Epidemics 32 (2020) 100393

Contents lists available at ScienceDirect

Epidemics

journal homepage: www.elsevier.com/locate/epidemics

Influencing public health policy with data-informed mathematical models of infectious diseases: Recent developments and new challenges

Amani Alahmadi[a], Sarah Belet[a,k], Andrew Black[b,k], Deborah Cromer[c], Jennifer A. Flegg[d,*], Thomas House[e,f,**], Pavithra Jayasundara[g], Jonathan M. Keith[a,k], James M. McCaw[d,h,*], Robert Moss[h], Joshua V. Ross[b,k,***], Freya M. Shearer[h], Sai Thein Than Tun[i], James Walker[b], Lisa White[i], Jason M. Whyte[j,k], Ada W.C. Yan[l], Alexander E. Zarebski[m]

[a] School of Mathematics, Faculty of Science, Monash University, Melbourne, Australia
[b] School of Mathematical Sciences, University of Adelaide, Adelaide, Australia
[c] Kirby Institute for Infection and Immunity, UNSW Sydney, Sydney, Australia and School of Mathematics and Statistics, UNSW Sydney, Sydney, Australia
[d] School of Mathematics and Statistics, University of Melbourne, Melbourne, Australia
[e] Department of Mathematics, University of Manchester, Manchester, UK
[f] IBM Research, Hartree Centre, Sci-Tech Daresbury, Warrington, UK
[g] School of Public Health and Community Medicine, UNSW Sydney, Sydney, Australia
[h] Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, Melbourne, Australia
[i] Big Data Institute, Nuffield Department of Medicine, University of Oxford, UK
[j] Centre of Excellence for Biosecurity Risk Analysis (CEBRA), School of BioSciences, University of Melbourne, Melbourne, Australia
[k] Australian Research Council Centre of Excellence for Mathematical and Statistical Frontiers (ACEMS)
[l] MRC Centre for Global Infectious Disease Analysis, Department of Infectious Disease Epidemiology, School of Public Health, Imperial College London, London, UK
[m] Department of Zoology, The University of Oxford, Oxford, UK

ARTICLE INFO

Keywords:
Parameter identifiability
Data challenges
Computational methodology
Prior knowledge
Policy and communication
Bayesian analysis

ABSTRACT

Modern data and computational resources, coupled with algorithmic and theoretical advances to exploit these, allow disease dynamic models to be parameterised with increasing detail and accuracy. While this enhances models' usefulness in prediction and policy, major challenges remain. In particular, lack of identifiability of a model's parameters may limit the usefulness of the model. While lack of parameter identifiability may be resolved through incorporation into an inference procedure of prior knowledge, formulating such knowledge is often difficult. Furthermore, there are practical challenges associated with acquiring data of sufficient quantity and quality. Here, we discuss recent progress on these issues.

1. Introduction

Despite progress on many fronts, infectious diseases remain a key threat to human health worldwide (Heesterbeek et al., 2015). From 1-12 July 2019, we participated in the workshop "Influencing public health policy with data-informed mathematical models of infectious diseases" at the MATRIX institute in Victoria, Australia (MATRIX Institute, 2019). Much of the discussion and scientific work at this event concerned the challenges identified five years ago following the Infectious Disease Dynamics 2013 programme at the Isaac Newton Institute, particularly those related to the integration of multiple datasets (De Angelis et al., 2015). In this paper, we return to several of the challenges identified by De Angelis et al. (2015) and consider both recent progress and perspectives for this rapidly developing field.

One key challenge relates to the structure of the underlying assumed mechanistic model and the observed data; in particular, whether the model parameters can be estimated given the model and the observations, and whether we can obtain analytical insights into "parameter identifiability", a property of a model that must be satisfied for precise parameter inference to be possible. If parameter identifiability is an

* Corresponding authors at: School of Mathematics and Statistics, University of Melbourne, Melbourne, Australia.
** Corresponding author at: Department of Mathematics, University of Manchester, Manchester, UK.
*** Corresponding author at: School of Mathematical Sciences, University of Adelaide, Adelaide, Australia.
E-mail addresses: jennifer.flegg@unimelb.edu.au (J.A. Flegg), thomas.house@manchester.ac.uk (T. House), jamesm@unimelb.edu.au (J.M. McCaw), joshua.ross@adelaide.edu.au (J.V. Ross).

https://doi.org/10.1016/j.epidem.2020.100393
Received 24 November 2019; Accepted 25 April 2020
Available online 17 May 2020

issue, then can/should the model be reparameterised, and how should this be done? A natural question that arises here is whether we can measure something else in the process (included in the model yet or not) that can help resolve the issue, or build in existing (that is, prior) knowledge in a structured manner. For identifiable models, particularly if these are very complicated, there is a question of what we can reasonably do with current inference methods, for example Markov chain Monte Carlo (MCMC) or Maximum Likelihood estimation.

In this paper, we focus mainly on a Bayesian approach to parameter inference, which has been commonly adopted in the infectious disease modelling field since the work of O'Neill and Roberts (1999). This approach has proliferated for various reasons. One is the difficulty in interpreting an epidemic in terms of frequentist statistical theory, as a small sample from a larger population, making the Bayesian approach to parameter estimation more philosophically natural (MacKay, 2003). Another advantage of Bayesian methods is their ability to accommodate incomplete observations of the epidemic (O'Neill and Roberts, 1999), by treating missing data as latent variables. Bayesian methods facilitate data assimilation and uncertainty quantification in a natural and unified framework (De Angelis et al., 2015). Finally, computational methods are rapidly advancing in this space (see Section 5) making application of those methods to real-world data sets increasingly feasible.

In Bayesian inference, parameters θ are considered as random variables and the aim of the inference is to estimate their distribution. The posterior distribution, $p(\theta|y)$, is derived from the likelihood, $p(y|\theta)$, which comes from a probability model for the observed data y, and the prior, $p(\theta)$, which encodes knowledge available before the current data was observed:

$$p(\theta|y) \propto p(y|\theta)p(\theta)$$

The essence of Bayesian inference is to update what you believe about the parameters through the observation of data. This then poses another major challenge: how do we use prior knowledge in a consistent and convenient way within models of infectious disease? Within a Bayesian setting, it is natural to specify priors on the model parameters themselves. However, experts typically cannot easily quantify their beliefs about the parameters directly; rather they will have knowledge of (and so be able to construct a prior on) observables associated with the underlying process, such as the expected peak prevalence or the duration of the outbreak.

Statistical inference based on data generated by a single type of observation process is routine, but challenges remain when performing inference based on multiple observation processes, and/or different types of data from a variety of sources. Another challenge relates to issues around data and specifically how multiple types of data, drawn from alternative sources, can be included in the modelling framework. For example, when modelling a nascent epidemic, we might have access to case notification data, special studies (such as First Few Hundred studies), and phylogenetic data. Statistical models that integrate multiple data sources are beginning to gain traction in infectious disease modelling (De Maio et al., 2018; Campbell et al., 2019).

There is a need across all of the above issues to develop computational algorithms that can help end users to automate some of these processes. With increased volumes of and access to data (from multiple sources), algorithms need to be efficient and make use of recent advancements in computational hardware. However, there is still an important place for expert human input to gain mechanistic model insight (rather than relying, say, on machine learning techniques only, which also have their place, but are not a focus of this paper). This insight will be enhanced by addressing each of the challenges we have focused on, including considerations of parameter identifiability, the construction of priors around process observables, and the integration of data from multiple sources.

Finally, the field of mathematical epidemiology is intimately tied to the life sciences, epidemiology, and the practice of public health itself. To make an impact, that is, to contribute to policy with the purpose of reducing the burden of disease and saving lives, mathematical epidemiologists need to consider the context in which their work exists. Towards the end of the article, we provide some commentary on this broader context, and how it may influence our practice.

2. Identifiability

The mere ability to fit a model to data does not guarantee that the model's parameters can be uniquely determined; parameters may not be identifiable. As an example, consider the well-known deterministic SIR model:

$$\frac{dS}{dt} = -\frac{\beta SI}{N}, \frac{dI}{dt} = \frac{\beta SI}{N} - \gamma I, \frac{dR}{dt} = \gamma I \tag{1}$$

At the start of the epidemic, we can make the approximation that susceptibles are not depleted ($S(0) \approx N$), which results in an approximation described by:

$$I(t) \approx \exp((\beta - \gamma)t) \tag{2}$$

By inspection we see that the function $I(t)$ does not change if β and γ are changed, provided the difference $r = \beta - \gamma$ remains the same. Consequently, even if $I(t)$ is perfectly observed, only the value of r can be inferred, not the values of β and γ themselves. Given observation of prevalence then, β and γ are considered unidentifiable, while r is considered to be identifiable. For non-trivial models, identifiability analysis – using analytic and/or simulation-based techniques – is required to determine which model parameters are identifiable.

There are two main aspects to identifiability. "Structural identifiability" (Bellman and Åström, 1970; Whyte, 2013) concerns whether different parameter vectors produce different probability distributions of observed data. Structural identifiability ensures true parameter values can be inferred under idealised conditions: that the model is an exact representation of the system under study, and that the observations uniquely determine the probability distribution of the data. (This latter condition is only possible given an infinite number of observations.) Structural identifiability is thus a property of the model, not the specific data observed. For example, a continuous-time *deterministic* compartmental model can only be structurally identifiable if different parameter vectors generate distinct output trajectories, regardless of how precisely those trajectories are determined by a specific data set. A continuous-time *stochastic* compartmental model is structurally identifiable if different parameter vectors result in different probability distributions for the observables.

On the other hand, "practical identifiability" (Godfrey, 1983) concerns whether or not parameter values can be uniquely, precisely and accurately determined for realistic measurement frequencies, quantity and quality, and in light of discrepancies between the model and the real-world process under study. Practical identifiability is thus less rigorously defined, and dependent on the specific data observed. For example, when applying likelihood-based inference methods to a continuous-time deterministic compartmental model with stochasticity in the observation process only, the model is practically identifiable if one set of parameter values maximises the likelihood, given measurements for a realistic number of measurement time points (and replicates where relevant).

Structural identifiability is typically assessed using analytic or numerical methods, whereas practical identifiability analysis is often assessed by undertaking a simulation/re-estimation study. In such a study, one selects a particular parameter vector, uses it to simulate data from the model subject to noise, conducts parameter inference, and then investigates the features of parameter estimates to determine if estimates adequately approximate assumed values.

Within structural identifiability, a distinction is made between global identifiability, whereby a unique parameter vector in the whole parameter space can be determined, versus local identifiability, where a

unique parameter set can be determined in the neighbourhood of the true parameter set. Within practical identifiability, there is the distinction between *a priori* identifiability – where the results of identifiability analysis apply to all realistic data sets – and identifiability given a particular data set.

It may be taken as given that unidentifiability of some number of model parameters will greatly reduce the insight that can be drawn from fitting a model to data, and this is indeed often the case. However, the situation is more nuanced. One must consider whether the desired scientific insight rests upon the values of parameters themselves ("are the parameter values of intrinsic interest?"); or lies in use of the model to make predictions; or in testing competing mechanistic hypotheses. For these latter situations, non-identifiability is not necessarily as significant an issue. We now discuss two types of challenges associated with identifiability. First, what challenges lie in determining *a priori* which parameters may be identifiable given a data collection process? And second, what challenges are faced when interpreting the results of fitting an unidentifiable model to data?

2.1. Challenges in determining identifiability

Identifiability analysis has not yet seen widespread uptake in biological system modelling, as described by Nguyen et al. (2016). The authors noted that (Page 2): "… the booming works on mathematical models in biological and medical research over the last years have been accompanied with a disproportionately low amount of assessments on parameter validity …". They drew this conclusion by interrogating publication records in PubMed Central from 1990 to 2015. The authors noted that publications featuring models composed of ordinary differential equations were much more common than those which also listed keywords relating to some form of identifiability. See Nguyen et al., 2016, Fig. 1 and its associated supplementary text for details.

Part of the challenge in increasing the uptake of identifiability analysis is raising awareness of the many pitfalls that can occur in the absence of such analysis, and thus the necessity of performing such an analysis. We note that some authors in the field have recognised the value of analytical or numerical methods of scrutinising models, and have advocated for their inclusion in modelling practice (see Boianelli et al., 2015 for a relatively recent example). Many studies use simulation/re-estimation methods as described above to determine whether model parameters can be estimated, but do not explicitly describe this as identifiability analysis. The clear labelling of identifiability analyses as such, and references to established methods, would raise awareness of the benefits of identifiability analysis and change community perceptions both regarding the need to conduct such analyses, and the ease of doing so.

However, even where researchers are aware of the importance of identifiability analysis, several practical problems can arise. First, ideally, a model should be practically identifiable *a priori*, but this is difficult to establish. Instead, most methods assess practical identifiability for a given (simulated or actual) data set or a given set of 'true' parameter values (Raue et al., 2009; Yan et al., 2019). On the other hand, there is a proliferation of methods to assess structural identifiability (as reviewed by Chis et al., 2011), but structural identifiability is necessary but not sufficient for practical identifiability. A challenge arises then in either improving methods for *a priori* practical identifiability analysis and/or making them more accessible, or developing methods to combine the results of structural and practical identifiability analyses for selected data sets.

Identifiability analysis presents a major technical barrier, especially to the non-specialist. The barrier is perceived to be particularly high for structural identifiability analysis, as it involves manipulation of model equations rather than simulation and parameter estimation, the latter of which are more readily accessible skills, already used in fitting models to data. A challenge for the field is to promote the use of automated tools for structural identifiability analysis. Many tools have already been developed in the context of systems biology (Bellu et al., 2007; Chis et al., 2011; Meshkat et al., 2014; Karlsson et al., 2012), but awareness of their utility in the epidemiological community remains low, and these are often implemented in proprietary software, limiting their accessibility.

Difficulties of structural identifiability analysis have encouraged the alternative of practical identifiability analysis using numerical methods, although as discussed above, structural and practical identifiability analyses serve slightly different purposes and are not strictly interchangeable. The simulation and estimation processes for practical identifiability analysis are relatively straightforward (although may be time-consuming), but automated tools would still lower the barrier for their application. More importantly, decisions on how to conduct such a study and interpret results are not straightforward, and there may remain issues with convincing journals, editors and reviewers that simulation/re-estimation studies are worthy of publication in and of themselves and indeed required before data analyses are undertaken and reported.

Returning to the actual conduct of practical identifiability, a number of questions arise. For example, what parameter values should be used to simulate data? How many parameter vectors should be used to provide confidence that results can be considered general? When interpreting parameter confidence intervals or posterior distributions, how narrow do they have to be for us to claim that a parameter is

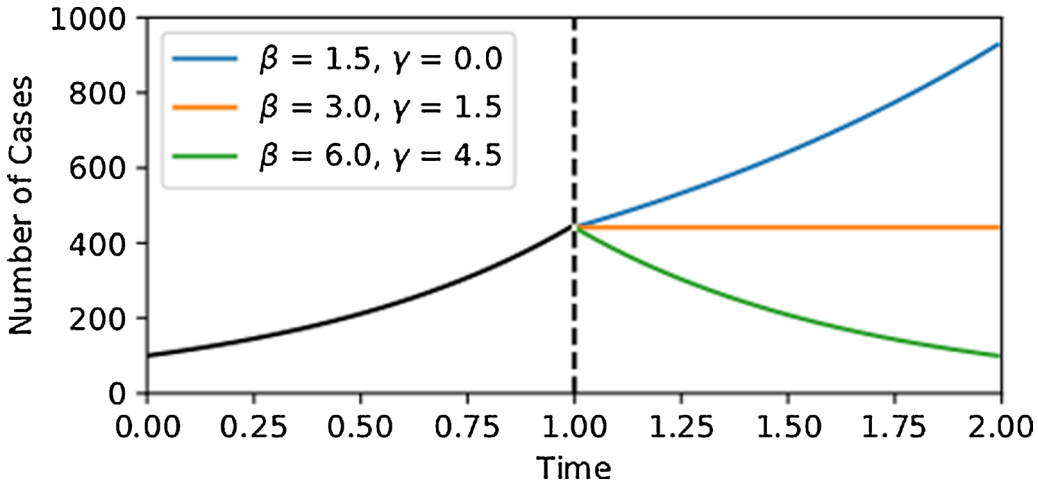

Fig. 1. Solutions of Equation (1) giving the number of infectious people, with an intervention at Time = 1 that halves the pre-intervention β for the post-intervention period. Multiple parameter pairs reproduce the pre-intervention data (black line), yet distinct parameter pairs (β, γ) produce differing post-intervention predictions.

identifiable? Proposing robust numerical methods to provide insight into undesirable model features, and guide their remediation, remains a challenge for the community.

A review of structural identifiability analysis methods and their suitability for different model structures in infectious disease modelling would help guide practitioners in their choice of methods, as has been conducted for systems biology models (Chis et al., 2011). Although the suitability of analysis methods is obviously independent of the physical interpretation of model equations, a similar review for common model structures in infectious disease modelling would enable easier comparison between reviewed models and a particular model of interest.

2.2. Challenges in interpreting the results of fitting an unidentifiable model to data

If a model is unidentifiable, obtaining meaningful results from fitting to data and interpreting these results can be extremely challenging. Appropriate interpretation begins with an acknowledgment that the adequacy of the inference depends on how the model is to be used.

If the aim of fitting the model to data is to determine values of parameters that are of intrinsic interest, then one can examine whether it is possible (or likely) that changes to the model or to planned data collection will remedy any lack of identifiability. For example, does holding some parameters constant (or imposing strong priors on some parameters in the Bayesian context) or acquiring additional data result in an identifiable model? Reparameterisation is unlikely to help in this situation, as the reparameterised model may be identifiable, but the original unidentifiable parameters of interest will no longer appear in the model. However, information on identifiable parameters can guide reparameterization, enabling one to make stronger claims about the values of the new parameters, the biological interpretation of which can then be investigated. One advantage of testing a model for global structural identifiability is that it can reveal the parameter combinations ("observational parameters" (Jacquez and Greif, 1985)) which can be determined uniquely under the idealised conditions of the test. Used with a method appropriate for the model class (e.g. one employing the notion of "structural equivalence" for linear state-space models (Vajda, 1984), knowledge of these combinations can guide the reparameterization of the model into one that is globally identifiable.

On the other hand, if the aim of parameter inference is to make model predictions – either for the unperturbed system or in the context of an intervention – we may seek to propagate parameter uncertainty through our model so as to produce a range of predictions, allowing us to quantify prediction uncertainty. It is possible that although parameters are not individually identifiable, parameter sets consistent with observations make similar (or identical) predictions, or that the quantitative behaviour of certain subsets or functions of parameters are well determined despite lack of parameter-level identifiability (e.g. Yan et al., 2019). However, if interventions act by changing the values of unidentifiable parameters or unidentifiable combinations thereof, predictions are unlikely to be consistent. For example, consider the SIR model and suppose that there exists an intervention which halves β for all time after the intervention is applied to System (2). Fig. 1 shows that alternative values of $(\beta/2, \gamma)$ lead to a wide range of predictions for I over this post-intervention time period. As such, uncertainty over the true values of β and γ creates doubt over the benefit of the proposed intervention. Additional data collection and/or fixing model parameters may be required, as discussed above. Reparameterisation is unlikely to be helpful in this context, as the new parameters will be identifiable but not linked directly to the intervention (for example, reparameterising system (2) as $I(t) = I(0)\exp(rt)$, removing β). An exception is if new parameters are subject to stronger priors than the original parameters, enabling more precise inference of the values of the original parameters.

In summary, addressing the challenges associated with identifiability will provide the disease modelling community with a systematic means of comparing models and evaluating their usefulness. In turn, we expect this to enable progress on the discipline's fundamental challenges in using models to direct resources towards ensuring better health outcomes.

3. Incorporating prior knowledge

The ability to identify parameters, or at least to have distributions on parameters which capture our full knowledge of the disease process, is dependent upon use of prior knowledge. When analysing a (new) data set within the Bayesian framework, we must specify a prior distribution on the parameters of the model. This provides a natural way to incorporate existing information (obtained from the literature, past experience etc.) about plausible values for the parameters. The prior distribution also offers a way to incorporate information about observable quantities (i.e., properties of the system that can be measured and expressed as a function of model parameters). This requires a clear distinction between knowledge of the real-world system (and our understanding of it) and knowledge of the model we are using to represent it (Gelman et al., 2017; Craig et al., 1997).

The Bernstein-Von Mises theorem tells us the likelihood will (asymptotically) come to dominate over the prior as the amount of data increases provided relatively mild technical conditions are met (Kleijn and van der Vaart, 2012). However, one should not use this as an excuse to neglect the choice of prior distribution. In cases where data is limited, there is a risk of a (potentially incorrect) prior dominating the analysis; the result of the analysis will not reflect the data (and desired distribution over parameters of interest) but rather the (potentially incorrect) prior information.

Furthermore, when analysing infectious disease data, it is unusual to have large amounts of high-quality data, and hence it is very attractive to supplement our data with prior knowledge. A weakly informative prior can assist with some of the statistical identifiability issues discussed in the section above on Identifiability, but if the prior distribution is poorly specified (even if a non-informative prior is used) it can lead to misleading results (Gelman et al., 2017). A desirable solution is to select a prior distribution that concentrates prior probability on plausible parameter values but does not dominate the posterior distribution, i.e. it allows the information in the data to determine the outcome of the inference. But doing so may be difficult.

Mechanistic models of infectious disease dynamics are often expressed in terms of parameters far removed from the aspects of the process that are observed. It is arguably rare for scientists to have good knowledge of model parameters; rather, they have, or could construct, an informed view of various system observables. To make explicit the difference between model parameters and observables, consider a mathematical model of influenza infection within a host. *A priori*, one may not know plausible orders of magnitude for the rate parameters (of the mathematical model), but likely they will know that an influenza infection resolves in days, rather than hours or months.

Given that scientists often have a better grasp on the value of observable quantities rather than the model parameters themselves, it makes more sense to articulate our prior belief in terms of these observable quantities. But of course, we then have a task to translate between the two. This challenge is not universal across models. Consider logistic growth describing the size of a population through time, $N(t)$, with the following differential equation:

$$\frac{dN}{dt} = rN\left(1 - \frac{N}{K}\right)$$

It is simple to parameterise this model in terms of the growth rate, r, and the carrying capacity, K, both of which are natural quantities to observe. Consequently, if one were interested in developing a prior distribution for this model, one would only need to specify the distribution of plausible values to observe for these quantities. Considering the alternative equation

$$\frac{dN}{dt} = rN - cN^2$$

the observable consequences of different values of the parameter c are less obvious, and hence it is unclear what plausible values may be. This situation is common in mathematical epidemiology: the observable quantities are themselves a function of the solution to the model, and most disease dynamic models do not have any closed analytic form for these quantities. Consequently, it is infeasible to develop a clearly interpretable parameterisation – in terms of writing down a prior based on past observations – of these models.

The problem extends beyond just the parameterisation used. It is tempting to think that it is sufficient to use published estimates of individual, model-inferred parameters to construct a prior distribution. However, even when this is possible, issues will still arise. Reported parameter estimates are model dependent. The quintessential example being use of an SIR or an SEIR model, whereby inference may lead to the same estimates of the basic reproduction number but different estimates of the rates of infection and recovery. Understanding the marginal distributions of parameters is insufficient to construct a plausible prior distribution when correlations determine model behaviour.

If we are serious about incorporating prior knowledge into future analyses, then it seems sensible that we should turn to fields where this has already successfully been implemented. Priors informed by expert opinion (obtained via expert elicitation) or the results of previous studies have proved popular in the field of ecological where they have found diverse applications (Hemming et al., 2017; Low Choy et al., 2009). However, the use of expert elicitation to inform priors has become much more popular among ecologists over the past decade (Drescher et al., 2013). Using expert opinion in conjunction with Bayesian inference is an attractive option for researchers in ecology, where the types of data collected are likely to involve a high degree of uncertainty (Kuhnert et al., 2010; Martin et al., 2005), or are difficult or expensive to collect (O'Hagan, 1998; Martin et al., 2012).

Martin et al. (2005) solicited expert opinion for a study regarding the effect of livestock grazing on various Australian birds. In cases where experts agreed with each other, the resulting credible sets for parameters are typically tightened, this can be an effective and cost-effective way to improve estimates. This improvement was particularly noticeable in situations where existing data was weak but the precision, or agreement, between expert information was high. Just as important is the case of incorporating expert knowledge where there was noticeable disagreement between experts; in this situation, results did not differ significantly from analysis where expert opinion was not used.

Expert opinion can also improve confidence in parameter estimates obtained from analysing these data, although care should be taken to ensure the elicitation of this information is carried out correctly (Martin et al., 2005; Morgan, 2014). There is a need to be rigorous in the selection of experts and the execution of the elicitation (Drescher et al., 2013). Kadane and Wolfson (1998) and Chaloner and Duncan (1983) have developed methodology which abstracts much of the mathematical detail, simplifying the elicitation process. However, these methodologies only cover certain types of models, and are limited in scope and the degree to which they scale with the complexity of the model being considered. We believe it would be beneficial for the modelling community to extend this work, and develop statistical methodology that enables "automatic elicitation" for a wider range of models.

While peer-reviewed literature provides a rich source of 'prior knowledge' and would typically form the basis for determination of the prior, caution should be exercised due to potential systematic effects in publication practice that can make it difficult to reliably source and account for all (published) primary sources (Reich et al., 2011). In that paper, Reich et al. demonstrated that, due to publication and referencing practices, so called 'medical facts' can become enshrined as truth in the absence of strong and sufficient empirical evidence. This strongly suggests that when using the peer-reviewed literature to establish a prior, one must proceed carefully, being sure to establish a process to identify relevant primary source literature and avoid the pitfalls identified by Reich et al.

Given the importance of the prior distribution and that there may be substantial amounts of knowledge about observable quantities of the process being modelled, how might one go about using this knowledge to specify a prior distribution? An idealised work-flow to utilise prior knowledge may consist of the following steps: (1) determine a set of relevant observable quantities which characterise the system and for which there is some quantitative understanding; (2) construct a prior distribution on the parameters of the model which reflects this understanding; and (3) carry out the remainder of the inference process as normal. There are three aspects to this work-flow which are challenging: (1) determining appropriate observable quantities; (2) representing this information in the prior distribution of parameters for an arbitrary model; and (3) devising a way to do this which is not prohibitively computationally expensive. We now expand on these three challenges.

3.1. Which observable quantities?

The first aspect is the most specific to the particular application, in that it requires some knowledge of the system being modelled to know what observable quantities of the process characterise it. When multiple observable quantities are being used, strong correlations between the quantities leads to redundancy; ideally independent observable quantities would be used. Above, the time required to resolve an influenza infection was given as an example of an observed quantity. This quantity may be expressed in terms of the solution to a particular model, even if it is not one of the parameters of the model. For example, if one were modelling the number of people hospitalised with influenza during an epidemic, one might use the total number of patients hospitalised as an observable quantity. Historical records of hospitalisation could be used to estimate, a distribution for this for previous epidemics. In the SIR model, the final size is determined by the initial condition and basic reproduction number, hence information about the final size can constrain the prior distribution for these parameters (Miller and Joel, 2012).

3.2. How to represent prior knowledge?

The second aspect involves the process of taking a representation of the uncertainty in, potentially several, observable quantities, and translating this into a prior distribution over the parameters of a model. This is a non-trivial task, even when the distributions on the observable quantities are self-consistent (and it is easy to construct examples where this is not the case). Moreover, when there are correlations between observable quantities this can further complicate matters if one is not content to assume a joint distribution with independent components.

3.3. Can choosing a prior distribution be made easier?

Finally, the third aspect involves finding a way to efficiently choose a prior distribution. As discussed above, extensive research from the field of psychology suggests that eliciting information from experts in a defensible way is labour-intensive. However, there are alternatives to obtaining information from domain experts. With the rise of "big data" there will be an increasing amount of data to be mined to inform prior distributions. The use of additional data sources brings its own challenges, e.g., accounting for correlations between the data sets as discussed further in the subsequent section on Data Challenges.

While solutions to these problems would improve our ability to carry out inference, there is another equally important conceptual contribution from this work-flow. Decoupling prior knowledge of observable quantities and prior distributions from specific mathematical models allows us to create prior distributions which can be shared

between models and capture the same information, independent of how that information translates into a distribution on the parameters of that particular model. This itself is a powerful idea, as it increases the portability of parameter estimates as they are no longer attached to the particular model with which they were obtained. For example, suppose models X and Y share observable quantities but do not have the same set of parameters; it is possible to use parameter estimates obtained with X to construct an equivalent prior for Y, where "equivalent" refers to having the same distribution on observable quantities. Moreover, since such observable quantities will often involve a combination of the model parameters, even if the prior distributions of the observable quantities are independent, the prior distribution they enforce on the parameters may have a rich correlation structure.

Spending considerable effort on choosing a prior distribution can seem indulgent and, provided there is sufficient data, it will often have only a small effect on the results. Debate over the choice of prior has inspired many developments in Bayesian statistics, and there are cases where it has a strong effect on the inferences drawn (Moss et al., 2019a). In this section we have described some open problems regarding the practical application of informative prior distributions while attempting to motivate why, despite these difficulties, they can still be very beneficial. In the words of Judea Pearl, "It is plain silly to ignore what we know" (Pearl, 2001), however this comes with the caveat that it relies on this knowledge being appropriate and accurate.

4. Challenges posed by data

As discussed by De Angelis et al. (2015), precise and accurate model inference relies on the volume and quality of data, with quality encompassing both variability and bias. If the observation process is well-characterised, parameter estimation can still be performed, although the precision of these estimates will be limited by the quality of the data. Combining data from multiple sources can increase the precision and accuracy of estimates but presents its own challenges. In this section we will first discuss challenges in characterising the observation process for a single epidemiological data set, before moving on to those for combining multiple data sets.

4.1. Challenges within a single epidemiological data set

When data measuring one quantity is used as a proxy for another quantity, the relationships involved should be well-characterised to reduce bias in inferred quantities. For example, the proportion of individuals reporting influenza-like illness (ILI) in a weekly community survey can be used as a proxy for influenza prevalence in the general population (Adler et al., 2014; Carlson et al., 2010). Here, ILI is used as a proxy for influenza virus infection; however, influenza infection prevalence may be overestimated due to ILI caused by non-influenza pathogens, or underestimated due to asymptomatic infections. Moreover, prevalence in the survey population is used as a proxy for prevalence in the general population, but demographics of the survey population may not reflect that of the general population. To reduce bias during inference, the relationship between reported ILI and influenza virus infection, and between the survey and general population, should be explicit in the observation process. For example, the former can be achieved by specifying a reporting probability conditional on influenza infection, and a background observation probability due to illnesses other than influenza that may vary over time.

The biases which are likely to affect inference, and are thus important to model, are likely to differ by both situation and data type. For example, when performing inference during outbreak scenarios, if lags in data collection are either ignored or mis-specified, inference may be poorly affected (Azmon et al., 2014; Moss et al., 2019a). However, even for the same data type (such as incidence data), such lags may not drastically affect inference in endemic scenarios. Reporting rates are also more likely to vary during the course of an outbreak, as indicated by Flutracking data (Carlson et al., 2010). Many studies calculating time-varying effective reproduction numbers (e.g. Rosello et al., 2015) are not robust to time-dependent reporting rates, so if these methods are used, time-dependent reporting rates should be included in the observation process.

Where available, data sources covering the same timeframe as the primary data should be used to inform biases. For example, community survey data can be used to infer changes in healthcare seeking behaviours and testing practices over the course of an epidemic, which can then be used to improve epidemic forecasts using a different data source (Peppa et al., 2017; Moss et al., 2019b). In other situations, sensible observation models and/or parameter values can be obtained using historical data. For example, observation noise can be estimated for previous epidemics, and resampled from when proposing parameters to fit to data from a new epidemic (Ertem et al., 2018). Where the relationship between historical and current observation processes is unclear, rather than assuming that the historical and current observation processes are the same, a better approach may be to use historical parameter values to inform values of the current observation process. For example, different outbreaks of the same pathogen may occur in different locations and in different populations, and it is unclear how observation biases translate across outbreaks. When we know the direction in which a parameter will change but not by how much – for example, assuming that testing rates will increase in a pandemic – a historical parameter value could be used as the lower bound for a prior distribution on the testing rate or, more conservatively, to construct a prior distribution where only a small proportion of probability mass is below this value. Another example is when an intervention increases the testing probability. This scenario requires particular care, as increased testing may increase observed prevalence even when an intervention is effective, and inference ignoring increased testing may incorrectly conclude that the intervention increases prevalence (Ali et al., 2015). Conversely, increased testing (motivated, say, with the aim of ascertaining every case possible) could decrease test-positivity, if specimens are collected indiscriminately and testing denominators are unavailable, and this could lead to under-estimates of prevalence.

When inference is conducted on “incidentally available” data rather than data collected for the particular inference study, extra attention has to be paid to modelling of the observation process. This is especially an issue in outbreaks, as surveillance protocols are developed alongside the unfolding of the outbreak. Hay et al. (2018) documented that case definitions for microcephaly became more stringent as the Zika outbreak in 2015-2016 developed, and that many cases were reclassified. Either a combination of behavioural change and overreporting of cases under early definitions, or increased Zika surveillance between the two epidemic waves, were required to explain changes in reported microcephaly incidence. Communication between field workers, policymakers and modellers becomes especially important in this context, and local modellers have the opportunity to inform the data collection process (see Policy and Communication section).

On the other hand, closer ties between designers of data collection protocols and developers of inference methods have enabled the collection of data sets designed to infer particular model parameters, as identifiability of model parameters may depend on the study design (see Identifiability section). For example, in the 2009 influenza pandemic, first few hundred (FF100) studies were conducted specifically to understand the transmissibility and severity of the disease during the early stages of the epidemic (McLean et al., 2010). Since the collection of this dataset, model-based inference methods have been developed to infer hospitalisation rates and within-household transmission in real time (Black et al., 2017).

4.2. Challenges when using multiple epidemiological datasets

When modelling transmission of an infectious disease there are often multiple different epidemiological data types available with

which to infer model parameters. For example, when inferring key characteristics of a nascent epidemic, we might have access to confirmed case counts, syndromic surveillance data and special studies (such as FF100), but each carries its own underlying biases and there is no guarantee that they provide a self-consistent view of disease activity (e.g. Thomas et al., 2015). While there are established approaches to performing inference with each of these types of data, we rarely use methods that can simultaneously consider all available data although there are notable exceptions (Corbella, 2019). Consequently, typically the result is either multiple competing parameter estimates – which leads to obvious challenges for decision-makers – or a single estimate that ignores some of the available information. The alternative is to analyse all available, relevant data using a single joint model. Ideally, the joint model is able to integrate the different data sources in a way that retains the strengths of each, without losing information. De Angelis et al. (2015) note that the motivation to combine information from multiple data sources arises from both a perception that this will produce more `defendable', robust outputs and a recognition that comprehensive outbreak analysis requires multiple data types. Here we will focus on data-integrating models for multiple epidemiological time series data but note that models for integrating these and other epidemiological data types and/or phylogenetic datasets have recently gained traction in outbreak analysis (De Maio et al., 2018; Campbell et al., 2019).

Integrating multiple different datasets can increase the precision and accuracy of parameter estimates and enable a greater range of relevant (unobservable) quantities to be estimated (Birrell et al., 2018). For example, sharing information across abundant, low-quality surveillance data (*i.e.*, high volume, but unknown or poorly characterised observation processes) and a subset of high-quality surveillance data (*i.e.*, well-characterised observation processes, but low volume) can enable the estimation of nuisance parameters, like reporting biases. This approach has previously been demonstrated in ecology (Fithian et al., 2015), where data structures are not dissimilar to disease data. Another advantage of joint inference is the automatic weighting of information from different data sources. When writing out a joint model, explicitly describing the observation models for each data type (including parameters for reporting biases) provides an objective way to weight their respective utility.

An important challenge when constructing a joint model is the handling of dependencies between data sources. Here we separate these dependencies into two types: 1) data sources observe the same underlying epidemiological process, and 2) observation processes themselves are dependent (for example, individuals may be captured by two or more surveillance systems).

Understanding dependencies in the observation process will require close collaboration with data collectors and public health policy-makers (Muscatello et al., 2017; Doms et al., 2018). In order to understand the magnitude of dependencies in the observation processes between outbreak surveillance datasets, it is important to know precisely how each dataset is assembled and to map out all possible research and health system pathways that could lead to an individual being counted in one or more dataset(s). For example, during an influenza pandemic, households participating in FF100 studies are likely to be recruited from routine case notification systems and would therefore be counted in both FF100 data and case notifications. Further, if we wanted to add information from community survey data such as from Flutracking (Carlson et al., 2010; Moss et al., 2019b), we must consider how duplications of these data may arise in FF100 studies and/or case notification datasets.

While it may be possible to write out a single model that links all available datasets, there could be practical hurdles to performing inference for such models. It is important to consider whether inferring parameters from multiple datasets simultaneously, and thus adding substantial model complexity, is worthwhile from both a computational and a decision-making perspective, particularly if supporting decision-making in real-time is a goal. For example, Shubin and colleagues (Shubin et al., 2016) made simultaneous use of data from community and hospital surveillance systems in their transmission model of pandemic A(H1N1)pdm09 influenza, but the computation time for inference is reported in months, which is not practical for real-time use (note that this was not the goal of their analysis). Moreover, if the goal of such modelling studies is to inform public policy (in real-time or otherwise), the model structure and appropriate interpretation of its outputs will need to be clearly communicated to decision-makers, and more complicated models may be more difficult to translate (see Section 6).

It should also be noted that integrating multiple, low-quality data sources with a single high-quality dataset, will not necessarily provide benefit over analysing the high-quality dataset alone. For example, Moss et al. (2017) found that simultaneously using data from three different surveillance systems only improved retrospective seasonal influenza forecasts under certain circumstances, and could even reduce forecasting performance, when compared to forecasts generated using a single data source. They hypothesised that the synthesis of data from multiple surveillance systems may only provide benefit if each data source captures distinct, but complementary, aspects of the epidemiological or observation process.

5. Computational methodology

In common with other areas of mathematical biology, methods for fitting complex epidemic models have progressed a lot recently, driven by advances in Bayesian computational statistics (Green et al., 2015). These methods can be classified in a number of ways and our taxonomy reflects our personal biases and preference for mechanistic, stochastic, models. While non-mechanistic models are useful for some forecasting problems where relatively large amount of historical data are available (Brooks et al., 2018), small data sets, as would be available in the event of an outbreak, can only be interpreted in a mechanistic setting, and likewise the testing and forecasting of various intervention strategies. A fundamental difficulty with inference from outbreak data is that most of the underlying process is unobservable, hence the need to infer or integrate over a large amount of missing data to sample from the parameter posterior. One way of classifying existing algorithms is according to which part of the calculation handles the missing data. This impacts how suitable they are to be parallelised and hence handle larger problems as well as incorporate other evidence, including multiple datasets (Brooks et al., 2011; Birrell et al., 2018).

For models of small, closed populations such as households, continuous-time Markov chain (CTMC) models have found success, due to the size of the state-space being small enough to leverage numerical solutions for calculating the likelihood (Black et al., 2017). For most models, in larger populations, these methods break down due to the increased size of the state space. The oldest methods for exact inference are so called data-augmented (or auxiliary variable) MCMC (DA-MCMC) (O'Neill and Roberts, 1999). These typically infer the missing data as well as the parameters as a single Markov chain from which an expression for the likelihood is trivial to evaluate. Samplers are also easy to construct using a combination of Gibbs and Metropolis-Hastings steps. Data-augmented methods are highly flexible, allowing the use of non-Markovian models, non-homogeneous mixing and detailed spatial information (Touloupou et al., 2018; Stockdale et al., 2017). The downsides are common difficulties with convergence and mixing that get worse as the amount of missing data to be inferred grows (McKinley et al., 2014). Efficient use is reliant upon conjugate priors (allowing the posterior to be specified explicitly), so incorporating more general, informative priors (as discussed earlier) can be challenging. Finally, DA-MCMC is fundamentally a serial algorithm, so its use on large datasets becomes slow and parallelism is not easily exploited, beyond running multiple chains.

Although almost all useful epidemic models are analytically

intractable, they are typically very simple to simulate. Approximate Bayesian computation (ABC) uses simulations for fitting models where the likelihood is intractable (Kypraios et al., 2020), but where the simulated data can be compared with summary statistics. This is probably the most simply implemented method in this class but comes at the cost of introducing some approximation into the posterior. Other methods are exact in that they use an estimate of the likelihood, but still target the correct posterior. The use of sequential Monte Carlo (SMC) methods (which perform estimation sequentially through data) in epidemic modelling is very natural due to the prevalence of time series data and the need to fit dynamical models (Doucet et al., 2001). Pseudo-Marginal methods such as particle marginal Metropolis-Hastings exploit the unbiased likelihood estimate obtained from a particle filter to also perform inference for the underlying parameters (Andrieu et al., 2010; Brooks et al., 2015). In comparison with data-augmented methods, these methods can be seen as integrating over the missing data in the estimation of the likelihood, so the Markov chain targeting the parameter posterior is greatly simplified.

The key to the efficient operation of these pseudo-marginal algorithms is keeping the variance of the likelihood estimate within tight bounds (Doucet et al., 2015; Sherlock et al., 2015), otherwise the mixing of the Markov chain targeting the posterior can become very poor. An advantage of pseudo-marginal methods is that they are parallelised quite naturally, so modern computing hardware can be leveraged to reduce the variance in the likelihood estimate by simply using larger numbers of particles of averaging the estimates of independent particles. Although these methods represent the current state of the art in this area, there are still challenges to be overcome. They are not 'online' in that the computational expense (and hence run-time) increases as the length of the time series increases (Kantas et al., 2015). Non-Markovian models remain challenging and the overall efficacy is restricted by the ability to produce simulations that are in some way close to the observed data. Current research employs variance reduction techniques to reduce the variance of the likelihood estimate, and in particular importance sampling has been used to produce realisations that match data closely (Black, 2019; McKinley et al., 2014). In the case where the model is not strictly non-identifiable, but there is a complex posterior distribution over the parameters, then MCMC methods based on Riemannian geometry can be used (House et al., 2016).

5.1. Model Selection

Using data to accurately infer parameters of epidemic models is an important step for informing public health policy, forecasting and understanding the dynamics of diseases. However, if the models are inappropriate these tools generate misinformation. Information criteria such as AIC, BIC, DIC are the most common method for deciding on the best model. These information criteria are used widely by both frequentists and Bayesians due to their asymptotic properties and their often ease of calculation. However, in cases where there is little data and reasonable prior understanding of the epidemic process, these criteria may fall short. Simply put, this is because AIC, BIC, and DIC are intrinsically non-Bayesian as their formulations do not account for prior knowledge of model parameters.

The gold standard for selecting between models while accounting for prior information is to either calculate Bayes factors or the model evidence (Kass and Raftery, 1995). These approaches are sometimes avoided due to their computational difficulty; they either require calculation of the normalising constant of the posterior distribution, or calculation of a ratio of normalising constants. Although this problem is difficult via classic methods such as reversible-jump MCMC (Green, 1995), there are increasingly efficient methods for performing this kind of model selection. One approach is SMC^2 (Drovandi and McCutchan, 2016; Chopin et al., 2013), which allows for model selection to be performed during the inference process. Although this is an attractive and efficient method, the stochastic error in model selection estimates is not well understood. An alternative is importance sampling-based methods (Gelfand and Dey, 1994; Touloupou et al., 2018), which give unbiased estimates for model selection along with estimates of error. While these methods are computationally intensive, they are made efficient if parameter inference is performed *a priori*. Further, they are embarrassingly (i.e. trivially) parallelisable, that is, they are able to take advantage of modern computational architecture.

Another recent approach to model selection has been through the use of classification methods, and in particular Random Forests (Pierre et al., 2016). This approach has the benefit of needing only to simulate from the model, and efficiency of the classification algorithms themselves. This is particularly important if one is to consider the optimal design for model selection (Markus and David, 2018), a particularly computationally-expensive pursuit, and for which heuristics based upon the Random Forest approach have been proposed recently (Cope and Ross, 2020).

6. Policy and communication

Until now, we have concerned ourselves with some of the key technical challenges in modern day mathematical epidemiology. But for our work to contribute to public health policy, we must consider the broader scientific, social and political environment in which it exists. This is a broad topic and one not unique to health, for example ecology provides a highly-relevant exemplar discipline in which modelling has had a sustained and meaningful impact on decision making (e.g. Ball et al., 2009). A central tenet of the approach in that field is that we must distinguish, from the outset, between science for knowledge and discovery's sake, and science for the express purpose of contributing to the decision-making process. Within this context, we make the following observations.

Mathematical models are developed by researchers from a broad range of backgrounds, many of whom do not necessarily have the knowledge required to translate the intricacies of structural identifiability analysis and Bayesian approaches to parameter estimation into practice. In general, more complex models tend to be favoured by policy makers as they are perceived as being more "realistic", and indeed such models likely have more internal validity. Yet they are, in all likelihood, less general and have weaker external validity, in terms of providing unbiased predictions for other related scenarios or situations. Complexity can be a desired quality of a model even if the data are not available to support the model structure. Thus, there is a tension between transparency (which correlates with incompleteness but also with generality) and realism (which correlates with complexity and opacity, but also completeness and lack of generality). Highly complex models are in danger of projecting a false sense of accuracy, with the portrayed accuracy enhancing their attractiveness to policymaking stakeholders, but with untenable policy recommendations as a potential outcome of this cycle (Cooper, 2006). And even if accurate for the scenario at hand, findings drawn from them will be less generalisable (because they are more specific). Clearer communication of the pitfalls of the "unconscious use of a non-identifiable model" (Siekmann et al., 2012) will help to avoid this trap.

As the epidemiological modelling discipline globalises and nascent modelling groups form in countries previously lacking this capacity, our community should consider not only training modellers to model, but also explore how models are used in the policy environment. It is important to acknowledge that this environment will differ, and in strong ways, across different social, economic and political contexts. We propose that such topics are included in future training activities offered by established groups. We encourage success stories of models influencing public health policy (or otherwise) be published. The practice of modelling goes far beyond the technical details of the mathematics, statistics and biology.

Policy makers are not the only ones with whom the modellers need to communicate effectively. During the lifecycle of the model

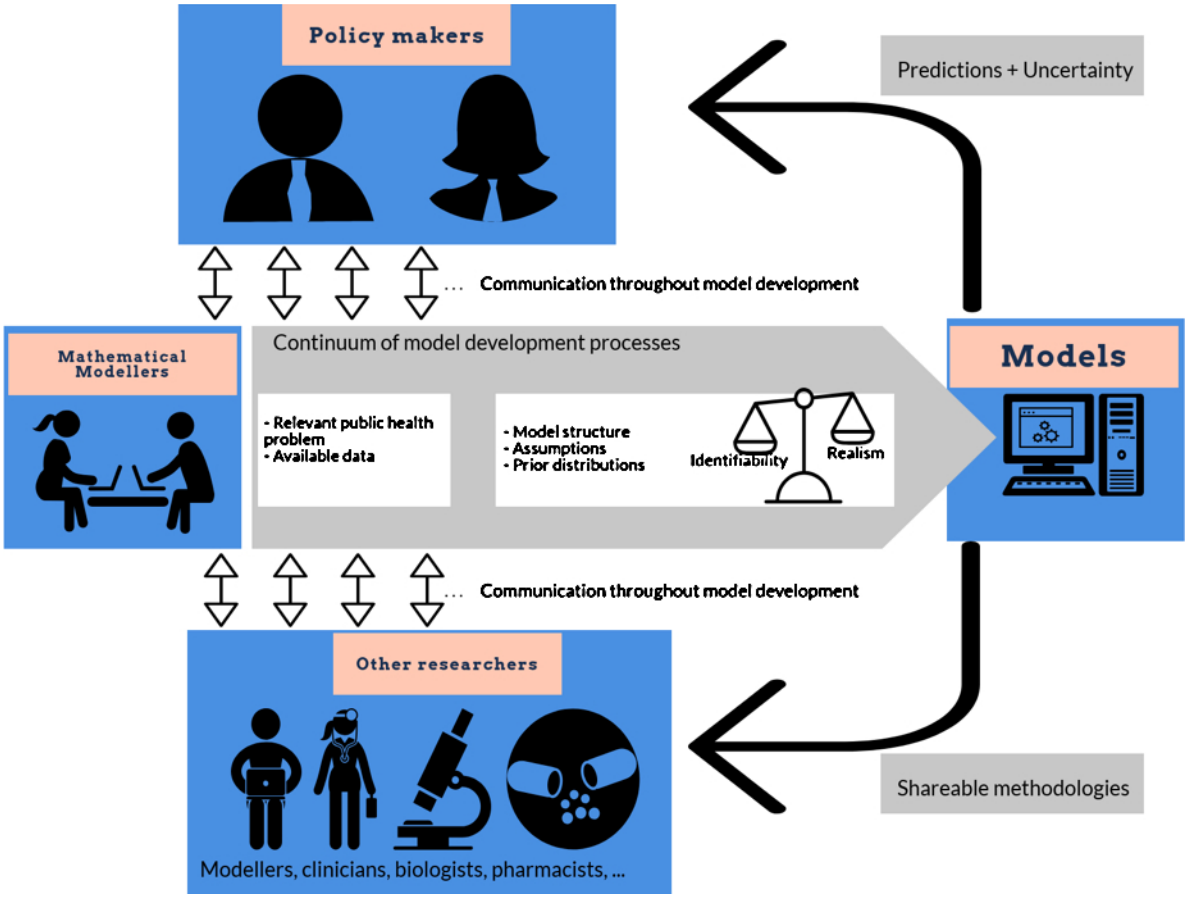

Fig. 2. Essential communications between three key groups in the model development and outcome dissemination.

development, we need to understand the intricacies of the problem, the available data and the existing views on reasonable values for parameters. Therefore, communicating and involving the experts in the respective biological, medical and health fields should lead to better calibrated and validated models.

Effective interactions among these three key groups (modellers, researchers from other disciplines and policymakers) relies on effective communication, with several themes and foci for improvement (Fig. 2).

6.1. Developing a shared understanding of the problem

Modelling that strives to inform policy is best conducted in consultation with policymakers. Such engagement enables modellers and policymakers to gain a mutual understanding of the policy question to be addressed and define specific modelling objectives. It is also an opportunity to improve modeller understanding of the policy context and stakeholder understanding of the capabilities and limitations of mathematical models within this context. Examples of productive engagement between policymakers and modellers exist in many fields, including infectious disease epidemiology (Lee et al., 2013; Probert et al., 2016; Knight et al., 2016; Qualls et al., 2017; Moss et al., 2019b). Recently, researchers have explored the use of participatory approaches to modelling, where policymakers and modellers "co-develop" models and applications (Gaydos et al., 2019). These examples suggest that a variety of approaches to stakeholder engagement, e.g., in terms of the frequency and timing of consultations within the cycle of model development, can lead to successful outcomes. Modelling conducted in the absence of stakeholder engagement is at risk of being underused, misused or inappropriate for addressing a specific policy question (Glasser et al., 2011; Muscatello et al., 2017; Doms et al., 2018).

6.2. Understanding the available data

Data is required to calibrate, validate or fit the mathematical models. As part of the planning phase of the model development, knowledge of what data are or will become available is essential. If the data is not publicly available, it is essential that communications are made with the proprietor of the data and have prior agreement with them on data usage.

Modellers have to interact with researchers from other disciplines such as biology, medicine, pharmacology etc. to make more sense of the available data. In these instances, improved communication of modelling ideas to a non-mathematical audience would allow for better discussion on whether the important biological aspects are captured from the model and how the data can be used to estimate the model parameters.

If multiple datasets are available, the use of all datasets in a single data-integrating model may be appealing to both modellers and policymakers – under the assumption that more data will result in greater precision. As discussed earlier, a data-integrating model may not provide benefit over the use of any single dataset unless each data source captures distinct, but complementary, aspects of the epidemiological or observation process (Moss et al., 2017). When deciding whether data integration is an appropriate and feasible approach to addressing a specific policy question, it is important for modellers to understand how datasets may be dependant and/or complementary. This process will be most insightful when done in close consultation with data-collectors and data-users.

6.3. Prior distributions informed by experts

Constructing models to aid decision making in public health

requires striking an appropriate balance between accurately representing the system being modelled, and making simplifying assumptions. It is often helpful to construct these models using domain knowledge. The Bayesian framework, with its concept of a prior distribution, provides a natural way to incorporate this information (Gelman et al., 2004). However, representing this knowledge into parameter values is not always straight-forward (Kuhnert et al., 2010). The process of comprehensively eliciting prior knowledge requires substantial effort from both the elicitor and the responder. However, as seen in the field of ecology, this can be a worthwhile approach (Kuhnert et al., 2005; Martin et al., 2005; Choy et al., 2009). Care must be taken to mitigate the impact of psychological biases such as 'anchoring' and the 'conjunction fallacy' which can lead to poor representation of prior knowledge (Kynn, 2008; Kuhnert et al., 2010). Moreover, if the responder is unfamiliar with probabilistic concepts this can also hamper the elicitation process (Morgan, 2014). Substantial effort in the fields of both statistics and psychology has produced a number of guidelines to assist in carrying out successful elicitations (Spetzler et al., 1975, Kynn, 2008; Choy et al., 2009; Kuhnert et al., 2010; Morgan, 2014; O'Hagan and Oakley, 2019).

As discussed in Section 3 above, there are also technical issues relating to the incorporation of the elicited information.

6.4. Sharable methodologies and best-practice for open and transparent modelling

To prevent reinventing previous work, modellers can provide adaptable, user-friendly routines or packages with worked examples supported by published theory and community-wide consensus. Reproducible methodologies and model source code can be shared on freely available code repositories such as GitHub [https://github.com/], Bitbucket [https://bitbucket.org], et cetera. There should be sufficient documentation provided with the source code in order to ensure reproducibility. Anonymised or simulated data could also be shared if necessary. Knowledge on the various types of licenses that are available to protect intellectual property while maintaining reproducibility is essential for modellers.

Model predictions could be shared through interactive web-applications such as Shiny apps [http://www.rstudio.com/shiny/] and plotly/Dash apps [https://plot.ly]. These web-applications can also help during the iterative process of communicating the model to policymakers and end-users, and improving the model based on their feedback. Computationally expensive models may not be able to run in real-time. In such cases, results could be pre-generated and stored as data which can then be retrieved for a specific scenario. Examples of some shiny apps based on models, together with their source code can be found here (Tun et al., 2017; Celhay et al., 2019).

6.5. Communicating model predictions in policymaking

When communicating model-based insights to policymakers, it is important to present those results in scientific and statistically rigorous language but also in a clear and transparent way to a non-technical audience. Ideally, results should be presented in a way that policymakers and stakeholders can quickly understand to assist their decisions. Various types of communication such as written briefings, informal meetings, and technical interfaces could be used as necessary. Uncertainty in the predictions should also be communicated adequately. The key to successful communication in this context however is to recall that policy development is primarily concerned with **decision-making**, rather than scientific discovery *per se*. With this in mind, scientific findings can be presented in a way that focuses on the decision to be made, providing the scientific evidence as the (rigorous and transparent) basis on which advice is provided.

7. Conclusions

From the very earliest attempts to represent disease dynamics mathematically, the relationship between epidemic models and data has been both of clear importance and a major challenge (Abbey, 1952). Herein we have focused on key challenges with a particular focus on using modelling to inform public health policy:

(i) Model parameters may not be able to be identified uniquely, either in general or for the specific data available in an application. Awareness of this is critical to ensure that robust policy conclusions are drawn from models;
(ii) Prior knowledge can be highly valuable, but it must pertain appropriately to the disease dynamics of current interest. Specification of prior knowledge on disease process observables assists in consistent and readily-interpretable specification; and,
(iii) Data must be modelled appropriately accounting for the observation process. Increasingly we have access, and computational resources, to exploit multiple datasets. However, the dependencies that arise through the underlying epidemiological system and in the sampling process in the observation models must be accounted for to draw robust conclusions.

We have discussed recent progress and new perspectives on each of these challenges, along with recent computational advances in methods with a particular focus on inference. Each of these areas remain active research topics, where advances are critical to improve the robustness, appropriateness and sophistication of model-based, policy-relevant outputs.

Acknowledgements

We thank the mathematical research institute MATRIX in Australia where part of this research was performed. AEZ is supported by the Oxford Martin School

References

Abbey, H., 1952. An examination of the Reed-Frost theory of epidemics. Hum Biol. 24 (September (3)), 201–233.
Adler, A.J., Eames, K.T., Funk, S., Edmunds, W.J., 2014. Incidence and risk factors for influenza-like-illness in the UK: online surveillance using Flusurvey. BMC Infect. Dis. 14 (May 1), 232. https://doi.org/10.1186/1471-2334-14-232.
Ali, H., Cameron, E., Drovandi, C.C., McCaw, J.M., Guy, R.J., Middleton, M., El-Hayek, C., Hocking, J.S., Kaldor, J.M., Donovan, B., 2015. Wilson DP; Australian chlamydia incidence estimation group. A new approach to estimating trends in chlamydia incidence. Sex Transm. Infect. 91 (November (7)), 513–519. https://doi.org/10.1136/sextrans-2014-051631.
Andrieu, C., Doucet, A., Holenstein, R., 2010. Particle Markov chain Monte Carlo methods. J. R. Stat. Soc. 72, 269–342. https://doi.org/10.1111/j.1467-9868.2009.00736.x.
Azmon, A., Faes, C., Hens, N., 2014. On the estimation of the reproduction number based on misreported epidemic data. Stat. Med. 33 (March 30 (7)), 1176–1192. https://doi.org/10.1002/sim.6015.
Ball, I.R., Possingham, H.P., Watts, M., 2009. In: Moilanen, A., Wilson, K.A., Possingham, H.P. (Eds.), Marxan and relatives: Software for spatial conservation prioritisation. Chapter 14: Pages 185-195 in Spatial conservation prioritisation: Quantitative methods and computational tools. Oxford University Press, Oxford, UK.
Bellman, R., Åström, K.J., 1970. On structural identifiability. Mathematical Biosciences 7 (3-4), 329–339. https://doi.org/10.1016/0025-5564(70)90132-X.
Bellu, G., Saccomani, M.P., Audoly, S., D'Angiò, L., 2007. DAISY: a new software tool to test global identifiability of biological and physiological systems. Comput. Methods Programs Biomed. 88 (October (1)), 52–61.
Birrell, Paul J., De, Angelis, Daniela, Presanis, Anne, M., 2018. Evidence synthesis for stochastic epidemic models. Statist. Sci. 33https://doi.org/10.1214/17-STS631. no. 1, 34-43.
Black, A.J., Geard, N., McCaw, J.M., McVernon, J., Ross, J.V., 2017. Characterising pandemic severity and transmissibility from data collected during first few hundred studies. Epidemics 19 (June), 61–73. https://doi.org/10.1016/j.epidem.2017.01.004.
Black, A.J., 2019. Importance sampling for partially observed temporal epidemic models. Stat. Comput. 29, 617. https://doi.org/10.1007/s11222-018-9827-1.
Boianelli, A., Nguyen, V., Ebensen, T., Schulze, K., Wilk, E., Sharma, N., Meyer-Hermann, M., 2015. Modeling influenza virus infection: a roadmap for influenza research.

Viruses 7 (10), 5274–5304. https://doi.org/10.3390/v7102875.
Brooks, Steve, Gelman, Andrew, Jones, Galin, Meng, Xiao-Li, 2011. Handbook of Markov Chain Monte Carlo. Chapman and Hall/CRC, New York.
Brooks, L.C., Farrow, D.C., Hyun, S., Tibshirani, R.J., Rosenfeld, R., 2015. Flexible Modeling of Epidemics with an Empirical Bayes Framework. PLoS Comput. Biol. 11 August 28 (8):e1004382. doi: 10.1371/journal.pcbi.1004382.
Brooks, L.C., Farrow, D.C., Hyun, S., Tibshirani, R.J., Rosenfeld, R., 2018. Nonmechanistic forecasts of seasonal influenza with iterative one-week-ahead distributions. PLoS Comput. Biol. 14 June 15 (6): e1006134. doi: https://doi.org/10.1371/journal.pcbi.1006134.
Campbell, F., Cori, A., Ferguson, N., Jombart, T., 2019. Bayesian inference of transmission chains using timing of symptoms, pathogen genomes and contact data. PLoS Comput Biol. 15 (March 29 (3)), e1006930. https://doi.org/10.1371/journal.pcbi.1006930.
Carlson, S.J., Dalton, C.B., Durrheim, D.N., Fejsa, J., 2010. Online Flutracking survey of influenza-like illness during pandemic (H1N1) 2009, Australia. Emerg Infect Dis. 16 (December (12)), 1960–1962. https://doi.org/10.3201/eid1612.100935.
Chaloner, K.M., Duncan, G.T., 1983. Assessment of a beta prior distribution: PM elicitation. J. R. Stat. Soc. 32 (1), 174–180.
Chis, O.T., Banga, J.R., Balsa-Canto, E., 2011. Structural identifiability of systems biology models: a critical comparison of methods. PLoS One. 6 (11), e27755. https://doi.org/10.1371/journal.pone.0027755.
Chopin, N., Jacob, P.E., Papaspiliopoulos, O., 2013. SMC^2: An efficient algorithm for sequential analysis of state space models. J. R. Stat. Soc. 75 (3), 397–426.
Choy, S.L., O'Leary, R., Mengersen, K., 2009. Elicitation by design in ecology: using expert opinion to inform priors for Bayesian statistical models. Ecology 90, 265–277. https://doi.org/10.1890/07-1886.1.
Celhay, O.J., Silal, S.P., Maude, R.J., et al., 2019. An interactive application for malaria elimination transmission and costing in the Asia-Pacific [version 2; peer review: 1 approved, 1 approved with reservations]. Wellcome Open Res. 4, 61. https://doi.org/10.12688/wellcomeopenres.14770.2.
Cooper, B., 2006. Poxy models and rash decisions. Proc. Natl. Acad Sci. U. S. A. 103 (August 15 (33)) 12221-2. Epub 2006 Aug 7.
Cope, R.C., Ross, J.V., 2020. Identification of the relative timing of infectiousness and symptom onset for outbreak control. J. Theor. Biol. 486, 110079.
Corbella, Alice, 2019. Statistical Inference in Stochastic/deterministic Epidemic Models to Jointly Estimate Transmission and Severity (Doctoral thesis). https://doi.org/10.17863/CAM.41539.
Craig, P.S., Goldstein, M., Seheult, A.H., Smith, J.A., 1997. Pressure matching for hydrocarbon reservoirs: a case study in the use of bayes linear strategies for large computer experiments. Case Studies in Bayesian Statistics. Springer, pp. 37–93.
De Angelis, A.M., Presanis, P.J., Birrell, G., Tomba, Scalia, House, T., 2015. Four key challenges in infectious disease modelling using data from multiple sources. Epidemics 10, 83–87.
De Maio, N., Worby, C.J., Wilson, D.J., Stoesser, N., 2018. Bayesian reconstruction of transmission within outbreaks using genomic variants. PLoS Comput. Biol. 14 (April 18 (4)), e1006117. https://doi.org/10.1371/journal.pcbi.1006117.
Gaydos, D.A., Petrasova, A., Cobb, R.C., Meentemeyer, R.K., 2019. Forecasting and control of emerging infectious forest disease through participatory modelling. Phil. Trans. R. Soc. B. 374 (1776). https://doi.org/10.1098/rstb.2018.0283.
Doms, C., Kramer, S.C., Shaman, J., 2018. Assessing the Use of influenza forecasts and epidemiological modeling in public health decision making in the United States. Sci Rep. 8 (August 17 (1)), 12406. https://doi.org/10.1038/s41598-018-30378-w.
Doucet, Arnaud, Freitas, Nando de, Gordon, Neil (Eds.), 2001. Sequential Monte Carlo Methods in Practice. Springer-Verlag, New York. https://doi.org/10.1007/978-1-4757-3437-9.
Doucet, A., Pitt, M.K., Deligiannidis, G., Kohn, R., 2015. Efficient implementation of Markov chain Monte Carlo when using an unbiased likelihood estimator. Biometrika 102 (June 2), 295–313. https://doi.org/10.1093/biomet/asu075.
Drescher, M., Perera, A.H., Johnson, C.J., Buse, L.J., Drew, A., Burgman, M.A., 2013. Towards rigorous use of expert knowledge in ecological research. Ecosphere. 4 (July 13 (7)), 1–26. https://doi.org/10.1890/ES12-00415.1.
Drovandi, C.C., McCutchan, R.A., 2016. Alive SMC^2: Bayesian model selection for low--count time series models with intractable likelihoods. Biom 72, 344–353. https://doi.org/10.1111/biom.12449.
Ertem, Z., Raymond, D., Meyers, L.A., 2018. Optimal multi-source forecasting of seasonal influenza. PLoS Comput Biol. 14 (September 4 (9)), e1006236. https://doi.org/10.1371/journal.pcbi.1006236.
Fithian, W., Elith, J., Hastie, T., Keith, D.A., 2015. Bias correction in species distribution models: pooling survey and collection data for multiple species. Methods Ecol. Evol. 6, 424–438. https://doi.org/10.1111/2041-210X.12242.
Gelfand, A.E., Dey, D.K., 1994. Bayesian model choice: asymptotics and exact calculations. J. R. Stat. Soc. 56 (3), 501–514.
Gelman, Andrew, Carlin, John B., Stern, Hal S., Rubin, Donald, B., 2004. Bayesian Data Analysis, 2nd ed. Chapman and Hall/CRC.
Gelman, A., Simpson, D., Betancourt, M., 2017. The prior can often only be understood in the context of the likelihood. Entropy 19, 555.
Glasser, J.W., Hupert, N., McCauley, M.M., Hatchett, R., 2011. Modeling and public health emergency responses: lessons from SARS. Epidemics 3 (Mar. (1)), 32–37. https://doi.org/10.1016/j.epidem.2011.01.001.
Godfrey, Keith., 1983. Compartmental Models and their Application." *Compartmental Models and their Application.* Academic Press.
Green, P.J., 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. Biometrika 82 (4), 711–732. https://doi.org/10.1093/biomet/82.4.711.
Green, P.J., Łatuszyński, K., Pereyra, M., Robert, C.P., 2015. Bayesian computation: a summary of the current state, and samples backwards and forwards. Stat. Comput. 25, 835. https://doi.org/10.1007/s11222-015-9574-5.
Markus, Hainy, David, J., 2018. Price, Olivier Restif, Christopher Drovandi, Optimal Bayesian Design for Model Discrimination Via Classification. [arXiv:1809.05301].
Hay, J.A., Nouvellet, P., Donnelly, C.A., Riley, S., 2018. Potential inconsistencies in Zika surveillance data and our understanding of risk during pregnancy. PLoS Negl. Trop. Dis. 12 (December 10 (12)), e0006991. https://doi.org/10.1371/journal.pntd.0006991.
Heesterbeek, H., Anderson, R.M., Andreasen, V., Bansal, S., De Angelis, D., Dye, C., Eames, K.T.D., Edmunds, W.J., Frost, S.D.W., Funk, S., Hollingsworth, T.D., House, T., Isham, V., Klepac, P., Lessler, J., Lloyd-Smith, J.O., Metcalf, C.J.E., Mollison, D., Pellis, L., Pulliam, J.R.C., Roberts, M.G., Viboud, C., Isaac Newton Institute IDD Collaboration, 2015. Modeling infectious disease dynamics in the complex landscape of global health. Science 347 6227 aaa4339.
Hemming, V., Burgman, M.A., Hanea, A.M., McBride, M.F., Wintle, B.C., 2017. A practical guide to structured expert elicitation using the IDEA protocol. Methods Ecol. Evol. 9 (July 30), 169–180.
House, T., Ford, A., Lan, S., Bilson, S., Buckingham-Jeffery, E., Girolami, M., 2016. Bayesian uncertainty quantification for transmissibility of influenza, norovirus and Ebola using information geometry. J. R. Soc. Interface 13 (121), 20160279.
Jacquez, John A., Greif, Peter, 1985. Numerical parameter identifiability and estimability: Integrating identifiability, estimability, and optimal sampling design. Math. Biosci. 77 (1-2), 201–227.
Kadane, J.B., Wolfson, L.J., 1998. Experiences in elicitation. J. R. Stat. Soc. 47 (1), 3–19.
Kantas, Nikolas, Doucet, Arnaud, Singh, Sumeetpal, S., Maciejowski, Jan, Chopin, Nicolas, 2015. On Particle methods for parameter estimation in state-space models. Statist. Sci. 30 (3), 328–351. https://doi.org/10.1214/14-STS511.
Karlsson, J., Anguelova, M., Jirstrand, M., 2012. An efficient method for structural identifiability analysis of large dynamic systems. IFAC Proc. Vol. 45 (16), 941–946.
Kass, R.E., Raftery, A.E., 1995. Bayes factors. J. Am. Stat. Assoc. 90 (430), 773–795. https://doi.org/10.1080/01621459.1995.10476572).
Kleijn, B.J.K., van der Vaart, A.W., 2012. The Bernstein-Von-Mises theorem under misspecification. Electron. J. Stat. 6, 354–381. https://doi.org/10.1214/12-EJS675. The Institute of Mathematical Statistics and the Bernoulli Society.
Knight, G.M., Dharan, N.J., Fox, G.J., Stennis, N., Zwerling, A., Khurana, R., Dowdy, D.W., 2016. Bridging the gap between evidence and policy for infectious diseases: how models can aid public health decision-making. Int. J. Infect. Dis. 42, 17–23. https://doi.org/10.1016/j.ijid.2015.10.024.
Kuhnert, P.M., Martin, T.G., Mengersen, K., Possingham, H.P., 2005. Assessing the impacts of grazing levels on bird density in woodland habitat: a Bayesian approach using expert opinion. Environmetrics 16, 717–747. https://doi.org/10.1002/env.732.
Kuhnert, P.M., Martin, T.G., Griffiths, S.P., 2010. A guide to eliciting and using expert knowledge in Bayesian ecological models. Ecol. Lett. 13, 900–914. https://doi.org/10.1111/j.1461-0248.2010.01477.x.
Kypraios, T., Neal, P., Prangle, D., 2017. A tutorial introduction to Bayesian inference for stochastic epidemic models using Approximate Bayesian Computation. Math. Biosci. 287, 42–53. https://doi.org/10.1016/j.mbs.2016.07.001.
Kynn, Mary., 2008. The' Heuristics and Biases' Bias in Expert Elicitation. J. R. Stat. Soc. 171 (1), 239–264 *Series A (Statistics in Society)*.
Lee, B.Y., Haidari, L.A., Lee, M.S., 2013. Modelling during an emergency: the 2009 H1N1 influenza pandemic. Clin. Microbiol. Infect. 19, 1014–1022.
Low Choy, S., O'Leary, R., Mengersen, K., 2009. Elicitation by design in ecology: using expert opinion to inform priors for Bayesian statistical models. Ecology 90 (January 1 (1)), 265–277. https://doi.org/10.1890/07-1886.1.
Martin, T.G., Kuhnert, P.M., Mengersen, K., Possingham, H.P., 2005. The power of expert opinion in ecological models using Bayesian methods: Impact of grazing on birds. Ecol. Appl. 15, 266–280. https://doi.org/10.1890/03-5400.
Martin, T.G., Burgman, M.A., Fidler, F., Kuhnert, P.M., Low-Choy, S., McBride, M., Mengersen, K., 2012. Eliciting expert knowledge in conservation science. Conserv. Biol. 26 (January 26 (1)), 29–38. https://doi.org/10.1111/j.1523-1739.2011.01806.x.
MATRIX Institute, 2019. https://www.matrix-inst.org.au/events/influencing-public-health-policy-with-data-informed-mathematical-models-of-infectious-diseases/ (accessed July 2019).
MacKay, D.J.C., 2003. Information Theory, Inference, and Learning Algorithms. Cambridge University Press.
McKinley, T.J., Ross, J.V., Deardon, R., Cook, A.R., 2014. Simulation-based Bayesian inference for epidemic models. Comput. Stat. Data Anal. 71, 434–447. https://doi.org/10.1016/j.csda.2012.12.012.
McLean, E., Pebody, R.G., Campbell, C., Chamberland, M., Hawkins, C., Nguyen-Van-Tam, J.S., Oliver, I., Smith, G.E., Ihekweazu, C., Bracebridge, S., Maguire, H., Harris, R., Kafatos, G., White, P.J., Wynne-Evans, E., Green, J., Myers, R., Underwood, A., Dallman, T., Wreghitt, T., Zambon, M., Ellis, J., Phin, N., Smyth, B., McMenamin, J., Watson, J.M., Pandemic (H1N1), 2010. 2009 influenza in the UK: clinical and epidemiological findings from the first few hundred (FF100) cases. Epidemiol. Infect. 138 (November (11)), 1531–1541. https://doi.org/10.1017/S0950268810001366.
Meshkat, N., Kuo, C.E., DiStefano 3rd, J., 2014. On finding and using identifiable parameter combinations in nonlinear dynamic systems biology models and COMBOS: a novel web implementation. PloS one 9 (10), e110261. https://doi.org/10.1371/journal.pone.0110261.
Miller, Joel, C., 2012. A note on the derivation of epidemic final sizes. Bull. Math. Biol. 74 (9), 2125–2141. https://doi.org/10.1007/s11538-012-9749-6. https://pubmed.ncbi.nlm.nih.gov/22829179.
Morgan, M.G., 2014. Use (and abuse) of expert elicitation in support of decision making for public policy. Proc. Natl. Acad. Sci. USA 111 May 20 (20): doi: 10.1073/

pnas.1319946111. Epub 2014 May 12. 7176-84.
Moss, R., Zarebski, A., Dawson, P., McCAW, J.M., 2017. Retrospective forecasting of the 2010-2014 Melbourne influenza seasons using multiple surveillance systems. Epidemiol. Infect. 145 (January (1)), 156–169 Epub 2016 Sep 27. PubMed PMID: 27671159.
Moss, R., Zarebski, A.E., Dawson, P., Franklin, L.J., Birrell, F.A., McCaw, J.M., 2019a. Anatomy of a seasonal influenza epidemic forecast. Commun. Dis. Intell. 43 (Mar 15). https://doi.org/10.33321/cdi.2019.43.7.
Moss, R., Zarebski, A.E., Carlson, S.J., McCaw, J.M., 2019b. Accounting for Healthcare-Seeking Behaviours and Testing Practices in Real-Time Influenza Forecasts. Trop. Med. Infect. Dis. 4 January 11 (1). pii: E12. doi: 10.3390/tropicalmed4010012.
Muscatello, D.J., Chughtai, A.A., Heywood, A., Gardner, L.M., Heslop, D.J., MacIntyre, C.R., 2017. Translation of real-time infectious disease modeling into routine public health practice. Emerg. Infect. Dis. 23 (May (5)). https://doi.org/10.3201/eid2305.161720.
Nguyen, Van Kinh, et al., 2016. Analysis of practical identifiability of a viral infection model. PLOS One 11 (12), e0167568. https://doi.org/10.1371/journal.pone.0167568. 30 Dec.
O'Hagan, A., 1998. Eliciting expert beliefs in substantial practical applications. J. R. Stat. Soc. 47 (1), 21–35.
O'Hagan, T., Oakley, J., 2019. SHELF: The Sheffield Elicitation Framework. Visited Jul 31. http://www.tonyohagan.co.uk/shelf/.
O'Neill, P.D., Roberts, G.O., 1999. Bayesian inference for partially observed stochastic epidemics. J. R. Stat. Soc. 162, 121–129. https://doi.org/10.1111/1467-985X.00125.
Pearl, J., 2001. Bayesianism and Causality, or, Why I am Only a Half-Bayesian. In: Corfield, D., Williamson, J. (Eds.), Foundations of Bayesianism. Applied Logic Series. Springer, Dordrecht vol 24.
Peppa, M., John Edmunds, W., Funk, S., 2017. Disease severity determines health-seeking behaviour amongst individuals with influenza-like illness in an internet-based cohort. BMC Infect. Dis. 17 (March 31 (1)), 238. https://doi.org/10.1186/s12879-017-2337-5.
Pierre, Pudlo, Jean-Michel, Marin, Arnaud, Estoup, Jean-Marie, Cornuet, Mathieu, Gautier, Christian, P. Robert, 2016. Reliable ABC model choice via random forests. Bioinformatics 32 (March 6), 15. https://doi.org/10.1093/bioinformatics/btv684. 859-15866.
Probert, W.J.M., Shea, K., Fonnesbeck, C.J., Runge, M.C., Carpenter, T.E., Dürr, S., Garner, G., Harvey, N., Stevenson, M.A., Webb, C.T., Werkman, M., Tildesley, M.J., Ferrari, M.J., 2016. Decision-making for foot-and-mouth disease control: objectives matter. Epidemics 15, 10–19. https://doi.org/10.1016/j.epidem.2015.11.002.
Qualls, N., Levitt, A., Kanade, N., Wright-Jegede, N., Dopson, S., Biggerstaff, M., Reed, C., Uzicanin, A., Frank, M., Holloway, R., Koonin, L., Rasmussen, S., Redd, S., Hurst, C., Rainey, J., Jernigan, D., 2017. Community mitigation guidelines to prevent pandemic influenza — United States. MMWR. Recommendations Rep. 66, 1–34. https://doi.org/10.15585/mmwr.rr6601a1.
Raue, A., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U., Timmer, J., 2009. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. Bioinformatics 25 (August 1 (15)), 1923–1929. https://doi.org/10.1093/bioinformatics/btp358.
Reich, N.G., Perl, T.M., Cummings DAT, Lessler J, 2011. Visualizing clinical evidence: citation networks for the incubation periods of respiratory viral infections. PLoS ONE 6 (4), e19496. https://doi.org/10.1371/journal.pone.0019496.
Rosello, A., Mossoko, M., Flasche, S., Van Hoek, A.J., Mbala, P., Camacho, A., Funk, S., Kucharski, A., Ilunga, B.K., Edmunds, W.J., Piot, P., Baguelin, M., Tamfum, J.J., 2015. Ebola virus disease in the Democratic Republic of the Congo, 1976-2014. Elife 4 November 3 pii: e09015. doi: 10.7554/eLife.09015.
Shubin, M., Lebedev, A., Lyytikäinen, O., Auranen, K., 2016. Revealing the true incidence of pandemic A(H1N1)pdm09 influenza in finland during the first two seasons - an analysis based on a dynamic transmission model. PLoS Comput. Biol. 12 (March 24 (3)), e1004803. https://doi.org/10.1371/journal.pcbi.1004803.
Sherlock, Chris, Thiery, Alexandre, Roberts, H., Gareth, O., Rosenthal Jeffrey, S., 2015. On the efficiency of pseudo-marginal random walk Metropolis algorithms. Ann. Statist. 43 (1), 238–275. https://doi.org/10.1214/14-AOS1278.
Siekmann, I., Sneyd, J., Crampin, E.J., 2012. MCMC can detect nonidentifiable models. Biophys. J. 103 (December 5 (11)), 2275–2286. https://doi.org/10.1016/j.bpj.2012.10.024.
Spetzler, Carl S., Carl-Axel, S., Von Holstein, Staël, 1975. Exceptional Paper—Probability Encoding in Decision Analysis. Manage. Sci. 22 (3), 340–358.
Stockdale, J.E., Kypraios, T., O'Neill, P.D., 2017. Modelling and Bayesian analysis of the Abakaliki smallpox data. Epidemics 19 (June), 13–23. https://doi.org/10.1016/j.epidem.2016.11.005.
Thomas, E.G., McCaw, J.M., Kelly, H.A., Grant, K.A., McVernon, J., 2015. Quantifying differences in the epidemic curves from three influenza surveillance systems: a nonlinear regression analysis. Epidemiol. Infect. 143 (January (2)), 427–439. https://doi.org/10.1017/S0950268814000764.
Touloupou, P., Alzahrani, N., Neal, P., Spencer, S.E.F., McKinley, T.J., 2018. Efficient model comparison techniques for models requiring large scale data augmentation. Bayesian. Anal. 13 (2), 437–459. https://doi.org/10.1214/17-BA1057.
Tun, S.T.T., von Seidlein, L., Pongvongsa, T., Mayxay, M., Saralamba, S., Kyaw, S.S., Chanthavilay, P., Celhay, O., Nguyen, T.D., Tran, T.N., Parker, D.M., Boni, M.F., Dondorp, A.M., White, L.J., 2017. Towards malaria elimination in Savannakhet, Lao PDR: mathematical modelling driven strategy design. Malar. J. 16 (November 28 (1)), 483. https://doi.org/10.1186/s12936-017-2130-3.
Vajda, Sàndor., 1984. Structural equivalence and exhaustive compartmental modeling. Math. Biosci. 69 (1), 57–75.
Whyte, J.M., 2013. Inferring global a priori identifiability of optical biosensor experiment models. In 2013 IEEE International Conference on Bioinformatics and Biomedicine 17–22. https://doi.org/10.1109/BIBM.2013.6732453. IEEE.
Yan, A., Zaloumis, S.G., Simpson, J.A., McCaw, J.M., 2019. Sequential infection experiments for quantifying innate and adaptive immunity during influenza infection. PLoS computat. Boil. 15 (1), e1006568. https://doi.org/10.1371/journal.pcbi.1006568.