Comment

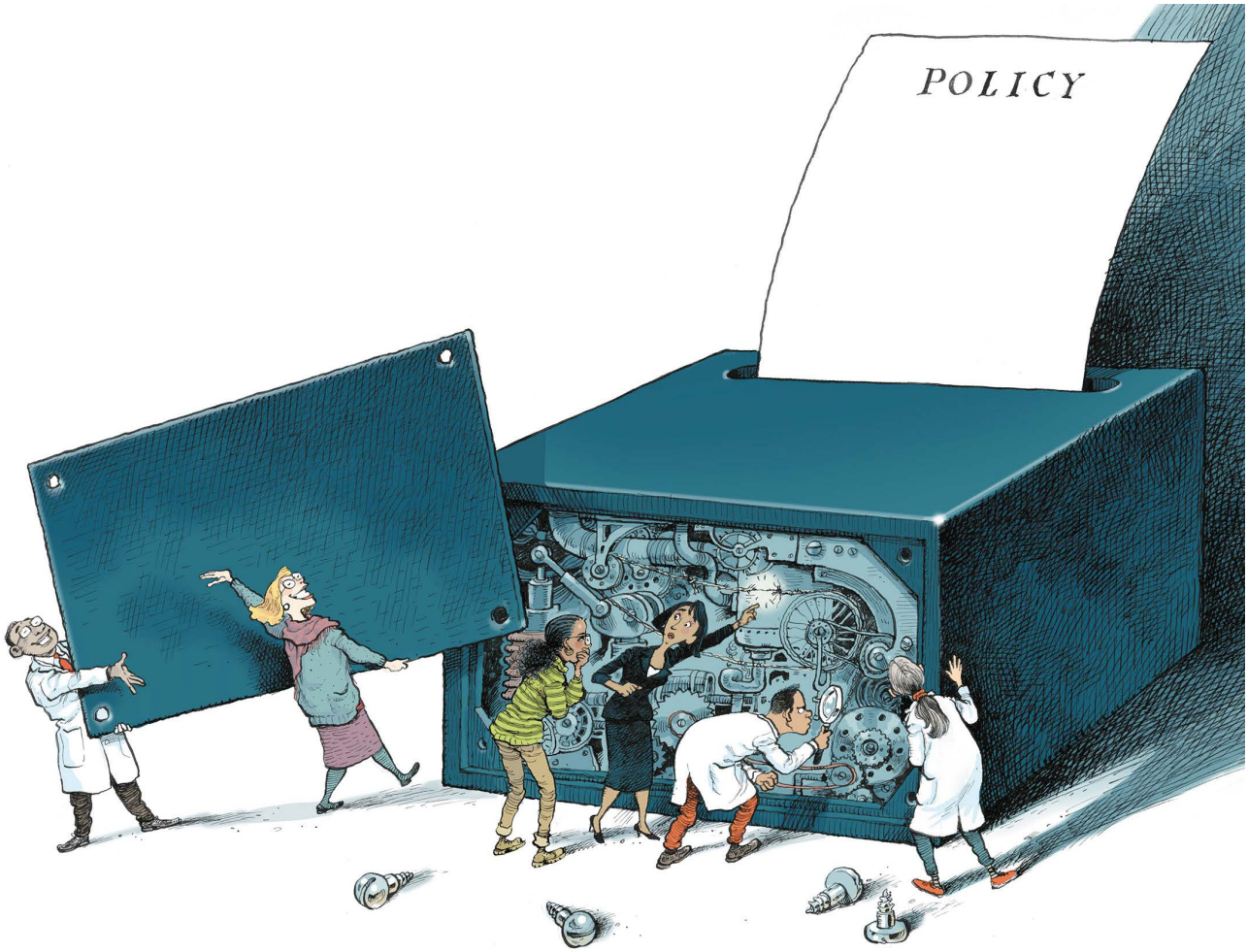

ILLUSTRATION BY DAVID PARKINS

Five ways to ensure that models serve society: a manifesto

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

Pandemic politics highlight how predictions need to be transparent and humble to invite insight, not blame.

The COVID-19 pandemic illustrates perfectly how the operation of science changes when questions of urgency, stakes, values and uncertainty collide – in the 'post-normal' regime.

Well before the coronavirus pandemic, statisticians were debating how to prevent malpractice such as *p*-hacking, particularly when it could influence policy[1]. Now, computer modelling is in the limelight, with politicians presenting their policies as dictated by 'science'[2]. Yet there is no substantial aspect of this pandemic for which any researcher can currently provide precise, reliable numbers. Known unknowns include the prevalence and fatality and reproduction rates of the virus in

© 2020 Springer Nature Limited. All rights reserved.

populations. There are few estimates of the number of asymptomatic infections, and they are highly variable. We know even less about the seasonality of infections and how immunity works, not to mention the impact of social-distancing interventions in diverse, complex societies.

Mathematical models produce highly uncertain numbers that predict future infections, hospitalizations and deaths under various scenarios. Rather than using models to inform their understanding, political rivals often brandish them to support predetermined agendas. To make sure predictions do not become adjuncts to a political cause, modellers, decision makers and citizens need to establish new social norms. Modellers must not be permitted to project more certainty than their models deserve; and politicians must not be allowed to offload accountability to models of their choosing[2,3].

This is important because, when used appropriately, models serve society extremely well: perhaps the best known are those used in weather forecasting. These models have been honed by testing millions of forecasts against reality. So, too, have ways to communicate results to diverse users, from the Digital Marine Weather Dissemination System for ocean-going vessels to the hourly forecasts accumulated by weather.com. Picnickers, airline executives and fishers alike understand both that the modelling outputs are fundamentally uncertain, and how to factor the predictions into decisions.

Here we present a manifesto for best practices for responsible mathematical modelling. Many groups before us have described the best ways to apply modelling insights to policies, including for diseases[4] (see also Supplementary information). We distil five simple principles to help society demand the quality it needs from modelling.

Mind the assumptions

Assess uncertainty and sensitivity. Models are often imported from other applications, ignoring how assumptions that are reasonable in one situation can become nonsensical in another. Models that work for civil nuclear risk might not adequately assess seismic risk. Another lapse occurs when models require input values for which there is no reliable information. For example, there is a model used in the United Kingdom to guide transport policy that depends on a guess for how many passengers will travel in each car three decades from now[5].

One way to mitigate these issues is to 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. This often reveals that the uncertainty in predictions is substantially larger than originally asserted. For example, an analysis by three of us (A.Saltelli, A.P., S.L.P.) suggests that estimates of how much land will be irrigated for future crops varies more than fivefold when extant models properly integrate uncertainties on future population growth rates, spread of irrigated areas and the mathematical relationship between the two[6].

"The best way to keep models from hiding their assumptions, including political leanings, is a set of social norms."

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

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 many are seduced by the idea of adding complexity in an attempt to capture reality more accurately. As modellers incorporate more phenomena, a model might fit better to the training data, but at a cost. Its predictions typically become less accurate. As more parameters are added, the uncertainty builds up (the uncertainty cascade effect), and the error could increase to the point at which predictions become useless.

The complexity of a model is not always an indicator of how well it captures the important features. In the case of HIV infection, a simpler model that focuses on promiscuity turned out to be more reliable than a more involved one based on frequency of sexual activity[5]. The discovery of the existence of 'superspreading events' and 'superspreader' people with COVID-19 similarly shows how an unanticipated feature of transmission can surprise the analyst.

One extreme example of excess complexity is a model used by the US Department of Energy to evaluate risk in disposing of radioactive waste at the Yucca Mountain repository. Called the total system performance assessment, it comprised 286 sub-models with thousands of parameters. Regulators tasked it with predicting "one million years" of safety. Yet a single key variable — the time needed for water to percolate down 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. Whereas an engineer is called to task if a bridge falls, other models tend to be developed with large teams and use such complex feedback loops that no one can be held accountable if the predictions are catastrophically wrong.

Mind the framing

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

Modellers know that the choice of tools will influence, and could even determine, the outcome of the analysis, so the technique is never 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 cost–benefit model could not predict realistically the mechanisms of beach erosion by waves or the effectiveness of beach replenishment by human intervention. It could be easily manipulated to boost evidence that certain coastal-engineering projects would be beneficial[7]. A fairer assessment would have considered how extreme storm events dominate in erosion processes.

Shared approaches to assessing quality need to be accompanied by a shared commitment to transparency. Examples of terms that promise uncontested precision include: 'cost–benefit', 'expected utility', 'decision theory', 'life-cycle assessment', 'ecosystem services', and 'evidence-based policy'. Yet all 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

© 2020 Springer Nature Limited. All rights reserved.

compensation — for the victims of aeroplane crashes, for instance. Although it might have a place in choosing the best public-health policy, it can produce a questionable appearance of rigour and so disguise political decisions as technical ones[8].

The best way to keep models from hiding their assumptions, including political leanings, is a set of social norms. These should cover how to produce a model, assess its uncertainty and communicate the results. International guidelines for this have been drawn up for several disciplines. They demand that processes involve stakeholders, accommodate multiple views and promote transparency, replication and analysis of sensitivity and uncertainty. 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 have not been widely adopted[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 the relationships between them. This makes it possible to communicate how different framings and assumptions map into different inferences, rather than just a single, 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 in decision makers as the delivery of quantitative results.

Examples of models that have adhered to these guidelines can be found in forecasting flooding risk, and in the management of fisheries. These included stakeholders' insights and intuitions about both inputs and desired ends.

Mind the consequences

Quantification can backfire. Excessive regard for producing numbers can push a discipline away from being roughly right towards being precisely wrong. Undiscriminating use of statistical tests can substitute for sound judgement. By helping to make risky financial products seem safe, models contributed to derailing the global economy in 2007–08 (ref. 5).

Once a number takes centre-stage with a crisp narrative, other possible explanations and estimates can disappear from view. This might invite complacency, and the politicization of quantification, as other options are marginalized. In the case of COVID-19, 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 modellers tell the United Kingdom it will see 510,000 deaths[9] if no steps are taken to mitigate the pandemic, some might imagine a confidence of two significant digits. Instead, even the limited uncertainty analysis run by the modellers — based on just one parameter — reveals a range of 410,000–550,000 deaths. Similarly, the World Health Organization predicts up to 190,000 deaths for Africa (see go.nature.com/3hdy8kn). That number corresponds to a speculative scenario in which 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. Although thought experiments are useful, they should not be treated as predictions.

Opacity about uncertainty damages trust. A message from the field of sociology of quantification[10] is that trust is essential for numbers to be useful[8]. Full explanations are crucial.

Mind the unknowns

Acknowledge ignorance. For most of the history of Western philosophy, self-awareness of ignorance was considered a virtue, the worthy object of intellectual pursuit — what the fifteenth-century philosopher Nicholas of Cusa called learned ignorance, or *docta ignorantia*. Even today, communicating what is not known is at least as important as communicating what is known. Yet models can hide ignorance.

> **"Although thought experiments are useful, they should not be treated as predictions."**

Failure to acknowledge this can artificially limit the policy options and open the door to undesired surprises. Take, for instance, those that befell the heads of governments when the economists in charge admitted that their models — by design — could not predict the last recession. Worse, neglecting uncertainties could offer politicians the chance to abdicate accountability. Experts should have the courage to respond that "there is no number-answer to your question", as US government epidemiologist Anthony Fauci did when probed by a politician.

Questions not answers

Mathematical models are a great way to explore questions. They are also a dangerous way to assert answers. Asking models for certainty or consensus is more a sign of the difficulties in 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. It follows, in our view, that good modelling cannot be done by modellers alone. It is a social activity. The French movement of *statactivistes* has shown how numbers can be fought with numbers, such as in the quantification of poverty and inequalities[11].

A form of societal activism on the relationship between models and society is offered by US-based engineer-entrepreneur Tomás Pueyo. He is not an epidemiologist, but writes about COVID-19 models and explains in plain language the implications of 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 help to 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.

The authors

Andrea Saltelli is a professor at the Center for the Study of the Sciences and the Humanities, University of Bergen, Norway, and at Open Evidence Research, Open University of Catalonia, Barcelona, Spain.
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.
e-mail: andrea.saltelli@uib.no

The full list of affiliations is available online at go.nature.com/2ce8uqt. Supplementary information accompanies this Comment (go.nature.com/2ce8uqt).

1. Mayo, D. G. *Statistical Inference as Severe Testing.* (Cambridge Univ. Press, 2018).
2. Devlin, H. & Boseley, S. 'Scientists criticise UK government's 'following the science' claim' (*The Guardian*, 23 April 2020).
3. Stirling, A. 'How politics closes down uncertainty'. Available at https://go.nature.com/3kjvutz.
4. Behrend, M. R. *et al. 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., Lo Piano, S. & Saltelli, A. *Geophys. Res. Lett.* **47**, e2020GL087360 (2020).
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 Univ. Press, 1996).
9. Ferguson, N. M. *et al. Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand* (Imperial College London, 2020).
10. Espeland, W. N. & Stevens, M. L. *Eur. J. Sociol.* **49**, 401–436 (2008).
11. Bruno, I., Didier, E. & Vitale, T. *Partecipazione conflitto* **7**, 198–220 (2014).

© 2020 Springer Nature Limited. All rights reserved.