**Stochastic SIHR model:**

We introduced Susceptible-Infected and infectious without isolation-Hospitalized in isolation-Removed (SIHR) model with four compartments: Susceptible ($S$), infected and infectious without isolation ($I)$, hospitalized in isolation ($H)$, removed ($R)$. The corresponding dynamic system of these four compartments were described by ordinary differential equations:

$\frac{dS\left(t\right)}{dt}=-α\frac{I\left(t\right)}{N}S(t)$,

$\frac{dI\left(t\right)}{dt}=α\frac{I\left(t\right)}{N}S(t)-βI\left(t\right)$,

$\frac{dH\left(t\right)}{dt}=βI\left(t\right)-γH\left(t\right)$,

$\frac{dR\left(t\right)}{dt}=γH\left(t\right)$,

$N=S\left(t\right)+I\left(t\right)+H\left(t\right)+R(t)$,

where $α$ is the average number of contacts per person per time. In assuming the incubation period and the hospitalization period exponentially distributed, $1/β$ is the mean of incubation period, $1/γ$ is the mean of the hospitalization period. $N$ is the total number of host population.

Based on time-varying SIHR model, let $Ω=\left(α\_{0}, d, m,γ\right) , $we introduced randomness to SIHR model: Let $S\_{t}, I\_{t}, H\_{t} , R\_{t}$ be the observed values of $S, I, H, R$ at the time $t$ and $A\_{t}=(S\_{t}, I\_{t}, H\_{t} , R\_{t})$ is assumed to be a multidimensional Markov process, i.e., for $g<t$,

$p\left(A\_{s\leq g}\right)=p\left(A\_{g}\right)=Mul\left(\frac{A\_{t}\left(A\_{g},Ω\right)}{N},N\right)$,

where $A\_{t}\left(A\_{g},Ω\right)$ is the vector at the time $t$ determined by the deterministic dynamic system above given the initial value $A\_{g}$ and parameters $Ω$ and $β$. $Mul(∙|p,n)$ is the density of multinomial distribution with the incident rate vector $p$ and the total number $n$.

****

**Figure 1.** The diagram of transfer process of four compartments. Susceptible ($S$), infected and infectious without isolation ($I)$, hospitalized in isolation ($H)$, and removed ($R)$.

Notice that the observable data is $H\_{1:T}, R\_{1:T}, N$, where $T$ is the length of time. Treat $I\_{1:T}$ as the missing data, once the prior distribution of $Ω $and $I\_{1}$: $π\left(Ω\right)$, $π(I\_{1})$ were given, we apply the Bayesian method combined with MCMC (Markov Chain Monte Carlo) to simulate the posterior distribution:

$π\left(H\_{1:T},R\_{1:T},N,β\right)∝\prod\_{i=1}^{T-1}Mul\left(\frac{A\_{i+1}\left(A\_{i},Ω\right)}{N},N\right)π\left(Ω\right) π(I\_{1})$*.*

The point estimates of $Ω $and $I\_{1:T}$ were presented as the median of the posterior distribution while 95% credible intervals were constructed with 2.5% and 97.5% quantiles.



**Figure 2.** The time-variant $α$.

**COVID-19 prevalence**

Definition of COVID-19 cases (per 100,000)：

$$COVID-19 cases\left(per 100, 000\right)=\frac{Cumulative confirmed cases}{Total population of this area}×100,000$$



**Figure 3.** COVID-19 cases (per 100, 000) in South Korea, Italy, and the United States from February 18 to April 23. On February 18, the 31st confirmed case in South Korea triggered the rapid spread in this country, so we used it as the starting date of wide spread outside mainland China.

**Table 1.** COVID-19 cases (per 100,000) in South Korea, Italy and the United States from February 18 to April 23.

|  |  |  |  |
| --- | --- | --- | --- |
| **Date** | **South Korea** | **Italy** | **The United States** |
| 2/18/2020 | 0.061  | 0.005  | 0.005  |
| 2/19/2020 | 0.100  | 0.005  | 0.005  |
| 2/20/2020 | 0.203  | 0.005  | 0.005  |
| 2/21/2020 | 0.398  | 0.005  | 0.005  |
| 2/22/2020 | 0.845  | 0.055  | 0.010  |
| 2/23/2020 | 1.175  | 0.218  | 0.010  |
| 2/24/2020 | 1.626  | 0.380  | 0.010  |
| 2/25/2020 | 1.907  | 0.467  | 0.016  |
| 2/26/2020 | 2.462  | 0.618  | 0.017  |
| 2/27/2020 | 3.448  | 0.872  | 0.018  |
| 2/28/2020 | 4.562  | 1.078  | 0.018  |
| 2/29/2020 | 6.149  | 1.468  | 0.019  |
| 3/1/2020 | 7.293  | 1.863  | 0.021  |
| 3/2/2020 | 8.463  | 2.859  | 0.027  |
| 3/3/2020 | 10.124  | 3.442  | 0.032  |
| 3/4/2020 | 10.973  | 4.205  | 0.038  |
| 3/5/2020 | 11.885  | 5.192  | 0.049  |
| 3/6/2020 | 12.871  | 6.486  | 0.071  |
| 3/7/2020 | 13.745  | 7.729  | 0.103  |
| 3/8/2020 | 14.276  | 10.061  | 0.135  |
| 3/9/2020 | 14.598  | 12.261  | 0.174  |
| 3/10/2020 | 14.667  | 15.148  | 0.218  |
| 3/11/2020 | 15.139  | 16.999  | 0.305  |
| 3/12/2020 | 15.362  | 20.581  | 0.402  |
| 3/13/2020 | 15.576  | 25.409  | 0.557  |
| 3/14/2020 | 15.785  | 29.166  | 0.696  |
| 3/15/2020 | 15.934  | 35.128  | 0.910  |
| 3/16/2020 | 16.078  | 41.186  | 1.149  |
| 3/17/2020 | 16.242  | 46.210  | 1.542  |
| 3/18/2020 | 16.424  | 52.033  | 1.986  |
| 3/19/2020 | 16.720  | 60.206  | 3.198  |
| 3/20/2020 | 16.890  | 67.770  | 4.372  |
| 3/21/2020 | 17.177  | 78.363  | 5.964  |
| 3/22/2020 | 17.368  | 88.485  | 8.239  |
| 3/23/2020 | 17.493  | 99.820  | 11.907  |
| 3/24/2020 | 17.642  | 105.577  | 14.116  |
| 3/25/2020 | 17.837  | 114.246  | 16.784  |
| 3/26/2020 | 18.040  | 122.850  | 21.028  |
| 3/27/2020 | 18.218  | 133.095  | 26.138  |
| 3/28/2020 | 18.503  | 142.854  | 31.866  |
| 3/29/2020 | 18.708  | 152.720  | 37.946  |
| 3/30/2020 | 18.860  | 161.336  | 43.676  |
| 3/31/2020 | 19.104  | 168.631  | 50.374  |
| 4/1/2020 | 19.301  | 175.273  | 57.664  |
| 4/2/2020 | 19.475  | 182.616  | 65.860  |
| 4/3/2020 | 19.643  | 190.325  | 77.631  |
| 4/4/2020 | 19.826  | 198.647  | 87.822  |
| 4/5/2020 | 19.984  | 205.833  | 94.890  |
| 4/6/2020 | 20.076  | 212.961  | 102.503  |
| 4/7/2020 | 20.168  | 218.905  | 113.893  |
| 4/8/2020 | 20.271  | 224.875  | 121.723  |
| 4/9/2020 | 20.347  | 230.259  | 131.188  |
| 4/10/2020 | 20.400  | 237.202  | 142.666  |
| 4/11/2020 | 20.459  | 243.727  | 152.911  |
| 4/12/2020 | 20.521  | 251.479  | 160.793  |
| 4/13/2020 | 20.570  | 258.238  | 169.137  |
| 4/14/2020 | 20.623  | 263.445  | 178.632  |
| 4/15/2020 | 20.675  | 268.353  | 186.810  |
| 4/16/2020 | 20.718  | 272.758  | 197.783  |
| 4/17/2020 | 20.761  | 279.010  | 206.575  |
| 4/18/2020 | 20.796  | 284.779  | 215.470  |
| 4/19/2020 | 20.812  | 290.545  | 224.926  |
| 4/20/2020 | 20.837  | 295.577  | 232.701  |
| 4/21/2020 | 20.855  | 299.303  | 240.939  |
| 4/22/2020 | 20.876  | 303.810  | 250.723  |
| 4/23/2020 | 20.892  | 309.375  | 258.032  |

Through April 23, 2020, the cumulative confirmed cases were 10,702 in South Korea, 187,327 in Italy and 849,094 in the United States, while their corresponding COVID-19 cases (per 100, 000) were 20.892, 309.375 and 258.032 in South Korea, Italy, and the United States, respectively.

Comparing the three countries, we can find that South Korea had the highest COVID-19 cases (per 100,000), followed by Italy, and the United States had the lowest from February 18 to March 9. On March 10 the COVID-19 cases (per100,000) in Italy surpassed that in South Korea, rising rapidly.

**Model validity**



**Figure 4.** The prediction of cumulative confirmed cases in South Korea.



**Figure 5.** The prediction of cumulative confirmed cases in Italy.



**Figure 6.** The prediction of cumulative confirmed cases in the United States.

The data of 14 days in three countries were used as a training set to predict the number of cumulative confirmed cases in the next three days. We found that the number of predicted cumulative confirmed cases was very close to the number of actual cumulative confirmed cases, indicating that our model had realistic prediction