

The impact of containment measures and air temperature on mitigating COVID-19 transmission: non-classical SEIR modeling and analysis

Di Liu^{†,1}, Qidong Tai^{†,2}, Yaping Wang³, Miao Pu³, Sikai Ge⁴, Tingting Ji⁵, Lei Zhang^{*,2}, Bo Su^{*,6}

†: Contributed equally

*: Joint correspondence authors

Supplementary methods

The stochastic ScEIQRsh epidemiological model

To simulate the epidemiological data of COVID-19 outbreak and control across provinces in Mainland China from January 2020 to March 2020, we developed an early spreading, early non-pharmaceutical-intervention stochastic model, denoted as ScEIQRsh model. The flow diagram of ScEIQRsh model was demonstrated in Fig. 1.

Compartments of the stochastic ScEIQRsh model:

Sc: the contactable susceptible subpopulation for the infectors, comprised of their family, relatives, co-workers, friends, and some strangers who could be contacted by the infector under the interventive social prevention. The initial Sc is defined as a continuous random variable with Gaussian distribution in the model.

E: the exposed individuals who are in the incubation period after effective contact with the infectors.

I: the infectors, either immigrant or local reproductive infectors who are still outside of the public health measures.

Q: The close contacts of infectors found out by epidemiological survey, and notified to be in self-quarantine at home, in the hotel or indicated isolating room for 14-day medical observation. The Q value in this model is the daily reported cumulative close contacts or entry of medical observation minus the daily cumulative dismissal of medical observation for each province.

Rs: The self-recovery individuals who have never been diagnosed and hospitalized because of mild symptoms, or asymptomatic infection, and thus were not recorded in the daily official epidemic reports.

Rh: The cumulative individuals who were etiologically diagnosed (mostly SARS-CoV-2 RNA rtPCR positive in oropharyngeal swabs, and in consistent with other clinical symptoms) and hospitalized in isolate wards. The cumulative number includes any hospitalizing, or dead, or cured COVID-19 patients.

Removed: The removed means any infectors who have been deprived of ability to propagate, either by gain of immunity (Rself, cured in Rh) or by public health measures (infectors in Q, the hospitalizing in Rh), or death (the dead in Rh). The removed in this model is the sum of cumulative Rself, Rh and the positive cases in Q.

Parameters of the stochastic ScEIQRsh model

β : the effective transmission rate for the contactable susceptible (S_c).

σ : the progression rate from exposed to being infectious, which is the reciprocal of the incubation period (days) in the propagate chain.

γ : the remove rate for R_{self} , which is the reciprocal of the apparent period of being propagating for a self-recovery infector in the propagate chain.

η : the remove rate for R_h , which is the reciprocal of the apparent period of being propagating for a hospitalized infector in the propagate chain.

κ : the average number of traceable close contacts for each confirmed case, which was investigated and notified to be in quarantine by the epidemiological survey group of each province. All the found close contacts were assumed to be in 14-day quarantine.

ρ : the virus-positive rate of the individuals in quarantine for each province.

ω : the dismiss rate of virus-negative individuals in quarantine

Formulation and Parameters setting of ScEIQRsh model

The simultaneous differential equations system for the stochastic ScEIQRsh model is as following:

$$(1) \frac{dE}{dt} = \beta S_c I / (S_c + E + I + R_s + R_h + \rho Q) - \kappa \eta I - \sigma E$$

$$(2) \frac{dI}{dt} = \sigma E - \eta I - \gamma I$$

$$(3) \frac{dQ}{dt} = \kappa \eta I - \omega Q$$

$$(4) \frac{dR_h}{dt} = \eta I + \rho \omega Q$$

$$(5) \frac{dR_s}{dt} = \gamma I$$

* In equation (1), unlike the classical SIR or SEIR model, the denominator of flow-in velocity is ($S_c + E + I + R_s + R_h + \rho Q$) instead of N or any other constants. The denominator means all the transmitted individuals in the system. The classical SIR or SEIR model assumes that the infectors mix up with different susceptible every day, and ScEIQR model assumes that the infectors mix up with the fixed contactable susceptible every day.

The parameters setting: β , σ , γ and η were set as random variables with Gaussian distribution; κ , ω , ρ was set as random variables with uniform distribution. Parameters range setting: $\beta[1,19]$, $\sigma[0.27,0.5]$, $\gamma[0.04,0.3]$, $\eta[0.001,0.999]$, $\kappa[0,350]$, $\omega[0.07,0.6]$, $\rho[0,0.1]$.

S_c was set as a random variable with Gaussian distribution and the range setting was $S_c(0,0.002N)$. N denotes the total population of the province. The other initial compartment values were estimated as the following: initial $R_h = H_0$, initial $R_s = 0$, initial $Q = Q_0$, initial $I = H_0 * (1 - \eta) / \eta$, initial $E = \text{initial } I / \sigma$. H_0 , Q_0 denotes the cumulative

hospitalized cases, close contacts in quarantine at Day 0 (23rd, Feb, 2020) reported by public health administration of the province. If H0 or Q0 is missing for some province, H0 or Q0 will be given an assumed number.

Simulation and Model Fitting

The data process was performed on R (version 3.6.1), and the “deSolve” R package was used as the solver of differential equations. The R source code can be found in GitHub.

Briefly, model parameters and Sc were random sampling with Metropolis-Hastings (M-H) algorithm, one of Markov Chain Monte Carlo (MCMC) method. The proposal distribution for accept-reject is a Bernoulli distribution from the comparison of cost function of curve fitting in iteration (better or not).

Both curves of Rh and Q was simultaneously fitted with the raw data by least squares method, and the cost function was SSE/SST (Sum of Squares for Error/Sum of Squares for Total). The optimized parameters were documented with 100000 iterations of 0.1 step size from 0 to 60 days with burn-in of 50000 iterations for every province of Mainland China. The expected value and standard deviation for each parameter were calculated.

Other indexes

Sc(c): restriction factor, the proportion of contactable susceptible (Sc) over the total population of a province under the interventive social prevention, which is simply calculated as Sc/N.

utI%: the proportion of the self-recovery removed, including asymptomatic infections or any infection without hospitalization and report, which were estimated as $\gamma/(\eta+\kappa\rho\eta+\gamma)$.

SurveyQ: an estimation for the quality of epidemical survey, which is calculated as $\kappa\rho$.

Incubation period: The incubation period was the time elapsed from exposure to SARS-COV-2 to the symptoms firstly apparent, calculated with $1/\sigma+1/\eta$.

Communicable period: The time for untraceable infectors with contagious among susceptible, calculating with $1/\gamma$.

Rvq: the index to estimate overall effect of non-pharmaceutical-intervention, which was defined as the ratio of transmission velocity to remove velocity when Q reached the max. If $Rvq>1$, the intensity of non-pharmaceutical-intervention is not enough to stop propagating of infectors, the infectors will increase. If $Rvq<1$, the intensity of non-pharmaceutical-intervention can stop propagating of infectors, the infectors will decrease. The derivation process is as follows:

The transmission velocity: $\beta \cdot Sc \cdot I / (Sc + E + I + Rs + Rh + \rho Q)$

The remove velocity: $\gamma I + \eta I + \kappa \rho I$

Thus, the ratio of velocity (in/out):

$$Rv = (\beta \cdot Sc \cdot I / (Sc + E + I + Rs + Rh + \rho Q)) / (\gamma I + \eta I + \kappa \rho I)$$

And the following equation can be inferred:

$$R_v = \beta / (\gamma + \eta + \kappa \eta \rho) \cdot (c \cdot S_c) / (cN - S_c(1-c) - c(1-\rho)Q)$$

In the above equation, all the parameters are constants during the spreading and control period except Q , which R_v is changing dependent on Q . Hence, we defined R_{vq} as the value of R_v when Q approaches the maximum, and R_{vq} was used to estimate the overall effect of non-pharmaceutical-intervention.

$$R_{vq} = \beta / (\gamma + \eta + \kappa \eta \rho) \cdot (c \cdot S_c) / (cN - S_c(1-c) - c(1-\rho)Q_{max})$$

The air temperature of every province during COVID-19 outbreak and spreading

The historical meteorological data were collected from china meteorological administration. The daily mean air temperature was calculated from the Jan 15, 2020 to Feb 15, 2020, i.e., from a week before the Jan 23, 2020 to 3 weeks after that. In this period, COVID-19 began spreading and controlled by NPI in most of provinces of China, except Hubei.

The epidemical data and epidemical survey of 31 provinces

The daily confirmed and quarantined cases were used for fitting the model. Cases confirmed with laboratory test of the same sample identified two targets positive detection with real-time reverse-transcription-polymerase-chain-reaction (RT-PCR) assay or high-throughput sequencing, one is ORF, another one is the N protein of SARS-CoV-2. The quarantined cases were the population close contact the confirmed cases, asymptomatic infections and suspected cases identified with epidemiological investigation within 24 hours. The close contacts were including:

- 1) living, studying, working in the same house with the confirmed or suspected cases.
- 2) medical staff, family members or other persons who have close contact with cases in the process of diagnosis, treatment, or nursing.
- 3) take the same transport and have close contact with infector, including caregivers, peers or other passengers(31).

The quarantined people were isolated in home or in a specific place for observation their symptoms for 14 days. If the lab testing results of close contacts of confirmed cases and asymptomatic infections are negative during the medical observation period, they still need to continue quarantine until the end of the observation of 14 days. For those who are in close contact with suspected cases, can be relieved of medical observation when suspected cases are excluded of infection.

The provinces of China

There are 34 provincial-level administrative divisions of China, including 23 provinces, 4 municipalities (Beijing, Tianjin, Shanghai, Chongqing), 5 autonomous regions (Guangxi, Inner Mongolia, Tibet, Ningxia, Xinjiang) and 2 special administrative regions (Hong Kong, Macau). 23 provinces are including Anhui, Fujian, Gansu, Guangdong, Guizhou, Hainan, Hebei, Heilongjiang, Henan, Hubei, Hunan, Jiangsu, Jiangxi, Jilin, Liaoning, Qinghai,

Shaanxi, Shandong, Shanxi, Sichuan, Yunnan, Zhejiang, Taiwan. Seven geographical regions were classified of Mainland China, named as North China, Northeast China, East China, Central China, South China, Northwest China, Southwest China. The 29 provinces were included in our study, and they were separated into each geographical region and represented by numbers as following:

- 1) North china: Beijing, Tianjin, Hebei, Shanxi, Inner Mongolia (1-5);
- 2) Northeast China: Liaoning, Jilin, Heilongjiang (6-8);
- 3) East China: Shanghai, Jiangsu, Zhejiang, Anhui, Fujian, Jiangxi, Shandong (9-15) and Taiwan;
- 4) Central China: Henan, Hubei, and Hunan (16-18);
- 5) South China: Guangdong, Hainan, Guangxi (19-21), Hongkong, and Macau;
- 6) Northwest China: Shaanxi, Gansu, Ningxia, Xinjiang (22-25) and Qinghai;
- 7) Southwest China: Chongqing, Sichuan, Guizhou, Yunnan (26-29) and Tibet.