Extended SIR prediction of the epidemics trend of COVID-19 in Italy and compared with Hunan, China

Jia Wangping 1,2*, Han Ke 1*, Song Yang 1, Cao Wenzhe 1, Wang Shengshu 1, Yang Shanshan 1, Wang Jianwei 1, Kou Fuyin 1, Tai Penggang 1, Li Jing 1, Liu Miao 1, He Yao 1

1Institute of Geriatrics, Beijing Key Laboratory of Aging and Geriatrics, National clinical research center for geriatrics diseases, Second Medical Center of Chinese PLA General Hospital, 28 Fuxing Road, Beijing, 100853, China;

2Department of Military Medical Technology Support, School of Non-commissioned Officer, Army Medical University, Shijiazhuang, Hebei, China.

* Correspondence:

Prof. He Yao, MD, PhD
E-mail: yhe301@x263.net

Liu Miao, MD, PhD
Email: liumiaolmbxb@163.com

Keywords: COVID-19, coronavirus, Italy, prediction, epidemics trend.

Abstract

Background: Coronavirus Disease 2019 (COVID-19) is currently a global public health threat. Outside of China, Italy is one of the most suffering countries with the COVID-19 epidemic. It is important to predict the epidemics trend of COVID-19 epidemic in Italy to help develop public health strategies.

Methods: We used time-series data of COVID-19 from Jan 22, 2020 to Mar 16, 2020. An infectious disease dynamic extended susceptible-infected-removed (eSIR) model, which covers the effects of different intervention measures in dissimilar periods, was applied to estimate the epidemic trend in Italy. The basic reproductive number was estimated using Markov Chain Monte Carlo methods and presented using the resulting posterior mean and 95% credible interval (CI). Hunan, with similar total number of populations in Italy, was used as a comparative item.

Results: In the eSIR model, we estimated that the basic reproductive number for COVID-19 was respectively 4.10 (95% CI: 2.15–6.77) in Italy and 3.15 (95% CI: 1.71–5.21) in Hunan. There would be totally 30,086 infected cases (95% CI: 7,920–81,869) under the current country blockade and the endpoint would be Apr 25 (95% CI: Mar 30 to Aug 07) in Italy. If the country blockade is imposed 5 day later, the total number of infected cases would expand the infection scale 1.50 times.

Conclusion: Italy's current strict measures can efficaciously prevent the further spread of COVID-19 and should be maintained. Necessary strict public health measures be implemented as soon as
possible in other European countries with a high number of COVID-19 cases. The most effective strategy needs to be confirmed in further studies.

1 Introduction

The Corona Virus Disease 2019 (COVID-19) started in Wuhan, China in December and quickly spread to China and many countries and regions in the world (1-3). The COVID-19 outbreak made assessment as a pandemic by the World Health Organization (WHO) on March 11. It is currently a global public health threat and more than 100 countries including Italy, Iran, the United States, South Korea, and Japan are suffering from COVID-19. Outside of China, Italy is one of the most suffering countries with the COVID-19 epidemic. As of Mar 16, the cumulative number of confirmed cases in Italy reached 27980, ranking second in the world, the total confirmed deaths, 2158, and the fatality rate, as high as 7.71%, which has become one of the highest among the major epidemic countries. However, few studies have assessed the epidemic status in Italy (4, 5).

Global public health measures are required to cope with the rapid spread of the epidemic. China has taken precise and differentiated strategies, including self-quarantine of residents in Wuhan and other areas, community-based prevention and control. These measures have played an important role in preventing and controlling the epidemic. Previous studies have shown that due to the isolation of Wuhan, the overall epidemiological progress in mainland China has been delayed by three to five days and the number of internationally transmitted cases has been reduced by nearly 80% (6). Italy detected the first two cases of imported COVID-19 on Jan 31. After that, Italy was the first country to declare a state of emergency. Since then, various measures have been implemented to control the spread of COVID-19. It is vital to evaluate the role of Italian quarantine measures for decision-making.

Mathematical modeling is helpful to predict the possibility and severity of disease outbreak and provide key information for determining the type and intensity of disease intervention. The SIR model and its modifications such as SEIR model have been widely applied to the current outbreak of COVID-19. Tang et al. estimated the infectivity of COVID-19 based on a classical susceptible-exposed-infected-removed (SEIR) epidemiological model (7). Wu et al proposed an extended SEIR model to forecast the spread of 2019-nCoV both within and outside of mainland China (3). However, these studies assumed that the exposed population were not infectious, which may be not suitable in COVID-19. Yang Z et al. predicted that China's epidemic will peak in late February and end in late April by a combination of SEIR model and a machine-learning artificial intelligence (AI) approach (8). However, this study and the above studies did not consider the phase-adjusted preventive measures and time-varying parameters, which may affect the accuracy of predictions.

We adopted extended susceptible-infected-removed (eSIR) model (9), which covers the effects of different epidemic prevention measures in different periods and helps to achieve the following specific objectives:

AIM 1: Compare the epidemic development of COVID-19 in Italy with provinces with similar total number of populations in China.

AIM 2: Predict the epidemiological trend of COVID-19 in Italy via a modified and calibrated model.

AIM 3: Assess the effectiveness of Italian existing and assumed precautionary quarantine measures.

2 Methods
### 2.1 Data sources

In this study, we used the publicly available dataset of COVID-19 provided by the Johns Hopkins University (10). This dataset includes many countries’ daily count of confirmed cases, recovered cases and deaths. As time-series data, it is available from 22 January 2020. Besides, we also gathered and cross-checked data in DXY.cn (11), a website providing real-time data of the COVID-19. These data are collected through public health authorities’ announcements and are directly reported public and unidentified patient data, so ethical approval is not required.

### 2.2 Prediction models

The reproduction number, $R_0$, reflects the transmissibility of a virus spreading under no control, representing the average number of new infections generated by each infected person (12). The COVID-19 is likely to decline and eventually disappear if $R_0 \leq 1$. To estimate trends and calculate the $R_0$, we used an extended SIR model (eSIR model) with a time-varying transmission rate (9). In brief, this model is one extension to the basic epidemiological model——susceptible-infected-removed (SIR) model. $S$, $I$ and $R$ represent the number of susceptible, infected, and recovered individuals accordingly, and $N = S + I + R$ is the total population. The transmission rate is constant in the SIR model. It should be noted that in actual situations, the speed of transmission can be changed through many interventions, such as personal protective measures, community-level isolation and city blockade. As is shown below, the eSIR model adds transmission modifier $\pi(t)$ to the SIR model, so it allows a time-varying probability of the transmission rate.

An extended SIR model with time-varying transmission rates.

$$\frac{dS}{dt} = -\pi(t) \frac{\beta SI}{N}$$

$$\frac{dS}{dt} = \pi(t) \frac{\beta SI}{N} - \gamma I$$

$$\frac{dR}{dt} = \gamma I$$

$$R_0 = \frac{\beta}{r}$$

The transmission rate modifier $\pi(t)$ can be specified according to actual interventions in different times and regions. According to Chinese government isolation measures and previous study, we set $\pi(t) = 0.9$ if $t \in$ (Jan 23, Feb 04], city blockade; $\pi(t) = 0.5$ if $t \in$ (Feb 4, Feb 8], enhanced quarantine;
\( \pi(t) = 0.1 \) if \( t > \text{Feb 8} \), more enhanced quarantine in Hunan. In the opinion of Italy government isolation measures, we set \( \pi(t) = 0.9 \) if \( t < \text{Mar 4} \), no concrete quarantine protocols; \( \pi(t) = 0.5 \) if \( t \in (\text{Mar 04, Mar 09}) \), some cities blockade and enhanced quarantine; \( \pi(t) = 0.1 \) if \( t > \text{Mar 09} \), country blockade and more enhanced quarantine in Italy. We also assumed that the Italy government took country blockade and intensified quarantine on Mar 05 or Mar 15.

We did prediction with an R software package—eSIR which can output the Markov Chain Monte Carlo (MCMC) estimation, inference, and prediction under the extended SIR models with time-varying transmission modifier \( \pi(t) \). The model can also yield the turning points of the epidemiological trend of COVID-19. The first turning point was defined as the mean predicted time when the daily proportion of infected cases becomes smaller than the previous ones. The second turning point was defined as the mean the predicted time when the daily proportion of removed cases (i.e. both recovered and dead) becomes larger than that of infected cases. Besides, an end point was defined as the time when the median proportion of current infected cases turn to zero.

We did all analyses in R (version 3.6.2).

3 Results

3.1 Epidemic development of COVID-19 in Italy compared with Hunan

Figure 1 demonstrates daily new COVID-19 cases and epidemic distribution of COVID-19 in Hunan, China and Italy. The number of new cases and confirmed cases show an exponential trend since Feb 21 in Italy from while the number of new cases turns to zero from Feb 29 in Hunan.

3.2 Prediction of the epidemics trend of COVID-19 in Italy compared with Hunan

Table 1 summarizes the posterior mean and credible intervals of R0, incidence, and endpoint in Hunan and Italy according to the eSIR model. R0 for COVID-19 was respectively 3.15 (95% CI: 1.71–5.21) in Hunan and 4.10 (95% CI: 2.15–6.77) in Italy. There would be totally 3 988 infected cases (95%CI: 786–9 065) in Hunan. There would be totally 30 086 infected cases (95%CI: 7 920–81 869) under the current country blockade in Italy. Figure 2 indicates an epidemiological trend of COVID-19 under existing preventions in Hunan, China and Italy. The first and second turning point in Hunan appeared on Feb 04 and Feb 09. The first and second turning point in Italy is Mar 10 and Mar 11. The predictions suggest that the endpoints of the COVID-19 epidemics in Hunan and Italy will come on Feb 23 (95%CI: Feb 19 to Mar 23) and Apr 25 (95%CI: Mar 30 to Aug 07), separately.

3.3 Prediction of the epidemics trend of COVID-19 in Italy assumed precautionary quarantine measures.

Table 2 summarizes the posterior mean and credible intervals of R0, incidence, and endpoint according to different beginning of country blockade in Italy. The mean number of COVID-19 infected cases in Italy would be 10636 (95%CI: 2357–23326) and the endpoint would arrive on Mar 16 (95%CI: Mar 09 to Apr 22) if the beginning time of country blockade were Mar 05. Furthermore, that would be 44993 (95%CI: 16676–96759) and Apr 27 (95%CI: Apr 01 to Jul 27) if the beginning of country blockade were Mar 15.

4 Discussion
This impact of the COVID-19 response (overall quarantine regulations, social distancing, and isolation of infections) in China is encouraging for the many other countries (13). We compared the situation in Hunan, China, which has the similar population as Italy to predict. The spread of COVID-19 in Hunan Province appeared relatively early and has now entered a phase of no inflections, which helps to observe the entire course of the epidemic. Besides, due to the similarity of population size and geographical location adjacent to Hubei, Hunan's public health measures can provide useful guidance for Italy in preventing the further spread of COVID-19.

In our study, the eSIR model with R software package was used to evaluate the impact of intervention measures on Italian COVID-19 epidemic. In previous studies, estimation of the epidemic of an infectious disease is often performed using constant parameter (14-17). The advantage of eSIR model is that it combines time-varying isolation measures and expands SIR model to adapt to the time-varying transmission rate in the population. Lili Wang et al. found that COVID-19 outside Hubei in China has been so far much less severe (9). But they did not perform each province’s analyses. The first and second points in our study are respectively Feb 04 and Feb 09, which are the same as these outside Hubei in China. Furthermore, the actual number of infected cases (1018) is included in the predicted number of infected cases (786-9065) and the endpoint (Feb 29) is included in the predicted endpoint (Feb 19 to Mar 23) in our study, which also reflects stability and accuracy of the eSIR model. Combining the above data and methods, these findings show that eSIR model is more suitable for predicting the epidemic trend of COVID-19.

Our results showed that R0 was estimated to be respectively 3.12 (95% CI, 1.83-3.01) and 3.74 (95% CI, 2.09-6.02) in SIR model and eSIR model in Italy; 2.97 (95% CI, 1.80-5.01) and 3.15 (95% CI, 1.71-5.21) in SIR model and eSIR model in Hunan. Li Qun et al estimated R0 to be 2.2 (95% CI, 2.09-6.02) among the first 425 patients in Wuhan, China (18). Other studies estimated R0 to be 1.4-2.5 (19), 2.68 (95% CI 2.47-2.68) (3), 3.6-3.8 (20), 6.47 (95% CI 5.71-7.23) (7). Ying Liu et al found that the estimated mean R0 for COVID-19 is around 3.28, with a median of 2.79 and IQR of 1.16 by reviewing R0 of COVID-19 in 12 studies. (21) We estimated R0 to be 3.15 (95% CI 1.71-5.21) in Hunan, which is in agreement with these findings. But R0 in Italy was estimated to be 4.10 (95% CI: 2.15–6.77), which is larger than that in Hunan. This needs to be confirmed by further studies.

This study showed that COVID-19 spread rapidly throughout Italy after Feb 21. Possible reasons for such rapid growth of infections include: (1) more timely caution and preventative measures were not taken, (2) The number of infections during Jan 31-Feb 20 could be under-reported due to underdiagnosis, given subclinical or asymptomatic cases. The incubation period for COVID-19 is thought to be within 14 days following exposure, with most cases occurring approximately four to five days after exposure (18, 22, 23). So it seems impossible to maintain totally two or three cases during Jan 31-Feb 20 in Italy. In addition, the rapid increase in the number of infections after Feb 21 might reflect a belated realization of the spread of COVID-19.

Previous studies have shown that more rigorous government control policies were associated with a slower increase in the infected population (6, 16, 24-27). Our report also showed that rigorous measures can effectively prevent the further spread of COVID-19 in Hunan. Based on our model, Italy should still maintain all levels of quarantines like China by Apr 25 (Mar 30-Aug 07). Furthermore, Tianyi Qiu et al found that delaying the lockdown by 1-6 days in Wuhan would expand the infection scale 1.23-4.94 times and the epidemic would be finally out of control if lockdown imposed 7 days later (17). Our study also shows that taking government control earlier can decrease the number of infected cases. In addition, from China’s experience, various control measures, including the early detection and isolation of individuals with symptoms, traffic restrictions, medical
tracking, and entry or exit screening, can well prevent the further spread of COVID-19. These measures are in line with the latest recommendations by the World Health Organization and a previous study in Spain (28). But the most effective strategy still needs to be confirmed by further studies. Consequently, it is better and necessary to apply strict public health measures in other European countries with a high number of COVID-19 cases.

Our study has some limitations. Firstly, it is based on the assumption that rigorous measures like China have been taken in Italy, although this study uses the new model to obtain dynamic results, which is instructive for the prevention and control of the epidemic in Italy. Secondly, the suspected cases and the daily number of hospitalized cases are not available, so they are not considered in the eSIR model. Thirdly, some unforeseeable factors may affect these estimated data in our study such as super-spreaders exist.

In conclusion, the current study is the first to provide a prediction for epidemic trend after strict prevention and control measures were implemented in Italy. Our study suggests that rigorous measures like China should still be maintained in Italy by Apr 25 (Mar 30-Aug 07) to prevent further spread of COVID-19.

5 Conflict of Interest

We declare no competing interests.

6 Author Contributions

JW, HK, LM and HY contributed to the study design. JW, HK, and LM contributed to the writing of the manuscript. JW, HK, SY, CW contributed to the data analysis. SY, CW and WJ contributed to the data compilation. WJ, YS, WS and HY contributed to critical review. TP, KF and LJ contributed to the literature search. YS, WJ, and KF contributed to the design of tables and figures.

7 Funding

The study was funded by Army Logistics Emergency Scientific Research Project; Emergency scientific research of the army and the emergency scientific research of Chinese PLA General Hospital (20EP008).

8 Acknowledgments

We greatly appreciated the technical assistance provided by Zhenxing Cheng; Institute of Blue and Green Development, Shandong University, Weihai, 264209, PR China.

9 Reference


### Tables

#### Table 1. R0, incidence and endpoint in Hunan and Italy according to the eSIR model.

<table>
<thead>
<tr>
<th>Model</th>
<th>R0</th>
<th>Incidence (%)</th>
<th>Endpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>95% CI</td>
<td>Mean</td>
</tr>
<tr>
<td>Hunan</td>
<td>3.15</td>
<td>1.71-5.21</td>
<td>0.0066</td>
</tr>
<tr>
<td>Italy</td>
<td>4.10</td>
<td>2.15-6.77</td>
<td>0.0498</td>
</tr>
</tbody>
</table>

Inf: The endpoint appears more than 200 days after t₀.

#### Table 2. R0, incidence and endpoint in Italy according to different beginning of country blockade.

<table>
<thead>
<tr>
<th>Begin time of country blockade</th>
<th>R0</th>
<th>Incidence (%)</th>
<th>Endpoint</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>95% CI</td>
<td>Mean</td>
</tr>
<tr>
<td>Mar 05</td>
<td>3.26</td>
<td>1.83-5.34</td>
<td>0.0176</td>
</tr>
<tr>
<td>Mar 15</td>
<td>4.24</td>
<td>2.44-6.61</td>
<td>0.0745</td>
</tr>
</tbody>
</table>
**Figures**

**Figure 1. Epidemic development of COVID-19 in Hunan, China and Italy.**

A, B: Daily new COVID-19 cases in Hunan, China and Italy.
C, D: Epidemic distribution of COVID-19 in Hunan, China and Italy.

**Figure 2. Epidemiological trend of COVID-19 under existing preventions in Hunan, China and Italy according to eSIR model.**

A: Hunan, China; B: Italy.

The black dots left to the blue vertical line denote the observed proportions of the infected and removed compartments on the last date of available observations or before. The blue vertical marks is the current time up to which we have observed data (t₀). The green and purple vertical lines denote the first and second turning points, respectively. The cyan and salmon color area denotes the 95% credible interval of the predicted proportions of current infected cases before and after t₀, respectively. The gray and red curves are the posterior mean and median curves.

Epidemiological trend of COVID-19 under existing preventions of Hunan, China and Italy in eSIR model.
Figure 1. Epidemic development of COVID-19 in Hunan, China and Italy.
Figure 2. Epidemiological trend of COVID-19 under existing preventions in Hunan, China and Italy according to eSIR model.