Received: 15 January 2022 | Accepted: 30 June 2022

DOI: 10.1111/1467-9566.13535

SOCIOLOGY OF HEALTH & ILLNESS

ORIGINAL ARTICLE

Uncomfortable science: How mathematical models, and consensus, come to be in public policy

Tim Rhodes[1,2] | **Kari Lancaster**[2]

[1]London School of Hygiene and Tropical Medicine, London, UK

[2]University of New South Wales, Sydney, Australia

Correspondence
Tim Rhodes, London School of Hygiene and Tropical Medicine, 15-17 Tavistock Place, London, WC1E 8HT, UK.
Email: tim.rhodes@lshtm.ac.uk

Funding information
Australian Research Council, Grant/Award Number: DP210101604

Abstract

We explore messy translations of evidence in policy as a site of 'uncomfortable science'. Drawing on the work of John Law, we follow evidence as a 'fluid object' of its situation, also enacted in relation to a hinterland of practices. Working with the qualitative interview accounts of mathematical modellers and other scientists engaged in the UK COVID-19 response, we trace how models perform as evidence. Our point of departure is a moment of controversy in the public announcement of second national lockdown in the UK, and specifically, the projected daily deaths from COVID-19 presented in support of this policy decision. We reflect on this event to trace the messy translations of "scientific consensus" in the face of uncertainty. Efforts among scientists to realise evidence-based expectation and to manage the troubled translations of models in policy, including via "scientific consensus", can extend the dis-ease of uncomfortable science rather than clean it up or close it down. We argue that the project of evidence-based policy is not so much in need of technical management or repair, but that we need to be thinking altogether differently.

KEYWORDS
COVID-19, evidence, evidence-based policy, mathematical models, scientific consensus, translation

INTRODUCTION

Mathematical models can play a key role in generating evidence for use in policy and planning, especially in the face of risk and uncertainty. The COVID-19 pandemic, like pandemics before it, is no exception (Brooks-Pollock et al., 2021). In the UK, models and projections have featured in Government planning across multiple sectors, from health to environment to security, framed by an "evidence-based policy" approach which has gained momentum over the past 20 years (Cabinet Office, 1999; Cairney, 2022; Performance and Innovation Unit, 2000). In uncertainty, models perform a bridge to knowing by generating forecasts and scenarios of what might or could happen. This affords anticipatory governance through projections; calculations of unknown futures made in the present to help shape policy decisions and precautionary actions (Adams et al., 2009; Samimian-Darash, 2016). A common distinction in the use of models as evidence in policy is between more abstracted projections generated as scenarios, often of more distant futures and possibilities, and more empirically grounded predictions and forecasts in the shorter-term (Huppert & Katriel, 2013; Lipsitch et al., 2011). An evidence-based approach idealises using scientific knowledge in a technical–rational decision-making process to produce better policy outcomes (Smith, 2013). In such an approach, models are imagined to progressively attune to their contexts through iterative empirical grounding over time (Glasser et al., 2011). Policy is assumed to be informed by, if not consequent upon, emergent modelled evidence and how this is translated via scientific consensus (Brooks-Pollock et al., 2021). A prominent claim in the COVID-19 pandemic in the UK, for instance, is that Government policy decisions are "following the science" (Cairney, 2021; Evans, 2022).

Models in pandemic

Precisely how scientific knowledge generated through models is made to work as evidence in policy in situations of uncertainty is, however, a key question. In the UK COVID-19 pandemic, the work of translating modelling as evidence for policy has been undertaken via several committees and advisory groups. In particular, the UK Government has sought scientific advice from its SPI-M Committee (Scientific Pandemic Influenza Group on Modelling), an advisory body comprising mathematical modellers and other scientists, and SAGE Committee (Scientific Advisory Group in Emergencies), the body responsible for translating "scientific consensus", including from SPI-M, into evidence for policy advice to Government Cabinet Office. SPI-M was established to assemble evidence on the epidemiology of influenza and other pandemics, and in the COVID-19 pandemic has generated routine calculations of epidemic growth along with short-term forecasts and longer-term scenarios in response to Government commissions. SPI-M produces publicly available consensus statements as evidence produced from multiple models. These also inform the development of Reasonable Worst Case Scenarios (RWCS) for use by Government in planning. Models also "go public" with the communication of projected infections, hospitalisations and deaths: a daily feature of the UK pandemic in 2020 and 2021, supplemented by televised Government press briefings.

Models and projections have thus become sites of public deliberation, as well as contestation, in the UK policy response, including in relation to unprecedented national lockdowns. Published debate has concentrated on matters of method, accuracy and uncertainty (Cairney, 2021; Pawson, 2021; Brooks-Pollock et al., 2021; Salteli et al., 2020; Sasse et al., 2020). There is an emerging social studies of science literature investigating how models perform as evidence in the COVID-19 pandemic (Anderson, 2021; Evans, 2022; Rhodes & Lancaster, 2020, 2022). Yet little attention has been paid to the practices of translation through which modelling is put to use as evidence and of the productive capacity of "scientific consensus" itself.

An evidence-making event

On 31 October 2020, a second national lockdown in the UK was announced via a televised briefing. This second lockdown came into force on 5 November and lasted 4 weeks. Following calls, including among scientists advising the Government, for more stringent infection control measures throughout September and October, the policy decision was presented by the Prime Minister as a means to prevent 'medical and moral disaster' for the National Health Service 'beyond the raw loss of life' (Kmietowicz, 2020). Projections of daily deaths from COVID-19 derived from mathematical models were presented in support of the policy decision. These projections became contested. In the analysis which follows, we reflect on this moment of controversy and its hinterlands through the qualitative interview accounts of modelling scientists involved. We do this to trace how models and consensus are afforded agency to perform as evidence in policy. We thus pay attention to modelling evidence as *performed* and not merely *translated*.

From evidence translated to evidence performed

Evidence-based approaches in policy idealise a mode of governance which presents decisions as made and justified on the basis of reliable evidence. Evidence is usually imagined here as independent, anterior and definite (Pawson, 2006; Smith, 2013). Yet presumptions of evidence translating in policy as if stable and free of context quickly become troubled in practices (Cairney, 2017; Colebatch, 2009). The metaphor of "translation" is itself critiqued as fundamentally delimiting (Greenhalgh & Wieringa, 2011). Messy translation draws attention to three linked problematics of evidence-based policy: first, that policy problems pre-exist their articulation in policy responses; second, that independent and objective evidence can be reliably generated to shape policy responses; and third, that evidence can be bridged smoothly into policy action (Lancaster & Rhodes, 2020). Rather than questioning these onto-epistemological assumptions, efforts to navigate the mess of translation generally seek to "bridge" the presumed divides of science and policy, by adapting evidence in relation to its implementation contexts and by proposing technical solutions to navigating the "real world" challenges of translation (Oliver & Cairney, 2019). Here, the focus is on how to make evidence translate better to perform decisions that can be claimed as "evidence-based" or "evidence-informed" when governing decisions in practice are rarely simply consequential upon evidence (Oliver & Cairney, 2019).

Public health emergencies, such as pandemics, present particular challenges in the performance of evidence-based policy, given the simultaneous entanglement of evidence and intervention in situations of time-compressed and politicised decision-making (Lancaster et al., 2020; Rutter et al., 2020). Mathematical models are often presented as promising a bridge

as a "boundary object" in evidence translation (Star, 2010; Wheeler et al., 2018). But models also embody messy translations in their iterative movements between data and abstraction, prediction and projection, quantification and qualification and knowns and unknowns (Callon & Numiesa, 2005; Christley et al., 2013; Sismondo, 1999). Rather than imagining evidence being brought in to policy or seeing evidence and policy as separate domains to be bridged through translation, there is a need to attend to evidence 'as part of the "stuff happening" of policy making' (Greenhalgh & Russell, 2006, p. 34).

An alternative approach, therefore, is to treat evidence as performed. Here, frameworks of evidence-based intervention are envisaged as networks of ideas, technologies and people that are sustained and extended through their continued performance in discursive and material practices (Latour, 2005; Law, 2004). Rather than presumed anterior, stable and independent, evidence and intervention entangle and "come to be", in their events of implementation (Rhodes & Lancaster, 2019). Evidence itself becomes a *fluid thing* performed *in its situation*, rather than fixed a priori as an object to be taken up and "used" (Lancaster, 2016; Law & Singleton, 2005). Here, practices of "evidence-based" policy become visible as matters of situated performance in actor networks, networks that could be made otherwise (Latour, 2005; Law, 2004). Models—and how they are made as evidence in policy—are neither separate nor alone but always entangled in, and afforded agency by, their actor network and implementation event (Callon & Numiesa, 2005; Latour, 2005; Rhodes & Lancaster, 2021, 2022).

We can therefore treat evidence-based policy as a performance that locates to a hinterland of practices, which while messy and problematic, appear too difficult or costly to undo (Law, 2004). By hinterland of practices, we refer to the logics, claims, discourses, methods and other inscriptions that standardise, routinise and reproduce "evidence-based policy", even in the face of mess and uncertainty (Law, 2004, 2009). Consensus is one device in modelling knowledge that ostensibly works to smooth the "mess" of translation for policy (Beatty & Moore, 2010; Horst & Irwin, 2010; Latour, 1987). Scientific consensus performs as a benchmark for scientific knowledge, for constituting "*the* evidence", upon which a policy can be legitimately grounded (Jasanoff, 2010). Consensus can present as a united front of joint agreement (Beatty & Moore, 2010; Melo-Martín & Intermann, 2018), a coordination of evidence assembled into a singular space (Law, 2004; Mol, 2002), that appears to mask or tame the un-ease of uncertainty (Horst & Irwin, 2010; Latour, 1987). Yet, as we shall see, consensus practices are not neutral in their effects and neither do they fundamentally alter or resolve the dis-ease which materialises evidence-based policy as an "uncomfortable science".

CASE STUDY METHODS

This analysis draws on qualitative research investigating how mathematical models generate evidence in UK's COVID-19 policy. We undertook depth interviews with 29 mathematical modellers and other scientists engaged in the UK COVID-19 response. All interviews were undertaken by TR, remotely via Teams, between May 2021 and December 2021. Interviews generally lasted between 75 and 90 min and adopted a conversational approach to co-produce an account. All interviews were audio-recorded with consent and transcribed verbatim. Key themes included working as a scientist in a pandemic, generating evidence in relation to policy, communicating modelling evidence, key events in models and modelling, deliberation and consensus. We sampled a diversity of mathematical modellers and modelling teams within and beyond UK Government expert bodies. Roughly half of our sample (14) were engaged in SPI-M. Our analysis enables low inference description through

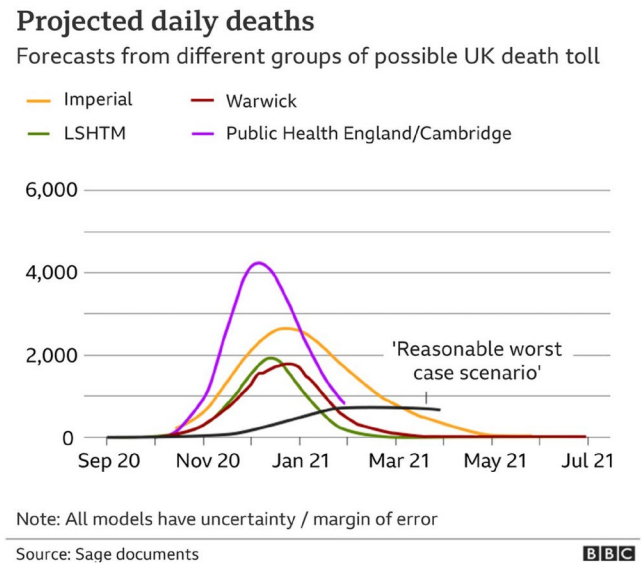

FIGURE 1 Projected Covid-19 deaths, BBC, 31 October 2020

close engagement with interview accounts (Seale, 1999) but to protect against deductive disclosure; we do not summarise participant biographical details. The study received ethics approval from the London School of Hygiene and Tropical Medicine Observational Ethics Committee.

We take as our point of departure for this analysis the Government public briefing of 31 October 2020 used to announce a second national lockdown in the UK in response to the COVID-19 pandemic. Our analysis begins by introducing a graph and linked projections generated from mathematical models that were televised nationally (Figure 1). We reflect on the projections presented to explore, through the interview accounts of mathematical modellers, the event and its evidence-making hinterland. We first engage with modellers' accounts of "messy translation" in relation to the projections presented, before situating accounts of models and consensus in relation to the policy decision to lockdown. This leads us to reflect on how the scientific consensus generated through models deliberated upon at SPI-M performs in policy in the face of uncertainty.

Our analysis is not oriented to representing the accuracy of 'truth claims' but instead investigates their performance. This means that we are primarily interested in the objects, materials and ideas that come into being through the eventuation of the narration, rather than accentuating interviews as a means to capturing an outside reality (Bacchi & Goodwin, 2016). Accordingly, we make no inference regarding the causality of policy decisions and events, but follow how models, evidence, policy and their relations are performed in the "in here" of accounts (Law, 2004). We acknowledge that we, and our methods, are co-producers in this story. We offer our analysis as a mode of "evidence-making intervention" in deliberations about models in policy (Rhodes & Lancaster, 2019, 2022).

A GRAPH AND AN OUTLIER

Figure 1 is a graph of 'projected daily deaths' from COVID-19 in the UK from four modelling groups who participate in the SPI-M committee that generates evidence for use by the UK Government. The graph appeared on the BBC national news, on Saturday 31 October 2020, as

evidence in support of the Prime Minister's announcement for a second national lockdown. The outlier, the worst-case scenario among four estimates, projected a December peak of 4000 deaths a day based on an assumption of 1000 a day by early November, nearly twice as many projected in the second worst-case scenario. These 'death scenarios' were reported by some as 'apocalyptic' (Blanchard & Leatham, 2020), leading to the 'nuclear option' (Kuenssberg, 2020) of an unwanted lockdown because there was 'no alternative' (Kmietowicz, 2020). The outlying projections became publicly questioned as 'overestimates', reported to be 'four' or 'five' times 'too high' (Howdon & Heneghan, 2020; Mahase, 2020). Some media depicted the Government committees as 'misleading the public' by 'cherry-picking the scariest data' (Blanchard & Leatham, 2020; Howdon & Heneghan, 2020). The graph televised by the BBC was based on scenarios presented at SPI-M some 3 weeks before, with the upper projections of daily deaths since having been revised downwards.

MESSY TRANSLATIONS

According to one scientist advising the Government, 'The most important thing in an emergency is to have clear concise messaging and communications to avoid confusion and avoid misunderstanding at all costs'. The event of 31 October, by this account, constituted a failure in evidence translation. Indeed, modellers interviewed said 'it was terrible communication', a 'definite miscommunication of science at a rather critical point', 'a textbook example of how science communication should not be done'. Some cast the event as an 'abuse' of models residing in 'politics'. In this interpretation, the event signalled 'a failure of policy, and potentially then, a failure of the advice, and by extension, a failure of the evidence'. This is an account which enacts evidence-based policy as failing to work as it should.

During interviews, modellers reflected on this event and the projections presented (Figure 1). First, it was emphasised that these were 'not projections in terms of what we think would happen' but 'a reasonable worst-case of what might happen'. The estimates had been produced specifically to 'illustrate a new variant' and 'how bad it could be', but became 'sold outside scientific circles as a forecast'. The original graph presented at SPI-M stated clearly, in capitals, that 'THESE ARE SCENARIOS, NOT PREDICTIONS OR FORECASTS', unlike the version the BBC televised. Second, these were not current estimates but 'leaked from a month before'. They were, we are told, an 'inflection' of their situation at the time 'when we weren't sure'. Third, the media event drew particular attention to the worst-case outlier, despite SPI-M and SAGE having 'signed-off' an updated set of 'more statistically rigorous' projections as 'consensus' for Government to use. These accounts do not trouble "consensus" so much as its *translation*.

How then, in the accounts of those involved, might a Government press briefing come to use leaked out-dated worst-case estimates rather than the current "consensus" produced by its committee of scientific advisors? 'Politics' enters as one immediate answer: 'They decided they would make more of a case'; 'That was done for purely political reasons' and a 'political calculation'. For some scientists, the outlier estimate gathering attention was possibly 'rubbish', yet in policy performed well: it 'looked worse'. The veracity of the projection is not the matter of concern here: 'The numbers are a side-effect'. Rather, projections are made to work, in the moment, because of their narrative fit, to perform a pandemic that is "worse" and "big" and that warrants unprecedented policy change (See also, Rhodes & Lancaster, 2022). This is the idea that: 'When politicians want to do something drastic, they use projections of very bad events'.

Here then, 'some lines on a graph' are put to use as evidence as a 'game changer'. Projections are governing in this moment not as precise calculations but as narrative qualifications which "come to be" inside a policy event:

> There was nothing inherently wrong with the projections but the story to which they were being used to tell wasn't the story for which they had necessarily been produced. (14)

> The Government was in a tailspin because the NHS was falling over. I slightly think that the modelling at that point was pretty irrelevant, and that a full-scale lockdown was going to happen just for the purposes of the NHS more than anything else. (6)

Yet, according to modellers, performing the policy narrative in this way was not needed. For modellers advising Government, there was a 'consensus view' on the epidemic 'out there' that was sufficiently evidenced as bad enough: 'They really didn't have to sex it up in any way'. We are told that discounting the outlier to concentrate on the other three projections in Figure 1 'in actual fact' is close 'to what happened'. For modellers, the evidence *itself* is presumed to circulate independently of the policy event and was telling a sufficiently grim story of worsening pandemic: 'I mean, our projections were [really] bad anyway, you know, there is going to be a shitload of deaths unless you take some action'; 'You're still talking about tens of thousands of deaths, it's still loads of deaths coming in the next few weeks'.

Critically, such messy translation, in the account of scientists, is not good for science or government: 'All of us were very annoyed because it reflected badly on us'. The event was felt to reproduce circulating narratives of doubt, of models as 'false' and 'wrong', thereby 'undermining science'. With divergence presented 'without explanation' in its new situation, the graph 'appeared to show that the models were just basically predicting wildly different outcomes, and therefore, weren't working'. Models were being put to use as evidence in uncomfortable ways, troubling the performance of science as if careful, measured, ordered and capable of consensus. Here, the same projection, the same graph, becomes fluid and multiple, transforming in relation to its implementation events, some 3 weeks apart. This is evidence transformed, an ontological movement in translation felt too far to be comfortable:

> The decision was presented as being, "look, this is what the modellers are telling us, everyone is going to die, so we have to lockdown, even though I really, really don't want to". Making us to be the bad guys and the shroud wavers and the rest of it. That was not a good point. [...] That was not SPI-M consensus, that had not gone to SAGE. That was not a considered "what's your question, and this is the evidence we have for the answer". This was the preliminary draw of some lines on a graph. (2)

SITUATING THE WORK OF MODELS AND CONSENSUS

Let us explore further how modellers account for the event of 31 October as a problem of evidence translation. Our starting point here is the invocation of the epidemiological model as a promise of evidence-based intervention in shaping policies of preparedness through a process which is

idealised as an iterative one, where 'you adapt your control policy as you resolve uncertainty as time goes on'. One modeller imagined this as a 'virtuous circle':

> You develop a model, you get some answers, you take that back to the policy teams, and they tell you why you're wrong, and you go back round the loop ad infinitum or until someone tells you to stop.
>
> (13)

The emergency situation is presented as a force of disruption to this evidence-based process, because in conditions of absented or compressed time models can instead become 'fire and forget missiles':

> Now, that conceptual framework is completely broken in a time space when you've got a week to give an answer. [...] You give an answer and a decision is made. And they don't come back, and you don't go round that loop and look for better data or whatever. It's sort of a fire and forget missile, rather than this virtuous loop.
>
> (13)

The emergency situation, and absented time specifically, is here presented as an account of the ideal state of evidence-based policy-making 'breaking-down', with evidence not 'breaking-through' as it should. Accordingly, this means that there is work to be done to navigate the mess of evidence translation to make evidence "work better". For instance, in the months before the press briefing of 31 October, modellers had reached out to the Government to generate a new RWCS to inform planning. Without commission from the Cabinet Office, SPI-M had not produced a new RWCS for the Government since June. This meant, by September, that 'actual cases' had increased above those of the RWCS that the Government was using to shape the policy: 'We were above it, so the real life was above what they were planning to, which isn't the way you should be running things'. The SPI-M consensus statement of 23 September warns, in bold, that 'the epidemic is close to breaching the agreed RWCS', and by 7 October declares, again in bold, that 'we are breaching the number of infections and hospital admissions in the RWCS' (SPI-M-O, 2020a, b). Here, a modeller captures the sense of growing concern with policy detaching from the presumed actuality of case projections:

> I remember July we started being, like, well cases are up-ticking, and SPI-M started hammering at it. And then August it was unequivocally high. And then September, we were like, well, this is the alarm now, at this point something had to be done.
>
> (18)

To energise dialogue, modellers had advised the Cabinet Office, in late September, of their new reasonable worst-case projections. Again, this was a proactive move. Rather than wait, the feeling at SPI-M was that 'we are just going to have to get on with it'. These are the projections that made their unexpected public appearance on 31 October. Yet the Cabinet Office 'never got back'. Through September and October, modellers suggest they were 'unaware of what the strategy was'. The optimum time to act on the basis of emergent evidence was escaping: 'Once you have decisive evidence about something like waves being too high, you don't need models anymore, you are seeing the deaths'.

The account presented here is one of "playing the game" of evidence-based policy-making, while tinkering to make it work better, even as it is breaking down. There is said to be generalised disconnect between models and policy through this time: 'The science is never actually used. It was ignored'; 'They [Government] never listened to the science'. In the months before the second national lockdown, there was said to be 'overwhelming consensus' among scientists of 'exponential growth leading to the most terrible second wave'. There was 'no diversity of opinion', and there 'was never any doubt'. SPI-M had been modelling, in September, the potential effects of "circuit breaker" interventions. But the consensus of the need to act was apparently not translating well: 'We were ignored going into the second wave':

> We published in last September that if we don't have a circuit-breaker, if we don't introduce some form of restrictions, then there will be a huge number of cases over the winter. And there was a huge number of cases over the winter. So that was basically ignored.
>
> (6)

The failure of "scientific consensus" to translate into precautionary action is presented as defying evidence-based expectation, even while accepting the contingency of politicised decisions: 'I just don't know why there was no proper plan'; 'Why do you want to put yourself in a situation where you have thousands of people dying? Mitigate the situation!'. Consensus, for some advising the Government, could not have been any more evident:

> It was clear where we were going and that was well communicated. That was well communicated to the Ministers. That was well communicated via unofficial channels. That was communicated via academic freedom, via the media, to say this is where it was going. And all of that didn't make any difference at all to the political decision-making. [...] It's very, very hard to see what happened in Cabinet, but a more clear example of the modelling saying, "Well, this is going to happen and how bad it will be", you couldn't find. And yet, out the other side, decisions were made as they were.
>
> (14)

Scientists here, are neither separate to their policy situation nor unaffected by playing the game of making evidence work. They entangle as affected users of evidence, not mere knowledge producers (Lahsen, 2005). We see this in their proactive efforts, within the generous constraints of Government policy advice infrastructures, to advocate for "evidence" and its "use". The claim being performed is that this is evidence-based policy not as we know it, nor as it should be, even while this messy situation is not a complete surprise. Models invoked as 'fire and forget missiles' in a 'broken' evidence-based system is a narrative of 'failure' which tames the limits of expectation in consensus being capacitated to work. As one modeller comments: 'For most of 2020, we would produce something and it would kind of disappear into the ether, and we'd have no idea what kind of impact it had'. The problem of translation failure is presented as 'working in a vacuum' and 'not knowing what was going to happen'. There is not enough policy attachment in the models:

> From a scientific, from an advisory point of view, it makes providing the evidence impossible, and particularly the modelling evidence. We don't know what we're trying to model.
>
> (2)

> When coming up with a reasonable worst-case scenario you really need to have some input from the policy guys, because it's really got nothing to do with the epidemiology and everything to do with what they're going to do, what policy.
>
> (4)

In this account of policy detachment, lockdown decisions can dissent from, rather than follow, scientific consensus:

> You have a Prime Minister who unequivocally just did not want to do any restrictions on this, and any bit of evidence that could go to the contrary they chose. The scientific consensus was there. SAGE had consensus and that is what I think should have been listened to. They're meant to be your advisor as a Government.
>
> (20)

Crucially, scientists advising the Government have to *live with the mess*. They are embodied in the situation, all the while they are playing the game, and even while performing science and policy as if 'independent'. Here is an account reflecting on the projections used to justify the second lockdown (Figure 1) which emphasises "letting go" of evidence to accept a "discounted" version of science:

> What I didn't want to do is jump all over the media, and say, "No, no, no, take those figures down. They are horrible, they are wrong". And they are horrible, and they are not right, but the progression of the public argument we didn't want to muddy. [...] It didn't matter whether those numbers were big or bigger, the disaster was on us. Correcting them [the projections] was the right thing to do. [...] But it wasn't worth having it out in public because the right decision was made for the wrong reasons.
>
> (14)

THE UNCOMFORTABLE SCIENCE OF TROUBLED CONSENSUS

In this account of bumpy science-policy interface, evidence and consensus are not stable things that are made in models *for* policy, but *fluid things*, that "come to be" *in* policy (Lancaster & Rhodes, 2022). The fluid attachments of models in policy generate troubles for evidence-based science and its performance. We can consider the troubles of consensus as an "uncomfortable science". Scientific consensus acts as a key bridging mechanism to transform models as evidence in COVID-19 policy. SPI-M gathers models and their projections together to reassemble these as evidence, expressed as a singular 'consensus view' for onwards travel, via SAGE and the Cabinet Office. In the accounts of modellers, consensus-making performs deliberation across 'different'

and 'independent' models, with 'greater confidence' reached when models 'give similar answers' (SPI-M-O, 2020c). This performs 'scrutiny':

> One of the key things about the consensus idea is that ideally we have three different groups producing independent work. They bring it to the Committee, we go through it, essentially as a rapid review. We have a motto on SPI-M, which is "Tell me why I am wrong". That enables the Secretariat to then come up with what the Committee thinks, the consensus view.
>
> (2)

Whilst enacting a joint agreement, deliberation across models engages directly with the uncertainty and difference of projections, which is especially important when learning from outliers (as per Figure 1):

> We need to actually properly scrutinise this, and actually say, "Well, okay, this one is an outlier, but do we have evidence that suggests that this outlier might actually, be, in inverted commas, the "correct" one?". [...] It's a bit nerve-wracking, especially at the beginning, because you submit your predictions or scenarios, and you don't particularly want to be the modeller, the model, who's got a huge number when everyone else agrees. But, invariably there are some models that are outliers, and then there's a discussion that ensues about what are the assumptions that were made, trying to really disentangle why one model looks different from others.
>
> (10)

The consensus-making process, and the consensus statements published by SPI-M, are described as forms of 'protection' against individual or outlying projections taking flight in policy without evidence-based scrutiny. This is said to be a lesson of the lockdowns, where corroboration 'ensures that nothing is only supported by one piece of evidence', with consensus enabling a 'layer of security' before 'models get escalated up the chain'. Consensus statements are 'not always unanimous', and crucially, 'include the uncertainty'. Here, they work to manage against uncertainty getting 'omitted' or 'pushed under the carpet', as projections are 'communicated onwards', often 'recreated' in policy communications as 'a bunch of single lines' (as in the event of 31 October; see Figure 1). The SPI-M documentation on the projections that performed evidence for lockdown at the press briefing of 31 October warned, in bold, that 'these indicative scenarios are not precise predictions of what is going to happen' and 'show preparatory work' (SPI-M-O, 2020c). They also caveat that long-term scenarios generate 'wide uncertainty', get 'more unreliable the further into the future they look' and offer 'qualitative information' to aid government planning.

The performance of consensus then, locates to a hinterland of evidence-based assumption. Consensus performs evidence for policy in a transportable form through a process resembling the "standardised packages" of science idealising evidence as independent, derived through careful deduction and ultimately, singular in its presentation (Law, 2004). Consensus-making also performs "being at ease", affording "protection", in uncertainty (Beatty & Moore, 2010; Horst & Irwin, 2010; Oreskes, 2004). While serving to simplify the communication of complexity and

difference that circulates inside what is presented as "the evidence", performing consensus in the name and game of evidence-based policy-making is not without trouble.

Troubles constituting consensus

Consensus trouble was largely voiced in epistemological terms, as a problem of performing "the evidence" that is most useful or correct, as well as certain enough. When the ideal state and 'virtuous loop' of evidence-based iteration is challenged by absented time, short-cuts in the doing of science become inevitable, with 'mistakes' possible. This is one aspect affecting an uncomfortable science:

> We try to make it as error-proof as possible, but it's difficult. It's especially difficult to do that in a short timeframe obviously. So there is always the worry that something is wrong, or that you've done something wrong. [...] Often times we'll have wound up doing, getting the simulation results quite late, and it doesn't feel like I've had a whole lot of time to interpret them and understand whether you feel really confident. (28)

In an emergency situation, science as a tool in policy becomes presented as an uncertain "leap of faith", where the invitation is to 'act on the faith that what the model is predicting is going to become reality'. With outcomes across multiple models unknown, and independent or observed data comparisons not possible, expert judgement becomes the ultimate arbiter (Thompson & Smith, 2019). Consensus-making is a process of negotiated judgement for *living with* irreducible (or not yet reducible) uncertainty (Christley et al., 2013).

Modellers entangle with, rather than step back from, the uncertainty of their models (Christley et al., 2013). As discussed in interviews: they navigate the "leap of faith" of evidence translation by accentuating the empirical limits of projections; they emphasise the policy value of treating models as "qualifications" rather than as precise "calculations"; they caution against uncertainty getting lost as models transform in their use as evidence; and they emphasise generating policy advice through expert judgement over models that are too uncertain or full of ungrounded assumptions. The consensus performed by SPI-M then, is viewed 'as strongly scientifically-based *as possible*', given the irreducible uncertainties of the emergency situation. In this view, consensus is as '*close as you're going to get* to mitigating the challenges of 'inherently subjective' judgements. Uncertain models do not speak robustly or comfortably for themselves but create incredible interpretive latitude. They constitute a site of qualitative deliberation, translated as the 'consensus *view*'. Clearly, irreducible uncertainty cannot be risk-managed away through calculus, and neither, more fundamentally, is it simply an effect of the emergency situation. Modellers playing the game of evidence-based policy in uncertainty find themselves grappling with 'the true nature of science', which at best is viewed as only ever 'partially correct':

> You are cast in this rapid situation where you have to produce work on a very, very, very short timescale from very, very limited data. And that is science. But people will hold that science to the bar of what they expect from science. And this disconnect, or this lack of knowledge, or this lack of communication, is so important, because you aren't expecting this coronavirus work to be correct. [...] When you enter the pandemic, you are confronted with the true nature of science, as the slow-moving,

> subjective, partially correct way of understanding the world: "These guys have been working in epidemiology their whole lives, and the best they can give us is some guesses, many of which turn out to be wrong".
>
> (18)

Here, in this presentation of science as 'slow-moving' and 'partially correct', we can see modelling presented as engaging with uncertainty but not incertitude; a practice which works within, and which reproduces, the hinterlands of expectation performed by evidence-based science without actualising such expectation (Law, 2004; Leach et al., 2021). This so-called 'partially correct' game of evidence-based science, even if felt not to be 'true' to the situation, continues nonetheless. Models are quintessentially capable of living with incertitude rather than in need of narrating this away (Savransky, 2017), yet in policy reside, uncomfortably, within a network of expectation that performs empirical fit and correspondence to actuality as the primary markers of success. The troubles of navigating between the evidence-based ideals of making 'definite statements that correspond to definite realities' and the 'uncertainty, fuzziness and undecidability' of actual science (Latour & Woolgar, 1986; Law, 2004, p. 31) become especially visible in times of emergency.

Troubles using consensus

Moreover, as we have heard, scientific consensus might not translate in policy as idealised. For all the performance of this as a failure of evidence-based policy not working as it should (see above), this is not a surprise (Cairney, 2017, 2022). Modellers also see the unpredictability of policy decision-making as beyond mitigation:

> The decisions get made all over the place. We imagine this as an efficient decision making machine. But the closer you get to Government the more you realise that the actual decision happens all over the place, all the time, especially in a crisis.
>
> (7)

What is significant, therefore, is that the performance of evidence-based policy is reproduced nonetheless. When modellers speak of scientific consensus having been 'ignored' or 'miscommunicated', of 'not being listened to', of advice having 'failed', of decisions 'pinned' unreasonably to evidence and of the 'virtuous loop' in science-policy translation having 'completely broken', the focus becomes how to correct the 'misfires' of evidence travelling into policy. Here, the problem is performed as not 'the evidence' but the infrastructure smoothing its travel. Indeed, SPI-M was said to be 'a beautifully constructed mechanism, that has only one flaw in it, and that's the people who make policy'.

Correcting the model-policy alignment is generally cast as an effort to bring modelling closer inside the policy and the policy closer inside modelling. Part of this process is attuning 'the ask' of policy to models. This is an effort to preserve and do 'good science'; to model questions that are 'actually addressable with modelling' and to 'put a stop' to models that are felt unneeded or that cannot be done well, as they are 'taken into the policy domain to hang a decision around'. Models are said to detach from policy because 'policy has a tendency to ask questions which make no

sense, and they also have a tendency to just grab the nearest number and not understand its context'. This is learning generated by the lockdowns; that 'policy has to be in the model'; and that modelling in the face of uncertainty has to be protected from becoming a 'veneer for science', especially when there is evidence enough to make decisions (like "facts" such as "deaths"). Models and policy are said to have become closer through the pandemic, with their connections less bumpy: 'We know each other a lot more'; 'Now we are much more familiar'. Since the second lockdown, some say there has been 'threshold change'. There is now a 'roadmap', 'dialogue' and 'better understanding', which helps to better 'tailor the advice'.

We can see in these accounts that the promise of a smoother translation being realised performs models and policy as *getting closer yet staying apart*. Performing "independence" remains fundamental to preserving the authority of models as "evidence-based". Modellers advising the Government accordingly say 'our job is to provide evidence for policy', and that 'policy decisions are what Government does'. Science and policy are reproduced as if separate to the other, with models and consensus doing the bridging work. Efforts to clean up the problem of messy translation thus sustain an expectation that evidence pre-exists its implementation event; as ready-made for transfer. This affects an uncomfortable science because scientists creating evidence for policy navigate between the "sacred" and the "profane" (Colebatch, 2009); between acting as independent experts engendering trust, authority and legitimacy in the policy-making process and the day-to-day uneasy actualities of 'fire and missile' evidence-making where policy decisions can dissent from scientific consensus.

Affects and political concerns

The messy evidence translations of lockdowns, with policy seemingly detaching from consensus in dangerous ways, generates affective disturbance and conflicts of political concern. The technical, political and personal entangle, even as efforts are made to present them as apart when acting the role of an independent and objective knowledge producer (Evans, 2022; Lahsen, 2005). Take, for instance, this account:

> You've done all this stuff and the Government ignored it, and you go into a second lockdown because you refuse to do a two-week circuit breaker. My uncle died, really sad, 58, with no pre-existing conditions. He wouldn't have died if he was in New Zealand or Denmark. And for what benefit? The UK economy was tanked anyway, whether or not we did it, or not, so at what point am I helping out these Government decisions, rather than resigning from SPI-M, out of protest saying "I am not going to do it", or the other scientists resigning "We're not going to work for a Government like this". [...] But then, of course, equally you have a moral obligation to present the truth, despite an unequivocally shoddy Government.
>
> (18)

Playing the game of science as if independent of its social situation is especially uneasy when models and evidence are made so public and political. While presented as if neutral and value-free, scientific consensus is also brought to life in a 'blame game' in which models are said to have been depicted publicly as if 'running the show' with policy 'following the science'. Indeed,

the political performativity of evidence translation is a lesson of lockdown decisions for modellers, with consensus put to use in policy events to 'give the idea that the policy is just following the modelling, you know, that we're making decisions'. With evidence afforded an unpredictable life that is public and political, there is a fear of 'backlash', of getting 'burnt' and of 'being thrown under the bus', among some who advise the Government. The uncomfortable science of modelling for policy hits home the situated realities of evidencing as never neutral nor independent, or free of affect and politics, despite this game being played. As one modeller reflects:

> This is awful really. I'm almost becoming a politician. When they [media] are pushing me for numbers, or saying for instance "Do you think we should unlock?", I'm deliberately not answering those questions because I know that that's going to lead to a headline. You end up like getting sore sitting on the fence because you are really worried that they will cherry-pick and quote that as headline. So, I'm just not giving them anything. I think it's really sad because actually we should be able to go in the media and say actually what we think.
>
> (11)

CONCLUSION

We have reflected on how mathematical models are brought to life in policy. Rather than seeing evidence as bridging the gap between science and policy, as if these were independent and apart, we have traced how evidence becomes in its implementation (Rhodes & Lancaster, 2019). In this view, evidence does not pre-exist its implementation, as if "out there", ready-made for translation, but is enacted in the "here-and-now" of situated translation practices. Models and projections thus "come to be" as *evidence of their situation* (Savransky & Rosengarten, 2016). This means that models in policy are not things merely *translated*, but are things *performed*, and *transformed*, in their implementation events. This thinking is helpful because it sticks with evidence as *actually eventuated in practices* rather than as idealised in evidence-based intervention imaginaries.

Holding on: Acting as if

Situating evidence as a becoming of its implementation event hits home the uncomfortable realities of holding on to the idea of science and policy *as if* they were ever anything other than performances made in practices, and *as if* they unfold in an ordered linear, technical and consequential relationship. An 'uncomfortable science' is affected by living the troubles, and playing the game, of idealised evidence-based expectation. This trouble can be traced to a hinterland of expectation of evidence as singular, stable, independent and definite. While these troubles become especially visible in situations of an emergency, the roots of uncomfortable science run deeper. Our case study has focussed on mathematical models in the COVID-19 lockdown policy, but this is as much a story about the failures and hinterlands of the project of evidence-based policy more generally. Importantly, we have noted how efforts to risk-manage evidence-based troubles, including through practices of scientific consensus, can extend the dis-ease of uncomfortable science rather than clean it up or close it down. Consensus-making is itself a troubled device of translation; it seeks to navigate yet reproduces the mess of evidence-based policy translation.

Performing models and policy *as if* independent from the other, evidence *as if* stable and singular, and policy *as if* consequential, does not make evidence translation smoother or easier but *extends dis-ease*. Uncomfortable science invites thinking altogether differently about evidence and its performance:

> If we feel uncomfortable without clear, definite and singular accounts of clear, definite and singular structures, then that is how it is. However, if we are able and willing to tolerate the uncertainties and the specificities of enactment, flux and resonance, then we find that we are confronted with a quite different set of important puzzles about the nature of the real and how to intervene in it.
>
> (Law, 2004, p. 141)

Letting go: Another translation is possible

Following Law's provocation, our analysis is suggesting that another mode of evidence translation is possible; one that works with a 'different set of important puzzles about the nature of the real and how to intervene in it'. This is more than managing the mess of evidence-based policy assumption, more than deploying techniques and strategies to smooth such troubles away by learning to play the game (Oliver & Cairney, 2019). We need to move beyond noticing evidence in policy as messy and performative (Boswell, 1998; Cairney, 2017, 2022), to questioning, more fundamentally, the onto-epistemological logics, which reproduce evidence-based expectation. We can ask then, what potential is lost by treating the hinterland of evidence-based expectation as too difficult or costly to disinvest from, despite being broken in practice? Indeed, what might be gained by alternatively treating models and evidence as fluid objects that are made, and made successful, in the moment of their implementations?

This is partly a proposal which invites letting go of the 'phantastic mathematical objects and achievables of model-land' (Thompson & Smith, 2019, p. 13), to incorporate multiple other forms of knowledge when deliberating on pandemic futures (Christley et al., 2013; Leach et al., 2021; Saltelli et al., 2020). It puts deliberation across multiple forms of evidence, including beyond models, more transparently at the centre of evidence-making negotiations, not to work through or beyond contingency but to more comfortably live and engage with it (Rhodes & Lancaster, 2020). But more fundamentally, it invites thinking beyond the promise of certitude realised through progressive empirical fit in favour of living with the incertitude of unfolding situations in-the-now. This is an ontological move, a model of action which shifts from an "evidence-based" to an "evidence-making" approach, for it accentuates models, mathematical and otherwise, as modes of *qualification* and *speculation* on futures that could be made possible and not merely futures that appear probable (Savransky, 2017).

What does a mode of speculation afford that an orientation to empirically grounded prediction does not? Speculation is different to prediction, and sits uncomfortably alongside the use of models predicated as tools of evidence-based forecast. Because speculation invites deliberation on alternatives and scenarios, it shifts from a primary focus on what is likely to be or what might be (probabilities) to considering what could be (possibilities). Probabilities close down possibilities, whereas speculations open them up. Success here resides in the effects that such interventions make in the here and now rather than in their capacity to predict or pre-empt with more or less accuracy (Halewood, 2017). Because speculation is a mode of intervention which deliberates more openly in the making of policy futures, it can question the taken-for-granted and familiar to imagine more radical departures beyond extensions of the present (Savransky, 2017). This means that speculations

are not to be viewed as a substitute for predictions and forecasts of what might happen when these are not empirically doable, but rather, constitute a different mode of evidencing and intervening.

The question then, is how to create space for multiple models to act as sites of deliberation in relation to policy alternatives, including when a situation demands an immediate response. The COVID-19 pandemic has indeed energised the modelling of new and unprecedented policy scenarios. This indicates the speculative potential of models not only to invite policy change in keeping with the immediate and familiar—as reproduced in the evidence-based logics of epidemiological responses which promise certainty through progressive empirical fit—but also to experiment with more radical jumps and alternatives. First, holding on to the promise of certainty in an evidence-based policy approach which purports to "follow the science" might slow down or delay policy action in situations when certitude or consensus is beyond reach (Evans, 2022). Second, some letting go of an orientation towards evidence-based prediction, by affording models of possibility over probability, as befits a situation of incertitude, also encourages projection in relation to multiple matters of social and political concern and not only an emphasis on epidemiological fact (Latour, 2004; Law, 2009). Here, for instance, models might break from their relatively narrow bioepidemiological logics—projecting how virus reproduces in susceptible populations—to incorporate the social and economic dynamics of different pandemic and lockdown futures (Anderson, 2021). Third, this implies a broadening of expertise beyond the epidemiological model, and beyond the performance of "consensus", with models acting instead as modes of deliberative intervention in the generation of possible futures, including as part of scientific advisory infrastructures guiding policy.

Models and policy work more speculatively and experimentally in emergencies than the probabilistic and empiricist logics of evidence-based policy imagine. Rather than holding on to a "sacred" account of evidence-based policy troubled in practice (Colebatch, 2009), engaging with the situated actualities of evidencing in policy makes for a science less uncomfortable with itself as well as opens up speculative potential. Living with the realities of incertitude and fluidity is never going to be easy, but this is our situation. Let us not pretend otherwise.

AUTHOR CONTRIBUTIONS

Tim Rhodes: Conceptualisation (Equal); Data curation (Lead); Formal analysis (Lead); Writing—original draft (Lead); Writing—review & editing (Lead). **Kari Lancaster**: Conceptualisation (Equal); Data curation (Supporting); Formal analysis (Supporting); Writing—original draft (Supporting); Writing—review & editing (Supporting).

ACKNOWLEDGEMENTS

We thank our participants who generously gave their time to work in dialogue with us. This project is part supported by an Australian Research Council Discovery Project grant (DP210101604) and part supported through the London School of Hygiene and Tropical Medicine. We are grateful for support from the UNSW SHARP (Professor Tim Rhodes) and Scientia (Associate Professor Kari Lancaster) schemes.

DATA AVAILABILITY STATEMENT

The qualitative interview data linked to this analysis are not made available in order to protect against deductive disclosure of the participants involved.

ORCID

Tim Rhodes https://orcid.org/0000-0003-2400-9838
Kari Lancaster https://orcid.org/0000-0002-2411-7970

REFERENCES

Adams, V., Murphy, M., & Clarke, A. E. (2009). Anticipation: Technoscience, life, affect, temporality. *Subjectivity*, *28*(1), 246–265. https://doi.org/10.1057/sub.2009.18

Anderson, W. (2021). The model crisis, or how to have critical promiscuity in the time of Covid-10. *Social Studies of Science*, *51*(2), 167–188. https://doi.org/10.1177/0306312721996053

Bacchi, C., & Goodwin, S. (2016). *Poststructural policy analysis*. Palgrave.

Beatty, J., & Moore, A. (2010). Should we aim for consensus? *Episteme*, *7*(3), 198–214. https://doi.org/10.3366/epi.2010.0203

Blanchard, S., & Leatham, X. (2020, November 2). *Apocalyptic forecast of 4,000 coronavirus deaths a day could be FIVE TIMES too high and had already been proven wrong when government revealed it at weekend*. Daily Mail.

Boswell, J. (1998). What makes evidence-based policy such a useful myth? *Governance*, *31*(2), 199–214. https://doi.org/10.1111/gove.12285

Brooks-Pollock, E., Danon, L., Jombart, T., & Pellis, L. (2021). Modelling that shaped the early COVID-19 pandemic response in the UK. *Philosophical Transactions Royal Society B*. *376*(1829).

Cabinet Office. (1999). *Modernising government*. The Stationery Office.

Cairney, P. (2017). Evidence-based best practice is more political than it looks: A case study of the 'Scottish Approach'. *Evidence and Policy*, *13*(3), 499–515. https://doi.org/10.1332/174426416x14609261565901

Cairney, P. (2021). The UK Government's COVID-19 policy: What does "guided by the science" mean in practice? *Frontiers in Political Science*, *3*, 624068. https://doi.org/10.3389/fpos.2021.624068

Cairney, P. (2022). The myth of 'evidence-based policymaking' in a decentred state. *Public Policy and Administration*, *37*(1), 46–66. https://doi.org/10.1177/0952076720905016

Callon, M., & Numiesa, F. (2005). Economic markets as calculative and calculated collective devices. *Organization Studies*, *26*(8), 1229–1250. https://doi.org/10.1177/0170840605056393

Christley, R. M., Mort, M., Wynne, B., Wastling, J. M., Heathwaite, A. L., Pickup, R., Austin, Z., & Latham, S. M. (2013). "Wrong, but useful": Negotiating uncertainty in infectious disease modelling. *PLoS One*, *8*(10), e76277. https://doi.org/10.1371/journal.pone.0076277

Colebatch, H. K. (2009). *Policy* (3rd ed.). Open University Press.

Evans, R. (2022). SAGE advice and political decision-making: 'Following the science' in times of epistemic uncertainty. *Social Studies of Science*, *52*(1), 53–78. https://doi.org/10.1177/03063127211062586

Glasser, J. W., Hupert, N., McCauley, M. M., & Hatchett, R. (2011). Modelling and public health emergency responses: Lessons from SARS. *Epidemics*, *3*(1), 32–37. https://doi.org/10.1016/j.epidem.2011.01.001

Greenhalgh, T., & Russell, J. (2006). Reframing evidence synthesis as rhetorical action in the policy making drama. *Healthcare Policy*, *1*(2), 34–42. https://doi.org/10.12927/hcpol.2006.17873

Greenhalgh, T., & Wieringa, S. (2011). Is it time to drop the 'knowledge translation' metaphor? A critical literature review. *Journal of the Royal Society of Medicine*, *104*(12), 501–509. https://doi.org/10.1258/jrsm.2011.110285

Halewood, M. (2017). Situated speculation as a constraint on thought. In A. Wilkie, M. Savransky, & M. Rosengarten (Eds.), *Speculative research* (pp. 52–66). Routledge.

Horst, M., & Irwin, A. (2010). Nations at ease with radical knowledge: On consensus, consensusing and false consensusness. *Social Studies of Science*, *40*(1), 105–126. https://doi.org/10.1177/0306312709341500

Howdon, D., & Heneghan, C. (2020, November 2). *SAGE models overestimation of deaths*. Centre for Evidence-Based Medicine, University of Oxford. https://www.cebm.net/covid-19/the-innacuracoes-in-the-sage-models/

Huppert, A., & Katriel, G. (2013). Mathematical modelling and prediction in infectious disease epidemiology. *Clinical Microbiology and Infection*, *19*(11), 999–1005. https://doi.org/10.1111/1469-0691.12308

Jasanoff, S. (2010). Science and society: Testing time for climate science. *Science*, *328*(5979), 695–696. https://doi.org/10.1126/science.1189420

Kmietowicz, Z. (2020). Covid-19: "There is no alternative", says Johnson, announcing new restrictions for England. *British Medical Journal*, *371*, m4247. https://doi.org/10.1136/bmj.m4247

Kuenssberg, L. (2020, October 31). *Coronavirus: Boris Johnson launches the nuclear option he swore to avoid*. Guardian.

Lahsen, M. (2005). Seductive simulations? Uncertainty distribution around climate models. *Social Studies of Science*, *35*(6), 895–922. https://doi.org/10.1177/0306312705053049

Lancaster, K. (2016). Performing the evidence-based drug policy paradigm. *Contemporary Drug Problems*, *43*(2), 142–153. https://doi.org/10.1177/0091450916633306

Lancaster, K., & Rhodes, T. (2020). What prevents health policy being 'evidence-based'? New ways to think about evidence, policy and interventions in health. *British Medical Bulletin*, *135*(1), 38–49. https://doi.org/10.1093/bmb/ldaa026

Lancaster, K., & Rhodes, T. (2022). The thing we call evidence: Toward a situated ontology of evidence in policy. In S. Ehlers & S. Esselborn (Eds.), *Evidence in action between science and society*. Routledge.

Lancaster, K., Rhodes, T., & Rosengarten, M. (2020). Making evidence and policy in public health emergencies: Lessons from COVID-19 for adaptive evidence-making and intervention. *Evidence & Policy*, *16*(3), 477–490. https://doi.org/10.1332/174426420x15913559981103

Latour, B. (1987). *Science in action*. Harvard University Press.

Latour, B. (2004). Why has critique run out of steam? From matters of fact to matters of concern. *Critical Inquiry*, *30*(2), 225–248. https://doi.org/10.1086/421123

Latour, B. (2005). *Reassembling the social: An introduction to actor network theory*. Oxford University Press.

Latour, B., & Woolgar, S. (1986). *Laboratory life: The construction of scientific facts*. Princeton University Press.

Law, J. (2004). *After method*. Routledge.

Law, J. (2009). Seeing like a survey. *Cultural Sociology*, *3*(2), 239–256. https://doi.org/10.1177/1749975509105533

Law, J., & Singleton, V. (2005). Object lessons. *Organization*, *12*(3), 331–335. https://doi.org/10.1177/1350508405051270

Leach, M., MacGregor, H., Ripoll, S., Scoones, I., & Wilkinson, A. (2021). Rethinking disease preparedness: Incertitude and the politics of knowledge. *Critical Public Health*, *32*(1), 82–96. https://doi.org/10.1080/09581596.2021.1885628

Lipsitch, M., Finelli, L., Heffernan, R. T., Leung, G. M., & Redd, S. C. (2011). Improving the evidence base for decision making during a pandemic: The example of 2009 influenza A/H1N1. *Biosecurity and Bioterrorism*, *9*, 89–115.

Mahase, E. (2020). Cobid-19: England's daily deaths are still "likely to exceed" first wave peak despite projection changes. *British Medical Journal*, *371*, m4321. https://doi.org/10.1136/bmj.m4321

Melo-Martín, I., & Intermann, L. (2018). *The fight against doubt: How to bridge the gap between scientists and the public*. Oxford University Press.

Mol, A. (2002). *The body multiple: Ontology in medical practice*.

Oliver, K., & Cairney, P. (2019). The dos and don'ts of influencing policy: A systematic review of advice to academics. *Palgrave Communications*, *5*(1), 21. https://doi.org/10.1057/s41599-019-0232-y

Oreskes, N. (2004). Beyond the Ivory Tower: The scientific consensus on climate change. *Science*, *306*(5702), 1686. https://doi.org/10.1126/science.1103618

Pawson, R. (2006). *Evidence-based policy*. Sage.

Pawson, R. (2021). The coronavirus response: Boxed in by models. *Evaluation*, *27*(2), 149–167. https://doi.org/10.1177/1356389020968579

Performance and Innovation Unit. (2000). *Adding it up: Improving analysis and modelling in central government*. Performance and Innovation Unit.

Rhodes, T., & Lancaster, K. (2019). Evidence-making interventions in health: A conceptual framing. *Social Science and Medicine*, *238*, 112488. https://doi.org/10.1016/j.socscimed.2019.112488

Rhodes, T., & Lancaster, K. (2020). Mathematical models as public troubles in COVID-19 infection control: Following the numbers. *Health Sociology Review*, *29*(2), 177–194. https://doi.org/10.1080/14461242.2020.1764376

Rhodes, T., & Lancaster, K. (2021). Excitable models: Projections, targets, and the making of futures without disease. *Sociology of Health and Illness*, *43*(4), 859–880. https://doi.org/10.1111/1467-9566.13263

Rhodes, T., & Lancaster, K. (2022). Making pandemics big: On the situational fit of Covid-19 mathematical models. *Social Science and Medicine*, *301*, 114907. https://doi.org/10.1016/j.socscimed.2022.114907

Rutter, H., Wolpert, M., & Greenhalgh, T. (2020). Managing uncertainty in the Covid-19 era. *British Medical Journal*, *370*, m3349. https://doi.org/10.1136/bmj.m3349

Saltelli, A., Bammer, G., Bruno, I., Charters, E., Di Fiore, M., Didier, E., Nelson Espeland, W., Kay, J., Lo Piano, S., Mayo, D., Pielke Jr, R., Portaluri, T., Porter, T. M., Puy, A., Rafols, I., Ravetz, J. R., Reinert, E., Sarewitz, D., Stark, P. B., … Vineis, P. (2020). Five ways to ensure that models serve society: A manifesto. *Nature*, *582*(7813), 482–484. https://doi.org/10.1038/d41586-020-01812-9

Saminmian-Daresh, L. (2016). Practicing uncertainty: Scenario-based preparedness exercises in Israel. *Cultural Anthropology*, *31*(3), 359–386. https://doi.org/10.14506/ca31.3.06

Sasse, T., Haddon, C., & Nice, A. (2020). *Science advice in a crisis*. Institute for Government.

Savransky, M. (2017). The wager of an unfinished present: Notes on speculative pragmatism. In A. Wilkie, M. Savransky, & M. Rosengarten (Eds.), *Speculative research* (pp. 25–38). Routledge.

Savransky, M., & Rosengarten, M. (2016). What is nature capable of? Evidence, ontology and speculative medical humanities. *Medical Humanities*, *42*(3), 166–172. https://doi.org/10.1136/medhum-2015-010858

Seale, C. (1999). *The quality of qualitative research*. Sage.

Sismondo, S. (1999). Models, simulations, and their objects. *Science in Context*, *12*(2), 247–260. https://doi.org/10.1017/s0269889700003409

Smith, K. (2013). *Beyond evidence-based policy in public health*. Palgrave Macmillan.

SPI-M-O. (2020a). *Consensus statement on Covid-19*, September 23 2020.

SPI-M-O. (2020b). *Consensus statement on Covid-19*, October 7 2020.

SPI-M-O. (2020c). *Long term winter scenarios preparatory working analysis*, September 23 2020.

Star, S. L. (2010). This is not a boundary object: Reflections on the origin of a concept. *Science, Technology and Human Values*, *67*, 1530–1540.

Thompson, E. L., & Smith, L. A. (2019). Escape from model-land. *Economics: Open Access*, *13*(2019–2040), 1–15. https://doi.org/10.5018/economics-ejournal.ja.2019-40

Wheeler, K. G., Robinson, C. J., & Bark, R. H. (2018). Modelling to bridge many boundaries. *Regional Environmental Change*, *18*(6), 1607–1619. https://doi.org/10.1007/s10113-018-1304-z

How to cite this article: Rhodes, T., & Lancaster, K. (2022). Uncomfortable science: How mathematical models, and consensus, come to be in public policy. *Sociology of Health & Illness*, 1–20. https://doi.org/10.1111/1467-9566.13535