The epidemic trend of COVID-19 in SAARC countries: a predictive modelling and analysis

Sarita Bhatt [*, **], Radha Krishna Joshi [*, **], Tika Ram Lamichhane[*] and Madhav Prasad Ghimire[*, **]

[*]Central Department of Physics, Tribhuvan University, Kirtipur, Kathmandu, Nepal.

[**]Condensed Matter Physics Research Center, Lumbini, Nepal.

Abstract: This paper aims to integrate novel coronavirus daily cases in SAARC countries; India, Pakistan, Bangladesh, Nepal, Sri Lanka, Afghanistan, Maldives and Bhutan to forecast the epidemic trend of COVID-19 by using logistic model. The recent trend of coronavirus cases were analyzed from the COVID-19 epidemiological data for SAARC countries from 23 January 2020 to 31 May 2021. The final size, growth rate parameter and point of inflection of COVID-19 for each countries were calculated by fitting the logistic curve with the cumulative cases. The graphical patterns of COVID-19 daily cases reflect that its second wave impact is more devastating than the first wave in SAARC countries. The increasing trend of COVID-19 cases in these countries was well described by logistic model with coefficient of determination greater than 0.96. The predictive final size of the second wave infections is maximum for India which is 19.8 million with growth rate parameter of 0.08 and inflection time of 68 days whereas the predictive final size is minimum for Afghanistan which is 0.041 million with growth rate parameter of 0.06 and inflection time of 71 days. The logistic model is helpful in predicting the trajectory of the infected cases in a country if the current scenario of this type of infectious disease remains same. Also, it helps the government to frame policy decisions and necessary actions that controls the transmission of COVID-19 in the South Asian region.

Keywords: COVID-19 cases; Logistic model; Best fit; Growth rate parameter; Inflection point; Predictive cases.

Introduction

Coronaviruses are a large family of respiratory viruses that cause disease in mammals and birds[1]. A coronavirus consists of non-segmented positive-sense, enveloped and single-stranded ribonucleic acid (RNA) genome having size of about 30 kilobase (kb) which is found as one of the longest knowngenome[2, 3]. It is classified into four genera: alpha, beta, gamma, and delta with many stereotypes. The alpha and beta types are found infectious to humankind (CoVs)[3] whereas the gamma and delta types infect birds[4, 5]. The alpha coronaviruses are HCoV-229E and HCoVNL63, and the beta coronaviruses are HCoV-HKU1, HCoV-OC43, severe acute respiratory syndrome coronavirus (SARS-CoV), Middle East respiratory syndrome coronavirus (MERS-CoV), and SARS-CoV-2 which infect humans[6, 7]. Coronavirus disease (COVID-19) which ranks among the 10 deadliest plagues in history is caused by SARS-CoV-2. As an extremely pathogenic virus, it was first detected in Wuhan, China in December 2019 and subsequently spread worldwide. It was declared a global health emergency in 2020 by World Health Organization (WHO)[8]. SARS-CoV-2 main protease takes part in viral transcription and replication and acts as an active target for the drug candidates as explained by different interactions[9, 10]. Recently, Pelinovsky et al. (2020) used both generalized and simple logistic model to their study and analyzed COVID-19 data in few countries such

Author for correspondence: Tika Ram Lamichhane and Madhav Prasad Ghimire, Central Department of Physics, Tribhuvan University, Kathmandu, Nepal.

Email: tika.lamichhane@cdp.tu.edu.np; madhav.ghimire@cdp.tu.edu.np

Received: 26 Aug 2021; First Review: 10 Sep 2021; Second Review: 23 May 2022; Accepted: 25 May 2022.

Doi: https://doi.org/10.3126/sw.v15i15.45669

as Austria, Switzerland, Italy, Turkey and South Korea. They also compared the simple logistic with generalized logistic growth model which best fits the data of coronavirus disease 2019 with coefficient of determination greater than 0.8. The previous findings and relevant interpretations revealed that the most appropriate growth curves for these countries were from the generalized logistic model but not from the crude logistic model[8,11,12].

Attanayake et al. (2020) used Gompertz, logistic, Weibull and exponential models in order to interpret the COVID-19 epidemic data and described the growth rate of COVID-19 cases for Sri Lanka, Italy, USA, and Hubei province of China. They forecasted the epidemic trend of COVID-19 by calculating the final size of the pandemic, growth rate parameter and coefficient of determination (COD) along with Akaike information criterion (AIC), Bayesian information criterion (BIC) and residual mean squared error (RMSE). The results revealed that order of more to less appropriate model for the selected countries is logistic, Weibull and Gompertz model, respectively. By studying the growth curve for each country, they interpreted the outbreaks in Italy and Sri Lanka. Currently, the coronavirus spreading rate is slower in Italy and Sri Lanka, however it is expected to be stopped in Hubei, China[13]. Jain et al. (2020), used phenomenological models, such as generalized logistic growth, logistic growth and generalized growth model for COVID-19 pandemic in India. They estimated growth rate in early stage and epidemic final size with 95% confidence interval by plotting growth curve which signified the growth of cumulative number of cases and the future trend of pandemic[14]. Wang et al. (2020) fitted the epidemiological data for COVID-19 using logistic model for the estimation of trend of the pandemic[15]. They explored global data of COVID-19 outbreak and analyzed for several countries like Brazil, India, Russia, Peru, and Indonesia. Thus, the hybrid Logistic and Prophet model has been proved to be highly effective for forecasting the epidemic trend of COVID-19. The physical parameters (growth rate, point of inflection and final size of infections) estimated from the data modeling under proper statistical analysis are expected to be useful for policy making in restricting the community transmission of deadly viruses[12].

Methodology

Dataset sources

To evaluate the trend of COVID-19 in SAARC countries, the data were collected from the site of Worldometer and Ministry of Health and Population, Nepal[16, 17] from 23 January 2020 (the initial day of the COVI-19 case appeared in Nepal) to 31 May 2021. The data includes total number of cases and daily new cases which are used to model out and forecast the epidemic trend of COVID-19 in SAARC countries.

Logistic method

Modelling of infectious disease helps us to predict future impact of disease on the basis of current trend[18]. The most common approaches for the study of infectious disease are mathematical, statistical and computational modelling. We have used a logistic model proposed by Verhulst in 1838[19]. The logistic model has approximately exponential growth nature at the initial phase, continuous growth with a reduced rate and finally reaches maximum resulting in a sigmoid curve defined by equation (1)[20, 21].

$$s(t) = \frac{k}{1 + e^{-\alpha(t-t_0)}} \quad \dots (1)$$

Here, $s(t)$ represents cumulative number of cases, k is maximum value or final size of the model as $t \to \infty$, α is growth rate parameter and t_0 corresponds to inflection point which marks a critical turning point at which transmission of the disease begins to decline showing a flat curve. The final size of the cumulative cases were estimated with 95% confidence interval. In this paper, the values of four parameters k, a, t_0 and R^2 were calculated by using logistic regression tool of Desmosgraphing calculator 2021.

Model evaluation

The fitting ability of logistic model is given by coefficient of determination (COD). The COD represents proportion of variance in dependent variable due to independent variable which can be calculated by using equation (2)[22,23]:

$$R^2 = 1 - \frac{\sum_j (S_j - \hat{S}_j)^2}{\sum_j (S_j - \bar{S})^2} \quad \ldots(2)$$

Here, S_j represents the actual cumulative confirmed cases, $\hat{S}_j$ is the predicted cumulative confirmed cases, $\bar{S}$ is an average of actual commutative confirmed cases. The value of R^2 lies in between 0 and 1. The value of R^2 near to 0 cannot predict the change in dependent variable due to the change in independent variable, i.e. the results obtained from the fitting are not statistically significant. The value of R^2 near to 1 signifies that the results obtained from the fitted equation are significant.

Results and discussion

Although there is no exact definition for a second or third wave, a new wave of pandemic is defined as a rise of new infection after initial fall in the number of cases. Most of the countries in the world are recently experiencing fast evolving new wave of pandemic with new strain of COVID-19 which is more transmissible than previous one[24, 25]. This new wave can be assumed as a second or third wave with respect to last wave seen for daily cases versus time. South Asian region is among the most densely populated region compared to other region in the world with poor economy[24]. Thus, this can be one reason that SAARC countries are vulnerable countries and are considered the epicenter of second wave of COVID-19 pandemic. Figure 1 shows the waves of COVID-19 cases in SAARC countries. Starting of new wave of pandemic differ from one country to another. In most of these countries, the new wave starts between March and April, 2021. Pakistan and Afghanistan have already experienced the small magnitude of second wave and are seen to be switched into peak of third wave. India, Nepal, Sri Lanka, Bangladesh and Maldives have the greater impact of COVID-19 pandemic appearing with the larger magnitude of second wave after about 400 days from 23 January 2020 (the date of first case appeared in South Asian region, i.e. in Nepal) as shown in Figure 1.

The first COVID patient in Bhutan was confirmed on 6 March 2020. After the first case reported, there was a gradual increase in the number of cases due to which country's borders were sealed and national lockdown was declared[26]. Finally, Bhutan conquered the rise in cases on March 2021 with zero active cases from 5 to 18 March (longest period). The cases, however, started to increase from April 2021 showing very small peak at the end of May 2021. Since there is very small rise in cumulative cases, it seems difficult to study the logistic model for such case with

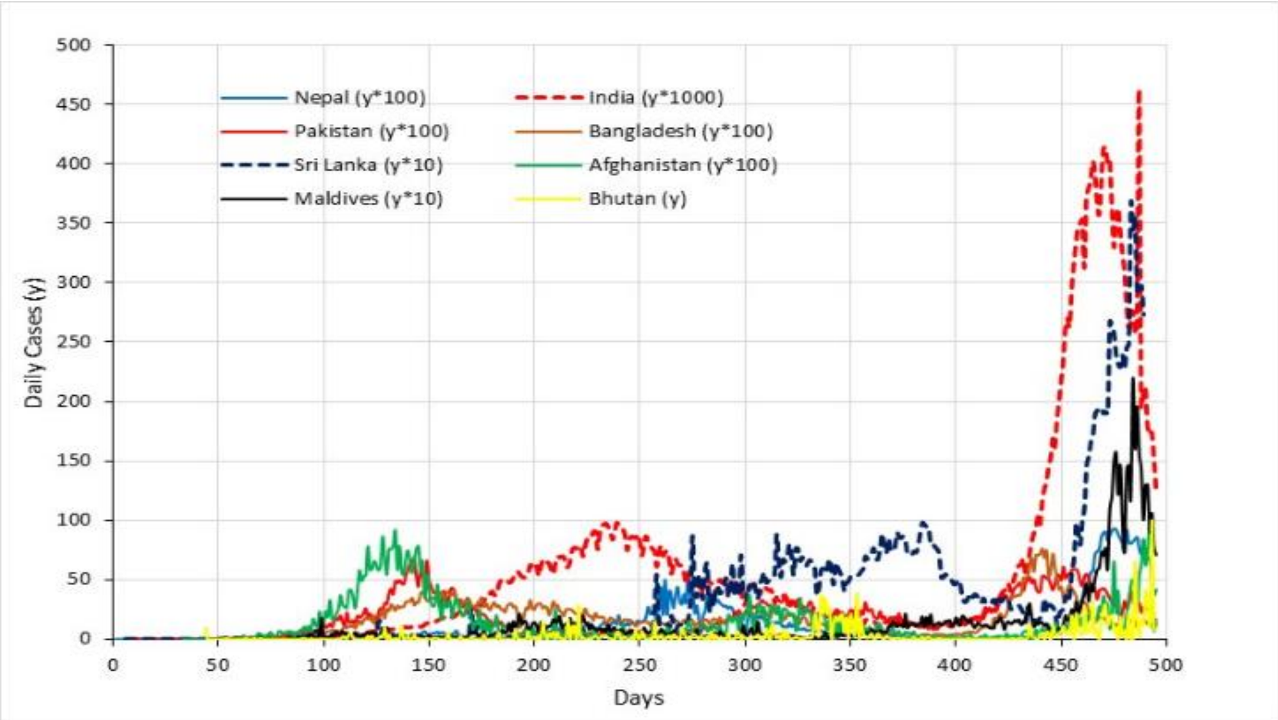

Figure 1: COVID-19 daily cases in SAARC countries from 23 January 2020 to 31 May 2021.

very low coefficient of determination[22]. As of April 2021, over 93% of adult population have completed vaccinating themselves. Among them, 85% have vaccinated by first dose within a week. Thus, while the pandemic continues to worsen in south Asia, this country has managed to control the pandemic[27].

Table 1. The different parameters of the model described by equation 1 with 95% confidence interval of k values for the COVID-19 second wave infections

Countries	k	α	t_0(days)	R^2	*Conf. Interval*
Afghanistan	41352	0.06	71	0.99	38802, 43902
Bangladesh	90499	0.12	14	0.97	83517, 97418
India	19800000	0.08	68	0.99	18614357, 20985643
Maldives	47087	0.13	35	0.99	43425, 50749
Nepal	307291	0.13	43	0.99	285075,329511
Pakistan	319206	0.07	37	0.99	297352,341060
Sri Lanka	108312	0.11	34	0.99	100163,116461

Figure 2 illustrates the fitting of cumulative cases with the logistic model which is obtained to describe the epidemic trend in the countries. The thick line in Figure 2 represents the observed cumulative confirmed cases whereas thin solid curve represents that estimated by logistic model. The parameters of a time dependent logistic model defined by equation 1 and calculated for each country are listed in Table 1.

From Table 1, we observe that the values of four parameters differ from one country to another following different growth trend. The early phase of pandemic have similarities in the nature of curve. This means that as the time rises, the new cases are found to increase in daily basis creating an alarming situation. India is the world's second populated country and has the highest number of COVID-19 cases in the second wave of pandemic. The final pandemic size estimated from this model has its highest value for India and lowest for Afghanistan. Meanwhile, the highest value of point of inflection is for Afghanistan followed by India, Nepal, Pakistan, Maldives, Sri Lanka and Bangladesh.

The logistic curve has been fitted very well with observed cumulative cases in each of the SAARC countries. The estimated curve after 31 May 2021 shows that there is rapid increase in the cases which is approximately exponential up to time (t_0) which is corresponding to the inflection point

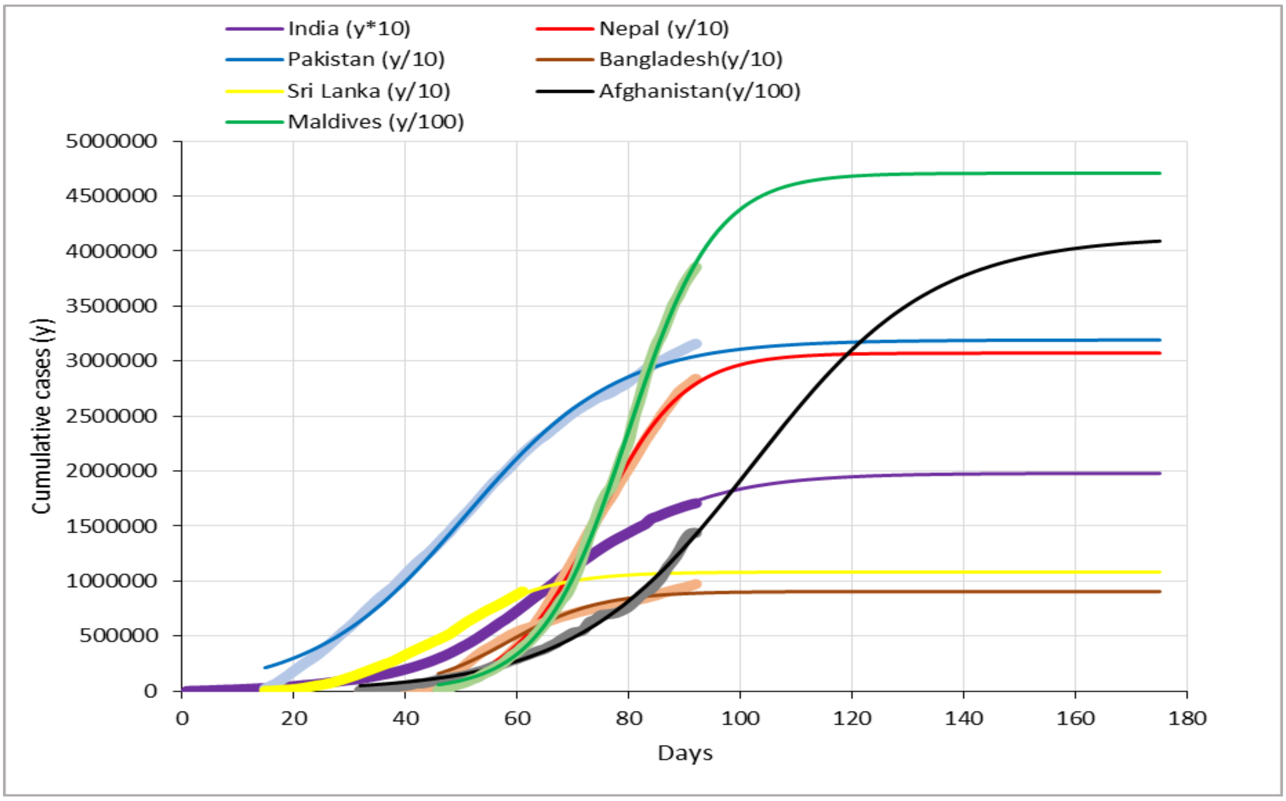

Figure 2: Estimated cumulative cases (thin curves) fitted with observed cumulative cases (thick curves) for the COVID-19 new wave in SAARC countries from 1 March 2021 to 31 May 2021 using a logistic model.

and at maturity, growth ends showing flat curve as shown in Figure 2. Thus, this model predicts not only the number of cases accurately for long time but also the new cases per day. Also, the logistic model suggests prediction of future trend which will be an alarming if same situation is continued. It encourages us to take the precaution and safety measure before situation get worst.

As an impact of second wave of COVID-19 pandemic starting from mid-March 2021, the estimated final size of cumulative cases for India is about 20 million with inflection time of 68 days and that for Nepal is about 3 lakhs with the inflection time of 43 days denoted by purple and red color curves in Figure 2.From the observations it has been found that India has alarming situation among all the SAARC countries. The situation is getting more flared as each day passes. For India, the peak of second wave seen on 6th May 2021. During these time COVID-19 cases reached more than 4 lakhs per day. This proves the current situation is more alarming and devastating than the first wave caused by the double mutant coronavirus. This mutant virus is found to have more efficient transmission rate with lesser incubation period[28]. Other factors for sudden spike in cases in countries like India could be failing to implement lockdown policy. Further the biggest spike in cases is recorded during festivals such as Holi and Kumbha Mela, local assembly and state elections in India that occurred in March 2021. This has caused adverse effect on the health care system that includes shortage of oxygen cylinder, ventilators, ICU, beds in hospitals, medicine, vaccines, etc. The impact of COVID-19 second wave has been observed to be more intensive in South Asian region than the first wave. According to Daun et al. (2020), predicted size of the first wave outbreak in India was around 2.5 million with the inflection point arriving July 26, 2020[27].

In our study of logistic regression analysis, India has the predicted size of second wave infections of about 20 million with the inflection point on 68th day from its starting on mid-March 2021. Since Nepal shares open boarder with India, many Nepalese went India for job in cities like Delhi, Mumbai and other states where daily cases are rapidly increasing. When COVID-19 impact started to rise again after the first wave in India, most of the Nepalese people returned home through the open boarder. This caused significant rise of COVID-19 cases in Nepal resulting in the shortage of oxygen cylinder, ventilators, ambulance, hospital beds and medicines[29]. This type of scenario appears in most of the SAARC countries getting impact of the second wave caused by double mutant coronavirus. Timely vaccination, effective implementation of lockdown, making the country's healthcare system stronger and the people's awareness by wearing mask, washing hands and maintaining social distance can reduce the growing impact of COVID-19 pandemic.

Limitations of study

This study does not consider the effect of quarantine, isolation, lockdown and governmental measures such as social-distancing policy and vaccination implemented to control COVID-19 transmission. Mathematical modelling assumes that the growth rate of COVID-19 in a particular region remains constant, however in actual practice, it remains changing with time.

Conclusion

In the SAARC countries, the new impact of COVID-19 pandemic caused by doubly variant coronavirus appeared in March 2021. India is the most affected and Bhutan is the least affected by this pandemic to the date of 31 May 2021. Observing the daily cases, the SAARC countries having the decreasing order of impact of this infectious disease are: India, Nepal, Sri Lanka, Pakistan, Bangladesh, Maldives, Afghanistan and Bhutan, respectively. For all SAARC countries except Bhutan, the curve of logistic model is well fitted with the observed cumulative cases resulting the coefficient of determination greater than 0.95. Out of these seven countries, final size of the second wave pandemic is the highest for India with estimated cumulative cases 19.8 million (95% confidence interval: 18.61 to 20.98 million), growth rate parameter 0.08 arriving inflection point on 71thday of COVID-19 second wave impact. The final size of the estimated cumulative cases is the lowest for

Afghanistan which is 0.041 million (95% confidence interval: 0.038 to 0.044 million) with growth rate parameter 0.06 and inflection time of 71 days. For Nepal, the estimated final size of the cases is 0.32 million (95% confidence interval: 0.28 million to 0.32 million) with inflection time of 43 days. It is found that the COVID-19 epidemic curve reached peak for Bangladesh in 14 days earlier than other countries. Though the growth rate is higher, infection period of second wave in Bangladesh is the least. Also, the peak occurred in 71 days and in 68 days for Afghanistan and India, respectively. The peak of pandemic occurred later in Afghanistan than in India. In case of Nepal, the peak arrived in 43 days which is earlier than Afghanistan and India and later than Bangladesh, Sri Lanka, Pakistan and Maldives. This type of data modelling provides the scientific idea to the policymakers about the future status of the infectious disease on the basis of present trend. By analyzing the results obtained from the explicit mathematical modelling, one can set a strategy for the control of this disease that provides awareness among the individuals.

Conflict of interest

The authors declare that there is no conflict of interest.

Funding information

There is no funding available for this research.

References

1. Sohrabi, C., Alsafi, Z., O'Neil, N., Khan, M., Kerwan, A. and Al-Jabir, A., et al. 2020. World health organization declares global emergency: A review of the 2019 novel coronavirus (Covid-19). *Int. J. Surg.* **76**: 71-76.
2. Toit, D. 2020. Outbreak of a novel coronavirus. *Nat. Rev. Microbiol.* **18**: 123–124.
3. Yang, P., and Wang, X. 2020. Covid-19: A new challenge for human beings. *Cell. Mol. Immunol.* **17**: 555–557.
4. Jahangir, M. A., Muheem, A., and Rizvi M. F. 2020. Coronavirus (Covid-19): history, current knowledge and pipeline medications. *Int. J. Pharm. Pharmacol.* **4**: 1–9.
5. Zhu, W., Yang, J., Lu S., Lan, R., Jin, D. and Luo, X. I., et al. 2021. Beta- and novel delta-coronaviruses are identified from wild animals in the Qinghai-Tibetan plateau, China. *Virol. Sin.* **18**: 402–411.
6. Wille, M., Holmes, E. C. 2020. Wild birds as reservoirs for diverse and abundant gamma- and delta coronaviruses. *FEMS Microbiol. Rev.* **44**: 631-644.
7. Luo, X., Zhou G. Z., Zhang, Y., Zou, L. P. and Yang, Y. S. 2020. Coronaviruses and gastrointestinal diseases. *Mil. Med. Res.* **7**: 1–6.
8. Lamichhane, T. R. and Ghimire, M. P. 2020. Research on Covid-19 from biophysical perspective. *Tribhuvan University Journal.* **34**:1-14.
9. Lamichhane, T. R. and Ghimire, M. P. 2020. Structural Analysis of COVID-19 Main Protease and its Interaction with the Inhibitor N3. *ChemRxiv. https://doi.org/10.26434/chemrxiv.12400604.v1*
10. Lamichhane, T. R. and Ghimire, M. P. 2021. Evaluation of SARS-CoV-2 main protease and inhibitor interactions using dihedral angle distributions and radial distribution function. *Heliyon*.**7**: e08220.
11. Pelinovsky, E., Kurkin, A., Kurkina, O., Kokoulina, M. and Epifanova, A. 2020. Logistic equation and Covid-19. *Chaos Soliton Fract.* **140**: 110241.
12. Joshi, R. K., Bhatt, S., Lamichhane, T. R. and Ghimire, M. P. 2021. Analysis of second wave of COVID-19 cases in Nepal with a logistic model. *Asian J. Med. Sci.***12**: 20-26.
13. Attanayake, A. M. C. H., Perera, S. S. N. and Jayasinghe, S. 2020. Phenomenological modelling of Covid-19 epidemics in Sri Lanka, Italy, the United States, and Hubei province of China. *Comput. Math. Methods Med.* **2020**: 1–15.
14. Jain. M., Bharti, P. K., Kataria, P. and Kumar, R. 2020. Modelling logistic growth model for Covid-19 pandemic in India. *ICCES. IEEE.* 784–789.
15. Wang, P., Zheng, X., Li, J. and Zhu, B. 2020. Prediction of epidemic trends in Covid-19 with logistic model and machine learning technics. *Chaos Soliton Fract.* **39**: 1–7.
16. Worldometer (Accessed: 1 June 2021). URL: https://www.worldometers.info/coronavirus
17. Ministry of Health and Population, Nepal (2021). URL:https://covid19.mohp.gov.np/#/
18. Woolhouse, M. 2020. How to make predictions about future infectious disease risks. Philos. *Trans. R. Soc. B.* **366**: 2045-2054.
19. Verhulst, P. F. 1838. Notice sur la loique la population poursuitdans son accroissement. Corresp. *Math. Phys.* **10**: 113–126.
20. Hao, Y., Xu, T., Hu, H., Wang, P., and Bai, Y. 2020. Prediction and analysis of corona virus disease 2019. *PloS One*.**15**: e0239960.
21. Torrealba-Rodriguez, O., Conde-Gutierrez, R. A. and Hernandez-Javier, A. L. 2020. Modeling and prediction of Covid-19 in Mexico applying mathematical and computational models. *Chaos SolitonFract.* **138**: 109946.
22. Tjur, T. 2012. Coefficients of determination in logistic regression models a new proposal: The coefficient of discrimination. *Amer. Statist.* **63**: 366–372.
23. Wibowo, F. W. 2021. Prediction modelling of Covid-19 outbreak in Indonesia using a logistic regression model. *J. Phys.: Conf. Ser.* **1830**: 012015.

24. oore, J. P. and Offit, P. A. 2021. SARS-COV-2 vaccines and the growing threat of viral variants. *JAMA*. **325**: 821-822.
25. Shaheen, I. 2013. SARS-COV-2 vaccines and the growing threat of viral variants. *IOSR-JHSS.* **15**: 1–9.
26. Tsheten, T., Wangchuk, S., Wangmo, D., Clements, A. C. A. Gray, D. J.and Wangdi, K. 2021. Covid-19 response and lessons learned on dengue control in Bhutan. *J. Med. Entomol.* **58**: 502–504.
27. Dorji, T., Tamang, S. T., Wangmo, D., Clements, A. C. A., Gray, D. J. and Wangdi, K., Bhutan's experience with covid-19 vaccination in 2021.*BMJ Glob. Health.* **6**: e005977.
28. Matta, S., Rajpal, S., Chopra, K. K. and Arora, V. K. 2021. Covid-19 vaccines and new mutant strains impacting the pandemic. *Indian J. Tuberc.* ***68***:171–174.
29. Baniya, J., Bhattarai, S., Pradhan, V. and Thapa, B. J. 2020. Visibility of invisible: Covid-19 and Nepal-India migration. *Tribhuvan University Journal*. **34**: 101–115.
30. Duan, Q., Wu, J., Wu, G. and Wang, Y. G. 2020. Predication of inflection point and outbreak size of COVID-19 in new epicenters. *arXiv preprint:2007.0747*