STUDY PROTOCOL

Mathematical models on COVID-19 in India: A systematic review protocol [version 1; peer review: awaiting peer review]

Sezal Panchal¹, Denny John²,³, Geetha R. Menon⁴, Narassima M.S.⁵, Tushar Shaw²

¹Institute of Tropical Medicine and International Health, Charite - Medical University Berlin, Berlin, 13353, Germany
²Faculty of Life and Allied Health Sciences, Ramaiah University of Applied Sciences, Bengaluru, Karnataka, 560054, India
³Center for Public Health Research, MANT, Kolkata, West Bengal, 700078, India
⁴ICMR-National Institute of Medical Statistics, New Delhi, Delhi, 110029, India
⁵Operations & Analytics, Great Lakes Institute of Management, Chennai, Tamil Nadu, 603102, India

Abstract

Background: More than 278 million cases and more than 5.4 million deaths due to coronavirus disease (COVID-19) were reported worldwide by the end of 2021. More than 34 million cases and more than 478,000 deaths have been reported in India. Epidemiologists, physicians and virologists are working on a number of conceptual, theoretical or mathematical modelling techniques in the battle against COVID-19.

Protocol: This systematic review aims to provide a comprehensive review of published mathematical models on COVID-19 in India and the concepts behind the development of mathematical models on COVID-19, including assumptions, modelling techniques, and data inputs. Initially, related keywords and their synonyms will be searched in the Global Literature on Coronavirus Disease database managed by World Health Organisation (WHO). The database includes searches of bibliographic databases (MEDLINE, Scopus, Web of Science, EMBASE etc.), preprints (MEDRXIV), manual searching, and the addition of other expert-referred scientific articles. This database is updated daily (Monday through Friday). Two independent reviewers will be involved in screening the titles and abstracts at the first stage and full-texts at the second stage, and they will select studies as per the inclusion and exclusion criteria. The studies will be selected for their quality, transparency, and ethical aspects, using the Overview, Design concepts, Details (ODD) protocol and International Society for Pharmacoeconomics and Outcomes Research-Society for Medical Decision Making (ISPOR-SMDM) guidelines. Data will be extracted using standardized data extraction tools and will be synthesized for analysis. Disagreements will be resolved through discussion, or with a third reviewer.
Conclusions: This systematic review will be performed to critically examine relevant literature of existing mathematical models of COVID-19 in India. The findings will help to understand the concepts behind the development of mathematical models on COVID-19 conducted in India in terms of their assumptions, modelling techniques, and data inputs.

Keywords
COVID-19, mathematical model, transparency, ethics

This article is included in the Emerging Diseases and Outbreaks gateway.

Corresponding author: Denny John (djohn1976@gmail.com)

Author roles: Panchal S: Data Curation, Formal Analysis, Investigation, Methodology, Validation, Writing – Original Draft Preparation, Writing – Review & Editing; John D: Conceptualization, Data Curation, Formal Analysis, Investigation, Methodology, Project Administration, Resources, Supervision, Validation, Writing – Review & Editing; Menon GR: Conceptualization, Investigation, Project Administration, Validation, Writing – Review & Editing; M.S. N: Conceptualization, Investigation, Validation, Writing – Review & Editing; Shaw T: Validation, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: The author(s) declared that no grants were involved in supporting this work.

Copyright: © 2022 Panchal S et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Panchal S, John D, Menon GR et al. Mathematical models on COVID-19 in India: A systematic review protocol [version 1; peer review: awaiting peer review] F1000Research 2022, 11:532 https://doi.org/10.12688/f1000research.109518.1

First published: 17 May 2022, 11:532 https://doi.org/10.12688/f1000research.109518.1
Introduction
The sudden outbreak of a new pathogen called the coronavirus disease (COVID-19) in Wuhan province has threatened the world population within a short period of its occurrence. The COVID-19 pandemic has been exhausting the health care resources not only in poor or developing countries but in developed countries as well. By December 2021 more than 278 million cases including 5.4 million deaths were reported by World Health Organisation. The paucity of health resources in terms of financing and availability, and large population in India has put forth more challenges compared to high-income countries. Initial non-pharmaceutical preventive measures, such as lockdowns, social distancing and sanitation measures were successful in managing the spread of disease but in later stages, COVID-19 caused depletion of medical supply, shortage of health care workers, hospital beds, intensive care units (ICUs), diagnostic safety kits and oxygen cylinders. Cumulatively with the existing living and health conditions like lack of nutrition, limited access to clean water and sanitation, communicable and non-communicable diseases, it has consecutively created a helpless situation for the government. Moreover, factors such as mass migration of workers, unemployment, the education system, and management of critical non-COVID patients have become a concern amidst the pandemic.

Mathematical models are potential tools in providing a better understanding of disease spread dynamics and transmission. These mathematical models can help us to understand the epidemic, the size and duration of the pandemic wave, and the extent of illness in co-morbid conditions in countries that are struggling with health resources. A prerequisite for a model is that it should provide predictions corresponding to reality. During the past year and a half, several mathematical models have been published for COVID-19 in low-and-middle income countries (LMICs). China has published a few models for quantitative prediction of infection. Other models were developed to predict the effect of non-pharmaceutical measures on epidemic dynamics. In India, a number of mathematical models on COVID-19 were proposed and published in peer-reviewed journals and as grey literature during this period of time.

This systematic review aims to provide a comprehensive review of existing mathematical models on COVID-19 in India that are published from January 2020 to January 2022. The identified studies from the systematic review would be helpful in filling this research gap. The review additionally helps to understand the concept behind the development of mathematical models on COVID-19 conducted in India in terms of their assumptions, modelling techniques, and data inputs. The review will also aim to identify, where feasible, the reliability of the various mathematical models in predicting the COVID-19 pandemic in India. These insights might help to develop a methodology and the potential use of these models in predicting epidemic outbreaks in a limited resourced setting for future pandemics. A preliminary search on the Cochrane database of systematic review, PROSPERO, MEDLINE, and Implementation reports was conducted, and no systematic reviews on the topic were identified.

Protocol
Method and design
The study design is a systematic review of mathematical models on COVID-19 in India. The method has been developed and reported in compliance with Preferred Reporting Items for Systematic Reviews and Meta-analysis Protocol (PRISMA-P). Please see Reporting guidelines for the completed checklist. The search for the final review will be documented and reported as per PRISMA-S.

Objectives
1) To perform a comprehensive review of existing mathematical models on COVID-19 and to assess the reported number of cases of infections, the peak of infections, mortalities, and spread of the epidemic in India.

2) To identify the concept behind the development of the mathematical models on COVID-19, for example, assumptions, modelling techniques and data inputs.

3) If possible, to check the reliability of mathematical models (i.e., the closeness of the predictions with the actual data) in predicting the real epidemic situation in a limited resource country.

Research questions
What are the concepts behind the development of the mathematical models on COVID-19?

What are the assumptions, modelling techniques, and data inputs and the qualities, transparency and ethical considerations of mathematical modelling on COVID-19 in India?
Eligibility criteria

Inclusion criteria

Research articles adopting infectious disease modelling, mathematical modelling, autoregressive integrated moving average (ARIMA) modelling,22,23 regression modelling,24 agent-based25 network,26 or simulation models27 on COVID-19 published in peer-reviewed journals and preprint servers focusing on the population of India will be assessed for inclusion in the review. A study will be selected if it presents a mathematical or statistical model of COVID-19 and reports the following parameters - an incubation period, basic reproduction number (R₀) infectious period, fatality, peak time, peak size, total infection number, or elimination time.

Exclusion criteria

This review will be limited to English language studies. and studies with the following criteria will be excluded from the review:

1. Articles on mathematical modelling of COVID-19 in countries other than India.
2. Articles on non-COVID outcomes.
3. Articles where the abstract or full text is not available.
4. Articles not conducting and reporting mathematical models will be excluded.
5. Articles will be excluded if they only present on evaluating intervention strategies without offering parameter estimates or trajectory projection.
6. Review papers, empirical studies, emergency response articles, microbiological studies, disease surveillance studies focused on treatment or vaccination in the host, studies that are not disease-specific, studies not involving population-wide spread, studies on computer viruses, social media modelling, internet modelling, or phone modelling.
7. Reviews and non-original papers.

Search strategy

The comprehensive search strategy has been developed in consultation with an information specialist. The developed search strings will be used to search in the database, i.e. WHO (Global literature on coronavirus disease) and will be supplemented by a manual search at Semantic Scholar for relevant English language articles published from 1 January 2020 to January 2022. Additionally, cross-referencing of included studies focussed on India from previously published systematic reviews on the topic, and forward and backward citations of included studies will be conducted to identify more studies. The global literature on coronavirus disease database includes searches of bibliographic databases (MEDLINE, Scopus, Web of Science, Europe PMC, EMBASE), pre-prints (MEDRXIV), clinical trial registry (ICTRP), manual searching, and the addition of other expert-referred scientific articles, and this database is updated daily (Monday through Friday).28

Search terms related to COVID-19, methodology, and population were identified and used concurrently as: “SARS-CoV-2,” “Coronavirus Disease 2019,” “COVID-19,” “2019-nCoV,” “coronavirus,” OR “pneumonia” AND “model,” “modelling,” “modelling,” “dynamic,” “estimation,” “prediction,” OR “transmission” AND “India,” OR “Republic of India,” OR “India,” “Indian”. A study will be selected if it presents a mathematical or statistical model of COVID-19 and reports the following parameters - an incubation period, basic reproduction number (R₀) infectious period, fatality, peak time, peak size, total infection number, or elimination time. Reference lists will be manually searched and onward citation searching will be conducted using WHO for all included studies.

The search strategy will be internally reviewed using CADTH PRESS 2015 guidelines.29

Study selection

Studies identified through search strategy across identified databases and grey literature will be imported to Covidence (version 2.0) software (Rayyan is an example of a free alternative that can be used to replicate the study) and will be first
screened at the title and abstract level by two authors (SP and DJ). After both the reviewers screen the articles, the following criteria will be used for categorizing: (1) both authors agree on inclusion; (2) one author recommends inclusion; (3) both authors are unsure; (4) one author recommends exclusion and the other is unsure or (5) both authors agree on exclusion. Full-text articles for abstracts classified as 1 or 2 will be retrieved. Those classified as 3 or 4 will be discussed with the third reviewer for inclusion in the full-text review. Records classified as 5 will be excluded.

The full texts will be reviewed by the reviewers (SP and DJ) as per the inclusion/exclusion criteria. Reasons for exclusion at the full-text stage will be reviewed and recorded. The screening decisions will be reported using PRISMA-2020 guidelines.30

Assessment of methodological quality and risk of bias
For quality assessment for selected studies, a modified critical appraisal checklist prepared by Fone et al.31 will be used. The checklist is provided in Extended data.41 For epidemiological models32 a modified risk of bias tool will be used to assess the risk of bias of individual studies. This bias tool is also provided in Extended data.41 Assessment for risk of bias for components such as model setting and population, appropriateness of modelling methodology and structure, fitting methodology, and reporting the conflicts of interest, which are essential for assessing the reproducibility of the model, alignment of the model and research question, will be performed. The Professional Society for Health Economics and Outcomes Research- Society for Medical Decision Making (ISPOR-SMDM)33 for good practices guidelines will be referred to for the task.

Identification of ethical risks associated with the development of mathematical models and their implementation is important. An ethical framework will be considered for the accountability of scientists for the communication and translation of mathematical models to policymakers for a better understanding of the strengths and weaknesses of scientific evidence. Moreover, ethical framework for mathematical models helps to understand the ethical and socio-economic impact of biased and unpredictable events. A biomedical ethics-based evaluation will be conducted using parameters mentioned in Appendix 4 in Extended data.41

Data extraction
Data will be extracted from included articles into a piloted, standardized Excel database by two independent reviewers. Reference lists will be manually searched and further online citations of all included studies will be searched using the Web of Science. The following data will be extracted from each article: the date of publication, location/setting and study population (urban or rural) and duration, age, gender, the density of population, number of people in every household and locality, study duration and sample size. The source of data for population, risk of exposure at workplace and work-from-home capabilities if industry-specific constraints are to be included, as well as the source of data for population, source of data for epidemiology and travel. Detailed demography about the population is also being extracted.

We will also extract data about on following parameters: percentage of pre-symptomatic transmissions, pre-symptomatic transmission period, percentage of asymptomatic patients, serial interval, incubation period, the onset of symptoms/illness onset to diagnosis, onset of symptoms/illness onset to hospital admission, hospital stay length, time from hospital admission to death/discharge, the onset of symptoms/illness onset to death, the onset of symptoms/illness onset to discharge/recovery, the proportion of patients who require ventilator support, duration of ventilator support, percentage of deaths, percentage of discharged, the percentage in hospital, percentage of patients requiring oxygen support percentage of ICU admissions, the onset of symptoms/illness onset to ICU admission, ICU stay length, percentage of deaths from ICU, percentage of discharged from ICU, the percentage in hospital from ICU, and percentage transferred from ICU to general hospital wards. Other outcomes, namely magnitude of infection, confirmed cases, peak time, mortality due to infection, further consequences of disease, validation and performance of each model will be assessed. The secondary outcomes and implementation of the models in preparedness for epidemic will also be studied.

In the case where we encounter a model that has been considering vaccination and acquired immunity from the previous infection, we will consider parameters such as the proportion of people who are immune to infections for the following reasons: (a) already infected and recovered, (b) vaccinated (single double/double dose).

The assumption involved in the development of models and the outcome(s) is to be predicted. Other extracted data will include sample size, mean, standard deviation (SD), confidence interval, median, interquartile range (IQR) and fitted distribution used in estimation, missing data, model fitting and calibration approaches. The method or strategies used for checking model performance and evaluation will also be studied. If any uncertainty of missing data is found, the corresponding author will be contacted for additional information or missing data.
We will record how the sensitivity analysis was performed, as well as the data bias considerations and finally the results and interpretation and discussion for the model.

**Data synthesis**

A narrative summary of included literature will be produced with the qualitative synthesis of extracted data. The selected studies will be evaluated for case data sources (epidemiological, population and travel), modelling approaches, compartments used, population mixing assumptions, model fitting and calibration approaches, sensitivity analysis used and data bias considerations. We will use the Bio-surveillance Analytics Resource Directory (BARD)\textsuperscript{34} framework to systematically characterize the models. Additionally, the final included studies will be subjected to an ethical framework for mathematical models for policymaking.\textsuperscript{35} The certainty of the evidence will be assessed for four primary outcomes: incidence, onward transmission, mortality, and resource use. Covidence systematic review software will be used for conducting the systematic review. The systematic review will be documented as a publication and policy brief for decision-makers. We will classify the models based on their theoretical types, epidemiological and population data types, and validation data types. A summary of the evidence table from all the included articles will be prepared.

Since mathematical models of COVID-19 use different methods there would be variations in data inputs and assumptions and hence we do not envisage conducting estimations of summary effect measures. However, if we find two or more studies that can be pooled we will use pooled estimates using random-effects meta-analysis of a number of infections, number of confirmed cases, and mortality. In the absence of meta-analysis, we aim to use synthesis without meta-analysis (SWiM) approaches for grouping of studies into intervention analysis (social distancing, testing, travel ban, personal hygiene & sanitation, and therapeutic interventions/treatments, epidemiological data classification, population data classification, travel data classification, and validation data classification). We will also focus on the modelling methods used, estimates of epidemiological impact, and reporting standards. We will also identify the limitations and gaps in each of the models as per SWiM guidelines.

Since COVID-19 does not seem to affect children and teens in the same way as adults, any obtained data on paediatric patients will be analysed separately.\textsuperscript{36,37}

**Statistical analysis**

If possible, for various parameters identified from the included studies, we will conduct meta-analysis using the statistical software \texttt{R} version 3.6.2 (\texttt{meta, metaphor and dmter} packages). For parameters reporting mean and standard deviation (SD), or median and interquartile range (IQR), a meta-analysis of single means will be conducted, and where mean and SD are not reported, they will be estimated from median and IQR.\textsuperscript{36,38} For those parameters presented as percentages, a meta-analysis of proportions will be used. To take into account the variability between and within studies, the random-effects model will be fitted with the Restricted Maximum Likelihood Method (REML). In order to meet the normality assumption underlying the meta-analysis, the natural logarithm transformation will be applied. The null hypothesis of no variance among studies will be tested using the Q-statistic, and the degree of heterogeneity will be quantified using the $I^2$ index. Outliers and influencers diagnoses will also be performed.

Meta-analysis results will be presented as pooled mean or percentage and its associated 95% confidence interval (CI) provided by the meta-analysis for parameters for two or more studies. For each parameter, forest plots will be developed and presented to visualize all the included studies.

The publication bias will be assessed through generation of a funnel plot if at least 10 studies are included in meta-analysis. The symmetry of funnel plot will be tested by Egger test; Meta-analysis: publication bias.\textsuperscript{39}

We will use the GRADE approach to rate the quality of evidence for the primary outcomes, where effectiveness of interventions has been estimated using mathematical modeling.\textsuperscript{40}

**Ethics and dissemination**

Ethical approval is not required for this study. The completed systematic review will be submitted for publication in a peer-reviewed journal.

**Study status**

A protocol of the systematic review has been registered on PROSPERO (registration number: CRD42022299112). Relevant searches have been completed in academic and non-academic databases, and study screening is currently ongoing.
Conclusions

This systematic review will be performed to identify, and critically review published mathematical models on COVID-19 in India. Understanding of the concept behind the development of COVID-19 mathematical models in India in terms of their assumptions, modelling techniques and data inputs could help the policymaker, scientist and physicians to promote best practices in mathematical modelling. If possible the review will aim to rule out the reliability of mathematical models to predict the real epidemic situation.

Data availability

Underlying data

No underlying data are associated with this article.

Extended data


This project contains the following extended data:

- Annexure-1 Critical Appraisal Sheet.docx.
- Appendix-2 Risk of Bias.docx (tool for assessment of epidemiological modelling studies).
- Data extraction sheet.xlsx (data extraction tool for epidemiological burden studies).
- Appendix 4. docx (list of parameters for biomedical ethics-based evaluation).
- PRISMA-S.docx (template for PRISMA-S checklist that will be completed after the final review).
- Final Search strategy_Mathematical models on Covid-19 in India.docx.

Reporting guidelines


Data are available under the terms of the Creative Commons Attribution 4.0 International license (CC-BY 4.0).

Acknowledgements

The authors wish to thank Ms Vasumathi Sriganesh, Founder & CEO, QMed Foundation, Mumbai, for providing inputs for the development of the search strategy.

References

8. COVID-19: Legal Impact in Mexico; measures issued by various authorities | White & Case LLP: [cited 2021 Sep 9]. Reference Source


The benefits of publishing with F1000Research:

• Your article is published within days, with no editorial bias
• You can publish traditional articles, null/negative results, case reports, data notes and more
• The peer review process is transparent and collaborative
• Your article is indexed in PubMed after passing peer review
• Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com