Modeling and Predicting Corona Contagion Dynamics in China, USA, Brazil & Ethiopia

Thomas Wetere Tulu[1, 2, *], Ieng Tak Leong[1], Zunyou Wu[3]

[1]Department of Mathematics, University of Macau, Macau, China

[2]Computational Science Program, Addis Ababa University, Addis Ababa, Ethiopia

[3]National Center for AIDS/STD Control and Prevention, Chinese Center for Disease Control and Prevention, Beijing, China

Email address:
thomaswetere@yahoo.com (T. W. Tulu)
*Corresponding author

Abstract: The COVID-19 pandemic is a global pandemic of coronavirus disease 2019, caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). The outbreak was first identified in Wuhan, China, in December 2019. In this article, we investigate the problem of modelling the trend of the current Coronavirus disease 2019 pandemic in China, USA, Ethiopia and Brazil along time. Two different models were developed using Bayesian Markov chain Monte Carlo simulation methods. The models fitted included Poisson autoregressive as a function of a short-term dependence only and Poisson autoregressive as a function of both a short-term dependence and a long-term dependence. The models can be employed to understand the contagion dynamics of the COVID-19, which can heavily impact health, economy and finance. The result indicates whether disease has an upward/downward trend, and where about every country is on that trend, all of which can help the public decision-makers to better plan health policy interventions and take the appropriate actions to control the spreading of the virus.

Keywords: Mathematical Modeling, Poisson Autoregressive, COVID-19, Markov Chain Monte Carlo, Simulation

1. Introduction

COVID-19 virus infection is a severe infectious disease occurring at unprecedented speed which has become the global public health treat now [1]. What makes the disease the worst of all is no specific effective treatment available, its dynamics is not much researched and understood. It is an infectious disease that is caused by the severe acute respiratory syndrome known as corona virus. COVID-19 is a viral disease quickly spreading its roots in various parts of the world. Its symptoms include fever, sore throat, coughing, and difficulty in breathing. While the virus was identified first in December 2019 in the city Wuhan, which is the capital of Hubei province in China [2], it now a global pandemic and has affected more than ten million people. Many of the countries had made mild preparations knowing that the diseases would ultimately catch up with them [3] due to inadequate information on its short- and long-term effects. Some of the common signs of COVID-19 includes fever, shortness of breath and dry coughs [4]. Other uncommon symptoms include muscle pains, mild diarrhea, abdominal pain, sputum production, loss of smell, as well as sore throat.

According to Word Health Organization's statistics as at March 30, 2020, [5] the mortality rate of persons who had been diagnosed cases was on average of 4.6 percent and it ranges from 0.2 percent to higher level of 15 percent depending on the age group, which also depends on the health status of the predisposed person, location of residence as well as immunity system among other factors.

The diseases are mostly spread through via respiratory droplets, which are produced whenever a person coughs or sneezes. In addition, an individual can also contract

the virus through touching a contaminated place or surface before touching their face [6]. This novel virus can easily survive on the ground or surfaces for up to a maximum of 72 hours; when the exposed person makes a contact with an infected person, symptoms can easily show up from two to a maximum of fourteen days, with the expected mean number of days as five days. The World Health Organization has reported a record increase in coronavirus cases around the world for the third time in eight days, with the total number of infections rising by 230,370 in a 24-hour period in July, 2020. This definitely shows the pandemic is still a great point of concern for the world.

Though some of the recommended measures that can be taken as prevention of infection includes social distancing and frequent hand washing, covering sneezes and coughs with inner elbow, using mask, a tissue and keeping away unwashed hands from one's face as per WHO [7, 11-16], it is rational to study the dynamics of the disease pandemic through modeling techniques.

2. Modeling

2.1. Stochastic Modeling Methods

An overview of COVID-19 conceptual framework

In an overview of the COVID-19 Conceptual framework subsection, it is important to know how COVID-19 infections move from one stochastic state to another during the period of infection.

In any given population, it comprises of four states of nature namely; Susceptible (S (t)), Infectious (I (t)), Recoveries (R (t)) and Deaths (D (t)). The notation t is the time during the study.

Briefly, susceptible individuals ($S(t)$) may become infected ($I(t)$) after the contact with Covid-infected individuals who are capable of infecting others including nurses, doctors etc. at hospitals and with a chance of infecting others before being recovered/removed from the disease ($R(t)$) or die of Covid-19 and then join ($D(t)$).

2.2. Model Development

Following Agosto and Giudici [8], the number of new cases y_t reported at time t is assumed to follow Poisson distribution.

$y_t \sim \text{Poisson}(\lambda_t), P(Yi = yi) = \frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!}$ Where y_i=1, 2, 3 …

$\log (\lambda_t)= \alpha+ \beta\log (1 + y_{t-1})$ (Model 1)
$\log (\lambda_t)= \alpha+ \beta\log (1 + y_{t-1}) + \gamma\log (\lambda_{t-1})$ (Model 2)

In each model, the inclusion of 1 in log (1 +y_{t-1}) allows to address the problem generated by zero values, α represents the intercept term and β expresses the short-term dependence of the expected number of cases reported at time t, λ_t, on these observed in the time t–1. The γ component in model 2 corresponds to a trend component and, more specifically, it represents the long-term dependence of λ_t on all past counts of the observed process.

3. Model Estimation Method and Simulation

Both models were implemented from a Bayesian perspective using Gibbs sampling MCMC simulation methods with the relevant code using WinBUGS software [6].

Monte Carlo simulations are used to model the probability of different outcomes in a process that cannot easily be predicted due to the intervention of random variables. As the spread of Corona virus is a random process, the Monte Carlo algorithm is used to simulate the Markov Chain process of which the transfer matrix changes over time.

In a Markov Chain process the physical state at time t + 1 depends only on the state at time t.

Define the state matrix as: X (t) =S(t), I(t), R(t), D(t) to represent the compartments in a population. Then, the initial state matrix X(0) is obtained as: X(0)=(1-I_0, I_0, 0, 0). According to Markov Chain theory the transition matrix can be given as: P(t)={ P(i, j) }$_{4x4}$ where P(i, j) is the transition probability from state i to state j for i, j an element of {1, 2, 3, 4}.

$$P(t) = \begin{bmatrix} \text{Variable} & S & I & R & D \\ S & P(1,1) & P(1,2) & P(1,3) & P(1,4) \\ I & P(2,1) & P(2,2) & P(2,3) & P(2,4) \\ R & P(3,1) & P(3,2) & P(3,3) & P(3,4) \\ D & P(4,1) & P(4,2) & P(4,3) & P(4,4) \end{bmatrix}$$

Where,
P(1,4)=P(2,1)=P(3,2)=P(3,4)=P(4,1)=p(4,2)=P(4,3)=0 as there is no transition and P(1,1)+ P(1,2)+P(1,3)=P(2,2)+P(2,3)+P(2,4)=P(3,1)+P(3,3)=P(4,4)=1.

Finally, the state matrix is given by:

$$X(t)=X(0)\prod_{t=0} P(t)$$

For every model, an initial 10,000 iterations were run as a "burn-in" to reach convergence [9]. This was evaluated by testing the convergence statistic for two parallel MCMC chains with different initial values. Following that, an additional 50,000 iterations were run for parameters estimation purposes. To this end, the prior distributions for all the regression parameters were specified as α, β, γ~N (0, 10⁶)

4. Data Source

Data are available from the website of World health organization, WHO website for all the countries used here [10].

5. Result and Discussion

The overall temporal distribution of daily counts of COVID-19 cases is presented in Figures 1-4 below. The data

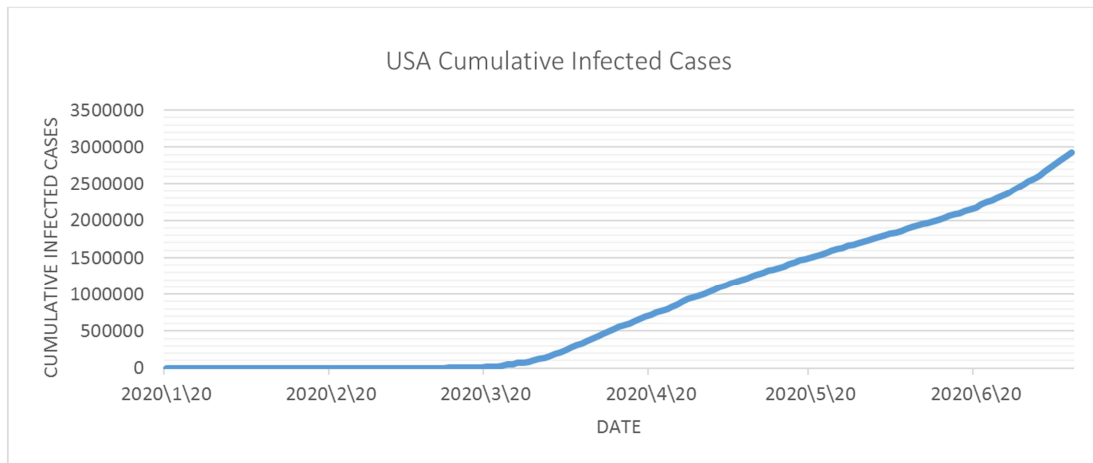covers the period from January 11 to July 15, 2020.
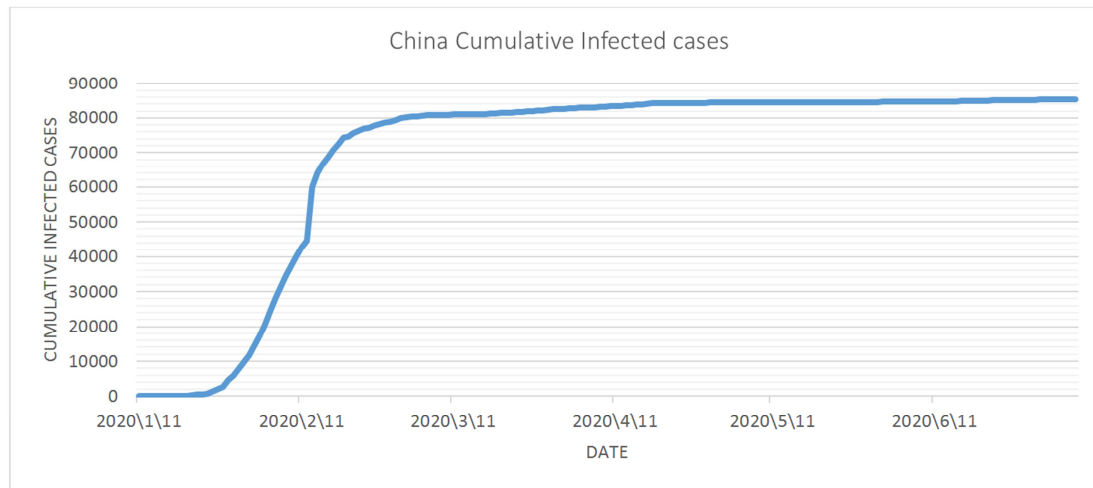

Figure 1. Cumulative infected cases in USA.

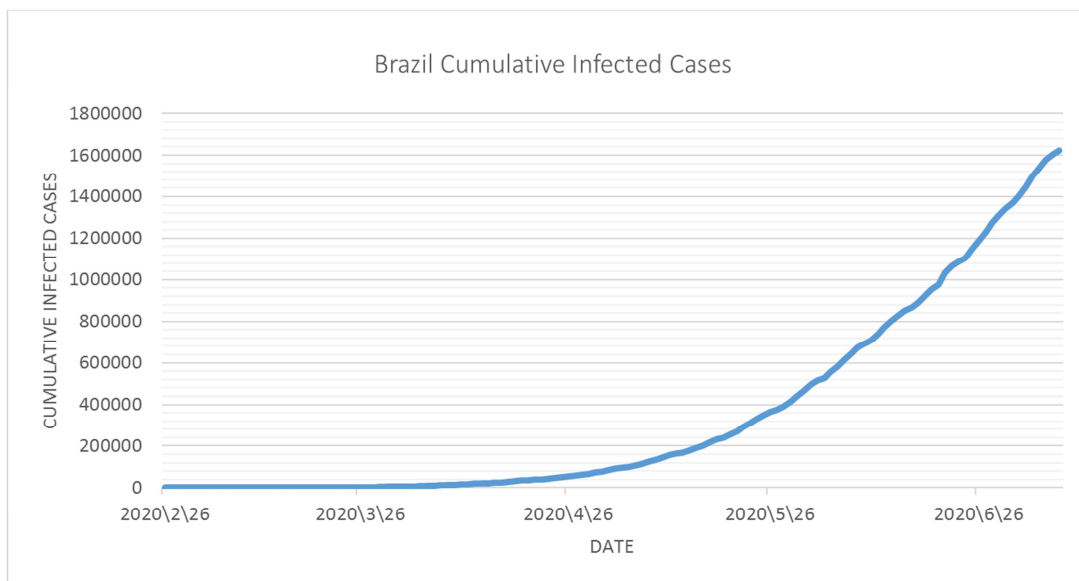*Figure 2.* Cumulative infected cases in China.

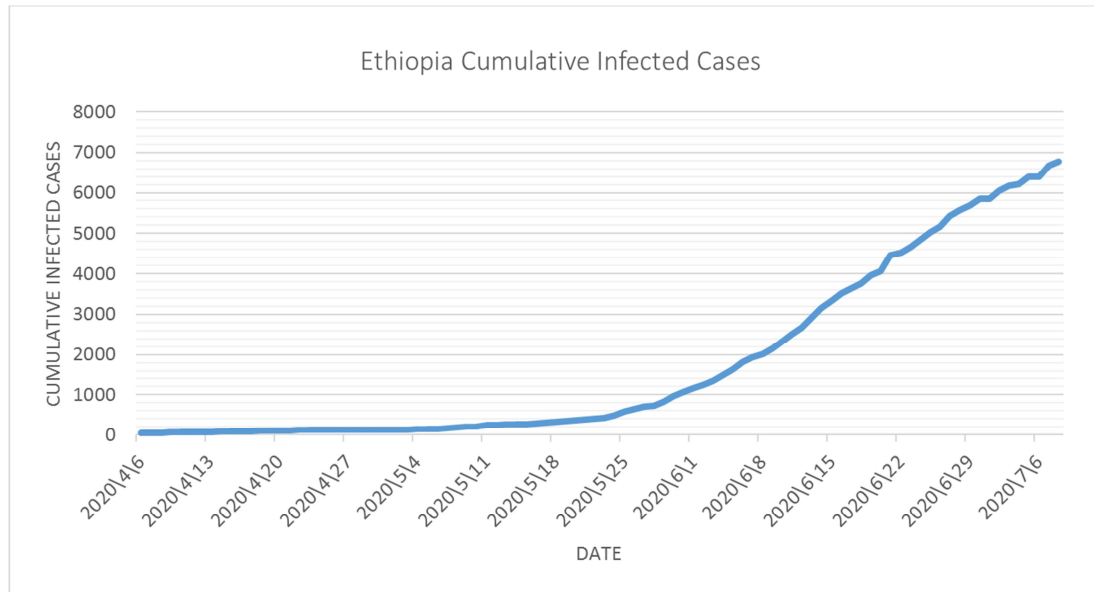*Figure 3.* Cumulative infected cases in Brazil.

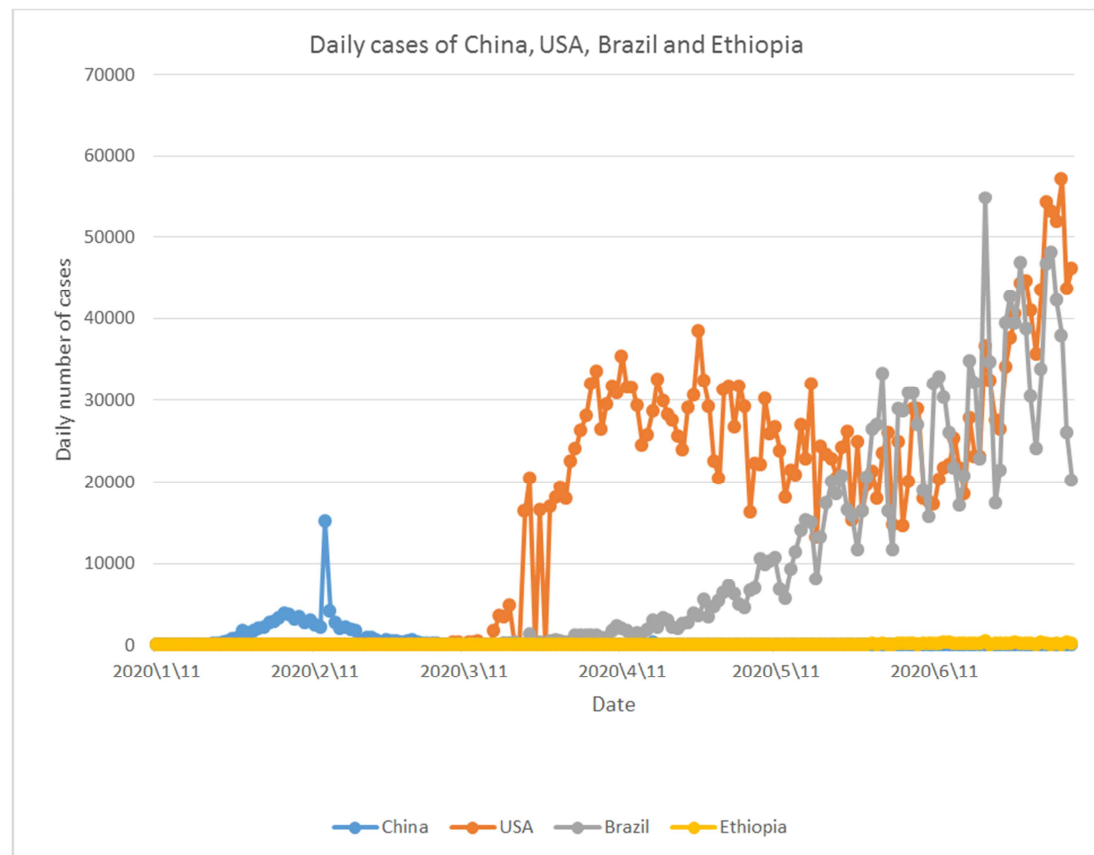**Figure 4.** *Cumulative infected cases in Ethiopia.*

Figure 5. *Daily infected cases in China, USA, Brazil and Ethiopia.*

As it is clearly seen from figures above the cumulative infected case is increasing in an alarming rate in all countries except China which has almost a slope very close to zero.

5.1. Estimated Coefficients of the Poisson Autoregressive Models

Table 1 below shows the estimated coefficients of the Poisson autoregressive for both models together with their associated 95% credible intervals.

We notice that all coefficients had the expected positive sign as well as their credible intervals excluding zero, indicating the presence of a short-term dependence for model 1 and both a short-term dependence and a long-term trend for model 2.

Table 1. *Model estimates with standard errors and p-values.*

Country	Parameter	Model 1	Model 2	Std Error (p-value)
China	α	0.4334	0.2531	0.1427 (0.0779)
	β	0.8924	0.5291	0.0698 (0.0000)
	γ	-	0.4083	0.06956 (0.0000)
	RMSE	NA	NA	
	R_0	1.107		
USA	α	3.2599	2.1033	0.7917 (0.0092)
	β	0.6807	0.4430	0.09502 (0.0000)
	γ	-	0.3516	0.095353 (0.00037)
	RMSE	NA	NA	
	R_0	3.76		
Brazil	α	0.5033	0.5203	0.2313 (0.0267)
	β	0.94482	0.8755	0.10089 (0.0000)
	γ	-	0.07129	0.0990 (0.4731)
	RMSE	NA	NA	
	R_0	3.268		
Ethiopia	α	0.9849	0.7851	0.29929 (0.01098)
	β	0.7755	0.4577	0.11890 (0.0002)
	γ	-	0.5372	0.11349 (0.001713)
	RMSE	7.56	5.34	
	R_0	1.47		

5.2. Discussion

Table 1 above shows that all estimated autoregressive coefficients are significant with all p-values less than that of α, confirming the presence of both a short-term dependence and a long-term trend in all the four countries. From the estimated values it is clearly seen that, the estimate of γ parameter is lower than that of β in the case of China, USA and Brazil showing the downward trend. In the case of Ethiopia β parameter is lower than that of γ showing that the pandemic did not reach the pick yet or the presence of the upward trend.

Agosto and Giudici [8] drew a similar conclusion from their analysis of the Chinese data, which covers the period from January 20 to March 15, 2020.

Besides, the results in table above show that, if the expectation of new cases for yesterday was close to 0, 1000 new cases observed yesterday generate about 400 new expected cases today in the case of Ethiopia.

Basic reproduction number of COVID-19 is crucial parameter during a pandemic which is used to estimate the risk of COVID-19 outbreak and evaluate the effectiveness of implemented measures. For the first time, we report estimates of the basic reproduction number R_0 of COVID-19 outbreak in comparison between China, USA, Brazil and Ethiopia using SIR model by fitting the model to official reported data from WHO. According to the result USA has the highest R_0 and Ethiopia has the lowest. This automatically implies the average daily number of an infected person making infection is very high in USA than other countries so that urgent solution should be applied. Due to short onset time, current estimates of R_0 for COVID-19 in these countries might be biased. Moreover, the basic reproduction number is continuously modified during a pandemic by accurate assumptions introduced and becomes more reliable as more

data and information released.

6. Conclusion

In this article, we analyzed the trend of the current Coronavirus disease 2019 (COVID-19) pandemic in China, USA, Brazil and Ethiopia along time through modeling. We developed two different models of increasing complexity using Bayesian MCMC simulation methods, and found that the Poisson autoregressive as a function of both a short-term dependence and a long-term dependence provides the best fit to the data from WHO. The use of Poisson autoregressive model that allows to capture short and long term memory effects can greatly improve the estimation of number of new cases and can indicate whether disease has an upward/downward trend, and where about every country is on that trend, all of which can help the public decision-makers to better plan health policy interventions and take the appropriate actions to control the spreading of the virus. The model is applicable to other countries and more time periods as data becomes available.

Finally, we strongly recommend that there should be an increase in the level of awareness of Covid19 especially in developing countries. Government and NGO's should institutionalize best practices in care and support for people living with Covid-19 and also ensure the availability of retroviral drugs for infected people. Other precautionary measures such as blood screening before transfusion, total hygiene and distance should be observed in order to reduce the pandemic.

Competing Interests

The authors declare that they have no competing interests.

Authors' Contributions

The authors declare that the study was conducted in collaboration with the same responsibility. All authors read and approved the final manuscript.

Acknowledgements

References

[1] Balkhair AA. COVID-19 Pandemic: A New Chapter in the History of Infectious Diseases. *Oman Med J.* 2020; 35 (2): e123.

[2] Yunlu Wang, Menghan Hu, Qingli Li, Xiao-Ping Zhang, Guangtao Zhai, Nan Yao. Abnormal respiratory patterns classier may contribute to large-scale screening of people infected with COVID19 in an accurate and unobtrusive manner. arXiv preprint arXiv: 2002.05534; 2020.

[3] Zixin Hu, Qiyang Ge, Li Jin, Momiao Xiong. Artificial intelligence forecasting of COVID-19 in China. arXiv preprint arXiv: 2002.07112; 2020.

[4] Fan Wu, Su Zhao, Bin Yu, Yan-Mei Chen, Wen Wang, Zhi-Gang Song, Yi Hu, Zhao-Wu Tao, Jun Hua Tian, Yuan-Yuan Pei, et al. A new coronavirus associated with human respiratory disease in china. Nature. 2020; 579 (7798): 265–269.

[5] World Health Organization. Global surveillance for COVID-19 disease caused by human infection with the 2019 novel coronavirus, interim guidance; 2020.

[6] Spiegelhatler, D. J., Thomas, A., Best, N. G., & Lunn, D. (2003). WinBUGS User manual: Version 1.4. Cambridge: MRC Biostatistics Unit.

[7] WHO: Coronavirus Cases.: //covid19.who.int/?gclid=EAIaIQobChMI4saYtNe_6gIVB7e WCh0uRAcgEAAYASAAEgK8O_D_BwE.

[8] Agosto, A. and Giudici, P, A Poisson Autoregressive Model to Understand COVID-19 Contagion Dynamics (March 9, 2020). Available at SSRN: https://ssrn.com/abstract=3551626.

[9] Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulationusing multiple sequences. Statistical Sciences, 7, 457–472.

[10] NovelCoronavirus2019. https://ourworldindata.org/coronavirus-testingnovel-coronavirus-2019.

[11] WHO. Coronavirus disease (COVID-19) situation report 165. Geneva: World Health Organization. July 3, 2020. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200703-covid-19-sitrep-165.pdf?sfvrsn=b27a772e_6 (accessed July 3, 2020).

[12] Rosenberg ES, Dufort EM, Udo T, et al. Association of treatment with hydroxychloroquine or azithromycin with in-hospital mortality in patients with COVID-19 in New York State. JAMA 2020; 323: 2493–502.

[13] Arabi YM, Deeb AM, Al-Hameed F, et al. Macrolides in critically ill patients with Middle East respiratory syndrome. Int J Infect Dis2019; 81: 184–90.

[14] Cao B, Wang Y, Wen D, et al. A trial of lopinavir-ritonavir in adults hospitalized with severe Covid-19. N Engl J Med 2020; 382: 1787–99.

[15] Marijon E, Karam N, Jost D, et al. Out-of-hospital cardiac arrest during the COVID-19 pandemic in Paris, France: a population-based, observational study. The Lancet Public Health.

[16] Worldometer. COVID-19 coronavirus pandemic. 2020. https://www.worldometers.info/coronavirus/ (accessed May 28, 2020).