CHAPTER 3

PREDICTIVE ANALYTICS OF THE COVID-19 OUTBREAK UNDER UNCERTAINTY OF THE DISEASE SPREADING

Norhayati Rosli, Noryanti Muhammad,
Muhammad Fahmi Ahmad Zuber

ABSTRACT

COVID-19 pandemic was identified in Wuhan, China in 2019 and has spread at a tremendous rate affecting all countries over the world. Understanding the spreading disease is crucial; hence, the dynamic behaviour of the disease can be predicted. This paper is aimed to model the COVID-19 outbreak by extending the deterministic susceptible-infected-recovered-death (DSIRD) into a stochastic SIRD (SSIRD) model. Infectious rate parameter of the DSIRD model is perturbed with Brownian motion to reflect the uncertainty of the COVID-19 outbreak. Fourth order stochastic Runge-Kutta (SRK4) method is used to simulate the model. Parameter estimation is estimated using the Markov Chain Monte Carlo (MCMC) method. The simulated results for three ASEAN countries of Malaysia, Indonesia and Singapore indicate that SSIRD model is consistent with the infected COVID-19 data; hence, shows the model is adequate in explaining the behaviour of the infectious disease.

Keywords Mathematical model, COVID-19, Pandemic, Stochastic Runge-Kutta, Markov Chain Monte Carlo

---

Norhayati Rosli, Noryanti Muhammad, and Muhammad Fahmi Ahmad Zuber
Centre for Mathematical Sciences, Universiti Malaysia Pahang, Lebuhraya Tun Razak, 26300 Gambang, Pahang

Noryanti Muhammad
Centre of Excellence for Data Science & Artificial Intelligence, Universiti Malaysia Pahang, Lebuhraya Tun Razak, 26300 Gambang, Pahang

Norhayati Rosli (Corresponding Author)
e-mail: norhayati@ump.edu.my

Faculty of Computing, UMP Research Book Series: Emerging Technologies during the Era of Covid-19 Pandemic,
Vol. 1, [insert doi here later]

INTRODUCTION

Coronavirus disease (COVID-19) that was first detected in Wuhan, China in December 2019 has resulted in a tremendous disaster worldwide. There have been 526,182,662 confirmed cases of COVID-19, including 6,286,057 deaths till May 2022 (World Health Organisation, 2022, May 31). Most of the infected patients have experienced mild to moderate symptoms and have recovered without special treatment. However, some of the infected people become seriously ill and need medical attention. Mathematical models are important tools to gain understanding in informing evidence-based decisions (Norris et al., 2018). It is crucial for researchers and practitioners to develop a mathematical model of the COVID-19 outbreak and predict the severity of the epidemic trajectories that can alarm public health and the severe impacts on the healthcare system. Dynamics of the disease over time provide an overview of the epidemiological situation and identify whether epidemic control measures have a measurable impact (Bick et al., 2021). With regards to the studies carried out so far, numerous researchers have reported the deterministic model about the transmission of the pandemics. For the incomplete list are Wang et al., 2020; Imai et al., 2020; Anastassopoulou et al., 2020; Xiong & Yan, 2020; Singh & Adhikari, 2020; Azar et al., 2020; Tang et al., 2020a, 2020b; Roda et al., 2020; Leung et al., 2020; You et al., 2020; Yu, X., 2020; Zhuang et al., 2020; Zhao et al., 2020.

The deterministic classical compartmental mathematical models such as Susceptible-Exposed-Infected-Remove (SEIR) and Susceptible Infected-Remove (SIR) are frequently employed to describe the disease outbreak. While these models are useful, the models did not include the influence of stochasticity. Many individuals within the populations are assumed to react unpredictably in response to the pandemic (David et al. 2022). Hence, there is a need to consider stochastic models in capturing these effects.

This paper is aimed to model the COVID-19 outbreak by extending the deterministic susceptible-infected-recovered-death

(DSIRD) into a stochastic SIRD (SSIRD) model. Infectious rate parameter of the DSIRD model is perturbed with Brownian motion to reflect the unpredictable response of the susceptible populations to the COVID-19. Fourth order stochastic Runge-Kutta (SRK4) method is used to simulate the model. Then, the parameter estimation is estimated using MCMC method. The reproduction number is calculated for the control measures imposed within a certain period.

MATHEMATICAL MODELLING

DSIRD is developed by converting the disease flow diagram (Figure 1) into a system of ordinary differential equations (ODEs).

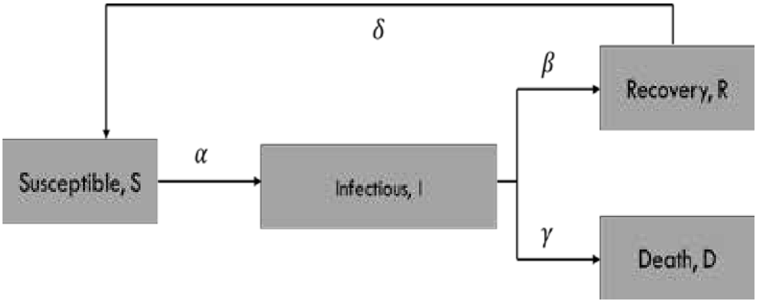

Figure 1: Diagram of the Disease Flow Compartment.

The population is divided into four compartments i.e., susceptible (*S*), infected (*I*), recovery (*R*) and death (*D*). The entire population will initially become susceptible (*S*) (exposed to the disease) once the outbreak starts. The disease spreading amongst susceptible people is subjected to the uncertainty and depending on the obedience of the citizen to the control measure imposed by the authority. The infectious rate is high if there is inconsistent compliance with the rules. At the initial time, the number of susceptible people is measured by taking the difference between the number of populations, *N,* and the infected people. The population is assumed to be constant, which is there is no change in the birth and death rate. The susceptible people at initial time,

t_0, is denoted by S_0. Initial infected people is given by I_0. Susceptible people is removed from susceptible compartment at a rate of $\frac{\alpha}{N}$ once they got infected. It means that the number of infectious people is increasing at a rate of $\frac{\alpha}{N}$ and has been isolated as an infected person, $I(t)$. Once recovered, they are moved to a recovery compartment, $R(t)$, with a rate of β. $D(t)$ is the number of the death person at time t. Death person is removed from the infected compartment with a rate of γ. Those who recovered could be reinfected and is transformed to the susceptible compartment with a rate of ∂. Individual is assigned to one of the following disease states (susceptible or infected or recovery or death) at a time, *t.* DSIRD model describing the compartmental diagram of the disease flow in Figure 1 is given by

$$\frac{dS}{dt} = -\frac{\alpha IS}{N} + \delta R$$
$$\frac{dI}{dt} = \frac{\alpha IS}{N} - \beta I - \gamma I$$
$$\frac{dR}{dt} = \beta I - \delta R$$
$$\frac{dD}{dt} = \gamma I \tag{1}$$

Infectious rate parameter, α, is subjected to the uncontrolled factors of the inconsistent compliance of the public to the control measures imposed by the government and the emergence of the new variant. The DSIRD model (1) can be transformed into SSIRD by perturbing the infectious rate parameter α with a Brownian motion, $W(t),\ t \geq 0$, such that

$$\frac{\alpha}{N} \to \frac{\alpha}{N} + \frac{dW}{dt} \tag{2}$$

This led to a system of SSIRD model, as given in Equation (3).

$$
\begin{aligned}
dS &= \left(-\frac{\alpha IS}{N} + \delta R \right) dt + \left(-\frac{\alpha}{N} \right) IS\, dW \\
dI &= \left(\frac{\alpha IS}{N} - \beta I - \gamma I \right) dt + \frac{\alpha}{N} IS\, dW \\
dR &= \left(\beta I - \delta R \right) dt \\
dD &= \gamma I dt
\end{aligned}
\qquad (3)
$$

The Brownian motion, $W(t),\ t \geq 0$, can be generated using pseudo random numbers of Box-Muller method. The epidemiology parameters of Equation (3) need to be estimated and the model is simulated by using SRK4 method. Next subsection shows the MCMC and SRK4 methods of the parameter estimation and model simulation, respectively.

Parameter estimation

The epidemiology parameters of SSIRD in Equation (3) are estimated using the Markov Chain Monte Carlo (MCMC) method of the Metropolis-Hasting (MH) algorithm. MH algorithm considers two types of distribution i.e. the stationary distribution, $\pi(x)$, and the conditional density, $Q(y|x).$ It generates a collection of states based on the desired distribution, $P(x).$ MH algorithm utilises the concept of Markov process, such that $\pi(x) = P(x).$ If the current state of the chain is $x,$ the chain will then move to a new candidate of the sample $y.$ The following algorithm calculates the transition value of the Markov chain (x_t) at time t and its value at time $t+1.$

```
Metropolis-Hasting Algorithm
Initialization.
    1. Choose an initial state of $x_0$.
    2. Let the initial time, $t = 0$.
Perform Iteration
    1. Generate a random candidate state, $y$ from the conditional density $Y_t \sim Q(y|x_t)$.
    2. Calculate the probability of acceptance
           $p(x,y) = \min\left\{\frac{P(y)}{P(x_t)}\frac{Q(x_t|y)}{Q(y|x_t)},1\right\}$.
    3. Generate a uniform random number of $\rho \in [0,1]$.
        i) If $\rho \le p(x,y)$, then accept the new state and set $x_{t+1} = y$.
        ii) If $\rho > p(x,y)$, then reject the new state and set $x_{t+1} = x_t$.
    4. Increase time by setting $t = t+1$.
```

The reported COVID-19 ASEAN data, which date from 25 January 2020 to 24 January 2021, are fitted to the SSIRD (3) for parameter estimation. The data are obtained from Asia COVID-19 tracker portal. Three different ASEAN countries are considered, involving Malaysia, Indonesia and Singapore.

Numerical algorithm

The results are simulated using four stage stochastic Runge-Kutta method (SRK4) of order 1.5. SRK4 is coded in MATLAB and Graphical User Interface (GUI) of SSIRD model is developed. SRK4 method is given by

$$Y_i = y_n + \Delta \sum_{j=1}^{4} a_{ij}^{(0)} f\left(Y_j\right) + \sum_{j=1}^{4} \left(b_{ij}^{(1)} J_1 + b_{ij}^{(2)} \frac{J_{10}}{\Delta} \right) g\left(Y_j\right), \quad i = 1, \ldots, 4$$

$$y_{n+1} = y_n + \Delta \sum_{i=1}^{4} \alpha_i^{(0)} f\left(Y_i\right) + \sum_{i=1}^{4} \left(\gamma_i^{(1)} J_1 + \gamma_i^{(2)} \frac{J_{10}}{\Delta} \right) g\left(Y_i\right) \tag{4}$$

where Δ is a step size, n is the number of iteration, $\mathbf{A} = \left(a_{ij}\right)_{4\times4}$ and $\mathbf{B} = \left(b_{ij}\right)_{4\times4}$ are matrices of real elements, and $\alpha^T = \left(\alpha_1, \alpha_2, \alpha_3, \alpha_4\right)$ and $\gamma^T = \left(\gamma_1, \gamma_2, \gamma_3, \gamma_4\right)$ are row vectors in $\Re^4$.

The stochastic components are generated by the terms

$$J_1 = \int_{t_n}^{t_{n+1}} dW = W\left(t_{n+1}\right) - W\left(t_n\right) \quad \text{and} \quad J_{10} = \int_{t_n}^{t_{n+1}} \int_{t_n}^{t} dW(s)\, dt.$$

The random variable J_{10} is approximated via $\frac{J_{10}}{\Delta} = \frac{\sqrt{\Delta}}{2}\left(N_1 + \frac{N_2}{\sqrt{3}} \right)$

where N_1 and N_2 are standard normal distribution. SRK4 scheme of strong order of 1.5 has been introduced by Burrage and Burrage (1999) and can be written in tableaus form as

$$\mathbf{A} = \begin{bmatrix} 0 & & & \\ \frac{1}{2} & 0 & & \\ 0 & \frac{1}{2} & 0 & \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad \mathbf{B}^{(1)} = \begin{bmatrix} 0 & & & \\ -0.72429163 & 0 & & \\ 0.042373534 & -0.19944379 & 0 & \\ -1.5784755 & 0.84010034 & 1.73837510 & 0 \end{bmatrix}$$

$$\mathbf{B}^{(2)} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 2.700200410 & 0 & 0 & 0 \\ 1.75261649 & 0 & 0 & 0 \\ -2.918524118 & 0 & 0 & 0 \end{bmatrix}, \quad \alpha^T = \begin{bmatrix} \frac{1}{6} & \frac{1}{3} & \frac{1}{3} & \frac{1}{6} \end{bmatrix}$$

$$\gamma^{(1)T} = \begin{bmatrix} -0.78007 & 0.073637 & 1.4865 & 0.21992 \end{bmatrix}$$

$$\gamma^{(2)T} = \begin{bmatrix} 1.69395 & 1.63610 & -3.02400 & -0.306049 \end{bmatrix}$$

The numerical algorithm governing SRK4 method is presented in Figure 2.

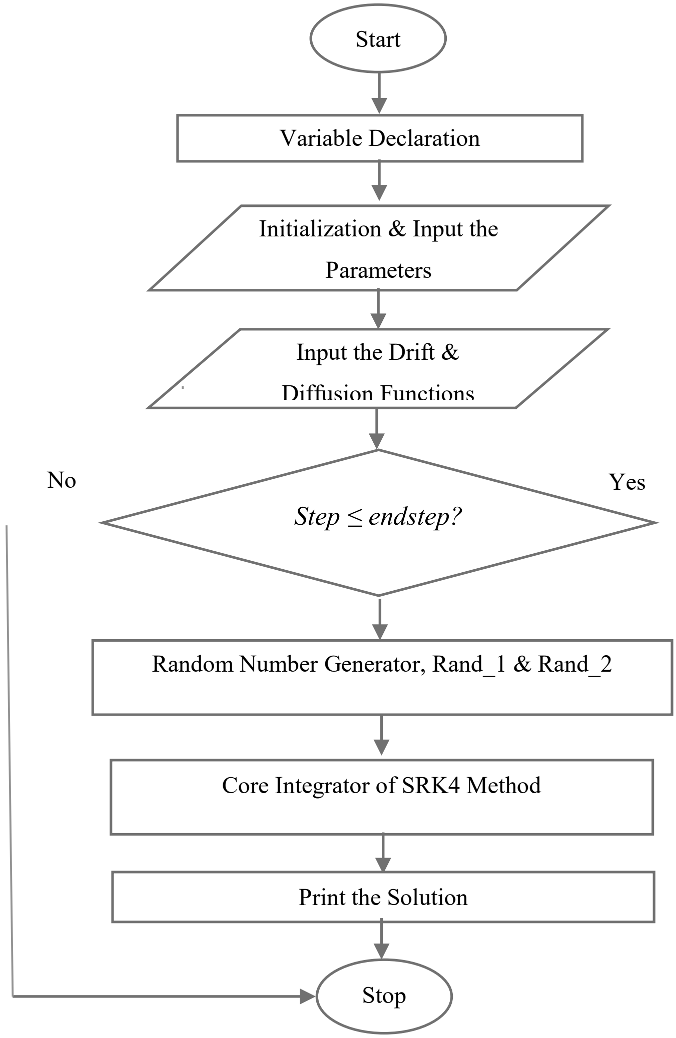

Figure 2: Algorithm of SRK4 Method.

RESULTS AND DISCUSSION

In this section, we present the numerical simulation of SSIRD model (3) for three ASEAN countries; Malaysia, Indonesia and Singapore. SSIRD model in equation (3) is simulated for 100 simulations via SRK4 method and the epidemiological parameters are estimated using MCMC method in MATLAB. COVID-19 data and the pandemic curves are visualised in Figure 3 to 5. The simulated results for Malaysia COVID-19 data are illustrated in Figure 3.

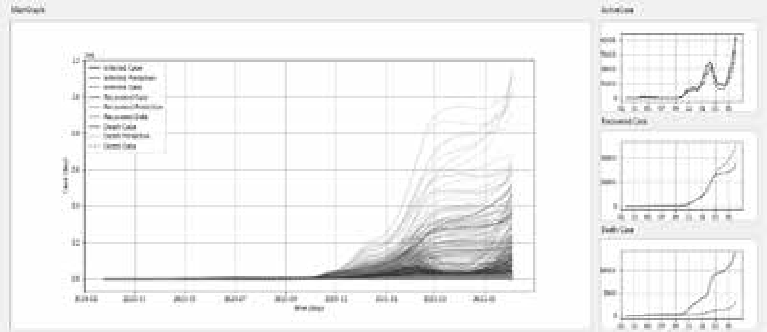

Figure 3: GUI of SSIRD model for Malaysia
(January 2020 – May 2021).

The numerical results are obtained by performing 100 simulations of the SSIRD model in Equation (3), as indicated in the “Main Graph” of Figure 3. The average of the sample paths is computed, and the results are displayed in “Active Case”, “Recovered Case” and “Death Case”. Malaysia manages to control the disease outbreak from January to September 2020. During this period, various non-pharmaceutical intervention (NPI) measures were implemented, including movement control order (MCO), not permitting interstate travel, closure of the non-essential economic activities and shorter operating hours for certain sectors with limited number of workers. However, once certain NPI is lifted, the number of cases has been increased. The National COVID-19 Immunisation Programed started in February 2021. In May 2021, 80% of the herd immunity is yet to be achieved while certain measures have been lifted. This gives rise to the number of active cases in May 2021.

The numerical results for Indonesia COVID-19 data are displayed in Figure 4. Indonesia has the highest number of populations in ASEAN, with the reported population in 2020 is 273.5 million. The COVID-19 trajectories for Indonesia are obtained by performing 100 simulations of SSIRD model (3), as indicated in the "Main Graph" of Figure 4. The average of the sample paths is computed, and the results are depicted in "Active Case", "Recovered Case" and "Death Case" of Figure 4.

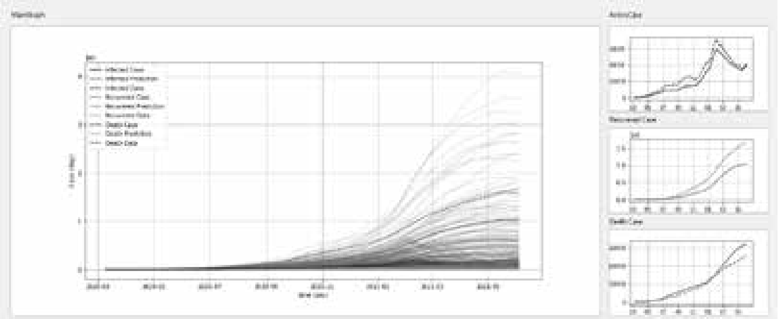

Figure 4: GUI of SSIRD model for Indonesia (March 2020 – May 2021).

Based on Figure 4, we can see that the number of infected cases increased gradually for the period of March 2020 to November 2020. However, from December 2020 to February 2021, the trajectory shows a drastically increasing trend in the number of infected cases. Indonesia has implemented large scale social restrictions, which include closing of the schools and public places, limiting travel to and from restricted provinces and regions, and restricting public transport. The active cases were decreased from January to May 2021, when Indonesia government for a certain region, such as Java and Bali, had enforced the Communities Activities Restrictions Enforcement (CARE) to prevent the disease from spreading. COVID-19 vaccination programme in Indonesia has started earlier than Malaysia, which is on 13 January 2021. In terms of the total doses given, Indonesia has ranked third in ASEAN and fifth in the world.

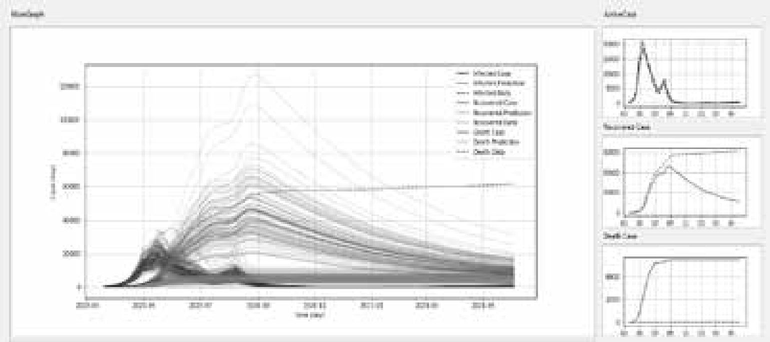

Figure 5: SSIRD model (3) for Singapore
(March 2020 – May 2021).

The numerical results for Singapore COVID-19 data are displayed in Figure 5. The number of populations in Singapore that was reported in 2020 is 5.686 million. The simulated result for Singapore COVID-19 outbreak is obtained by performing 100 simulations of SSIRD model in Equation (3), as indicated in the "Main Graph" of Figure 5. The average of the sample paths is computed, and the results are depicted in "Active Case", "Recovered Case" and "Death Case" of Figure 5.

The Singapore government is able to manage the spreading of the virus in the community through a contact tracing system. However, the number of infectious cases increased exponentially from March to June 2020, when the virus infected migrant workers. Infectious cases keep rising due to the in-dormitory living conditions for foreign workers and safe distancing measures turn out to be inadequate. Singapore manages to control the disease outbreak by imposing a lockdown in the locality of the foreign workers. At the same time, massive testing has been stepped up to detect and isolate positive cases. The measures imposed by the Singapore government had successfully reduced the number of positive cases amongst migrant workers.

CONCLUSION

The COVID-19 outbreak is uncontrolled and subjected to uncertainty. The model should consider these factors by perturbing the Brownian motion into the deterministic model DSIRD. In this paper, we have extended the DSIRD model into its stochastic counterpart and the epidemiological parameters were estimated using MCMC method. The pandemic curves of three ASEAN countries, which include Malaysia, Indonesia and Singapore, have been simulated. In summary, the proposed stochastic SIRD model shows that the proposed model is satisfactory and pertinent to be used for predicting the COVID-19 outbreak under uncertainty of the disease spreading. Presently, the performance of the model is investigated in detail and another scenarios compartment is considered to improve the model.

ACKNOWLEDGEMENT

The authors would like to thank the Ministry of Higher Education for providing financial support under Fundamental research grant No. FRGS/1/2019/STG06/UMP/02/2 (University reference RDU1901139) and Universiti Malaysia Pahang for additional financial support under Postgraduate Research Grant PGRS210327

REFERENCES

Anastassopoulou, C., Russo, L., Tsakris, A., & Siettos, C. (2020). Data-based analysis, modelling and forecasting of the COVID-19 outbreak. *PLoS ONE*, *15*(3). https://doi.org/10.1371/journal.pone.0230405

Azar, K. M. J., Shen, Z., Romanelli, R. J., Lockhart, S. H., & ... (2020). Disparities In Outcomes Among COVID-19 Patients In A Large Health Care System In California: Study estimates the COVID-19 infection fatality rate at the US county …. *Health* ….

Burrage, K. and Burrage, P., (1996). High strong order explicit runge-kutta methods for stochastic ordinary differential equations. Applied Numerical Mathematics, 22(1), 81-101. https://doi.org/10.1016/S0168-9274(96)00027-X

Bick, A., Blandin, A., & Mertens, K. (2021). Work from Home Before and After the COVID-19 Outbreak. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3786142

Imai, N., Shang, Y., Inthavong, K., Tu, J., Dorigatti, I., Cori, A., Riley, S., Ferguson, N. M., Xiao, Q., Cetto, R., Doorly, D. J., Bates, A. J., Rose, J. N., Mcintyre, C., Comerford, A., Madani, G., Tolley, N. S., & Schroter, R. (2020). Report 2: Estimating the potential total number of novel Coronavirus cases in Wuhan City, China. *Annals of Biomedical Engineering*, *48*(January).

Leung, K., Wu, J. T., Liu, D., & Leung, G. M. (2020). First-wave COVID-19 transmissibility and severity in China outside Hubei after control measures, and second-wave scenario planning: a modelling impact assessment. *The Lancet*, *395*(10233). https://doi.org/10.1016/S0140-6736(20)30746-7

Norris, S. L., Egger, M., Johnson, L., Althaus, C., Schöni, A., Salanti, G., & Low, N. (2018). Developing WHO guidelines: Time to formally include evidence from mathematical modelling studies. *F1000Research*, *6*. https://doi.org/10.12688/f1000research.12367.2

Roda, W. C., Varughese, M. B., Han, D., & Li, M. Y. (2020). Why is it difficult to accurately predict the COVID-19 epidemic? Infectious Disease Modelling, 5. https://doi.org/10.1016/j.idm.2020.03.001

Singh, R., & Adhikari, R. (2020). Age-structured impact of social distancing on the COVID-19 epidemic in India. http://arxiv.org/abs/2003.12055

Tang, B., Wang, X., Li, Q., Bragazzi, N. L., Tang, S., Xiao, Y., & Wu, J. (2020a). Estimation of the Transmission Risk of 2019-nCov and Its Implication for Public Health Interventions. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3525558

Tang, B., Wang, X., Li, Q., Bragazzi, N. L., Tang, S., Xiao, Y., & Wu, J. (2020b). Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions. *Journal of Clinical Medicine*, *9*(2). https://doi.org/10.3390/jcm9020462

Wang, W., Wu, Q., Yang, J., Dong, K., Chen, X., Bai, X., Chen, X., Chen, Z., Viboud, C., Ajelli, M., & Yu, H. (2020). Global, regional, and national estimates of target population sizes for covid-19 vaccination: Descriptive study. *The BMJ*, *371*. https://doi.org/10.1136/bmj.m4704

World Health Organisation. (2021, March 11). WHO Coronavirus (COVID-19) Dashboard. https://covid19.who.int/

Xiong, H., & Yan, H. (2020). Simulating the Infected Population and Spread Trend of 2019-nCov Under Different Policy by EIR Model. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3537083

You, C., Deng, Y., Hu, W., Sun, J., Lin, Q., Zhou, F., Pang, C. H., Zhang, Y., Chen, Z., & Zhou, X. H. (2020). Estimation of the time-varying reproduction number of COVID-19 outbreak in China. *International Journal of Hygiene and Environmental Health*, *228*. https://doi.org/10.1016/j.ijheh.2020.113555

Yu Professor, X. (2020). Updated Estimating Infected Population of Wuhan Coronavirus in different policy scenarios by SIR Model.

Zhuang, Z., Zhao, S., Lin, Q., Cao, P., Lou, Y., Yang, L., Yang, S., He, D., & Xiao, L. (2020). Preliminary estimates of the reproduction number of the coronavirus disease (COVID-19) outbreak in Republic of Korea and Italy by 5 March 2020. *International Journal of Infectious Diseases*, *95*. https://doi.org/10.1016/j.ijid.2020.04.044