional to the square of the entire population \(N\) when \(N\) is large (contact number is proportional to \(N(N-1)/2 \propto N^2\) if \(N\) is large). We regard \(f \propto N - N_{\text{join}}\), where, when the lockdown is started, \(N_{\text{join}}\) denotes the number of individuals locked down. In our model, the entire population is not locked down right away.

In later analysis we use the following fact: the joining population \(N_{\text{join}}\) is a function of time \(x\) and analytically it can be written as,

\[ N_{\text{join}} = ax + bx^2 + cx^3 + \cdots. \]

As the first approximation we take only the first
term, which is correct for small $x$, but may not be correct when $x$ gets larger. Interestingly, the fit to the daily infection by an exponential of a polynomial will be found later to be sufficient if we take a polynomial up to the third order, thus indicating that $\kappa$ can be well fitted by a second order polynomial and justifying the approximation $N_{\text{gain}} = ax$.

$\rho$: Natural recovery and/or cure rate of COVID-19. It is not known accurately so far but we can estimate it as will be shown later.

$\sigma$: Rate of finding the infected among the symptomatic individuals. The infected found must be separated from the public and included in $n_T$.

$\sigma$ can be defined and estimated as follows:

$$\sigma = \frac{\text{number of tested individuals per day} \times \text{positivity rate of the test}}{\text{number of infections}}.$$  \[6\]

In the following analysis we use only the fact that $\sigma$ is proportional to the product of the number of tests and the positivity rate. $\sigma$ can depend on time in general. The effective efficiency of the test is usually called the “positivity rate”.

“Tested individuals” in Eq. [6] refers in practice to individuals tested by PCR, for which the efficiency is discussed in Appendix 6.

To increase the value of $\sigma$ there are at least two distinct approaches, since it is proportional to number of tests and to the positivity rate of the test. For example, China and South Korea adopted the policy of increasing the number of tests while Japan instead opted to increase the positivity rate by tracing the “clusters”. Either way, they were relatively successful in suppressing the initial outbreak of the COVID-19 compared to other countries.

We must aim at

$$\kappa \leq 0 \rightarrow \lambda f^2 \leq \rho + \sigma$$

$$\rightarrow \text{reproduction number } R_t = \frac{\lambda f^2}{\rho + \sigma} \leq 1. \ [7]$$

Usually, the reproduction rate is expressed as $R_t = \lambda f^2 / \rho$ ($= \beta / \gamma$) in SIR model parameters that appear in Appendix 1. The appearance of $\sigma$ in the denominator of Eq. [7] is the feature of our model that leads to the effectiveness of the “search and find method” in suppressing the spread of virus.

Incidentally, although the parameter $R_t$ is the most commonly used and convenient parameter to discuss the spread of virus, other parameters have been proposed in the literature as described in Appendix 4.

8. Rough estimation of $\rho$: Here we use the data from the cruise ship Diamond Princess (https://www.statista.com/statistics/1099517/japan) that harbored in Yokohama with infected passengers. The total number of those infected, and who remained in Japan and disembarked early March, is 672. Among them, 638 recovered by March 19. This proportion exceeds the conventional estimate of about 80%, but we adopt the value above. Then we obtain $638/20 \approx 32$ as the number of those cured per day. Therefore, we get,

$$\rho = 32/672 = 0.05. \ [8]$$

To see the implication of this value, suppose $\kappa = 0.15$, $\sigma = 0$ and $f = 1$ in the formula,

$$\kappa = \lambda f^2 - \rho - \sigma. \ [5]$$

Then we get

$$\lambda = \kappa + \rho = 0.2. \ [9]$$

Next, we calculate how much smaller we should make the individual contact rate $f$ to make $\kappa \leq 0$, under the assumption of $\sigma = 0$. Then,

$$\kappa = 0.2 f^2 - 0.05 = 0, \ [10]$$

$$f = \sqrt{\frac{0.05}{0.2}} = 0.5. \ [11]$$

Many governments around the world are trying to achieve this individual contact rate by locking down the cities and/or ordering their citizens to stay home. This is rather difficult although not impossible. Moreover, $\kappa \geq -0.05$ in the case of $\sigma = 0$, and it takes a long time (months) to reduce the infection. Therefore, it is important to have a non-zero $\sigma$.

9. From the above definition of $\sigma$ we can answer the question “how are $U$ and $dn/dx$ related?”, namely,

$$\frac{dn}{dx} = \sigma U = \sigma U_0 \exp \left[ \int_0^t \kappa(t) \, dt \right]. \ [12]$$

When $\sigma$ is constant i.e., independent of time, the daily new infection $dn/dx$ depends on $\sigma$ as $\sigma e^{-\sigma x}$ indicating that $dn/dx$ increases initially when $\sigma$ (test number or test positivity rate) is increased due to the factor $\sigma$, but eventually it decreases due to the factor $e^{-\sigma x}$.

In passing we note that the analyses presented in the following sections do not distinguish between symptomatic and asymptomatic infections, which may not always be appropriate. In Appendix 5, we show an analysis that accounts for this distinction.

We are now ready to analyze the situation of COVID-19 in several geographical locations based on our model described in this section. We first analyze three large cities and then several countries.
3. Analysis of Tokyo data

From March 23 to April 23 the numbers of new infections per day (https://stopcovid19.metro.tokyo.lg.jp/en/) are:

\{16, 17, 41, 47, 40, 63, 68, 13, 78, 66, 97, 89, 116, 143, 83, 79, 144, 178, 187, 197, 166, 91, 161, 126, 148, 201, 181, 107, 102, 123, 132, 134\} (Fig. 3). [13]

We analyze all new infections per day without distinguishing among infections with or without identified sources. The same will be true for other data we analyze (Los Angeles and New York) later in this paper. The unique Japanese effort to trace the source of infection as a cluster is captured as an increase in the efficiency of the test in Eqs. [5] and [6], thus increasing the value of \( \sigma \).

First, we point out that \( \sigma \) is proportional to the number of tests per day as its defining formula Eq. [6] shows. For Tokyo, the number of tested individuals per day at the time of this initial analysis is approximately 300, independent of time. We see later that Los Angeles has an increasing number of tested individuals per day and New York has a constant value of 8900 per day. The value of \( \sigma \) is quite interesting when compared to New York data, which we analyze later in this note.

We fit this Tokyo time series using the Mathematica least squares method, assuming the form of an exponential of a third order polynomial:

\[
\kappa = 0.196892 - 0.00885x + 2.003 \times 10^{-5}x^2. \tag{15}
\]

This must be equated with,

\[
\kappa = \lambda(1 - ax)^2 - 0.05 - \sigma. \tag{16}
\]

This gives the following values of the parameters in our model:

\[
\begin{align*}
        a &= 0.0045, \tag{17a} \\
      \lambda &= 0.98, \tag{17b} \\
        \sigma &= 0.73. \tag{17c}
\end{align*}
\]

If we incorporate the time delay between infection and presentation of symptoms, which is estimated to be 3 to 5 days with the highest activity 8 days after the infection, we must consider the time shift between \( dn/dx \) and \( U \). Elaborate discussion of the time gap between infection and appearance of symptoms is given in Appendix 2. We ignore this effect for the time being but it must be included when we discuss more accurate timing.

The flattening of \( U \) occurs somewhat after we reach \( \kappa = 0 \). This corresponds to the maximum of the daily infection number.

\[
\kappa = 0.196892 - 0.00885x + 2.003 \times 10^{-5}x^2 = 0. \tag{18}
\]

Solving this equation, we get \( x = 23.5 \).

The number of infections in Tokyo flattened around April 17. At that time,

\[
f^2 = 0.8, \tag{19}
\]

which is much larger than the government’s goal of \( f^2 = 0.2 \).

We plot here our fit to the number of daily infections in Tokyo in Fig. 4.

We plot in Fig. 5 the flattening of Tokyo infection number \( n \), which may occur somewhat later.
according to the above model. Flattening does not require strengthening of social distancing.

Incidentally, for $\kappa = 0$:

$$\lambda (1 - ax)^2 - 0.05 - \sigma = 0.$$  \[20\]

This is where $dU/dx = 0$ with $U$ the total number of unidentified infections.

At this point we have

$$R_t = \frac{\lambda f^2}{\gamma + \sigma} = \frac{\lambda (1 - ax)^2}{\gamma + \sigma} = 1.$$  \[21\]

$R_t$ is the standard SIR model notation and the appearance of $\sigma$ in the denominator is a particularly noteworthy feature of our model.

Have we reached the government goal of $f^2 = 0.2$ by May 2? The answer is the following:

Assume that $f$ had reached the value of 0.7 already on March 23, a 30% decrease in the contact rate $f$. Then $f^2$ was 0.49 at the date.

$$(1 - ax)^2 = (1 - 0.0045 \times 40)^2 = (0.82)^2 \approx 0.67.$$  \[22\]

Therefore, we have,

$$f^2$$ on May 2 is $0.49 \times 0.67$

$$= 0.33 - \text{not too far from } 0.2,$$  \[23\]

$$R_t = \frac{\lambda f^2}{\gamma + \sigma} = \frac{0.98 \times 0.33}{0.05 + 0.73} = 0.41.$$  \[24\]

If there was no decrease of $f^2$ on March 23, namely $f^2 = 1$ on that day rather than 0.49, we get,

$$R_t = \frac{\lambda f^2}{\gamma + \sigma} = \frac{0.98 \times 0.67}{0.05 + 0.73} = 0.84.$$  \[25\]

This is, of course, consistent with the previous calculation that $R_t$ reached 1 on the 23rd day after March 23.

We now plot the following function in Fig. 6,

$$dn/dx = \exp\left[2.75961 + 0.196892x - 0.0044258x^2 + 6.67625 \times 10^{-6}x^3\right],$$  \[26\]

all the way to May 10 to see if our parametrization works well.

$x = 0$ corresponds to March 24 and $x = 46$ to May 10. The value of May 10 announced by the local Tokyo Government is a bit above 20, which agrees remarkably well with this diagram and confirms the applicability of our parametrization.

We solve

$$\log [dn/dt] = 2.75961 + 0.196892x - 0.0044258x^2 + 6.67625 \times 10^{-6}x^3 = 0.$$  \[27\]

The answer is,

$$x = 60.3 \rightarrow 24$$ of May.  \[28\]

After this date Tokyo will have zero new infections, in principle.

3.A. Analysis of later data of Tokyo. In this subsection we analyze the Tokyo data from May 24 to June 26 with special attention to the relaxing of social distancing on May 25.

The daily new infection number from May 24 to June 26 is (https://stopcovid19.metro.tokyo.lg.jp/en/):

$$\{14, 8, 10, 11, 15, 21, 14, 5, 13, 34, 12, 34, 12, 28, 20, 26, 14, 13, 12, 18, 22, 25, 47, 48, 27, 16, 41, 35, 39, 35, 29, 31, 55, 48, 54\}.$$  \[29\]

This is depicted in Fig. 7.

There seems to be no way to fit this data unless we take into account the effect of asymptomatic infections, because if we only consider infections with symptoms it would appear that new daily infections have died away towards the end of May.

The value of $\lambda$ on May 25 is calculated to be,
Here we used the value $\lambda = 0.18$ which takes into account the time delay between infection and the onset of symptoms (see Table 1.c. at the end of section 5.A. and Appendix 2). Therefore, we fit the data with,

$$\kappa = \lambda_25(1 + a_{25}x)^2 - 0.05.$$  \[31\]

Note that there is no contribution from $\sigma$ in this formula, as shown in Appendix 5, case (2) (natural recovery is the same for infections with or without symptoms). Also, the sign of $a_{25}$ must be positive to account for relaxation of social distancing regulations on May 25.

This gives,

$$dn/dx = 6 \exp[0.045x + 0.023ax^2 + 0.015a^2x^3].$$  \[32\]

We plot this function in Fig. 8 for $a = 0.02$. The fit to the histogram is shown in Fig. 9.

The analysis above indicates that there exists a realistic danger of breakout of COVID-19 spread in Tokyo. The only way to suppress the breakout seems to be to apply the search and find method for the asymptomatic infections, if one wants to avoid another lockdown.

The pooling method for PCR swabs may be suitable if the number of tests becomes large: combine swabs from ten individuals, for example, and test them as one swab. Only when the test is positive individual swabs should be put under PCR test.

### 4. Analysis of Los Angeles data

Next, we analyze the data from Los Angeles (https://corona-virus.la/data).

The data shows that the daily new infection numbers from March 27 to April 13 are the following:

\begin{align*}
257, 344, 332, 342, 548, 513, 534, 521, 711, 663, \\
420, 550, 620, 429, 475, 456, 323, 239.
\end{align*}

We plot this data in Fig. 10. We fit this by a function $\exp[p(x)]$ where $p(x)$ is a third order polynomial. The result is,

$$p(x) = 5.41365 + 0.17285x - 0.00541924x^2$$

$$- 0.000205763x^3.$$  \[34\]

We plot this fit in Fig. 11. We regard this fit as $dn/dx = \sigma \exp[\int_0^t \kappa(t)dt]$ where we parametrize $\kappa$ as follows:

$$\kappa = K - \rho - \sigma,$$  \[5\]

$$K = \lambda(1 - ax)^2, \quad \rho = 0.05.$$  \[35\]

As for $K$, this parametrization shows that the contact rate decreases linearly as $1 - ax$. $\lambda$ is the infection rate on March 27 and $(1 - ax)^2$ describes the rate of decrease of this number during this period. $\rho$ is a universal number and we use the value 0.05 obtained from the data of cruise ship Diamond Princess.
For \( \sigma \) we calculate it using the data published in the above report:

The data (https://corona-virus.la/data) shows that the accumulated test numbers during this period are given as:

\[
\text{test number} = \{6741, 10028, 14033, 16646, 21372, 21746, 24685, 28931, 33082, 37430, 41189, 42215\}. 
\]  

We fit this as a polynomial up to the third order and the result is:

\[
3958.73 + 3131.95x - 3.30069x^2 + 1.38928x^3. 
\]  

By differentiating with respect to \( x \), we get the daily test number:

\[
Y = 3131.95 - 6.62x + 4.2x^2. 
\]  

To calculate \( \sigma \) we use the formula written in equation:

\[
\sigma = \text{(number of tests)} \times \text{(positivity rate)} / \text{(those who have symptoms)}. 
\]  

Nevertheless, since we do not know for sure the positivity rate of the tests nor the number of those who have symptoms in LA we simply use,

\[
\sigma = \frac{\alpha}{3132} (3131.95 - 6.62x + 4.2x^2) \\
= \alpha(1 - 0.0021135x + 0.00134x^2), 
\]  

with \( \alpha \) a constant to be decided later.

This gives,

\[
\kappa = \lambda(1 - ax)^2 - 0.05 \\
- \alpha(1 - 0.0021135x + 0.00134x^2). 
\]  

We fit (using Mathematica) the histogram in Fig. 10 as shown in Fig. 11 by,

\[
A \exp \left[ \int_0^x dx \left\{ \lambda(1 - ax)^2 - 0.05 \\
- \alpha(1 - 0.0021135x + 0.00134x^2) \right\} \right]. 
\]  

A here is equal to \( \sigma \) but we use the average value \( A = \langle \sigma \rangle \) for simplicity and leave it just as a constant.

We then have,

\[
\kappa = dp/dx = 0.17285 - 0.0108x - 0.000618x^2 \\
= \lambda(1 - ax)^2 - 0.05 \\
- \alpha(1 - 0.0021135x + 0.00134x^2). 
\]  

Solving this equation, we get,

\[
\alpha = 0.483, \\
\lambda = 0.70428, \\
a = 0.0084. 
\]  

This result reveals various important facts. We point out, as the most important fact, that the policy of so-called social distancing by lockdown of a city does not seem to be efficient enough without increasing the number of tests in the above LA data analysis.

4.A. Analysis of Los Angeles data from March 27 to May 12. The data is given (https://corona-virus.la/data) as:

\[
dn/dx = \{257, 344, 332, 342, 548, 513, 534, 521, 711, 663, 420, 550, 620, 429, 475, 456, 323, 239, 670, 472, 399, 567, 642, 334, 1491, 1400, 1318, 1081, 1035, 607, 440, 900, 597, 1541, 733, 1065, 691, 781, 568, 1638, 851, 815, 883, 1011, 454, 591, 961\}. 
\]  

We plot this data in Fig. 12.
We can fit this by the following function,
\[ \frac{dn}{dx} = \exp[5.94616 + 0.002852x + 0.001579x^2 + 0.0000294x^3] \]
This function is shown in Fig. 13 and the fitting is shown in Fig. 14.

We solve,
\[ \lambda(1 - ax)^2 - 0.05 - \alpha(1 - 0.0021135x + 0.00134x^2) = 0.00285 + 0.00316x - 0.0000883x^2, \]
and obtain,
\[ \lambda = 0.131, \quad a = -0.0114, \quad \alpha = 0.07862. \]
The value of \( a \) is negative. This means that herd contact rate did not decrease but rather increased during this period. The only way that Los Angeles is sustaining the present situation is by increasing the number of tested individuals rather intensively.

5. Analysis of New York data

Next, we analyze the data from New York State. From March 23 to April 15 the daily number of infection (https://www1.nyc.gov/site/doh/covid-19-data.page) is:

\[ \begin{align*}
3477, 4365, 4705, 4872, 4953, 3350, 3424, 5973, 5136, 5706, 5229, 4647, 3979, 3113, 6118, 5706, 5229, 4647, 3979, 3113, 148519, 162464, 187166, 178917, 184122, 191868, 197273, 210610, 219737, 229040, 237474, 243074, 251840.
\end{align*} \]

We show it in Fig. 15, which is rather similar to the Los Angeles data except for the scale.

Another set of data is available on the accumulated number of test frequency from March 23 to April 21:

\[ \begin{align*}
51406, 63752, 70857, 78936, 83000, 87414, 97409, 104096, 110606, 127788, 135861, 148519, 162464, 187166, 178917, 184122, 191868, 197273, 210610, 219737, 229040, 237474, 243074, 251840.
\end{align*} \]

We plot this in Fig. 16. This looks linear and we fit it as follows:
This means that the daily test number is 8944.52.

We fit the daily infection number depicted in Fig. 15. The contribution of the test to \( \kappa \) is just a constant times 8944.52 which is also a constant. We designate it as \( \sigma \). The infection rate \( \lambda f^2 \) is put to be \( \lambda(1 - ax)^2 \) as in the case of LA but we assume \( (ax)^2 \ll ax \) which will be justified later.

Then \( \kappa = \lambda(1 - 2ax) - 0.005 - \sigma \) and the daily number is \( \sigma \exp \int_0^x \kappa dx \).

Fitting was done using the Mathematica least squares method and we get,

\[
\kappa = 0.0611162 - 0.00566x. \tag{52}
\]

This fit is shown in Fig. 17. We get,

\[
\kappa = \lambda(1 - 2ax) - 0.005 - \sigma = 0.0611162 - 0.00566x. \tag{53}
\]

Adopting the value of \( \lambda \) obtained in the analysis of Los Angeles data:

\[
\lambda = 0.704. \tag{54}
\]

We get,

\[
\sigma = 0.594, \quad a = 0.0040. \tag{55}
\]

We see that testing represented as \( \sigma \) here is crucial as in the case of Los Angeles: \( \sigma \) is as large as \( \lambda \). We also see that \( (ax)^2 \ll ax \) since \( 0 \leq x \leq 20 \).

The reason we ignored the \( (ax)^2 \) is that fitting the Fig. 15 data by a third order polynomial yields a negative coefficient for the third order term \( x^3 \), which is impossible unless we have a \( x^2 \) term in \( \gamma \). The latter can be achieved by including the increasing medical effort and/or increasing death rate. Both of these contribute to reducing the daily detected infection number \( dn/dx \) because the coefficient of \( x^2 \) in \( \lambda(1 - ax)^2 \) is positive thus contributing a positive number to \( x^3 \) term to the daily infection number. We simplified the process by ignoring the \( x^2 \) term in the fitting. The consequence is that we have one free parameter in our fit. We fixed it by adopting the Los Angeles value for \( \lambda \), which could be more or less common within a country.

5A. Analysis of New York data from March 23 to May 11. The new infections per day (https://www1.nyc.gov/site/doh/covid/covid-19-data.page) is,

\[
\{3477, 4365, 4705, 4872, 4953, 3350, 3424, 5973, 5136, 4883, 5533, 5396, 3666, 3565, 6118, 5706, 5229, 4647, 3979, 3113, 3526, 3555, 2117, 2351, 3768, 3059, 3457, 2832, 2465, 1589, 1010, 2304, 2712, 2334, 2003, 1863, 1047, 776, 1532, 1503, 1382, 927, 1040, 612, 401, 586\}. \tag{56}
\]

This is shown in Fig. 18.
We fit this by our original function which was used to analyze the data from March 23 to April 17:

\[ dn/dx = \exp[8.1789 + 0.0611162x - 0.00283x^2]. \]  

This function is depicted in Fig. 19 and the fitting in Fig. 20.

The fit is qualitatively good but it seems it is substantially lower than the actual value after mid-April. Therefore, we try to fit the whole data set by a new function of exponential of a certain third order polynomial:

The result is,

\[ dn/dx = \exp[8.2887 + 0.0347x - 0.0019x^2 + (5.8924)^{-6}x^3]. \]  

This function is depicted in Fig. 21 and the fit in Fig. 22.

Solving the equation,

\[ \kappa = \lambda(1 - ax)^2 - 0.05 - \sigma \]
\[ = 0.0347 - 0.00386x + 0.0000177x^2. \]  

We get,

\[ a = 0.00917, \]  
\[ \lambda = 0.210, \]  
\[ \sigma = 0.125. \]  

We compare the situation of three cities in Tables 1a, 1b and 1c. The effect of \( \Delta t \) is remarkable. The value of \( \lambda \) (infection rate) is the smallest in Tokyo and larger in LA and in NY. The value of \( \sigma \) is comparable between Tokyo and New York but Los Angeles has a much larger value when taking into account its time dependence.
The most significant difference between Tokyo and the above two U.S. cities is the scale of infection. In particular, NY has 42 times more infections than Tokyo. To understand it we note that,  

1. Spreading does exponentially grow. This means that the initial response to the virus is very important. Tiny difference of the parameters on top of the exponential function makes a large difference at later time as NY—Tokyo difference indicates. Japan adopted the “cluster model” and traced the infection.  

2. Remembering that  \( \sigma = \frac{\text{number of tested individuals per day} \times \text{positivity rate of the test}}{\text{number of infections}} \) two countries, South Korea and Japan were successful initially in making  \( \sigma \) larger compared to other countries for different reasons and methods: SK dramatically increased the number of tested individuals. On the other hand, Japan increased the positivity rate of the test by tracing the sources to “clusters” although it does not work that well after large number of infections became untraceable.  

6. Analysis of Japan data (All Japan data from March 26 to May 1)  

The data is given (https://www.worldometers.info/coronavirus/country/japan/) as follows:
We now turn to the predictions we can make based on the above analysis:

(1) Date of \( f_2 \neq 0 \).

\[
0.207 - 0.012x + 0.00010x^2 = 0,
\]
\[
x = 20 \rightarrow \text{April 15.}
\]

(2) Behavior of \( R_t \).

\[
R_t = \frac{0.384(1 - 0.016x)^2}{0.05 + 0.127} = 2.16(1 - 0.016x)^2.
\]

This function is shown in Fig. 26.

(3) The current (May 1) value of \( f_2 \).

\[
f_2(\text{May 1}) = (1 - 0.016 \times 36)^2 = 0.180.
\]

We now turn to the predictions we can make based on the above analysis:

(1) Date of \( \kappa = 0 \).

\[
x = 20 \rightarrow \text{April 15.}
\]

(2) Behavior of \( R_t \).

\[
R_t = \frac{\lambda f^2}{\gamma + \sigma} = \frac{0.384(1 - 0.016x)^2}{0.05 + 0.127} = 2.16(1 - 0.016x)^2.
\]

This function is shown in Fig. 26.

(3) The current (May 1) value of \( f_2 \).

\[
f_2(\text{May 1}) = (1 - 0.016 \times 36)^2 = 0.180.
\]

We plot the total number of infections in Japan starting from March 26 for 50 days in Fig. 27.

6.A. What is the solution towards early recovery? For this purpose, we consider two cases:

(1) \( f^2 \) reaches 0.4 and stays there afterwards. We first calculate on which day this was realized in our model:

\[
(1 - 0.016x)^2 = 0.4.
\]

This gives \( x = 23 \), and this is 23 days after March 26—April 18.

\[
\begin{align*}
\kappa &= \lambda f^2 - \gamma - \sigma \\
&= 0.384 \times 0.40 - 0.05 - 0.127 = -0.023. \quad [69]
\end{align*}
\]

On this day the value of \( dn/dt \) is given by,

\[
501 = \exp[4.39 + 0.207 \times 23 - 0.0062 \times 23^2 + 0.000034 \times 23^3]. \quad [70]
\]

We plot \( 501 \exp[-0.023x] \) after \( x = 23 \) in Fig. 28. When can we reach \( dn/dx < 1 \), meaning no new infection?

\[
501 \exp[-0.023x] = 1. \quad [71]
\]

This gives \( x = 270.287 \) and means 9 months after March 26.
(2) $f^2$ reaches the value of 0.2 and stays there afterwards. This happens on the day given by,

$$(1 - 0.016x)^2 = 0.2.$$ \[72\]

This gives $x = 35$ after March 26 and the date is April 30.

$$\kappa = \lambda f^2 - \gamma - \sigma = 0.384 \times 0.2 - 0.05 - 0.127 = -0.10.$$ \[73\]

On this day the value of $dn/dx$ is given by,

$$dn/dx = \exp[4.29 + 0.2072 \times 35 - 0.00616 \times 35^2 + 0.000338 \times 35^3] = 233.$$ \[74\]

We plot the function $233 \exp[-0.1x]$ in Fig. 29.

We find the day when $dn/dx$ goes below 1,

$$233 \exp[-0.1x] = 1,$$ \[75\]

$$x = 54.5.$$ \[76\]

This is 54 days after April 30, namely, June 23.

6.B. The effect of test number increase. We now double the value of $\sigma$ from 0.127 to 0.254 and see how it changes the situation.

Case (1) $f^2 = 0.4$.

$$\kappa = \lambda f^2 - \gamma - \sigma = -0.023 - 0.127 = -0.150.$$ \[77\]

We plot $501 \exp[-0.150x]$ in Fig. 30.

We get the date of $dn/dx = 1$,

$$501 \exp[-0.150x] = 1.$$ \[78\]

This gives $x = 41.4$. This is 41 days after April 30, namely, June 11.

Case (2) $f^2 = 0.2$.

$$\kappa = \lambda f^2 - \gamma - \sigma = -0.150 - 0.127 = -0.277.$$ \[79\]

We plot $dn/dx = 233 \exp[-0.277x]$ in Fig. 31.

The date of $dn/dx = 1$ is given by,

$$233 \exp[-0.277x] = 1.$$ \[80\]

This gives $x = 19.7$.

We have $dn/dx = 1$ on 20 days after April 30, that is, May 20, showing that it is possible to terminate the social distancing policy before the end of May. Maybe we should stress that this is the only case among the four cases considered above in which it is safely possible to terminate the Japanese governmental emergency status at the end of May. Collaboration of social distancing and the increase of PCR test numbers is essential. A premature relaxation of the social distancing will lead to further spreading of the virus. The actual date depends on how large the time gap between the infection and the appearance of symptoms as will be discussed in Appendix 2 and mentioned in section 4.

The following two diagrams of Fig. 32 and Fig. 33 illustrate the effect of doubling of the value of $\sigma$ to the daily new infections of Tokyo. To be more precise about the dates we must subtract the incubation period of average 7 days in each case.
We summarize the above result with this subtraction included in the Table 2. Table 2 shows that safely lifting of emergency status at the end of May is not possible irrespective of whether $f^2$ has reached the value of 0.4 or 0.2 although, if we adopt more moderate criterion than $dn/dt F 1$, we may lift the emergency at the end of May if $f^2 0.2$. Whichever value of $f^2$, we can escape from the emergency situation towards the end of May or early June in case of $f^2 0.4$ provided that we double the value of $\sigma$. In fact, in case of $f^2 0.2$ with doubled $\sigma$, we did not have to extend the emergency status for the month of May, if we had doubled the value of $\sigma$ before this time.

The reader is reminded that to achieve the doubling of $\sigma$ requires doubling of:

$$\sigma \propto \text{number of test cases times positivity rate of the test}.$$ 

6.C. Another way to find a path to recovery in Japan. We extend our polynomial fit to early May. $x 60$ corresponds to May 5 in Fig. 34. Then we solve,

$$4.298 + 0.207x - 0.00616x^2 + 0.000033x^3 = 0. \quad [81]$$

to find when the daily new infection becomes zero.

We find,

$$x = 71.29 \rightarrow 5 \text{ of June}.$$ 

How can we make this date arrive earlier?

We assume we can increase the value of $\sigma$ on the date $x \rightarrow x + \Delta t$ as follows:

We take into account the time lag between the infection and the appearance of the symptom in this analysis. We have from Appendix 2,

$$\frac{\sigma_{\Delta t}}{1 - \sigma_{\Delta t}} = \sigma \quad \text{or} \quad \sigma_{\Delta t} = \frac{\sigma}{1 + \Delta t \sigma}. \quad [82]$$

The current value of $\sigma = 0.127$, and, if we adopt the value of $\Delta t$ to be its maximum value 14, we get,

$$\sigma_{\Delta t} = 0.046. \quad [83]$$

Suppose we increase the value of $\sigma_{\Delta t} = 0.046$ by 20% on the date $x$

$$\sigma(\text{after date } x) = \frac{1.2\sigma_{\Delta t}}{1 - 1.2\sigma_{\Delta t}\Delta t} = 0.243. \quad [84]$$

This shows an interesting fact: By an increase of only 20% in the value of $\sigma_{\Delta t}$, which is proportional to the
product of test number and positivity rate, we can almost double the value of $\sigma$.

After the date $t$ we must add $-(0.243-0.127)x$ to the polynomial,

$$P = 4.298 + 0.207x - 0.00616698x^2$$
$$+ 0.00003387x^3. \quad [85]$$

That is, $P$ becomes,

$$P_T = 4.298 + 0.091x - 0.006167x^2$$
$$+ 0.00003387x^3. \quad [86]$$

Adopting the date $t$ to be May 15 when $dn/dx = 32.3$, we have a new $dn/dx$ after this date,

$$dn/dx = 32.3/0.097 \times \exp[4.2987 + 0.091x - 0.006167x^2 + 0.00003387x^3]. \quad [87]$$

It is depicted in Fig. 35.

We obtain the date when $dn/dx < 1$ by solving:

$$1 = (32.3/0.097) \times \exp[4.2987 + 0.091x - 0.006167x^2 + 0.0000339x^3]. \quad [88]$$

We get,

$$x = 62.41 \rightarrow 27 \text{ of May}. \quad [89]$$

This is an improvement from June 5 and we will be able to exit emergency status towards the end of May.

7. Analysis of U.S. data from March 17 to May 2

We now proceed to analyze the U.S. data.

The daily new infection of U.S. is given (https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html) as:

$\{2797, 3419, 4777, 3528, 5836, 8821, 10934, 10115,$
$$13487, 16916, 17965, 19332, 18251, 22635, 22562,$$
$$27043, 26315, 18819, 9338, 63455, 43438, 34347,$$
$$31534, 31705, 33251, 33288, 29145, 24117, 26385,$$
$$27259, 29164, 29002, 29916, 26008, 29468, 26527,$$
$$25868, 37144, 30226, 33119, 29355, 23459,$$
$$23901, 26512, 30787, 30326, 29671\}. \quad [90]$$

This is depicted in Fig. 36.

It is clear that we can divide the U.S. data into two periods: March 17 to April 4 and April 5 to May 2. Something happened on April 5 which should be investigated.

(1) Analysis of the first period. The initial period (March 17 to April 4) is shown in Fig. 37. This can be fitted by,

$$\exp[7.87537 + 0.0891674x + 0.0198095x^2$$
$$- 0.00106288x^3], \quad [91]$$

leading to the following $\kappa$,

$$\kappa = \lambda(1-ax)^2 - 0.05 - \sigma(x)$$
$$= 0.0891674 + 0.04x - 0.0032x^2. \quad [92]$$

We note that $\sigma(x)$ must contain $x^2$ term to make $x^2$ coefficient negative.
We put,
\[ \sigma(x) = s + ux + vx^2. \]  \[93\]

Then we get,
\[ \lambda - 0.05 - s = 0.089, \]  \[94a\]
\[ -2a\lambda - u = 0.04, \]  \[94b\]
\[ \lambda a^2 - v = -0.0032. \]  \[94c\]

There is no way to determine all the parameters unless the information on the test numbers is provided, as in the case of New York where we used the Los Angeles value of \( \lambda \).

In this case, we may try to fit with the exponential of a second order polynomial, with the result,
\[ \frac{dn}{dx} = \exp[7.38432 + 0.350848x - 0.0120767x^2], \]  \[95\]
leading to,
\[ \kappa = \lambda(1 - 2ax) - 0.05 - \sigma = 0.351 - 0.036x, \]  \[96\]
\[ 2a\lambda = 0.036, \]  \[97\]
\[ \lambda - \sigma = 0.401. \]  \[98\]

We still need information on test numbers.

To estimate roughly these parameters, we use the Los Angeles value for \( \lambda \) assuming this parameter should be rather universal:
\[ \lambda = 0.704. \]  \[99\]

Then we get,
\[ a = 0.026, \]  \[100\]
\[ \sigma = 0.303. \]  \[101\]

(2) Analysis of the second period. The daily infection number is shown in Fig. 38. If we ignore bins of the first two days, April 5 and April 6 when something must have happened (actually several days before taking into account the time delay) the bins are more or less flat with the value around 30000 infections per day.

This implies
\[ \frac{d^2n}{dx^2} = \sigma \frac{dn}{dx} = \kappa \sigma = 0, \]  \[102\]
\[ \kappa = \lambda f^2 - \gamma - \sigma = 0, \]  \[103\]
\[ R_t = \frac{\lambda f^2}{\gamma + \sigma} = 1. \]  \[104\]

This indicates the following:

(1) U.S. data shows that the \( f^2 \) reached the non-zero minimum value and is staying with that value in the whole second period.

(2) Since there is no way to decrease \( f^2 \) any further we can think of two ways to decrease \( R_t \):

(A) Increase the value of \( \sigma \), which requires the increase of tested individuals without decreasing the positivity rate (random testing, for example, would decrease it).

(B) Increase the value of \( \gamma \) by medical method most probably by adopting some efficient drug.

In each case or combination of these two methods we must not increase \( f^2 \).

In order to roughly estimate the value of \( f^2 \) we use the values obtained in analyzing the first period:
\[ \lambda = 0.704, \quad \gamma = 0.05, \quad \sigma = 0.303. \]  \[105\]

Then we have,
\[ \kappa = \lambda f^2 - \gamma - \sigma = 0.704f^2 - 0.05 - 0.304 = 0. \]  \[106\]
This gives \( f^2 = 0.44 \).

7.A. Further analysis of U.S. data (March 17 to May 28). Towards the end of May, U.S. data shows slight decline in the number of daily new infections. We analyze this situation:

The data is given (https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html) as:

\{2797, 3419, 4777, 3528, 5836, 8821, 10934, 10115, 13487, 16916, 17965, 19332, 18251, 22635, 22562, 27043, 26315, 18819, 9338, 63455, 43438, 43447, 31534, 31705, 33251, 33288, 29145, 24157, 26385, 27259, 29164, 29002, 29916, 26008, 29468, 26527, 25868, 37144, 30226, 33119, 29355, 23459, 23901, 26512, 30787, 30326, 29671, 29763, 19138, 22303, 23366, 30861, 25996, 26660, 23792, 18106, 21467, 20869, 27191, 22977, 31967, 13284, 24481, 23405, 20869, 27191, 22977, 31967, 13284, 24481, 23405, 20869, 27191, 22977, 31967, 13284, 24481, 23405, 20869, 27191, 22977, 31967, 13284, 24481, 23405\}

![Fig. 38. New infections per day in the second period.](image-url)
We fit this by an exponential of the third order polynomial,
\[ p = 7.984 + 0.1876x - 0.004451x^2 + 0.0009031x^3, \]
as shown in Fig. 40.

From this we obtain,
\[ \kappa = \lambda(1 - ax)^2 - 0.05 - \sigma \]
\[ = 0.188 - 0.009x + 0.000096x^2. \]

This gives,
\[ a = 0.021, \]
\[ \lambda = 0.214, \]
\[ \sigma = -0.024, \]
with,
\[ R_t = 1 + \frac{\kappa}{\sigma + \rho} \]
\[ = 1 + (0.188 - 0.009x + 0.000096x^2) \times 38.5. \]

Figure 40 shows that our fitting curve already indicates that there is a dangerous symptom of outbreak although the actual histogram does not show it.

8. Case of Sweden from April 1 to April 30

Among the major European countries Italy, France and Germany are moving towards the recovery. But, as far as the data shows, U.K. and Sweden are still struggling with the constant or even increase of the new infections. U.K. is the worst and it is very similar to the situation in U.S. Sweden started with controlled initial stage and the total number did not increase as much as other countries but the current situation does not seem to be going well.

Daily new infection number from April 1 to April 30 is given (https://www.worldometers.info/coronavirus/country/sweden/) by,
\[ \{512, 519, 612, 365, 387, 376, 487, 726, 722, 544, 466, \]
\[ 332, 465, 497, 482, 613, 676, 606, 563, 392, 545, \]
\[ 682, 751, 812, 610, 463, 286, 695, 681, 790\}. \]

This is shown in Fig. 41.

We fit this by a function \(\exp[\text{linear function of date}]\). The result is,
\[ \frac{dn}{dx} = \exp[6.13547 + 0.0096879x]. \]

The fitting of the data with this function is shown in Fig. 42. This leads to,
\[ \kappa = \lambda - 0.05 - \sigma = 0.0097. \]

This is a very small positive number and unless we increase the value of \(\sigma\), the number of new infections will stay or increase.

We understand that Sweden adopted the policy, at least partially, taking into account what is called...
“herd immunity model” which we discuss in Appendix 3. The above analysis implies that its success is not confirmed. Rather, its difficulty is implied in some serology test (antibody test) performed in many parts of the world.

9. African COVID-19 situation from March 31 to April 28

Data is available in the form of histogram from WHO (https://covid19.who.int). The rough numbers are taken from the histogram every 2 days apart starting from March 31 to April 28:

{March 31, April 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28} → {300, 330, 400, 400, 450, 570, 350, 600, 650, 760, 800, 1000, 1000, 800, 780}.

We plot this in Fig. 43.

This is fitted by the following exponential of a polynomial with the help of Mathematica:

\[
\frac{dn}{dx} = \exp(5.542201 + 0.1318x - 0.003069x^2).
\]

Combining the fitting curve with Fig. 43 histogram we get Fig. 44. In this case we have,

\[
\kappa = \lambda(1 - 2ax)^2 - 0.05 - \sigma
\]
\[
= 0.132 - 0.00614x,
\]
\[
2a\lambda = 0.006,
\]
\[
\lambda - 0.05 - \sigma = 0.132.
\]

Again, using the Los Angeles value of \( \lambda = 0.704 \), we get,

\[
a = 0.004, \quad \sigma = 0.522.
\]

10. Brazilian situation of COVID-19 from March 27 to May 5

Among the Latin American countries Brazil has the largest case of infections followed by Peru and Mexico. Brazil is still in the state of sharp increase of new infections per day as the following data (https://www.worldometers.info/coronavirus/country/brazil/) shows:

{March 27, 30, April 2, 5, 8, 11, 14, 17, 20, 23, 26, 29, May 2, 5} → {502, 323, 1076, 852, 133, 1089, 1832, 217, 1927, 3735, 3379, 6276, 4970, 9493}.

We plot this in the histogram in Fig. 45.

This can be fitted by an exponential of a polynomial:

\[
\frac{dn}{dx} = \exp(6.24 - 0.0775x + 0.0213x^2).
\]

We show it in Fig. 46.

If we adopt the Los Angeles value of \( \lambda = 0.704 \), we get,
Note that the sign of $a$ is negative meaning that the contact rate is increasing. How we can avoid a disaster in Brazil seems to be most urgent.

11. Situation of COVID-19 in China

In the Eurasian continent Russia and India are as dangerous as Brazil in the data of new infections. China seems to have controlled the situation which is unique and we analyze the Chinese situation here. The daily new infections are given from January 30 to March 1 (https://www.worldometers.info/coronavirus/country/china) as:

$$\{1500, 1772, 2066, 2275, 2622, 2901, 3067, 3250, 3260, 3281, 3171, 2904, 2015, 14108, 5090, 2641, 2008, 2048, 1888, 1749, 391, 889, 823, 943, 746, 554, 560, 480, 423, 413, 411, 356\}.$$  \[124\]

This is depicted in Fig. 47.

$$a = -0.03,$$  \[123a\]

$$\sigma = 0.576.$$  \[123b\]

We fit this with exp[polynomial] and the result is,

$$dn/dx = \exp\left[6.7945 + 0.3398x - 0.0230x^2 + 0.000359x^3\right].$$  \[125\]

We plot the fitting in Fig. 48.

We can now calculate the parameters of our model:

$$\kappa = \lambda(1 - ax)^2 - 0.05 - \sigma$$

$$= 0.3398 - 0.046x + 0.00108x^2.$$  \[126\]

We get,

$$a = 0.047,$$  \[127a\]

$$\lambda = 0.49,$$  \[127b\]

$$\sigma = 0.101.$$  \[127c\]

12. Policy issues

We now discuss some policy issues involved.

1) Why do some countries have smaller numbers of infections? The number of infections
increases exponentially, and human policy/culture can affect only the argument of this exponential function. This indicates that a small difference in the policy/culture can make a huge difference in the consequences. To explain the situation, we take a simple case of linear exponent approximation:

Assume that total infection \( N \sim e^{\lambda x} \) and \( a \sim 0.8/\) day in one place \( A \) and \( 0.5/ \) day in another place \( B \) at the initial stage. After 20 days we have,

\[
\frac{N_A}{N_B} \sim e^{(0.8-0.5) \times 20} = e^{0} \sim 403.
\]  

[128]

Of course, \( B \) can take much better policy later and the ratio may decrease but this example shows that it is no surprise that some countries like Japan or South Korea, which adopted reasonable "search and find policy" at the beginning, have much less total numbers compared to U.S., for example.

(2) What happens following the reopening after lockdown? We take Tokyo case as an example, where we have,

\[
\begin{align*}
\lambda &= 0.98, &\text{[129a]} \\
a &= 0.004, &\text{[129b]} \\
\sigma &= 0.73. &\text{[129c]}
\end{align*}
\]

We have,

\[
e^{(\lambda-\rho-\sigma)x} = e^{0.2x} \quad \text{and} \quad e^{(\lambda_2-\rho-\sigma_2)x} = e^{0.003x}.
\]  

[130]

This shows that, whether or not we take the time gap into account, a second wave is coming. The same conclusion is obtained for New York and Los Angeles.

Above is the case of total opening but we may have a partial opening with finite \( f \) in \( K = \lambda f^2 \). In the above case, to make

\[
\kappa = 0.98 f^2 - 0.05 - 0.73 = 0,
\]  

[131]

we get,

\[
f = 0.9.
\]  

[132]

This may be achievable by some restrictions on the citizen’s activities at the time of opening.

In case of NY we have,

\[
\kappa = 0.219 f^2 - 0.05 - 0.125 = 0,
\]  

[133]

\[
f = 0.89.
\]  

[134]

The value is almost the same as that of Tokyo.

10% reduction of social (or physical) contact from the normal time may or may not be possible. In case of whole Japan, we have,

\[
\begin{align*}
\lambda &= 0.384, &\sigma &= 0.127, &\text{[135]} \\
\kappa &= 0.384 f^2 - 0.05 - 0.127 = 0. &\text{[136]}
\end{align*}
\]

We get,

\[
f = 0.68.
\]  

[137]

In this case, 32% reduction of social (physical) contact compared to the normal state is required. The obviously better way to avoid this is to increase \( \sigma \) by combining massive antibody tests with antigen tests and/or PCR tests.

This is especially true when non-symptomatic infection is dominating the spread as is shown in case of Tokyo in section 3.A.

13. Conclusion

It is efficient to strengthen the test opportunities without decreasing its positivity rate and isolate those who were tested positive for SARS-CoV-2 rather than isolate everyone by adopting the policy of social distancing, although social distancing is very helpful.

This is especially true because the number of asymptomatic infections is possibly much larger than that with symptom as demonstrated in New York or Los Angeles data and discussed in Appendix 5 of this paper. Unless we perform relatively larger scale test (first antibody test and then antigen or PCR test), we will be forced to lock down a city or a country again and again until efficient vaccine is invented.

As shown in Eq. [7] in section 2, the effect of testing and of recovery rate appears in the denominator of \( R_t \) implying the importance of these two effects. This is probably the most important aspect of our model that is distinct from the conventional interpretation of SIR model.

The analysis shows that U.S. is in trouble unless it can increase the test numbers substantially or some effective medicine is invented. U.K. also needs substantial increase of testing although we do not show the analysis here and Sweden needs it to a lesser degree. The Swedish policy of herd immunity is discussed in Appendix 3. The African continent seems to be responding reasonably well so far but ambiguity remains. Brazil seems to be heading for a disaster unless some decisive measure is taken. In Eurasian continent, Russia and India are still far from controlling the situation although we do not show the analysis in this paper.

Appendix 1. Derivation of our model based on the standard epidemiology model:

SIR model[2,9]

Here I explain how we can reach the above model starting from the “standard model” in the field of mathematical epidemiology. We use standard
notations rather than the our notations used in the above main sections: \( \rho \to \gamma, \ x \to t, \ U \to I, \) and \( n \to I_h. \)

SIR model is defined to be a set of differential equations for \( S: \) susceptible, \( I: \) infected and \( R: \) recovered:

\[
\begin{align*}
\frac{dS}{dt} &= -\beta IS/N, \quad [1.1] \\
\frac{dI}{dt} &= \beta IS/N - \gamma I, \quad [1.2] \\
\frac{dR}{dt} &= \gamma I. \quad [1.3]
\end{align*}
\]

Here \( \beta IS/N \) is the rate of decrease of the susceptible \( S \) and \( \gamma \) is the recovery rate and \( S + I + R = N. \) We assume \( N \) is a constant (total population of a body we want to discuss). We modify the model in the following way:

**1** Approximation. At least in case of COVID-19 we can safely assume that \( S \gg I \) or \( R. \) Therefore, we can safely put,

\[
S/N \simeq 1. \quad [1.4]
\]

Then the above equations become linear,

\[
\begin{align*}
\frac{dS}{dt} &= -\beta I, \quad [1.5] \\
\frac{dI}{dt} &= \beta I - \gamma I, \quad [1.6] \\
\frac{dR}{dt} &= \gamma I. \quad [1.7]
\end{align*}
\]

**2** Modification.

2.1) In SIR model it is usually assumed that \( \beta \) and \( \gamma \) are independent of time. But \( \beta \) depends on the social distancing, for example, and therefore depends on time. \( \gamma \) can be time dependent through increasing medical effort although natural recovery power is independent of time.

We can think of \( \beta \) as follows: it can be written as \( \lambda f^2 \) where \( \lambda \) is the infection rate at certain time and \( f \) is the contact rate of all the members of the population we consider. Both parameters can depend on time. The other contribution comes from the test.

2.2) More important modification is the following: We divide \( I \) into two parts: one is the infected but not recognized so and is at large. We write this as \( I_u. \) The other one is the infected and segregated.

\[
I = I_u + I_h. \quad [1.8]
\]

\( I_h \) is well protected and will not contribute to the first equation. The equations become:

\[
\begin{align*}
\frac{dS}{dt} &= -\beta I_u, \quad [1.9] \\
\frac{dI}{dt} &= \beta I_u - \gamma I, \quad [1.10] \\
\frac{dR}{dt} &= \gamma I. \quad [1.11]
\end{align*}
\]

The second of this set of equations is important and the first and third equations are just the definitions of \( S \) and \( R. \)

\( I_h \) comes out of \( I \) by the PCR test, and assuming test capacity can cover the entire symptomatic population, we have

\[
dI_{h}/dt = \sigma I_u. \quad [1.12]
\]

This is equivalent to Eq. [12] of section 2.

Equation \( dI_{h}/dt = \beta I_u - \gamma I \) becomes,

\[
dI_{u}/dt + dI_{h}/dt = dI_{u}/dt + \sigma I_u = \beta I_u - \gamma(I_u + I_h) = (\beta - \gamma)I_u - \gamma I_h. \quad [1.13]
\]

Substituting Eq. [1.12] to Eq. [1.13], we get,

\[
dI_{u}/dt = (\beta - \sigma - \gamma)I_u - \gamma \int_0^t \sigma I_u dt. \quad [1.14]
\]

We have the integro-differential equation for the \( I_u. \)

When we can ignore the last term, we reach the equation:

\[
dI_{u}/dt = (\beta - \sigma - \gamma)I_u. \quad [1.15]
\]

This is what was used in the above analysis (Eq. [4] with Eq. [5] of section 2). The last term in Eq. [1.14] turned out to be small because of the small value for \( \gamma \) but should be taken into account eventually.

One can include the effect of herd immunity (\( \pm \)something else that prevents infection) by assuming \( S \neq N \) as follows:

\[
S/N = 1 - m_i/N \equiv 1 - \phi, \quad [1.16]
\]

where \( m_i \) is number of those who acquired herd immunity or are protected from infection by some unspecified mechanism.

Then the equation for \( I \) becomes,

\[
dI/dt = \beta I(1 - \phi) - \gamma I. \quad [1.17]
\]

Therefore, we must change our \( \beta \) to,

\[
\beta = \beta(1 - \phi). \quad [1.18]
\]

This leads to,

\[
(\beta - \sigma - \gamma) \to (\beta' - \sigma - \gamma) = (\beta(1 - \phi) - \sigma - \gamma). \quad [1.19]
\]

This simply reduces the effect of \( \beta \) up to 14.9% according to the NY data.

The time dependence of the herd immunity is given by the first equation of the SIR model:

\[
dS/\ dt = -\beta I_u. \quad [1.15]
\]

This leads to,

\[
\frac{N d\phi}{dt} = \beta I_u = \beta I_{u0} \exp \left[ \int_0^t \kappa(t) dt \right]. \quad [1.20]
\]

The last equation comes from our model:
We must be changed to, supposed to be around two weeks.

This shows that the herd immunity increases just like the total infection \( n \).

**Appendix 2. The effect of lag between the infection and onset of symptoms**

(This appendix is written in collaboration with Koichi Yazaki.)

Here we discuss how to incorporate the interval of time subsequent to infection until onset of symptoms. We use \( t \) as the time variable instead of \( x \).

We start with,

\[
\frac{dn}{dt} = \sigma U(t - \Delta t).
\]

[2.1]

This means that the time of testing when a patient presents with symptoms is \( \Delta t \) later than the time when that patient was actually infected. \( \Delta t \) is supposed to be around two weeks.

Then the equation,

\[
dU(t)/dt = \kappa(t) U(t) = (K - \sigma - \gamma) U(t),
\]

[2.2]

must be changed to,

\[
dU(t)/dt = \kappa(t) U(t) = (K - \gamma) U(t) - \sigma U(t - \Delta t)
\]

\[
= \kappa U(t) - \sigma (U(t - \Delta t) - U(t)).
\]

[2.3]

We solve this perturbatively,

\[
U(t) = C(t) \exp \left[ \int_0^t \kappa(t) dt \right].
\]

[2.4]

We get,

\[
(dC/dt) \exp \left[ \int_0^t \kappa(t) dt \right] = \sigma (U(t) - U(t - \Delta t))
\]

\[
= \sigma U_0 \left( \exp \left[ \int_0^t \kappa(t) dt \right] - \exp \left[ \int_{t-\Delta t}^t \kappa(t) dt \right] \right). \quad [2.5]
\]

Then,

\[
C = 1 + \sigma U_0 \int_0^t dt \exp \left( 1 - \exp \left[ - \int_{t-\Delta t}^t \kappa(t) dt \right] \right).
\]

[2.6]

and,

\[
\frac{dn}{dt} = \sigma U(t - \Delta t)
\]

\[
= \sigma \exp \left[ \int_0^{t-\Delta t} \kappa(t) dt \right]
\]

\[
\times \left\{ 1 + \sigma U_0 \int_0^{t-\Delta t} dt \right\}
\]

\[
\times \exp \left( 1 - \exp \left[ - \int_{t-\Delta t}^t \kappa(t) dt \right] \right). \quad [2.7]
\]

To see the behavior of this contribution at large \( t \) \((t \gg \Delta t)\) we approximate this in the following way:

\[
\frac{dn}{dt} = \sigma U(t - \Delta t)
\]

\[
= \sigma \exp \left[ \int_0^{t-\Delta t} \kappa(t) dt \right]
\]

\[
\times \left\{ 1 + \sigma U_0 \Delta t \int_0^{\Delta t} \exp \kappa(t) dt \right\}. \quad [2.8]
\]

Fitting the left hand side by the right hand side can be done with,

\[
\kappa(t) = \lambda (1 - at)^2 - \sigma - \gamma. \quad [2.9]
\]

Following analysis was performed by Koichi Yazaki. We start from Eq. \([2.3]\):

\[
\frac{dU(t)}{dt} = \kappa U(t) - \sigma (U(t - \Delta t) - U(t)). \quad [2.10]
\]

Suppose we expand \( U(t - \Delta t) \) and take only up to the first order in \( \Delta t \), we get,

\[
(1 - \sigma \Delta t) \frac{dU(t)}{dt} = \kappa U(t), \quad [2.11]
\]

Then,

\[
\frac{dU(t)}{dt} = \kappa' U(t), \quad [2.12]
\]

with,

\[
\kappa' \equiv \kappa/(1 - \sigma \Delta t), \quad [2.13]
\]

and,

\[
\frac{dU(t)}{dt} = \sigma \exp \left[ \int_0^{\Delta t} \kappa' dt \right]. \quad [2.14]
\]

On the other hand, from Eq. \([2.1]\), we have,

\[
\frac{dn}{dt} = \sigma U(t - \Delta t) = \sigma \exp \left[ \int_0^{\Delta t} \kappa' dt \right]
\]

\[
\equiv \langle \sigma \rangle \exp [F(t - \Delta t)]. \quad [2.15]
\]

We get,

\[
\kappa'(t - \Delta t) = dF(t - \Delta t)/dt \equiv dP(t)/dt. \quad [2.16]
\]

Then we obtain,

\[
\frac{dP}{dt} = \kappa'(t - \Delta t) = \kappa(t - \Delta t)/(1 - \sigma \Delta t)
\]

\[
= \lambda (1 - at)^2 - \gamma - \sigma = \lambda_{\Delta t}(1 - a_{\Delta t}(t - \Delta t))^2
\]

\[
- \gamma_{\Delta t} - \sigma_{\Delta t}/(1 - \sigma_{\Delta t} \Delta t)
\]

\[
= \frac{\lambda_{\Delta t}(1 + a_{\Delta t} \Delta t)^2}{1 - a_{\Delta t} \Delta t} \left( 1 - \frac{a_{\Delta t}}{1 + a_{\Delta t} \Delta t} \right)^2
\]

\[
= \frac{\lambda_{\Delta t}}{1 - a_{\Delta t} \Delta t} \left( 1 - \frac{a_{\Delta t}}{1 + a_{\Delta t} \Delta t} \right)^2
\]
which leads to,

\[
\lambda_{\Delta t}(1 + a_{\Delta t}\Delta t)^2/(1 - \sigma_{\Delta t}\Delta t) = \lambda, \tag{2.18a}
\]

\[
a_{\Delta t}/(1 + a_{\Delta t}\Delta t) = a, \tag{2.18b}
\]

\[
(\gamma_{\Delta t} + \sigma_{\Delta t})/(1 - \sigma_{\Delta t}\Delta t) = \gamma + \sigma. \tag{2.18c}
\]

Alternatively we have,

\[
a_{\Delta t} = a/(1 - \Delta t a), \tag{2.19a}
\]

\[
\sigma_{\Delta t} = \sigma/(1 + \Delta t \sigma), \tag{2.19b}
\]

\[
\gamma_{\Delta t} = \gamma/(1 + \Delta t \gamma), \tag{2.19c}
\]

\[
\lambda_{\Delta t} = (1 - \Delta t a)^2/\lambda. \tag{2.19d}
\]

We must also shift the date by \(\Delta t\) backward because what happened in \(dn/dt\) happened in \(U \Delta t\) days before.

This correction was pointed out in section 5. We can also continue to use \(\lambda, a, \gamma, \sigma\) rather than \(\lambda_{\Delta t}, a_{\Delta t}, \gamma_{\Delta t}, \sigma_{\Delta t}\) in our formula.

We list the values of these parameters in the form of table created by Koichi Yazaki (Tables 3.a and 3.b).

### Appendix 3. Herd immunity model

Here we discuss a model which is in sharp contrast to our model described above:

**Herd immunity model.** This model asserts that acquiring herd immunity of the entire population is the best way to stop the spreading of the virus. The policy was loosely taken in Sweden based on this model and, as was shown above, it is not a big success at least up to now. If Sweden will continue this policy remains to be seen. Here we describe this model and its hidden or explicit assumptions.

First, we point out that, when we justified our model starting from standard SIR model, we assumed that the susceptible population \(S\) is large compared to \(I\) (infected) or \(R\) (recovered). The herd immunity model asserts that \(S\) can be reduced substantially by the herd immunity in sharp contradiction with our model assumption.

We explain the model based on a recent paper by a Kyoto group\(^{10}\):

The points they are making are summarized as follows:

1. There are two types of SARS-CoV-2 virus S-type and L-type as found by Chinese researchers. S-type is the original type and L-type is the mutated type (in replicase/transcriptase) and the latter is more preferred for translation.

2. Initially the S-type was spread, and a sizable population got the herd immunity against the virus.

3. This sizable population can resist against the powerful L-type when it comes up due to acquired herd immunity.

4. The calculation goes as follows:
   Assume reproduction number \(R_t\) is 2.2 for S-type and 2.17 for L-type. Then we get \(1 - 1/R_t = 0.55\) for L-type and 0.54 for S-type. This means that 54% of the population is infected for S-type and 55% of the population is infected by the L-type. But if the S-type infection comes first and get the herd immunity, they claim, since 0.55 − 0.54 = 0.01 (more precisely 0.006), only 0.6% will be infected by L-type.

5. Initial spread of S-type can be measured by its negative correlation against the yearly

| \(\Delta t\) | \(\lambda\) | \(a\) | \(\sigma\) |
|---|---|---|---|
| 0 | 0.3743 | 0.01648 | 0.1171 |
| 2 | 0.2886 | 0.01704 | 0.08775 |
| 4 | 0.2331 | 0.01764 | 0.06824 |
| 6 | 0.1914 | 0.01828 | 0.05519 |
| 8 | 0.1614 | 0.01898 | 0.04560 |
| 10 | 0.1380 | 0.01973 | 0.03828 |
| 12 | 0.1190 | 0.02054 | 0.03253 |
| 14 | 0.1033 | 0.02142 | 0.02792 |

| \(\Delta t\) | \(\lambda\) | \(a\) | \(\sigma\) |
|---|---|---|---|
| 0 | 0.9781 | 0.004523 | 0.7312 |
| 2 | 0.4123 | 0.004566 | 0.2854 |
| 4 | 0.2665 | 0.004608 | 0.1708 |
| 6 | 0.1996 | 0.004651 | 0.1185 |
| 8 | 0.1612 | 0.004695 | 0.08863 |
| 10 | 0.1362 | 0.004740 | 0.06939 |
| 12 | 0.1187 | 0.004785 | 0.05601 |
| 14 | 0.1057 | 0.004831 | 0.04621 |
influenza spread. Children are the source of S-type spread.

6. It is important to perform antibody test to large population to find out the scale of herd immunity.

Here is the summary of assumptions made in the above analysis.

1. The value of reproduction number \( R_t \) may have been bigger than the current value but the value 2.2 for L-type and 2.17 for S-type, especially the difference of these two numbers, is almost an assumption. The current value of \( R_t \) is already below 1 (0.9) in case of Japan as shown in section 6 consistently with the government report. This can provide us with how the herd immunity ratio behaves as a function of time.

2. Admitting the negative correlation between influenza and the COVID-19, how can this lead to the value of 54% infection of S-type?

As for the item 6, we already have some data from New York and Santa Clara county in California: New York gives 14.9% (https://www.6sqft.com/new-york-covid-antibody-test-preliminary-results/) and Santa Clara (https://www.sccgov.org/sites/covid19/Pages/covid19) gives 2.8 to 4.2%, far below the value 55% adopted in the above paper. I interpret the importance of these numbers not by excluding the non-local (in time variable) index may have its own meaning. Here we compare this index with commonly used indices as follows:

We observe that the dangerous character of COVID-19 is that \( dN/\text{dt} \) seems to behave like \( \exp[\text{polynomial}] \) and the polynomial is higher than the first order. It increases rapidly and decreases rapidly, too.

The indicator must clearly show how this polynomial is changing as a function of time. The above expression integrates over this \( \exp[\text{polynomial}] \) and obscures the contribution of the exponentiated term, although it may reflect some other feature of COVID-19.

Suppose \( dN/\text{dt} = \exp P \) and, instead of the above expression, we use,

\[
\int_0^t \frac{dN}{\text{dt}} \frac{dN}{\text{dt}} dt = \int_0^t \frac{dN}{\text{dt}} \frac{dN}{\text{dt}} dt.
\]

We have positive \( \frac{dN}{\text{dt}} \) (daily new infection). It is obvious that this number is between 0 and 1. This non-local (in time variable) index may have its own meaning. Here we compare this index with commonly used indices as follows:

\[
\kappa = \frac{d^2N}{dN/\text{dt}^2} = dP/\text{dt}.
\]

In our model,

\[
dN/\text{dt} = \sigma U = \sigma \exp \left[ \int_0^t \frac{dP}{\text{dt}} dt \right] = \sigma \exp \left[ \int_0^t \kappa dt \right].
\]

Our model also asserts that,

\[
\kappa = \lambda f^2 - \gamma - \sigma.
\]

Another index widely used is,

\[
R_t = \frac{\lambda f^2}{\gamma + \sigma},
\]

although this expression itself is unique to our model.

Japanese government decided that the measure of exit from the emergency is,

\[
\int_{-7}^d \frac{dn}{\text{dt}} dt \leq 0.5 \times 10^{-5} N.
\]
In our model we have,
\[
\int_{d-7}^{d} \frac{dn}{dt} dt = \int_{d-7}^{d} dt \sigma \left[ \int_{0}^{t} \kappa dt \right] \leq 0.5 \times 10^{-5} N,
\]
where \( N \) is the total population of the target area, Tokyo, New York, Los Angeles, or Japan etc.

This can be written using the \( K_d \) discussed above:
\[
K_d = \frac{\int_{d-7}^{d} \frac{dn}{dt} dt}{\int_{0}^{d} \frac{dn}{dt} dt} \leq 0.5 \times 10^{-5} N = \frac{n(d)/N}{n(d)/N}.
\]

Appendix 5. Asymptomatic infections (This appendix is by Jiro Arafune with some comments of h.s.)

We have unidentified infections (\( U \) in our notation; see for example https://www.fukuishimbun.co.jp/articles/4/1071784, http://www.kansensho.or.jp/uploads/files/topics/2019ncov/covid19.casereport-200409.5.pdf (both in Japanese)). There are two kinds among them: those officially with symptoms and those officially without. The latter can be further divided into two groups. One kind is those who have infection with symptoms too mild to satisfy the official criterion for testing. It depends on the strictness of the criteria for testing. For example, Japan had a very strict criterion such as 4 consecutive days of body temperature higher than 37.5°C. Under such a strict criterion we have large number of individuals belonging to \( U \) who don’t contribute to \( n(t) \), the group of officially identified infections. Another kind is the infection of those who have developed the immunoglobulin against the SARS-CoV-2 but harbor the virus transiently before completely eliminating it or maintaining it permanently in the body. According to the data from New York and a part of California there are at least several percent of the population who belong to this category. It is possible that they transmit the virus to large numbers of people before they are no longer infectious.

We take two examples to estimate the number of unidentified symptomatic infections that is designated as \( U \) in our model: Japan and Tokyo. We use \( t \) as the time variable.

(1) Japan case. We calculate \( \sigma_{\Delta t} = \sigma/(1 + \Delta t \sigma) \) with \( \sigma = 0.127 \) (section 6) and \( \Delta t = 10 \).

We find \( \sigma_{\Delta t} = 0.056 \) and this gives,
\[
U = \frac{1}{\sigma_{\Delta t}} \frac{dn}{dt} = \frac{1}{0.056} \frac{dn}{dt} = 17.2 \exp[4.29867 + 0.20722t - 0.0061699t^2 + 0.0003387t^3],
\]
with \( t = 0 \) corresponding to March 26. We plot this in Fig. 49 from May 1 to May 31. This shows that \( U \) was above 1000 on May 1 but will decrease to small number towards the end of May.

(2) Tokyo case. \( \sigma_{\Delta t} = 0.127 \) (section 5.A.) gives,
\[
U = \frac{1}{\sigma_{\Delta t}} \frac{dn}{dt} = \frac{1}{0.127} \frac{dn}{dt} = 7.9 \exp[2.75961 + 0.196892t - 0.0044258t^2 + 6.67625 \times 10^{-6} t^3].
\]

We plot this in Fig. 50 from May 1 to May 24 (when \( dn/dt \) goes to 1).

This plot shows that there were about 100 unidentified symptomatic infections on May 1 but it will disappear on May 24 when \( dn/dt \) becomes less than 1. In fact, taking into account the time lag, unidentified infection with symptoms may have disappeared already by May 14.

But these numbers do not seem to include the asymptomatic infections as we will show later in this appendix. There may be more non-symptomatic cases.
It is important to realize that the remaining unidentified infections without symptoms can cause a lot of damage without further intervention. Therefore, we treat this issue here. We can divide \( U(t) \) into two groups; those with symptoms, and those without symptoms (due to two reasons as given above) even after infected.

We denote the number of the former and that of the latter, as \( U_1 \) and \( U_2 \), respectively (we neglect the effect of the finiteness of the time gap between the infection and the test).

We modify the original differential equations \( dU / dt = \kappa U \) in the following:
\[
\begin{align*}
\frac{dU_1}{dt} &= (K_{11} - \rho_1 - \sigma_1) U_1 + K_{12} U_2, \quad \text{(5.3a)} \\
\frac{dU_2}{dt} &= (K_{22} - \rho_2 - \sigma_2) U_2 + K_{21} U_1. \quad \text{(5.3b)}
\end{align*}
\]

The conservation of the infection by the PCR test demands that,
\[
\frac{dn}{dt} = \sigma_1 U_1 + \sigma_2 U_2. \quad \text{(5.4)}
\]

For simplicity, we assume,
\[
K_{11} = K_{22} = K_{12} = K_{21} \equiv K. \quad \text{(5.5)}
\]

And, by definition,
\[
\sigma_2 = 0. \quad \text{(5.6)}
\]

Equations [5.3a] and [5.3b] become,
\[
\begin{align*}
\frac{dU_1}{dt} &= (K - \rho_1 - \sigma_1) U_1 + K U_2 \equiv \kappa_1 U_1 + K U_2, \quad \text{(5.7a)} \\
\frac{dU_2}{dt} &= (K - \rho_2 - \sigma_2) U_2 + K U_1 \equiv \kappa_2 U_2 + K U_1. \quad \text{(5.7b)}
\end{align*}
\]

From these equations we get,
\[
\frac{d^2U_{1,2}}{dt^2} - (\kappa_1 + \kappa_2) \frac{dU_{1,2}}{dt} - (K^2 - \kappa_1 \kappa_2) U_{1,2} = 0. \quad \text{(5.8)}
\]

By putting,
\[
U_{1,2} = \xi_{1,2} e^{\lambda_{1,2} t} + \zeta_{1,2} e^{\lambda_{1,2} t}. \quad \text{(5.9)}
\]

We get,
\[
\lambda_{\pm} = \frac{1}{2} \left( 2K - \rho_1 - \sigma_1 - \rho_2 \pm \sqrt{(\rho_1 + \sigma_1 - \rho_2)^2 + 4K^2} \right). \quad \text{(5.10)}
\]

And therefore,
\[
\begin{align*}
\xi_{1,2} &= \frac{K}{\lambda_+ - \kappa_{1,2}} \xi_{2,1}, \quad \text{(5.11a)} \\
\zeta_{1,2} &= \frac{K}{\lambda_- - \kappa_{1,2}} \zeta_{2,1}. \quad \text{(5.11b)}
\end{align*}
\]

The “constants” \( \kappa \) and \( K \) may in fact vary with time; at the very least, their values at the initial stage are different from those at a later stage.

(1) Initially, when infections are increasing, \( \lambda_+ \) is positive; \( \lambda_- \) can be positive or negative.

(2) Later when the \( U_1 \) infection is decreasing,
\[
\lambda_+ \leq 0 \quad \text{and} \quad \lambda_- \leq 0. \quad \text{(5.12)}
\]

Generally, none of the constants \( \xi_{1,2} \) and \( \zeta_{1,2} \) vanish unless \( K \) vanishes.

\( \lambda_- \) is more negative than \( \lambda_+ \) so that later behavior is dominated by \( \lambda_+ \).

Therefore, we get,
\[
\frac{U_2}{U_1} \approx \frac{\xi_{2} e^{\lambda_{1,2} t}}{\xi_{1} e^{\lambda_{1,2} t}} = \frac{\lambda_+ - \kappa_1}{K} = \frac{1}{2} \left( -\kappa_1 + \kappa_2 + \sqrt{(\kappa_1 - \kappa_2)^2 + 4K^2} \right) = \frac{K}{-\kappa_1 + \kappa_2} = -\left( K - \rho_1 - \sigma_1 \right) + \left( K - \rho_2 \right) = \rho_1 + \sigma_1 - \rho_2. \quad \text{(5.13)}
\]

We get from these equations,
\[
\frac{U_2}{U_1} = \frac{\rho_1 + \sigma_1 - \rho_2 + \sqrt{(\rho_1 + \sigma_1 - \rho_2)^2 + 4K^2}}{2K}. \quad \text{(5.15)}
\]

We consider two cases:

(1) \( \rho_2 \geq \rho_1 + \sigma_1 \). This case corresponds to the assumption that those without symptoms have much stronger resistance to the infection due to stronger personal immunity which may also result in faster recovery.

In this case we have,
During the lockdown $K$ may be small ($K = \lambda t^2$) and we have,
\[ \frac{U_2}{U_1} \sim \frac{K}{\rho_2}, \]
[5.17]
But after opening the lockdown $K$ is no longer small and if it is larger than $\rho_2$, we get,
\[ \frac{U_2}{U_1} \sim \frac{\rho_1}{\rho_2}. \]
[5.18]

In this case we have,
\[ \lambda_+ = K \frac{U_2}{U_1} + \kappa_1 = K \frac{U_2}{U_1} + K - \rho_1 - \sigma_1, \]
[5.19]
\[ \approx K - \rho_1 - \sigma_1: \text{during lockdown}, \]
[5.20]
\[ \approx 2K - \rho_1 - \sigma_1: \text{after opening lockdown}. \]
[5.21]

(2) If natural recovery does not change very much between $U_1$ and $U_2$, as was shown in at least some results of the study (https://www.fukuishimbum.co.jp/articles/-/1071784, http://www.kansensho.or.jp/uploads/files/topics/2019ncov/covid19_casereport_200409_5.pdf (both in Japanese)), we may put,
\[ \rho_1 = \rho_2. \]
[5.22]
Then we have,
\[ \frac{U_2}{U_1} \sim \frac{\sigma_1 + \sqrt{\sigma_1^2 + 4K^2}}{2K}. \]
[5.23]

This formula also describes the situation at the time of reopening after lockdown: During the lockdown when $\sigma_1 \gg K$, we have,
\[ \frac{U_2}{U_1} \sim \frac{\sigma_1}{K}. \]
[5.24]
In this case of $\rho_1 \sim \rho \sim \rho_2$, and $\sigma_1 > \rho > K$, contrary to the above case (1), we have a quite different situation: $U_2/U_1 \gg 1$, after some time. This means the number of asymptomatic infections will be much larger than the symptomatic ones later. This may be important to take into consideration to estimate the second wave of COVID-19.
\[ \lambda_+ = K \frac{U_2}{U_1} + \kappa_1 = K \frac{U_2}{U_1} + K - \rho_1 - \sigma_1 = K - \rho_1, \]
[5.25]
showing that the contribution of $\sigma_1$ does not appear in this formula unlike the case (1). The actual situation could be something in between with a mixture of case (1) and case (2).

Also, the transition time of reopening after lockdown and the time lag between infection and expression of symptoms must be taken into account. Here we considered extreme cases to demonstrate the importance of asymptomatic infections.

This result means two things:

(1) During the opening of the lockdown it is important to find those infected but without symptoms by combining the antibody test, antigen test and PCR test;

(2) It is also important to loosen the criteria for PCR test as the Japanese government accepted. Otherwise, we will be forced to the lockdown status again and again.

**Appendix 6. Sensitivity of PCR test**

There exist large number of studies on this subject. For example, there is a report from FDA (Food and Drug Administration) entitled, “Accelerated Emergency Use Authorization (EUA) Summary Covid-19 RT-PCR Test (Laboratory Corporation of America)” (https://www.fda.gov/media/136151/download).

On the subject we refer to a few which serve for our purpose.

(1) **Research in vitro.** The main purpose of this research is to determine the minimum number of SARS-CoV-2 “molecules” per µL that can be detected by the Polymerase Chain Reaction (PCR) method. It depends on which part of the RNA sequence of the virus is used and how much of the complementary primer is used. It also depends on the duration of the polymerase chain reaction that multiplies copies of the sequence. Here we quote one study from the abovementioned FDA report, which claims that 6.5 copies of RNA/µL can be detected.

(2) **Statistical analysis.** Sensitivity in the field is significantly different from the above studies in vitro. It depends on where the swab is taken: nasal, throat or sputum. It also depends crucially on the time when the test is performed: how many days after onset of symptoms. Statistical research by the Oxford group based on 298 tests across 30 patients (150 nasal and 148 throat swabs) shows the following:

(i) Positive test probability declines significantly as time passes from the onset of symptoms: 94.39% on day 0 to 67.15% on day 10 for nasal swab and similarly for the throat. This indicates that the amount of virus de-
creases with this percentage either in nasal or throat.

(ii) False-negative probability is maximal (≈18%) 4 days after symptom onset. This suggests that the sensitivity of the PCR test in the field is above 80%.

(iii) From this result we may conclude that the PCR test is 80% sensitive within 10 days from the outset of symptoms and the power of test and also that of transmission declines significantly after approximately 10 days from symptom outset.

Acknowledgment

I appreciate greatly a friend of mine Jiro Arafune (former director general of Institute for Cosmic Ray Research of the University of Tokyo) for discussions and suggestions. I also thank Professor Toshimichi Ikemura who was one of my colleagues at The Graduate University for Advanced Studies and an expert of virus research. I would also like to express my deep appreciation to Nobuhiro Go and Koichi Yazaki both of whom are highly recognized scientists and classmates of mine from the Physics Department of the University of Tokyo. They gave me excellent advice and constructive guidance. Other classmates, Yasushi Arai, Reiji Sano and Yuko Tokieda, encouraged me strongly for which I express my sincere gratitude. Finally, I would like to thank Jonathan Miller of OIST for his various useful comments and his suggestion to publish the draft paper.

References

1) Adam, D.C., Wu, P., Wong, J.Y., Lau, E.H.Y., Tsang, T.K., Cauchemez, S. et al. (2020) Clustering and superspreading potential of SARS-CoV-2 infections in Hong Kong. Nat. Med. 26, 1714–1719.

2) Schuchat, A. (2020) Public health response to the initiation and spread of pandemic COVID-19 in the United States. February 24–April 21, 2020. MMWR Morb. Mortal. Wkly. Rep. 69, 551–556.

3) Furuse, Y., Sando, E., Tsuchiya, N., Miyahara, R., Yasuda, I., Ko, Y.K. et al. (2020) Clusters of coronavirus disease in communities, Japan, January–April 2020. Emerg. Infect. Dis. 26, 2176–2179.

4) Hammer, L., Dubbel, P., Capron, I., Ross, A., Jordan, A., Lee, J. et al. (2020) High SARS-CoV-2 attack rate following exposure at a choir practice—Skagit County, Washington, March 2020. MMWR Morb. Mortal. Wkly. Rep. 69, 606–610.

5) Jang, S., Han, S.H. and Rhee, J.-Y. (2020) Cluster of coronavirus disease associated with fitness dance classes, South Korea. Emerg. Infect. Dis. 26, 1917–1920.

6) Lloyd-Smith, J.O., Schreiber, S.J., Kopp, P.E. and Getz, W.M. (2005) Superspreading and the effect of individual variation on disease emergence. Nature 438, 355–359. There is also a theoretical support for the existence of super spreaders in the framework of network model of social connection.

7) Odagaki, T. (2020) Analysis of the outbreak of COVID-19 in Japan by SIQR model. Infect. Dis. Model 5, 691–698.

8) Kermack, W.O. and McKendrick, A.G. (1927) A contribution to the mathematical theory of epidemics. Proc. R. Soc. Lond. A 115, 700–721.

9) Nishiura, H. and Inaba, T. (2006) Prediction of infectious disease outbreak with particular emphasis on the statistical issues using transmission model. Statistical Math. 54, 461–480 (in Japanese with English abstract).

10) Kamikubo, Y. and Takahashi, A. (2020) Epidemiological tools that predict partial herd immunity to SARS coronavirus 2 (preprint). doi:10.1101/2020.03.25.20043679.

11) Nakano, T. and Ikeda, Y. (2020) Novel indicator of change in COVID-19 spread status (preprint). doi:10.1101/2020.04.25.20080200.

12) Wikramaratna, P., Paton, R.S., Ghafari, M. and Lourenco, J. (2020) Estimating the false-negative test probability of SARS-CoV-2 by RT-PCR (preprint). doi:10.1101/2020.04.05.20033355.

(Received July 13, 2020; accepted Oct. 17, 2020)