

# Understanding unreported cases in the 2019-nCov epidemic outbreak in Wuhan, China, and the importance of major public health interventions

Z. LIU<sup>(a)\*†</sup>, P. MAGAL<sup>(a,b)‡†</sup>, O. SEYDI<sup>(c)</sup>, AND G. WEBB<sup>(d)†</sup>

<sup>(a)</sup> *School of Mathematical Sciences, Beijing Normal University,  
Beijing 100875, People's Republic of China*

<sup>(b)</sup> *Univ. Bordeaux, IMB, UMR 5251, F-33400 Talence, France.  
CNRS, IMB, UMR 5251, F-33400 Talence, France.*

<sup>(c)</sup> *Département Tronc Commun, École Polytechnique de Thiès, Sénégal*

<sup>(d)</sup> *Mathematics Department, Vanderbilt University, Nashville, TN, USA*

February 4, 2020

## 1 Abstract

We develop a mathematical model to provide epidemic predictions for the 2019-nCov epidemic in Wuhan, China. We use reported case data up to January 31, 2020 from the Chinese Center for Disease Control and Prevention and the Wuhan Municipal Health Commission to parameterize the model. From the parameterized model we identify the number of unreported cases. We then use the model to project the epidemic forward with varying level of public health interventions. The model predictions emphasize the importance of major public health interventions in controlling 2019-nCov epidemics.

## 2 Author summary

An epidemic outbreak of a new human coronavirus 019-nCov, has occurred in Wuhan, China. For this outbreak, the unreported cases and the disease transmission rate are not known. In order to recover these values from reported medical data, we present the mathematical model (4.1) for outbreak diseases. By knowledge of the cumulative reported symptomatic infectious cases, and assuming (1) the fraction  $f$  of asymptomatic infectious that become reported symptomatic infectious cases, (2) the average time  $1/\nu$  asymptomatic infectious are asymptomatic, and (3) the average time  $1/\eta$  symptomatic infectious remain infectious, we estimate the epidemiological parameters in the model (4.1). We then make numerical simulations of the model (4.1) to predict forward in time the severity of the epidemic. We observe that public health measures, such as isolation, quarantine, and public closings, greatly reduce the final size of the epidemic, and make the turning point much earlier than without these measures. We observe that the predictive capability of model (4.1) requires valid estimates of the parameters  $f$ ,  $\nu$  and  $\eta$ , which depend on the input of medical and biological epidemiologists. Our results can contribute to the prevention and control of the 2019-nCov epidemic in Wuhan.

## 3 Introduction

An epidemic outbreak of a new human coronavirus, termed the novel coronavirus 2019-nCov, has occurred in Wuhan, China. The first cases occurred in early December, 2019, and by January 29, 2020 more than 7000 cases had been reported in China [1]. Early reports advise that 2019-nCov transmission

\*Research was partially supported by NSFC and CNRS (Grant Nos. 11871007 and 11811530272) and the Fundamental Research Funds for the Central Universities.

†Corresponding authors.

‡Research was partially supported by CNRS and National Natural Science Foundation of China (Grant No.11811530272)

may occur from an infectious individual, who is not yet symptomatic [2]. Evidently, such asymptomatic infectious cases are not reported to medical authorities. For epidemic influenza outbreaks, reported cases are typically only a fraction of the total number of the symptomatic infectious individuals. For the current epidemic in Wuhan, it is likely that intensive efforts by Chinese public health authorities, have reduced the number of unreported cases.

Our objective is to develop a mathematical model, which recovers from data of reported cases, the number of unreported cases for the 2019-nCov epidemic in Wuhan. For this epidemic a modeling approach has been developed in [3], which did not consider unreported cases. Our work continues the investigation in [4] and [5] of the fundamental problem of parameter identification in mathematical epidemic models. We address the following fundamental issues concerning this epidemic: How will the epidemic evolve in Wuhan with respect to the number of reported cases and unreported cases? How will the number of unreported cases influence the severity of the epidemic? How will public health measures, such as isolation, quarantine, and public closings, mitigate the final size of the epidemic?

## 4 Results

### 4.1 The model and data

Our model consists of the following system of ordinary differential equations:

$$\begin{cases} S'(t) = -\tau S(t)[I(t) + U(t)], \\ I'(t) = \tau S(t)[I(t) + U(t)] - \nu I(t) \\ R'(t) = \nu_1 I(t) - \eta R(t) \\ U'(t) = \nu_2 I(t) - \eta U(t). \end{cases} \quad (4.1) \quad \boxed{4.1}$$

Here  $t \geq t_0$  is time in days,  $t_0$  is the beginning date of the epidemic,  $S(t)$  is the number of individuals susceptible to infection at time  $t$ ,  $I(t)$  is the number of asymptomatic infectious individuals at time  $t$ ,  $R(t)$  is the number of reported symptomatic infectious individuals at time  $t$ , and  $U(t)$  is the number of unreported symptomatic infectious individuals at time  $t$ . This system is supplemented by initial data

$$S(t_0) = S_0 > 0, I(t_0) = I_0 > 0, R(t_0) = 0 \text{ and } U(t_0) = U_0 \geq 0. \quad (4.2) \quad \boxed{4.2}$$

The parameters of the model are listed in Table 1.

Symbol	Interpretation	Method
$t_0$	Time at which the epidemic started	fitted
$S_0$	Number of susceptible at time $t_0$	fixed
$I_0$	Number of asymptomatic infectious at time $t_0$	fitted
$U_0$	Number of unreported symptomatic infectious at time $t_0$	fitted
$\tau$	Transmission rate	fitted
$1/\nu$	Average time during which asymptomatic infectious are asymptomatic	fixed
$f$	Fraction of asymptomatic infectious that become reported symptomatic infectious	fixed
$\nu_1 = f \nu$	Rate at which asymptomatic infectious become reported symptomatic	fitted
$\nu_2 = (1 - f) \nu$	Rate at which asymptomatic infectious become unreported symptomatic	fitted
$1/\eta$	Average time symptomatic infectious have symptoms	fixed

Table1

Table 1: *Parameters of the model.*

We use three sets of reported data to model the epidemic in Wuhan: First, data from the Chinese CDC for all of China (Table 2), second, data from the Wuhan Municipal Health Commission for Hubei Province (Table 3), and third, data from the Wuhan Municipal Health Commission for Wuhan Municipality (Table 4). These data vary, but represent the epidemic transmission in Wuhan, from which almost all the cases originated in the larger regions.

Date	January	20	21	22	23	24	25	26	27	28	29
Confirmed cases (cumulated) for China		291	440	571	830	1287	1975	2744	4515	5974	7711
Mortality cases (cumulated) for China			9	17	25	41	56	80	106	132	170

Table2

Table 2: Reported case data Jan. 20, 2020 - Jan. 29, 2020, reported for all of China by the Chinese CDC [1].

Date	January	23	24	25	26	27	28	29	30	31
Confirmed cases (cumulated) for Hubei		549	729	1052	1423	2714	3554	4586	5806	7153
Mortality cases (cumulated) for Hubei		24	39	52	76	100	125	162	204	249

Table3

Table 3: Reported case data Jan. 23, 2020 - Jan. 31, 2020, reported for Hubei Province by the Wuhan Municipal Health Commission. [6].

Date	January	23	24	25	26	27	28	29	30	31
Confirmed cases (cumulated) for Wuhan		495	572	618	698	1590	1905	2261	2639	3215
Mortality cases (cumulated) for Wuhan		23	38	45	63	85	104	129	159	192

Table4

Table 4: Reported case data Jan. 23, 2020 - Jan. 31, 2020, reported for Wuhan Municipality by the Wuhan Municipal Health Commission. [6].

## 4.2 Comparison of the model (4.1) with the data

For influenza disease outbreaks, the parameters  $\tau$ ,  $\nu$ ,  $\nu_1$ ,  $\nu_2$ ,  $\eta$ , as well as the initial conditions  $S(t_0)$ ,  $I(t_0)$ , and  $U(t_0)$ , are usually unknown. Our goal is to identify them from specific time data of reported symptomatic infectious cases. To identify the unreported asymptomatic infectious cases, we assume that the cumulative reported symptomatic infectious cases at time  $t$  consist of a constant fraction along time of the total number of symptomatic infectious cases up to time  $t$ . In other words, we assume that the removal rate  $\nu$  takes the following form:  $\nu = \nu_1 + \nu_2$ , where  $\nu_1$  is the removal rate of reported symptomatic infectious individuals, and  $\nu_2$  is the removal rate of unreported symptomatic infectious individuals due to all other causes, such as mild symptom, or other reasons.

The cumulative number of reported symptomatic infectious cases at time  $t$ , denoted by  $CR(t)$ , is

$$CR(t) = \nu_1 \int_{t_0}^t I(s) ds. \quad (4.3) \quad 4.3$$

Our method is the following: We assume that  $CR(t)$  has the following special form:

$$CR(t) = \chi_1 \exp(\chi_2 t) - \chi_3. \quad (4.4) \quad 4.4$$

We evaluate  $\chi_1, \chi_2, \chi_3$  using the reported case data in Table 2, Table 3 and Table 4. We obtain the model starting time of the epidemic  $t_0$  from (4.4):

$$CR(t_0) = 0 \Leftrightarrow \chi_1 \exp(\chi_2 t_0) - \chi_3 = 0 \Rightarrow t_0 = \frac{1}{\chi_2} (\ln(\chi_3) - \ln(\chi_1)).$$

We fix  $S_0 = 11.081 \times 10^6$ , which corresponds to the total population of Wuhan. We assume that the variation in  $S(t)$  is small during the period considered, and we fix  $\nu, \eta, f$ . By using the method in section 6.1, we can estimate the parameters  $\nu_1, \nu_2, \tau$  and the initial conditions  $U_0$  and  $I_0$  from the cumulative reported cases  $CR(t)$  given (4.4). We then construct numerical simulations and compare them with data.

The evaluation of  $\chi_1, \chi_2$  and  $\chi_3$  and  $t_0$ , using the cumulative reported symptomatic infectious cases in Table 2, Table 3 and Table 4, is shown in Table 5 and in Figure 1 below.

Table 5

Name of the parameter	$\chi_1$	$\chi_2$	$\chi_3$	$t_0$
From Table 2 for China	0.16	0.38	1.1	5.12
From Table 3 for Hubei	0.23	0.34	0.1	-2.45
From Table 4 for Wuhan	0.36	0.28	0.1	-4.52

Table 5: Estimation of the parameters  $\chi_1$ ,  $\chi_2$ ,  $\chi_3$  and  $t_0$  by using the cumulated reported cases in Table 2, Table 3 and Table 4.

**Remark 4.1** According to the Table 2, Table 3 and Table 4, the time  $t = 0$  will correspond to December 31. So in Table 5, the value  $t_0 = 5.12$  means that the starting time of the epidemic is January 5, the value  $t_0 = -2.45$  means that the starting time of the epidemic is December 28, and  $t_0 = -4.52$  means that the starting time of the epidemic is December 26.

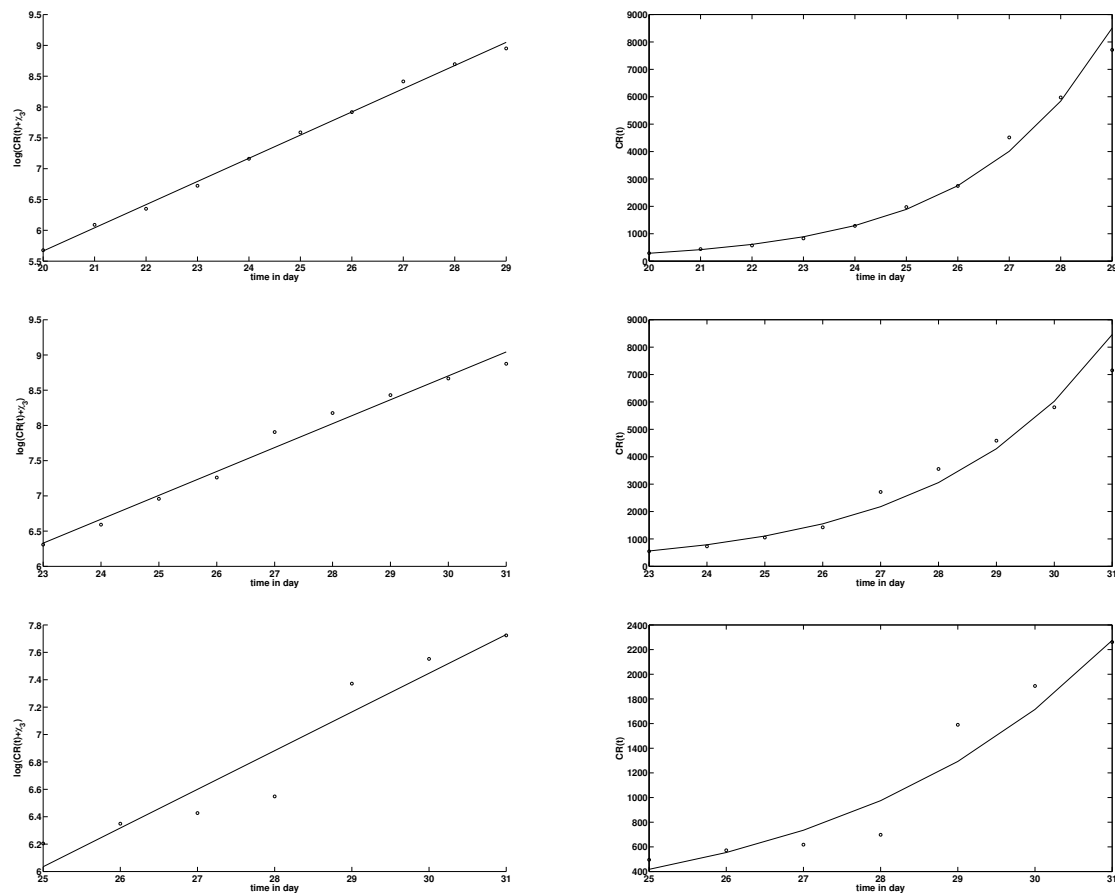


fig1

Figure 1: In the left side figures, the dots correspond to  $t \rightarrow \ln(CR(t) + \chi_3)$ , and in the right side figures, the dots correspond to  $t \rightarrow CR(t)$ , where  $CR(t)$  is taken from the cumulated confirmed cases in Table 2 (top), in Table 3 (middle) and in Table 4 (bottom). The straight line in the left side figures corresponds to  $t \rightarrow \ln(\chi_1) + \chi_2 t$ . We first estimate the value of  $\chi_3$  and then use a least square method to evaluate  $\chi_1$  and  $\chi_2$ . We observe that the data for China in Table 2 and Hubei in Table 3 provides a good fit for  $CR(t)$  in (4.4), while the data for Wuhan in Table 4 does not provide a good fit for  $CR(t)$  in (4.4).

**Remark 4.2** As long as the number of reported cases follows (4.1), we can predict the future values of  $CR(t)$ . For  $\chi_1 = 0.16$ ,  $\chi_2 = 0.38$  and  $\chi_3 = 1.1$ , we obtain

Jan.30	Jan.31	Feb.1	Feb.2	Feb.3	Feb.4	Feb.5	Feb.6
8510	12390	18050	26290	38290	55770	81240	118320

The actual number of reported cases for China are 8,163 confirmed for January 30, 11,791 confirmed for January 30, and 14,380 confirmed for February 1. So the exponential formula (4.4) overestimates the number reported after day 30.

From now on, we fix the fraction  $f$  of symptomatic infectious cases that are reported. We assume that between 80% and 100% of infectious cases are reported. Thus,  $f$  varies between 0.8 and 1. We assume  $1/\nu$ , the average time during which the patients are asymptomatic infectious varies between 1 day and 7 days. We assume that  $1/\eta$  the average time during which a patient is symptomatic infectious, varies between 1 day and 7 days. So, we fix  $f, \nu, \eta$ . Since  $f$  and  $\nu$  are known, we can compute

$$\nu_1 = f\nu \text{ and } \nu_2 = (1 - f)\nu. \quad (4.5) \quad \text{AUX4.5}$$

Moreover by following the approach described in the supplementary, we should have

$$I_0 = \frac{\chi_1 \chi_2 \exp(\chi_2 t_0)}{f\nu} = \frac{\chi_3 \chi_2}{f\nu}, \quad (4.6) \quad \text{AUX4.6}$$

$$\tau = \frac{\chi_2 + \nu}{S_0} \frac{\eta + \chi_2}{\nu_2 + \eta + \chi_2}, \quad (4.7) \quad \text{AUX4.7}$$

and

$$U_0 = \frac{\nu_2}{\eta + \chi_2} I_0 = \frac{(1 - f)\nu}{\eta + \chi_2} I_0. \quad (4.8) \quad \text{AUX4.8}$$

By using the approach described in the supplementary material, the basic reproductive number for model (4.1) is given by

$$R_0 = \frac{\tau S_0}{\nu} \left( 1 + \frac{\nu_2}{\eta} \right).$$

By using  $\nu_2 = (1 - f)\nu$  and (4.7) we obtain

$$R_0 = \frac{\chi_2 + \nu}{\nu} \frac{\eta + \chi_2}{(1 - f)\nu + \eta + \chi_2} \left( 1 + \frac{(1 - f)\nu}{\eta} \right). \quad (4.9) \quad \text{AUX4.9}$$

## section5

### 4.3 Numerical simulations

We can find multiple values of  $\eta, \nu$  and  $f$  which provide a good fit for the data. For application of our model,  $\eta, \nu$  and  $f$  must vary in a reasonable range. For the corona virus 2019-nCov epidemic in Wuhan at its current stage, the values of  $\eta, \nu$  and  $f$  are not known. From preliminary information, we use the values

$$f = 0.8, \eta = 1/7, \nu = 1/7.$$

By using the formula (4.9) for the basic reproduction number, we obtain from the data in Table 2, that  $R_0 = 4.13$ . Using model (4.1) and the values in Table 5, we plot the graph of  $t \rightarrow CR(t), t \rightarrow U(t)$  and the data for the confirmed cumulated cases in Figure 2, according to Table 2 for China, Table 3 for Hubei and Table 4 for Wuhan. We observe from these figures that the data for China and Hubei fit the model (4.1), but the data for Wuhan do not fit the model (4.1) because the model (4.4) is not a good model for the data for Wuhan in Table 4. The data for Wuhan do not fit an exponential function.

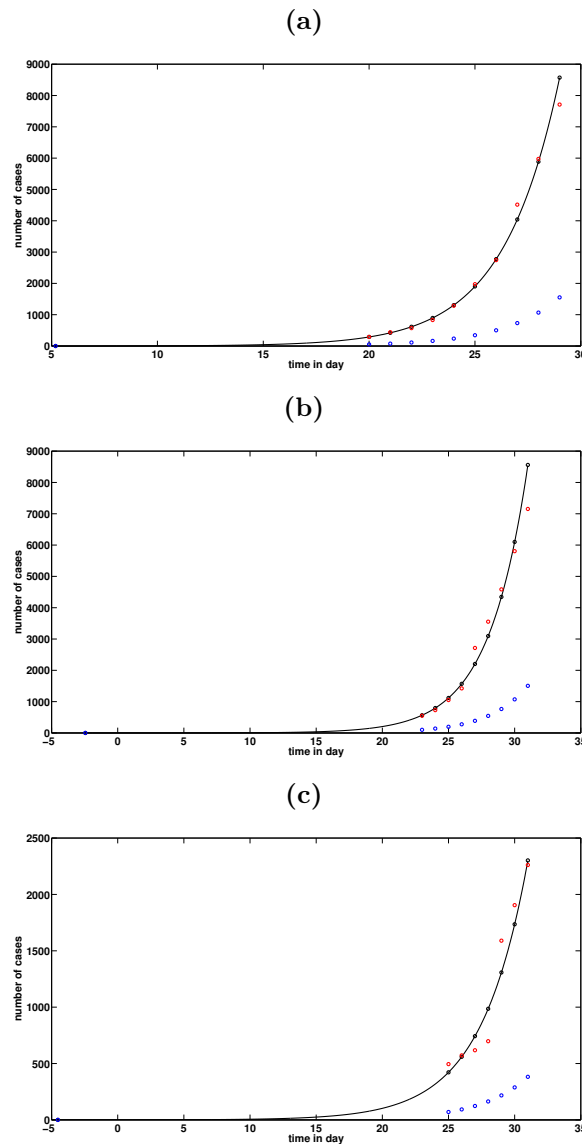


fig2

Figure 2: In these figures we use  $f = 0.8$ ,  $\eta = 1/7$ ,  $\nu = 1/7$  and  $S_0 = 11.081 \times 10^6$ . The remaining parameters are derived by using (4.6)-(4.8). In Figure (a), we plot the number of  $t \rightarrow CR(t)$  (black solid line) and  $t \rightarrow U(t)$  (blue dotted) and the data (red dotted) corresponding to the confirmed cumulated case for all China in Table 2. We use  $\chi_1 = 0.16$ ,  $\chi_2 = 0.38$ ,  $\chi_3 = 1.1$ ,  $t_0 = 5.12$  and  $S_0 = 11.081 \times 10^6$  which give  $\tau = 4.44 \times 10^{-08}$ ,  $I_0 = 3.62$ ,  $U_0 = 0.2$  and  $R_0 = 4.13$ . In Figure (b), we plot the number of  $t \rightarrow CR(t)$  (black solid line) and  $t \rightarrow U(t)$  (blue dotted) and the data (red dotted) corresponding to the confirmed cumulated case for Hubei province in Table 3. We use  $\chi_1 = 0.23$ ,  $\chi_2 = 0.34$ ,  $\chi_3 = 0.1$  and  $t_0 = -2.45$  and  $S_0 = 11.081 \times 10^6$  which give  $\tau = 4.11 \times 10^{-08}$ ,  $I_0 = 0.3$ ,  $U_0 = 0.02$  and  $R_0 = 3.82$ . In Figure (c), we plot the number of  $t \rightarrow CR(t)$  (black solid line) and  $t \rightarrow U(t)$  (blue dotted) and the data (red dotted) corresponding to the confirmed cumulated cases for Wuhan in Table 4. We use  $\chi_1 = 0.36$ ,  $\chi_2 = 0.28$ ,  $\chi_3 = 0.1$  and  $t_0 = -4.52$  and  $S_0 = 11.08 \times 10^6$  which give  $\tau = 3.6 \times 10^{-08}$ ,  $I_0 = 0.25$ ,  $U_0 = 0.02$  and  $R_0 = 3.35$ .

In what follows, we plot the graphs of  $t \rightarrow CR(t)$ ,  $t \rightarrow U(t)$ , and  $t \rightarrow R(t)$  for Wuhan by using model (4.1). We define the turning point  $t_p$  as the time at which the red curve (i.e., the curve of the non-cumulated reported infectious cases) reaches its maximum value. For example, in the figure below, the turning point  $t_p$  is day 54, which corresponds to February 23 for Wuhan.

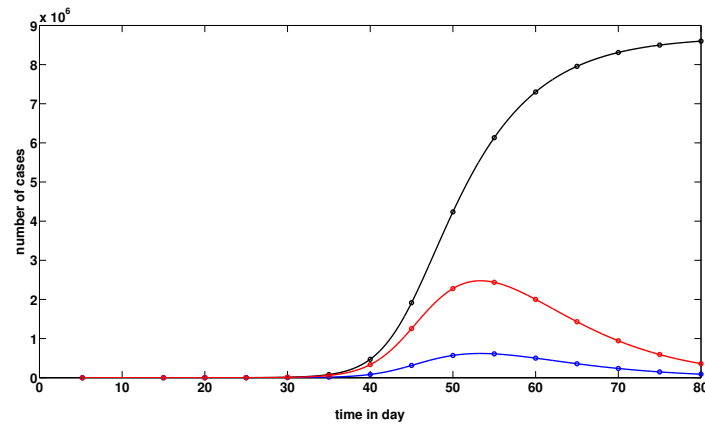


fig3

Figure 3: In this figure we plot the graphs of  $t \rightarrow CR(t)$  (black solid line),  $t \rightarrow U(t)$  (blue solid line) and  $t \rightarrow R(t)$  (red solid line). We use  $f = 0.8$ ,  $\eta = 1/7$ ,  $\nu = 1/7$ , and  $S_0 = 11.081 \times 10^6$ . The remaining parameters are derived by using (4.6)-(4.8). We obtain  $\tau = 4.44 \times 10^{-08}$ ,  $I_0 = 3.62$  and  $U_0 = 0.2$ . The cumulated number of reported cases goes up to 8.5 million people and the turning point is day 54. So the turning point is February 23 (ie. 54 – 31).

In the following we take into account the fact that very strong isolation measures have been imposed for all China since January 23. Specifically, since January 23, families in China are required to stay at home. In order to take into account such a public intervention, we assume that the transmission of 2019-nCov from infectious to susceptible individuals stopped after January 25. Therefore, we consider the following model: for  $t \geq t_0$

$$\begin{cases} S'(t) = -\tau(t)S(t)[I(t) + U(t)], \\ I'(t) = \tau(t)S(t)[I(t) + U(t)] - \nu I(t) \\ R'(t) = \nu_1 I(t) - \eta R(t) \\ U'(t) = \nu_2 I(t) - \eta U(t) \end{cases} \quad (4.10) \quad 4.10$$

where

$$\tau(t) = \begin{cases} 4.44 \times 10^{-08}, & \text{for } t \in [t_0, 25], \\ 0, & \text{for } t > 25. \end{cases} \quad (4.11) \quad 4.11$$

The figure below takes into account the public health measures, such as isolation, quarantine, and public closings, which correspond to model (4.10) and (4.11). By comparison of Figure 4-(a) with Figure 3, we note that these measures greatly mitigate the final size of the epidemic, and shift the turning point about 24 days before the turning point without these measures.

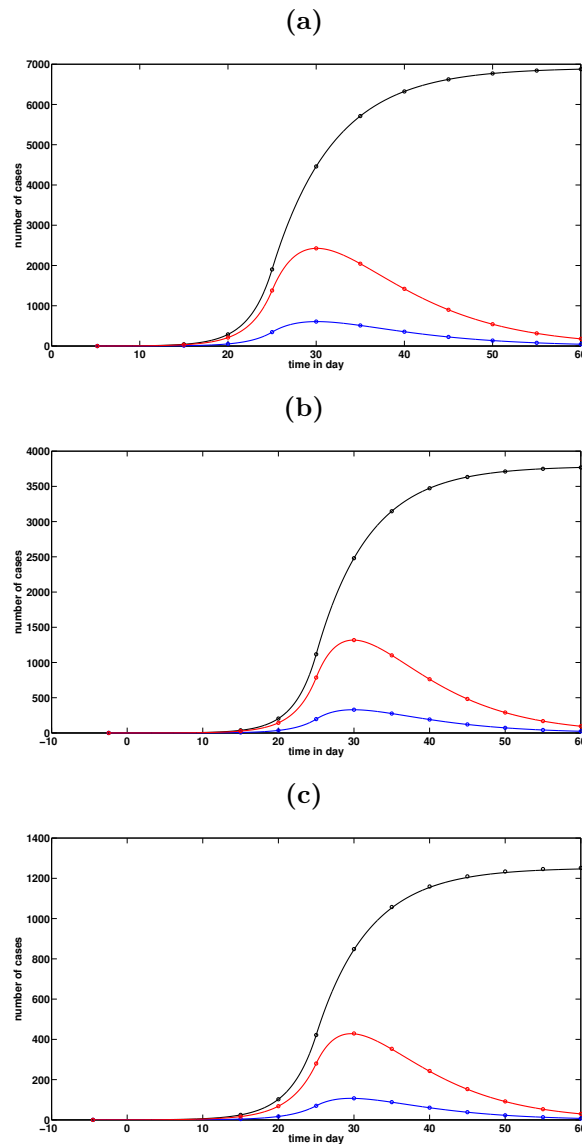


fig4

Figure 4: In this figure we plot the graphs of  $t \rightarrow CR(t)$  (black solid line),  $t \rightarrow U(t)$  (blue solid line) and  $t \rightarrow R(t)$  (red solid line). We use again  $f = 0.8$ ,  $\eta = 1/7$ ,  $\nu = 1/7$ , and  $S_0 = 11.081 \times 10^6$ . In Figure (a), we use  $\chi_1 = 0.16$ ,  $\chi_2 = 0.38$ ,  $\chi_3 = 1.1$ ,  $t_0 = 5.12$  for the parameter values for China which give  $\tau = 4.44 \times 10^{-08}$  for  $t \in [t_0, 25]$  and  $\tau = 0$  for  $t > 25$ ,  $I_0 = 3.62$ ,  $U_0 = 0.2$ . In Figure (b), we use  $\chi_1 = 0.23$ ,  $\chi_2 = 0.34$ ,  $\chi_3 = 0.1$  and  $t_0 = -2.45$ , for the parameter values obtained from the data for Hubei province, which give  $\tau = 4.11 \times 10^{-08}$  for  $t \in [t_0, 25]$  and  $\tau = 0$  for  $t > 25$ ,  $I_0 = 0.3$ ,  $U_0 = 0.02$ . In Figure (c), we use  $\chi_1 = 0.36$ ,  $\chi_2 = 0.28$ ,  $\chi_3 = 0.1$ , and  $t_0 = -4.52$  for the parameter values obtained from the data for Wuhan, which give  $\tau = 3.6 \times 10^{-08}$  for  $t \in [t_0, 25]$  and  $\tau = 0$  for  $t > 25$ ,  $I_0 = 0.25$ ,  $U_0 = 0.02$ . The cumulated number of reported cases goes up to 7000 in Figure (a), 4000 in Figure (b) and 1400 in Figure (c), and the turning point is around January 30 in Figures (a), (b) and (c).

## 5 Discussion

As a consequence of our study, we note that public health measures, such as isolation, quarantine, and public closings, greatly reduce the final size of this epidemic, and make the turning point much earlier than without these measures. With our method, we fix  $\eta$ ,  $\nu$  and  $f$  and get the same turning point for the three data sets in Table 2, Table 3 and Table 4. We choose  $f = 0.8$ , which means around 80% of cases are reported in the model, since cases are very well documented in China. Thus, we only assume that a small fraction, 20% were not reported. This assumption may be confirmed later on.

We also vary the parameters  $\eta$ ,  $\nu$  and  $f$ , and we do not observe a strong variation of the turning point.



Nevertheless, the number of reported case are very sensitive to the data sets, as shown in the figures. The formula (4.4) for  $CR(t)$  is very descriptive until January 26 for the reported case data for China and Hubei, but is not reasonable for Wuhan data. This suggests that the turning point is very robust, while the number of cases is very sensitive. We can find multiple values of  $\eta$ ,  $\nu$  and  $f$  which provide a good fit for the data. This means that  $\eta$ ,  $\nu$  and  $f$  should also be evaluated using other methods. The values  $1/\eta = 7$  days and  $1/\nu = 7$  days, are taken from information concerning earlier corona viruses, and are used now by medical authorities [2].

The predictive capability of models (4.1) and (4.10) requires valid estimates of the parameters  $f$  (fraction of asymptomatic infectious that become reported symptomatic infectious), the parameter  $1/\nu$  (average time asymptomatic infectious are asymptomatic), and the parameter  $1/\eta$  (average time symptomatic infectious remain infectious). In Figure 5, we graph  $R_0$  as a function of  $f$  and  $1/\nu$  for the data in Table 2, to illustrate the importance of these values in the evolution of the epidemic. The accuracy of these values depend on the input of medical and biological epidemiologists.

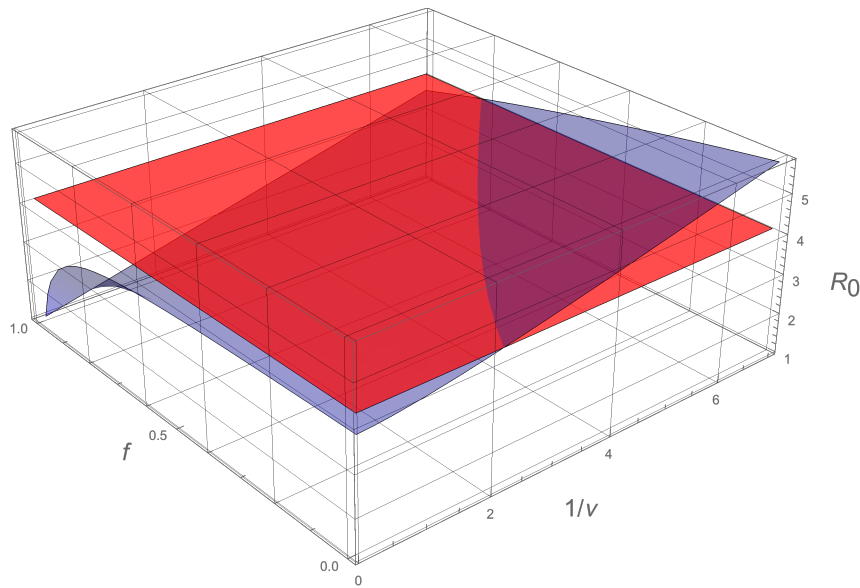


fig5

Figure 5: In this figure we use  $1/\eta = 7$  days and we plot the basic reproductive number  $R_0$  as a function of  $f$  and  $1/\nu$  using (4.9) with  $\chi_2 = 0.38$ , which corresponds to the data for China in Table 2. If both  $f$  and  $1/\nu$  are sufficiently small,  $R_0 < 1$ . The red plane is the value of  $R_0 = 4.13$ .

In influenza epidemics, the fraction  $f$  of reported cases may be significantly increased by public health reporting measures, with greater efforts to identify all current cases. Our model reveals the impact of an increase in this fraction  $f$  in the value of  $R_0$ , as evident in Figure 5 above, for the 2019-nCov epidemic in Wuhan.

## 6 Supplementary material

### 6.1 Method to estimate the parameters of (4.1) from the number of reported cases

From now on, we fix  $f, \nu, \eta$ .

**Step 1:** Since  $f$  and  $\nu$  we know that

$$\nu_1 = f\nu \text{ and } \nu_2 = (1 - f)\nu.$$

**Step 2:** By using equation (4.3) we obtain

$$CR'(t) = \nu_1 I(t) \Leftrightarrow \chi_1 \chi_2 \exp(\chi_2 t) = \nu_1 I(t) \quad (6.1) \quad \boxed{6.1}$$

and

$$\frac{\exp(\chi_2 t)}{\exp(\chi_2 t_0)} = \frac{I(t)}{I(t_0)},$$

and therefore

$$I(t) = I_0 \exp(\chi_2(t - t_0)). \quad (6.2) \quad \boxed{6.2}$$

Moreover by using (6.1) at  $t = t_0$

$$I_0 = \frac{\chi_1 \chi_2 \exp(\chi_2 t_0)}{f \nu} = \frac{\chi_3 \chi_2}{f \nu}. \quad (6.3) \quad \boxed{6.3}$$

**Step 3:** In order to evaluate the parameters of the model we replace  $S(t)$  by  $S_0 = 11.081 \times 10^6$  in the right-hand side of (4.1) (which is equivalent to neglecting the variation of susceptibles due to the epidemic, which is consistent with the fact that  $t \rightarrow CR(t)$  grows exponentially). Therefore, it remains to estimate  $\tau$  and  $\eta$  in the following system:

$$\begin{cases} I'(t) = \tau S_0 [I(t) + U(t)] - \nu I(t) \\ U'(t) = \nu_2 I(t) - \eta U(t). \end{cases} \quad (6.4) \quad \boxed{6.4}$$

By using the first equation we obtain

$$U(t) = \frac{1}{\tau S_0} [I'(t) + \nu I(t)] - I(t),$$

and therefore by using (6.2) we must have

$$I(t) = I_0 \exp(\chi_2(t - t_0)) \text{ and } U(t) = U_0 \exp(\chi_2(t - t_0)),$$

so by substituting these expressions into (6.4) we obtain

$$\begin{cases} \chi_2 I_0 = \tau S_0 [I_0 + U_0] - \nu I_0 \\ \chi_2 U_0 = \nu_2 I_0 - \eta U_0. \end{cases} \quad (6.5) \quad \boxed{6.5}$$

**Remark 6.1** Here we fix  $\tau$  in such a way that the value  $\chi_2$  becomes the dominant eigenvalue of the linearized equation (6.10) and  $(I_0, U_0)$  is the positive eigenvector associated to this dominant eigenvalue  $\chi_2$ . thus, we apply implicitly the Perron-Frobenius theorem. Moreover the exponentially growing solution  $(I(t), U(t))$  that we consider (which is starting very close to  $(0, 0)$ ) follows the direction of the positive eigenvector associated with the dominant eigenvalue  $\chi_2$ .

By dividing the first equation of (6.5) by  $I_0$  we obtain

$$\chi_2 = \tau S_0 \left[ 1 + \frac{U_0}{I_0} \right] - \nu$$

and hence

$$\frac{U_0}{I_0} = \frac{\chi_2 + \nu}{\tau S_0} - 1. \quad (6.6) \quad \boxed{6.6}$$

By using the second equation of (6.5) we obtain

$$\frac{U_0}{I_0} = \frac{\nu_2}{\eta + \chi_2}. \quad (6.7) \quad \boxed{6.7}$$

By using (6.6) and (6.7) we obtain

$$\tau = \frac{\chi_2 + \nu}{S_0} \frac{\eta + \chi_2}{\nu_2 + \eta + \chi_2}. \quad (6.8) \quad \boxed{6.8}$$

By using (6.7) we can compute

$$U_0 = \frac{\nu_2}{\eta + \chi_2} I_0 = \frac{(1 - f)\nu}{\eta + \chi_2} I_0. \quad (6.9) \quad \boxed{6.9}$$

## section4

## 6.2 Computation of the basic reproductive number $R_0$

In this section we apply results in Diekmann, Heesterbeek and Metz [7] and Van den Driessche and Watmough [8]. The linearized equation of the infectious part of the system is given by

$$\begin{cases} I'(t) = \tau S_0[I(t) + U(t)] - \nu I(t), \\ U'(t) = \nu_2 I(t) - \eta U(t). \end{cases} \quad (6.10) \quad \boxed{6.10}$$

The corresponding matrix is

$$A = \begin{bmatrix} \tau S_0 - \nu & \tau S_0 \\ \nu_2 & -\eta \end{bmatrix}$$

and the matrix  $A$  can be rewritten as

$$A = V - S$$

where

$$V = \begin{bmatrix} \tau S_0 & \tau S_0 \\ \nu_2 & 0 \end{bmatrix} \text{ and } S = \begin{bmatrix} \nu & 0 \\ 0 & \eta \end{bmatrix}.$$

Therefore, the next generation matrix is

$$VS^{-1} = \begin{bmatrix} \frac{\tau S_0}{\nu} & \frac{\tau S_0}{\eta} \\ \frac{\nu_2}{\nu} & 0 \end{bmatrix}$$

which is a Leslie matrix, and the basic reproductive number becomes

$$R_0 = \frac{\tau S_0}{\nu} \left( 1 + \frac{\nu_2}{\eta} \right). \quad (6.11) \quad \boxed{6.11}$$

By using (6.8) we obtain

$$R_0 = \frac{\chi_2 + \nu}{S_0} \frac{\eta + \chi_2}{\nu_2 + \eta + \chi_2} \frac{S_0}{\nu} \left( 1 + \frac{\nu_2}{\eta} \right)$$

and by using  $\nu_2 = (1 - f)\nu$  we obtain

$$R_0 = \frac{\chi_2 + \nu}{\nu} \frac{\eta + \chi_2}{(1 - f)\nu + \eta + \chi_2} \left( 1 + \frac{(1 - f)\nu}{\eta} \right). \quad (6.12) \quad \boxed{6.12}$$

## 7 Acknowledgments

We thank Arnaud Ducrot for discussion and technical assistance.

## 8 Author contributions

Z.L, O.S, P.M and G.W conceived and designed the study. P.M and O.S analyzed the data, carried out the analysis and performed numerical simulations, Z.L and G.W conducted the literature review. All authors participated in writing and reviewing of the manuscript.

**Keywords:** corona virus, reported and unreported cases, isolation, quarantine, public closings; epidemic mathematical model

## References

CDC

- [1] Chinese Center for Disease Control and Prevention. [http://www.chinacdc.cn/jkzt/crb/zl/szkb\\_11803/jszl\\_11809/](http://www.chinacdc.cn/jkzt/crb/zl/szkb_11803/jszl_11809/)

NEJM

- [2] New England Journal of Medicine, Letter to the Editor, DOI: 10.1056/NEJMc2001468, January 30, 2020.

- [3] Biao Tang, Xia Wang, Qian Li, Nicola Luigi Bragazzi, Sanyi Tang, Yanni Xiao, Jianhong Wu, Estimation of the transmission risk of 2019-nCov and its implication for public health interventions [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3525558](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3525558)
- [4] P. Magal and G. Webb, The parameter identification problem for SIR epidemic models: Identifying Unreported Cases, *J. Math. Biol.* (2018). <https://doi.org/10.1007/s00285-017-1203-9>
- [5] A. Ducrot, P. Magal, T. Nguyen, G. Webb. Identifying the Number of Unreported Cases in SIR Epidemic Models. *Mathematical Medicine and Biology*, (2019)
- [6] Wuhan Municipal Health Commission. <http://wjw.wuhan.gov.cn/front/web/list3rd/yes/802>
- [7] O. Diekmann, J. A. P. Heesterbeek and J. A. J. Metz, On the definition and the computation of the basic reproduction ratio  $R_0$  in models for infectious diseases in heterogeneous populations, *Journal of Mathematical Biology* **28**(4) (1990), 365-382.
- [8] P. Van den Driessche and J. Watmough, Reproduction numbers and subthreshold endemic equilibria for compartmental models of disease transmission, *Mathematical Biosciences* **180** (2002) 29-48.