Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/mjafi

Original Article

Predictive models of COVID-19 in India: A rapid review

Atul Kotwal [a], Arun Kumar Yadav [b,*], Jyoti Yadav [c], Jyoti Kotwal [d], Sudhir Khune [e]

[a] MG (Med), South Western Command, C/o 56 APO, India
[b] Associate Professor, Department of Community Medicine, Armed Forces Medical College, Pune, India
[c] Independent Researcher, 1006, Sandalwood Building, Green Valley, Wanowarie, Pune, India
[d] Professor, (Pathology), Sir Ganga Ram Hospital, New Delhi, India
[e] Resident, Department of Community Medicine, Armed Forces Medical College, Pune, India

ARTICLE INFO

Article history:
Received 27 April 2020
Accepted 2 June 2020
Available online 17 June 2020

Keywords:
Systematic review
Mathematical modeling
COVID-19
India

ABSTRACT

Background: The mathematical modelling of coronavirus disease-19 (COVID-19) pandemic has been attempted by a wide range of researchers from the very beginning of cases in India. Initial analysis of available models revealed large variations in scope, assumptions, predictions, course, effect of interventions, effect on health-care services, and so on. Thus, a rapid review was conducted for narrative synthesis and to assess correlation between predicted and actual values of cases in India.

Methods: A comprehensive, two-step search strategy was adopted, wherein the databases such as Medline, google scholar, MedRxiv, and BioRxiv were searched. Later, hand searching for the articles and contacting known modelers for unpublished models was resorted. The data from the included studies were extracted by the two investigators independently and checked by third researcher.

Results: Based on the literature search, 30 articles were included in this review. As narrative synthesis, data from the studies were summarized in terms of assumptions, model used, predictions, main recommendations, and findings. The Pearson's correlation coefficient (r) between predicted and actual values ($n = 20$) was 0.7 ($p = 0.002$) with $R2 = 0.49$. For Susceptible, Infected, Recovered (SIR) and its variant models ($n = 16$) 'r' was 0.65 ($p = 0.02$). The correlation for long-term predictions could not be assessed due to paucity of information.

Conclusion: Review has shown the importance of assumptions and strong correlation between short-term projections but uncertainties for long-term predictions. Thus, short-term predictions may be revised as more and more data become available. The assumptions too need to expand and firm up as the pandemic evolves.

* *Corresponding author.*
E-mail address: arunyadavpsm@gmail.com (A.K. Yadav).

https://doi.org/10.1016/j.mjafi.2020.06.001

Introduction

Outbreaks of infectious diseases encompassing entire nations or civilizations are known to humankind since antiquity. To list a few from the long list, biblical pharaonic plagues in Ancient Egypt (1715 BC)[1], the 'cocoliztli' epidemics in Mesoamerican native population during the 16th century,[2] bubonic plague in Europe (1348), pandemic influenza (1918–1919) affecting America, Europe, Asia, and Africa, less severe influenza pandemics in 1957 and 1963, Human Immuno-deficiency virus (HIV) (1981), Severe acute respiratory syndrome (SARS) (2003), pandemic H1N1 (2009), Middle East respiratory syndrome (MERS) (2012), avian influenza (H7N9), and Ebola (2014–16).[3]

The quest of scientists and researchers to predict the dynamics and progress of a novel epidemic/pandemic through the population results in use of various techniques and approaches of mathematical modeling and in turn leads to a plethora of models with varying assumptions and approaches.

The initial mathematical model credit goes to Bernoulli et al. [4] who analyzed the mortality due to smallpox in England, wherein he showed that inoculation against the virus would increase the life expectancy at birth by approximately three years.

Later, the foundations of mathematical modeling for infectious diseases were established by Kermack et al.[5] The early models classified the persons as susceptible, infected (infectious), and recovered (SIR).[6] Further improvements saw more complex compartmental models, utilization of age structure, stochastic transmission models, and so on. Over the years and with each epidemic/pandemic, newer approaches and softwares including machine learning are being used for mathematical modeling. The three main categories of infectious disease models are as follows: statistical based; mathematical/mechanistic state space; and empirical/machine learning based.[7]

Mathematical models for infectious diseases and their statistical tools have become an integral part of the inputs for planning control and mitigation measures. These models provide us opportunities to test various strategies in simulations before applying them in populations or individuals.[8] Mathematical modeling for infectious diseases uses many sources of data and various assumptions. The predictions made by the model's virtual world should be relevant to reality, and the model needs to be as close to the reality in real world. Robust predictions can never be expected from vague/incomplete/wrong assumptions. Efforts toward simplification, approximation, idealization, and abstraction lead to all models as partial descriptions of the mechanisms operating in reality. Thus, robustness of each model needs to be assessed based on whether its assumptions approximately correspond to reality or not.

SARS-Cov-2, emerged on the horizon in late December 2019 when few local health authorities in China reported clusters of patients with pneumonia of unknown cause. The surveillance mechanism established during 2003 SARS outbreak helped in identification of the pathogen (SARS-CoV-2).[9,10]

The SARS-CoV-2 infection has spread in 210 countries as of 24 April 2020 with 2,697,316 cases and 188,857 fatalities.[11] In India, the first case was reported on 30 January 2020, and the numbers gradually increased till 03 March after which the per day increase has been faster with current tally of 21,700 infected and 686 dead.[12]

The mathematical modeling of this evolving pandemic in India has been attempted by a wide range of researchers from the very beginning of cases in India. An initial analysis of these models regarding India revealed large variations in scope, assumptions, prediction on numbers, the course of the pandemic in India, effect of various interventions, effect on health care services, and so on.

The literature search did not reveal any review of available models and thus, this study was conducted as a rapid review of the mathematical models used for prediction of Coronavirus disease (COVID-19) in India for narrative synthesis and to assess correlation between predicted and actual value of cases in India.

Materials and Methods

A review protocol was prepared and uploaded in Prospero for registration (Registration ID - CRD42020180513). All articles on mathematical models on the COVID-19 on India were included in the study with predefined inclusion and exclusion criteria. Because this review pertains to different types of mathematical modeling, it did not fit into any types of present guidelines available for systematic review, and thus it is being titled as a rapid review, and narrative synthesis was conducted as first objective. However, the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines were followed to the extent possible.

Inclusion and exclusion criteria

To be included in this rapid review, eligible studies had to meet the following criteria: (a) study of predictive modeling; and (b) studies carried out only for India or as part of multi-countries predictive model with India as one of the countries. The studies that were not included in the rapid review were as follows: (a) perspective studies without modeling; and (b) studies or reviews without modeling. Studies published in English were only included in the rapid review.

Literature search (search strategy)

A comprehensive two-step search strategy was formulated and adopted. First, the literature searches through databases (Medline, google scholar, MedRxiv and BioRxiv) were carried out. All the articles submitted to these databases for COVID were searched. The literature search was carried out till 10:00 AM on 22 April 2020 (IST). The search strategy for Medline using Pubmed has been provided as Supplementary File No.1. In second step, hand searching of the articles was carried out, and known modelers were also contacted for unpublished modeling of the Indian COVID data.

Data extraction

The data from the included studies were extracted on data extraction form by the two investigators, (2) and (3) independently. In case of the discordance in the data, the same was resolved with discussion involving the third senior researcher (1). The extracted data were tabulated in the form of two tables. The data were extracted for the following variables: type of mathematical models; software used; profession of modelers; effect of lockdown studied or not; assumptions used; peak infected numbers; data and data sources. Main summary measures were peak infection rate and predicted value for the number of COVID cases. Because the data used for mathematical modeling is based on the hard data acquired from different sources, the predicted number may change in individual study, based on mathematical models used and assumptions taken. The data on peak infected infection if feasible would be averaged out in models giving predictions on full cycle on epidemics. In few studies, based on the mathematical modeling, the predicted value was calculated if not provided in the manuscript. Predicted values were plotted against actual values of the same date of epidemic. The relationship between predicted and actual value was explored using coefficient of determination. For the non-quantitative variables, qualitative synthesis was attempted.

The statistical software StataCorp. 2013, Stata Statistical Software: Release 13, College Station, TX: StataCorp LP was used for statistical analysis. The p value of less than 0.05 was taken as statistically significant.

Results

Based on the literature search, 30 studies were selected for inclusion in the rapid review.[13–42] The PRISMA chart is as shown in Fig. 1. The study characteristics including variables studied, lockdown effect, date of data collection, and peak infected numbers are shown in Table 1.

The data extracted from 30 research articles showed that the modeling on the data available in public domain started as early as 21 March 2020.[13] The latest data used for modeling was of 13 April 2020.[40] Mathematical techniques used for modeling were also varied. Types of models used have been depicted in Fig. 2, which shows that most studies (17, 56%) were published using SIR model or its variant. The assumptions made by different models regarding R_0 (R Naught), infectious period, recovery time, serial interval, and so on, are given in Table 2. Varied professionals have indulged in modeling on data available in various public platform for India, these included doctors or medical researchers (10, 34%), mathematician/physicist/engineer (13, 45%), biostatistician (2, 7%), and others (4, 14%) similar to in horticulture, and so on. Most common statistical software used was R (11, 37%) followed by Python (6, 20%) and Matrix laboratory (MATLAB) (4, 13%). A total of 9 studies (30%) did not mention any software.

Few of the models provided only predictive models without predicting the course of pandemic in India,[16,19,28,36,38] whereas few others predicted the entire course with peak infected values of pandemic in India.[13,14,16,40]

Of 30 mathematical models, 20 have predicted the number of cases in future. Of these 20, four predictions were outliers and hence were not plotted.[19,26,29,32] The scatter diagram for predicted and actual values at a particular timeline of sixteen studies in model is shown in Fig. 3. The Pearson's correlation coefficient for short-term predictions is 0.7 ($p = 0.002$), indicating strong correlation and the coefficient of determination (R^2) is 0.49, signifying that 49% variation in actual data is being

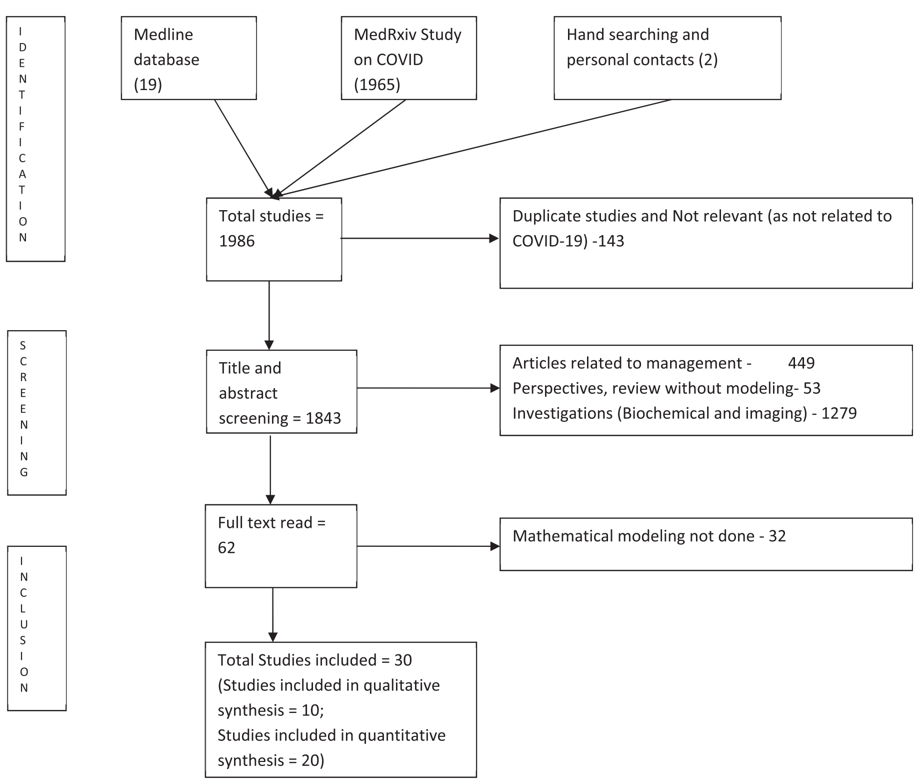

Fig. 1 – PRISMA chart.

Table 1 – Study characteristics of mathematical modeling studies.

S No	Reference	Characteristics and factors studied	Last date of data collection	Peak infected numbers/predicted number
1.	13	‘q’ metric - varying value of quarantine, infected population, peak numbers, age stratified ICU admissions, and fatality rate	21 March 2020	364 million cases and 1.56 million deaths overall with peak by mid-July 2020 with q1 matrix
2	14	Mathematical framework using exponential and sigmoid type function. Result of different states using the mathematical framework	13 April 2020	Values 10^4 and 10^5, Peak of cases expected in end May/June, respectively in different states
3	15	Quarantine (hidden nodes) and effect of lockdown and relaxations. Lockdown in multiple phases is studied	29March 2020	197,200 cases after relaxation of 6 days after 15 days lockdown
4	16	Rate of growth is different in different state. Effect of climate and population density was studied	10 April 2020	NA
5	17	Prediction of the number of deaths	26 March 2020	Projected death rate (n) is 211 and467 at the end of the 5th and 6th week from 26 March
6	18	Short-term forecasting for maximum cases and new cases	02 April 2020	12,500 cases on 20 April 2020
7	19	Present situation of India, theoretical aspects of R(0)	28 March 2020	Not mentioned
8	20	Long- and short-term effects of initial 21 days lockdown and study alternative explanation for slower growth rate like temperature	07 April 20	9181 cases on April 30
9	21	Impact of social distancing measures - workplace non-attendance, school closure, lockdown—and their efficacy with duration was investigated	25 March 2020	167 million on 02 July 2020
10	22	Parameters and indicators that quantify the growth and spread of diseases. Infected population, peak infected number of cases, were calculated	07 April 2020	22,000 in last week of April. By July India will get over COVID-19
11	23	Predictions; R0; and Public health preparedness. Reproduction number, Herd Immunity, Requirements for hospitalization and ICU, cumulative cases	03 April 2020	2,49,635 cases and 18,739 deaths until the end of April
12	24	Peak date and total number of infections considering the lockdown	11April 2020	2.2×10^5cases on 31 May 2020
13	25	Model fitting and predictions for number of cases for next two weeks.	30 March 2020	5300–6135 cases till 13 April 2020
14	26	Effect of social distancing, infected population, peak numbers, and peak date	31 March 2020	17,525,869 peak cases in third week of June
15	27	Estimated parameters such as R(0), infected population, peak numbers, mean serial interval, daily epidemic growth rate, doubling time, CFR	12 April 2020	Mid-July to early August 2020 with around 12.5% of population will be infected
16	28	Effect of power-law behavior: transition from exponential regime to power law may act as an indicator of flattening of curve	7 April 2020	Not mentioned
17	29	Estimation of new cases and effect of lockdown	01 April 2020	31 days to all population in unconstrained environment
18	30	Analysis of age and sex of COVID-19 cases, using SIR model range of contact rate and public health intervention was assessed.	04 April 2020	5583 to 13,785 active cases by 14 April 2020
19	31	Forecasting COVID-19 for number of new cases, deaths, and drop down in recovery rate	28 March 2020	5200 or 6378 cases and 197 deaths by 29 April 2020
20	32	Predictions for COVID-19 outbreak in India. Modeled along with social distancing. Peak infected number of cases	30 March 2020	13,000 final cases by end of May
21	33	Effect of lockdown, short-term predictions, effect of social distancing, effect of religious event, identification of prominent clusters	08 April 2020	86,864 cases by 02 May 2020
22	34	Estimating the final epidemic size for COVID-19	08 April 2020	Range of final cases between 16,916 and 36,323
23	35	COVID-19 case data of 5 countries, short-term forecasting, case fatality rate considering lockdown	04 April 2020	800 cases on 14 April
24	36	Effect of travel restriction and quarantine, delay in introduction of infection in India and estimated infected cases, percent reduction in hypothetical peak	26 February 2020	46% of infected travelers would not be detected by thermal screening at airport exit and entry

25	[37]	Impact of lockdown, contact and non-contact transmissions on infection dynamics, premature withdrawal of lockdown is likely to promote a rapid and sharper infection peak, infected population, impact of lockdown duration on primary and secondary peak dynamics	04 April 2020	Not mentioned
26	[38]	Impact of various parameters such as weather, vaccination (BCG), lockdown on COVID-19.	09 April 2020	Not mentioned
27	[39]	Cases in 6 countries including 3 Indian states with power law exponent with future predictions	01 Apr 2020	2412 to 10,307 cases as on 14 April 2020
28	[40]	Detection rate of SARS-CoV-2 infections based on data on age distribution, infection fatality rates, reported death, and confirmed case	08 April 2020	1,59,939 infections instead of 5480 confirm cases on 08 April 2020
29	[41]	Interpret existing data with respect to other countries and deviation trend due to lockdown	10 April 2020	Entire population 1st week of July 2020
30	[42]	Risk assessment of COVID-19 pandemic in India and impact of lockdown	03 April 2020	50,00,000 cases in June 2020

SIR, susceptible, infected (infectious), and recovered;

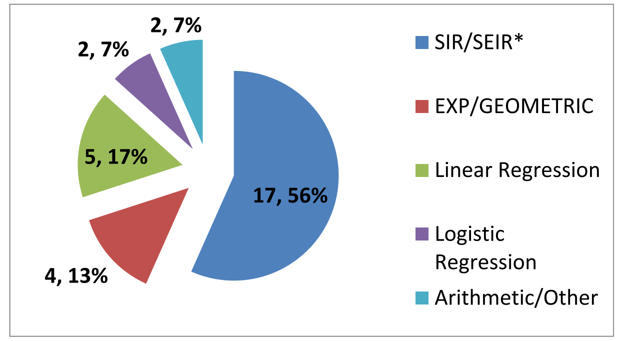

Fig. 2 – Type of mathematical modeling used for modeling COVID-19, *SIR, susceptible, infectious, recovered; SEIR, susceptible, exposed, infectious, recovered.

accounted for by the predicted value. The Pearson's correlation between SIR or its variant models (total 12) was 0.65 ($p = 0.02$), indicating moderate to strong correlation.[43] For other type of models' subgroup analysis could not be performed due to less numbers of studies in subgroups.

Discussion

In India, the first case was detected on 30 January 2020 and the number of cases from 02 February 2020 to 01 March 2020 remained three. The cases are being regularly reported from 02 March onward. There was a sudden increase in number of cases on 04 March 2020 due to various reasons, one of those being change in testing policy.[13] Thereafter, the data regarding increase in testing, cases and deaths is consistent and amenable to model. The earliest SEIR modeling was performed on the data from 05 March 2020 to 23 March 2020.[13] The majority were based on SIR or its modification, which was first introduced by Kermack et al.[5] and since then is popular for modeling of infectious diseases. All SIR (or its modifications) have certain assumptions, many of which act as model limitations. The commonest ones being fixed, homogenous population, random mixing, compartmentalization, not catering to change in population dynamics and agent characteristics during the epidemics. Although the approach is flexible to cater for all the assumptions, it increases in complexity and interpretation and moreover, many a times data are not available on aforementioned assumptions.

Arithmetic, geometric, and exponential progressions are other methods of prediction. Linear regression models also include its variations such as Lasso and ridge regression. Although they are easy to understand and good for short-term predictions, their inherent properties preclude them from being accurate for long-term predictions as is evident from our review of these models.[21,36] Even techniques such as autoregression integrated moving average, used alone or in combination with wavelet transformation may be improved upon by use of repressor.[27,33] However, because they are based on time series, any deviations from the past may not be captured by these models.

Table 2 – Model, assumptions, data sources and software's used.

S No	Reference	Model used	Assumptions/estimated	Data source	Software used
1.	13	SEIR (Susceptible, exposure, infectious and recovered) (Modified for effect of quarantine)	N = 1375.98 million, Incubation period = 5.1, Infectious period = 7, R0 = 2.28, Growth rate of the epidemic in India = 1.15. Herd immunity may be achieved when 55–65% of population infected	Web site: World meters MoHFW	MATLAB/Simulink Release 2018b, MS Excel with Sim Voi
2.	14	SIR, Social distancing matrix, Bayesian error propagation analysis	All cases to be symptomatic (less severe effect) R0 = 2.108	Web site: World meters, population pyramid sites	Python
3.	15	Arithmetic Progression; Tree-based model structure	RO = 1.9 One infective node infects another infective node in 2.3 days, Recovery rate = 4 days.	Web site: World meters, WHO	Not mentioned
4.	16	Susceptible-Infectious-Quarantined-Recovered (SIQR)	RO = 1.55, Epidemic doubling time = 4.10 days; Incubation Period = 05 days; Infected to quarantine ratio = 10.45	Web site: World meters, WHO	Not mentioned
5.	17	SIR model and tanh model	No assumptions regarding R0	MOHFW, census registrar	R
6.	18	SIRD (susceptible, Infectious, recovered, death) model and Sequential Bayesian method (SBM)	R0 = 1.42–1.85, Mean serial interval = 3.9 days, Index case can infect 2.8 individuals, mean recovery time = 14 ± 5.3 days, doubling time = 4.30 days.	Web site: World meters	R software and Package ggplot2
7.	19	Exponential growth model	RO = 2.56, herd immunity as 61%, Serial Interval = 4.4 days	Web site:MoHFW, WHO, covid19india.org	Not mentioned
8.	20	Multiple and linear regression analyses	No assumptions regarding R0, Projected death rate (n) is 211 and 467 at the end of the 5th and 6th week, respectively w.e.f. 26 Mar 20. CFR = 1.650	Website: covid19india.org and WHO	Python 3.8.2 software&excel with XL-STAT statistical software
9.	21	Lasso regression	No assumptions regarding R0	Web site: MoHFW, covid19india.org	Prophet Python
10.	22	SIR model	R0 = 2.6	Web site: MoHFW	Not mentioned
11.	23	Exponential fit models and polynomials equations	NA	Web site: World meters	Python
12.	24	Geometric progression	R0 = 2.26, Rate of infection = 1.92 days. Recovery time = 14 days	Web site: World meters	Not mentioned
13.	25	SIR model	RO = 2.4–2.9. Median age of COVID-19 patients = 36 yrs. CFR = 3.8%,75.0% of the deceased were also males	Website: covid19india.org	Microsoft Office Excel 2007
14.	26	SEIR (Modified for effect of social distancing)	N = 133.92 crores 1/time incubation = 1/5 1/time infection = 1/7 R0 = 1.8 and 2.2 Studied as varying value of Rho (0–1) with 1 as no intervention and 0 as complete lockdown.	Web site: World meters MoHFW	Python
15.	27	Autoregression integrated moving average model (ARIMA), SIR and Richard's model	No assumptions regarding R0	Web site: Johns Hopkins Corona Virus Resource Center	R
16.	28	Exponential model, logistic model, SIR Model	R0 = 1.504 Initial doubling times = 4.8 days	Web site: John Hopkins University Coronavirus Data Stream	MATLAB
17.	29	SEIR & Regression model	R0 = 2.02	Web site: John Hopkins University Coronavirus Data Stream	R

18.	[30]	Exponential and polynomial regression modeling	No Assumptions regarding R0 Death rate = 3% Doubling rate = 4.1–7.1 days	Web site: MoHFW& John Hopkins University Coronavirus Data Stream	R
19.	[31]	Exponential, logistic, SIR, generalized SEIR (SEIQRDP) Model	Infection ratio = 4% DR (%) = 3.28 CFR = 2–3%	Web site: John Hopkins University Coronavirus Data Stream	MATLAB
20.	[32]	Regression based predictive model	No assumptions regarding R0	Web site: World meters covid19india.org	R
21.	[33]	Hybrid model approach (ARIMA& Wavelet transformation)	No assumptions regarding R0	Web site: World meters ourworldindata.org/coronavirus	R
22.	[34]	SEIR (modified for quarantine)	R0 = 1.5 to 4.98	Web site: WHO, DG of Civil aviation; Statistics' and reports	Not mentioned
23.	[35]	SEIR models	50% relative contribution of non-contact transmission increases R0 by 15–35%, a 150% relative contribution can double it	Web site: CDC COVID-19 report, 20, WHO report, 2019	MATLAB R2016b package
24.	[36]	Linear regression correlation, Pearson's correlation	No Assumptions regarding R0	Web site: WHO site, Historical Weather	R
25	[37]	SIR (susceptible, infected, recovered) model	No assumptions regarding R0	Website: covid19india.org	Not mentioned
26.	[38]	Model proposed by Bommer and Vollmer	India's detection rate = 3.6% below the world average of 6%. Maharashtra (1.8%)	Website: covid19india.org, ICMR	Not mentioned
27.	[39]	Logistic model	No assumptions regarding R0	Web site: World meters, Wikipedia	R
28.	[40]	logistic model	No Assumptions regarding R0	Web site: covid19india.org, Wikipedia	R
29.	[41]	SEIR Model	R0 = 1.4 to 3.9. Death rate = 1–3%	Web site: MOHFW, India	Python and R-Programming languages
30.	[42]	eSIR	R0 = 2 (no intervention) R0 = 1.5 Moderate intervention	Website: Johns Hopkins University	R

SIR, susceptible, infected (infectious), and recovered; SEIR, susceptible, exposed, infectious, recovered.

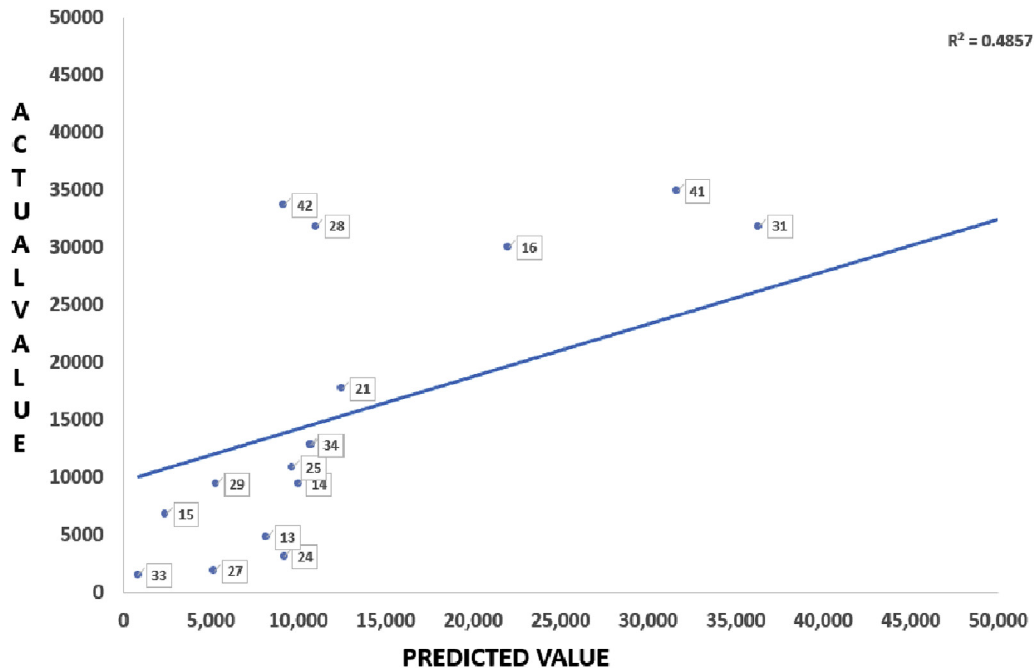

Fig. 3 – Predicted and actual value of number of total cases for COVID cases in India. #The number in box is reference number of studies.

There is a huge variation among the models in the numbers, which may be attributed to different assumptions by the models and because of mathematical models predicted for different time periods. Hence it was not possible to synthesize the pooled results. It is extremely important to understand the assumptions in the models. Our review showed that few of models did not explicitly mention their assumptions,[17,21,29] whereas some had too few[15] or too many[13] assumptions in the models. The review brings out another interesting fact about the wide varying assumptions used for modeling, for example, the value of R^0 varied from 1.4 to 4.98. Such assumptions over wide range have implications on the number of cases which the models predict. These assumptions reflect uncertainty about the disease especially in an evolving pandemic.

The study found a fair correlation for short-term predictions, thus emphasizing the need for corrections of predictive models as more and more data become available. We opine that long-term predictions may be difficult as predictive models are based on parsimonious inputs for sake of better understanding, which with assumptions may not simulate real life scenarios. However, these short-term predictions are equally important for the health planners, decision makers, and so on, for arrangement of adequate resources to tackle epidemics.

Complex or hybrid models with explicit assumptions encompassing important ones such as effect of non-pharmacological interventions, age structure, interactions, stochasticity, quarantine, isolation, socioeconomics etc, are required especially in an evolving epidemic as unique as COVD-19. Most of the models did not incorporate uncertain data, which is an important paradigm of epidemiology. However, this could be attributed to less data to use for the models to begin with and is not a comment on the approach or the methodology adopted.

Another important contribution of mathematical models is the qualitative information generated by each model, which provides a range of inputs to the planners at various levels. This review has provided narrative synthesis of 30 models and can be used by modelers, planners and researchers.

The rapid availability of models with a large number of those being non-peer reviewed as well as well as availability to the lay press and their own interpretation is fraught with the danger of models getting into disrepute. We as researchers and planners need to look beyond the straightforward answers from the models (magnitude, numbers, mortality) and instead use models to try to implement policies which may change the predictions by various scenarios for the greater public good. In addition, models should be interpreted in the context of the entire system, such as including other medical conditions, social, economic, cultural and ethical considerations. It should be taken as just one of the inputs for planning purposes.

One of the recent examples of widening the scope is to use eight stages of infection: susceptible (S), infected (I), diagnosed (D), ailing (A), recognized (R), threatened (T), healed (H), and extinct (E), collectively termed SIDARTHE.[44] Now with more data availability, the future models for India may also look at further refinements using different approaches and tools for better use of quantitative outputs of the models.

Because mathematical modeling involves equations and predictions are made by solving them, there is little scope of subjectivity. The risk of bias as seen in other epidemiological studies may not be quantified. Hence it differs from other rapid review in this aspect. Explicit assumptions and the basis of the assumptions should be included in every predictive modeling study. Owing to varied assumptions and mathematical models, it becomes difficult to synthesize the results. Another important limitation is to check for the quality of studies of the mathematical modeling. The consensus may evolve over period of time but as of now there is lack of scale for quantifying quality of study in mathematical modeling.

Conclusion and recommendations

This review has clearly shown the importance of assumptions and strong correlation between short-term projections but uncertainties for long-term predictions. The results for long-term predictions could not be synthesized as very few studies have provided the same. The short-term predictions may be revised as more and more data become available. The assumptions too will expand and firm up as the pandemic evolves because at the start of pandemic, data are sparse and making correct assumptions is difficult. Models with more realistic assumptions may be developed subsequently. There is a case for state-specific models in our country owing to the large variation in assumptions for each state.

Disclosure of competing interest

The authors have none to declare.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.mjafi.2020.06.001.

REFERENCES

1. Trevisanato SI. The biblical plague of the Philistines now has a name, tularemia. *Med Hypotheses*. 2007;69:1144–1146. https://doi.org/10.1016/j.mehy.2007.02.036.
2. Acuna-Soto R, Stahle DW, Therrell MD, Gomez Chavez S, Cleaveland MK. Drought, epidemic disease, and the fall of classic period cultures in Mesoamerica (AD 750–950). Hemorrhagic fevers as a cause of massive population loss. *Med Hypotheses*. 2005;65:405–409. https://doi.org/10.1016/j.mehy.2005.02.025.
3. WHO | Disease outbreaks by year. WHO. Available at: http://www.who.int/csr/don/archive/year/en/[accessed April 23, 2020].
4. Bernoulli D, Blower S. An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it. *Rev Med Virol*. 2004;14:275–288. https://doi.org/10.1002/rmv.443.
5. Kermack WO, McKendrick AG, Walker GT. A contribution to the mathematical theory of epidemics. In: *Proceedings of the Royal Society of London Series A, Containing Papers of a Mathematical and Physical Character*. vol. 115. 1927:700–721. https://doi.org/10.1098/rspa.1927.0118.
6. Anderson RM, May RM. *Infectious Diseases of Humans: Dynamics and Control*. Oxford, New York: Oxford University Press; 1992.
7. Siettos CI, Russo L. Mathematical modeling of infectious disease dynamics. *Virulence*. 2013;4:295–306. https://doi.org/10.4161/viru.24041.
8. Huppert A, Katriel G. Mathematical modelling and prediction in infectious disease epidemiology. *Clin Microbiol Infect*. 2013;19:999–1005. https://doi.org/10.1111/1469-0691.12308.
9. Zhu N, Zhang D, Wang W, et al. A novel coronavirus from patients with pneumonia in China, 2019. *N Engl J Med*. 2020;382:727–733. https://doi.org/10.1056/NEJMoa2001017.
10. Li Q, Guan X, Wu P, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. *N Engl J Med*. 2020. https://doi.org/10.1056/NEJMoa2001316, 0:null.
11. Coronavirus update (live): 2,697,316 cases and 188,857 deaths from COVID-19 virus pandemic - worldometer. Available at: https://www.worldometers.info/coronavirus/[accessed April 24, 2020].
12. MoHFW | Home. Available at: https://www.mohfw.gov.in/ [accessed April 1, 2020].
13. Chatterjee K, Chatterjee K, Kumar A, Shankar S. Healthcare impact of COVID-19 epidemic in India: a stochastic mathematical model. *Med J Armed Forces India*. 2020;76:147–155. https://doi.org/10.1016/j.mjafi.2020.03.022.
14. Singh R, Adhikari R. *Age-structured Impact of Social Distancing on the COVID-19 Epidemic in India*. ArXiv:200312055 [Cond-Mat, q-Bio]. 2020.
15. (11) (PDF) modeling and predictions for COVID 19 spread in India. ResearchGate. Available at: https://www.researchgate.net/publication/340362418_Modeling_and_Predictions_for_COVID_19_Spread_in_India [accessed April 24, 2020].
16. Tiwari A. Modelling and analysis of COVID-19 epidemic in India. *MedRxiv*. 2020. https://doi.org/10.1101/2020.04.12.20062794.
17. Roy A, Kar S. Nature of transmission of Covid19 in India. *MedRxiv*. 2020. https://doi.org/10.1101/2020.04.14.20065821.
18. Rajendrakumar AL, Nair ATN, Nangia C, et al. Epidemic landscape and forecasting of SARS-CoV-2 in India. *MedRxiv*. 2020. https://doi.org/10.1101/2020.04.14.20065151.
19. Rai B, Shukla A, Dwivedi LK. COVID-19 in India: predictions, reproduction number and public health preparedness. *MedRxiv*. 2020. https://doi.org/10.1101/2020.04.09.20059261.
20. Ghosal S, Sengupta S, Majumder M, Sinha B. Linear Regression Analysis to predict the number of deaths in India due to SARS-CoV-2 at 6 weeks from day 0 (100 cases - March 14th 2020). *Diabetes Metab Syndr*. 2020;14:311–315. https://doi.org/10.1016/j.dsx.2020.03.017.
21. Virk JS, Ali SA, Kaur G. Recent update on COVID-19 in India: is locking down the country enough? *MedRxiv*. 2020. https://doi.org/10.1101/2020.04.06.20053124.
22. Bhola J, Venkateswaran VR, Koul M. Corona epidemic in Indian context: predictive mathematical modelling. *MedRxiv*. 2020. https://doi.org/10.1101/2020.04.03.20047175.
23. Verma MK, Asad A, Chatterjee S. COVID-19 epidemic: power law spread and flattening of the curve. *MedRxiv*. 2020:2020. https://doi.org/10.1101/2020.04.02.20051680.
24. (11) (PDF) COVID-19: mathematical modeling and predictions. ResearchGate. Available at: https://www.researchgate.net/publication/340375647_COVID-19_Mathematical_Modeling_and_Predictions [accessed April 24, 2020].
25. Mazumder A, Arora M, Bharadiya V, et al. Study of epidemiological characteristics and in-silico analysis of the effect of interventions in the SARS-CoV-2 epidemic in India. *MedRxiv*. 2020. https://doi.org/10.1101/2020.04.05.20053884.
26. Patrikar S., Poojary D., Basannar DR., Kunte R. Projections for COVID-19 and Evaluation of Epidemic Response: Strategies for India. Personal Communication. [n.d].
27. Kumar P, Singh RK, Nanda C, et al. Forecasting COVID-19 impact in India using pandemic waves Nonlinear Growth Models. *MedRxiv*. 2020. https://doi.org/10.1101/2020.03.30.20047803.
28. Ranjan R. Predictions for COVID-19 outbreak in India using Epidemiological models. *MedRxiv*. 2020. https://doi.org/10.1101/2020.04.02.20051466.

29. Gupta R, Pandey G, Chaudhary P, Pal SK. SEIR and Regression Model based COVID-19 outbreak predictions in India. *MedRxiv*. 2020. https://doi.org/10.1101/2020.04.01.20049825.
30. Gupta R, Pal SK, Pandey G. A comprehensive analysis of COVID-19 outbreak situation in India. *MedRxiv*. 2020. https://doi.org/10.1101/2020.04.08.20058347.
31. Ranjan R. Estimating the final epidemic size for COVID-19 outbreak using improved epidemiological models. *MedRxiv*. 2020. https://doi.org/10.1101/2020.04.12.20061002.
32. Bhardwaj R. A predictive model for the evolution of COVID-19. *MedRxiv*. 2020. https://doi.org/10.1101/2020.04.13.20063271.
33. Chakraborty T, Ghosh I. Real-time forecasts and risk assessment of novel coronavirus (COVID-19) cases: a data-driven analysis. *MedRxiv*. 2020. https://doi.org/10.1101/2020.04.09.20059311.
34. Mandal S, Bhatnagar T, Arinaminpathy N, et al. Prudent public health intervention strategies to control the coronavirus disease 2019 transmission in India: a mathematical model-based approach. *Indian J Med Res*. 2020. https://doi.org/10.4103/ijmr.IJMR_504_20.
35. Roy S. COVID-19 pandemic: impact of lockdown, contact and non-contact transmissions on infection dynamics. *MedRxiv*. 2020. https://doi.org/10.1101/2020.04.04.20050328.
36. Ray S, Roy M. Susceptibility and sustainability of India against CoVid19: a multivariate approach. *MedRxiv*. 2020. https://doi.org/10.1101/2020.04.16.20066159.
37. Bhattacharya S, Islam MM, De A. Search for trends of Covid-19 infection in India, China, Denmark, Brazil, France. Germany and the USA on the basis of power law scaling. *MedRxiv*. 2020. https://doi.org/10.1101/2020.04.03.20052878.
38. Goli S, James KS. How much of SARS-CoV-2 Infections is India detecting? A model-based estimation. *MedRxiv*. 2020. https://doi.org/10.1101/2020.04.09.20059014.
39. Mondal S, Ghosh S. Fear of exponential growth in Covid19 data of India and future sketching. *MedRxiv*. 2020. https://doi.org/10.1101/2020.04.09.20058933.
40. Mondal S, Ghosh S. Possibilities of exponential or Sigmoid growth of Covid19 data in different states of India. *MedRxiv*. 2020. https://doi.org/10.1101/2020.04.10.20060442.
41. Biswas S, Mukherjee M. Risk assessment of nCOVID-19 pandemic in India: a mathematical model and simulation. *MedRxiv*. 2020. https://doi.org/10.1101/2020.04.10.20060830.
42. 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*. 2020. https://doi.org/10.1101/2020.04.15.20067256.
43. Akoglu H. User's guide to correlation coefficients. *Turk J Emerg Med*. 2018;18:91–93. https://doi.org/10.1016/j.tjem.2018.08.001.
44. Giordano G, Blanchini F, Bruno R, et al. Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. *Nat Med*. 2020:1–6. https://doi.org/10.1038/s41591-020-0883-7.