

# Relations of parameters for describing the epidemic of COVID—19 by the Kermack—McKendrick model

Toshihisa Tomie

Changchun University of Science and Technology, Changchun, Jilin, China

[tomie@cust.edu.cn](mailto:tomie@cust.edu.cn)

## Abstract

In order to quantitatively characterize the epidemic of COVID—19, useful relations among parameters describing an epidemic in general are derived based on the Kermack—McKendrick model. The first relation is  $1/\tau_{grow} = 1/\tau_{trans} - 1/\tau_{inf}$ , where  $\tau_{grow}$  is the time constant of the exponential growth of an epidemic,  $\tau_{trans}$  is the time for a pathogen to be transmitted from one patient to uninfected person, and the infectious time  $\tau_{inf}$  is the time during which the pathogen keeps its power of transmission. The second relation  $p(\infty) \approx 1 - \exp(-(R_0 - 1)/0.60)$  is the relation between  $p(\infty)$ , the **final size** of the disaster defined by the ratio of the total infected people to the population of the society, and the **basic reproduction number**,  $R_0$ , which is the number of persons infected by the transmission of the pathogen from one infected person during the infectious time. The third relation  $1/\tau_{end} = 1/\tau_{inf} - (1 - p(\infty))/\tau_{trans}$  gives the decay time constant  $\tau_{end}$  at the ending stage of the epidemic. Derived relations are applied to influenza in Japan in 2019 for characterizing the epidemic.

## 1. Introduction

We reported the understanding of the present status and forecasting of pneumonia by COVID—19 in China which is supposed to have originated in Wuhan by analyzing the data up to February 11 (ref.1). In ref.1, we clarified that the behavior of the epidemic was different in different regions and that the outbreak was well described by a Gaussian as was for influenza in Japan. We reported the following; 1. the epidemic in China passed the peak in the beginning of February, 2. the date of the epidemic peak was different by region and was the latest in Wuhan. For more than 10 days after our forecast, the epidemic closely followed our forecast. The new patients outside Wuhan decreased to less than 4 % of that at the peak around February 5.

Although the epidemic in China is near the end, patients by COVID—19 are found in many other countries, and COVID—19 is still a big fear of the people over the world. In order to forecast the epidemic of COVID—19 in other countries, we need to theoretically characterize the COVID—19 epidemic by fitting a model calculation to the data observed in China. We choose the Kermack—McKendrick model (ref.2) for our analysis. The Kermack—McKendrick model was proposed as early

**NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.**

as 1927, but still, it is the basis of many modified models for describing epidemics.

## 2. Present status of COVID—19 in China

Figure 1 shows the changes of newly infected people of COVID—19 in China. The data were taken from refs. 3 and 4. Up to 40<sup>th</sup> day, *i.e.*, February 9, the outbreak of the epidemic was described well by a Gaussian as was clarified in ref.1. The fitting Gaussians are  $850 \cdot \exp(-((x-33)/8)^2)$  for mainland China except Hubei province,  $1100 \cdot \exp(-((x-34.2)/7)^2)$  for Hubei province except Wuhan, and  $1850 \cdot \exp(-((x-38)/7)^2)$  for Wuhan. In ref.1, we mentioned that influenza in Japan shifted from a Gaussian to an exponential decay at the skirt of the epidemic.

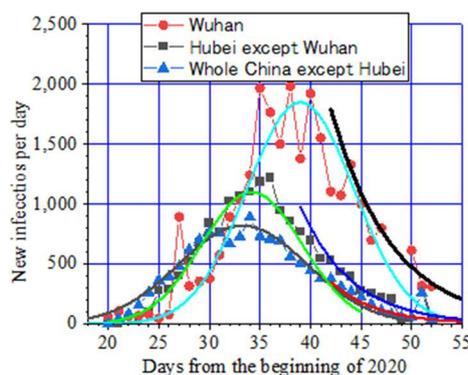


FIG.1; Present status of COVID—19 in China. The shapes of the epidemic in three regions changed from a Gaussian to an exponential decay around February 10.

The decay of COVID—19 in China was forecasted by assuming an exponential decay. Reference 1 was written on February 12. Since that day, the epidemic of COVID—19 in China has been changing as forecasted. On February 20, the number of newly infected people was only one-thirtieth of that at the peak in mainland China except Hubei province. The fitted exponential decay curves are  $350 \cdot \exp(-(x-42)/4.5)$  for mainland China except Hubei province,  $800 \cdot \exp(-(x-40)/5)$  for Hubei province except Wuhan, and  $2700 \cdot \exp(-(x-40)/5)$  for Wuhan.

We want to analyze the epidemic of COVID—19 in detail by applying a model. As a preparation of the analysis, useful relations of parameters describing an epidemic are derived in the present paper.

## 3. Derivation of the relations among parameters

Useful relations of parameters of an epidemic are Eqs. (7) to (10) in the following, which are derived from the Kermack—McKendrick model (ref.2). The number of susceptible people in the group is set as  $S_0$ , the number of infected persons is  $I(t)$ , the number of persons who have been infected and recovered to obtain immunity or have died is  $R(t)$ , the number of persons who is susceptible but not yet infected is  $S(t)$ .  $S_0 = S(t) + I(t) + R(t)$ . When the **transmission power** of the disease is set as  $\beta$  and the **recovery rate** from the infection is  $\gamma$ , the epidemic of the disease is given by the following three differential equations.

$$dS(t)/dt = -\beta S(t)I(t) \quad \text{----- (1)}$$

$$dR(t)/dt = \gamma I(t) \quad \text{----- (2)}$$

$$dI(t)/dt = \beta S(t)I(t) - \gamma I(t) \quad \text{----- (3)}$$

We define the **initial transmission time constant**,  $\tau_{trans}$ , after which a pathogen is transmitted to one susceptible host in the early stage of the epidemic and the **infectious time of** a pathogen,  $\tau_{inf}$ , after which the pathogen loses the power of transmission as follows,

$$\tau_{trans} = 1/\beta S_0 \quad \text{----- (4)}$$

$$\tau_{inf} = 1/\gamma \quad \text{----- (5)}$$

The **basic reproduction number**,  $R_0$ , which is the number of persons infected by the transmission of a pathogen from one infected person during the infectious time, is given by

$$R_0 = \tau_{inf} / \tau_{trans} \quad \text{----- (6)}$$

In the following, we see that an epidemic starts with exponential growth with a time constant  $\tau_{grow}$  given by

$$1/\tau_{grow} = 1/\tau_{trans} - 1/\tau_{inf} \quad \text{----- (7)}$$

and decays exponentially with a time constant  $\tau_{end}$  given by

$$1/\tau_{end} = 1/\tau_{inf} - (1-p(\infty))/\tau_{trans} \quad \text{----- (8)}$$

Here,  $p(\infty)$  is the **final size** of the disaster which is defined by the ratio of the total number of the infected people,  $R(\infty)$ , to the population of the society,  $S_0$ . As shown later,  $p(\infty)$  is approximated as,

$$p(\infty) \approx 1 - \exp(-(R_0 - 1)/0.60) \quad \text{----- (9)}$$

By using  $R_0$ , Eq. (7) is rewritten as

$$\tau_{grow} = \tau_{inf} / (R_0 - 1) \quad \text{----- (10)}$$

$R_0$ ,  $\tau_{inf}$ , and  $p(\infty)$  are parameters describing an epidemic. The above relations, Eqs. (7) to (10), are derived as follows.

In the beginning of the epidemic,  $I(t)$  and  $R(t)$  are negligible compared to  $S_0$ , and Eq.(3) becomes

$$dI(t)/dt \approx (\beta S_0 - \gamma) I(t), \quad \text{----- (11)}$$

and we get

$$I(t) = \exp(t/\tau_{grow}) I(0). \quad \text{----- (12)}$$

Thus, Eq. (7) is derived.

The condition of spreading the disease is  $\tau_{grow} > 0$ , and then,  $R_0 > 1$ . For this condition,  $S_0$  should be larger than the **transmission threshold**  $S_{th}$  given by

$$S_{th} = \gamma/\beta \quad \text{----- (13)}$$

As infection proceeds,  $S(t)$  reduces. When  $\beta S(t) - \gamma$  becomes negative, the number of infected persons  $I(t)$  begins to decrease according to Eq. (3), and eventually,  $I(t)$  becomes 0. When  $I(t)$  is 0, the reduction of the number of the uninfected susceptible persons  $S(t)$  stops, as seen in Eq. (1), and  $S(t)$  converges to  $S(\infty)$ .

By dividing Eq. (3) by Eq. (1), we get

$$dI(t)/dS(t) = -1 + S_{th}/S(t) \quad \text{----- (14)}$$

and then we get

$$I(t) = I(0) + S_0 - S(t) + S_{th} \ln (S(t)/S_0) \quad \text{----- (15)}$$

Because  $I(\infty)=0$  and  $I(0) \ll S_0$ , we get

$$S(\infty) - S_0 = S_{th} \ln (S(\infty)/S_0). \quad \text{----- (16)}$$

We write the ratio of the population of the recovered person  $R(t)$  to  $S_0$  as  $p(t)$ ,

$$p(t) = R(t)/S_0. \quad \text{----- (17)}$$

By replacing  $S(\infty)/S_0$  by  $p(\infty)$ , Eq. (16) becomes

$$\exp (-p(\infty) R_0)=1-p(\infty). \quad \text{----- (18)}$$

In deriving Eq. (18), we employed the relation

$$R_0 = S_0 / S_{th}. \quad \text{----- (19)}$$

which is derived from Eqs. (4), (5), (6), and (13). The number  $p(\infty)$  is referred as the “**final size**” and Eq. (18) is referred as the “**final size equation**”.

At the shrinking stage of the spread of the disease,  $S(t)$  converges to  $S(\infty)$ . Then, Eq. (3) becomes,

$$dI(t)/dt = -(\gamma - \beta S(\infty)) I(t) \quad \text{----- (20)}$$

which gives exponential decay of the patients.

$$I(t)/dt = \exp (-t/\tau_{end}) I_{grow} \quad \text{----- (21)}$$

The time constant of the decay  $\tau_{end}$  is given by

$$\begin{aligned} 1/\tau_{end} &= \gamma - \beta S(\infty) \\ &= \gamma - \beta S_0(1-p(\infty)) \\ &= 1/\tau_{inf} - (1-p(\infty))/\tau_{trans} \end{aligned} \quad \text{----- (22)}$$

Thus, Eq. (8) is derived.

The final size equation was numerically solved and the result is shown by the solid curve in Fig.2. As  $R_0$  increases from 1,  $p(\infty)$  increases from zero and saturates at a large  $R_0$ . When  $R_0$  is larger than 2.5, more than 90 % of people in the group are infected. As shown by the dotted curve in Fig.2,  $p(\infty)$  can be approximated by the equation,  $1 - \exp(-(R_0-1)/0.60)$ . Thus, Eq. (9) is derived.

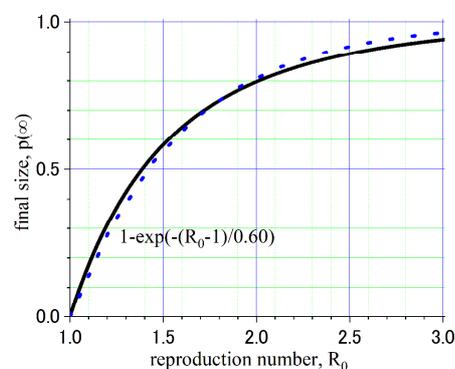


FIG.2: Numerically calculated final size  $p(\infty)$  of an epidemic as a function of re-production number,  $R_0$ .

#### 4. Applying the model to the epidemic of influenza in Japan

We apply the above model to the epidemic of influenza in Japan as cited in ref.1 as a reference for the general epidemic. Figure 3 shows the influenza epidemic in Japan over the past decade (ref.5). We choose influenza in 2019, which is referred as JpnInf2019 in this paper. In ref.1, we fitted a Gaussian to JpnInf2019. As shown by the red solid curve in Fig. 4, the model described above reproduces JpnInf2019 better than a simple Gaussian. The parameters were  $\tau_{trans} = 0.52$  weeks and  $\tau_{inf} = 1$  week.

From the values of  $\tau_{trans}$  and  $\tau_{inf}$  for fitting, the reproduction number  $R_0 = 1.92$  and the final size

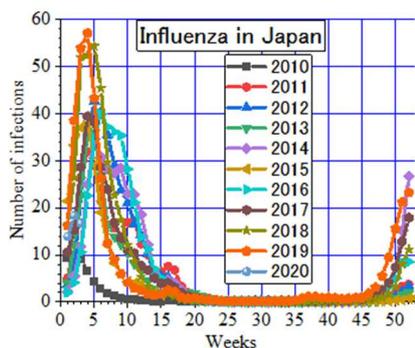


FIG.3: Epidemic of influenza in Japan over the past decade.

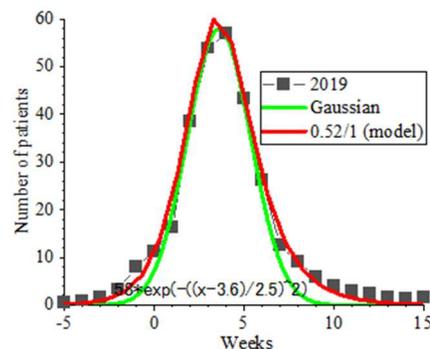


FIG.4: Reproducing JpnInf2019 by the model with  $\tau_{trans}=0.52$  weeks and  $\tau_{inf}=1$  week.

estimated by Eq. (18) is  $p(\infty) = 0.77$ . Ministry of Health, Labor and Welfare of the Japanese government reported that the estimated number of patients who consulted doctors was about 12,000, 000 over the country (ref. 6). This estimated value was only 10 % of the population of Japan and far smaller than  $p(\infty) = 77\%$ . There are three possibilities for the reason. The first is that most patients did not go to a hospital and the real number of patients was 92,400, 000. The second is that only 1/8 of the area in Japan was the reach of the pathogen. The third is a combination of the first and the second reasons. Thus, we realize the evaluation of the final size  $p(\infty)$  is difficult.

While reproduction of the epidemic by the model is fairly good near the peak of the epidemic, but at the far skirt, deviation of the data from the model is not negligible as seen in Fig.5 which shows a log plot of Fig.4. In the model fitting, the time constant was 1.15 weeks at the rising and the decay time constant was 1.7 weeks. Hence,  $\tau_{grow}$  calculated by Eq. (7) was 1.1 weeks. But in JpnInf2019, the reported number of patients increased exponentially with a time constant of 1.5 weeks and decreased exponentially at the ending stage with a time constant of 2.5 weeks. By increasing  $\tau_{trans}$  to 0.61 weeks

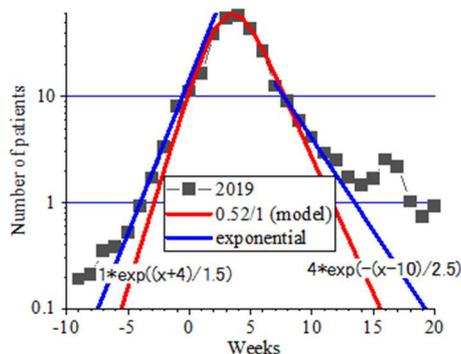


FIG.5: A log plot of Fig.4. At the skirt, the increase and the decrease of the epidemic is slower than the model.

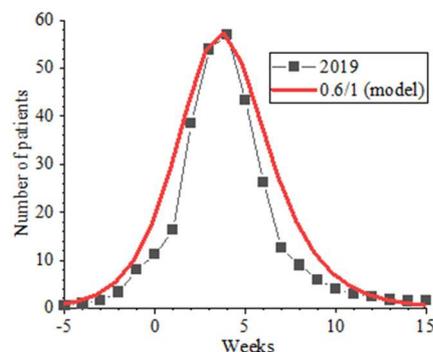


FIG.6: If the observed  $\tau_{grow}$  and  $\tau_{end}$  are used, the model cannot reproduce the main part of the epidemic.

from 0.52 weeks, the time constant of the rising of the epidemic can be increased to 1.5 weeks in the model calculation, but the width of the epidemic in the model calculation was too wider than the real one as shown in Fig.6.

In all the cases of the epidemic, there should be a difference between the reported number of patients and the real number of infected people. Not all infected people will go to a hospital and not all patients are not inspected by medical institutes. In analyzing statistics, we assume the ratio of the reported number to the real number is constant during the epidemic, but often the assumption can be wrong. We expect the ratio is smaller in the beginning and at the ending of an epidemic. Then, the time constant of the “apparent rise” and “apparent ending” will be shorter than the real ones. However, this expectation was the opposite in Fig.5.

The slower increase and the slower decrease at the skirt of the epidemic could suggest the transmission power  $\beta$  of the virus may change in time, which, we think, is not plausible. At present, we do not know the reason why the simple Kermack–McKendrick cannot reproduce the whole epidemic including the skirt parts. Thus, from the analysis of JpnInf2019 by the model described above, we learn that it is important to remember that the time constant of the “apparent” increase of the epidemic in the early stage does not re-produce the whole epidemic.

The above information helps greatly to understand the to-be-planned analysis of COVID—19.

## 5. Summary

Useful relations of parameters of an epidemic are derived by following the Kermack–McKendrick model. The first relation is  $1/\tau_{grow} = 1/\tau_{trans} - 1/\tau_{inf}$ . The final size of the disaster was numerically calculated and we found it is given by  $p(\infty) \approx 1 - \exp(-(R_0 - 1)/0.60)$ . The third relation  $1/\tau_{end} = 1/\tau_{inf} - (1 - p(\infty))/\tau_{trans}$  gives the decay time constant  $\tau_{end}$  at the last stage of the epidemic.

By applying the model, we found that the epidemic of influenza in Japan in 2019 was re-produced by the parameters;  $\tau_{trans} = 0.52$  week and  $\tau_{inf} = 1$  week and that  $\tau_{grow}$  observed in the early stage can be different from  $\tau_{grow}$  for re-producing the overall epidemic.

## References

- 1) Toshihisa Tomie, “Understanding the present status and forecasting of COVID—19 in Wuhan”, medRxiv preprint <https://doi.org/10.1101/2020.02.13.20022251>
- 2) W.O.Kermack and A.G.McKendrick, “A contribution to the mathematical theory of epidemics”, Proc. R. Soc. London A **115** (1927) 700.
- 3) National Health Commission of the People’s Republic of China, Prevention of Epidemics (in Chinese)

[http://www.nhc.gov.cn/xcs/yqtb/list\\_gzbd.shtml](http://www.nhc.gov.cn/xcs/yqtb/list_gzbd.shtml)

4) Health Commission of Hubei Province, Anti-Prevention New Coronary Disease Infection, Information (in Chinese)

<http://wjw.hubei.gov.cn/bmdt/ztzl/fkxxgzbdgrfyq/xxfb/index.shtml>

5) National Institute of Infectious Diseases, IDWR Bulletin Data 4th Week 2020

<https://www.niid.go.jp/niid/ja/data.html>

“Number of reports at fixed disease points- Comparison with the past 10 years- “ (in Japanese)

6) National Institute of Infectious Diseases and

Ministry of Health, Labor and Welfare Tuberculosis Infectious Disease Division,

Page 7 in “Influenza in this year 2018/2019” (in Japanese)

<https://www.niid.go.jp/niid/images/idsc/disease/influ/fludoco1819.pdf>