 BMJ Open

Are COVID-19 models blind to the social determinants of health? A systematic review protocol

Ava John-Baptiste,1,2,3,4 Marc S Moulin,1 Shehzad Ali1,2,3,5

ABSTRACT

Introduction Infectious disease models are important tools to inform public health policy decisions. These models are primarily based on an average population approach and often ignore the role of social determinants in predicting the course of a pandemic and the impact of policy interventions. Ignoring social determinants in models may cause or exacerbate inequalities. This limitation has not been previously explored in the context of the current pandemic, where COVID-19 has been found to disproportionately affect marginalised racial, ethnic and socioeconomic groups. Therefore, our primary goal is to identify the extent to which COVID-19 models incorporate the social determinants of health in predicting outcomes of the pandemic.

Methods and analysis We will search MEDLINE, EMBASE, Cochrane Library and Web of Science databases from December 2019 to August 2020. We will assess all infectious disease modelling studies for inclusion of social factors that meet the following criteria: (a) focused on human spread of SARS-CoV-2; (b) modelling studies; (c) interventional or non-interventional studies; and (d) focused on one of the following outcomes: COVID-19-related outcomes (eg, cases, deaths), non-COVID-19-related outcomes (ie, impacts of the pandemic or control policies on other health conditions or health services), or impact of the pandemic or control policies on economic outcomes. Data will only be extracted from models incorporating social factors. We will report the percentage of models that considered social factors, indicate which social factors were considered, and describe how social factors were incorporated into the conceptualisation and implementation of the infectious disease models. The extracted data will also be used to create a narrative synthesis of the results.

Ethics and dissemination Ethics approval is not required as only secondary data will be collected. The results of this systematic review will be disseminated through peer-reviewed publication and conference proceedings.

PROSPERO registration number CRD42020207706.

INTRODUCTION

In the early months of the pandemic, COVID-19 was described by New York Governor Andrew Cuomo as the ‘great equaliser’ as the new coronavirus swept through society without bias. A shared lack of immunity meant that regardless of prestige, wealth, fame or age, all were thought to be equally at risk. However, epidemiological researchers quickly demonstrated that this was, and remains to be, far from the truth. The pandemic has put existing health inequalities under the microscope and highlighted that COVID-19 has disproportionately affected marginalised racial, ethnic, and socioeconomic groups through its unequal health burden and its disparity of economic losses.

A recent investigation found that among 158 counties in large US metropolitan areas, those with predominantly non-white populations had infection rates approximately eight times higher, and a mortality rate more than nine times higher than counties with a mostly white population. Similarly, people living in the poorest neighbourhoods of England and Wales have been found to be twice as likely to die from COVID-19 as those in more affluent areas. Examples of similar trends can be found throughout history. Perhaps most notably in terms of size and scale, data collected from the 1918 influenza pandemic demonstrated that non-white minorities had higher all-cause mortality and influenza-related mortality rates when compared with their white counterparts. These disparities are a symptom of deeper societal and health
system inequities, including disproportionate exposure through high-risk jobs, prevalence of comorbidities, and inequitable access to testing and treatment.

Over the past century, infectious disease models have played a fundamental role in helping public health policymakers implement infection control interventions, as well as shape the theories used in disease mitigation. Infectious disease models have been described as, ‘a scientific approach to formulating an explanation for an observed phenomenon and then testing this formulation to project the outcome of various experiments under pertinent conditions’ (p1). Models provide a framework for experts to study a number of important factors related to the SARS-CoV-2 pandemic, including the following:

1. Viral transmission dynamics—models use observed data such as the number of confirmed cases, to estimate difficult-to-observe transmission dynamics, such as the incubation period and percentage of asymptomatic spread in the population.
2. Anticipate the future course of an outbreak—models forecast future incidence, prevalence, morbidity and mortality, based on input parameters.
3. Guide public health planning and infectious disease control—models help researchers simulate real-world possibilities and thus appraise the potential impact of disease-control strategies such as mask wearing, school closures, testing and contact-tracing on the population.
4. Models can estimate a range of outcomes, those directly related to COVID-19 or secondary effects of the disease and control measures, on population health, health service utilisation, and economic outcomes such as cost and quality of life.

Infectious disease models can only be fit for purpose if the models are appropriately conceptualised and model parameters are based on the best available evidence. Model conceptualisation requires decisions about the model structure, including which health states to include, division of the population into subgroups and selection of characteristics to define the subgroups. Thus, models can incorporate the social determinants of health, such as income, education, employment and ethnicity. However, structuring models to reflect social determinants increases complexity and sometimes, due to limited availability of data, raises questions about structural uncertainty. Complexity and uncertainty may lead modellers to simplify and favour average population models foregoing appraisal of different types of individuals or important subgroups. Current evidence highlights the importance of social factors in predicting the course of the pandemic and in analyses of policy options, although it is unclear to what extent current infectious disease models are based on an average population approach and ignore social factors. To the best of our knowledge, this limitation has yet to be explored in the context of the COVID-19 pandemic. Therefore, the present systematic review will aim to answer the following research question: ‘To what extent do COVID-19 models incorporate the social determinants of health?’ Our objective is to systematically review human COVID-19 modelling studies (predictive, forecasting) to identify what proportion of models incorporate social factors and for included modelling studies, to identify how social factors were incorporated into the conceptualisation of the infectious disease models. The work will provide an important reflection on COVID-19 modelling practices and can lead to development of a blueprint for other modellers to incorporate social factors in the future. We hope that this review will lead to infectious disease models that explicitly acknowledge the role of social determinants in making better predictions as well as mitigating inequalities, and therefore result in better infectious disease outcomes.

**METHODS**

**Protocol and registration**

This systematic review will be conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-analyses guidelines for the identification, screening, eligibility and inclusion of retrieved research studies. The current protocol is registered in the International Prospective Register of Systematic Reviews. The anticipated start date for the proposed systematic review is November 2020. The full results will be submitted for peer-reviewed publication immediately following completion of quality assessment and the drafting of a manuscript in August 2021.

**Search strategy and eligibility criteria**

MEDLINE, EMBASE, Cochrane Library and Web of Science databases will be searched. The present systematic review is designed to have broad eligibility criteria to identify all COVID-19 modelling studies published from December 2019 to 14 August 2020. Studies will be initially eligible if they: (1) have a human SARS-CoV-2 population; (2) are a modelling study (eg, mechanistic, mathematical, network, simulation, auto-regressive, Markov or other model); and (3) investigate one of the following outcomes: COVID-19 disease-related outcomes (eg, cases, hospitalisations, deaths, recoveries), non-COVID-19 disease-related outcomes (eg, impact on other health conditions such as mental health), impacts on health services (eg, bed occupancy, ventilators, surgical delays), impact of policies (eg, lockdowns, social distancing) or interventions (vaccines, treatments, etc) on COVID-19 or societal outcomes. Eligible studies will then be assessed for components of social determinants. This will allow the research team to calculate the percentage of COVID-19 modelling studies that incorporated social determinants into their predicted outcomes of the pandemic (ie, studies that included social factors (numerator) and all eligible studies (denominator)). Data will only be extracted from models that included social determinants. The following studies will be excluded: (1) within-host biological studies, (2) non-human SARS-CoV-2 studies, (3) phylogenetic/genetic studies, (4) environmental/meteorological studies without health impact analysis on
of the articles will be screened independently by two review members to identify studies that potentially meet the inclusion criteria outlined above. Conflicts will be resolved by a third independent reviewer. The full texts of these potentially eligible studies will then be retrieved and independently assessed for eligibility by two review team members. Any disagreements between them over the eligibility of particular studies will be resolved by a third reviewer. Data will then be extracted from the studies that incorporated social factors. A flow chart for this systematic review can be found in figure 1.

Data extraction and quality assessment
Data will be extracted into Covidence by one reviewer and confirmed by a second reviewer. Disagreements will be resolved with a discussion between the two reviewers and if necessary, the final decision will be made by a third reviewer. If data related to the outcomes are missing, this will be reported. Authors will not be contacted to request missing data. The following data will be extracted from included studies: author, publication year, geography (country, province/state, city/town, region, low/medium/high-income country), setting (eg, general population, long-term care home, hospital, workplace, etc), population size, goals (predicting the pandemic, impact on other health conditions, impact on health services, impact of policy or interventions, economic analyses), type of model (mathematical, mechanistic, etc), social and economic factors, and approach to incorporating social and economic factors. An example extraction table containing the characteristics of included studies in the systematic review can be found in table 1.

There is currently no standardised tool for assessing the quality of infectious disease modelling studies. We plan to adapt the principles of best practices for infectious disease modelling, laid out by Pitman et al into a tool for quality assessment of infectious disease models.24

Evidence synthesis
The extracted data will be used to create a narrative synthesis of the results. Based on our broad inclusion criteria, it is likely that the study designs and data will be too heterogeneous for meta-analyses to be appropriate. We expect to report on the number (per cent) of identified infectious disease models that incorporated social factors. We will report on the social factors considered, and indicate the number (per cent) of included models that considered each social factor. The narrative synthesis will describe how social factors were incorporated into the conceptualisation of the infectious disease models, estimate policy impact across social groups (if relevant) and describe which model types (eg, compartmental, network models) were used when social factors were considered. We will compare models in terms of the approaches used to incorporate social factors; identify trends and see if an increasing number of published models incorporated
social factors over time; and based on our knowledge of published epidemiological estimates that quantify the association between social factors and COVID-19, compare incorporation of social factors in COVID-19 models to this evidence base.

**DISCUSSION**

To the best of our knowledge, this will be the first systematic review that aims to identify the frequency and approach used in COVID-19 models to incorporate the

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Month/year</th>
<th>Geography</th>
<th>Setting</th>
<th>N</th>
<th>Goal(s)</th>
<th>Model type</th>
<th>Social factors</th>
<th>Approach to social factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Smith et al</td>
<td>March 2020</td>
<td>Milan, Italy</td>
<td>Hospital</td>
<td>400</td>
<td>Predicting the pandemic</td>
<td>Compartmental</td>
<td>Income</td>
<td>TBD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High income</td>
<td></td>
<td></td>
<td>Case fatality</td>
<td></td>
<td>Employment</td>
<td></td>
</tr>
<tr>
<td>Jones et al</td>
<td>April 2020</td>
<td>Wuhan, China</td>
<td>General</td>
<td>2000</td>
<td>Impact of policy or intervention</td>
<td></td>
<td>Network</td>
<td>TBD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High income</td>
<td></td>
<td></td>
<td>Lockdown</td>
<td></td>
<td>Education</td>
<td></td>
</tr>
<tr>
<td>Harrison et al</td>
<td>May 2020</td>
<td>Seattle, USA</td>
<td>Long-term care</td>
<td>300</td>
<td>Impact on other health conditions</td>
<td></td>
<td>Mechanistic</td>
<td>TBD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>High income</td>
<td></td>
<td></td>
<td>Mental health</td>
<td></td>
<td>Employment</td>
<td></td>
</tr>
</tbody>
</table>

TBD, To be determined.
motivate enhancements in modelling best practices, and ultimately improve health for socially and economically vulnerable groups related to COVID-19 and other health conditions.

**Ethics and dissemination**

Ethical approval is not required as only secondary data will be collected. The results of this systematic review will be disseminated through peer-reviewed publication and conference proceedings.

**Contributors** AJ-B and SA secured the funding, designed the review, drafted the paper and approved the final version to be published. MM wrote the paper and approved the final version to be published.

**Funding** This project is funded by the Gordon and Betty Moore Foundation through grant GBMF6934 to Johns Hopkins University to support the work of the Society for Medical Decision Making COVID-19 Decision Modeling Initiative.

**Competing interests** None declared.

**Patient and public involvement** Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

**Patient consent for publication** Not required.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

**ORCID iD**
Marc S Moulin http://orcid.org/0000-0001-6654-4914

**REFERENCES**

8. Star L, Moghadas SM. The role of mathematical modelling in public health planning and decision making. Purple Paper, National Collaborative Center for Infectious Diseases 2010.