

RESEARCH ARTICLE

Open Access

Estimating the daily trend in the size of the COVID-19 infected population in Wuhan



Qiu-Shi Lin¹, Tao-Jun Hu² and Xiao-Hua Zhou^{1,3,4*}

Abstract

Background: The outbreak of coronavirus disease 2019 (COVID-19) has become a pandemic causing global health problem. We provide estimates of the daily trend in the size of the epidemic in Wuhan based on detailed information of 10 940 confirmed cases outside Hubei province.

Methods: In this modelling study, we first estimate the epidemic size in Wuhan from 10 January to 5 April 2020 with a newly proposed model, based on the confirmed cases outside Hubei province that left Wuhan by 23 January 2020 retrieved from official websites of provincial and municipal health commissions. Since some confirmed cases have no information on whether they visited Wuhan before, we adjust for these missing values. We then calculate the reporting rate in Wuhan from 20 January to 5 April 2020. Finally, we estimate the date when the first infected case occurred in Wuhan.

Results: We estimate the number of cases that should be reported in Wuhan by 10 January 2020, as 3229 (95% confidence interval [CI]: 3139–3321) and 51 273 (95% CI: 49 844–52 734) by 5 April 2020. The reporting rate has grown rapidly from 1.5% (95% CI: 1.5–1.6%) on 20 January 2020, to 39.1% (95% CI: 38.0–40.2%) on 11 February 2020, and increased to 71.4% (95% CI: 69.4–73.4%) on 13 February 2020, and reaches 97.6% (95% CI: 94.8–100.3%) on 5 April 2020. The date of first infection is estimated as 30 November 2019.

Conclusions: In the early stage of COVID-19 outbreak, the testing capacity of Wuhan was insufficient. Clinical diagnosis could be a good complement to the method of confirmation at that time. The reporting rate is very close to 100% now and there are very few cases since 17 March 2020, which might suggest that Wuhan is able to accommodate all patients and the epidemic has been controlled.

Keywords: COVID-19, Wuhan, Daily, Trend, Size, Infection

Background

As of 5 April 2020, the National Health Commission (NHC) of China has confirmed a total of 81 708 cases of COVID-19 in the mainland of China, including 265 severe cases and 3331 deaths. An additional total of 88 suspected cases were reported. Wuhan has 50 008 confirmed cases. The NHC has also received 890 confirmed

reports in Hong Kong Special Administrative Region, 44 in Macau Special Administrative Region, and 363 in Taiwan [1]. More than one million cases have been detected outside China.

Despite the considerable medical resources and personnel that have been dispensed to combat COVID-19 in Hubei province, hospital capacity was overburdened in the early stage of this epidemic. There was a shortage of hospital beds needed to accommodate the rising number of COVID-19 patients. In response to this growing crisis, Wuhan transformed hotels, venues, training centers and college dorms into quarantine and treatment centers for COVID-19 patients. Further, 13 temporary treatment

* Correspondence: azhou@math.pku.edu.cn

¹Beijing International Center for Mathematical Research, Peking University, Beijing 100871, China

³Department of Biostatistics, School of Public Health, Peking University, Beijing 100191, China

Full list of author information is available at the end of the article



© The Author(s). 2020 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

centers were built to provide over 10 000 beds [2]. Therefore, a careful and precise understanding of the potential number of cases in Wuhan is crucial for the prevention and control of the COVID-19 outbreak. Wu et al. [3] provided an estimate of the total number of cases of COVID-19 in Wuhan, using the number of cases exported from Wuhan to cities outside the mainland of China. However, since the number of cases is small, their estimate of the size of the epidemic in Wuhan may not be precise and has large variability. Using the number of cases exported from Wuhan to all cities, including cities in China outside Hubei province, You et al. [4] proposed a method to estimate the total number of cases of COVID-19 in Wuhan. However, their method can only give an estimate of the cumulative number of cases until a certain date.

In this article, we propose a new statistical method to estimate daily number of cases in Wuhan under a similar dynamic equation model as the one in reference [3]. Unlike the one in reference [3], our method can also handle the missing information on whether a case is exported from Wuhan.

Methods

The spread of COVID-19 outside Hubei province is relatively controlled given the adequate medical resources. We use the reported number outside Hubei as it is a fairly accurate representation of the actual epidemic situation. In this modelling study, we first estimate the epidemic size in Wuhan from 10 January to 5 April 2020, based on the confirmed cases outside Hubei province that left Wuhan by 23 January 2020. Since some confirmed cases have no information on whether they visited Wuhan before, we adjust the number of imported cases after taking these missing values into account. We then calculate the reporting rate in Wuhan from 20 January to 5 April 2020. Finally, we estimate the date when the first patient was infected.

Data

Data retrieved from publicly available records from provincial and municipal health commissions in China and ministries of health in other countries include detailed information for 10 940 confirmed cases outside Hubei province. An additional table in the [Supplementary Materials](#) shows these websites in more detail [see [Data_source.xlsx](#)]. Information on confirmed cases including region, gender, age, date of symptom onset, date of confirmation, history of travel or residency in Wuhan, and date of departure from Wuhan. We display demographic characteristics of these patients in Table 1. Among the 7500 patients with gender data, 3509 (46.8%) are female. The mean age of patients is 44.48 and the median age is 44. The youngest confirmed patient outside Hubei province was only 5 days old while the oldest is 97 years old (see Table 1).

Table 1 Demographic characteristics of patients with COVID-19 outside Hubei province

Age group (years)	Female (<i>n</i> = 3509)	Male (<i>n</i> = 3991)	No information (<i>n</i> = 3440)
0–20	97 (3) ^a	149 (4)	2
20–39	1076 (33)	1348 (36)	41
40–59	1425 (43)	1598 (43)	38
60–79	630 (19)	578 (15)	39
≥ 80	66 (2)	60 (2)	7
No information	215	258	3313

^a Number (%). The percentages do not take missing data into account

We display the epidemiological data categorized by the date of confirmation in Table 2. An imported case means a patient that had been to Wuhan and was detected outside Hubei province. A local case means a confirmed case that had not been to Wuhan. Among the total of 10 940 cases, 6903 (63.1%) have such epidemiological information. The number of imported cases reached its peak on 29 January 2020, and the fourth column of Table 2 shows that the proportion of imported cases declines over time. This might reflect the effect of containment measures taken in Hubei province to control the COVID-19 outbreak [5]. Meanwhile, the daily counts of local cases are over 300 from 2 February to 7 February 2020, which indicate that infections among local residents should be a major concern for authorities outside Hubei province.

The last column of Table 2 lists the mean time from symptom onset to confirmation for patients confirmed on each day. The median duration of all cases is 5 days, and the mean is 5.54 days. In general, the detection period decreased in the first week after 20 January 2020, but increased since then. The improvements in detection speed and capacity might cause the initial decline, and the rise may be due to more thorough screening, leading to the detection of patients with mild symptoms who would otherwise not go to the hospitals [6].

Assumptions

The proposed method relies on the following assumptions:

- 1) Between 10 January and 23 January 2020, the average daily proportion of departing from Wuhan is p .
- 2) There is a $d = d_1 + d_2$ -day window between infection and detection, including a d_1 -day incubation period and a d_2 -day delay from symptom onset to detection.
- 3) Patients are not able to travel d days after infection.

Table 2 Patient data categorized by the date of confirmation

Date of confirmation	Imported cases (<i>n</i> = 3039)	Local Cases (<i>n</i> = 3864)	Proportion ^a	No information (<i>n</i> = 4037)	Total cases (<i>n</i> = 10 940)	Onset to detection (days)
≤ 20 Jan ^b	23	1	96%	2	26	8.83
21 Jan	35	3	92%	7	45	8.60
22 Jan	55	6	90%	16	77	6.11
23 Jan	117	15	89%	52	184	5.37
24 Jan	165	18	90%	92	275	4.56
25 Jan	210	38	85%	110	358	4.15
26 Jan	198	49	80%	163	410	4.05
27 Jan	190	85	69%	208	483	5.00
28 Jan	283	104	73%	252	639	5.58
29 Jan	299	144	67%	280	723	5.43
30 Jan	291	214	58%	260	765	5.13
31 Jan	256	239	52%	308	803	5.05
1 Feb	160	252	39%	266	678	4.83
2 Feb	159	310	34%	269	738	5.55
3 Feb	164	410	29%	322	896	5.76
4 Feb	119	336	26%	294	749	5.99
5 Feb	107	376	22%	266	749	6.10
6 Feb	83	380	18%	258	721	5.92
7 Feb	53	363	13%	201	617	6.17
8 Feb	42	286	13%	202	530	6.49
9 Feb	30	235	11%	209	474	6.66

^a The proportion is the number of imported cases divided by the sum of imported and local cases

^b The count and average on the first row are taken over all cases confirmed by January 20, 2020

- 4) The proportion of imported cases in the patients with no information is the same as the observed proportion on each day.
- 5) Trip durations are long enough that a traveling patient infected in Wuhan will develop symptoms and be detected in other places rather than after returning to Wuhan.
- 6) All travelers leaving Wuhan, including transfer passengers, have the same risk of infection as local residents.
- 7) Traveling is independent of the exposure risk to COVID-19 or of infection status.
- 8) Recoveries are not considered in this method.

Assumptions 1–4 are used explicitly in the [Methods](#) section. They are fundamental assumptions for our statistical model. Other assumptions might also affect the result of our model, and we make some remarks about our assumptions.

- a) 10 January 2020 is the start of Chinese New Year travel rush, and 23 January 2020, is the date of Wuhan lockdown [5]. In the total of 10 940 cases, only 131 (1.2%) cases' date of departure from

Wuhan are not in this period. They are excluded from our analysis.

- b) If the true average daily proportion of leaving Wuhan is larger than the assumed *p*, this violation of Assumption 1 could lead to overestimation of the number of cases in Wuhan.
- c) If the average time from infection to detection is longer than the assumed *d* days, this violation of Assumption 2 would lead to an overestimation.
- d) If travelers have a lower risk of infection than residents in Wuhan, this violation of Assumption 6 would cause an underestimation.
- e) If infected individuals are less likely to travel due to the health conditions, this violation of Assumption 7 would cause an underestimation.

In the [Supplementary Appendix A](#), we perform the sensitivity analysis on the effect of some of the violations on our results.

Notations

Let Day *t*₀ denote the date of infection for the very first case. Let *N_t* be the cumulative number of cases that

should be confirmed in Wuhan by Day t . Other notations of our model are defined in Table 3.

The numbers T_t , I_t , and L_t are the observed data used in our model, t_c , r , and K are the parameters that determine how N_t changes over time.

Model

The growth trend of the size N_t of infected population is determined by the following ordinary differential equation:

$$\frac{dN_t}{dt} = \frac{r}{K} N_t (K - N_t), \quad r > 0, K > 0, \tag{1}$$

where K is the size of the population that are susceptible to COVID-19 in Wuhan, and r is a constant that controls the growth rate of N_t . This is the modified version of the famous SIR model [3, 10] in epidemiology. In the equation (1), the growth rate of N_t is proportional to the product of N_t and the number $K - N_t$ of people that are susceptible but not infected yet. It is a reasonable model for the epidemic transmission. At the beginning of this epidemic, when N_t is small, people have little knowledge of COVID-19, N_t grows at an exponential rate r . As N_t becomes larger, containment measures are taken to control it, the growth rate of N_t slows down, resulting in a sigmoid curve of N_t . Detailed explanations of the model (1) are given in the [Supplementary Appendix B](#). The model (1) has an analytical solution,

$$N_t = \frac{K}{1 + e^{-r(t-t_c)}} = K f_t, \tag{2}$$

where $f_t = \frac{1}{1 + e^{-r(t-t_c)}}$, and the derivative $\frac{dN_t}{dt}$ is maximized at $t = t_c$, $\frac{r}{2} = \frac{d \log N_t}{dt}$ is the growth rate of $\log N_t$ at time t_c , K is a parameter to be estimated.

Estimation

We use data on the confirmed cases who left Wuhan between 10 January and 23 January 2020, to estimate K . Under Assumption 2, cases infected on Day t will be detected on Day $t + d$, so the number of infected cases in Wuhan is N_{t+d} on Day t . If $t_0 \leq t \leq t_0 + d$, there should be no confirmed cases. If $t_0 + d < t \leq t_0 + 2d$, imported cases on Day t are infected in Wuhan on Day $t - d$. There are N_t infected cases in Wuhan on Day $t - d$, hence the number of imported cases x_t on Day t follows a binomial (N_t, p) distribution, where p is the assumed average daily probability of leaving Wuhan between 10 January and 23 January 2020. If $t > t_0 + 2d$, under Assumption 3, N_{t-d} patients are not able to travel, x_t has a binomial $(N_t - N_{t-d}, p)$ distribution. Let X_t be the cumulative number of imported cases by Day t , then

$$X_t = \sum_{k=1}^t x_k \sim \text{Binomial} \left(\sum_{k=t-d+1}^t N_k, p \right) + 2d, \quad t \geq t_0 \tag{3}$$

From equations (2) and (3), $X_t \sim \text{Binomial} \left(K \sum_{k=t-d+1}^t f_k, p \right)$. The parameter estimate \hat{K} is derived by maximizing the likelihood function

$$l(K) = \binom{K \sum_{k=t-d+1}^t f_k}{X_t} p^{X_t} (1-p)^{K \sum_{k=t-d+1}^t f_k - X_t}. \tag{4}$$

The lower and upper bound of the 95% confidence interval $[\hat{K}_l, \hat{K}_u]$ are values such that the cumulative distribution function $F(K) = \sum_{x=0}^{X_t} l(K)$ equals to 0.975 and 0.025, respectively. The reporting rate is the

Table 3 Notations for our model

Notation	Meaning	State	Reference
T_t	The reported number of confirmed cases outside Hubei province on Day t	Known	Official websites
I_t	The observed number of imported cases on Day t	Known	Official websites
L_t	The observed number of local cases on Day t	Known	Official websites
U_t	The number of cases with no information on Day t	Known	$U_t = T_t - I_t - L_t$
x_t	The number of imported cases on Day t	Estimated	
X_t	The cumulative number of imported cases by Day t	Estimated	$X_t = \sum_{k=1}^t x_k$
N_t	The cumulative number of cases that should be confirmed in Wuhan by Day t	Estimated	
t_0	The date of infection for the very first case	Estimated	
p	The daily probability of departing from Wuhan	Known	[7, 8]
d	The time from infection to detection	Known	[9]
t_c	The time when N_t increases at the fastest rate	Known	From our data
r	A parameter that determines the growth rate of N_t	Known	[3]
K	The size of the population that are susceptible to COVID-19 in Wuhan.	Estimated	

reported cumulative number of cases in Wuhan on Day t divided by our estimated number \widehat{N}_t . The estimate of the date t_0 of first infection is obtained by solving the equation $N_{t_0} = 1$.

Determining the number of imported cases x_t plays a crucial role in the modeling procedure. Note that not all cases have clear records on the history of travel or residency in Wuhan, we need to impute the missing values. Under Assumption 4, the proportion of imported cases in the U_t patients with no information is the same as the observed proportion $\frac{I_k}{I_k + L_k}$. Therefore,

$$x_t = I_t + U_t \times \frac{I_k}{I_k + L_k} = T_k \times \frac{I_k}{I_k + L_k}. \tag{5}$$

The average daily proportion of leaving Wuhan between 10 January and 23 January 2020 is estimated to be the ratio of daily volume of travelers to the population of Wuhan (14 million). More than 5 million people were estimated to leave Wuhan due to the Spring Festival and epidemic [7]. This number is mentioned by Wuhan Mayor in a press conference. We assume these passengers left Wuhan between the start of Chinese New Year travel rush on 10 January 2020, and the lockdown of Wuhan city on 23 January 2020. During the travel rush, 34% of the passengers traveled across 300 km [8]. Major

cities outside Hubei province are generally over 300 km from Wuhan. This would imply, on average, the daily probability p of traveling from Wuhan to places outside Hubei province would be $5 \times 0.34 / 14 / 14 = 0.009$. Li et al. estimated that the mean incubation period of 425 patients with COVID-19 was 5.2 days (95% CI: 4.1–7.0) [9]. The mean time from symptom onset to detection calculated from our data is 5.54 days, so we choose $d = d_1 + d_2 = 11$ days. On 29 January 2020, there was the maximum count of imported cases. Since x_t has a binomial $(N_t - N_{t-d}, p)$ distribution with constant p , $N_t - N_{t-d}$ also reaches its maximum at $t = 29$ January 2020. From the logistic function (2), t_c is the midpoint of t and $t - d$, that is $t - \frac{d}{2} = 24$ January 2020, which is shortly after the lockdown of Wuhan city [5]. Wu et al. estimated the epidemic doubling time as 6.4 days (95% CI: 5.8–7.1) as of 25 January 2020 [3]. From this result, we estimate that $\frac{r}{2} = \frac{d \log N_{t_c}}{dt} = \frac{\ln 2}{6.4} = 0.1$. Using these values for parameters p , d , t_c , and r , we can derive the maximum likelihood estimate $\hat{K} = 51\,273$, with 95% CI: 49\,844–52\,734.

Results

We estimate the number of cases that should be reported in Wuhan by 10 January 2020, as 3229 (95% CI: 3139–3321) and 51 273 (95% CI: 49 844–52 734) by 5

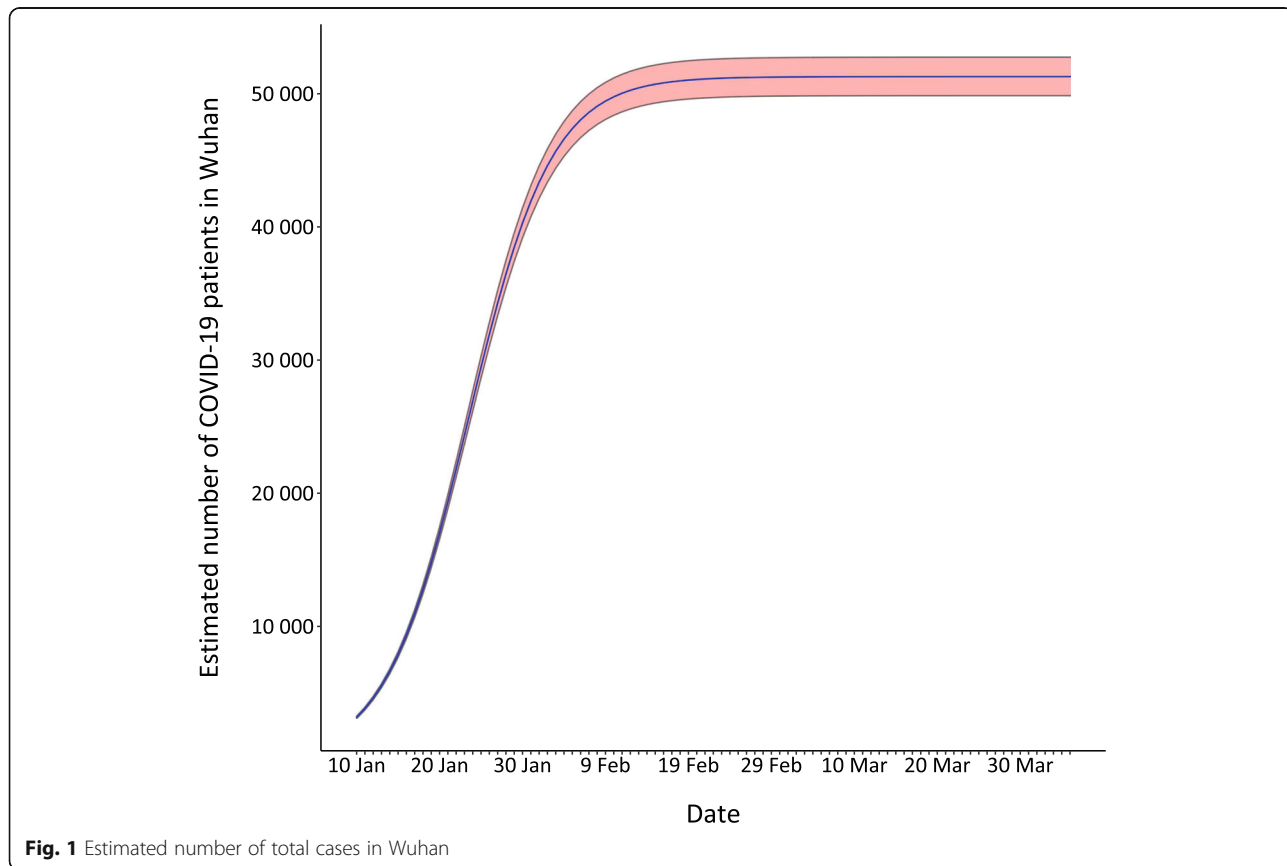


Fig. 1 Estimated number of total cases in Wuhan

April 2020. Figure 1 shows how the estimated number of cases in Wuhan increases over time, together with the 95% confidence bands.

As shown in Fig. 2, the reporting rate has grown rapidly from 1.5% (95% CI: 1.5–1.6%) on 20 January 2020 to 39.1% (95% CI: 38.0–40.2%) on 11 February 2020. It becomes 71.4% (95% CI: 69.4–73.4%) on 13 February 2020, and reaches 97.5% (95% CI: 94.8–100.3%) on 5 April 2020.

Table 4 gives the number of confirmed cases reported by Wuhan Health Commission, the estimated number and the reporting rate, as well as the 95% confidence intervals. By solving for t in the equation $N_t = 1$ with the expression of N_t given in (2), we obtain an estimate of the date of first infection as 30 November 2019.

Discussion

Most studies estimating the epidemic size of COVID-19 in Wuhan use the reported number of cases to predict the future trend. These researches ignore the possibility of considerable number of unreported cases in the early stage of this outbreak in Wuhan. We estimate the actual size of epidemic in Wuhan and predict the future trend based on information about COVID-19 cases outside Hubei province. Several recent studies share similar ideas that utilize external data to infer the number of

cases in Wuhan. You et al. [4] estimated a total of 3933 cases of COVID-19 in Wuhan (95% CI: 3454–4450) that had an onset of symptoms by 19 January 2020. Wu et al. [3] estimated that 75 815 individuals (95% CI: 37 304–130 330) have been infected in Wuhan as of 25 January 2020. This number far exceeds 50 008 cumulative cases reported in Wuhan, which seems not very reasonable. Nishiura et al. [11] estimated a total of 20 767 infected individuals as of 29 January 2020 based on a binomial model, which is simplified version of model (3), and eight confirmed cases on three chartered flights evacuating Japanese citizens from Wuhan. These results are estimates of the cumulative number of cases in Wuhan until a certain date and have wide confidence intervals due to limited data size. Using information of over 10 000 confirmed cases outside Hubei province, our statistical method can handle the problem of missing data and estimate the daily number of cases in Wuhan as shown in Fig. 1. Maugeri et al. [12] estimate a total of 8724 (95% CI: 8478–8921) infected cases and 92.9% (95% CI: 92.5–93.1%) unreported by 23 January 2020 with a proposed SEIRD model based on the reported number of deaths between 23 January and 9 February 2020. However, a total of 1290 cases were added to the death toll in Wuhan on 18 April 2020 by Wuhan government [13]. Thus, the number of deaths used in their

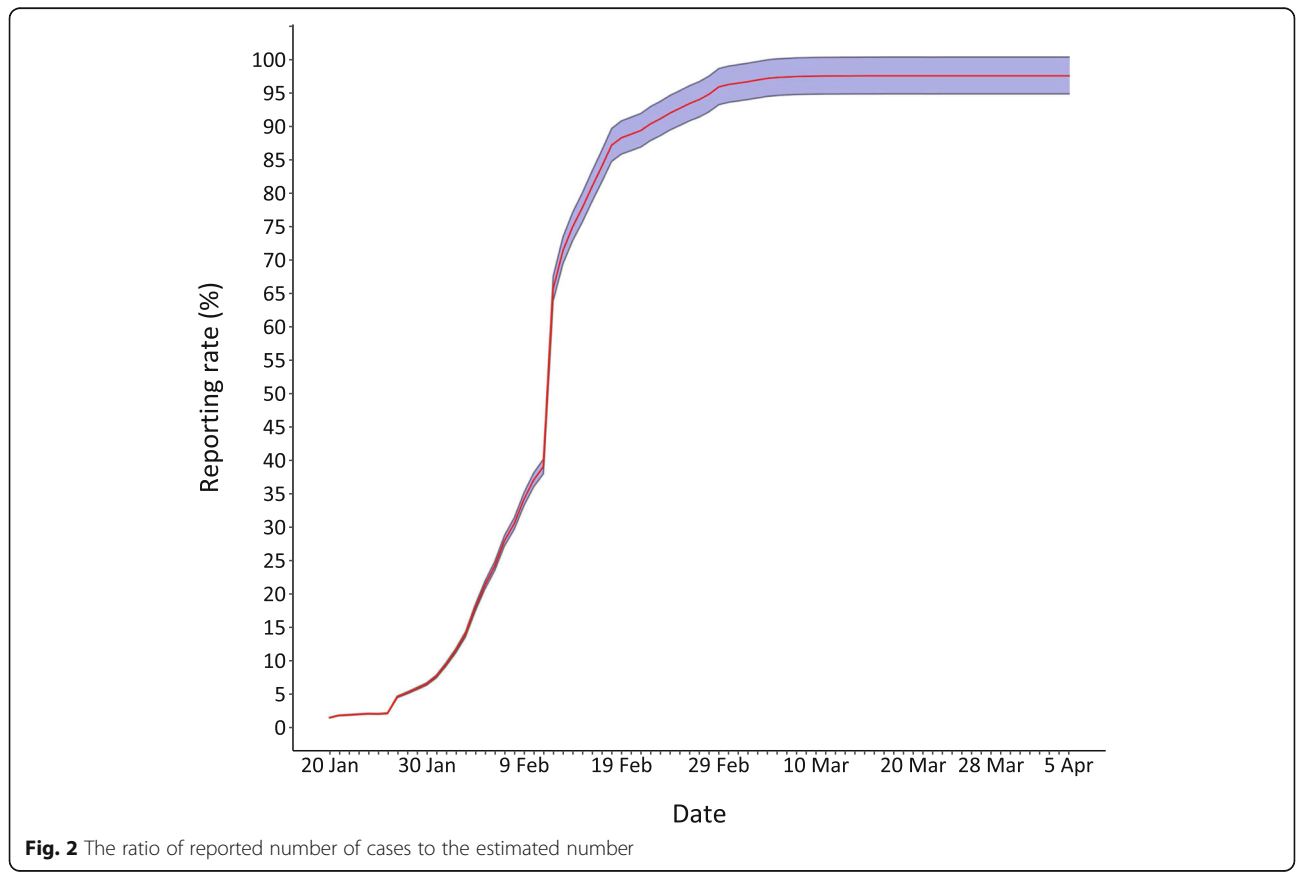


Fig. 2 The ratio of reported number of cases to the estimated number

Table 4 Model estimated number of cases and reporting rates

Date	Reported	Estimated number (95% CI)	Reporting rate (95% CI)
20 Jan	258	17 013 (16 539–17 498)	1.5% (1.5–1.6%)
25 Jan	618	29 453 (28 633–30 293)	2.1% (2.0–2.2%)
30 Jan	2639	40 292 (39 169–41 440)	6.6% (6.4–6.7%)
4 Feb	8351	46 601 (45 302–47 929)	17.9% (17.4–18.4%)
9 Feb	16 902	49 449 (48 071–50 858)	34.2% (33.2–35.2%)
11 Feb	19 558	50 036 (48 641–51 462)	39.1% (38.0–40.2%)
13 Feb	35 991	50 437 (49 031–51 874)	71.4% (69.4–73.4%)
18 Feb	44 412	50 962 (49 542–52 414)	87.2% (84.7–89.7%)
23 Feb	46 607	51 158 (49 732–52 616)	91.1% (88.6–93.7%)
28 Feb	48 557	51 231 (49 803–52 691)	94.8% (92.2–97.5%)
4 Mar	49 671	51 257 (49 829–52 718)	96.9% (94.2–99.7%)
9 Mar	49 965	51 267 (49 838–52 728)	97.5% (94.8–100.3%)
17 Mar	50 005	51 272 (49 843–52 733)	97.5% (94.8–100.3%)
25 Mar	50 006	51 273 (49 844–52 734)	97.5% (94.8–100.3%)
5 Apr	50 008	51 273 (49 844–52 734)	97.5% (94.8–100.3%)

research might not be accurate enough, leading to biases in their estimation. In the early stage of this epidemic, estimated numbers given by our method and existing researches are substantially larger than the reported number of confirmed cases. As of 5 April 2020, the reported cumulative number of cases in Wuhan is very close to the estimated number of our model, indicating the effectiveness of our method for long-term epidemic trend prediction. This method can effectively and accurately estimate the actual number of cases when the testing capability is insufficient. Similar statistical methods and ideas can be applied to other countries or regions that are still suffering from the outbreak of COVID-19 to support the prevention and control of this pandemic.

The major limitation of our methodology, as well as many other existing researches, is that time-varying parameters are not taken into consideration. Assumption 1 assumes that the daily probability of leaving Wuhan between 10 January and 23 January 2020, is approximately constant. Our estimate of traveling probability p might not be accurate due to the missing of exact daily number of traveling people from Wuhan to places outside Hubei province. We will try to improve the accuracy of p with more credible and precise transportation data in future research. Quarantine measures may have influences on some parameters in the epidemiological dynamic model (1), so that these parameters may change over time. It is a future research topic to allow time-varying parameters.

Conclusions

We provide a computationally efficient method of estimating the daily development of COVID-19 epidemic in Wuhan. The date of first infection is estimated as 30

November 2019. With the introduction of clinical diagnosis in the confirmation of COVID-19 in Wuhan, the reporting rate increases rapidly from about 40% to over 70% in only 2 days in February 2020. Clinical diagnosis could be a good complement to the method of confirmation in the early stage. The suspected cases in Wuhan declined to zero on 17 March 2020. Both the reported and estimated numbers show that there are very few cases since then. This might suggest the epidemic in Wuhan has been under control. The reporting rate is always increasing during this epidemic. As of 5 April 2020, the reporting rate is very close to 100%. Although the medical resources and testing capacity of Wuhan were insufficient at the beginning of this outbreak, Wuhan is now able to accommodate all patients with the assistance from the whole country and effective measures taken in the fight against COVID-19.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1186/s40249-020-00693-4>.

Additional file 1. The first file's name is Data_source.xlsx. It is a table that lists the official COVID-19 websites of provincial and municipal health commissions for every city in China. It includes province, city and source websites in both Chinese and English. Information of COVID-19 patients are collected from these websites.

Additional file 2. The second file's name is Appendix.docx. It contains two appendices with additional sensitivity analysis and detailed description of our method. They are only for reviewers' convenience, not for publication.

Acknowledgements

None.

Authors' contributions

QL and XZ designed the study. QL and TH collected and analyzed the data. QL, TH, and XZ interpreted the results and wrote the manuscript. The author(s) read and approved the final manuscript.

Funding

This research is supported by National Natural Science Foundation of China (Grant No. 82041023) and Zhejiang University special scientific research fund for COVID-19 prevention and control (Grant No. 2020XGX016).

Availability of data and materials

All data and materials used in this work were publicly available.

Ethics approval and consent to participate

The ethics approval and individual consent was not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Beijing International Center for Mathematical Research, Peking University, Beijing 100871, China. ²School of Mathematical Sciences, Peking University, Beijing 100871, China. ³Department of Biostatistics, School of Public Health, Peking University, Beijing 100191, China. ⁴Center for Statistical Science, Peking University, Beijing 100871, China.

Received: 7 March 2020 Accepted: 10 June 2020
Published online: 18 June 2020

References

1. National Health Commission Update on April 06, 2020. China CDC Weekly. <http://weekly.chinacdc.cn/news/TrackingtheEpidemic.htm#NHCApr06>. Accessed 6 Apr 2020.
2. Wang XD. Hubei ordered to admit all patients in hospitals. China Daily <https://www.chinadaily.com.cn/a/202002/09/WS5e3fba1ca3101282172760aa.html>. Accessed 9 Feb 2020.
3. Wu JT, Leung K, Leung GM. Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: a modelling study. *Lancet*. 2020;395(10225):689–97.
4. You C, Lin Q, Zhou X. An estimation of the total number of cases of NCIP (2019-nCoV)—Wuhan, Hubei Province, 2019–2020. *China CDC Weekly*. 2020; 2(6):87–91.
5. China declares lockdown in Wuhan on Thursday due to coronavirus outbreak. Tass. <https://tass.com/world/1111981>. Accessed 23 Jan 2020.
6. Xin W. Beijing to set up checkpoints in all residential communities. China Daily. <https://www.chinadaily.com.cn/a/202002/10/WS5e415cb1a3101282172766c4.html>. Accessed 1 Feb 2020.
7. 5 million-plus leave Wuhan: Mayor. China Daily. <https://www.chinadaily.com.cn/a/202001/27/WS5e2dcd01a310128217273551.html>. Accessed 27 Jan 2020.
8. Big data perspective: Wuhan in the Chinese New Year travel rush. Daily Economic News. <https://m.nbd.com.cn/articles/2020-01-22/1402239.html>. Accessed 22 Jan 2020. (In Chinese).
9. Li Q, Guan X, Wu P, Wang X, Zhou L, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N Engl J Med*. 2020;382(13):1199–207.
10. Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics. *Pro R Soc Lond A*. 1927;115(772):700–21.
11. Nishiura H, Kobayashi T, Yang Y, Hayashi K, Miyama T, Kinoshita R, et al. The rate of underascertainment of novel coronavirus (2019-nCoV) infection: estimation using Japanese passengers data on evacuation flights. *J Clin Med*. 2020;9(2):419.
12. Maugeri A, Barchitta M, Battiato S, Agodi A. Estimation of unreported novel coronavirus (SARS-CoV-2) infections from reported deaths: a susceptible-exposed-infectious-recovered-dead model. *J Clin Med*. 2020;9(5):E1350.
13. Cui J. Death toll in Wuhan revised in evaluation. China Daily. <https://www.chinadaily.com.cn/a/202004/18/WS5e9a3742a3105d50a3d17158.html>. Accessed 26 May, 2020.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

