ISSN:2394-2371
CODEN (USA):IJPTIL

RESEARCH PAPER

Regression and SIR model based COVID-19 pandemic forecasting in India

Tushant Agrawal[1] , Mayank Chhabra[2] *
[1]Amity Institute of Biotechnology, Amity University Rajasthan, Jaipur, India-303002
[2]Indian Institute of Health Management Research, IIHMR, Jaipur, India.
*Corresponding Author: **Mayank Chhabra**

ABSTRACT

Currently Covid-19 outbreak is spreading vigorously around the country with alarming growth, hence becoming a distress for country. As there is no vaccine or treatment for covid-19 available till now, so outbreak growth analysis should be monitored and requires attention by authorities to minimize the spread of disease by implementing necessary actions like Social distancing, lockdown and public hygiene practices. For outbreak like Covid-19 mathematical model plays an important role in gaining knowledge about the future of disease spread and in strategic decision making for controlling disease spread. This study aims to forecast the growth in number of confirmed cases, recovered cases and death for coming days in India using basic mathematical model. A simple effective mathematical SIR and Regression model was used to predict the future of Covid-19 pandemic by using time series data. Data was being analyzed till 25 May 2020 and prediction had been made for further number of new infected cases, recovered and death till 30 June 2020 and found a sudden growth in number of cases. The model study with realistic parameters set shows that the Covid-19 pandemic will be at its peak around the end of July and start decline around the month of August. We also estimate significant linear trend with best model fitting on COVID-19 cases in India. This study will help the Government and doctors in preparing their plans for the next month for controlling the spread. Based on the predictions for short-term interval, these models can be tuned for forecasting in long-term intervals.

Keywords: - COVID-19, Forecast, Prediction, Linear regression, SIR, R_0Values, Statistical Model.

INTRODUCTIO

One Now Days, the world faces a major crises, the COVID-19 outbreak. The pandemic disease becomes a great threat to the health and safety of people across the globe due to its rapid spreading and potential mortality. It might took over a week to reveal an infected individual to indications of new corona virus, even though may be able to infect more susceptible individuals throughout that time. At present, the rate of spreading the disease of corona virus in India has increased day by day with an estimate of more than 7000 cases. Jiang et al. [1] states that the case fatality rate from corona virus is predicted 4.5% and in the age group amid 70-79, it may be raise up to 8.0% and for >80 it has been reported to be 14.8%. The transmission from one person to another from droplets of respiratory among near contact with the average peoples infected by a patient being 1.5-3.5[2], in order to determine the essence of a contagious disease, just the reproductive number is a

***Corresponding Author:**
Mayank Chhabra
Indian Institute of Health Management Research, IIHMR University, Jaipur, India
E.Mail: mayankchhabra98@gmail.com
Article Published: April-June 2020

significant parameter and dynamic behaviour of the model has studied in terms of reproduction values. The initial action taken by governmental authorities to control COVID-19 infection is the implementation of lockdown for maintaining social distance. The purpose is to minimize disease spread rate by locking in high demographic areas, even if full monitoring cannot been attained. Presently, its evaluation closely monitored by governments, researchers, and the public alike. Since these measures have immense pressure on economy and is important for containing the Corona virus, quantitative estimates are imperative to learn the impact of spread, which will help in planning policies. This seems to be an effective way of managing the transmission of the ailments. One naive way to predict the immediate future of the progression of an epidemic is to fit the existing data with reasonable mathematical functions with a few free parameters. The development of complex algorithms and their scientific analysis would facilitated by the improvement of machine learning tools and technologies. The effective mathematical model that is simple yet quite effective to predict the future of the epidemic and the effect of lock-down / social isolation via a time-dependent coefficient of the model for India by using the existing data. This can subsequently predict the disease's biological basis. Yet even such a basic model is effective if there is no formal vaccine or adequate therapeutic supervision.

Taking into account the Indian context, the first case of corona virus ailments have seen in a student who came from Wuhan on 30, Jan 2020 in Kerala. [3-4] According to Govt. of India site, the total confirmed cases are 138885, in which Active Cases 77103, Deaths 4021 and Recovered 57720 individuals by 25 May. [5] Based on situation, we may say that the reason behind rising in cases because of migration of people from affected countries to India and on the contact of those infect person to normal person the disease has been transmitted in whole India. This paper consists of evaluating these metrics and thereby come up with quantitative estimates using customized Susceptible-Infected-Recovered (SIR), taking a corona data from MOHFW site for a period of 30 January until 25 May. As shown in Fig. 1 the Case Fatality Rate (CFR), Growth Rate (GR) and Daily Infection Rate (DIR) from 30 January until 25 May 2020, that how dynamically CFR, GR, and DIR changes with a daily basis death cases, confirmed cases and active cases.

The aim of the research was to determine the potential outcomes for a new global problem: the current outbreak of COVID-19 in India. Forecasting the disease trend and to provide quantitative recommendations using various statistical models. The main objectives of the current study are to demonstrate the forecasts that significantly associated uncertainty continues to grow in the confirmed COVID-19 cases. This paper offers a purposeful approach to the continuation of COVID-19 with a clear yet

effective system. Furthermore, a validation model for prediction of confirmed case, death cases and recovered cases of corona virus ailments for upcoming days with using regression algorithms was performed to validate the accuracy of SIR model.

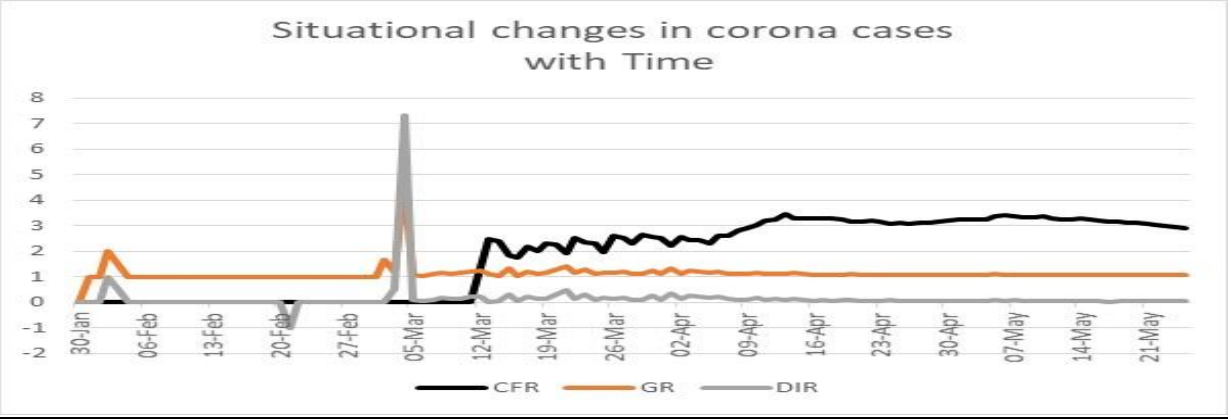


Fig. 1 An overview show of changes in Case Fatality Rate (CFR), Growth Rate (GR), and Daily Infection Rate (DIR)

METHODOLOGY

The prediction model is built from the publically accessible COVID-19 dataset and it was taken from various sites, the sources of the dataset are the Ministry of Health and Family Welfare, India (https://www.mohfw.gov.in, https://www.covid19india.org/). [5,6] The data contains the numbers of cumulative confirmed cases, recovered cases, and death cases from January 30, 2020, to May 25, 2020. This prediction model predicts and estimates the daily number of cumulative confirmed cases, recovered cases, and death cases from May 26, 2020, to June 30, 2020. Data are transformed for handling and preparation for study, stored in csv files format and uploaded on R studio 4.0.0 software with GA (Genetic Algorithms), deSolveand prophet packages. For better understanding, analysis, and calculation of data excel had been used. We have been targeted the data using through different approach of standard epidemiologic techniques of modelling for analysis, estimating and predicting ailments of corona cases that formed at any time. Such epidemiologic model is the susceptible-infected-recovered (SIR) model, which is process with a bunch of differential equations and the other one is regression analysis.

(1) SIR Model

SIR is mathematical model as well as a compartmental model that intermediate with number of susceptible S, number of infectious I, and the number of recovered peoples R. The mathematics behind this model is varies with different differential equation, which has given below: [7]

$$\frac{dS(t)}{dt} = -\frac{\beta}{N}S$$

(1a)

$$\frac{dI(t)}{dt} = (\frac{\beta}{N}S - \gamma)I$$

(1b)

$$\frac{dR(t)}{dt} = \gamma I$$

(1c)

Since, starting to process any modelling approach should be identified the dependent and independent variable. In this direction, we can be considered t (time) in days as an independent variable and other hand significantly two dependent variable which has proportional to each other. Therefore, any of both can be show the same outcomes of epidemic process. The below equation states (2) show the first depended variable with time and (3) focused with other depended variable with time.

S = S(t),I = I(t),R = R(t) (2)

s(t) = S(t)/N, i(t) = I(t)/N, r(t) = R(t)/N (3)

Two parameters are used in equation are beta (β) and gamma (γ) stands as rate of transmission and rate of average recovery which is determined by $1/D$ (where D is duration of infection).The entire population of an area at any time remains constant i.e. N = S+I+R. The parameter γ can calculated by estimating the average period of infectiousness and for β guess that each infected should be make a possible infecting contact within a period. Furthermore, find out R_0 values helping with below equation:

$$R_0 = \frac{\beta}{\gamma}\left(1 - \frac{I_0}{N}\right) \quad (4)$$

Above-mentioned equation, we can assume that the zero individual left at final as infected and further we could define the parameters according to equations. The sturdy part of this model is the calculation R_0 value. R_0 value tells of the diseases contagious. It is the key objective of a new case study by epidemiologists. Simply speaking, R_0 averages how many people can affected over a period by a single infected individual. If the $R_0 < 1$ value indicates a cessation of distribution, On the earlier basis of data the R0 value of COVID-19 has been noted amid 1.5 – 4. [8]

Currently COVID-19 by SIR model, we have considered the population of India approximate 138.72 crores, total confirmed cases, infected cases, and recovered cases. Our focus was to study the estimate time with maximum infected people of corona epidemic, so we could assume that after this maximum peak corona epidemic will be decline to end. Hence, the data taken before lockdown days with continuous different lockdown days until 25 May. We estimate the parameters β and γ for developing and estimating SIR model and its outcomes in three different ways. At firstly initial days data of corona epidemic of India was seen, secondly, the rapidly increasing in total confirmed cases report amid after lockdown time to 25 may 2020 data was reported, and third way to estimate infected cases of India using R software for whole data. The first two ways observed through the real data with estimating average period of infectiousness as well as each infected possible

by each infecting contact in a period [8]. The final way estimating through a package of R called GA package. Lastly the value of β and γ were determined that are respectively (i) beta = 0.26 and gamma = 0.12 (ii) beta = 0.52 and gamma = 0.40 (iii) beta = 0.60 and gamma = 0.39. Afterward, we have worked on mathematical equation and made a self-function program in R notebook for getting the outcomes for evolution of SIR model. For simplified calculation and validation we have used excel. We have been calculated R_0 from using different values of beta and gamma parameters respectively 2.16, 1.3, 1.25. This was done with this simplified SIR mathematical equation:

$S_n = S_{n-1}-((S_{n-1})*(\beta*I_{n-1})$

$I_n = I_n-1+(S_{n-1}/S)*(\beta*I_{n-1})-(I_{n-1}*Y)$

$R_n = R_{n-1}+(I_{n-1}*Y)$

Some assumption were considered related to rate changes of dependent variables, such as no one add up with susceptible cases, furthermore we should be ignored the birth and immigration. In addition, included point that recovered will not infect further and can be categorised in immune.

(2) Regression Model

Analysis of Regression is a type of forecasting modelling approach that explores the relation amid dependent and independent variable. This technique has used to anticipate the causal effect relationship between the variables and to model the time series. A significant method for data processing and analysis is regression analysis. In addition, many division of regression such are linear regression, polynomial regression, and logistic regression.

Currently a linear regression was used with R packages prophet for forecasting the values of upcoming cases of corona virus in India. Linear regression is easy and simplifies regression for predicting and analysis amid dependent and independent variable. Furthermore, the result from outcome was interpreted of linear regression in which a slope and intercept values will be define for new values or predicted values of output variable. Although, we have been using the specific variable with concurrently dependent and independent respective manner and concluded our test. The basic equation of linear regression is –

$Y = \beta_0 + \beta_1*x + \varepsilon$

The above equation should be create a straight line and used mostly in analysis. The above equation constant β_0, β_1 independent variable stands for intercept and slope and ε present as error rate.

RESULTS

Initially, the entire population is susceptible to the infection, and very negligible amount of people bring this infection into the country. As the disease spreads, infectious section starts increasing, along with the number of total recovered cases. In India, on 30 January 2020 the very first COVID-19 case has been reported. However, the number of people affected during the month of February was 3 and continued throughout the whole month. In the month of March 20, the major increase in disease spread

apparently started. Alterations in confirmed cases, cases of death and total recovered cases from 30 Jan 2020 to 25 May 2020 are shown in figure 2. However, March data suggest that the disease distribution has changed considerably.

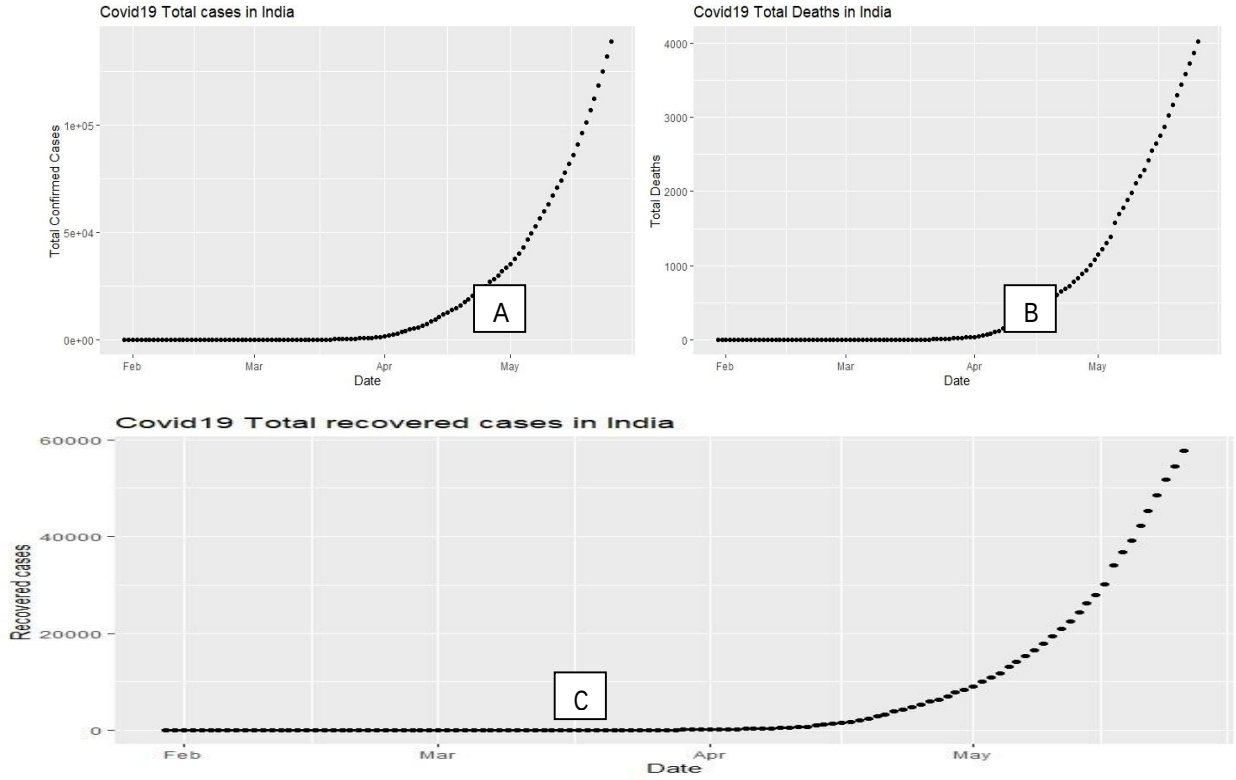


Fig. 2: Alteration in number of cases of COVID-19 in India (A) Total confirmed cases (B) total death cases (C) total recovered cases

Currently, we used data from 30 Jan to 25 May 2020 in order to forecast relevant COVID-19 pandemic from 26 May 2020 to 30 June 2020. Figure 3 and illustrate the predicted values with the actual values of total confirmed cases, total death cases and total recovered cases. Therefore, the dotted represents the actual values of cases and the blue line represents the forecast. The findings of reported cases showed strong correlations with real data days in the log basis. Thereby the log base confirmed cases predicted that are shown in fig.4.

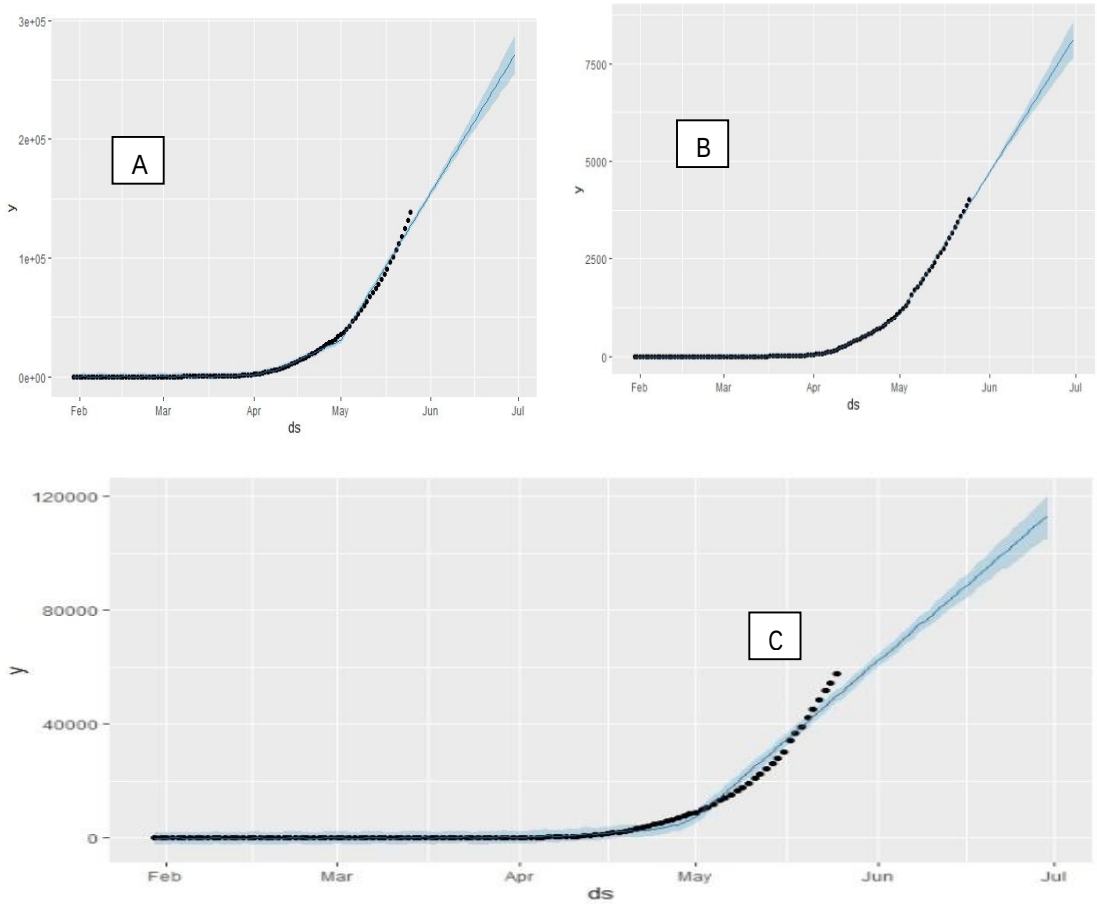


Fig. 3: Plot show the actual values with predicted values with days (A) y = Total Confirmed cases and ds = Date (B) y = Total death cases and ds = Date (C) y = Total Recovered cases and ds = Date

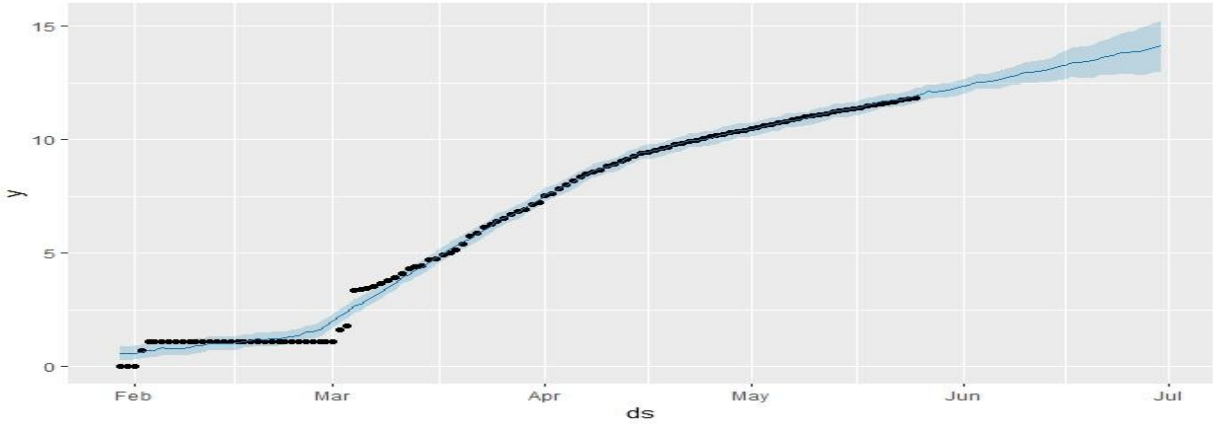


Fig. 4: *Show the linear growth trend in India between* the actual values with predicted values verses days (y = Total Confirmed cases in log values and ds = Date

In addition, the best fit among real values and predictive values has been evaluated in the implementation of the prediction and regression model; the effects of confirmed cases, death cases and total recovered cases with the prediction values demonstrate in figures 5.

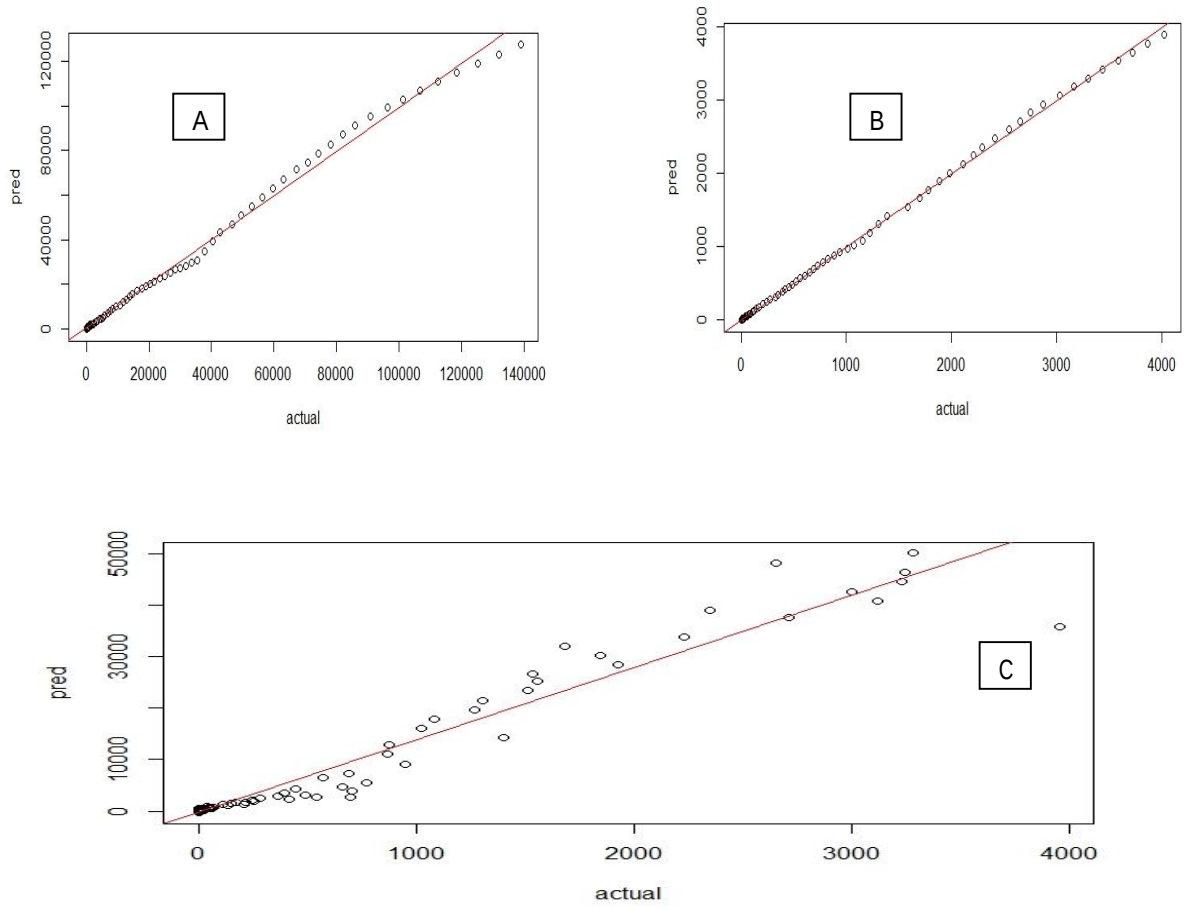


Fig.5: Fitting plot by linear regression (A) between actual (total death cases) and predict (predicted total death cases) (B) between actual (total confirmed cases) and predict (predicted total confirmed cases) (C) between actual (total recovered cases) and predict (predicted total confirmed cases)

The key input parameters were calculated with a simple linear regression analysis to establish the pattern for the all-variable projection in India. The model was shown to be highly capable of predicting (Table1).

Furthermore, in analysis to predict the day of maximum infected peak, we have demonstrate the SIR model plots and analyzes with different β and γ values considering the particular susceptible population, and we can conclude that this diseases will be in our control as soon as possible. The explanation of the SIR model in the current study is shown in Table 2 compared to different phases.

Table 1: Positive result of forecasting and linear regression model between actual and predicted cases

Parameters	Total confirmed cases	Log bases total confirmed cases	Total Death cases	New confirmed cases	New death cases	New Recovered	Recovered cases
Correlation (R)	0.9980	0.9982	0.9985	0.9882	0.9670	0.97418	0.97419
Multiple R square (R^2)	0.9961	0.9965	0.9994	0.9823	0.9351	0.949	0.9491
Adjusted R square	0.9961	0.9965	0.9994	0.9821	0.9345	0.9486	0.9486
F-statistic	29470	32870	182000	6365	1657	2143	2141
Residual standard error	2145	0.2403	26.91	234.9	11.8	193.9	2943

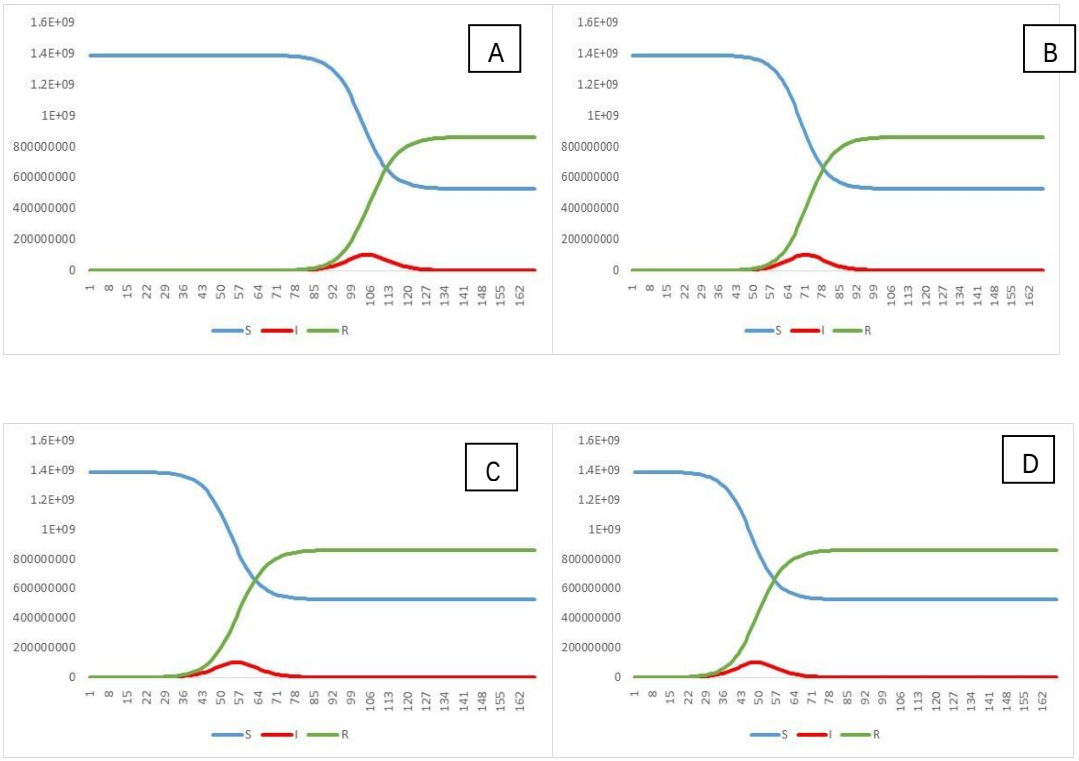
The percentages of maximum cases of infections showed promising results at estimated 18.73%, 3.00% and 7-7.4%, with 2.16, 1.3 and 1.55 R_0 values, respectively. In addition, for a SIR model, which is shown in fig.6 the best parameters, are selected.

Table 2: Observation and outcomes of SIR model

Stage L_0(A) Susceptible = 138297451 and Total days = 55			
Parameters	Beta = 0.26, Gamma = 0.12	Beta = 0.52, Gamma = 0.40	Beta = 0.60, Gamma = 0.39
Start Date	30-01-2020	30-01-2020	30-01-2020
IIC	1	1	1
EMI	260147523.3 (18.75%)	41611711.66 (2.99%)	102865790.8 (7%)
Day at EMI	161	168	105
R_0	2.16	1.3	1.552
Stage L_1 (B) Susceptible = 1387296846 and Total days =21			
Parameters	Beta = 0.26, Gamma = 0.12	Beta = 0.52, Gamma = 0.40	Beta = 0.52, Gamma = 0.40
Start Date	25-03-2020	25-03-2020	25-03-2020
IIC	606	606	606
EMI	260101506.6 (18.74%)	41595961.86 (2.99%)	100856201.2 (7.2%)
Day at EMI	112	112	73
R_0	2.16	1.3	1.552
Stage L_2 (C) Susceptible = 1387285519 and Total Days = 18			
Parameters	Beta = 0.26, Gamma = 0.12	Beta = 0.52, Gamma = 0.40	Beta = 0.55, Gamma = 0.44
Start Date	15-04-2020	15-04-2020	15-04-2020
IIC	11933	11933	11933
EMI	260073939.4 (18.74%)	41614486.24 (2.99%)	1028865416.7 (7.4%)
Day at EMI	90	85	56
R_0	2.16	1.3	1.552

Stage L_3 (D) Susceptible = 1387254616 and Total Days =14			
Parameters	Beta = 0.26, Gamma = 0.12	Beta = 0.52, Gamma = 0.40	Beta = 0.55, Gamma = 0.44
Start Date	04-05-2020	04-05-2020	04-05-2020
IIC	42836	42836	42836
EMI	260068039.6 (18.74%)	41613575.33 (2.99%)	102883854.7 (7.41%)
Day at EMI	90	85	49
R_0	2.16	1.3	1.552
Stage L_3 (E) Susceptible = 1387201283 and Total Days -14			
Parameters	Beta = 0.26, Gamma = 0.12	Beta = 0.52, Gamma = 0.40	Beta = 0.55, Gamma = 0.44
Start Date	18-05-20	18-05-2020	18-05-2020
IIC	96169	96169	96169
EMI	259938238.5 (18.73%)	41709560.24 (3.00%)	102934188.3 (7.41%)
Day at EMI	73	67	45
R_0	2.16	1.3	1.552

(IIC=Initial Infected Cases, EMI=Estimated Maximum Infected) [(a) Total population consider = 1387297451 (b) Average CFR & Growth rate between 30 January to 25 May 2020 = 1.83% & 1.21 (121%)]



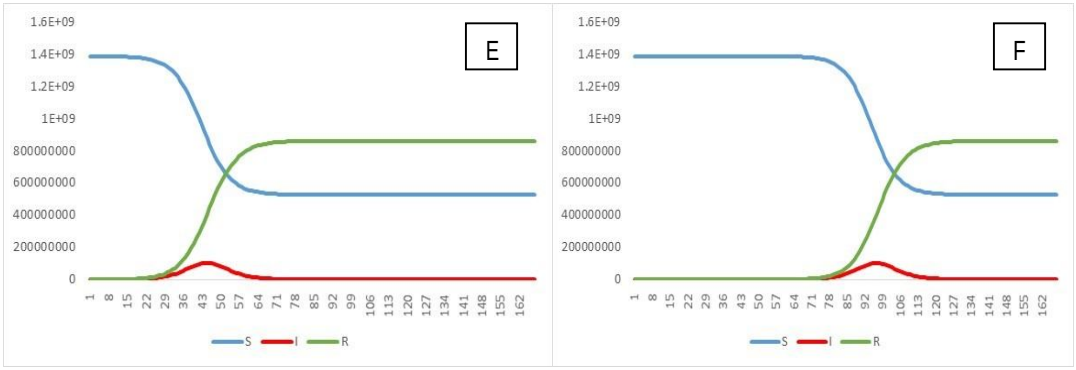


Fig 6: (A) IIC=1 start with l_0 stage (B) IIC=606 start with L_1 stage (C) IIC=11933 start with L_2 stage (D) IIC=42836 start with L_3 (E) IIC=96169 start with L_4. In all plot β=0.60 and γ=0.39 were taken (F) IIC=5 start with first week of March

In addition, correlation analysis with regression was carried out based on previous studies to determine the significance and subsequent severity of the connection between the outputs of all included variables and the inputs from certain data. The connection between output variables and all input variables showed a significant positive correlation (Table 3). Total confirmed cases, the death count, and recovered cases were determined from the regression model outcomes that were calculated in table 3. Thereby, established predicted values (output variable) with all possible input variable using linear regression model (Table 4). All predicted estimation based on table 3 outcomes.

Table 3: Regression analysis with correlation analysis show the all possible outcomes

	Summary	**Days**	**Total Confirmed cases**	**Death cases**	**Recovered cases**	**Active cases**
Total Confirmed cases	Correlation MultipleR^2 AdjustedR^2 P-value	0.7836, 0.6142, 0.6142, 1.25E-09	1	0.9985, 0.9971, 0.9971, 0.0967	0.9871, 0.9745, 0.9742, 1.72e-06	0.9947, 0.9896, 0.9895, 1.35e-06
Death cases	Correlation MultipleR^2 AdjustedR^2 P-value	0.7931 0.629, 0.6258, 4.74e-10	0.9985, 0.9971, 0.9971, 0.173	1	0.9792, 0.9589, 0.9586, 6.53e-05	0.9973, 0.9948, 0.9947, 4.73e-08
Recovered cases	Correlation MultipleR^2 AdjustedR^2 P-value	0.7064, 0.4991, 0.4947, 2.29e-07	0.9871, 0.9745, 0.9742, 1.47e-08	0.9792, 0.9589, 0.9586, 3.31e-07	1	0.9657, 0.9326, 0.932, 1.31e-08
Active cases	Correlation	0.8227,	0.9947,	0.9973,	0.9657,	1

MultipleR^2	0.677,	0.9896,	0.9948,	0.9326,
AdjustedR^2	0.672,	0.9895,	0.9947,	0.932,
P-value	2.89e-11	2.37e-05	5.08e-07	2.72e-05

Table 4: Prediction of all possible estimation of linear regression model for COVID-19 in India

Cases	Prediction Till 30 June	Prediction till 30 July
Total Confirmed Cases	400000-460000	760000-800000
Total Death Cases	8050-10000	15000-16500
Total Recovered Cases	200000-250000	350000-420000

In addition, an estimated and predicted outcome of the linear regression with the forecast package and SIR model has been detected and obtained. The results of the prediction are shown in Table 5 using two models and the expected values from the confirmed cases of corona, death and recovery cases by the regression analysis forecast, and the confirmation case of SIR by the component anticipated by the SIR model. We utilized 35 days of test set for both modelling techniques.

Table 6: Prediction of number of corona cases in 2019 in India for 35 days by SIR and regression model with forecast using R and excel

Date	Regression Prediction			SIR predict cases
	Confirmed cases	Death cases	Recovered	
26-05-2020	145271	4173	60865	146000
27-05-2020	151796	4328	64146	153000
28-05-2020	158469	4482	67454	159000
29-05-2020	165371	4639	70860	166000
30-05-2020	172425	4797	74382	174000
31-05-2020	179719	4960	78060	181000
01-06-2020	187215	5129	81772	189000
02-06-2020	194822	5308	85550	197000
03-06-2020	202527	5476	89464	205000
04-06-2020	210381	5650	93406	213000
05-06-2020	218462	5827	97447	221000
06-06-2020	226697	6005	101602	229000
07-06-2020	235171	6189	105914	238000
08-06-2020	243848	6378	110260	247000
09-06-2020	252634	6570	114672	256000
10-06-2020	261520	6765	119220	265000
11-06-2020	270554	6959	123796	275000
12-06-2020	279816	7155	128470	284000
13-06-2020	289231	7354	133260	294000
14-06-2020	298885	7557	138205	304000

15-06-2020	308742	7766	143185	314000
16-06-2020	318709	7979	148232	324000
17-06-2020	328775	8193	153413	335000
18-06-2020	338989	8407	158623	345000
19-06-2020	349431	8624	163931	356000
20-06-2020	360026	8843	169354	367000
21-06-2020	370861	9066	174934	378000
22-06-2020	381898	9295	180548	389000
23-06-2020	393045	9528	186228	401000
24-06-2020	404291	9762	192043	413000
25-06-2020	415686	9996	197887	424000
26-06-2020	427309	10233	203829	436000
27-06-2020	439084	10471	20988	449000
28-06-2020	451099	10715	216100	461000
29-06-2020	463316	10965	222348	474000
30-06-2020	475644	11217	228662	486000

Discussion

The exponential rise suggests that concurrent development has continued in the upcoming months since even the monitoring approaches implemented by the Indian Government seem to be notoriously strict and quite well functioning. However, the sufferers recovered can even be easily analyzed and the fatality rate can be controlled by implementing of linear trends. Hospital supplies and medical facilities can proceed at a rapid pace in order to develop the country for rapid rise. With recent initiatives and consultations, the Indian Government looks forward to flattening the curve.

Most analyses of COVID-19 Data have been carried out earlier. Throughout this direction, the analysis implemented by Das [9] was based on an epidemiological model and forecast basic reproduction rates at state and national levels. In contrast, Ray et al. [10] made a predictive case-count model in India. They addressed and predicted potential approaches over such a span of different intensities. Both research works have been designed with SIR models. In the evaluation and predictions of the alteration in the spread in COVID-19 disease, Gaurav et al. [11] demonstrated 2 machine learning models, SEIR and Regression. In this analysis, demonstrated a comprehensive analysis in India with varying phases of lock-down days and predict various kinds of probable consequences, showing a robust representation of the corona cases of present cases in India with growth possibilities. In addition, maximum day and peak infections are calculated by the developed SIR model and a forecast has been identified for confirmed cases for the coming days. The results of the regression analysis have shown that estimation and prediction for the next few days have been noteworthy. For all confirmed cases, death cases, and recovered cases, our forecast has been calculated by predicting time series

estimates.Figure 7 show a comparative study of total confirmed cases of COVID-19 in India with regression and SIR model prediction.

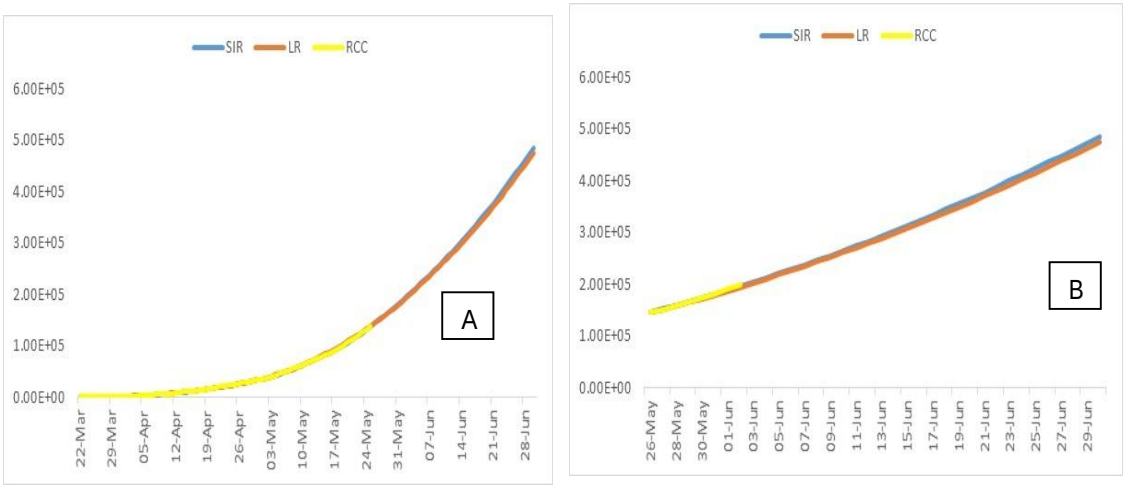


Fig. 7: Line plot show the comparative confirmed case with Real Confirmed Cases (RCC), linear regressed (LR) Confirmed case, SIR confirmed cases.

Prediction has been occupied by the results from the current study until June. In addition, the fast-growing corona virus patient implies that it will increase exponentially but may be steady at certain moments of day. Initially locked, initiatives like immigration bans relatively early in human existence, the use of the BCG to regulate tuberculosis for populations with prospective adverse influence of COVID-19 [12-13], malaria, and anti-malaria exposure[14], and easier propagation to hot and moist weather [15-16] managed to keep the number of cases below their current rating. Currently, however, no evidence is adequate to sustain these inferences, even though some of them are currently underway (17).

Different drug types could be used for corona disease prevention. Until now, vaccines against corona virus diseases have still not been discovered in India as well as worldwide potentials with strong test results. Plasma therapy has also been used to obtain preventive results in different regions of India. Physicians who provide essential care to health workers in this country must be covered by the scientific standards approved. The number of cases will rise exponentially due to the negligence of individuals and groups in the future. The high point is still reached too; the government must therefore be more vigilant and implement stringent measures.Furthermore, there should be a combative increasing trend in medical facilities throughout the country.

CONCLUSION

The study analyzed and predicted improvements in the propagation of the COVID-19 disease by using two SIR and linear regression machine

learning models. We examined the results and estimated that the percentage of individuals infected in India rapidly dramatically increased by a million. The reproduction value of $R_0$was then calculated using the SIR model to carry out a comprehensive study of corona disease at various stages of India's cases. Therefore, for the next 36 days from 26 May 2020 to 30 June 2020, we expected the number of confirmed cases of COVID-19 in India. The effects of social distancing between communities may be estimated after each lock-down phase, so India is at high risk of becoming a part of community transfer due to the reported infringement on the part of individuals as well as other demographic and socio-economic issues. When such an occurrence happens, the predictions made to the design study using the latest epidemiological models might be inaccurate.The model study with realistic parameters set shows that the Covid-19 pandemic will be at its peak around the end of July and start decline around the month of August, for predicted values the estimates were given in Table 4 .The data is changing dynamically commenting for long future can affect the accuracy of the model. The forecasts can, therefore, change along with real-time data changes every day. Consequently, the findings of this study can only be used for an empirical and reasonable comprehension of the essence of the pandemic, but no verdict or a new regulation is prudent.

REFERENCES

1. Jiang, F., Deng, L., Zhang, L., Cai, Y., Cheung, C.W. and Xia, Z., 2020. Review of the clinical characteristics of coronavirus disease 2019 (COVID-19). Journal of General Internal Medicine, pp.1-5.
2. Wu, Z. and McGoogan, J.M., 2020. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. Jama
3. Kerala reports first confirmed coronavirus case in India - India News. https://www.indiatoday.in/india/story/kerala-reports-first-confirmed-novel-coronavirus-case-in-india-1641593-2020-01-30. Accessed March 28, 2020.
4. Bhatnagar T, Murhekar M V, Soneja M, et al. Lopinavir/ritonavir combination therapy amongst symptomatic coronavirus disease 2019 patients in India: Protocol for restricted public health emergency use. *Indian J Med Res*. 2020;(PG-). doi:10.4103/ijmr.IJMR_502_20
5. MoHFW | Home. https://www.mohfw.gov.in/. Accessed May 25, 2020.
6. COVID-19 Tracker | India. https://www.covid19india.org/. Accessed May 25, 2020.

7. Herbert W Hethcote. The mathematics of infectious diseases. SIAM review, 42(4):599–653, 2000.
8. Mandal, S., Bhatnagar, T., Arinaminpathy, N., Agarwal, A., Chowdhury, A., Murhekar, M., Gangakhedkar, R.R. and Sarkar, S., 2020. Prudent public health intervention strategies to control the coronavirus disease 2019 transmission in India: A mathematical model-based approach. The Indian journal of medical research.
9. Das S. Prediction of COVID-19 Disease Progression in India : Under the Effect of National Lockdown. April 2020. http://arxiv.org/abs/2004.03147. Accessed April 21, 2020.
10. Ray D, Salvatore M, Bhattacharyya R, et al. Predictions, role of interventions and effects of a historic national lockdown in India's response to the COVID-19 pandemic: data science call to arms. medRxiv. April 2020:2020.04.15.20067256. doi:10.1101/2020.04.15.20067256
11. Gaurav Pandeya, Poonam Chaudharya, RajanGuptab, SaibalPalc, et al. SEIR and Regression Model based COVID-19 outbreak predictions in India.medRxiv. doi.org/10.1101/2020.04.01.20049825
12. Hegarty PK, Service NH, KamatAM,Dinardo A. BCG vaccination may be protective against Covid-19. 2020;(March). doi:10.13140/RG.2.2.35948.10880
13. BacilleCalmette-Guérin (BCG) vaccination and COVID-19. https://www.who.int/news-room/commentaries/detail/bacille-calmette-guérin-(bcg)-vaccination-and-covid-19. Accessed April 21, 2020.
14. Mayo Clinic is studying antimalarial drug for COVID-19 treatment | Healthcare Finance News. https://www.healthcarefinancenews.com/node/139936. Accessed April 21, 2020\
15. Why India has the upper hand against COVID-19 | World Economic Forum. https://www.weforum.org/agenda/2020/04/india-covid19-coronavirus-response-kerala-uttar-pradesh/. Accessed April 20, 2020.
16. Bukhari Q, Jameel Y. Will Coronavirus Pandemic Diminish by Summer? *SSRN Electron J.* March 2020. doi:10.2139/ssrn.3556998
17. Hydroxychloroquine Post Exposure Prophylaxis for Coronavirus Disease (COVID-19) - Full Text View - ClinicalTrials.gov. https://clinicaltrials.gov/ct2/show/NCT04318444. Accessed April 21, 2020

Cite this article as:

Agrawal T., Chhabra M. Regression and SIR model based COVID-19 pandemic forecasting in India. Int. J. Pharm. Technol. Biotechnol. 2020; 7(2): 08-23.