



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Why COVID-19 modelling of progression and prevention fails to translate to the real-world

Carl J. Heneghan^{*}, Tom Jefferson

The University of Oxford, UK

ARTICLE INFO

Keywords:

COVID-19
SARS-CoV-2
Reproduction number
Models

ABSTRACT

Mathematical models were used widely to inform policy during the COVID pandemic. However, there is a poor understanding of their limitations and how they influence decision-making. We used systematic review search methods to find early modelling studies that determined the reproduction number and analysed its use and application to interventions and policy in the UK. Up to March 2020, we found 42 reproduction number estimates (39 based on Chinese data: R_0 range 2.1–6.47). Several biases affect the quality of modelling studies that are infrequently discussed, and many factors contribute to significant differences in the results of individual studies that go beyond chance. The sources of effect estimates incorporated into mathematical models are unclear. There is often a lack of a relationship between transmission estimates and the timing of imposed restrictions, which is further affected by the lag in reporting. Modelling studies lack basic evidence-based methods that aid their quality assessment, reporting and critical appraisal. If used judiciously, models may be helpful, especially if they openly present the uncertainties and use sensitivity analyses extensively, which need to consider and explicitly discuss the limitations of the evidence. However, until the methodological and ethical issues are resolved, predictive models should be used cautiously.

1. Introduction

In the classic decision-making process for dealing with infectious diseases, three key variables are the burden of illness, the effects of interventions to minimise or prevent the burden and the likely harms and benefits accrued in the process of implementing them.

Models are used in simulating the interaction of all three variables, but here we will focus on the prediction of the likely progression of the pandemic (encapsulated in the concept of the reproduction number) and the assessment of preventive healthcare measures (Verma et al., 1981).

In the COVID pandemic, models have been widely used to predict how populations behave, respond to different interventions, and inform policy. However, the limitations of models for informing healthcare policies are often overlooked. An independent review of the UK response to the 2009 influenza pandemic, led by Dame Deirdre Hine, recommended the ‘Government Office for Science, working with lead government departments, should enable key ministers and senior officials to understand the strengths and limitations of likely available scientific advice as part of their general induction.’ Hine also reported that while ministers were keen to understand the likely outcomes of the H1N1 pandemic, ‘this led to unrealistic expectations of modelling, which could not be reliable in the early phases when there was insufficient data.’ (Gov.UK, 2009).

^{*} Corresponding author.

E-mail address: carl.heneghan@phc.ox.ac.uk (C.J. Heneghan).

Despite the Hine recommendations, we consider there is a poor understanding of the limitations of models and little discussion or thought given to how they influence decision-making. We analyse the limitations of modelling the progression and prevention in the COVID-19 pandemic, including the reproduction number's use and its application to interventions and policy in the UK.

1.1. The reproduction number

The basic reproduction number (R_0) is the average number of people infected by one person in an entirely susceptible population. The R is the reproduction number at any time during an epidemic and changes as the epidemic evolves (The Royal Society, 2020). The epidemic growth rate (r) refers to the number of new infections at an increasing or decreasing exponential rate. These terms are often used interchangeably to inform policy. In March 2020, the UK Government's Chief Scientific Advisor, Sir Patrick Vallance, said Britain's lockdown was having a "very big effect" on the R_0 , bringing it down to "below one" (Gov.UK press briefing, 2022). By May, UK Prime Minister Boris Johnson set out a Covid alert level system 'primarily determined by the R-value and the number of coronavirus cases,' which determined the level of social distancing measures (Gov.UK, 2020). The mainstream media widely reported the R number; for example, the BBC reported the R number as 'one of the big three,' which includes severity and the overall number of cases (BBC News, 2021).

1.2. Early estimates of the reproduction number

The COVID-19 pandemic led to a substantial increase in publications relating to the reproductive number (See Fig. 1). A search of PubMed using the terms "reproduction number OR "reproductive number" shows that over half (51%) of the publications have been indexed (2539/4959) since the COVID-19 pandemic began. Nearly seven times more publications were indexed in 2020 ($n = 1036$) than the 151 for the year the swine flu pandemic began (2009). In addition, a growing body of preprint publications and grey literature adds to the volume of studies reflecting large-scale reporting on the current pandemic.

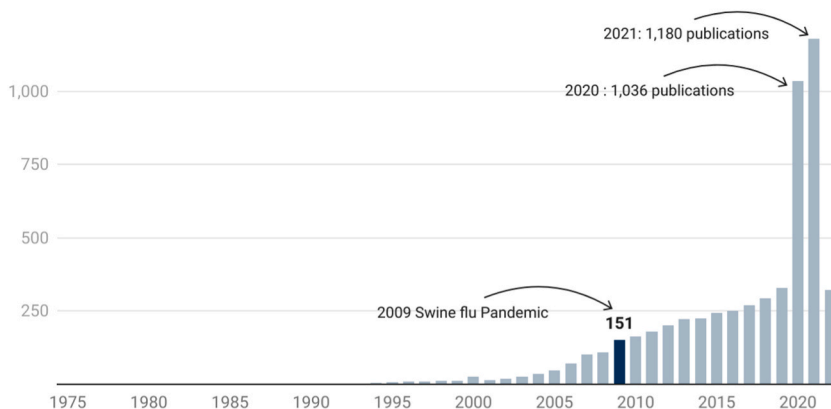
We searched PubMed, LitCovid and the WHO Covid-19 database for published studies that report the Reproduction number up to March 2020. We found 29 studies that included 42 estimates of the R_0 : 39 were based on Chinese data (one included overseas data), two on the Diamond Princess Cruise Ship and one from South Korean data (see Appendix 1, which provides the individual study references). Thirty-three estimates from China reported an average R_0 ranging from 2.1 to 6.47 and 95% confidence or credible intervals that varied substantially (see Fig. 2).

These early estimates are prone to biases because of the small amount of case data used over short time periods and the systematic differences that arise in identifying individuals who may differ from the population of interest (a form of ascertainment bias (Catalogue of Biases Collaboration Spencer and Brassey, 2017)). In addition, selection bias can occur due to inadequacies in testing and variation in those presenting for testing or because of poorly understood modes of agent transmission. Results are also distorted by systematic differences in the care provided or exposure to factors other than the infectious agent under investigation - a form of performance bias (Catalogue of Bias Collaboration Banerjee et al., 2019). Furthermore, most early estimates were based on limited data - only 80,000 cases were reported up to March 2020 in China, a tiny fraction of the COVID-19 cases, and it is unclear if these estimates are based on overlapping case data. Finally, the absence of pre-existing protocols setting out predefined methods will lead to reporting biases.

Further complexities are added as the R_0 varies based on the chosen method (Delamater et al., 2019a). For example, Song et al. modelled the early dynamics of the pandemic in China from 15 to 31 January 2020 using the maximum likelihood method, the

PubMed Timeline:

PubMed searches using the terms "reproduction number" or "reproductive number" reveal 4,959 publications since 1979 with 2,529 (51%) published in the COVID-19 outbreak.

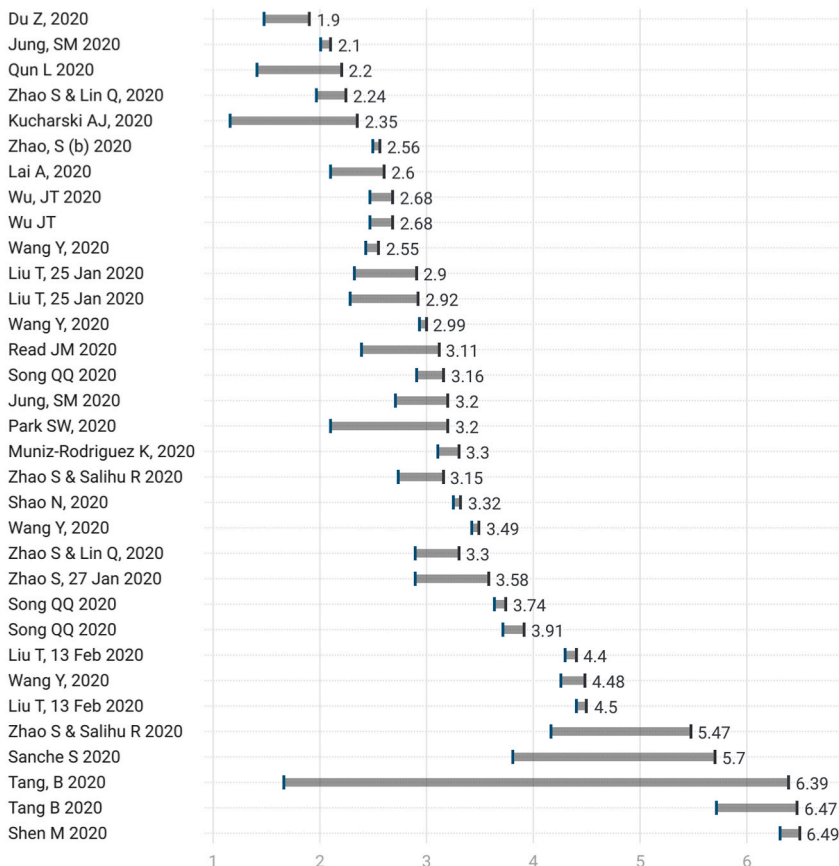


Created with Datawrapper

Fig. 1. Publications indexed in PubMed relating to the reproductive number over time.

Published estimates of the Reproduction Number (R0) from China for SARs-CoV-2 to March 2020

We searched Lit Covid, WHO COVID-19 and Google Scholar to March 2020. We report the published estimates for the R0 for COVID-19 (average 1.9 to 6.49)



The width of the bars represents credible intervals or 95% confidence intervals.
Created with Datawrapper

Fig. 2. Published estimates of the Reproduction Number (R0) from China to March 2020.

exponential growth method and the SEIR model. The estimates were 3.16 (95%CI: 2.90–3.43), 3.74 (95%CI: 3.63–3.87), and 3.91 (95%CI: 3.71–4.11) respectively (see Appendix 1).

The commonly used approach is the SIR compartmental model in which individuals are separated into mutually exclusive groups based on the population susceptibility (S) to infection, the rate at which infections (I) occur, and the rate of removal of infections (R) (Tolles and Luong, 2020). Modifications to the SIR model can include a period between being infected and becoming infectious - the exposed (E) period (SEIR model) and the loss of immunity in the population, plus changes in demography such as differences in births and deaths that give rise to the SEIRS model. Further modifications include the addition of certain factor such as quarantine (Q) restrictions. Early epidemic models commonly use exponential growth models without containment measures; however, this is less appropriate in the presence of measures (Tovissodé et al., 2020). Sources of further variation include the choice of model, the simulation method, the method of parameter estimation and the assumptions made (Skríp and Townsend, 2019).

Further issues arise with defining a case of COVID-19 that are often not accounted for in modelled estimates. We used evidence from Freedom of Information requests to show that UK bodies have a statutory duty to report positive cases based on PCR results; however, they do not have to advise which tests they are using or provide quantitative estimates of the PCR result, such as the cycle threshold (Ct) (Jefferson, 2022) Tests are, therefore, not standardised across the UK and few report thresholds for detection and positivity which is essential in distinguishing those who have active infection from those who have not. When they report thresholds for a positive result, the Ct can vary from 30 to 45 (meaning that a proportion of the donors probably and certainly are not infectious). Lack of awareness of the Ct values will overestimate the number of infectious cases, distorting the model estimates.

Fig. 2 shows this to be the case as the lack of overlap between the early estimates of R0, indicate the presence of significant statistical heterogeneity (I² = 99%) (Alimohamadi et al., 2020). As a general rule, an I² of 75%–100% indicates the presence of considerable

variation in effect estimates due to heterogeneity rather than chance (Cochrane Handbook for Systematic Reviews, 2022).

1.3. Problems generalising reproduction numbers

Internal validity is the degree to which a study is free from bias. In contrast, external validity is the extent to which you can generalize the findings of one study to other settings, situations, or populations. Several factors can affect the spread, and early estimates of transmission should be seen as specific to that setting and population at that moment in time. The high heterogeneity in early R_0 estimates questions its application outside the population from which the results were derived.

Limitations for using R_0 outside of the population and setting it was derived from have previously been noted (Ridenhour et al., 2014). A review of methods on the H1N1 Swine flu pandemic reported there is little evidence that an R_0 from one area can be applied to another, noting that studies from the same region provide highly variable results that are even more heterogeneous when early data are relied upon. Across eleven studies R_0 for H1N1 ranged from 1.03 to more than 2.9, corresponding to population attack rates between 6% and 93%.

A systematic review of the basic R_0 of measles found 18 studies providing 58 R_0 estimates that further illustrate the shortcomings: the R_0 varied from 1.43 to 770, substantially more often than the often-cited range of 12–18 (Guerra et al., 2017). The included studies used diverse methods to calculate R_0 , and estimates varied considerably by region and covariates. There have been attempts to estimate country-level reproduction numbers but they are limited by the accuracy of the contact data, the presence of missing data on covariates and projections in the early phase of the pandemic that do not account for the variations in non-pharmaceutical interventions (NPIs) and behaviour changes (Hilton and Keeling, 2020). Because R_0 is determined with heterogeneous modelling methods and the outputs are affected by numerous factors, it can easily be misinterpreted (Delamater et al., 2019b). When applying modelled estimates from one population or setting to another, care should be used.

1.4. Modelled estimates and the effectiveness of interventions

Imperial College's COVID-19 Response Team published on the 16 March 2020 the impact of NPIs to reduce COVID-19 mortality and healthcare demand (Ferguson et al., 2020). The authors modified an individual-based simulation model previously developed for pandemic influenza. The model outputs predicted 510,000 deaths in the UK in an 'unmitigated epidemic' scenario. Cited nearly 4000 times the report had a significant impact on UK and US policy.

On the 27 March 2020, UK policy redirected from a mitigation strategy to one of suppression, with restrictive measures coming into legal force. The same Imperial model state the unmitigated epidemic, would "predict approximately" 2.2 million deaths in the US. Nature report "it was shared with the White House and new guidance on social distancing quickly followed" (Adams, D, 2020)

Suppression (see box 1) aims for a decline in case numbers by reducing the R below one, whereas mitigation seeks to slow spread by reducing R, but not necessarily below one.

Data from England from May 2020 to April 2022 report the upper estimate of the R number never rose above 1.6 and the lower estimate above 1.2 (See Fig. 3). The effective reproduction number (R_e) is the number of people in a population who can be infected by an individual at any specific time. It varies as the population susceptibility changes either through immunity following infection or vaccination or as the demographics of the population change through births and deaths (Aronson JKA et al., 2020).

The Imperial College report stated in the mitigation scenario that 'even if all patients were able to be treated, we predict there would still be in the order of 250,000 deaths in GB, and 1.1–1.2 million in the US.' (Ferguson et al., 2020) Imperial's model predicted with suppression that deaths would come somewhere between 8700 and 39,000, assuming an R_0 of 2.4 (confidence interval estimate of 2.0–2.6). This R_0 was based on two studies cited in our review dataset: Li et al. 2020 (Li et al., 2020) and Riou J 2020 (Riou and Althaus, 2020). It is unclear why 2.4 was chosen as it is higher than these reports. The lower bound of the Imperial estimate does not reflect their estimates: Li Qun reports an estimate of 2.2 (95% CI, 1.4 to 3.9), and Riou J reports an R_0 around 2.2 (median value, with 90% high-density interval: 1.4–3.8). The report concluded that suppression was the only viable strategy. Five days after the UK lockdown began, advisors reported the UK would 'do well to keep deaths under 20,000.' (BBC News, 2022).

Death is widely used to assess the severity of pandemics. However, the accuracy of the cause of death assignment is not apparent in modelled outputs. In an analysis of 800 Freedom of Information requests, we found 'there was no consistency in the definition of cause of death or contributory cause of death across UK national bodies and in different bodies within the same nation.' (Jefferson, 2022)

Box 1

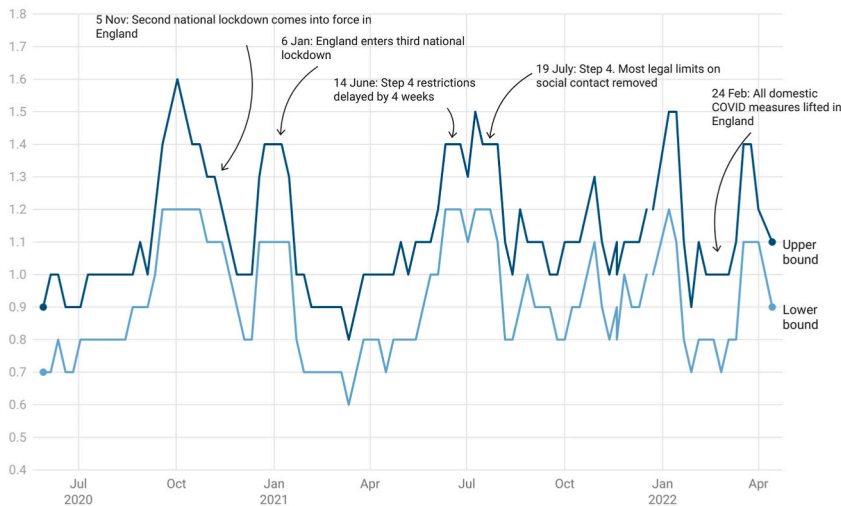
(Ferguson et al., 2020)

Suppression: 'to reduce the reproduction number (the average number of secondary cases each case generates), R, to below 1 and hence to reduce case numbers to low levels or (as for SARS or Ebola) eliminate human-to-human transmission.

Mitigation: 'to use NPIs (and vaccines or drugs, if available) not to interrupt transmission completely, but to reduce the health impact of an epidemic, akin to the strategy adopted by some US cities in 1918, and by the world more generally in the 1957, 1968 and 2009 influenza pandemics.'

The reproduction number (R) in England; May 2020 to April 2022

Estimates represent the transmission of COVID-19 from 2 to 3 weeks previously due to time delays between being infected, developing symptoms, and needing healthcare.



Lower and upper bound estimates for R (90% Confidence Interval) for England

Source: UK Health Security Agency; <https://www.gov.uk/guidance/the-r-value-and-growth-rate> · Created with Datawrapper

Fig. 3. The reproduction number (R) in England: May 2020 to April 2022.

Overall, we found 14 different ways of attributing the causes of death. The lack of consistent definitions leads to erroneous comparisons and estimates that differ depending on the definition used.

1.5. Sources of evidence for the effectiveness estimate

The Imperial report considered the impact of five different NPIs that included social distancing, stopping mass gatherings and closure of schools and universities implemented individually and in combination. The most effective interventions were predicted to be a combination of case isolation, home quarantine and social distancing of those over 70s. The “optimal mitigation scenario” was projected to result in 8-fold higher peak demand on critical care beds. However, it is unclear what the sources of the effect estimates incorporated in the model were. In March 2020, the Cochrane review of Randomised Controlled Trials of physical interventions against Influenza-Like illnesses (ILIs) and influenza included 67 RCTs and observational studies with a mixed risk of bias. The highest quality cluster-RCTs suggested respiratory virus spread can be prevented by handwashing; however, there was limited evidence that social distancing was effective (Jefferson et al., 2011).

This version included observational studies, but the 2020 version (Jefferson et al., 2020a) (also published as a preprint (Jefferson et al., 2020b)) removed these because there were sufficient published RCTs and Cluster RCTs ($n = 46$) to address the effectiveness of NPIs. Observational studies are subject to recall bias due to participants not accurately remembering previous events or exposures or omitting important details (Catalogue of Bias Collaboration Spencer et al., 2017). For example, in a systematic review of physical distancing, face masks, and eye protection to prevent transmission of SARS-CoV-2, we could not replicate many distance assumptions reported in the paper - only two studies provide precise estimates on distancing. While the review reported physical distancing to be effective, it further noted that ‘robust randomised trials are needed to better inform the evidence for these interventions.’ (Chu et al., 2020).

In October 2019, the World Health Organisation reported on ‘non-pharmaceutical public health measures for mitigating the risk and impact of epidemic and pandemic influenza’. The report included high-quality evidence from 18 systematic reviews about the effects of intervening and used the Grading of Recommendations Assessment, Development and Evaluation (GRADE) method to determine the direction (for or against) and the strength (the amount of certainty) of the recommendations (WHO, 2019). The evidence base on the effectiveness of NPIs in the community was judged as limited, and the overall quality of the evidence was considered very low for most interventions, further mandating the need for high-quality evidence on NPIs.

1.6. Reproduction numbers and reporting delays

Despite the widespread use of the R number, there is often a lack of a relationship with the timing of restrictions (see Fig. 3). This relationship is further affected by reporting lags, which can apply to transmission estimates from 2 to 3 weeks previously. Delays in reporting and dissemination can mean that modelled estimates are outdated by the time they are published. Approaches to regularly updating estimates lead to substantial changes in the outputs. For example, weekly projections on COVID-19 deaths by the (MRC Biostatistics Unit) reported on 12 October 2020 a projected 588 deaths for 30 October; on 21 October, this was revised down to 324

deaths for 31 October (Howdon et al. 2020).

In November 2020, we showed that slides used by the UK Government were out of date: the “worst-case scenario” of deaths based on modelling was weeks out of date, and errors were not unusual (The Telegraph, 2020). However, regularly updating models as data accumulates is resource intensive and, therefore, rarely done.

2. Discussion

Models have been widely used to influence UK policy in the COVID-19 pandemic. However, their use comes with several limitations. First, despite our best efforts, we could not find a taxonomy of different models and good practice guidelines for methods and reporting. The validation of models is also a scarce commodity. Initial overestimates of R_0 compared to the real-world highlight the problems with using early data and generalising context-specific R_0 numbers.

There has been a substantial increase in published modelling studies since the pandemic began. A systematic review of modelling studies estimating the risk of spread and effect of NPIs included 166 studies mostly modelled on data from China and Italy (Dhaoui et al., 2022). The review reported the calibration by fitting models to data was common, whereas the inclusion of uncertainties in disease characteristics and validation to external data was relatively uncommon. Yet, models should undertake robust validation or, at a minimum, explain why validation was not done (Walters et al., 2018).

Because most models are not validated and lack communication of their uncertainties (Eker, 2020) they are poorly understood. The biases and inherent limitations are often unaccounted for in policy decisions, despite clear lessons from previous pandemics.

In the past, models have been influential, leading to the stockpiling of the antiviral oseltamivir to interrupt or delay transmission of influenza H5N1 to “contain an emerging influenza pandemic in SE Asia” (Ferguson et al., 2005). The data upon which the model was based on an incomplete understanding of the circumstances of transmission, a massive exaggeration of the threat with a tone of certainty. The antiviral claimed properties were later revealed to be based on selective ghost-written pharmaceutical-sponsored publications (Cohen, 2009).

Modelling studies lack basic evidence-based standards that aid their quality assessment. They lack a quality framework such as GRADE that provides a systematic approach for grading the quality of the evidence and making recommendations. (<https://www.gradeworkinggroup.org/> Accessed 17 July) Models also lack reporting checklist, (EQUATOR network: www.equator-network.org/) such as CONSORT for RCTs or PRISMA for systematic reviews that seek to improve publication quality, facilitate complete and transparent reporting, and improve their interpretation. (CONSORT: www.consort-statement.org/) There are also no checklists for reporting the sources of heterogeneity, such as the ones that exist for pooling epidemic surveillance systems; (European Centre for Disease Prevention and Control, 2019) no registers for posting protocol methods, and evidence underpinning interventional effects are not readily identifiable. Models fit the available data and recreate the epidemic to that point (UK Government, 2021) and choose methods and assumptions that produce the best fit - hence the variation in methods across studies - however, post hoc methods changes or assumptions are also not identifiable. Hence the need for published protocols. The lack of reproducibility enhances the need for protocols: the US Government Accountability Office reported that ‘CDC’s guidelines and policy do not address the reproducibility of models or their code.’ Further undermining the scientific validity of models. (U.S. Government, 2020) The validity of modelling would be enhanced by developing standardized methods, quality criteria and reporting standards.

Perhaps the most perplexing aspect is the lack of high-quality evidence that underpins the modelled effect of intervening. For example, a societal economic evaluation of interventions to interrupt or slow down the spread of COVID-19 requires a reliable estimate of the impact of SARs-Cov-2; a reliable estimate of the costs and effects of the planned interventions alone or more likely together (multiple interventions) and a credible evaluation of the benefits, expressed as clinical outcomes (lives saved, cases or hospitalisations prevented).

Unfortunately, none of the above currently exists, nor is there an appetite for developing the evidence-base. Because we do not have robust COVID 19 evidence of the effect of NPIs, policy defaults to models - using a non-EBM approach to intervening. Since modelled assumptions inform policy decisions, this sets up a circular argument that means we now don’t require robust evidence as we have the modelled effects of intervening - and therefore, the high-quality evidence gap persists. For example, only two RCTs of face masks for COVID-19 have been published at the time of writing. This lack of evidence for NPIs is at odds with the substantial number of RCTs undertaken for pharmaceuticals, illustrating high-quality evidence can be developed during a pandemic.

Where do models fit in the hierarchical scale of evidence for decision-making? Probably they fit on the same rung as the data they use. Early in the pandemic, government advisors said that models were informed mainly by ‘educated guesswork, intuition and experience’. (Blakey R, 2020) If they use guesswork, they should be placed at the bottom along with expert opinion; if they use RCT data appropriately, higher up. In an ideal world, we’d have all the evidence to hand to inform policy judgments about what to do next. This is often not possible in the early stages of pandemics; however, as the pandemic evolves, robust data should drive decision-making, and high-quality evidence should be the mainstay for reducing uncertainties over whether to intervene.

3. Conclusion

If used judiciously, models may be helpful, especially if they openly present the uncertainties and use sensitivity analyses extensively, which need to consider and explicitly discuss the limitations of the evidence. Models are not an alternative to evidence generation and should be iterative. Until such time as the methodological and ethical issues have been resolved, predictive models should be used with caution: ‘by definition and design, models are not reality (Heesterbeek et al., 2015). The methodological shortcomings and the lack of generalizability of models mean they often fail to translate into the real world. Moreover, the policies that followed fail

to account for the uncertainties in modelling and did not necessarily reflect the high-quality evidence and emerging data.

Author statement

CH wrote the 1st draft and then TJ edited and contributed to subsequent drafts. CH and TJ added equally to the analysis and the final draft of the manuscript.

Funding

CH receives funding from the NIHR School for Primary Care Research Project 569 and is also funded by the NIHR Oxford BRC. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: TJ's competing interests are accessible at: <https://restoringtrials.org/competing-interests-tom-jefferson>. CH holds grant funding from the NIHR School of Primary Care Research, and previously received funding from the World Health Organization for a series of Living rapid review on the modes of transmission of SARs-CoV-2 reference WHO registration No2020/1077093. CH's Conflicts of interest and payments are accessible at: <https://www.phc.ox.ac.uk/team/carl-heneghan>

Data availability

Data will be made available on request.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jbior.2022.100914>.

References

- Alimohamadi, Y., Taghdir, M., Sepandi, M., 2020 May. Estimate of the basic reproduction number for COVID-19: a systematic review and meta-analysis. *J Prev Med Public Health* 53 (3), 151–157. <https://doi.org/10.3961/jpmph.20.076>.
- Aronson, J, et al., 2020. When will it be over? An introduction to viral reproduction numbers, R0 and Re, 17 July. <https://www.cebm.net/covid-19/when-will-it-be-over-an-introduction-to-viral-reproduction-numbers-r0-and-re/>, 2020.
- Catalogue of Bias Collaboration, Banerjee, A., Pluddemann, A., O'Sullivan, J., Nunan, D., 2019. Performance bias. In: *Catalogue Of Bias*, 14 July 2022. <https://catalogofbias.org/biases/performance-bias/>.
- Catalogue of Bias Collaboration, Spencer, E.A., Brassey, J., Mahtani, K., 2017. Recall bias. In: *Catalogue Of Bias*, 17 July 2022. <https://www.catalogueofbiases.org/biases/recall-bias>.
- Catalogue of Biases Collaboration, Spencer, E.A., Brassey, J., 2017. Ascertainment bias. In: *Catalogue of Bias*, 14 July 2022. <https://catalogofbias.org/biases/ascertainment-bias/>.
- BBC News, 2022. As it happened: UK will 'do well to keep deaths under 20,000' BBC news. <https://www.bbc.co.uk/news/live/52075063>.
- Blakey, R, 2020. Models behind coronavirus plans mostly educated guesses. <https://www.thetimes.co.uk/article/models-behind-coronavirus-plans-mostly-educated-guesses-fqmpmt623>.
- Chu, D.K., Akl, E.A., Duda, S., Solo, K., Yaacoub, S., Schünemann, H.J., 2020. & COVID-19 systematic urgent review group effort (SURGE) study authors. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis 395 (10242), 1973–1987. [https://doi.org/10.1016/S0140-6736\(20\)31142-9](https://doi.org/10.1016/S0140-6736(20)31142-9). *Lancet* (London, England).
- Cochrane Handbook for systematic reviews of interventions version 6.3. Chapter 10, 10.10 Heterogeneity, 2022, 14 July 2022. <https://training.cochrane.org/handbook/current/chapter-10#section-10-10>.
- Cohen, D., 2009 Dec 8. Complications: tracking down the data on oseltamivir. *BMJ* 339, b5387. <https://doi.org/10.1136/bmj.b5387>. Erratum in: *BMJ*. 2010;340:c406. PMID: 19995818.
- Coronavirus, 26 March 2021. What is the R number and how is it calculated?, 14 July 2022. <https://www.bbc.co.uk/news/health-52473523>.
- Delamater, P.L., Street, E.J., Leslie, T.F., Yang, Y.T., Jacobsen, K.H., 2019a. Complexity of the basic reproduction number (R0). *Emerg Infect Dis* 25 (1), 1–4.
- Delamater, P.L., Street, E.J., Leslie, T.F., Yang, Y.T., Jacobsen, K.H., 2019 Jan. Complexity of the basic reproduction number (R0). *Emerg Infect Dis* 25 (1), 1–4. <https://doi.org/10.3201/eid2501.171901>.
- Dhaoui, W. Van Bortel, Arsevska, E., Hautefeuille, C., Alonso, S. Tablado, Kleef, E.V., 2022. Mathematical modelling of COVID-19: a systematic review and quality assessment in the early epidemic response phase. *International Journal of Infectious Diseases* 116. <https://doi.org/10.1016/j.ijid.2021.12.260>. Page S110.
- Eker, S., 2020. Validity and usefulness of COVID-19 models. *Humanit Soc Sci Commun* 7, 54. <https://doi.org/10.1057/s41599-020-00553-4>.
- Ferguson, N.M., Cummings, D.A., Cauchemez, S., Fraser, C., Riley, S., Meeyai, A., Iamsirithaworn, S., Burke, D.S., 2005 Sep 8. Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature* 437 (7056), 209–214. <https://doi.org/10.1038/nature04017>.
- Ferguson, et al., 2020. Impact of Non-pharmaceutical Interventions (NPIs) to Reduce COVID-19 Mortality and Healthcare Demand. Imperial College COVID-19 Response Team, London, 17 July. <https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf>.
- Gov.UK press briefing, 2022. Foreign Secretary's statement on coronavirus (COVID-19): 30 March 2020 Press briefing. <https://www.gov.uk/government/speeches/foreign-secretarys-statement-on-coronavirus-covid-19-30-march-2020>.
- Guerra, F.M., Bolotin, S., Lim, G., Heffernan, J., Deeks, S.L., Li, Y., et al., 2017. The basic reproduction number (R0) of measles: a systematic review. *Lancet Infect Dis* 17, e420–e428.

- Heesterbeek, H., Anderson, R.M., Andreasen, V., Bansal, S., De Angelis, D., Dye, C., Eames, K.T., Edmunds, W.J., Frost, S.D., Funk, S., Hollingsworth, T.D., House, T., Isham, V., Klepac, P., Lessler, J., Lloyd-Smith, J.O., Metcalf, C.J., Mollison, D., Pellis, L., Pulliam, J.R., Roberts, M.G., Viboud, C., Isaac Newton Institute IDD Collaboration, 2015 Mar 13. Modeling infectious disease dynamics in the complex landscape of global health. *Science* 347 (6227), aaa4339. <https://doi.org/10.1126/science.aaa4339>.
- Hilton, J., Keeling, M.J., 2020 Jul 2. Estimation of country-level basic reproductive ratios for novel Coronavirus (SARS-CoV-2/COVID-19) using synthetic contact matrices. *PLoS Comput Biol* 16 (7), e1008031. <https://doi.org/10.1371/journal.pcbi.1008031>. <https://www.gradeworkinggroup.org/>. PMID: 32614817; PMCID: PMC7363110, 17 July.
- Howdon, D., Heneghan, C., November 2020. SAGE models overestimation of deaths. <https://www.cebm.net/covid-19/innacuracies-in-sage-models/>. Independent review into the response to the 2009 swine flu pandemic. <https://www.gov.uk/government/publications/independent-review-into-the-response-to-the-2009-swine-flu-pandemic>, 1 July 2010.
- Jefferson, T, et al., 2022.04.28. PCR testing in the UK during the SARS-CoV-2 pandemic – evidence from FOI requests. *MedRxiv*, 22274341. <https://doi.org/10.1101/2022.04.28.22274341>.
- Jefferson, T., Del Mar, C., Dooley, L., Ferroni, E., Al-Ansary, L.A., Bawazeer, G.A., et al., 2011. Physical interventions to interrupt or reduce the spread of respiratory viruses. *Cochrane Database of Systematic Reviews* 7, CD006207. <https://doi.org/10.1002/14651858.CD006207.pub4>.
- Jefferson, T., Del Mar, C.B., Dooley, L., Ferroni, E., Al-Ansary, L.A., Bawazeer, G.A., van Driel, M.L., Jones, M.A., Thorning, S., Beller, E.M., Clark, J., Hoffmann, T.C., Glasziou, P.P., Conly, J.M., 2020a. Physical interventions to interrupt or reduce the spread of respiratory viruses. *Cochrane Database of Systematic Reviews* 11, CD006207. <https://doi.org/10.1002/14651858.CD006207.pub5>, 03 May 2022.
- Jefferson, T., Jones, M.A., Al-Ansary, L., et al., 2020.03.30. Physical interventions to interrupt or reduce the spread of respiratory viruses. Part 1 - face masks, eye protection and person distancing: systematic review and meta-analysis. *medRxiv*, 20047217. <https://doi.org/10.1101/2020.03.30.20047217>.
- Li, Q., Guan, X., Wu, P., et al., 2020. Early transmission dynamics in wuhan, China, of novel coronavirus–infected pneumonia. *N Engl J Med* 2, 2.
- MRC Biostatistics Unit. Nowcasting and forecasting of the COVID-19 pandemic, 16 July. <https://www.mrc-bsu.cam.ac.uk/tackling-covid-19/nowcasting-and-forecasting-of-covid-19/>.
- Prime Minister’s statement on coronavirus (COVID-19), 14 July 2022. <https://www.gov.uk/government/speeches/pm-statement-on-coronavirus-11-may-2020>.
- Ridenhour, B., Kowalik, J.M., Shay, D.K., 2014. Unraveling R0: considerations for public health applications. *Am J Public Health* 104, e32–41.
- Riou, J., Althaus, C.L., 2020. Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020. *Euro Surveill* 25 (4), 1–5.
- Skip, L.A., Townsend, J.P., 2019 Jun 28. Modeling approaches toward understanding infectious disease transmission. *Immunoepidemiology* 227–243. https://doi.org/10.1007/978-3-030-25553-4_14. PMCID: PMC7121152.
- Special report: the simulations driving the world’s response to COVID-19 How epidemiologists rushed to model the coronavirus pandemic, 22 July. <https://www.nature.com/articles/d41586-020-01003-6>.
- The Royal Society, 24 August 2020. Reproduction number (R) and growth rate (r) of the COVID-19 epidemic in the UK: methods of estimation, data sources, causes of heterogeneity, and use as a guide in policy formulation, 14 July 2022. <https://royalsociety.org/-/media/policy/projects/set-c/set-covid-19-R-estimates.pdf>.
- The Telegraph, 2020. The Government’s use of data is not just confusing – the errors are positively misleading. <https://www.telegraph.co.uk/news/2020/11/05/governments-use-data-not-just-confusing-errors-positively-misleading/>.
- Tolles, J., Luong, T., 2020. Modeling epidemics with compartmental models. *JAMA* 323 (24), 2515–2516. <https://doi.org/10.1001/jama.2020.8420>.
- Tovissodé, Chénangnon Frédéric, et al., 20 Oct. 2020. On the use of growth models to understand epidemic outbreaks with application to COVID-19 data. *PLoS one* 15 (10), e0240578. <https://doi.org/10.1371/journal.pone.0240578>.
- UK Government, October 2021. Introduction to Epidemiological Modelling. UK.GOV. <https://www.gov.uk/government/publications/introduction-to-epidemiological-modelling/introduction-to-epidemiological-modelling-october-2021>.
- Understanding definitions and reporting of deaths attributed to COVID-19 in the UK – evidence from FOI requests. Tom Jefferson, Madeleine Dietrich, Jon Brassey, Carl Heneghan *medRxiv*, 2022.04.28, 22274344. <https://doi.org/10.1101/2022.04.28.22274344>.
- US Government, 2020. Accountability Office infectious disease modeling: opportunities to improve coordination and ensure reproducibility. <https://www.gao.gov/products/gao-20-372>.
- Verma, Bl, Ray, S.K., Srivastava, R.N., 1981 Jan-Mar. Mathematical models and their applications in medicine and health. *Health Popul Perspect Issues* 4 (1), 42–58. PMID: 10260952.
- Walters, C.E., Meslé, M.M.I., Hall, I.M., 2018 Dec. Modelling the global spread of diseases: a review of current practice and capability. *Epidemics* 25, 1–8. <https://doi.org/10.1016/j.epidem.2018.05.007>.
- World Health Organization. Non-pharmaceutical public health measures for mitigating the risk and impact of epidemic and pandemic influenza; 2019. Licence: CC BY-NC-SA 3.0 IGO. <https://apps.who.int/iris/bitstream/handle/10665/329438/9789241516839-eng.pdf>.
- European Centre for Disease Prevention and Control. Managing heterogeneity when pooling data from different surveillance systems. Stockholm: ECDC; 2019. Stockholm, October 2019 PDF ISBN 978-92-9498-383-1 doi: 10.2900/83039.