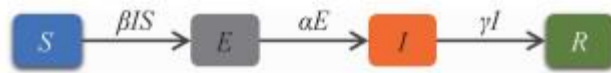


Dynamic model: The traditional SEIR model divides the population into four categories: vulnerable population (S), incubation period population (E), infected population (I), and evacuated population (R) (Figure 1), considering that the infected population has Contagious. Some scholars consider that the population in the incubation period is infectious in the SEIR model.

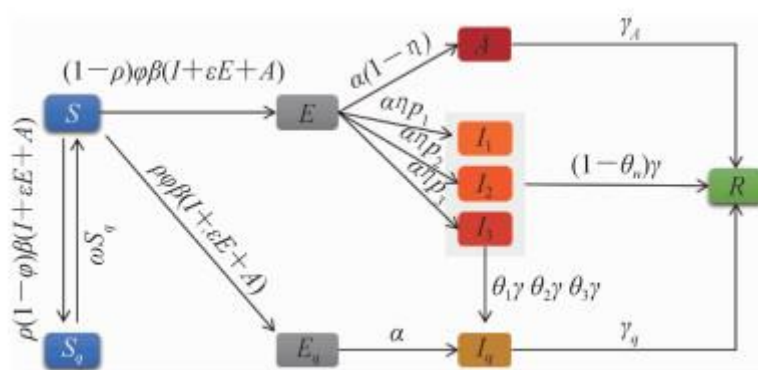


$$\begin{cases} \frac{dS}{dt} = -\beta S(t)I(t) \\ \frac{dE}{dt} = \beta S(t)I(t) - \beta S(t-\omega)I(t-\omega) \\ \frac{dI}{dt} = \beta S(t-\omega)I(t-\omega) - \beta S(t-\omega-\tau)I(t-\omega-\tau) \\ \frac{dR}{dt} = \beta S(t-\omega-\tau)I(t-\omega-\tau) \end{cases}$$

In the formula, τ is the infectious period. The infected person at time $(t-\omega)$ becomes infectious at time t after a period of ω ; while the infected person at time $(t-\omega-\tau)$ is infectious at time $(t-\tau)$, and at time t become a restorer.

The traditional SEIR model cannot satisfy our prediction of COVID-19. Because of this epidemic prevention work, group isolation strategies are generally adopted worldwide, and the corresponding S, E, I, and R populations have increased the corresponding quarantined populations. In addition, super infections, asymptomatic infections, etc. should all be considered.

So the extended model of SEIR was proposed



Due to the characteristics of the new coronavirus, this epidemic has its particularity: ① There are asymptomatic infected (A) and infectious, and ② People in the incubation period are also infectious. The traditional SEIR model does not consider these factors. Considering the current epidemic control measures (such as medical observation and isolation, close contact isolation, and isolation of infected persons) and the classification of infected persons (light, common, severe or critical), this study is further expanded on the basis of the SEIR model, adding Isolated susceptible

population (S_q), isolated incubation period population (E_q), isolated infected population (I_q), and consider the types of infected components (IC) (I_1 light type, I_2 normal type, I_3 heavy and critical), Referred to as SEIR expansion model; infected components, C ; asymptomatic infected, A ; quarantined, Q), a total of 10 warehouses (Figure 2).

$$\left\{ \begin{array}{l} \frac{dS}{dt} = - (1 - \rho)\beta\varphi(\varepsilon E + I + A)S - \rho\beta(1 - \varphi) \\ \quad (\varepsilon E + I + A)S - \rho\beta\varphi(\varepsilon E + I + A)S + \omega S_q, \\ \frac{dE}{dt} = (1 - \rho)\beta\varphi(\varepsilon E + I + \varepsilon A)S - \alpha E, \\ \frac{dA}{dt} = \alpha(1 - \eta)E - \gamma_A A, \\ \frac{dI_1}{dt} = \alpha\eta\rho_1 E - \gamma I_1, \\ \frac{dI_2}{dt} = \alpha\eta\rho_2 E - \gamma I_2, \\ \frac{dI_3}{dt} = \alpha\eta\rho_3 E - \gamma I_3, \\ \frac{dR}{dt} = \gamma_A A + (1 - \theta_1)\gamma I_1 + (1 - \theta_2)\gamma I_2 + \\ \quad (1 - \theta_3)\gamma I_3, \\ \frac{dS_q}{dt} = \rho\beta\varphi(\varepsilon E + I + A)S - \omega S_q, \\ \frac{dE_q}{dt} = \rho\beta(1 - \varphi)(\varepsilon E + I + A)S - \alpha E_q, \\ \frac{dI_q}{dt} = \alpha E_q + \theta_1\gamma I_1 + \theta_2\gamma I_2 + \theta_3\gamma I_3 - \gamma_q I_q, \\ \rho_1 + \rho_2 + \rho_3 = 1, I = I_1 + I_2 + I_3 \end{array} \right.$$

Parameter setting and model evaluation: Combine information from literature reports, official news, epidemiological surveys, and expert opinions to set parameters. The unknown parameters in the model can be estimated according to the daily cumulative number of confirmed cases described by the National Bureau of Statistics, using the Markov Chain Monte Carlo (MCMC) algorithm, combined with Gibbs sampling and non-informative prior, and then Make predictions. Simulate 200,000 times and burn in 100,000 times. The average deviation between the predicted value of the model and the actual value (BIAS) of the cumulative number of confirmed cases is used to evaluate the fitting effect and prediction effect of the model

parameter	meaning	Starting value	Search scope	Setting basis
S	Vulnerable population	100000	80000-120000	Expert Opinion
E	Incubation period	109	50-150	I(0)*Latent period length
I1	Not isolated mild patients	10	fixed	According to Iq(0), ρ , θ
I2	Normal patients not isolated	10	fixed	According to Iq(0), ρ , θ
I3	Not isolated critically ill patients	2	fixed	According to Iq(0), ρ , θ
A	Infectious asymptomatic infection	0	fixed	According to Iq(0), η
R	Evacuate (cured or died)	10	fixed	Official website initial data
Sq	Quarantined vulnerable population	0	fixed	The initial state is 0
Eq	Quarantined incubation period population	0	fixed	The initial state is 0
Iq	Quarantined infected person	0.5	fixed	Official website initial data
β	Contact rate of susceptible persons with incubation period and infected persons	9.08*0.00001	0.00001-0.0001	R0=2.2 infer initial data
ρ	Probability of being tracked	0.5	0.1-0.9	Expert Opinion
Φ	Infected constituent ratio among contacts	0.5	0.1-0.9	Expert Opinion
ε	Transmission coefficient for infected persons and incubation period	0.5	0.1-0.9	Expert Opinion
α	The rate at which people in the incubation period become infected	1/7.2	fixed	references
η	Proportion of symptomatic infections to infections	0.7	0.5-0.95	Expert Opinion
ρ_1	Composition ratio of mild infections	0.5	fixed	references
ρ_2	The proportion of ordinary infected persons	0.35	fixed	references
ρ_3	The proportion of severely infected persons	0.15	fixed	references
γ	Removal rate of infected persons (reciprocal course of disease)	0.06	fixed	references
γ_A	Probability of a mildly infected person being diagnosed and isolated	0.2	0.3-0.8	Expert Opinion
γ_q	Probability of an ordinary infected person being diagnosed and isolated	0.8	0.5-0.9	Expert Opinion
θ_1	Probability of severely infected people being diagnosed and isolated	1	fixed	Expert Opinion
θ_2	Probability of severely infected people being diagnosed and isolated	1	fixed	Expert Opinion
θ_3	Probability of release of quarantine for vulnerable population	0.1	0-1	National policy related
w				

1. Data source: The case data comes from the daily cumulative number of cases updated daily by the National Bureau of Statistics and the World Health Organization and information on cases and close contacts reported on related websites. The information collected includes the cumulative number of reported cases, the number of new cases per day, and the cumulative number of dismissed medical observation personnel.

2. Incubation period and generation interval estimation: the incubation period is the difference between the time of an infected person's infection and the time when clinical symptoms appear [6], the generation interval refers to the difference between the time of infection of an infected person and the time of infection of the next generation of infected persons [7]. The estimation of the incubation period is very important. It is an important basis for determining the isolation period of exposed persons and judging the infection time of a case. It is also a key parameter of the epidemiological prediction model; the generation interval represents the average time between the spread of the disease from one person to another person, The shorter the generation interval, the easier it is for the epidemic to show an outbreak pattern. The time of exposure to the source of infection in some cases cannot be directly observed, which makes it difficult to estimate the incubation period. Usually we cannot accurately observe the time of exposure to the source of infection and the appearance of symptoms, they may fall into certain intervals. However, the incubation period can be estimated by determining the difference between the earliest and latest exposure time of each case and the time when symptoms appear, and this time can be regarded as the interval cut-off estimate of the incubation period for each person [8].

R is an important parameter in the epidemic model. It represents the average number of people infected by a case during the disease period when all people are susceptible at the beginning of the disease. When $R > 1$, the epidemic will break out; otherwise, the epidemic will die out.

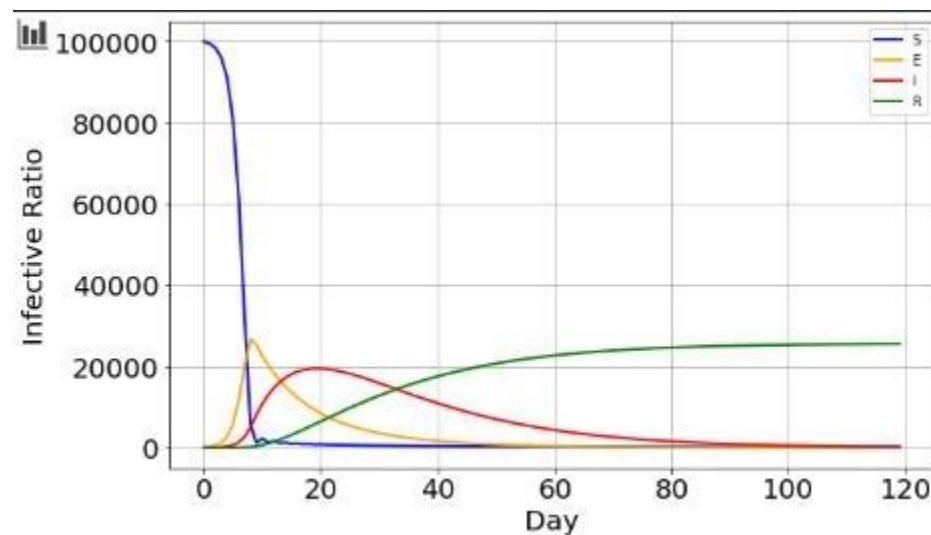
The definition R is the exponential growth rate of the cumulative number of infected cases, and C(t) is the cumulative number of COVID-19 cases. In the initial stage, the cumulative number of cases tends to increase exponentially, but how long this exponential growth stage will last requires people to judge. In this study, the size of the goodness of fit coefficient R2 was used to judge the duration of the initial exponential growth phase. Since the number of cases is an integer value, Poisson regression is used to estimate this parameter [14]. Refer to literature [13] to calculate R0:

$$R_0 = \frac{1}{M(-r)} = \frac{1}{\int_{a=0}^{\infty} e^{-ra} f(a) da}$$

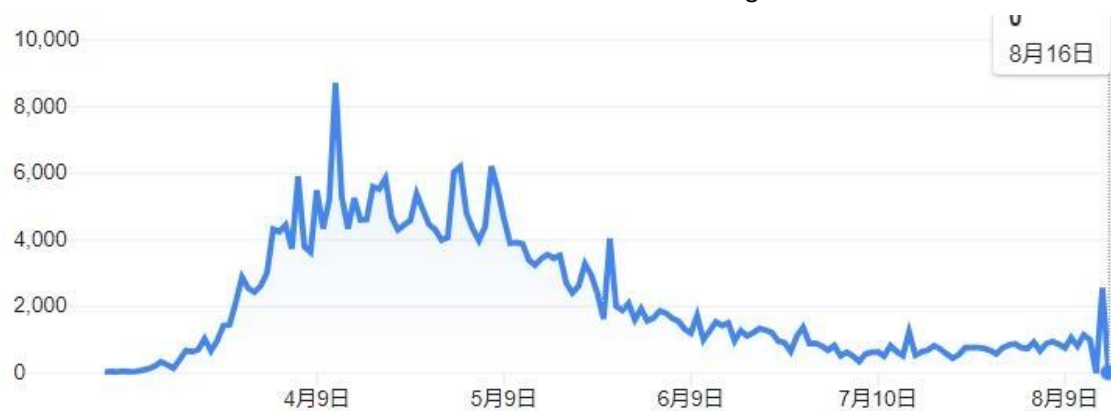
Where $M(-r)$ is the moment generating function of the generation interval distribution $f(a)$.

3. In the formula, β represents the transmission coefficient, k is the transfer rate from exposure to infection ($1/k$ on the value is the average length of the post-infection but non-infectious period), and γ represents the removal rate ($1/\gamma$ on the value is the The average length of the infectious period). Fitting the cumulative number of cases by the least squares method can estimate the 3 parameter values in the model within an appropriate parameter value range.

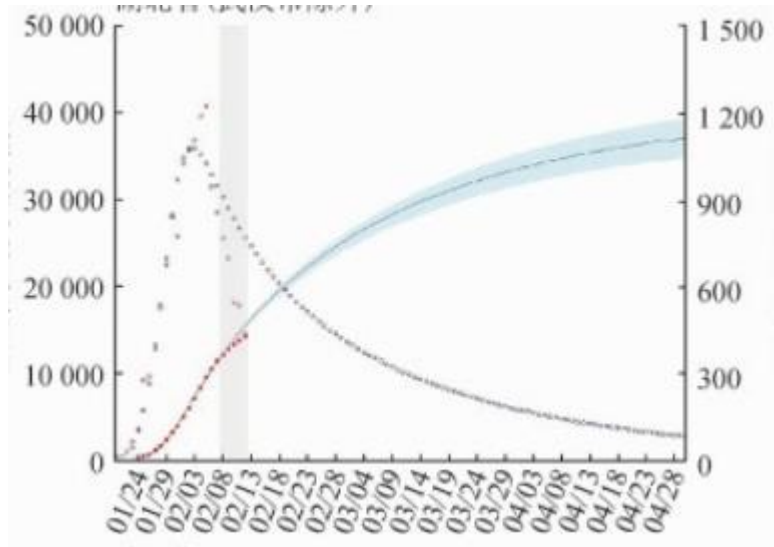
After setting the parameters for the local case in the UK, the visual output results are as follows:



The actual number of confirmed cases in the UK is shown in the figure below:



It basically stabilizes after 120 days and the fit of I is very high.



The estimated deviations of the model for the cumulative number of confirmed cases in the past 10 days are all within 5% (Table 3). The forecast deviation for the next 5 days is $3.8\% \pm 2.1\%$, and the maximum deviation is 6.5%.

This epidemic has a large variation in the incubation period and is contagious, and there is a specific proportion of asymptomatic or mildly infected people who have not been admitted to isolation. The traditional SEIR model does not fully consider these factors. The construction of the dynamic model requires the combination of disease characteristics and epidemic prevention and control measures. The improved SEIR expansion model proposed in this study fully considers the impact of these factors on the epidemic. The results show that the fitting effect is good, and the prediction deviation is within an acceptable range. 7.14 The forecast result is slightly higher, suggesting that in addition to conventional isolation measures (isolation of susceptible people, isolation of incubation period, isolation of infected persons), other comprehensive linkage prevention and control measures are effective, and it also indicates the need to iteratively update the dynamic model based on new data.

SEIR dynamic model can provide a reference for evaluating the effect of intervention measures. When the basic hypothesis is established, especially in the absence of external intervention, establishing a dynamic model based on early epidemic data or model prediction based on prior parameters can predict the development of the epidemic without intervention. However, when an epidemic occurs, corresponding prevention and control measures are generally taken. If the measures are appropriate, the epidemic will be controlled. At this time, the dynamic model results can be used as a reference to evaluate the effects of prevention and control measures.

The model must be dynamically adjusted while taking into account the evolution of the epidemic and the adjustment of prevention and control measures. The epidemic is fierce, and it is even more difficult to prevent and control the epidemic. Various measures have been adopted in various places to restrict internal personnel mobility, such as delays in starting work, delaying school opening, and traffic control. But the movement of people is imminent. The future increase in the mobility of personnel such as resumption of work and school will inevitably increase the difficulty

of epidemic control. The model should also consider population inflows and outflows, and isolation.

This model has limitations. First, most parameters (such as tracking rate, admission rate, isolation rate and other intervention measures related parameters) will dynamically change during the evolution of the epidemic. The current model has not considered this feature, and the random parameter dynamic model is worthy of further study. Second, the SEIR extended model considers too many parameters. In the case of fewer data points, the model may be over-fitted, which will affect the prediction effect and mislead the research and judgment of the epidemic. It is necessary to fix some parameters based on the literature results and comprehensive expert opinions to reduce the uncertainty of the model. Third, the official daily cumulative number of confirmed cases is based on the date of diagnosis of new coronary pneumonia. In fact, the number of daily cases can better reflect the epidemic pattern. Finally, with the resumption of work and school, the increase in intra-city mobility and inter-city mobility is inevitable, and these factors need to be considered in the later forecast.

From the analysis results, the inflection point of the epidemic has passed, but it will take a long time for the epidemic to ease and even end. We cannot relax our vigilance, and we should continue to strengthen prevention and control measures to avoid another outbreak. The SEIR extended model proposed in this paper can be used to predict the epidemic situation in the presence of prevention and control measures and evaluate the effects of prevention and control measures. It can also be further extended on the basis of this model to evaluate the impact of future population mobility on the epidemic.

[6]Cowling BJ, Park M, Fang VJ, et al. Preliminary epidemiological assessment of MERS-CoV outbreak in South Korea, May to June 2015[J]. Euro Surveill, 2015, 20(25): 7-13.

[7]Anderson RM, May RM. Infectious diseases of humans:dynamics and control[M]. Oxford, UK: Oxford University Press, 1991.

[8]Virlogeux V, Fang VJ, Park M, et al. Comparison of incubation period distribution of human infections with MERS-CoV in South Korea and Saudi Arabia[J]. Sci Rep, 2016, 6: 35839.

[13]Wallinga J, Lipsitch M. How generation intervals shape the relationship between growth rates and reproductive numbers[J]. Proc Biol Sci, 2007, 274(1609): 599-604.

[14]Hens N, Van Ranst M, Aerts M, et al. Estimating the effective reproduction number for pandemic influenza from notification data made publicly available in real time:a multi-country analysis for influenza A/H1N1v 2009[J]. Vaccine, 2011, 29(5): 896-904.