Five ways to make models serve society

Pandemic politics reminds that predictions must be social, transparent, and humble to invite insight, not blame.

Andrea Saltelli, Gabriele Bammer, Isabelle Bruno, Erica Charters, Monica Di Fiore, Emmanuel Didier, Wendy Nelson Espeland, John Kay, Samuele Lo Piano, Deborah Mayo, Roger Pielke Jr., Tommaso Portaluri, Theodore M. Porter, Arnald Puy, Ismael Rafols, Jerome R. Ravetz, Erik Reinert, Daniel Sarewitz, Philip B. Stark, Andrew Stirling, Jeroen van der Sluijs, Paolo Vineis

Well before the pandemic, statisticians were debating how to prevent statistical malpractice, such as p-hacking, particularly when it could influence policy. Now computer modelling has come into the media and policy limelight, with politicians presenting their policies as dictated by 'science' in the form of a mathematical model [FC REF][1]. Yet there is no aspect of this pandemic for which any researcher can currently provide precise, reliable numbers. Known unknowns include the prevalence of the virus in the population, its fatality and reproduction rates. Estimates of the proportion of asymptomatic individuals (crucial to understanding infection dynamics) have ranged from 5% to 80% [FC REF][2]. We know even less about the seasonality of infections, how immunity works, not to mention the impact of social distancing interventions in diverse, complex societies.

Mathematical models produce highly uncertain numbers predicting future infections, hospitalizations, and deaths under a variety of scenarios. Rather than using models to inform their understanding, political rivals often brandish them to support pre-determined agendas. To make sure their predictions do not become mere adjuncts to a political agenda, modellers, decision makers, and citizens need to establish new social norms such that modellers are not permitted to project more certainty than their models deserve, and politicians are not allowed to offload accountability to models of their choosing[1,3].

This is important because, when used appropriately, models serve society extremely well: perhaps the best known are weather forecasting models. To be sure, these have been honed over millions of forecasts tested against reality. But so have ways to communicate results to diverse users: from the Digital Marine Weather Dissemination System produced by the US National Oceanic and Atmospheric Administration and used by oceangoing vessels to the hourly forecasts accumulated by Weather.com's. Picnickers, airline executives and fishermen alike understand both that the modelling outputs are fundamentally uncertain and how to factor their predictions into decisions.

Here we present a plea for best practices for responsible mathematical modelling. Many groups before us have come together to describe the best way to apply modelling insights to policies, including modelling disease for policy[4]. Accounts of this work and the need along with background for it are detailed in the accompanying supplemental materials, organized according to paragraphs in this text. We distil five simple principles to help society demand the quality it needs from modelling.

1. Mind the assumptions: assess uncertainty and sensitivity

Models are often imported from one application to another, ignoring how reasonable assumptions in one situation can become nonsensical. Models that work for civil nuclear risk are not justified to assess seismic risk. Another lapse occurs when models require input values that lack any reliable information. A model used in the UK to guide transport policy depends on a guess for the number of passengers in each car three decades from now[5].

One way to mitigate these issues is to explicitly perform global uncertainty and sensitivity analyses. In practice, that means allowing all that is uncertain –variables, mathematical relationships and boundary conditions, to vary simultaneously as runs of the model produce its range of predictions. When done, this often reveals that the uncertainty in predictions is substantially larger than originally asserted. For example, an analysis by three of us (ASaltelli, AP, SLP) suggests that estimates of how much land will be irrigated for crops in future range over fivefold when extant models properly integrate uncertainties on future population growth rates, spread of irrigated areas, and the mathematical relationship between the two [6].

However, these global uncertainty and sensitivity analyses are often not done. Anyone turning to a model for insight should demand that such analyses are done and their results made accessible.

2. Mind the hubris: complexity can be the enemy of relevance

Most modellers are aware that there is a trade-off between the usefulness of a model and the breadth it tries to capture, but both modellers and politicians are seduced by adding more complexity in an attempt to better capture reality. As modellers incorporate more phenomena into a model, its predictions may improve, but at a cost: quality control becomes more difficult. As even more parameters are added, the uncertainty they carry builds up (a phenomenon known as the uncertainty cascade effect), and the error increases to the point where predictions become useless. A model's complexity does not ensure that all important features are captured. In the case of HIV infection, a simpler model that focused on promiscuity rather than frequency of sexual activity turned out to be more reliable[5]. The discovery of the existence of "superspreading events" and "superspreader" subjects for COVID-19 similarly shows how surprises may arise.

One extreme example of excess complexity is the Total System Performance Assessment model used by the US Department of Energy to evaluate risk in disposing radioactive waste at the Yucca Mountain repository. Composed of 286 sub-models with thousands of parameters, the model was tasked by regulators to predict "one million years" of safety. Yet the uncertainty range of a single key variable—the time needed for water to percolate to the underground repository level— was uncertain by three orders of magnitude, rendering the size of the model irrelevant[7].

Complexity is too often seen as an end in itself. Instead, the goal must be finding the optimum balance with error.

What's more, people trained in building models are often not drilled or incentivized for such analyses. While an engineer is called to task if a bridge falls, other models tend to be developed with large teams using such complex feedback loops that no one is held accountable if predictions are catastrophically wrong.

3. Mind the framing: match purpose and context

Results from models will at least partly reflect the interests, disciplinary orientations, and biases of developers. No one model can serve all purposes.

Modellers know that choice of tools will influence and may even determine the outcome of the analysis, so technique is never completely neutral. For example, the GENESIS model of shoreline erosion was used by the US Army Corps of Engineers to support cost-benefit assessments for beach preservation projects. The model could not predict realistically the mechanisms of beach erosion by waves nor the effectiveness of beach replenishment by human intervention but this way of framing the problem could easily be manipulated to meet the political desire for such projects[7].

Shared approaches to assessing quality need to be accompanied by a shared commitment to transparency. 'Cost-benefit,' 'expected utility', 'decision theory', 'life cycle assessment', 'ecosystem services', and 'evidence-based policy' are examples of terms that promise uncontested precision. All in fact presuppose a set of values about what matters – sustainability for some, productivity or profitability for others[3,8]. Modellers should not hide the normative values of their choices.

Consider the value of a statistical life, loosely defined as the cost of averting a death. It is already controversial for setting compensation, e.g. for the victims of an air crash. While it may have a place in discriminating among alternative public health policies, it can produce a false appearance of rigour and so disguise political decisions as technical ones.

The best defence to keep models from hiding their assumptions – including those which are political, is a set of social norms. [PQ] These should cover how to produce a model, assess its uncertainty, and communicate results. International guidelines for this have been drawn up in several disciplines; they demand that processes involve stakeholders, require enriching the set of framings so as to accommodate multiple views, and promote transparency, replication, and sensitivity and uncertainty analysis. Whenever a model is used for a new application with fresh stakeholders it must be validated and verified anew.

Existing guidelines for infectious disease modelling reflect these concerns, but the adoption is not widespread[4]. Simplified, plain-language versions of the model can be crucial: when a model is no longer a black box, those using it must react to assess individual parameters and relationships between them. This allows outcomes to be communicated in a way that shows how different framings and assumptions map into different inferences, rather than communicating a single over-simplified interpretation from an overly complex model. Or to put it in jargon: qualitative descriptions of multiple reasonable sets of assumptions can be as important in improving insight among decision makers as the delivery of quantitative results.

Positive examples along these dimensions can be found in forecasting flooding risk, and in the management of fisheries, where stakeholders' insights about both inputs and desired ends ensured the success of the modelling process. In both examples new models were coded to incorporate their intuitions.

4. Mind the consequences: quantification may backfire

Excessive regard for producing numbers can push a discipline away from being roughly right toward being precisely wrong, with statistical tests substituting for sound judgement[9]. Models contributed to crippling the global economy in 2007-8 by helping to make risky financial products seem safe[5].

Once a number takes the center-stage, with a crisp narrative, other possible explanations and estimates may disappear from view. This may invite complacency, and the politicization of quantification, as other options are marginalised. In the case of COVID, issues as diverse as availability of intensive care hospital beds, employment, and civil liberties are simultaneously at play, even if they cannot be simply quantified and then plugged into the models.

Spurious precision adds to a false sense of certainty. If we are told that the number of deaths in the UK in case of no mitigation actions will be 510 thousand (Imperial College model estimate[10]), we imagine a corresponding two digits level of confidence. Instead, the even limited analysis of the uncertainty run by the authors – based on just one parameter - reveals a range of 410-550 thousand deaths. An analysis by the WHO Africa predicts up to a maximum of 190 thousand deaths for that continent.

Analysis of the report reveals that this number corresponds to a highly speculative case, where ten uncertain input probabilities are increased by an arbitrary 10%, as if they were truly equally uncertain, with no theoretical or empirical basis for such a choice. While thought experiments are useful, they should not be treated as a prediction.

Opacity about uncertainty damages trust. A message from the field of sociology of quantification[11] is that trust is essential for quantification to be useful[8]. Full explanations are essential.

5. Mind the unknowns: acknowledge ignorance

For most of the history of Western philosophy, awareness of one's ignorance was considered a virtue, the worthy object of intellectual pursuit: learned ignorance, *docta ignorantia*, in the words of XV century's Nicolas of Cusa. Even today, communicating what is not known is at least as important as communicating what is known. Yet models can hide ignorance[3,5,7].

Failure to acknowledge 'not-knowing' may artificially limit the policy options. Worse, it may offer politicians the chance to abdicate accountability. Experts should have the courage to answer "there is no number-answer" to your question, as US government epidemiologist Anthony Fauci did when probed by a politician[12].

Questions not answers

COVID-19 illustrates perfectly the mutated conditions of operation of science when moved from normality to a post normal regime, where urgency, stakes, values and uncertainty collide.

Models are a great way of exploring these questions. They are a dangerous way to assert answers. Asking models for certainty or consensus is more a sign of the difficulties of making controversial decisions than it is a solution, and can invite ritualistic use of quantification.

Models' assumptions and limitations must be appraised openly and honestly. Process and ethics matter as much as intellectual prowess. All of which is to say that good modelling cannot be done by modellers alone. It is a social activity. The French movement of statactivistes has successfully shown how numbers can be fought with numbers, e.g. in the quantification of poverty and inequalities[13].

A form of societal activism on the relation between models and society is offered by Tomas Pueyo, not an epidemiologist, who maintains a blog for COVID-19 epidemiological models and explains in plain-language the implications of model uncertainties for policy options.

We are calling not for an end to quantification, nor for apolitical models, but for full and frank disclosure. Following these five points will preserve mathematical modelling as a valuable tool. Each contributes to the overarching goal of billboarding the strengths and limits of model outputs. Ignore the five and model predictions become Trojan horses for unstated interests and values. Model responsibly.

1. Devlin, H. & Boseley, S. Scientists criticise UK government's 'following the science' claim. *The Guardian* (2020).

2. Heneghan, C., Brassey, J. & Jefferson, T. COVID-19: What proportion are asymptomatic? - CEBM. *Centre for Evidence-Based Medicine* (2020). Available at: https://www.cebm.net/covid-19/covid-19-what-proportion-are-asymptomatic/. (Accessed: 29th April 2020)

3. Stirling, A. How politics closes down uncertainty - STEPS Centre. *STEPS Centre* Feb. 20 (2019). Available at: https://steps-centre.org/blog/how-politics-closes-down-uncertainty/. (Accessed: 31st March 2019)

4. Behrend, M. R. *et al.* Modelling for policy: The five principles of the Neglected Tropical Diseases Modelling Consortium. *PLoS Negl. Trop. Dis.* **14**, e0008033 (2020).

5. Kay, J. A. & King, M. A. *Radical uncertainty : decision-making beyond the numbers*. (W. W. Norton & Company, 2020).

6. Puy, A., Piano, S. Lo & Saltelli, A. Current models underestimate future irrigated areas. *Geophys. Res. Lett.* (2020). doi:10.1029/2020GL087360

7. Sarewitz, D., Pielke, R. A. & Byerly, R. *Prediction: Science, Decision Making, and the Future of Nature*. (Island Press, 2000).

8. Porter, T. M. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. (Princeton University Press, 1996).

9. Mayo, D. G. *Statistical Inference as Severe Testing*. *Statistical Inference as Severe Testing* (Cambridge University Press, 2018). doi:10.1017/9781107286184

10. Neil M.Ferguson *et al.* Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. *Imperial College London* (2020). Available at: https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf.

11. Espeland, W. N. & Stevens, M. L. A sociology of quantification. *Eur. J. Sociol.* **49**, 401–

436 (2008).

12. Fauci, A. There is no number-answer. *Twitter* (2020). Available at: https://twitter.com/marioricciard18/status/1237778247011663872.

13. Bruno, I., Didier, E. & Vitale, T. Editorial: Statactivism: forms of action between disclosure and affirmation. *Open J. Sociopolitical Stud.* **2**, 198–220 (2014).