Anchoring the mean generation time in the SEIR to mitigate biases in $\Re_0$ estimates due to uncertainty in the distribution of the epidemiological delays

A Preprint

Jair Andrade *
Data Science Institute and School of Computer Science
University of Galway
Galway
jair.andrade@nuigalway.ie

Jim Duggan
Data Science Institute and School of Computer Science
University of Galway
Galway
jim.duggan@universityofgalway.ie

January 4, 2023

Abstract

The basic reproduction number, $\Re_0$, is of paramount importance in the study of infectious disease dynamics. Primarily, $\Re_0$ serves as an indicator of the transmission potential of an emerging infectious disease and the effort required to control the invading pathogen. However, its estimates from compartmental models are strongly conditioned by assumptions in the model structure, such as the distribution of the latent and infectious periods (epidemiological delays). To further complicate matters, models with dissimilar delay structures produce equivalent incidence dynamics. Following a simulation study, we reveal that the nature of such equivalency stems from a linear relationship between $\Re_0$ and the mean generation time, along with adjustments to other parameters in the model. Leveraging this knowledge, we propose and successfully test an alternative parameterisation of the SEIR model that produces accurate $\Re_0$ estimates regardless of the distribution of the epidemiological delays, at the expense of biases in other quantities deemed of lesser importance. We further explore this approach's robustness by testing various transmissibility levels, generation times, and data fidelity (overdispersion). Finally, we apply the proposed approach to data from the 1918 influenza pandemic. We anticipate that this work will mitigate potentially large errors in the estimation of $\Re_0$.

***K*eywords** SEIR · Basic reproduction number · Distribution of the latent & infectious period · Generation time · Stan · Inference

1 Introduction

The analysis of any infectious disease's dynamics will inevitably lead to the basic reproduction number ($\Re_0$). Initially developed in the study of demographics (1), this quantity has been interpreted in the epidemiological context as the average number of secondary infections arising from the introduction of one infected individual into a totally susceptible population (2). The usefulness and importance of $\Re_0$ lie primarily in its *threshold*

*Corresponding author

phenomenon (3). That is, a pathogen can invade a totally susceptible population only if $\Re_0 > 1$ (4). Furthermore, the magnitude of R_0 gauges the transmission potential of an emerging infectious disease (3) and the effort required to control the invading pathogen (5). Thus, accurate estimation of R_0 is crucial for understanding and managing infectious diseases.

Another reason for the popularity of $\Re_0$ is that one can estimate it from epidemiological data (6) using a number of methods. For diseases that allow the assumption of endemic equilibrium and homogeneous mixing, one can follow Mollison's method (7) or Dietz's approach (8). The former requires prevalence data, whereas the latter leverages readily available information such as age at infection and average life expectancy. On the other hand, should the assumption of an infection leading to either immunity or death in a closed population be tenable, seroprevalence studies can inform the fraction of the population that acquired the disease during an epidemic, i.e. the *final epidemic size*. In their seminal paper, Kermack and McKendrick (9,10) formulated a relationship between the final epidemic size and R_0, from which the latter can be calculated. Unlike the previous methods, which require the epidemic to reach a steady state, R_0 may be determined from the intrinsic growth rate of the infected population (3,5) using incidence data of the early stages of the epidemic, as long as the growth of new cases exhibits *pure* exponential growth. Alternatively, we can employ the entire report of daily case notifications if R_0 is formulated as a function of a *compartmental* model's parameters (11–13). These models can be stochastic (14), semi-deterministic (15,16) or deterministic (9,17).

These compartmental models are said to be *mechanistic* (18), namely, their structures are based on a scientific understanding of infectious disease dynamics (19). The relevance of that mechanistic property lies in the role of the model. Rather than being a merely mathematical artefact to produce a desired output, the model also embeds a dynamic hypothesis of the underlying process that generates the observed data. Hence, the parameters, states and interactions that comprise a particular formulation represent their counterparts in the real world. If the model accurately captures the properties of the actual phenomenon, finding an adequate configuration (assign values to parameters) should yield a behaviour over time that resembles the observed trajectory. The values of such parameters can be obtained from individual-level observations (4) or via statistical inference (20–22), a process also known as *trajectory matching* or *model fitting*.

Furthermore, trajectory matching between simulated and observed behaviour can be regarded as a validation test on the dynamic hypothesis that links structure to behaviour (23). Nevertheless, one should understand this validation step as a *falsification* test (24). That is, if the model fails to reproduce the observed behaviour, it can certainly be rejected. On the contrary, obtaining an accurate match (or fit) does not immediately lead to accepting the dynamic hypothesis inasmuch as there may be other competing hypotheses that fit the data equally well. Indeed, this circumstance impacts the estimation of $\Re_0$ from compartmental models (and the intrinsic growth rate method), where different assumptions can yield accurate fits (25). However, estimates vary according to the specific assumptions in each fitting model (3,26).

For instance, the choice of the distribution of the latent and infectious periods (epidemiological delays) in the deterministic Susceptible-Exposed-Infectious-Recovered (SEIR) framework plays an essential role in the inference of $\Re_0$ (25). Briefly put, misspecifying the structure of such delays leads to biases in the estimates. That is, a systematic difference between true and estimated parameters. Although there are techniques (27,28) to construct models with realistic distributions, modellers do not know exactly which distribution to incorporate in their formulation. In view of this drawback, Wearing and colleagues (25) fitted various SEIR models (with different delay distributions) to a single incidence dataset to select the *best* structure based on a goodness-of-fit measure. Nevertheless, the results appear inconclusive. Notwithstanding that Krylova & Earn (29) assume their validity, no further research establishes the reliability of such an approach. This assessment immediately warrants the need for the work presented here: a systematic study oriented to determine whether it is possible to infer $\Re_0$ accurately from SEIR models fitted to incidence data in light of the uncertainty in the distributions of the epidemiological delays. We describe the steps of this study in the sections below. All the analysis is performed in R. The code is freely available at `https://github.com/jandraor/delays`.

2 Data Generating Process

In order to undertake a systematic study, experimenters must have access to a sizeable set of observations. In this case, they require multiple time series of daily case notifications of a particular disease under various conditions. Equally important, such conditions need to be known. Recall that this work aims to evaluate the possibility of inferring R_0 accurately. Then, it follows that such quantity must be known *a priori*. However, naturally-occurring outbreaks do not meet this last requirement. To further complicate matters, controlled experiments are less than desirable inasmuch as they would be costly to run, not to mention unethical. One

solution to circumvent this difficulty is leveraging the mechanistic property of compartmental models to employ them as *synthetic data generators* (30,31). Although appealing, this approach raises an immediate concern. Mathematical models, or in the words of Savage (32), *small worlds*, only capture a part of the reality (*large world*) that we aim to understand (33). However, no verification procedure confirms that a *small world* is able to capture the relevant aspects of the *large world*. In practice, we perform a wide range of validation procedures (23) in order to build *trust* in the model's ability to represent the key features of the studied phenomenon.

2.1 The system (latent) model

In the case of the SEIR framework, an extensive body of research permits us to build trust in its relevance. This framework has been widely applied to studying various infectious diseases, such as measles (29,34,35), COVID-19 (16,36–38), and influenza (12,17,22,39–41). Following this literature, it is reasonable to employ the SEIR as the basis for the synthetic data generator. In this work, we restrict our attention to the simplest versions of this family of models. The rationale for this decision is straightforward; conceptual models entail efficiency inasmuch as they facilitate the understanding and identification of the underlying causes of a particular result. Moreover, it is often the case that principles that stem from conceptual models apply to more elaborated extensions.

We begin by describing the *standard* SEIR model (Eq 1). This *within-host profile* (4) stratifies individuals as susceptible (S_t), exposed (E_t), infectious (I_t), and recovered (R_t), and describes the transitions between states ($S_t \to E_t \to I_t \to R_t$) in terms of differential equations. Susceptible individuals acquire infection, $S_t \to E_t$, through contact with infectious individuals, where the number of contacts is independent of the population size (N). Formally, one refers to this assumption as the frequency-dependent (or mass action) transmission: $\beta S_t I_t / N$. Here, β corresponds to the effective contact rate, or transmission parameter, which is given by the product of the contact rate and the transmission probability (11). The movement of individuals from the class E_t up to class R_t is modelled with transitions governed by constant per capita rates. In the case of the transition $E_t \to I_t$, the rate of onset of infectiousness (σ) modulates the flow of individuals from one state to the other. Mathematically, σE_t. Likewise, the recovery rate (γ) governs the transition $I_t \to R_t$, namely $\gamma I(t)$. Furthermore, we assume that the disease leads to permanent immunity and that the outbreak's time scale is much faster than the characteristic times for demographic processes (births and deaths), so their effects are not included. This last assumption implies that the population remains constant over the simulation period.

$$
\begin{aligned}
\dot{S} &= -\frac{\beta S_t I_t}{N} \\
\dot{E} &= \frac{\beta S_t I_t}{N} - \sigma E_t \\
\dot{I} &= \sigma E_t - \gamma I_t \\
\dot{R} &= \gamma I_t
\end{aligned}
\tag{1}
$$

Further, the assumption of becoming infectious ($E_t \to I_t$) at a constant rate and regardless of the time since infection is equivalent to formulating the latent period in terms of an exponential distribution. This formulation implies that, on average, individuals become infectious after σ^{-1} days of acquiring the disease. Similarly, assuming a constant recovery rate implies that the infectious period is exponentially distributed across the population with mean γ^{-1} (5). While mathematically convenient, the assumption of exponentially-distributed delays is epidemiologically unrealistic (25). Research has shown that the empirical distributions of the latent and infectious periods (42,43) have smaller variances than those predicted by the exponential distribution. In other words, the exponential distribution predicts that a many individuals leave their current state (E_t or I_t) almost immediately after entering and that a sizeable number of individuals have epidemiological delays that are much longer than the average. On the contrary, latent and infectious periods observed in the real world are closely centred about their mean (11).

A well-known mathematical procedure (44) to achieve tighter distributions within the compartmental framework is the subdivision of a class into stages arranged in series. For instance, one can divide the exposed class into i stages. Newly infected individuals enter the first exposed stage, E_t^1, pass through each in turn and become infectious upon leaving the *ith* stage (E_t^i). The progression between stages is assumed to occur at a constant per-capita rate ($i\sigma$), leading to an exponential waiting time with mean $\frac{1}{i\sigma}$ in each stage (11). This formulation implies that the lapse between infection and becoming infectious is described by the sum

of i independent exponential random variables with equal rates. Recall that this convolution results in a gamma-distributed random variable (45). Therefore, the subdivision of the exposed class into various stages is equivalent to formulate the latent period in terms of a gamma distribution with mean σ^{-1} and shape i. Similarly, one can divide the infectious class into j stages to formulating a gamma-distributed infectious period. Overall, we refer to this formulation as the SE^iI^jR framework (Eq 2). Notice that the standard SEIR corresponds to the SE^1I^1R instance. Moreover, as the parameter i increases, the distribution becomes more closely centred on its mean, to the extent that if $i \to \infty$, the variance is removed. That is, in the limit, all individuals have the same latent period (a similar argument applies to the infectious period). No less important, irrespective of the values of i and j that the SE^iI^jR may take, the basic reproduction number depends exclusively on the transmission rate and the mean infectious period (Eq 3). Lastly, it is noteworthy to mention that subdividing a class is a mathematical device that allows the incorporation of additional distributions in a system of differential equations, and the number of stages may not correspond to biological features of the infection process (34).

$$\begin{aligned}
\dot{S}_t &= -\frac{\beta S_t \sum_{k=1}^{i} I_t^k}{N} \\
\dot{E}_t^1 &= \frac{\beta S_t \sum_{k=1}^{i} I_t^k}{N} - i\sigma E_t^1 \\
\dot{E}_t^2 &= i\sigma E_t^1 - i\sigma E_t^2 \\
&\vdots \\
\dot{E}_t^i &= i\sigma E_t^{i-1} - i\sigma E_t^i \\
\dot{I}_t^1 &= i\sigma E_t^{i-j} - \gamma I_t^1 \\
\dot{I}_t^2 &= j\gamma I_t^{1-j} - \gamma I_t^2 \\
&\vdots \\
\dot{I}_t^j &= j\gamma I_t^{j-1} - j\gamma I_t^j \\
\dot{R}_t &= j\gamma I_t^j
\end{aligned} \tag{2}$$

$$\Re_0 = \beta\gamma^{-1} \tag{3}$$

Subsequently, we define the link between the SE^iI^jR and the incidence data reported by surveillance systems (Eq 4). Based on the literature (15,16,22,39), we posit that incidence ($\dot{C}$) is proportional to the rate at which individuals become infectious ($E_t^i \to I_t^1$). Such proportional effect or reporting rate (ρ) stems from the fact that individuals experience various degrees of symptom severity (46). In particular, individuals with low severity levels (asymptomatic and mild symptoms) may not seek health care attention, resulting in case reports that most likely miss a significant fraction of the infected individuals. As opposed to the continuous nature of differential equation models, case notifications occur at discrete times. To reconcile this *tension*, we define the report of new cases (x_{t^*}) as the change in the total number of cases (C_t) in one-day intervals.

$$\begin{aligned}
\dot{C}_t &= \rho i\sigma E_t^i \\
x_{t^*} &= C_{t^*+1} - C_{t^*}, t^* \in \mathbb{N}_0
\end{aligned} \tag{4}$$

As mentioned above, one of the key features of epidemiological modelling is that many infectious diseases conform to the SEIR profile (4). From a modelling perspective, this feature entails that, for instance, only parameter configuration differentiates measles and pandemic influenza models. In this work, we tailor the synthetic data generator to the latter, given that influenza pandemics are unpredictable but recurring events that can have significant global consequences (47). As a matter of fact, there have been four pandemics of influenza over the past 100 years, including the H1N1 pandemic in 1918, with 50 million deaths estimated (48). Adapting the SE^iI^jR framework to this choice involves the selection of plausible parameter values or *ground truths* (49). For simplicity, we restrict the synthetic data generator to eight instances: $i = \{1, 3\} \times j = \{1, 2, 3, 4\}$. These instances share constants σ, γ, β, ρ, and N, which are configured identically. In particular, we configure parameters σ and γ from the assumed values ($\frac{1}{2}$ for both) in the Cumberland case study (12,22). Following this choice, we select a value of β that yields a basic reproduction

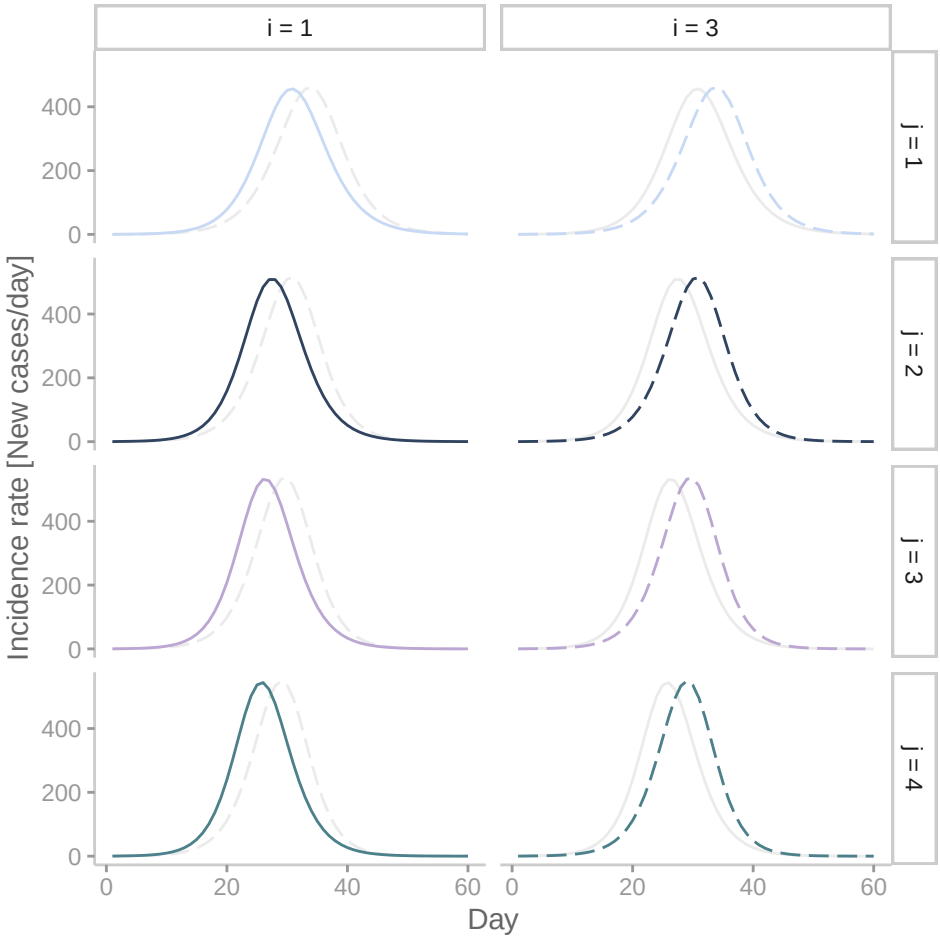


Figure 1: Incidence reports generated by various instances of the SE^iI^jR framework. In this plot, we present two distributions of the latent period and four distributions of the infectious period. The colour of a line corresponds to a particular value of j (infectious period distribution). Solid lines indicate that the incidence report stems from an SEIR model whose latent period is exponentially distributed ($i = 1$). Dashed lines indicate that the incidence report stems from an SEIR model whose latent period is gamma distributed ($i = 3$). In order to facilitate the comparison of two panels in a single row, we add the adjecent incidence report (grey lines).

number (*2.5*) within a plausible range (*2-4*) of pandemic influenza (50). Regarding ρ, we choose a value (*0.75*) consistent with reported estimates in the literature (12,22). The remaining constant, N, has only a scaling effect, and any particular value (*10,000* in this case) does not alter the model dynamics provided that $N = S_0 + E_0 + I_0 + R_0$, where $E_0 = \sum_{k=1}^{i} E_0^k$ and $I_0 = \sum_{k=1}^{j} I_0^k$. In relation to initial conditions, we assume that a *patient zero* triggers the outbreak of a novel influenza pathogen. In mathematical terms, $S_0 = N - 1$ and $I_0^1 = 1$. The remaining initial conditions of the within-host profile are set to zero.

Having delimited the SE^iI^jR framework and configured its instances, we run simulations (Fig 1) that illustrate the impact of the delay structure on the incidence dynamics. In agreement with the literature (4,25), note in Fig 1 that if we fix the distribution of the latent period (single column), incidence reports that stem from more tightly distributed infectious periods (larger j) reach the incidence peak earlier and end more abruptly. This difference in behaviour over time occurs despite the fact that the instances share identical $\Re_0$ and the same average latent and infectious periods. On the other hand, if we fix the infectious period (single row), decreasing the variance (increasing i) of the gamma-distributed latent period produces the opposite effect. Namely, tighter distributions of the latent period (larger i) push forward the peak time and extend the outbreak's duration.

2.2 Measurement model

Reflecting on the role of the output produced by the SE^iI^jR framework, notice that such *smooth* trajectories result from simulating a set of self-consistent causal assumptions (equations) that explains the reported incidence. Borrowing terminology from the state-space framework (18,51), we posit that such reports are predictions obtained from a *system* or *latent* process model. In practice, though, observed incidence by public health surveillance is somewhat different from predictions generated by system models. However the level of detail a mechanistic formulation can attain, a model is merely a partial representation of a more complex reality, and some elements are necessarily omitted. Consequently, it is required to equip the data-generating process with a structure that accounts for the discrepancies between model prediction and actual data. We refer to this structure as the measurement or *observational* model.

In epidemiology, it is recommended (19) to formulate the measurement model for the observation of new infections via the *Negative Binomial* distribution. This recommendation arises from the fact that such discrete probability distribution allows higher variability in the data compared to other alternatives (overdispersion), such as the *Poisson* and *Binomial* distributions while relaxing the assumption of equal variance across measurements (as opposed to the *Normal* distribution). Therefore, we define the observation of new cases (Y_{t^*}) in terms of a Negative Binomial distribution (Eq 5) specified by location and diffusion parameters. The former corresponds to the predicted incidence by the system model (x_{t^*}), whereas the latter (ϕ) modulates the concentration of measurements. Note that the inverse of the concentration parameter (ϕ^{-1}) represents overdispersion inasmuch as an increase in its magnitude leads to greater diffusion in the data.

$$Y_{t^*} \sim Nbin(x_{t^*}, \phi) \tag{5}$$

Defining a measurement model allows us to generate synthetic data via statistical simulation. We turn the measurement model into a sampling distribution (52) by fixing the right-hand side of Eq 5 to the predicted incidence from the system model (Fig 1) and ϕ^{-1} to a particular value. For each model instance, we generate *20 noisy* time series. We perform this process for two levels (*high* and *low*) of data fidelity, a feature measured by ϕ^{-1}. High-fidelity data ($\phi^{-1} = 0$) implies that the measurement model applies only a slight distortion on the original signal (incidence). Notice that this configuration of the Negative Binomial (with no overdispersion) is equivalent the Poisson distribution. Conversely, a positive value (overdispersion) of ϕ^{-1} (such as *1/3*) distorts the original signal to such an extent that one cannot easily discern the underlying incidence dynamics (low-fidelity data). In total, we generate *320* incidence reports. The reader can find the complete details in the electronic supplementary material S1. Fig 2 displays half of the generated synthetic reports. To facilitate the communication of results, we introduce the notation D^{ij}, which indicates the origin of a given set of time series. To provide an example, D^{14} indicates that the observed incidence was obtained from the SE^1I^4R instance.

3 Inference

Collecting a sizeable set of synthetic observations allows us to progress to the inference stage. Here, the goal is to assess the performance of various candidate models fitted to the synthetic data. Performance in this context refers to the ability to recover ground truths, particularly $\Re_0$, our quantity of interest. We fit such model candidates to the incidence reports following a *Bayesian* approach (22,53). That is, each candidate's unknown parameters are treated as random variables, which describe the knowledge (or uncertainty) about their actual values (54), expressed in terms of a probability distribution. This distribution is updated in light of new information summarised by a likelihood function. Recall that the likelihood function is the flip side of the sampling distribution (52). In other words, the left-hand side of the measurement model is fixed. Such updating process yields the target or posterior distribution, an *information device* whereby we derive answers for our inferential questions. We approximate the posterior distribution via sampling using the Hamiltonian Monte Carlo (HMC) algorithm (55) provided by the statistical package *Stan* (56).

3.1 Three unknowns (traditional)

Starting from a low degree of complexity, we initially restrict the inference analysis to D^{1j} high-fidelity observations (Fig 2). To fit each incidence report, we postulate four instances, $j = \{1, 2, 3, 4\}$, of the SE^1I^jR framework, which share identical mean latent and infectious periods. We refer to the approach of fixing the means of the epidemiological delays to values obtained from the literature, regardless of their distribution, as the *traditional* parameterisation. Moreover, it is assumed that the measurement model is fully known.

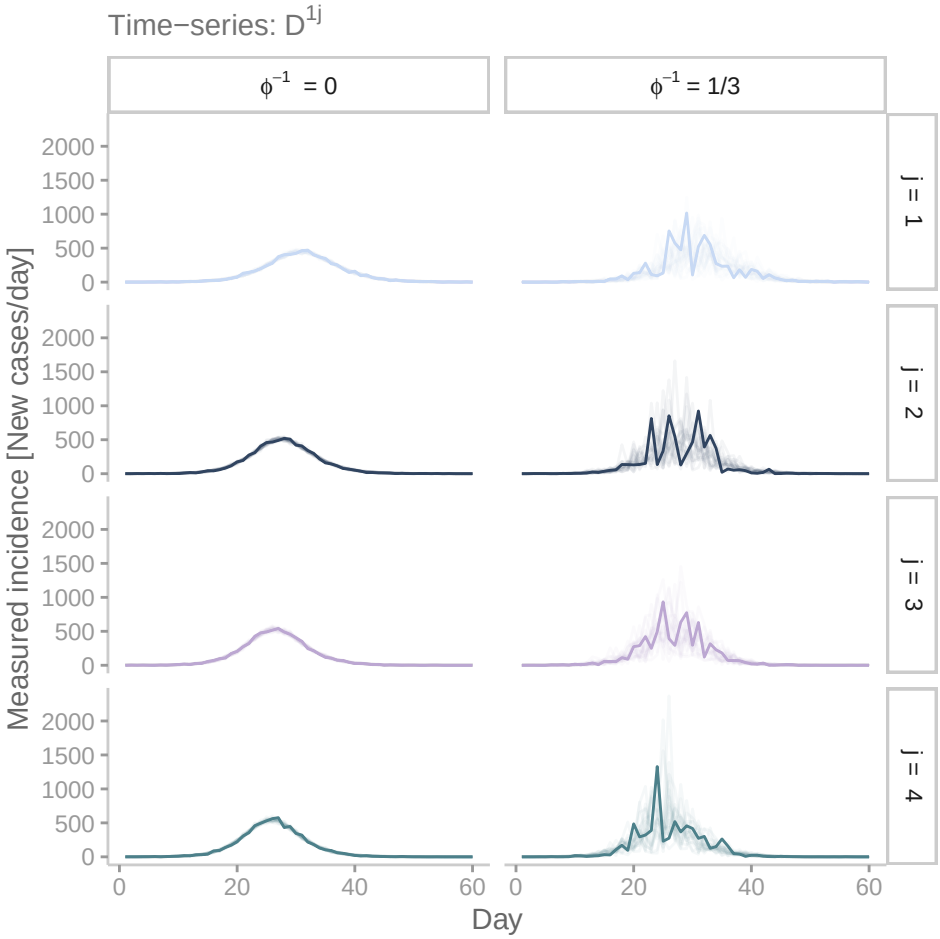


Figure 2: Synthetic data for the inference study. Each panel shows 20 noisy incidence reports, where one random time-series is highlighted. For instance, the top-left panel corresponds to sampled incidence reports from an SE^1I^1R with no overdispersion (Poisson) in the measurement model.

Consequently, discrepancies between estimated and actual values are ascribed to misspecification in the infectious period distribution. To avoid confusion between the origin of data and the fitting model, we denote the latter as M^{ij}. As a consequence, this design requires the estimation of *320* posterior distributions. Given this process's computational burden, we limit the number of random variables in each model to three: the transmission rate (β), the reporting rate (ρ) and the initial number of infected individuals in stage one (I_0^1). The remaining parameters and initial conditions are considered to be known, i.e. they are fixed to their actual values. Based on this setup, we fit each candidate to a given dataset using HMC sampling, with *four* Markov chains and 1000 iterations (plus 1000 for warm-up) each, checking for convergence and effective sample sizes. The complete set of results can be found in the electronic supplementary material *S2 §1*.

The results presented in Fig 3 replicate a finding reported in the literature (25,57): the existence of a subtle yet fundamental interaction between the assumed model structure and estimated $\Re_0$. Misspecifying the infectious period distribution by a tighter distribution (higher j) generates lower $\Re_0$ estimates (Fig 3A). Furthermore, regardless of the assumed distribution of the infectious period, *all candidate models fit the data equally well*. To emphasise the importance and implications of this observation, we compare inferred and actual *latent* incidences in Fig 3B. Recall that fitting a candidate model to a given incidence (y) produces a set of samples that describes the posterior distribution. Then, we use those samples to simulate the candidate's system model, thereby generating inferred latent incidences (lines in Fig 3B). This last step is equivalent to approximate expectations (22). Then, those lines are compared to x, the true latent incidence (Eq 4). Notice that by definition, we do not have access to x in practical applications, but by virtue of this simulation study, such an impediment is overcome. The comparison reveals a symmetry shared among the candidate models.

That is, any of these formulations can match the true latent incidence provided that β, ρ and I_0^1 are configured appropriately. It is important to remark that this symmetry is restricted to the latent incidence and does not extend to the dynamics of other states. For instance, candidates with different delay distributions that yield equivalent incidences will not reach the same long-term equilibrium, given the differences in their $\Re_0$.

Logically, such symmetry should render the approach of comparing *fit scores* impractical. A fit score, such as the Maximum Likelihood Estimate (MLE), measures the consistency between a dataset and the output generated by a model. Since the various candidates produce equivalent output, differences among MLEs will solely reflect the stochasticity (noise) of the measurement model rather than some structural variation in the system model. We empirically verify this conjecture by selecting the candidate with the largest MLE for each incidence report (see the electronic supplementary material *S2 §1.3.2*). We observe that M^{11} candidates attain the largest MLE in only *12* out of *20* times when matching D^{11} incidence reports. Even worse, M^{13} instances are always outperformed in fitting D^{13} datasets. Overall, no candidate passes the 60 % mark. In light of this evidence, one can safely conclude that score comparison is not a reliable approach to determining the correct distribution of epidemiological delays from incidence data. To further complicate matters, information criteria (such as *AIC* and *BIC*) and cross-validation methods cannot assist in this task, considering that model candidates produce equivalent output and share an equal number of unknown parameters. In fact, we spent considerable time carrying out Leave-Future-Out Cross-validation (*LFO-CV*) before noting the inherent symmetry in these structures.

3.2 Four unknowns

The reason for such inherent symmetry is an element that epidemic models implicitly specify: the time between the infection of a primary case and one of its secondary cases, also known as the *generation time* (58). This quantity's shape, in tandem with $\Re_0$, determines the initial dynamics of an infectious disease (5). Interestingly, these elements also characterise long-term behaviour. Krylova & Earn (29) found that SEIR models that account for demographic processes with different delay distributions produce equivalent dynamics of epidemiological transitions (e.g. from annual to biennial epidemic cycles) if they share identical $\Re_0$ and mean generation time (τ). This last quantity (Eq 6) is given as a function (58) of the average latent (σ^{-1}) and infectious (γ^{-1}) periods and the infectious period distribution (j).

$$\tau = \sigma^{-1} + \frac{j+1}{2j}\gamma^{-1} \tag{6}$$

In this analysis, we have, until now, fixed the mean generation time on each candidate model by excluding σ and γ from the inference process. Taking note of the effects on short and long-term dynamics that produce the interaction between τ and $\Re_0$, we now promote γ to the category of *estimated parameter* in order to explore the impact of a variable mean generation time. The reason for choosing γ as the extra parameter is based on the fact that it interacts with both quantities of interest (Eqs 3 and 6). This choice implies the need for estimating four parameters per model instance. To do so, we follow the approach described in the previous section. The reader can find the full set of results in the electronic supplementary material *S2 §2*. Unsurprisingly, given the extra degree of freedom, all candidates fit any of the incidence data equally well. In this design, though, the match between synthetic data and fitting model's output is achieved at the expense of less precision, although greater accuracy. *Precision* refers to the width of uncertainty intervals, and *accuracy* to whether the interval captures the actual value. To illustrate this phenomenon, in Fig 4, we present the results of fitting four candidates models (M^{1j}) to four incidence reports that stem from different distributions of the infectious period (D^{1j}). Here, we see that the range of $\Re_0$ widened (Fig 4A) compared to that presented in the previous section (Fig 3A).

Undoubtedly, the chief insight from allowing γ to vary is the unravelled interaction between $\Re_0$ and τ. We visualise this interaction by plugging samples of β and γ into Eqs 3 and 6 to obtain an approximation of the expected values of $\Re_0$ and τ. When these two quantities are displayed on a scatter plot (Fig 4B), a linear relationship appears, regardless of the data's origin or the fitting model's structure. The interpretation of such linear association indicates that for a given fitting model, infinite pairs of $\Re_0$ and τ yield equivalent incidence dynamics. However, each value of τ corresponds to exactly one value of $\Re_0$. This result explains the wide range of $\Re_0$ estimates obtained by Wearing and colleagues (25). By treating $\Re_0$ and τ as unknowns, the optimiser could settle on any section of the line drawn by these parameters.

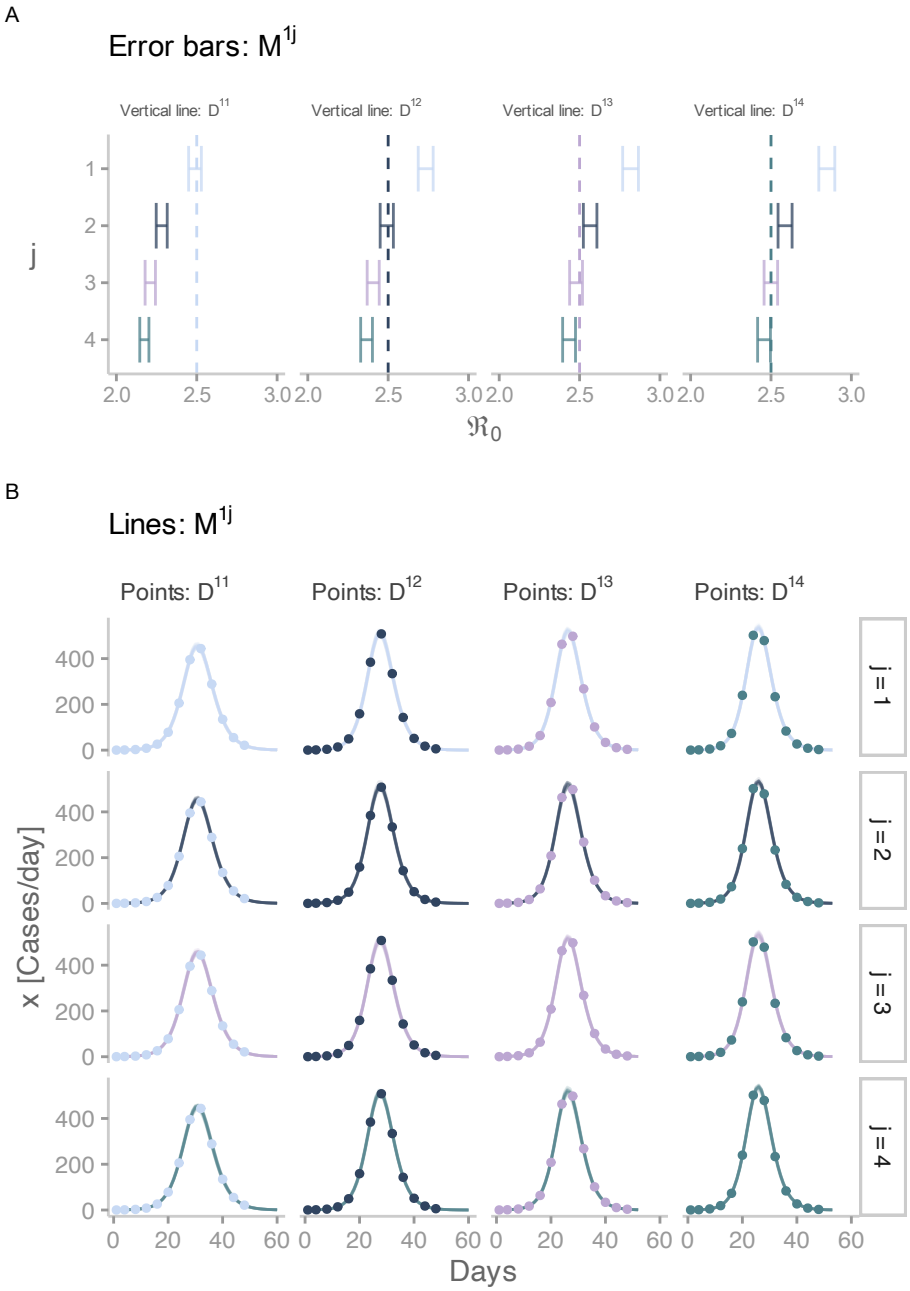


Figure 3: Inference results obtained from the three-unknown parameterisation. This plot shows the results of fitting four model candidates (M^{1j}) to four different incidence reports. Each report originated from a different infectious period distribution (D^{1j}). A) Comparison of estimates of the basic reproduction number by fitting model. Error bars correspond to 95 % credible intervals, and the vertical line denotes the true value. B) Comparison between inferred (lines) and actual (dots) latent incidences. Twenty time series represent inferred incidence. Given the high-fidelity data, all inferred incidences are nearly identical, giving the impression that there is only one line in each panel.

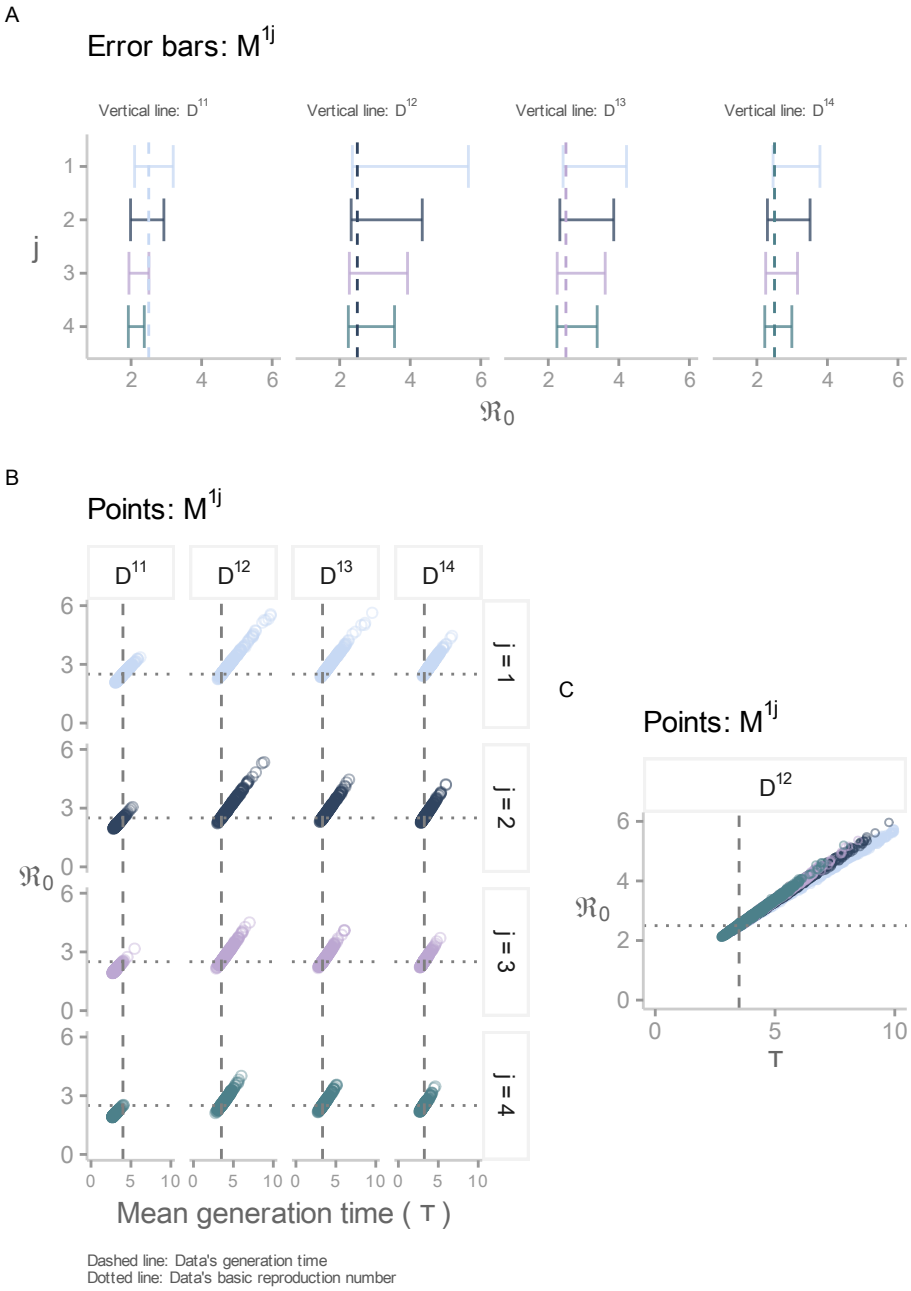


Figure 4: Inference results obtained from the four-unknown parameterisation. This plot shows the results obtained from fitting four models candidates (M^{1j}) to four different incidence reports. Each report originated from a different distribution of the infectious period (D^{1j}). A) Comparison of estimates of the basic reproduction number by fitting model. Error bars correspond to 95 % credible intervals, and vertical lines denotes the true value. B) Linear relationship between the basic reproduction number and the mean generation time. Fitting a candidate model to a specific incidence report results in a posterior distribution represented by samples. For each sample, we compute its respective $\Re_0$ and τ represented by a hollow dot. C) Linear relationship obtained from fitting the four candidates to a D^{12} incidence report. This plot is analogous to overlapping the second column in B.

3.3 Three unknowns (alternative)

More importantly, the linear relationships shown in Fig 4B reveal an intriguing insight. Notice that irrespective of the structure (M^{1j}) fitting data of any origin (D^{1j}), the true values of $\Re_0$ and τ (dotted and dashed lines, respectively) are subsumed into any of the linear associations. This observation leads us to consider whether it could be possible to accurately estimate $\Re_0$ from a model whose mean generation time is fixed to the true underlying value, but the shape of the epidemiological delays may differ from that of the data-generating process. To test this hypothesis, one could assign arbitrary values to σ and γ that yield (via Eq 6) the desired value of τ. Alternatively, one could reformulate the SE^iI^jR framework so that τ becomes a parameter of every model instance. We prefer the latter to demonstrate directly the leverage this parameter has on the dynamics of epidemiological models. Consequently, we combine Eqs 3 and 6 into 7, which expresses β as a dependent variable of four parameters: j, σ, τ, and $\Re_0$.

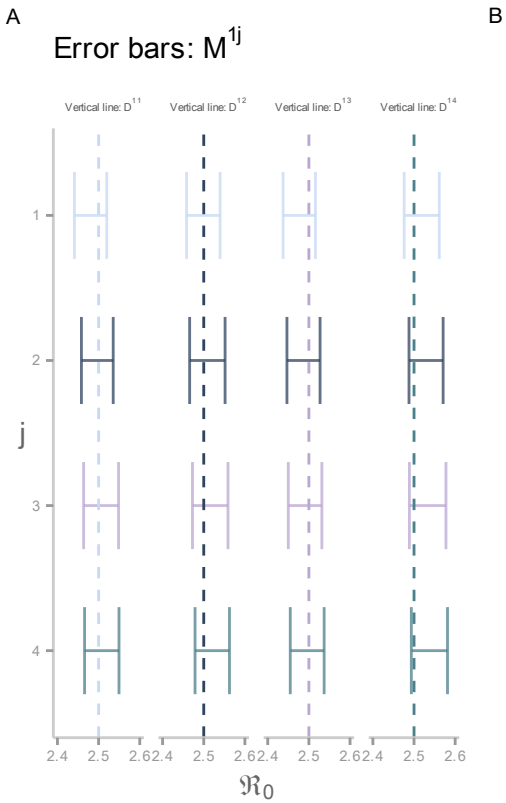
$$\begin{aligned} \beta &= \frac{2j(\tau - \sigma^{-1})}{\Re_0^{-1}(j+1)} \\ \gamma &= \beta \Re_0^{-1} \end{aligned} \tag{7}$$

Parameter j is given by the fitting model's structure, whereas σ and τ are fixed to the true values that produced the incidence reports. For instance, a D^{12} report stems from a structure whose σ and τ are equal to 0.5 and 3.5 (applying Eq 6), respectively. Therefore, an M^{14} candidate fitting this report has j, σ, and τ fixed to 4, 0.5, and 3.5, respectively. An immediate consequence of this procedure is the need to constrain γ in order to maintain logical consistency. Accordingly, we define γ as a function of β and $\Re_0$ (7). This approach is analogous to fixing γ to an arbitrary value that yields the desired τ. Such a value may not correspond to that of the data-generating process. Lastly, the remaining parameter, $\Re_0$, is deemed a random variable. Namely, subject to inference. We opt to estimate its inverse for a practical reason. Taking into account the threshold phenomenon and the fact that all incidence reports exhibit outbreak-like behaviour, any estimated value of $\Re_0$ must fall within the interval $(1, \infty)$. It then logically follows that its inverse ($\Re_0^{-1}$) spans over the range $(0, 1)$. This transformation permits the inference algorithm to operate in a much smaller parameter space, which enhances sampling efficiency.

We subsequently incorporate the redefined components (β and γ) into the SE^iI^jR framework to produce an alternative set of four candidate models with three unknowns: $\Re_0^{-1}$, ρ and I_0^1. Similarly as before, we estimate the posterior distribution for each candidate fitted to an incidence report. The reader can find the complete set results in the electronic supplementary material *S2 §3*. These results once more expose the intrinsic symmetry of SEIR formulations. Specifically, provided there is an adequate configuration, any candidate structure can accurately match the observed incidence despite differences in the infectious period distribution. Nevertheless, this alternative parameterisation exhibits a distinctive and crucial feature: the estimation of $\Re_0$ is less sensitive to the assumed distribution of the infectious delay. To support this claim, we present in Fig 5 the results of fitting the four alternative candidates to four incidence reports of dissimilar origin. Here, it can be seen (Fig 5A) that all candidates recover (via 95 % credible intervals) the underlying true $\Re_0$, notwithstanding the origin of the data or the fitting model.

Recovering the underlying $\Re_0$ is not exclusive to this sample of four datasets but is generalised across the *80* high-fidelity D^{1j} datasets. To summarise this insight, we borrow a concept from the *frequentist* tradition. Such a concept known as *coverage* (59) means that if one collects a large number of samples from the same process and constructs the corresponding confidence intervals, then a certain percentage of the intervals will contain or *cover* the true parameter. This percentage is given by the confidence level. For instance, if one fits a model to 100 datasets and estimates an equal number of confidence intervals at the 95% significance level, then 95 of those intervals will cover the true value. Admittedly, it is implicitly assumed that our 95% credible intervals (obtained from posterior distributions) are proportional to 95 % confidence intervals. To illustrate this procedure, we focus on the D^{11} incidence reports and the M^{12} fitting instance (second row in Fig 5B). M^{12} recovers the true $\Re_0$ in 19 out of 20 (95 %) times when fitting D^{11} reports, a result consistent with the definition of coverage. Indeed, estimated intervals for $\Re_0$ and ρ conform to this concept, where minor deviations are justified by the fact that coverage is defined asymptotically (infinite measurements). However, asymptotics does not account for the large deviance observed in the estimates of I_0. We explain this inconsistency in the section below where I_0 becomes more prominent.

To conclude this section, we report the analysis of the low-fidelity datasets (right column in Fig 2). The reader can find the results in the electronic supplementary material *S3*. Overall, we obtain similar insights in comparison to those derived from the high-fidelity datasets. In the absence of structural differences, it is



B

D^{ij}	M^{ij}	$\Re_0$	I_0	ρ
11	11	95%	95%	100%
11	12	95%	45%	95%
11	13	90%	10%	95%
11	14	85%	5%	95%
12	11	90%	65%	100%
12	12	90%	95%	100%
12	13	95%	85%	100%
12	14	95%	75%	100%
13	11	75%	50%	100%
13	12	95%	95%	100%
13	13	95%	95%	100%
13	14	100%	90%	100%
14	11	90%	20%	100%
14	12	90%	60%	100%
14	13	85%	80%	100%
14	14	85%	95%	95%

Figure 5: Inference results obtained from the three-unknown alternative parameterisation. This plot shows the results obtained from fitting four models candidates (M^{1j}) to four different incidence reports. Each report originated from a different distribution of the infectious period (D^{1j}). A) Comparison of estimates of the basic reproduction number by fitting model. Error bars correspond to 95 % credible intervals, and the vertical line denotes the true value. B) Coverage table. For instance, the third row corresponds to the candidate M^{13} fitting 20 D^{11} incidence reports. The third column of this row indicates that 18 out of the 20 95% credible intervals cover the true $\Re_0$ value. Namely, 90 % coverage.

unsurprising that the effect of larger noise in the signal (overdispersion) is confined to greater uncertainty in parameter estimates. More precisely, decreases in precision (wider credible intervals) can obscure or accentuate features of the inference process. On the one hand, overdispersion can mask structural deficiencies of fitting structures. For instance, noisier measurements cause the I_0 estimates from the *alternative three-unknown* parameterisation to conform to the expected coverage. This event should not occur based on the results obtained from the high-fidelity datasets. On the other hand, overdispersion can exacerbate identifiability issues. Under the *four-unknown* parameterisation, some $\Re_0$ estimates reach values up to 40. This result is a reminder that choosing an adequate number of unknowns is not a trivial decision. Setting more unknowns than the data can *tolerate* renders models unidentifiable. In this context, unidentifiability occurs because the incidence data does not provide enough information to update the prior distribution of γ. As discussed above, many values of γ are consistent with the observed incidence, an insight that holds for both levels of data fidelity. Finally, we note that overdispersion estimates are robust to the choice of the infectious period distribution.

3.4 Misspecifying the distribution of the latent period

Thus far, we have conducted the inference process assuming that the latent period distribution (i) is known. Lifting this constraint would strain our computational resources, producing a four-fold increase in the pool of candidates fitting a single report (assuming $i, j \in \{1, 2, 3, 4\}$). Instead of undertaking such costly exploration, one could leverage the fact that the mean generation time depends solely on the mean latent period rather

than its particular distribution (Eq 6). To test this idea, we compare the estimates obtained from candidate models with the *right* and *wrong* distribution of the latent period. We illustrate this process with the *80* D^{3j} low-fidelity ($\phi^{-1} = 1/3$) datasets. For each dataset, we fit eight candidates M^{ij} from the traditional three-unknown parameterisation (β, ρ, and I_0), where the distribution of the latent period can take the wrong ($i = 1$) and the right ($i = 3$) values, and the infectious period distribution varies as before, namely, $j \in \{1, 2, 3, 4\}$. The reader can find the complete results in the electronic supplementary material *S5*.

To facilitate the presentation of the results, we first focus on candidates M^{13} and M^{33} fitting one D^{33} incidence report. Fig 6A shows that both models predict similar, although not identical, latent incidence rates. Further inspection reveals that the slight difference in the predicted incidence due to dissimilar distributions of the latent period does not lead to variation in $\Re_0$ estimates. To corroborate this assessment, we expand the analysis to the eight candidates matching the same incidence report. The left-hand side of Fig 6B shows that $\Re_0$ estimates are sensitive to variation in the structure of the infectious period but are indifferent to the latent period distribution. In compliance with the literature, the more dispersed latent period ($i = 1$) leads to an earlier incidence peak compared to the tighter distribution ($i = 3$) in the context of identical $\Re_0$.

Nevertheless, the mechanism that enables models with heterogeneous distributions to produce analogous incidence dynamics remains unexplained. The right-hand side of Fig 6B, which displays I_0 estimates, provides the first hint. Here, we note that misspecifying the latent period distribution leads to biases in the predicted initial number of infected individuals. To explain this phenomenon, we draw on a broader view of the posterior distribution. It is commonplace to restrict inference analyses to one parameter at a time (i.e. marginal distributions), neglecting the information provided by the full posterior distribution. To redress this shortcoming, we visualise the full distribution via pair plots. Specifically, Fig 6C corresponds to the summary of the posterior distribution obtained from fitting M^{13} to one D^{33} incidence report. The upper triangular elements of this plot indicate that the three estimated parameters are strongly correlated. Especially β and I_0, or more compellingly, $\Re_0$ and I_0. Recall that the basic reproduction number is directly proportional to β. Therefore, although the mean generation time determines which $\Re_0$ corresponds to the observed incidence, I_0 (and ρ to a lesser extent) regulates the flexibility of $\Re_0$ to reach such *desired* value. Interestingly, I_0 provides such a degree of flexibility that unrealistic adjustments in its estimates allow us to equate dissimilar model structures. Notice that the only discernible difference between Figs 6C and 6D (M^{33} fitted to D^{33}) is seen in the marginal distributions of I_0. In fact, this phenomenon explains the *failure* of the alternative parameterisation to recover the true value of I_0 (Figs 5B).

In view of these symmetries, it is not unreasonable to expect that candidates from the *four-unknown* and the *alternative* parameterisations, too, are indifferent to the distribution of the latent period once I_0 and ρ correct for any misspecification. To verify this premise, we fit the parameterisations mentioned above to the D^{3j} low-fidelity incidence reports. As anticipated, the inference results indicate that the *four-unknown* parameterisation (S5 §2) uncover the linear association between τ and ρ due to the unidentifiability of γ. Likewise, the *alternative* parameterisation (S5 §3) recovers the true $\Re_0$ irrespective of the formulation of the epidemiological delays. Furthermore, we replicate these results using the D^{3j} high-fidelity datasets (see electronic supplementary information S4).

3.5 Sensitivity analysis

Another aspect that deserves careful attention is the choice of the measurement model (likelihood function). So far, fitting candidates have been amalgamated with the appropriate likelihood function. However, the choice of such a structure in an actual application is not straightforward. Even though this work does not aim to provide an exhaustive account of probability distributions, we explore the consequences of coupling the alternative parameterisation with Poisson and Normal density functions when fitting overdispersed incidence data. We perform such exploration by inferring $\Re_0$ from M^{1j} candidates fitted to the D^{3j} low-fidelity datasets discussed in the previous section. As expected, the results indicate that the Poisson distribution (see S5 §4) leads to overconfident (too precise) and biased (inaccurate) estimates in the context of overdispersion. Visually, we appreciate these features with narrow uncertainty intervals that do not cover the true value. Conversely, the Normal distribution (see S5 §5) offers more accurate estimates than the Poisson distribution, but coverage remains suboptimal compared to the Negative Binomial distribution. Admittedly, this assessment is based on a unique level of overdispersion, and generalisation requires further analysis.

Likewise, the synthetic data used for the analysis presented in the previous sections stems from models configured to identical $\Re_0$ and similar mean generation times (variation due to the infectious period distribution). Naturally, one wonders whether the usefulness of the alternative parameterisation holds under various conditions of $\Re_0$ and τ. In the absence of analytical solutions, we must resort once more to a simulation

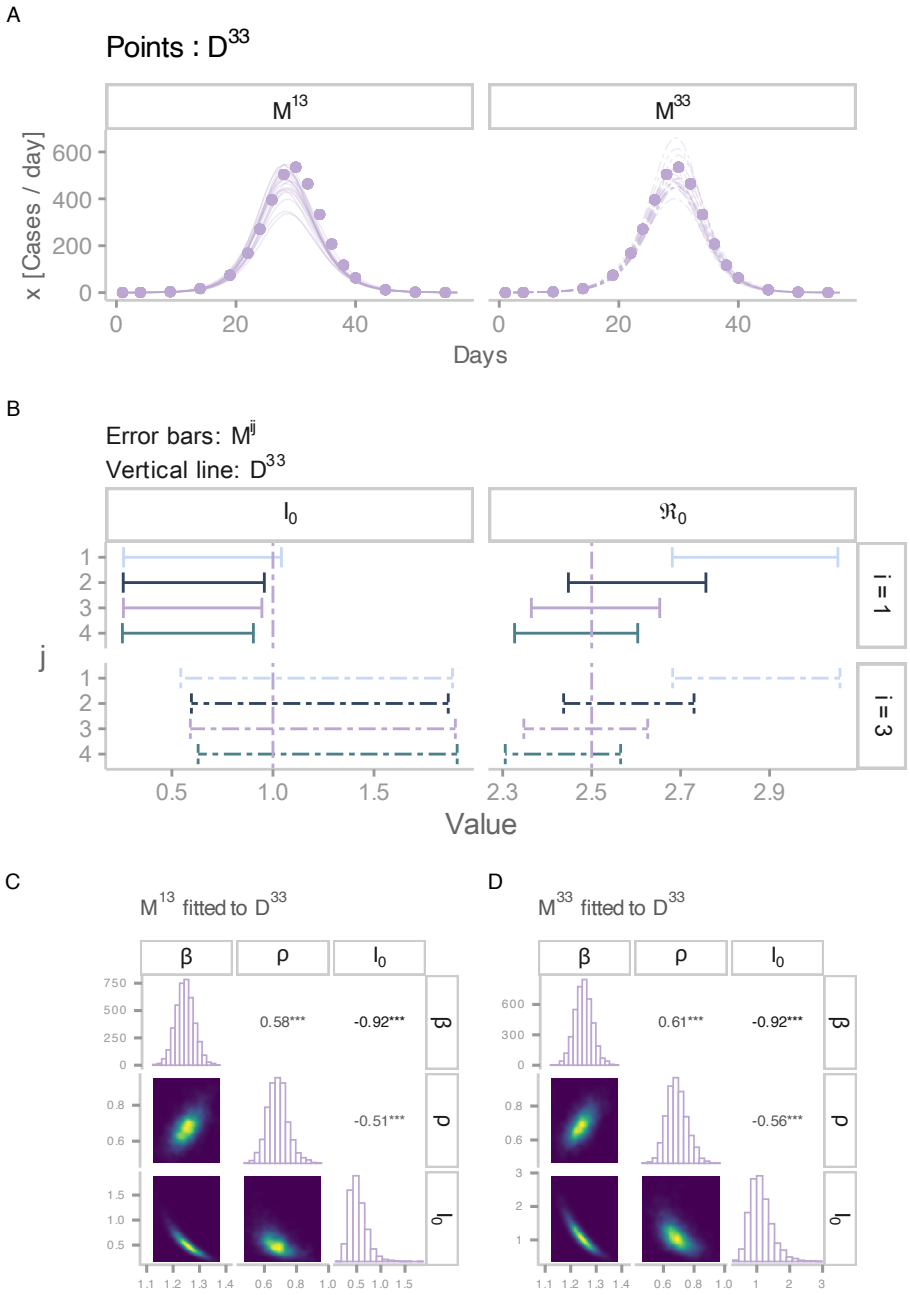


Figure 6: Comparison of inference results obtained from three-unknown models with the wrong and correct distributions of the latent period. A) Incidence fit from two candidate model matching a D^{33} incidence report. Candidate on the left has an exponentially-distributed (wrong) latent period, whereas candidate on the right has Gamma distributed latent period (correct). B) Comparison of estimates of the initial number of infectious individuals at stage 1 and the basic reproduction number by fitting model. Error bars correspond to 95 % credible intervals, and vertical lines denotes the true value. C) Joint posterior distribution from M^{13} candidate fitting a D^{33} report. The diagonal shows posterior marginal distributions. In the lower triangular part, each possible pairwise conditional distribution is displayed, whereas the upper triangular part presents the correlation among parameters. C) Joint posterior distribution from M^{33} candidate fitting a D^{33} report.

study. In particular, we repeat the workflow described in this paper for two additional configurations of τ and $\Re_0$ each (four in total). For simplicity, we restrict this sensitivity analysis to Poisson (high-fidelity) and overdispersed (low-fidelity) datasets derived from models with an exponentially-distributed latent period (D^{1j}). Additionally, we equip the fitting candidates with the appropriate likelihood function. The complete set of results is presented in the electronic supplementary information S6. We refer to each new configuration of $\Re_0$ and τ as *scenario*, and for simplicity, we characterise each scenario by the $\Re_0$ value and the mean generation time obtained from an exponentially-distributed infectious period (τ_e). For instance, the base case scenario, *Scenario 1*, corresponds to $\Re_0 = 2.5$ and $\tau_e = 4$ (results presented in Section 3.3).

For *Scenario 2*, we increase the reference mean generation time ($\tau_e = 8$), while keeping $\Re_0$ at 2.5. First, we focus on the high-fidelity datasets. Overall, the greater the divergence between the fitting model's distribution of the infectious period and the actual distribution that generated the data, the greater the loss in accuracy. Namely, lower coverage. To provide an example, the 95 % credible intervals constructed from M^{14} candidates fitting D^{11} incidence reports only attain coverage of 30% for $\Re_0$. Closer inspection, though, reveals that such accuracy loss is more *statistical* than *practical*. To support this statement, we calculate the average relative difference between the actual and estimated $\Re_0$, finding that misspecification of the infectious period distribution leads to a maximum average relative error of 2 %. In contrast, we would obtain discrepancies up to 15 % if we adopted the traditional approach. Simply put, it is costlier to misspecify the mean generation time than the mean infectious period. Furthermore, such slight differences in the alternative parameterisation are erased by overdispersion. Notice that the estimated coverage from the low-fidelity datasets is in agreement with its expected value. That is, overdispersion masks minor misspecification in the process model. Moreover, in *Scenario 3* ($\Re_0 = 2.5$, $\tau_e = 13$), we observe that further increasing of the mean generation time does not lead to significant drops in the coverage of $\Re_0$ under both levels of data fidelity. In a nutshell, it is reasonable to suggest that the alternative parameterisation is robust to various levels of the mean generation time.

Conversely, we cannot maintain the same assertion for various values of $\Re_0$. Indeed, Fig 4C provided the first hint. This plot shows that the straight lines do not overlap as $\Re_0$ reaches relatively high values. Consequently, in scenarios *4* ($\Re_0 = 9$, $\tau_e = 4$) and *5* ($\Re_0 = 17$, $\tau_e = 4$), we test the implications of larger transmissibility levels. The results indicate that as we increase the underlying $\Re_0$ for generating the data, the equivalency among fitting models dissipates and misspecification in the infectious period distribution leads to biased estimates of $\Re_0$. The size of such bias is proportional to the misspecification of the infectious period and the underlying $\Re_0$. This feature is primarily seen in the estimates derived from high-fidelity datasets, where coverage levels are low, and the average relative error between actual and estimated values cannot be overlooked. However, when we examine the posterior distributions obtained from fitting the low-fidelity data, it is seen that, once again, overdispersion masks misspecification in the process model, as evidenced by the high coverage levels. This is not to say that overdispersion is a desired feature in the data, but rather to emphasise that its presence hinders the attainment of precise estimates. Undoubtedly, having this understanding is of practical importance, given that it allows us to discern the necessary effort in data collection and model improvement.

4 Application to Influenza A

Leveraging the knowledge gained from the synthetic data, the last step in this work consists of exploiting the relationship between the basic reproduction number and the mean generation time to update the $\Re_0$ estimate of an outbreak of the 1918 influenza pandemic. The reader can find the full set of results in the electronic supplementary information S7. In particular, we focus on an outbreak that occurred in the city of Cumberland (Maryland) during the autumn of 1918, for which the U.S. Public Health Service organised special surveys (60) to determine the proportion of the population infected. Previous studies (12,22) employed the default heuristic of adopting an *SEIR* with exponentially-distributed epidemiological delays whose means were configured to values reported in the literature. Moreover, in these studies, the *SEIR* was coupled with the Poisson distribution resulting in a 95 % CI [2.5–2.6] for $\Re_0$. However, adopting a more realistic measurement model, such as the Negative Binomial distribution, produces lower and wider estimates: 95% CI [2.2, 2.4]. Further, if we jettison the assumption of an exponentially-distributed infectious period for a more realistic distribution, such as the gamma distribution, we even obtain lower estimates. For instance, a gamma-distributed infectious period with four stages (SEI^4R) returns a 95% CI of [2.0, 2.2]. As noted earlier, the estimates obtained from this default heuristic or traditional approach are sensitive to the uncertainty in the infectious period distribution. On the contrary, when we fix the mean generation time in the SEI^jR (alternative parameterisation) to a value (2.85 days) obtained from the literature (5,61), we derive nearly identical $\Re_0$ estimates (95% CI [2.0, 2.1]) regardless of the infectious period distribution (Fig 7). Notice that

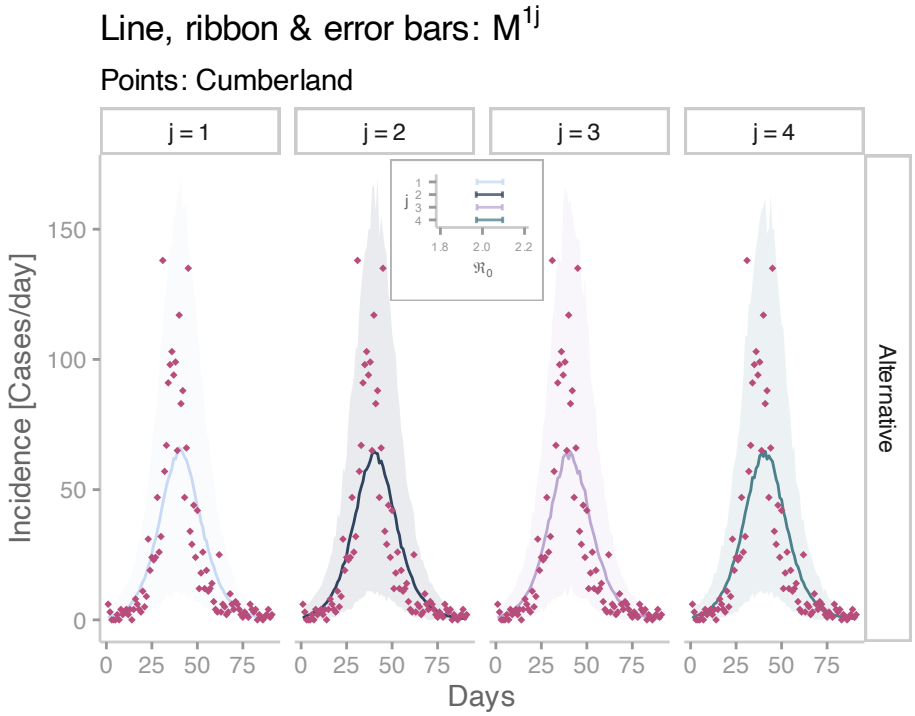


Figure 7: Application of the three-unknown alternative parameterisation to fit four candidates (M^{1j}) to the daily number of influenza cases (rhombi) detected by the U.S. Public Health Service in Cumberland (Maryland) during the 1918 influenza pandemic, from 22 September 1918 to 30 November 1918. Ribbons correspond to 95% credible intervals of the predicted measured incidence by the model. The solid line denotes the median. The box inside the plot shows the estimates of the basic reproduction number by fitting model. Error bars correspond to 95% credible intervals.

this estimate is similar to that obtained from the SEI^4R, bolstering the fact that the actual infectious period is far from being exponentially distributed.

5 Conclusion

The misspecification of various assumptions within the *SEIR* framework can jeopardise the estimation of $\Re_0$. In recognition of that risk, we ran a simulation study comprised of approximately 1000 synthetic datasets and 8000 model fits, whereby we identified the relative influence of some of those assumptions. Specifically, we found that fixing the mean generation time to a reliable estimate is of paramount importance. In contrast, one can be more lenient on the specification of the latent period distribution and the mean infectious period provided that other estimated parameters (I_0 and ρ) redress the misspecification. We leveraged this knowledge to formulate an alternative parameterisation that is more robust to the uncertainty of the epidemiological delays. However, there is a caveat with this alternative formulation. Although it exploits a local symmetry (incidence dynamics) of the SEIR framework, such symmetry does not extend to the other states of the system. Therefore, the usefulness of the alternative parameterisation is confined to the estimation of $\Re_0$, and it is not a substitute for other kinds of analyses. For instance, if, on the contrary, our quantity of interest were I_0, we would obtain unreliable estimates. Furthermore, the alternative SEIR with exponentially-distributed delays will be as overoptimistic as its *traditional* counterpart in predicting the critical vaccination proportion or the effectiveness of an imperfect VIH treatment in the context of within-host dynamics (57). Therefore, the alternative parameterisation is a mitigation strategy in the absence of complete information. Furthermore, its usefulness is abated by highly transmissible pathogens. Nevertheless, biases in the estimates due to large $\Re_0$ are only detected with high-fidelity data. That is, data with little or no overdispersion.

Despite the intense computational effort of simulation analyses, a single study cannot offer overarching statements. Further work is required to test the validity of these insights in stricter or more elaborated contexts. For instance, we assumed the complete availability of the incidence time-series throughout this study. This assumption restricts the validity of the approach to retrospective analyses. However, other situations exist where only a few incidence measurements are available to the modellers, such as the early phase of a pandemic response. Hence, it remains to be seen the effect of various levels of data availability on the performance of the suggested approach. Furthermore, for simplicity, we ignored age-related effects in the dynamics of the infectious disease as well as demographic stochasticity and time-varying contact rates. We expect that future research builds on the findings provided by this study and addresses the aforementioned challenges to construct ever more reliable models.

Acknowledgement

The project has received funding –through the School of Medicine, National University of Ireland Galway– from the European Union's Horizon 2020 research and innovation programme under grant agreement No 883285. The material presented and views expressed here are the responsibility of the author(s) only. The EU Commission takes no responsibility for any use made of the information set out. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

References

1. Dublin LI, Lotka AJ. On the true rate of natural increase. Journal of the American Statistical Association. 1925 Sep;20:305.
2. Anderson RM, May RM. Infectious Diseases of Humans: Dynamics and Control. Oxford University Press; 1992.
3. Heffernan JM, Smith RJ, Wahl LM. Perspectives on the basic reproductive ratio. Journal of The Royal Society Interface. 2005 Sep;2:281–93.
4. Keeling MJ, Rohani P. Modeling Infectious Diseases in Humans and Animals. Princeton University Press; 2011.
5. Wallinga J, Lipsitch M. How generation intervals shape the relationship between growth rates and reproductive numbers. Proceedings of the Royal Society B: Biological Sciences. 2007 Feb;274:599–604.
6. Kucharski A. The Rules of Contagion: Why Things Spread - and Why They Stop. Profile Books; 2020.
7. Mollison D. The structure of epidemic models. In: Epidemic models: Their structure and relation to data. Cambridge University Press; 1995. p. 17–33.
8. Dietz K. Transmission and control of arbovirus diseases. In: Epidemiology. Society for Industrial and Applied Mathematics; 1975. p. 104–21.
9. Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society of London Series A, Containing Papers of a Mathematical and Physical Character. 1927 Aug;115:700–21.
10. Diekmann O, Heesterbeek H, Britton T. Mathematical tools for understanding infectious disease dynamics. Princeton University Press; 2013.
11. Lloyd AL. Sensitivity of model-based epidemiological parameter estimation to model assumptions. In: Chowell G, Hyman JM, Bettencourt LMA, Castillo-Chavez C, editors. Mathematical and statistical estimation approaches in epidemiology. Dordrecht: Springer Netherlands; 2009. p. 123–41.
12. Vynnycky E, White R. An Introduction to Infectious Disease Modelling. Oxford University Press; 2010.
13. Brouwer AF. Why the spectral radius? An intuition-building introduction to the basic reproduction number. Bulletin of Mathematical Biology. 2022 Sep;84:96.
14. He D, Ionides EL, King AA. Plug-and-play inference for disease dynamics: measles in large and small populations as a case study. Journal of The Royal Society Interface. 2010;7(43):271–83.
15. Dureau J, Kalogeropoulos K, Baguelin M. Capturing the time-varying drivers of an epidemic using stochastic dynamical systems. Biostatistics. 2013 Jan;14(3):541–55.

16. Andrade J, Duggan J. Inferring the effective reproductive number from deterministic and semi-deterministic compartmental models using incidence and mobility data. PLOS Computational Biology. 2022 Jun;18(6):1–25.
17. Andrade J, Duggan J. An evaluation of Hamiltonian Monte Carlo performance to calibrate age-structured compartmental SEIR models to incidence data. Epidemics. 2020;33:100415.
18. Bretó C, He D, Ionides EL, King AA. Time series analysis via mechanistic models. The Annals of Applied Statistics. 2009;3(1):319–48.
19. Bretó C. Modeling and Inference for Infectious Disease Dynamics: A Likelihood-Based Approach. Statistical Science. 2018;33(1):57–69.
20. Ionides EL, Bretó C, King AA. Inference for nonlinear dynamical systems. Proceedings of the National Academy of Sciences. 2006;103(49):18438–43.
21. Endo A, van Leeuwen E, Baguelin M. Introduction to particle Markov-chain Monte Carlo for disease dynamics modellers. Epidemics. 2019;29:100363.
22. Andrade J, Duggan J. A Bayesian approach to calibrate system dynamics models using Hamiltonian Monte Carlo. System Dynamics Review. 2021;37(4):283–309.
23. Barlas Y. Formal aspects of model validity and validation in system dynamics. System Dynamics Review. 1996;12(3):183–210.
24. Oliva R. Model calibration as a testing strategy for system dynamics models. European Journal of Operational Research. 2003;151(3):552–68.
25. Wearing HJ, Rohani P, Keeling MJ. Appropriate models for the management of infectious diseases. PLoS Medicine. 2005 Jul;2:e174.
26. Gostic KM, McGough L, Baskerville EB, Abbott S, Joshi K, Tedijanto C, et al. Practical considerations for measuring the effective reproductive number, Rt. PLOS Computational Biology. 2020 Dec;16(12):1–21.
27. Hurtado PJ, Kirosingh AS. Generalizations of the "Linear Chain Trick": Incorporating more flexible dwell time distributions into mean field ODE models. Journal of Mathematical Biology. 2019 Oct;79(5):1831–83.
28. Greenhalgh S, Rozins C. A generalized differential equation compartmental model of infectious disease transmission. Infectious Disease Modelling. 2021;6:1073–91.
29. Krylova O, Earn DJD. Effects of the infectious period distribution on predicted transitions in childhood disease dynamics. Journal of The Royal Society Interface. 2013 Jul;10:20130098.
30. Gelman A, Vehtari A, Simpson D, Margossian CC, Carpenter B, Yao Y, et al. Bayesian workflow [Internet]. arXiv; 2020. Available from: `https://arxiv.org/abs/2011.01808`
31. King AA, Domenech de Cellès M, Magpantay FMG, Rohani P. Avoidable errors in the modelling of outbreaks of emerging pathogens, with special reference to Ebola. Proceedings of the Royal Society B: Biological Sciences. 2015;282(1806):20150347.
32. Savage LJ. The foundations of statistics. Dover Publications; 1972. (Dover books on mathematics series).
33. Pidd M. Tools for Thinking: Modelling in Management Science. Wiley; 1996.
34. Lloyd AL. Realistic distributions of infectious periods in epidemic models: Changing patterns of persistence and dynamics. Theoretical Population Biology. 2001;60(1):59–71.
35. Keeling MJ, Grenfell BT. Understanding the persistence of measles: Reconciling theory, simulation and observation. Proceedings of the Royal Society of London Series B: Biological Sciences. 2002;269(1489):335–43.
36. Davies NG, Klepac P, Liu Y, Prem K, Jit M, Pearson CAB, et al. Age-dependent effects in the transmission and control of COVID-19 epidemics. Nature Medicine. 2020;26(8):1205–11.
37. Gleeson JP, Brendan Murphy T, O'Brien JD, Friel N, Bargary N, O'Sullivan DJP. Calibrating COVID-19 susceptible-exposed-infected-removed models with time-varying effective contact rates. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. 2022;380(2214):20210120.

38. Rahmandad H, Lim TY, Sterman J. Behavioral dynamics of COVID-19: Estimating underreporting, multiple waves, and adherence fatigue across 92 nations. System Dynamics Review. 2021;37(1):5–31.
39. Chowell G, Ammon CE, Hengartner NW, Hyman JM. Transmission dynamics of the great influenza pandemic of 1918 in Geneva, Switzerland: Assessing the effects of hypothetical interventions. Journal of Theoretical Biology. 2006 Jul;241:193–204.
40. Chowell G, Nishiura H, Bettencourt LMA. Comparative estimation of the reproduction number for pandemic influenza from daily case notification data. Journal of The Royal Society Interface. 2007;4(12):155–66.
41. Vynnycky E, Edmunds WJ. Analyses of the 1957 (Asian) influenza pandemic in the United Kingdom and the impact of school closures. Epidemiology and Infection. 2008;136(2):166–79.
42. BAILEY NTJ. A statistical method of estimating the periods of incubation and infection of an infectious disease. Nature. 1954;174(4420):139–40.
43. SARTWELL PE. THE DISTRIBUTION OF INCUBATION PERIODS Of INFECTIOUS DISEASE. American Journal of Epidemiology. 1995 Mar;141(5):386–94.
44. ANDERSON D, WATSON R. On the spread of a disease with gamma distributed latent and infectious periods. Biometrika. 1980;67:191–8.
45. Blitzstein JK, Hwang J. Introduction to Probability, Second Edition. CRC Press; 2019. (Chapman & hall/CRC texts in statistical science).
46. Gamado KM, Streftaris G, Zachary S. Modelling under-reporting in epidemics. Journal of Mathematical Biology. 2014;69(3):737–65.
47. Organization WH. Pandemic influenza risk management: A WHO guide to inform and harmonize national and international pandemic preparedness and response. World Health Organization; 2017 p. 62 p.
48. Monto AS, Webster RG. Influenza pandemics: History and lessons learned. In: Textbook of influenza. 2013. p. 20–34.
49. Talts S, Betancourt M, Simpson D, Vehtari A, Gelman A. Validating bayesian inference algorithms with simulation-based calibration [Internet]. arXiv; 2018. Available from: `https://arxiv.org/abs/1804.06788`
50. Mills CE, Robins JM, Lipsitch M. Transmissibility of 1918 pandemic influenza. Nature. 2004 Dec;432(7019):904–6.
51. Arulampalam MS, Maskell S, Gordon N, Clapp T. A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. IEEE Transactions on Signal Processing. 2002;50(2):174–88.
52. Lambert B. A student's guide to Bayesian statistics. Los Angeles: SAGE; 2018.
53. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. Bayesian data analysis, third edition. Taylor & Francis; 2013. (Chapman & hall/CRC texts in statistical science).
54. McElreath R. Statistical Rethinking: A Bayesian Course with Examples in R and Stan. CRC Press LLC; 2020. (A chapman & hall book).
55. Betancourt M. A conceptual introduction to hamiltonian monte carlo [Internet]. arXiv; 2017. Available from: `https://arxiv.org/abs/1701.02434`
56. Carpenter B, Gelman A, Hoffman MD, Lee D, Goodrich B, Betancourt M, et al. Stan: A probabilistic programming language. Journal of statistical software. 2017;76(1).
57. Lloyd AL. The dependence of viral parameter estimates on the assumed viral life cycle: Limitations of studies of viral load data. Proceedings of the Royal Society of London Series B: Biological Sciences. 2001 Apr;268(1469):847–54.
58. Svensson Åke. A note on generation times in epidemic models. Mathematical Biosciences. 2007;208(1):300–11.
59. Freedman D, Pisani R, Purves R. Statistics. International student ed., 4. ed. New York: Norton; 2007.
60. Frost WH, Sydenstricker E. Influenza in maryland: Preliminary statistics of certain localities. Public Health Reports (1896-1970). 1919;34(11):491–504.

61. Hirotsu N, Ikematsu H, Iwaki N, Kawai N, Shigematsu T, Kunishima O, et al. Effects of antiviral drugs on viral detection in influenza patients and on the sequential infection to their family members—serial examination by rapid diagnosis (capilia) and virus culture. International Congress Series. 2004;1263:105–8.