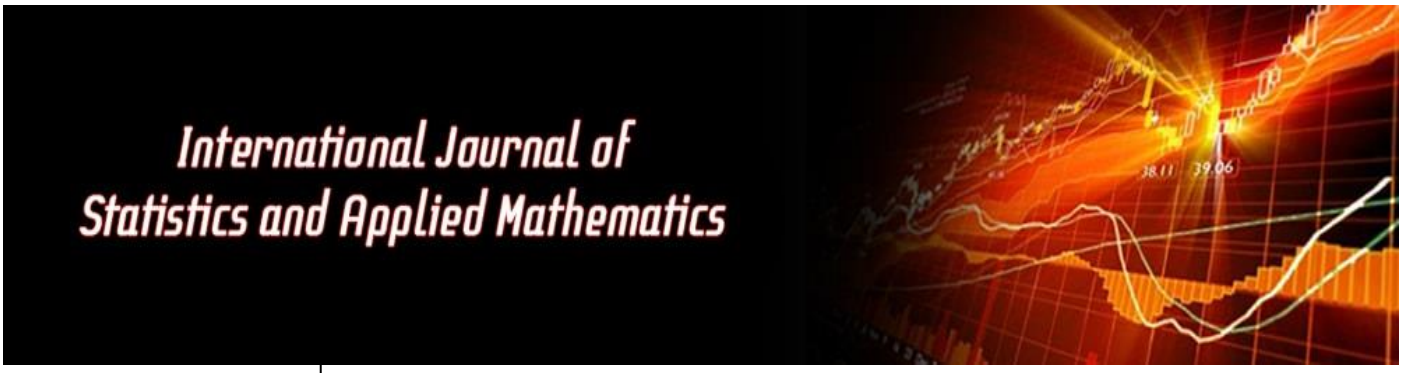

ISSN: 2456-1452
Maths 2020; 5(5): 20-24

www.mathsjournal.com
Received: 12-07-2020
Accepted: 14-08-2020

Martin Nyamu
Institute of Tropical and Infectious Diseases, University of Nairobi, Kenya

Benard Daniel Mutwiri
Institute of Tropical and Infectious Diseases, University of Nairobi, Kenya

Babu Lawrence
Data Manager Kemri Kisumu Kenya

Corresponding Author:
Benard Daniel Mutwiri
Institute of Tropical and Infectious Diseases, University of Nairobi, Kenya

Modelling of COVID-19 (Coronavirus) in Kenya Using SIR Model

Martin Nyamu, Benard Daniel Mutwiri and Babu Lawrence

DOI: https://doi.org/10.22271/maths.2020.v5.i5a.579

Abstract
The Novel Coronavirus (COVID-19), which originated from Wuhan city in Central China late 2019 is still spreading rapidly worldwide with the current infections standing at 16,899, 326 cases and 663, 540 deaths as at 29th July 2020 as captured in the World meter. Coupled with a rampant transmissibility rate, the virus has taken advantage of globalization and interconnectedness of the world as opposed to the previous years' pandemics and claimed a number of lives in a very short span. Indeed, despite the significant interventions and containment measures imposed by different nations, the virus seems to have exponentially doubled its prevalence with the peak yet to register in numerous countries, including Kenya. Therefore, the aim of this article was to model the COVID-19 (Coronavirus) cases in Kenya using an epidemiological model (SIR). Specifically, the article sought to expound the use of differential equations in infectious disease modelling, as well as, the importance of herd immunity and need for an urgent coronavirus vaccination. The study used a simple SIR model to predict and model the spread of Coronavirus in Kenya. The model was fitted using data of the reported incidences of coronavirus in Kenya (dating 13th March, 2020 to 21st July, 2020), extracted from the {coronavirus} package in R developed by Rami Krispin. Based on the analysis findings, the graphs postulated that the number of cases estimated by the SIR model slightly deviated from the exponential growth, though evidenced a significant exponent rise in the coronavirus cases in Kenya from March to July. Moreover, with a reproduction number of 1.20, the model suggested that about 16.67% of the population requires vaccination to stop the spread of the infection. However, with no intervention or measures to deter the spread of the pandemic, the peak in Kenya is expected to be reached by 30th of September with approximately 443,720 infections and 7,810 deaths (given a fatality rate of 1.8%). Overall, the findings of this study offer insight to the government of Kenya to impose appropriate and relevant measures and policy that would stop the spread of the virus to a larger population. Additionally, the model creates a public knowhow on the importance of the currently imposed containment measures and regulations to reduce the spread of viruses in Kenya.

Keywords: Coronavirus (Covid-19), SIR model, herd immunity, reproduction number (R_0), pandemic, parameter, kenya

1. Introduction
Despite significant universal investment in health prospects, global epidemics have been catastrophic and a threat to human life for several decades. For instance, the Black Death wiped-out about a third of the total population in Medieval Europe, and in 1918 Spanish Flu wiped more people than those who died during the First World War (Trilla *et al.*, 2008) [1]. In November 2002 to July 2003, SARS (COVID-2), a coronavirus that originated from Beijing, China, attacked various countries infecting and killing a significant number of people across the globe (World Health Organization, 2016) [13]. Additionally, the Swine Flu pandemic frightened the whole world, while Ebola and Zika virus outbreaks are still severe problems currently (Ryu, 2017) [7]. In 2012, Middle East Respiratory Syndrome (MERS-Cov) caused by novel Coronavirus killed 858 people from a proportion of 2,494 infected individuals; thus, recording a fatality rate of 34.4% (World Health Organization, 2016; Assiri, 2013) [13, 1]. Seemingly, the Novel Coronavirus (COVID-19), which originated from Wuhan city in Central China late in 2019 is still spreading rapidly worldwide with the current infections standing at 16,899, 326 cases and 663, 540 deaths as at 29th July 2020 (Worldometer, 2020) [14]. Coupled with a rampant transmissibility rate, the virus has taken advantage of globalization and interconnectedness of the world as opposed to the previous years' pandemics and claimed a

number of lives in s very short span. Indeed, despite the significant interventions and containment measures imposed by different nations, the virus seems to have exponentially doubled its prevalence with the peak yet to register in numerous countries, including Kenya. Since January 2020, a considerable number of global scientists have been analyzing the spread of coronavirus from different views, and with different strategies and technologies to develop the solution and mitigate its effects on the citizens.

Currently, mathematicians, statisticians, epidemiologists, and data scientists are at the forefront working to understand how the virus is spreading. The analytical commitments to predicting the virus trend is critical to aiding heath community; doctors, nurses, and governments, in dealing with the issues of epidemics. The availed statistical insights significantly inform optimal health decisions regarding the virus containment measures based on the notable prevalence rate. As such the main purpose of this article is to model the COVID-19 (Coronavirus) cases in Kenya using an epidemiological model (SIR). Specifically, the article sought to expound the use of differential equations in infectious disease modelling, as well as, the importance of herd immunity and need for an urgent coronavirus vaccination.

2. Methods

Data

The study employed the current Kenyan population based on World meter elaboration for the latest United Nations data (July 21, 2020) as the initial uninfected population (N) source. A vector was generated with the daily cumulative incidence for Kenya, from March 13, 2020 (when our daily incidence data started), through July 21. The study used a simple SIR model to predict and model the spread of Coronavirus in Kenya. The model was fitted using data of the reported incidences of coronavirus in Kenya (dating 13th March, 2020 to 21st July, 2020), extracted from the {coronavirus} package in R developed by Rami Krispin. The statistical analysis tool, RStudio, was used to model the data using simple SIR model (deterministic compartmental model), which divides the population into three compartments (Susceptible - Infected - Recovered) where:

- Susceptible (S) (not infected),
- Infectious (I), and
- Recovered (R) (that is, vaccinated or recovered with immunity).

In this case, the incidences predicted from SIR model were compared with the actual incidences dating 13th March, 2020. As such, the values were initialized as; total current population (N), Susceptible (S), Infected (I), and recovered (R).

2.1 The SIR Model and Its Optimal Parameters Values

The SIR model represents how an infection would spread through a population since it takes into consideration that some people will recover from the disease and never become susceptible. The SIR model assumes that individuals who recover from the infection become immune and cannot be infected again. These groups evolve as the virus continue to spread, that is, susceptible group S declines when people are infected and move to the infectious group I. As individuals recover or die, they move to the recovered group R. With the SIR model, as such the three differential equations were required to model the coronavirus outbreak dynamics. The transmission rate, β (beta), controls the transition from S to I, and the recovery rate, γ (gamma), controls the transition between I and R.

The following differential equations was used with SIR Model without demographic characteristics.

$$\frac{dS}{dt} = -\lambda S \quad \dots(i)$$
$$\frac{dI}{dt} = \lambda S - \gamma I \quad \dots(ii)$$
$$\frac{dR}{dt} = \gamma I \quad \dots(iii)$$

The first equation (i) indicates that the number of susceptible individuals (S) decreases with the number of newly infected individuals, where newly infected cases are the product of infection rate (β) and the number of susceptible individuals (S) who had contact with infectious individuals (I). The second equation (ii) suggests that the number of infectious individuals (I) increases with the newly infected individuals (βIS), minus the previously infected people who recovered (i.e., γI), which is the removal rate, multiplied by the infectious individuals (I).

Finally, the last equation (iii) states that the number of individuals in the recovered group (R) goes up with the increase in the recovery individuals from the infectious group (I) who either die or recover.

2.1.1 The SIR Model Assumptions

- The number of the infected individual rises at a rate proportional to both the infected and susceptible number of individuals. The number of susceptible people decreases at this same rate. The ratio involved in the transmission rate β (beta), is the same as in the SI model.
- The recovered individuals rise at a rate proportional to the infected number of individuals. The ratio involved is called the recovery rate, γ(gamma).
- A susceptible person who catches the disease becomes infectious immediately. The differential equations were first expressed as R functions with respect to the time before fitting the SIR model.

2.2 Optimal Values of the Unknown Parameters

Optimal values for the two unknown parameters β and γ in the model were obtained using an optimizer and a solver for the differential equations. The ordinary differential equations were solved using ode () function from {deSolve} package in R. And the optimal values were estimated by using optim () function from base R. The sum of squared differences between the infectious individuals (I) at time t and estimated individuals $\hat{I}(t)$ was minimized to obtain the residual sum of squares (RSS):

$$RSS(\beta, \gamma) = \sum_t (I(t) - \hat{I}(t))^2$$

The function to calculate the RSS was defined, given a set of values for β and γ. Finally, the SIR model was fitted to the data by finding the values for β and γ that reduces the residual sum of squares between the observed cumulative incidence observed in Kenya and the predicted cumulative incidence. Moreover, the function to calculate the RSS was defined, given a set of values for β and γ followed by the SIR model fit, where the values for β, and γ that reduces the residual sum of squares between the observed cumulative incidence observed in Kenya and the predicted cumulative

incidence were estimated. Besides, the model convergence was checked before fitting. The analysis results were summarized in both visual graphical and tabular summary statistics as presented in the next section.

3.0 Results and Discussion

Test for Model Convergence and Parameter Estimates

The convergence of the model was confirmed. Different estimates may be obtained for different choices for the initial values of the model. The aspect indicate that the fitting process is not stable; thus, a potential solution for a better fitting process is essential. From analysis of the model estimates, the fitted values for β and γ were 0.573 and 0.501 respectively. It was noted that the transition between S and I (transmission rate) controls the transition between I and R (Recovery rate). Nevertheless, those values do not have much meaning, though they are used to obtain the fitted number of people in each compartment of our SIR model for the dates to July 21 and then compare the fitted versus the observed data.

3.1 Graphical Representation and Prediction

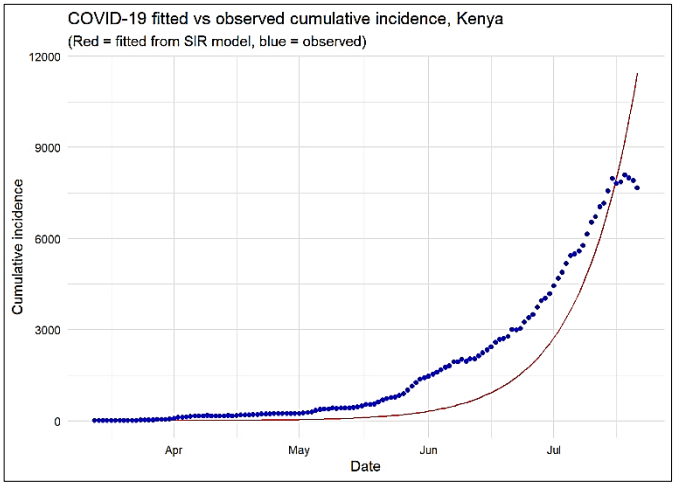

Fig 1: A graph of COVID-19 fitted versus observed cumulative incidence in Kenya

As shown in figure 1 above, the graph indicated that the number of observed confirmed cases slightly deviated from the number of confirmed cases expected by our model. This postulated that the pandemic in Kenya was growing exponentially from March to July. However, as time goes, the trend of the data will not follow the exponential phase in the long term as coronavirus continues to spread across all the counties in Kenya.

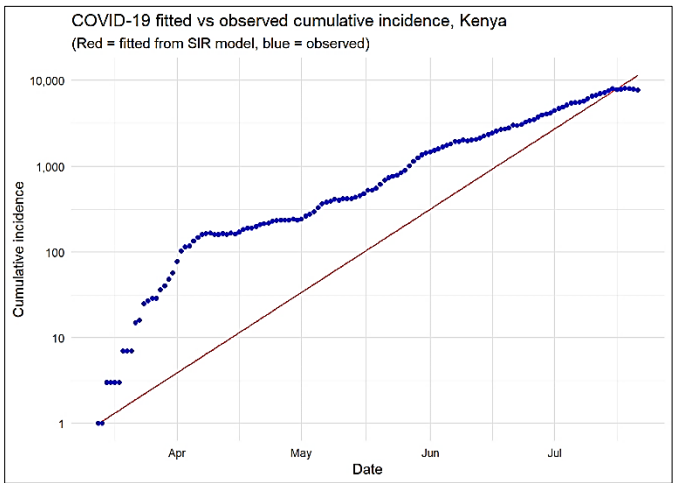

Fig 2: A graph of log scale of COVID-19 fitted versus observed Cumulative incidence

The graph above (Figure 2), is on a log scale on the y-axis called a semi-log plot or log-linear plot since only the y-axis is transformed with a logarithm scale. As such, this is essential to transform the variable into a log scale since it gives a more easily readable graph in terms of the disparity between the observed and expected number of confirmed cases, as well as, how the observed cases differ from the exponential trend. Seemingly, the plot suggested that at the beginning of the pandemic to July 2020, the number of confirmed cases stayed above what would be expected in an exponential phase. In particular, the number of confirmed cases increased day after day from March 13 and sustained the increment rate above an exponential rate to date (July 2020).

3.2 Reproduction Number (R_0)

The SIR model looked like a good fit to the observed cumulative incidence data in Kenya; thus, the fitted model was used to calculate the basic reproduction number R_0. The latter is also referred as basic reproduction ratio, and it depends on transmission rate (β) and recovery rate (γ). It is a mathematical term that indicates how contagious an infectious disease is. The R_0 gives the average number of people who will contract a contagious disease from one person with that disease (Cori *et al.*, 2013) [2]. The reproduction number:

- $R_0 < 1$ means each existing infection causes less than one new infection. In this case, the disease will decline and eventually die out.
- $R_0 = 1$ means each existing infection causes one new infection. The disease will stay alive and stable, but there won't be an outbreak or epidemic.
- $R_0 > 1$ means each existing infection causes more than one new infection. The disease will be transmitted between people, and there may be an outbreak or epidemic.

In the SIR model, the basic reproduction number R_0 can be calculated using the ratio of the transmission rate to the recovery rate:

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

The analysis revealed a reproduction number of 1.2 from the current cases of coronavirus in Kenya. The reproduction number of 1.2 was below values found by others for COVID-19 and for SARS and MERS, which are similar diseases also caused by a coronavirus (Liu *et al*., 2020) [6]. Moreover, in the literature, preliminary studies had estimated the reproduction number for COVID-19 to be between 1.5 and 3.5 (with a β close to 0.54 and γ close to 0.2) (Liu *et al*., 2020; You *et al*., 2020; Wang *et al*., 2020) [6, 15, 12]. The reproduction number (R_0) suggested that, on average, one infectious person spread the virus to about 2 uninfected individuals in Kenya. However, the available body of literature indicate that when the reproduction number goes below one, the pandemic disappears gradually (Wang *et al*., 2020) [12].

3.3 Herd Immunity

Herd immunity applies when a critical proportion of a susceptible population is immunized against a contagious disease, giving overall protection to the remainder of the unprotected community (herd). This minimizes the chance of an outbreak occurring and allows those who are not eligible for vaccines, such as infants and pregnant women, to also receive some protection from the disease (Fine *et al*., 2011) [4]. Herd immunity works because it is more difficult for diseases to spread between individuals if large numbers are already immune, as this breaks the chain of infection. The proportion of the population in need of effective immunization to prevent the sustained spread of the disease called herd immunity threshold has to be larger than $(1-\frac{1}{R_0})$ for simple models. From the reproduction number 1.2 obtained from our Kenyan data, about 16.67% the population need to be immunized to stop the spread of the infection. With an approximate of about 53.8 million Kenyan population, this translates into roughly 8.97 million people.

3.4 Prediction of Coronavirus Outbreak with no Intervention

The SIR model was employed to fit the data for the first 130 days of the available confirmed cases in Kenya to examine the effect of the outbreak if left to run its course without public health intervention.

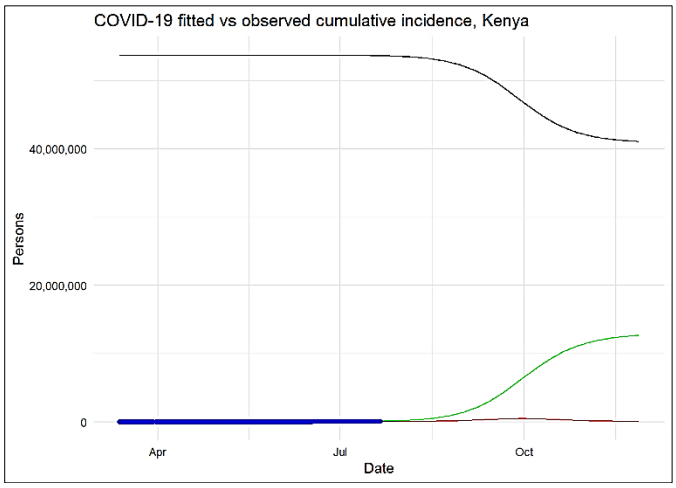

Fig 3: A Covid-19 graph when there was no intervention

The graph (Figure 3), above indicated that as the proportion of recovered people increases, the proportion of susceptible people decreases (See the green and black trend lines). The red line represents the trend of infected individuals and how their numbers change over time. Both the black and red lines are decreasing going down to the horizontal line but not yet reached zero, while the green line is decreasing at a slower rate. This means that the number of susceptible decreases as more people become infected, and the number of infected decreases as some people recover or die. More so, the graphical results suggested that the disease will die out before everyone gets infected since the red line is almost reaching zero. The graph in figure 4 below is same as figure 3 but log-scaled for the y-axis and with a legend for an explicit readability.

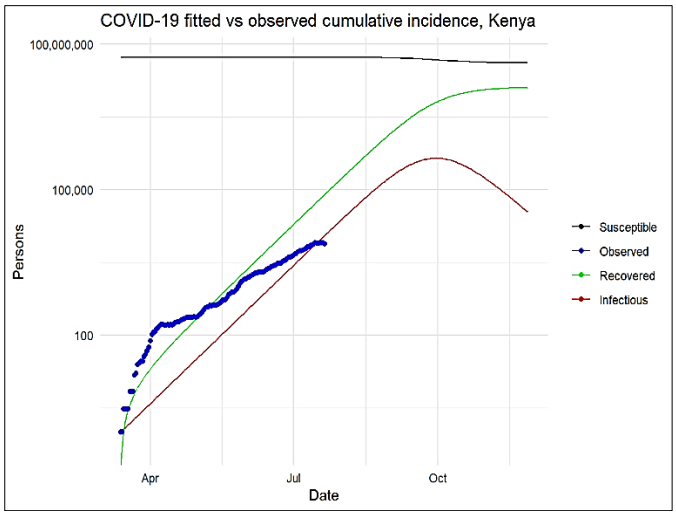

Fig 4: Log-linear graph

3.4.1 Model Deductions

The fitted SIR model alluded to interesting results pertaining the Peak of the Pandemic and the number of deaths. The predictions with the exact same settings and no intervention or measures undertaken to limit the spread of the pandemic indicated that the peak in Kenya is expected to be reached by probably 30th September 2020. In addition, the model adduced that in Kenya, if no measures or intervention is imposed, about 443,720 people will be infected by the end of September with about 7,810 deaths (with a fatality rate of 1.8%). Therefore, the above finding provides a rationale to government imposition of the current strict containment measures and regulations to reduce the spread of viruses. Indeed, the above predictions should be taken with a lot of caution, despite the limitations of the SIR model, that is, unrealistic assumptions such as no public health interventions, or strict measures imposed, fixed reproduction number extra. Nevertheless, more advanced projections can be developed with the help of {projection} package in R.

4.0 Conclusion and Recommendation

Overall, the study findings support the available body of literature and presents insight to the universal health fraternity across the globe to develop and actualize appropriate and relevant policy measures to mitigate the spread of the coronavirus. Moreover, the model alludes a need for careful and struct adherence to the proposed public health interventions by the public to evade catastrophic effects attested in the previous pandemics, such as SAR, Spanish Flu, Swine Flu, among others.

However, the model used in this study was limited to only three compartments (Susceptible, Infected, and Recovered); it did not consider the pandemic's latent period. Therefore, more sophisticated models are required to model the pandemic and reflect real-life transmission processes. Given that the

pandemic has exposed period, the SEIR model, which is an extension of the SIR model, would more appropriate. In the SEIR model, S is susceptible, E is the exposed or infected asymptomatic, I infected and symptomatic, and R is the Recovered. The SEIR model is a continuous-time dynamic model that assumes fixed transition rates. Some models allow varying transition rates, such as stochastic models that depend on individuals' characteristics and social networking. Furthermore, the SIR model used in this study assumed a constant reproduction number (R_0). Therefore, it would be pertinent to estimate the current effective reproduction number (Re) on a daily basis to keep track of the effectiveness and efficiency of public health interventions and probably estimate when the incidence curve will begin to decline. This can be achieved in a future study using {EpiEstim} package in R to estimate (Re) and allow to consider internal migration from one geographical region besides local transmission.

References

1. Assiri A, McGeer A, Perl TM, Price CS, Al Rabeeah AA, Cummings DA *et al*. Hospital outbreak of Middle East respiratory syndrome coronavirus. New England Journal of Medicine. 2013; 369(5):407-416.
2. Cori A, Ferguson NM, Fraser C, Cauchemez S. A New Framework and Software to Estimate Time-Varying Reproduction Numbers During Epidemics. American Journal of Epidemiology. 2013; 178(9):1505-1512. https://doi.org/10.1093/aje/kwt133
3. Dr. Holger K, Von Jouanne D. (2020, February 4). Epidemiology: How contagious is novel coronavirus (2019-nCoV)? Retrieved July 29, from, 2020. https://blog.ephorie.de/epidemiology-how-contagious-is-novel-coronavirus-2019-ncov
4. Fine P, Eames K, Heymann DL. Herd Immunity: A Rough Guide. Clinical Infectious Diseases. 2011; 52(7):911-916. https://doi.org/10.1093/cid/cir007
5. Khedkar PH, Patzak A. SARS-CoV-2: What do we know so far?. Acta Physiologica, 2020, e13470.
6. Liu Y, Gayle AA, Wilder-Smith A, Rocklöv J. The reproductive number of COVID-19 is higher compared to SARS coronavirus. Journal of travel medicine, 2020.
7. Ryu WS. New Emerging Viruses. Molecular Virology of Human Pathogenic Viruses, 2017, 289-302. https://doi.org/10.1016/B978-0-12-800838-6.00021-7
8. Tim C. (2020, February 18). Analysing COVID-19 (2019-nCoV) outbreak data with R - Part 1. Retrieved July 29, 2020, from https://timchurches.github.io/blog/posts/2020-02-18-analysing-covid-19-2019-ncov-outbreak-data-with-r-part-1/
9. Thompson RN, Stockwin JE, van Gaalen RD, Polonsky JA, Kamvar ZN, Demarsh PA *et al*. Improved inference of time-varying reproduction numbers during infectious disease outbreaks. Epidemics. 2019; 29:100356. https://doi.org/10.1016/j.epidem.2019.100356
10. Trilla A, Trilla G, Daer C. The 1918 “Spanish Flu” in Spain. Clinical Infectious Diseases. 2008; 47(5):668-673. https://doi.org/10.1086/590567.
11. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N. *et al*. Estimates of the severity of coronavirus disease 2019: A model-based analysis. The Lancet Infectious Diseases. 2020; 20(6):669-677. https://doi.org/10.1016/S1473-3099(20)30243-7
12. Wang Y, You XY, Wang YJ, Peng LP, Du ZC, Gilmour S *et al*. Estimating the basic reproduction number of COVID-19 in Wuhan, China. *Zhonghua liu xing bing xue za zhi= Zhonghua liuxingbingxue zazhi*. 2020; 41(4):476-479.
13. World Health Organization. Middle East respiratory syndrome coronavirus (MERS-CoV), 2016.
14. Worldometer. (2020, July 29). Coronavirus update (Live): 9,435,610 cases and 481,968 deaths from COVID-19 virus pandemic. Retrieved July 29, from, 2020. https://www.worldometers.info/coronavirus/?
15. You C, Deng Y, Hu W, Sun J, Lin Q, Zhou F, Zhou XH. Estimation of the time-varying reproduction number of COVID-19 outbreak in China. International Journal of Hygiene and Environmental Health, 2020, 113555.