



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

Journal Pre-proof

Uncertainty and error in SARS-CoV-2 epidemiological parameters inferred from population-level epidemic models

Dominic G. Whittaker, Alejandra D. Herrera-Reyes, Maurice Hendrix, Markus R. Owen, Leah R. Band, Gary R. Mirams, Kirsty J. Bolton, Simon P. Preston



PII: S0022-5193(22)00328-9
DOI: <https://doi.org/10.1016/j.jtbi.2022.111337>
Reference: YJTBI 111337

To appear in: *Journal of Theoretical Biology*

Received date: 30 June 2022
Revised date: 16 October 2022
Accepted date: 26 October 2022

Please cite this article as: D.G. Whittaker, A.D. Herrera-Reyes, M. Hendrix et al., Uncertainty and error in SARS-CoV-2 epidemiological parameters inferred from population-level epidemic models. *Journal of Theoretical Biology* (2022), doi: <https://doi.org/10.1016/j.jtbi.2022.111337>.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Uncertainty and error in SARS-CoV-2 epidemiological parameters inferred from population-level epidemic models

Dominic G. Whittaker^{a,1}, Alejandra D. Herrera-Reyes^{a,1}, Maurice Hendrix^{a,b},
Markus R. Owen^a, Leah R. Band^a, Gary R. Mirams^a, Kirsty J. Bolton^{a,*},
Simon P. Preston^a

^a*School of Mathematical Sciences, University of Nottingham, University Park, Nottingham, NG7 2RD, UK*

^b*Digital Research Service, University of Nottingham, University Park, Nottingham, NG8 1BB, UK*

Abstract

During the SARS-CoV2 pandemic, epidemic models have been central to policy-making. Public health responses have been shaped by model-based projections and inferences, especially related to the impact of various non-pharmaceutical interventions. Accompanying this has been increased scrutiny over model performance, model assumptions, and the way that uncertainty is incorporated and presented. Here we consider a population-level model, focusing on how distributions representing host infectiousness and the infection-to-death times are modelled, and particularly on the impact of inferred epidemic characteristics if these distributions are misspecified. We introduce an *SIR*-type model with the infected population structured by ‘infected age’, i.e. the number of days since first being infected, a formulation that enables distributions to be incorporated that are consistent with clinical data. We show that inference based on simpler models without infected age, which implicitly misspecify these distributions, leads to substantial errors in inferred quantities relevant to policy-making, such as the reproduction number and the impact of interventions. We consider uncertainty quantification via a Bayesian approach, implementing this for both synthetic and real data focusing on UK data in the period 15 Feb–14 Jul 2020, and emphasising circumstances where it is misleading to neglect uncertainty.

This manuscript was submitted as part of a theme issue on “Modelling COVID-19 and Preparedness for Future Pandemics”.

*Corresponding author

Email address: kirsty.bolton@nottingham.ac.uk (Kirsty J. Bolton)

¹Joint lead authors

1. Introduction

A simple deterministic discrete time Susceptible-Infected-Removed (‘SIR’) epidemic model, sometimes termed the ‘general epidemic model’ [11], is defined as follows. Suppose that S_i and I_i are the number of susceptible and infectious individuals, respectively, on day i ; and that between day i and $i + 1$ for some rate constants λ and θ , $\lambda S_i I_i$ of the susceptibles become infectious, and θI_i of the infectious individuals become ‘removed’ (by death or recovery, so that they are no longer susceptible nor infectious). Though very simple, this model embodies important epidemiological principles and exhibits reasonable dynamics, including exponential growth in the early stages and decline once the susceptible population is suitably depleted. The model also connects to key epidemiological parameters such as the basic reproduction number, \mathcal{R}_0 , defined as expected number of cases directly generated by one case in a population where all individuals are susceptible, for this model equal to λ/θ . The *SIR* model remains widely used for prediction and inference for epidemics, including SARS-CoV-2 [4, 8].

One of this simple model’s limitations, however, is that it does not distinguish the infectious individuals by how long they have been infected, termed the *infected age*. Consequently, it assumes that the infectiousness of infected individuals is constant, and therefore that the proportion of individuals lasting a particular number of days between infection and death follows a geometric distribution. These assumptions are strongly in conflict with clinical literature [16, 18, 22, 52], as we discuss in the following section. The principal contributions of this paper are to introduce a deterministic epidemic model which is structured by “infected age”, and to use this model as a basis to infer quantities important to policy-making. Similar models incorporating structure by infected age have been investigated by [12] and [23] in the the context of the “forward” problem, i.e. with the aim to understand the impact of infected-age structure on model dynamics. In contrast, our focus here is on the “inverse” problem, i.e., to understand the impact on quantities inferred from data, and in particular to highlight circumstances when ignoring infected-age structure gives misleading conclusions.

Quantities of policy-making relevance and of interest to infer include: the timing and impact of behavioural changes, especially relating to non-pharmaceutical interventions (NPIs) such as ‘lockdown’ measures; the effect of NPIs on epidemiological characteristics such as the effective reproduction number, \mathcal{R}_i (defined later); the timing and size of the peak in the number of new daily infections; and the number of individuals that remain susceptible at the end of an epidemic wave.

In Section 2 we describe the data available on the epidemic time-course and on changes in mobility patterns relating to response to NPIs. We describe an infected-age-structured model, which is the central model in the paper, and some simpler models for comparison of the types sometimes used for inferring characteristics of an epidemic. The results in Section 3 include an application of the model to compare competing hypotheses for epidemic decline; a simulation

study investigating the impact of using non-infected-age-structured models for inference when the data arise from an infected-age-structured model; and full Bayesian inference focusing on UK data for the SARS-CoV-2 outbreak in spring 2020, including for the quantities listed in the preceding paragraph. Section 4 contains a concluding discussion.

2. Data, Models and Methods

2.1. Data

We draw on epidemiological and clinical data for ancestral SARS-CoV-2 that was available following the first epidemic wave, as described further below.

2.1.1. Deaths associated with SARS-CoV-2

We focus on country-level SARS-CoV-2 mortality data for England and Wales for the period Feb-July 2020. The total population, denoted N in the following models, is 59.1 million [43]. Various time-course data are available for this period, including the number of confirmed cases by specimen date, though such data are influenced by rapid changes to testing capacity [10] and strategy [14] that make them challenging to connect to variables in an epidemic model. For this reason we focus exclusively on data for SARS-CoV-2-associated deaths collated by date of death for England and Wales by the UK Office for National Statistics. These deaths data count registered deaths where COVID-19 was mentioned on the death certificate [42], with the first death occurring on 8 Mar 2020. Although death data are available beyond July 2020, we restrict ourselves to considering inferences that may be made at the conclusion of the first epidemic wave. The deaths data are plotted later in results Figure 3.

2.1.2. Mobility data

The UK government announced on 23 Mar 2020 that a lockdown would be imposed entailing severe restrictions including that that no one should leave their place of residence except for some specific reasons [50]. Over the next 5 months this legislation was revised approximately monthly, sequentially expanding the permissible reasons to leave one's place of residence, whilst introducing guidance on COVID-secure workplaces [9] and mandating face coverings in some public spaces [51]. Publicly available data indicate the extent to which public behaviour altered in this period. Google mobility data, for example, indicate a relative change, with respect to a normal baseline, in activity in various categories across the UK [19], and provide a proxy for changes in mixing intensity in England & Wales over this period. Figure 1A shows the mean of 'workplace' and 'transit' categories computed daily from the UK Google mobility data, which motivates later modelling of a change in population mixing intensity with a step function to account for the introduction of NPIs.

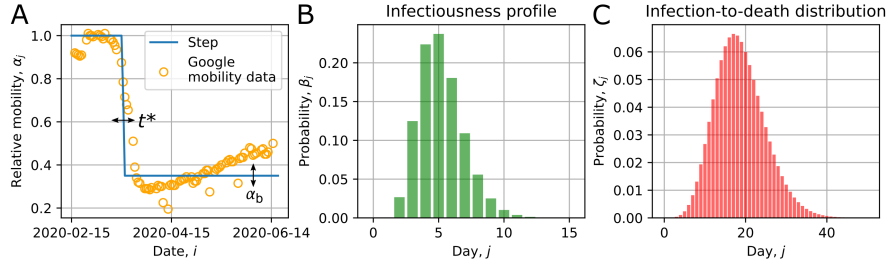


Figure 1: (A) Google mobility data and the parameterisation for a step function, defined in (3), for α_j , which describes changes in population mixing intensity. The data points are given by the mean of workplace and transit activity levels for the UK, with Saturday and Sunday removed (the four outliers being bank holidays). The blue step is given by equation (3) and contains two parameters which control the timing and severity of the drop in viral transmission due to lockdown measures: t^* and α_b , respectively. Plots (B) and (C) respectively show the infectiousness profile, over a 15 day period, and the infection-to-death distribution, shown over a 50 day period; these correspond to the distributions and default parameters listed in Table 1.

2.1.3. Clinical data on individuals' response to SARS-CoV-2 infection

The models in §2.2 involve epidemiological parameters drawn from the clinical literature that characterise individuals' infectiousness and mortality. These are summarised in Table 1, including 'default' values that are used in the later simulation study, and priors that are used for the Bayesian inference.

For most hosts infected with SARS-CoV-2, detectable viral load in the upper respiratory tract peaks at around 5–6 days following exposure [27]. For many hosts this approximately aligns with the delay from exposure to experiencing symptoms (the *infection-to-onset* period) [2], and thus there is a significant period of infectiousness prior to symptom onset [27, 1]. Some hosts remain asymptomatic throughout infection and experience similar viral load dynamics, besides somewhat faster viral clearance following peak viral load [36]. For the purposes of developing a transmission model, it is therefore important to describe a typical host *infectiousness profile*, β_j , subject to $\sum_{j=1}^{\infty} \beta_j = 1$, as depending on the time in days, j , since infection (rather than with respect to symptom onset).

The generation time interval, defined as the period from host infection to the generation of progeny infection, has a distribution that describes such an infectiousness profile [38]. We use the inferred gamma-distributed generation time from the Singapore cluster in Ganyani et al. [18], with mean 5.2 days and variance 2.96 days, to define a fixed infectiousness profile β_j when inferring the impact of lockdown in §3.4. Since the models in this paper are in discrete time, we discretise the gamma distribution by taking $\beta_1 = F(3/2)$ and $\beta_j = F(j + 1/2) - F(j - 1/2)$ for $j \geq 2$, where $F(\cdot)$ is the gamma cumulative distribution function. Estimates from other international data sets available in 2020 are congruent with the estimate from Ganyani et al. [18] [see, e.g. 49, 16]. Estimates for the generation interval for the UK during 2020 became available

in 2021: Challen *et al.* fit parametric distributions to infector-infectee data for the First Few Hundred cases monitored by PHE [6], while Hart *et al.* also fit a mechanistic model to household infector-infectee data that accounts for the role of symptom onset on infectiousness [25, 24]. Resulting point estimates for the realised mean generation interval (accounting for behaviour and susceptible depletion [17]) range from 3.6–4.9 days. To account for potential variability due to alternate estimates for the latent period, the role of an infected’s behaviour — in particular the efficacy and timing of isolation — in influencing the generation interval, and the potential for unrepresentative sampling of infector-infectee pairs, we draw on the range and reported confidence across early estimates to construct priors on the mean and variance of a gamma distributed infectiousness profile (see Table 1).

We refer to β_j as the infectiousness profile rather than the generation time distribution, as the latter has a small correction due to any overlap in the distributions for time from infection to death and infectiousness by infected-age due to the non-zero hazard of death while infectious (see §2.2.5). Figure 1B shows the ‘default’ infectiousness profile.

To infer the epidemic dynamics based on the death data, it is key to characterise accurately the *infection-to-death distribution*. We denote the probability mass function of the infection-to-death distribution by $\zeta = \{\zeta_j : j \geq 1, \zeta_j \geq 0, \sum_{j=1}^{\infty} \zeta_j = 1\}$, such that ζ_j is the proportion of individuals that die j days after infection, conditional on ultimately dying from SARS-CoV-2. The infection-to-death distribution is the convolution of infection-to-onset and onset-to-death distributions, for which separately there are data.

The infection-to-onset period is estimated from cases with well constrained exposure history [e.g. 2]. For consistency with our choice of infectiousness profile, we adopt the infection-to-onset period used by Ganyani *et al.* [18] when inferring the generation interval distribution above, namely a gamma distribution with mean 5.2 days and standard deviation 2.8 days. For onset to death, early data from Hubei, China, corrected for epidemic growth, suggested an average time 18 days and standard deviation 8.4 days [52]. A report published by the UK government provides distributions for the time from symptom onset to death for over 22,000 patients hospitalised in the UK during the first epidemic wave (up to August 1st 2020), by age range and sex [22]. The aggregate distribution is described by a gamma distribution with mean 14 days (std. dev. 9.8 days), and is stable when re-weighting the age- and sex-dependent distributions to match those of reported deaths in England and Wales over the same period [42], giving some confidence that this distribution is representative for the population we are modelling over the period of study. Similar estimates for the death delay are reported by Sherratt *et al.* [48] based on confidential UK data.

We model the infection-to-death distribution as a negative binomial distribution, chosen for an appropriate shape, computing point estimates for its parameters by matching moments to the convolution of the Ganyani *et al.* [18] infection-to-onset and Harrison *et al.* [22] onset-to-death distributions described above. This leads to the parameters shown in Table 1 and the distribution plot-

Table 1: Summary of key epidemiological parameters drawn from clinical literature, as described in §2.1.3, including default values used later to generate synthetic data (§3.2), and priors used in the Bayesian inference (§3.4). The default values correspond with the plots of the infectiousness profile and infection-to-death distribution shown in Figures 1B and C. The gamma distribution priors are parameterised as $\Gamma(\text{scale}, \text{shape})$, and the normal priors as $\mathcal{N}(\text{mean}, \text{std dev})$.

Quantity	Distn	Param	Symbol	Default	Prior
Infectiousness profile, $\{\beta_j\}$	Γ	mean	β_μ	5.2	$\Gamma(16, 1/3)$
		variance	β_{σ^2}	2.96	$\Gamma(2, 2.2)$
Infection-to-death distribution, $\{\zeta_j\}$	NB	mean	ζ_μ	18.69	$\mathcal{N}(20, 3)$
		dispersion	ζ_ϕ	0.0546	$\mathcal{N}(0.03, 0.025)$
Infection fatality rate	—	—	δ	0.00724	$\mathcal{N}(0.00724, 0.001)$

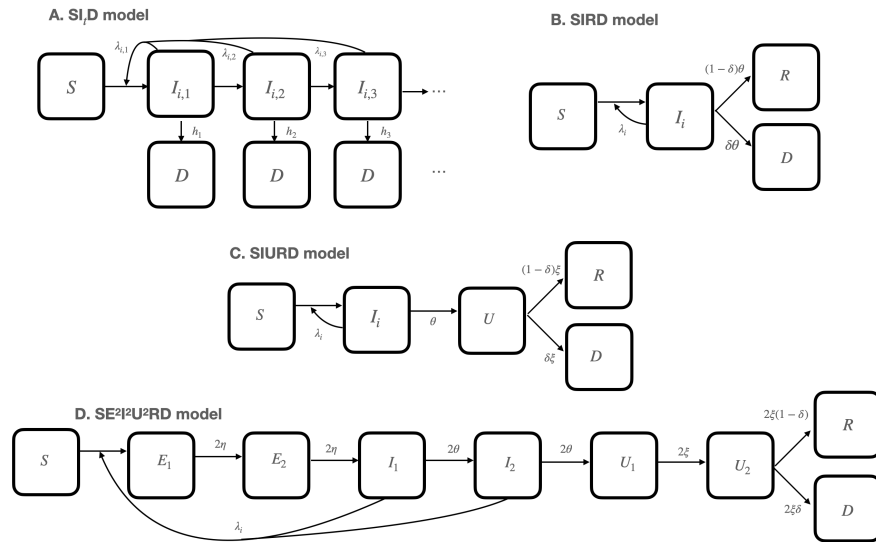
ted in Figure 1C. There remains uncertainty, however, in the infection-to-death distribution owing to uncertainty in the infection-to-onset period [see 2], the censoring effect of unknown symptom onset dates in the hospitalisation data [22], and regional variability (about which we lack data) that may influence the effective average time to death profiles [e.g. 26]. We hence characterise uncertainty in the infection-to-death distribution via priors on the mean and dispersion parameters. We choose priors, shown in Table 1, such that infection-to-death distributions with high prior probability are consistent with the distributions estimated by Harrison et al. [22] and Verity et al. [52] time-to-death distributions.

A further key parameter for inferring epidemic dynamics from death data is the *infection fatality rate* (IFR), δ , which is the risk of death for an infected host, neglecting other host covariates. Seroprevalence of antibody to SARS-CoV-2 — together with evidence regarding the preservation of measurable antibody [30] — provides an estimate of the integrated exposure history to SARS-Cov-2, and enables estimation of the δ . In the results sections we explore the impact of different assumptions about δ : in §3.1 we let δ be a free parameter; in §3.2 we fix its value to the central population-weighted estimate for England from [44]; and then in §3.3 and §3.4 we try to characterise knowledge about δ with a judicious prior (Table 1) that reflects uncertainty and bias that may arise for various reasons, such as non-representative serological surveys, non-uniform prevalence in different risk groups (e.g. in care homes versus the surrounding community) and waning antibody levels [44].

2.2. Epidemic models

In the following we present several simple deterministic epidemic models. The first, called SI_tD , is the central one in this paper and introduced with the goal of being as simple as possible whilst retaining structure by infected age. The subsequent $SIRD_{\Delta D}$, $SIURD$ and $SE^2I^2U^2RD$ models are simple models that are not directly structured by infected age. Figure 2 shows schematics of the various models.

Figure 2: Compartmental structure for models (A) SI_tD , (B) $SIRD/SIRD_{\Delta D}$ model, (C) $SIURD$, and (D) $SE^2I^2U^2RD$, as defined in §2.2. In (B) the $D_{\Delta D}$ compartment indicates deaths that will occur after an additional fixed delay ΔD . The $SIRD$ model is the special case of the $SIRD_{\Delta D}$ model with $\Delta D = 0$. Model (D) includes includes two compartments for each of the exposed (E_1, E_2), infectious (I_1, I_2), and non-infectious (U_1, U_2) states.



2.2.1. A model structured by infected age: SI_tD

We denote by $I_{i,j}$ the number of individuals who on day i became infected $j > 0$ days ago; and the total number of infected individuals on day i is thus $I_i = \sum_{j=1}^{\infty} I_{i,j}$. In this model, between day i and $i + 1$ the number of new infections is $I_{i+1,1} = S_i \sum_{j=1}^{\infty} \lambda_{ij} I_{i,j}$, for suitable λ_{ij} that in general depends on time i (to reflect changing transmission owing to changes in population mixing intensity, e.g., due to NPIs) and infected age j (to reflect non-constant infectiousness of those infected); and the number of new deaths amongst individuals with infected age j is $h_j I_{i,j}$, for suitable h_j . The general infected-age-structured model is therefore

$$S_{i+1} = S_i - I_{i+1,1}, \quad (1a)$$

$$I_{i+1,1} = S_i \sum_{j=1}^{\infty} \lambda_{ij} I_{i,j}, \quad (1b)$$

$$I_{i+1,j+1} = I_{i,j} - h_j I_{i,j}, \quad (1c)$$

$$D_{i+1} = D_i + \sum_{j=1}^{\infty} h_j I_{i,j}, \quad (1d)$$

for $i \in \{0, 1, 2, 3 \dots\}$ and $j \in \{1, 2, 3, \dots\}$.

In this model, in contrast to some common epidemic models including the non-infected-age-structured ones described later in this section, the I denotes individuals who have been “infected” but are not necessarily “infectious”, since individuals of infected age j contribute to infecting susceptibles if and only if $\lambda_{ij} > 0$. This removes the need for an E variable for “exposed but pre-infectious” individuals, or an R variable for “recovered” individuals.

The degrees of freedom in the model need to be controlled via some further modelling choices for the $\{\lambda_{ij}\}$ and $\{h_j\}$. We will write

$$\lambda_{ij} = N^{-1} \rho \alpha_i \beta_j. \quad (2)$$

Here, the infectiousness profile β_j , defined in §2.1.3, is a probability mass function where $\sum_{j=1}^{\infty} \beta_j = 1$, and α_i models the population mixing intensity relative to pre-epidemic social behaviour and is subject to the constraint $\alpha_0 = 1$. In practice α is a composite parameter capturing contact rates, social distancing (including mask wearing) and mobility. A simple model for the impact of introducing NPIs is to enable a sharp reduction in α_i at some change-point time t^* . Enabling a continuous t^* helps in Markov Chain Monte Carlo procedures described later, hence we model α_i to transition from 1 to a new baseline α_b at change-point time t^* via

$$\alpha_i = \mathbb{1}_{i < \lfloor t^* \rfloor} + \alpha_b \mathbb{1}_{i \geq \lceil t^* \rceil} + ((t^* - \lfloor t^* \rfloor) + \alpha_b (\lceil t^* \rceil - t^*)) \mathbb{1}_{i = \lfloor t^* \rfloor} \quad (3)$$

in which $\lfloor \cdot \rfloor$, $\lceil \cdot \rceil$ and $\mathbb{1}_{(\cdot)}$ respectively denote the standard floor, ceiling and indicator functions. An example of $\{\alpha_i\}$ is shown in Figure 1A.

In view of the constraints on $\{\alpha_i\}$ and $\{\beta_j\}$, we include $\rho > 0$ in (2) as a scale factor for the infectious pressure. Parameter $N = S_i + I_i + D_i$ is the population

size, which in this model is constant, because births and non-SARS-CoV-2 death are negligible on the time scales of interest and are thus neglected.

In the language of survival analysis, the $\{h_j : j \geq 1\}$ is the *hazard rate function* for the death of infecteds, such that h_j is the probability that an individual having survived $j - 1$ days post-infection will die on day j . In terms of the infection-to-death distribution, $\{\zeta_j\}$, and its corresponding cumulative distribution function, $P^\zeta = \{P_j^\zeta = \sum_{i=0}^j \zeta_i : j \geq 0\}$,

$$h_j = \frac{\delta \zeta_j}{1 - \delta P_{j-1}^\zeta}, \quad (4)$$

where δ is the infection fatality rate. Derivation of this expression is in Appendix A.

2.2.2. *SIRD and $SIRD_{\Delta D}$ models*

A simple and commonly used model [e.g. 4, 40] that does not maintain structure by infected age is as follows.

$$S_{i+1} = S_i - \lambda_i S_i I_i, \quad (5a)$$

$$I_{i+1} = I_i + \lambda_i S_i I_i - \theta I_i, \quad (5b)$$

$$R_{i+1} = R_i + (1 - \delta)\theta I_i, \quad (5c)$$

$$D_{i+1} = D_i + \delta\theta I_i. \quad (5d)$$

In this model (and those below), the number of new daily infecteds is $I_{i+1,1} = \lambda_i S_i I_i$, where

$$\lambda_i = N^{-1} \rho \alpha_i, \quad (6)$$

and $N = S_i + I_i + R_i + D_i$. There is a close parallel to (1) but with key differences: infectiousness of infected individuals is constant with respect to infected age (β_j is taken to equal 1), and the hazard of removal from being infected is also constant (h_j is taken to equal some constant θ). This model includes an R variable, because the assumption of constant infectiousness of I individuals necessitates a way other than death for an infected individual to cease being infectious. By analogy to (1), the constant hazard implies that the duration in infected state is geometrically distributed.

The $SIRD_{\Delta D}$ model is a generalisation of the $SIRD$ model that replaces equation (5d) with

$$D_{i+\Delta_D+1} = D_{i+\Delta_D} + \delta\theta I_i, \quad (7)$$

which introduces a non-negative integer ‘‘delay-to-death’’ parameter, Δ_D . Introducing a fixed delay in this way is a common modelling strategy to make infection-to-recovery and infection-to-death distributions distinct [e.g. 40] and the latter non-exponential. The basic $SIRD$ model is the special case with delay $\Delta_D = 0$. A generalisation of (7), in which Δ_D is not necessarily an integer, is

$$D_{i+[\Delta_D]+1} = D_{i+[\Delta_D]} + (1 - \Delta_D + [\Delta_D])\delta\theta I_i + (\Delta_D - [\Delta_D])\delta\theta I_{i-1}.$$

For example, if $\Delta_D = 7.2$ then 80% of the deaths will have a delay of 7 days and 20% will have a delay of 8 days.

2.2.3. An SIURD model

A different strategy besides incorporating a delay is to incorporate additional model states [see e.g. 47]. A model variation in this spirit is to incorporate a non-infectious state U that follows infection but precedes recovery or death, i.e.,

$$S_{i+1} = S_i - \lambda_i S_i I_i, \quad (8a)$$

$$I_{i+1} = I_i + \lambda_i S_i I_i - \theta I_i, \quad (8b)$$

$$U_{i+1} = U_i + \theta I_i - \xi U_i, \quad (8c)$$

$$R_{i+1} = R_i + (1 - \delta)\xi U_i, \quad (8d)$$

$$D_{i+1} = D_i + \delta\xi U_i, \quad (8e)$$

in which ξ is a rate constant.

2.2.4. An SE²I²U²RD model

The same strategy can be extended by including more states, for example an exposed (infected but not yet infectious) state, and by representing states using multiple compartments. Such an approach can enable the model dynamics to mimic the delay to peak infectiousness, and the delay between infection and death, and hence indirectly models infected ages [31]. The following model, chosen because it closely matches the approach of Kucharski et al. [37], uses two compartments for each of the E, I and U states.

$$S_{i+1} = S_i - \lambda_i S_i (I_i^{(1)} + I_i^{(2)}), \quad (9a)$$

$$E_{i+1}^{(1)} = E_i^{(1)} + \lambda_i S_i (I_i^{(1)} + I_i^{(2)}) - 2\eta E_i^{(1)}, \quad (9b)$$

$$E_{i+1}^{(2)} = E_i^{(2)} + 2\eta E_i^{(1)} - 2\eta E_i^{(2)}, \quad (9c)$$

$$I_{i+1}^{(1)} = I_i^{(1)} + 2\eta E_i^{(2)} - 2\theta I_i^{(1)}, \quad (9d)$$

$$I_{i+1}^{(2)} = I_i^{(2)} + 2\theta I_i^{(1)} - 2\theta I_i^{(2)}, \quad (9e)$$

$$U_{i+1}^{(1)} = U_i^{(1)} + 2\theta I_i^{(2)} - 2\xi U_i^{(1)}, \quad (9f)$$

$$U_{i+1}^{(2)} = U_i^{(2)} + 2\xi U_i^{(1)} - 2\xi U_i^{(2)}, \quad (9g)$$

$$R_{i+1} = R_i + 2(1 - \delta)\xi U_i^{(2)}, \quad (9h)$$

$$D_{i+1} = D_i + 2\delta\xi U_i^{(2)}, \quad (9i)$$

in which η is a rate constant.

2.2.5. Connection to epidemiological parameters

The foregoing models connect directly to some key epidemiological parameters. A parameter important for characterising whether the epidemic is growing

or declining is the time-varying reproduction number, \mathcal{R}_i , of which there are multiple variations. We adopt for \mathcal{R}_i in this paper the instantaneous reproduction number [17], which is the average number of secondary infecteds generated by a single infected assuming there are no population level changes in susceptibility or mixing behaviour during their infection. We adopt the ‘survival function’ approach to calculating \mathcal{R}_i ; performing the sum over infected age j of the number of new infecteds generated by a host at infected age j weighted by the proportion of hosts still infectious at infected age j [28]. For the SI_tD model:

$$\mathcal{R}_i = S_i \sum_{j=1}^{\infty} \lambda_{i,j} G_j = (S_i/N) \alpha_i \rho \sum_{j=1}^{\infty} \beta_j G_j. \quad (10)$$

where G_j is the survival function of an infected individual associated with the hazard h_j of removal from the I , defined as the probability of not having been removed by day j after infection, which equals $G_j = 1 - \delta P_{j-1}^{\zeta}$. The product $\beta_j G_j$ is the proportion of the \mathcal{R}_i secondary infections that are generated at an infected age j , and therefore when normalised is the (discretised) generation time distribution.

For the $SIRD$, $SIRD_{\Delta D}$, $SIURD$ and $SE^2I^2U^2RD$ models, we compute \mathcal{R}_i from the same definition, noting that infecteds must transit through I before reaching D ($G_j = 1$). Given residence in an infectious state, the number of new infecteds generated per infected ($\lambda_i S_i$) is independent of infected-age. The infectiousness profile is $\beta_j = A_j / \sum(A_j)$, in which A_j is the proportion of infecteds in an infectious state at infected-age j , which can be derived by considering the disease progression defined in equations (5), (8) or (9) as a Markov process (Appendix B). Hence β_j depends on the number of E and I states as well as the parameters governing the total time in the latent (η) and infected classes (θ). However, by construction, the average residence time in an infectious state is θ^{-1} for each of these models (Appendix B). Hence

$$\mathcal{R}_i = \lambda_i S_i \sum_{j=1}^{\infty} A_j = \lambda_i S_i \theta^{-1} = S_i N^{-1} \rho \alpha_i \theta^{-1}. \quad (11)$$

Estimates of \mathcal{R}_i can be used to infer the benefit of mitigating measures such as NPIs for reducing transmission. Of strategic interest is to infer $\text{argmin}_i \{\mathcal{R}_i < 1\}$, as this indicates the time i when NPIs were sufficient to put the epidemic into decline. Also related and important is the peak daily incidence, $\max_i \{S_i - S_{i+1}\}$, in which for each model $S_i - S_{i+1}$ is the number of new infecteds on day i . The basic reproduction number, written \mathcal{R}_0 , is the special case for infection seeded into a fully susceptible population ($S_i \approx N$) and prior to any mitigation ($\alpha_i = 1$).

By design, the SI_tD model explicitly incorporates the infectiousness profile, $\{\beta_j\}$, and the infection-to-death distribution $\{\zeta_j\}$, such that these can be chosen according to clinical data. For the other models in which these are not explicit,

it is helpful to understand what are the implied $\{\beta_j\}$ and $\{\zeta_j\}$. Calculation of these is in Appendix B.

2.2.6. Model initial conditions

In this paper, i indexes the number of days since 15 Feb 2020, which is day $i = 0$. We assume initially zero deaths, $D_0 = 0$, and zero recovered, $R_0 = 0$ (for models including R) and $S_0 = N - I_0$ for parameter I_0 which is to be inferred. For the infected-age-structured model, it is necessary to specify how the I_0 initial infecteds are distributed by infected age $\{I_{0,j}\}$. When the epidemic is growing exponentially the distribution of infecteds by infected age converges to an equilibrium distribution (see Appendix C), thus we assume convergence to this equilibrium distribution in the dynamics prior to day $i = 0$, and for the numerical calculations in this paper we use the equilibrium distribution as the initial condition at $i = 0$.

2.3. Observation model

The epidemic models above are deterministic. It is common to account for variability in the observed number of daily deaths, D_i^{obs} , on day i , in mechanistic models of infectious disease transmission via a negative binomial observation model [e.g. 41, 5]. The negative binomial model admits overdispersion, which is often present in count data on cases or deaths from an infectious disease owing to spatial and demographic heterogeneities, or other unmodelled processes [29]. We adopt the negative binomial model, assuming that

$$D_i^{\text{obs}} \sim \text{NB}(D_i, \phi), \quad (12)$$

independently for each i , where $\text{NB}(\mu, \phi)$ denotes the negative binomial distribution with mean μ and dispersion parameter ϕ defined such that the variance equals $\mu + \phi\mu^2$ [46].

2.4. Inference methods

We denote by Θ^* the free parameters that appear in the respective dynamical models, such that $D_i = D_i(\Theta^*)$, and by $\Theta = (\Theta^*, \phi)$ the vector of all the parameters including in the observation model (12). The data $\mathcal{D} = \{D_i^{\text{obs}}\}$ are the daily deaths indexed by time i . We denote by $P(\mathcal{D}|\Theta)$ the likelihood function for Θ under observation model (12). When adopting a Bayesian approach and specifying a prior distribution $P(\Theta)$ on Θ , the posterior distribution for Θ is then

$$P(\Theta|\mathcal{D}) \propto P(\mathcal{D}|\Theta)P(\Theta). \quad (13)$$

In the results sections below, where we compute point estimates of Θ we do so using the *maximum a posteriori* (MAP) estimator

$$\hat{\Theta}_{\text{MAP}} = \operatorname{argmax}_{\Theta} P(\Theta|\mathcal{D}), \quad (14)$$

then the inferred dynamics, and the epidemiological parameters inferred from them (defined in §2.2.5), are based on solutions of the respective epidemic

model with $\Theta^* = \hat{\Theta}_{\text{MAP}}^*$. The MAP estimator corresponds to the widely used maximum-likelihood estimator (MLE) $\hat{\Theta}_{\text{MLE}} = \text{argmax}_{\Theta} P(\mathcal{D}|\Theta)$ if the priors are uninformative, or more generally if $P(\Theta)$ is constant on a domain that contains $\hat{\Theta}_{\text{MLE}}$ and zero elsewhere. Elsewhere we target the full posterior distribution (13) by sampling from it using Markov chain Monte Carlo (MCMC). We used an adaptive Metropolis-Hastings MCMC algorithm [20] which involves, after a warm-up phase, adapting the covariance matrix of a multi-variate Gaussian proposal distribution according to the covariance of the accepted samples. Convergence to the stationary distribution is hastened with an initial maximisation of the posterior density with respect to Θ using a covariance matrix adaptation evolution strategy (CMA-ES) [21]. These methods were implemented using the PINTS python package [7].

For priors on Θ , those arising from the clinical data are detailed in Table 1 and §2.1.3, in addition to which we specify: $\rho \sim U(1, 10)$ and $I_0 \sim U(1, 5 \times 10^7)$ consistent with wide range of possible values for \mathcal{R}_0 and number of initial infecteds; $t^* \sim \mathcal{N}(31, \sigma = 3)$, where $i = 31$ corresponds to where the Google mobility data, shown in Figure 1A, first suggest a substantial reduction in mobility; and $\alpha_b \sim U(0, 1)$ and $\phi \sim U(0, 1)$ such that both are uniform on their possible ranges of values. For the *SIR*-type models, transition rates in equations (5), (8) and (9) have upper limits such that, for example, no more than the entirety of a compartment can transition out in a single time step. Consequently, we take $\theta \sim U(0, 1)$ for the *SIRD*, *SIRD Δ D* & *SIURD* models, $\xi \sim U(0, 1)$ for the *SIURD* model and $\theta, \eta, \xi \sim U(0, 0.5)$ for the *SE²I²U²RD* model. Due to the additional scaling by $1/\theta$ in the relationship between \mathcal{R}_i and ρ for *SIR*-type models (equation 11) plausible values for ρ are contracted and we use $\rho \sim U(0, 3)$.

3. Results

3.1. A change in mixing intensity, and not ‘herd immunity’, is necessary to explain the epidemic dynamics

A theory after decline from the initial epidemic peak was that the epidemic dynamics were affected little by NPIs and could be explained by ‘herd immunity’ [40], that is, that \mathcal{R}_i in (10) had reduced to below 1 because the number of remaining susceptibles, S_i , had sufficiently decreased to curtail the growth. The theory that depletion of susceptibles is adequate to explain the dynamics can be tested by fitting both models to the data; the ‘herd immunity’ theory corresponds to the a special case of the general model that has a step change in α_i but with the restriction of having constant $\alpha_i = 1$, achieved in the model by setting $\alpha_b = 1$. This supposes no effect from NPIs, and thus that the decline in the epidemic must be on account of the depletion of susceptibles. In this restricted model the parameter t^* is redundant, therefore the restricted model has two fewer parameters than the unrestricted model. To consider this, we use the *SI_tD* model, with the infectiousness profile and infection-to-death distribution parameters fixed to the values shown in Table 1, and noninformative priors on

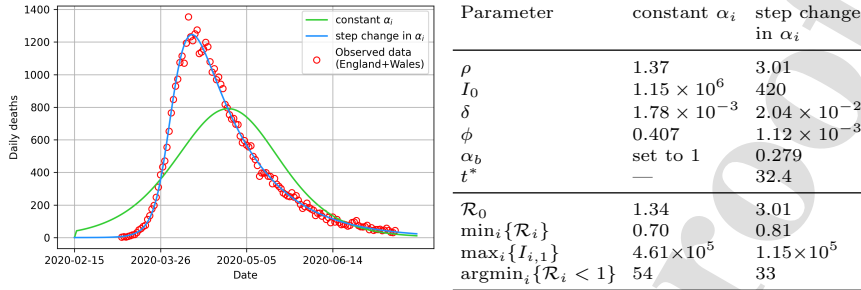


Figure 3: A comparison of the maximum likelihood parameter set SI_tD model predictions for cases in which ρ , I_0 , the infection fatality rate, δ , and NB ϕ are inferred and infectivity is constant (herd immunity), which is a special case of the model in which step function parameters t^* and α_b are also inferred (2 extra parameters; non-herd immunity). Inferred parameters are shown in the table to the right, together with emerging epidemiological quantities of interest below. The $\operatorname{argmin}_i \{\mathcal{R}_i < 1\}$ values of 54 vs. 33 translate to 9th April vs. 19th March 2020, respectively.

ρ , I_0 , t^* , α_b , δ , and ϕ , then fit the model to the England and Wales deaths data described in §2.1.1. We then do likewise for the restricted model, with the extra restriction that $\alpha_b = 1$ and with t^* removed.

The fitted models are shown in Figure 3, from which it is clear that the full model matches well to the data but the restricted model matches very poorly. The difference in the value of the maximised values of the log-posterior, $\log P(\Theta|\mathcal{D})$, for the two models is 260, which for models differing, as here, by two degrees of freedom is overwhelming evidence that the restricted model is inadequate.

Some values of epidemiological parameters in the fitted restricted model also seem implausible; for example, the fitted I_0 is greater than 1 million people, and the IFR is ~ 4 times smaller than the best estimate from O’Driscoll et al. [44] — see Figure 3 for all of the inferred parameter values.

The values of the maximised likelihoods, and visual inspection of the fits, make clear that a change in transmission dynamics over time, via α_i in the model, is necessary to explain the epidemic dynamics during the first outbreak. In other words, it was never plausible that ‘herd immunity’ was responsible for the end of the first outbreak, as confirmed by the subsequent resurgence of infections in autumn of 2020. We hence continue to use the step-change model of mixing intensity (3) in the following sections.

3.2. Inference for epidemiological parameters is unreliable if based on non infected-age structured models

To investigate the impact of model error on the values and uncertainties of inferred parameters — and in particular whether simpler epidemic models such as $SIRD$, $SIRD_{\Delta D}$, $SIURD$ and $SE^2I^2U^2RD$ can be used to infer accurately the epidemiological parameters — we generate synthetic data from the time-structured SI_tD model (§2.1.1) and observation model (12), with the model

Table 2: Details of the simulation study described in section §3.2, including parameter values, in addition to those in Table 1, used to create synthetic data from the SI_tD model; and MAP estimates of the model parameters and derived epidemiological parameters from fitting the various models to the synthetic data. The (–) indicate parameters not relevant to the particular model. The derived parameters are: the basic reproduction number \mathcal{R}_0 ; the minimum value of \mathcal{R} , $\min_i\{\mathcal{R}_i\}$; the index of the day at which \mathcal{R}_i first decreases below 1, $\operatorname{argmin}_i\{\mathcal{R}_i < 1\}$; the maximum number of new infections on any day, $\max_i\{S_i - S_{i+1}\}$; the mean, β_μ , of the infectiousness profile; and the mean of the infection-to-death distribution, ζ_μ . The (A) and (B) tables differ in that (B) involves extra parameters being fixed when the models are fitted, as described in §3.2.

(A)							
Model parameters	Synthetic data	Estimates					
		SI_tD	$SIRD$	$SIRD_{\Delta D}$	$SIURD$	$SE^2I^2U^2RD$ η free	$SE^2I^2U^2RD$ $\eta = 1/5.2$
I_0	860.0	787.2	33.4	3523.3	52.0	2148.2	2617.2
ρ	3.203	3.2534	0.6207	0.5986	1.2473	1.7519	0.9132
t^*	31.57	31.42	45.49	24.25	40.31	30.83	31.51
α_b	0.2814	0.2771	0.6266	0.6232	0.7913	0.0443	0.0637
θ	—	—	0.3879	0.3737	0.9911	0.1325	0.0937
η	—	—	—	—	—	0.1107	—
ξ	—	—	—	—	0.0413	0.0911	0.0939
ΔD	—	—	—	21.6614	—	—	—
ϕ	0.002	0.0014	0.0061	0.0054	0.002	0.0015	0.0015
Derived parameters							
\mathcal{R}_0	3.2	3.25	1.6	1.6	1.26	13.23	9.75
$\min_i\{\mathcal{R}_i\}$	0.79	0.79	0.88	0.88	0.87	0.52	0.54
$\operatorname{argmin}_i\{\mathcal{R}_i < 1\}$	32	32	46	25	41	31	32
$\max_i\{S_i - S_{i+1}\}$, ($\times 1000$)	295.8	304.9	197.3	213.1	423.5	624.4	670.1
β_μ	5.2	5.2	2.58	2.68	1.01	14.7	13.20
ζ_μ	18.69	18.69	3.58	24.68	26.21	28.57	27.53

(B)							
Model parameters	Synthetic data	Estimates					
		SI_tD	$SIRD$	$SIRD_{\Delta D}$	$SIURD$	$SE^2I^2U^2RD$	
I_0	860.0	787.179	68.469	1028.5878	125.5279	555.5233	
ρ	3.203	3.2534	0.424	0.418	0.4573	1.0403	
t^*	31.57	31.4183	46.6492	34.4785	39.6067	34.5997	
α_b	0.2814	0.2771	0.409	0.4142	0.3576	0.289	
ϕ	0.002	0.0014	0.0102	0.01	0.0041	0.0023	
Fixed parameters							
θ	—	—	1/5.2	1/5.2	1/5.2	1/2.974653	
η	—	—	—	—	—	1/2.974653	
ξ	—	—	—	—	1/12.49	1/11.744308	
ΔD	—	—	—	12.49	—	—	
Derived parameters							
\mathcal{R}_0	3.2	3.25	2.2	2.17	2.38	3.09	
$\min_i\{\mathcal{R}_i\}$	0.79	0.79	0.79	0.79	0.75	0.79	
$\operatorname{argmin}_i\{\mathcal{R}_i < 1\}$	32.0	32.0	47.0	35.0	40.0	35.0	
$\max_i\{S_i - S_{i+1}\}$, ($\times 1000$)	295.8	304.9	322.9	332.3	407.6	296.8	

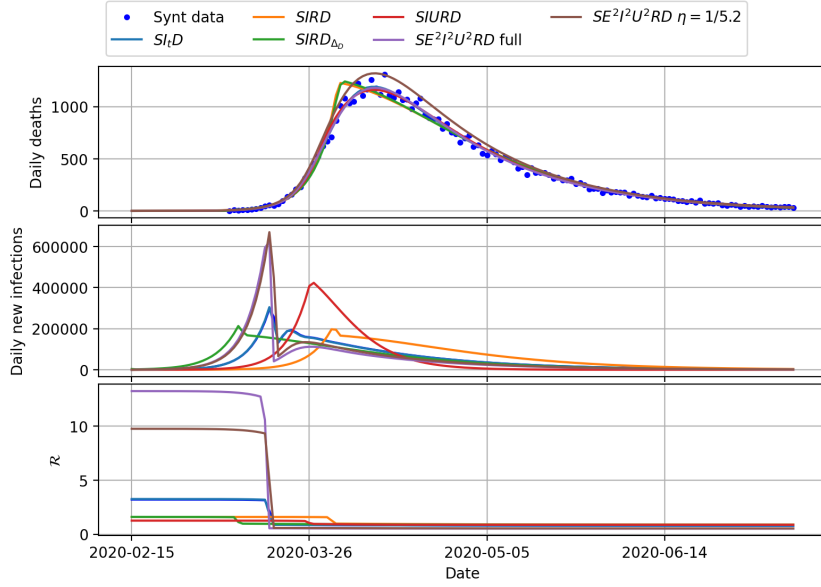
parameters as specified in Table 2, then fit the various models to compare the inferred parameter values with the true ones. We include the data-generating SI_tD in the set of models being fitted, primarily to check that the model parameters are indeed identifiable from the deaths data, but emphasise its advantage in being the data-generating model; the results in this section are hence not suited to advocating for SI_tD over the other models, but rather for understanding how precisely understood mis-specifications of the other models lead to errors in epidemiological quantities inferred from them.

When fitting the models, we fix the IFR, δ , to the default value from Table 1 and treat the other parameters as free. Hence the free parameters include those that determine the implied infectiousness profile and the infection-to-death distributions. The data-generating and fitted parameters values are summarised in Table 2A, and results are plotted in Figure 4.

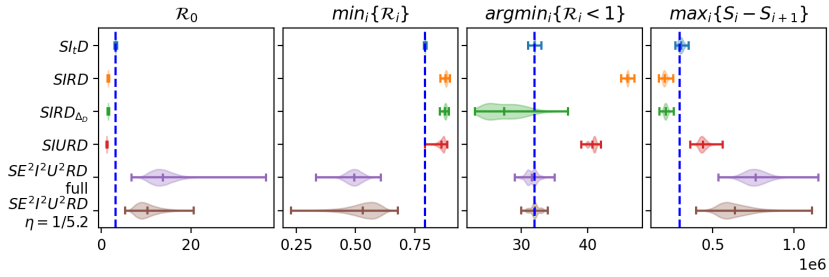
Figure 4A shows that each fitted model matches broadly well the deaths data. The $SIRD$ and $SIRD_{\Delta D}$ models provide similar and relatively reasonable fits to the synthetic death data, but the peak daily deaths in these model fits occurs almost one week before the true peak.

Each of the simpler models — $SIRD$, $SIRD_{\Delta D}$, $SIURD$ — substantially underestimate \mathcal{R}_0 compared to the true value, in each case at least by a factor of 1.6, but estimates from $SEI^2I^2U^2RD$ are vastly too large and with very high posterior variance; see Fig 4B. The latter also had the largest number of free parameters, so we also considered the case with η fixed to the value 1/5.2, which is an appropriate choice for this parameter in the sense that it fixes the model's onset-to-infection interval to match with the mean we assumed when setting the default infection-to-death distribution (see Section 2.1.3). Even with this parameter value fixed, there is little improvement in the inferred values for \mathcal{R}_0 . During the epidemic decline after introduction of NPIs, the misspecification of the model is less impactful on the inferred values for \mathcal{R}_i , but the inferred time, t^* , at which NPIs are inferred to impact mixing intensity is highly variable between models: for example it is inferred around 14 days too late for the $SIRD$ model, 7 days early for $SIRD_{\Delta D}$ and 9 days late for $SIURD$ (Table 2). Compared to the data-generating model, the inferred peak of new infections, $\max_i I_{i,1}$, and the inferred timing of this peak, $\operatorname{argmax}_i I_{i,1}$, are also very inaccurate. For $SIRD$ and $SIRD_{\Delta D}$ models the peak is approximately 36% too low and the timing is respectively 14 days late and 2 days early. For the $SIURD$ model the inferred peak is approximately 43% too high and 9 days too late. Because here we are fixing the IFR and are fitting to death data, all model versions return approximately the same remnant susceptible pool ($S_{i=\text{end}}/N \approx 0.88$) by design. This means that from equation (10) if mixing intensity returns to pre-epidemic levels following the first wave ($\alpha_i = 1$), and all other parameters are fixed, then $\mathcal{R}_i \approx 0.88\mathcal{R}_0$. This provides an upper bound on the median \mathcal{R}_i for a subsequent unmitigated outbreak equal to 2.8 in this simulated example. Because the simple SIR-like models underestimate \mathcal{R}_0 , they also substantially underestimate \mathcal{R}_i for a subsequent unmitigated outbreak, suggesting an upper bound of the median $\mathcal{R}_i \approx 1.4$ ($SIRD$, $SIRD_{\Delta D}$) or 1.1 ($SIURD$) which are highly misleading. In contrast, estimates of the median unmitigated \mathcal{R}_i for the

A



B



C

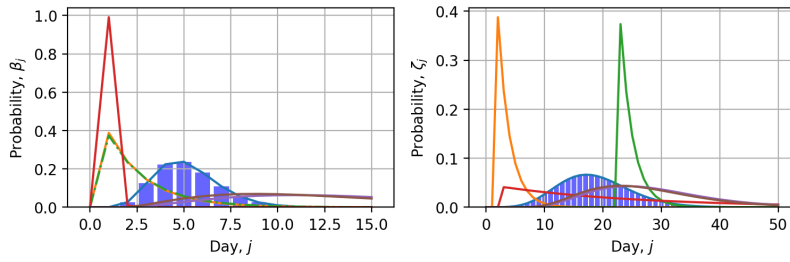


Figure 4: (A) Synthetic data from the SI_tD model generated using parameters in Table 1A and 2, and dynamics of fitted SI_tD , $SIRD$, $SIRD_{\Delta D}$, $SIURD$ and $SE^2I^2U^2RD$ models, corresponding to MAP estimates of parameters per Table 2A. (B) Violin plots of the posterior marginals for derived epidemiological parameters under the different models, with dashes representing (min, mean, max) and the vertical blue line showing the true values. (C) Infectiousness profiles and infection-to-death distributions for the fitted models.

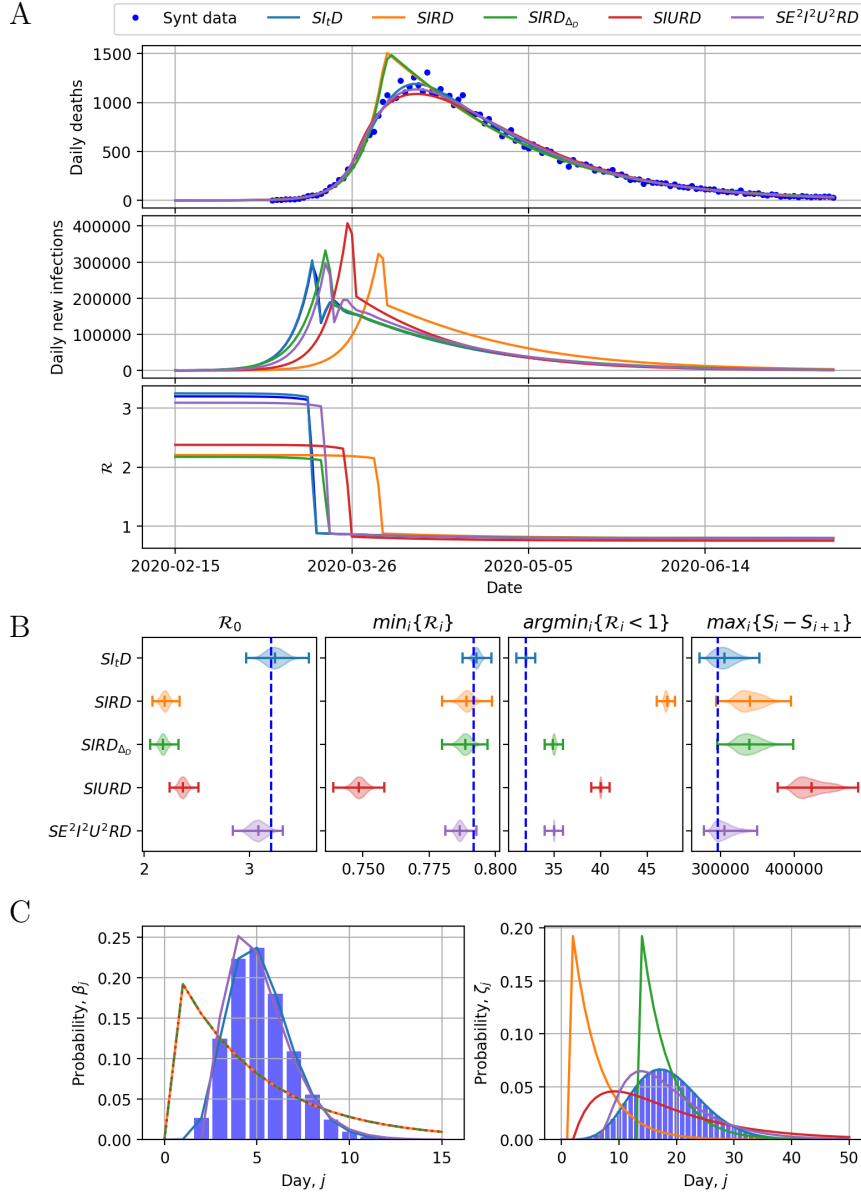


Figure 5: (A) Synthetic data from the SI_tD model (§2.2.1) generated using parameters in Table 1 and 2; and dynamics of fitted SI_tD , $SIRD$, $SIRD_{\Delta D}$, $SIURD$ and $SE^2I^2U^2RD$ models for the parameters indicated in Table 2B; this version has additional fixed parameters, whose values are chosen *a priori* such that for each model the infectiousness profile and infection-to-death distribution, shown in (C), matches as closely as possible to the data-generating ones. (B) Violin plots of the derived epidemiological parameters uncertainties for each model and the synthetic data true values (blue dotted lines).

$SEI^2I^2U^2RD$ models are vastly inflated. In the context of influencing policy-making, the errors arising from the simple SIR-like models are unacceptably large. Figure 4c shows that the implied infectiousness profile and infection-to-death distributions for the fitted models were all quite different to the true ones.

In the preceding investigation with synthetic data, the data-generating SI_tD model was at an advantage over the other models through having its infectiousness profile and infection-to-death distribution specified correctly, and without free parameters related to these that needed estimating. To what extent did the other models perform poorly on account of their extra parameters, rather than their model misspecification? To investigate, we fixed the values of those parameters to “optimal” values in the sense of matching moments of their implied infectiousness profile and infection-to-death distribution matched as closely as possible to the true ones, and with the consequence that the number of free parameters is then the same for each model, including for the data-generating SI_tD model. The parameters are summarised in Table 2B and explained as follows.

To match the mean of the infectiousness profile to the true one (recalling that the relationship between this profile and the model parameters for the various models is explained in Appendix B), for $SIRD$, $SIRD_{\Delta D}$ and $SIURD$ models we take the infectivity rate $\theta = 1/5.2$ where, per Table 1, 5.2 days is the mean of the assumed infectiousness profile. For the infection-to-death distribution, matching to the mean of 18.69 days for $SIRD_{\Delta D}$ entails taking $\Delta_D = 18.69 - 1 - 5.2 = 12.49$ and for $SIURD$ taking $\xi = 1/12.49$ respectively. For the $SE^2I^2U^2RD$ model there are many choices for η and θ that yield an implied infectiousness profile (Appendix B) with appropriate mean, however in general the variance is higher than that for the clinical infectiousness profile (§2.1.3). We therefore fix $\eta = \theta$, which minimises the variance in the implied infectiousness profile. We then select fixed values for θ & ξ to match the means of the clinical infectiousness and time-to-death profiles to those of the implied profiles.

The results of this second fitting, with extra fixed parameters, are shown in and the results are shown in Table 2B and Figure 5. For $SE^2I^2U^2RD$ in particular the results are much improved. Figure 5B shows that the derived epidemiological parameters are much more reliably inferred, although \mathcal{R}_0 is still slightly underestimated and $\text{argmin}_i\{\mathcal{R}_i < 1\}$ overestimated. The better performance is because the inclusion of the E state in this model, with the duplicated compartments for E, I and U, enables the fitted infectiousness profile and infection-to-death distribution to match reasonably closely to the true ones, as shown in Figure 5C. For each of the other models this is not the case, there is a much larger mismatch, as shown in Figure 5C, and a consequence is that the values of inferred epidemiological parameters, particularly \mathcal{R}_0 , remain highly unreliable.

In summary, a model such as $SE^2I^2U^2RD$ could plausibly lead to reasonable inference for epidemiological parameters, provided the number of pre-infectious, infectious and post-infectious compartments is judiciously chosen such that the

implied infectiousness profile and infection-to-death distribution can match suitably closely the clinical data; though if, as in this paper, inference is based on deaths data (§2.1.3) then it appears important that the parameters of such a model are estimated *a priori* from the clinical data.

The results in Table 2 & Figures 4 and 5 are generated from a single synthetic data set; however, the results are representative of the large number repetitions performed. Corresponding results from fitting these models to the real data are in the Supplementary Material (see Table S1 and Figures S1 & S2). For the real data we have no true parameters with which to compare the inferred parameters for the various models, but we do observe analogous disagreement in the inferred parameters, which we attribute to mis-specification of the infectiousness and infection-to-death distributions in the simpler SIR models.

3.3. Bayesian inference leads to small posterior uncertainty when conditioning on the infectiousness profile and infection-to-death distribution

In this section, we adopt a Bayesian approach to characterise the uncertainty in the inferred parameters and predictions. We fixed the infectiousness profile and infection-to-death distribution to their default values from Table 1. The priors for the other parameters are as specified in §2.4, and results for fitting the SI_tD model are shown in Figure 6. Figure 6A shows bi- and uni-variate marginal posterior distributions for the model parameters; these suggest relatively small posterior variance, and small posterior covariance except between ρ and α_b ; and I_0 and α_b . Figure 6B shows a good match between the deaths data and the model dynamics based on the maximum a-posteriori (MAP) model parameter values. Figure 6C shows corresponding inferred new infections, which peak in mid-March, owing to the inferred sharp change in α_i , then decline.

Posterior distributions for derived epidemiological parameters are shown in Fig 6E. The posterior for \mathcal{R}_0 is concentrated in the region ~ 3 to 3.5; $\min_i\{\mathcal{R}_i\}$ is concentrated around 0.795; almost all the posterior mass for $\operatorname{argmin}_i\{\mathcal{R}_i < 1\}$ is on $i = 32$, which is 18th March; and for $\max_i\{I_{i,1}\}$ there is somewhat more posterior uncertainty, with values between $\sim 220,000$ – $400,000$ plausible, and the posterior mode at $\sim 280,000$.

The change in α_i matches quite closely to the Google mobility data, both in the timing, t^* , and the extent of the drop, α_b , as shown in Figure 6D. These numbers are consistent with a reduction in mixing throughout March (before the official lockdown was mandated) due to voluntary changes in mixing intensity, perhaps driven by media coverage and government announcements prior to lockdown. In other words, the results suggest the Google mobility data may accurately reflect the timing and drop in population mixing around the period a lockdown was mandated.

Our inferred values of \mathcal{R}_0 are slightly higher than estimates based on an exponential growth model fitted to the death data [2.6 (2.4–2.9), Loneragan and Chalmers [39]], but consistent with other estimates for the UK [47]. Estimates of \mathcal{R}_i after lockdown are consistent with those from exponential decay models fitted to death data [39], and real-time estimates based on reported cases in the

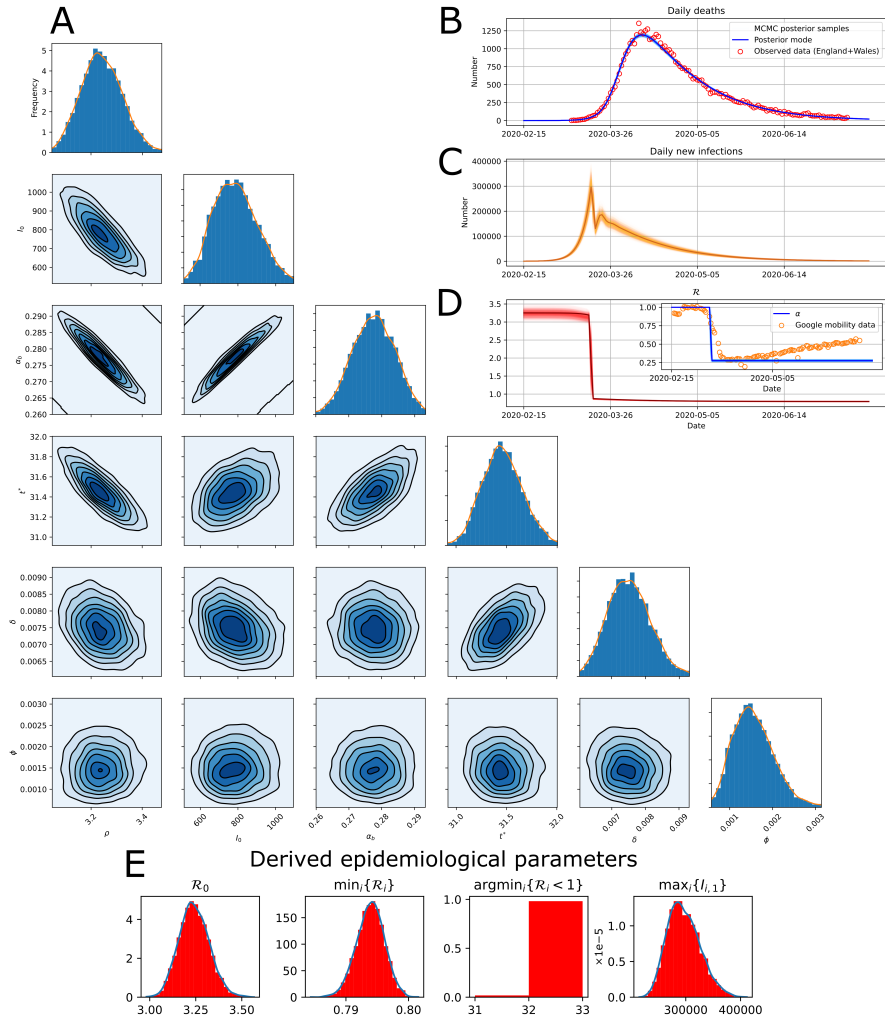


Figure 6: Results from the Bayesian inference described in §3.3. (A) Uni- and bivariate marginal posterior distributions for the model parameters. (B) Daily death data, with superimposed lines showing modelled deaths from 1000 posterior samples (transparent) and MAP parameter values (solid). (C) Inferred daily new infections based on MAP parameter values. (D) Posterior MAP estimate of R_i , with inset showing correspondence between inferred α and average of workplace and transit Google mobility data. (E) Posterior distributions for derived epidemiological parameters.

period April-June 2020 [see Royal Society SET-C [47] and references therein] or surveyed contact rates in late March 2020 [33].

3.4. *Incorporating uncertainty in the infectiousness profile and infection-to-death distribution leads to misleading conclusions owing to model error*

Here we repeat the analysis of the previous section, except this time incorporating uncertainty on the four parameters— β_μ , β_{σ^2} , ζ_μ , ζ_ϕ —describing the infectiousness and infection-to-death distributions. The priors on these parameters are as described in §2.1.3 and summarised in Table 1. Results analogous to Figure 6 are shown in Figure 7. These results show that posterior variances are somewhat larger but not substantially so, although MAP estimates are considerably different.

The MAP parameter gives visually a similarly good fit between model and observed data (Figure 7B), predicting daily new infections to peak in early- to mid-March, after which a sharp drop in new infections is observed (Figure 7C). The inferred mixing intensity in this case still agrees reasonably with the UK Google mobility data (Figure 7D), but is predicted to drop earlier and further than was inferred in §3.3, when the values of those four parameters were instead conditioned upon.

The posteriors for derived parameters show \mathcal{R}_0 is distributed largely between 5 and 7 and peaking at ~ 6 ; $\min_i\{\mathcal{R}_i\}$ is concentrated around 0.72, $\operatorname{argmin}_i\{\mathcal{R}_i < 1\}$ is some time between 11th and 17th March, with posterior peaking on 14th March, and $\max_i\{I_{i,1}\}$ somewhere between $\sim 250,000$ – $500,000$, peaking at $\sim 350,000$ (Figure 7E). These posterior distributions are somewhat in conflict with some informative priors (the influence of the priors can be seen in Supplementary Figure S3). For example, the peak infectiousness occurs ~ 3 days later than estimated by Ganyani et al. [18], and an \mathcal{R}_0 of around 6 is higher than most very early estimates of \mathcal{R}_0 [45]. However, reproduction numbers of around 6 have been reported for other European countries [3].

A possible explanation for the discrepancy in results between this section and §3.3, and with estimates from clinical data such as Ganyani et al. [18], is that it is a consequence of model error. To investigate this, we repeated the Bayesian inference of the present section and of §3.3 for the synthetic data shown in Figure 4. Doing so leads to posterior distributions that are consistent with the data-generating parameters, and consistent with each other, with larger variance under the approach of this section in which there is prior uncertainty on the four additional parameters. Results for these cases are in the Supplementary Material (Supplementary Figures S4 and S5). In other words, the results of the Bayesian inference are as we might expect when the data arise from the assumed model, supporting the possibility that the unexpected results for the real data are on account of model error.

If so, what are the possible sources of model error? The negative binomial observation model heavily penalises discrepancies between modelled and observed deaths when the number of daily deaths is small, especially when the over-dispersion parameter, ϕ is small. The MAP estimate of ϕ is indeed small—around 10^{-3} —especially so with the extra four degrees of freedom considered in

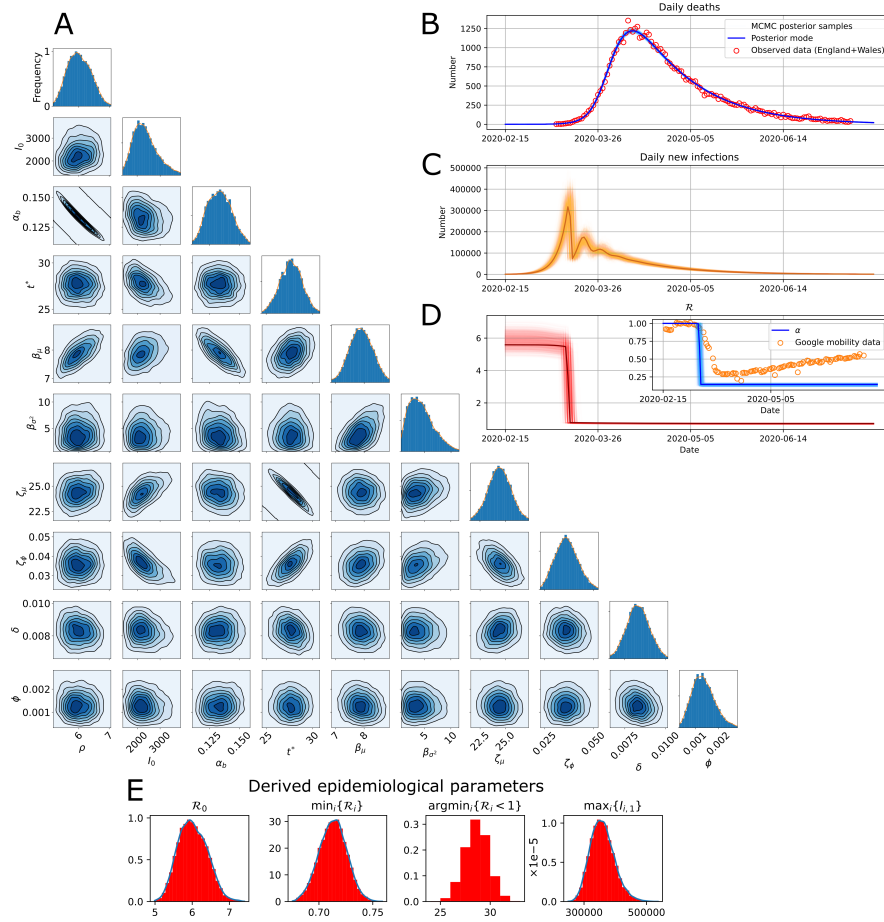


Figure 7: Results from the Bayesian inference described in §3.4. The interpretation of (A–E) are as for Fig 6. These results are analogous to those in Fig 6, except here the analysis involved placing priors on—rather than conditioning on—the parameters β_{μ} , β_{σ^2} , ζ_{μ} , ζ_{ϕ} of the infectiousness profile and infection-to-death distribution.

this section, which enable the deterministic part of the model to explain more of the variability in the data. A consequence is that the inference is dominated by the model having to match well the data in the period when the number of daily deaths is small but growing rapidly. This is the period when the deterministic component of the model, which assumes a homogeneously mixing population and neglects stochastic effects in the epidemic process, describes the dynamics least well. In summary: enabling the extra four degrees of freedom in this section appears to lead to an overfitted model, which leads to underestimation of the overdispersion parameter (and variance). The posterior consequently has low variance, but with inference dominated by the dynamics in a period when the model is likely to be least accurate, and the low-variance posterior is located in a region of parameter space that suggests model error is making the results unreliable.

The Supplementary Material contains results of some further exploration of this issue, including where we have fixed ϕ at two larger values, 0.01 and 0.1, (Supplementary Figures S6 and S7), which inflates posterior variance and shifts the MAP estimates to more plausible values, e.g. 4.5 for \mathcal{R}_0 ; and a case with the Negative Binomial observation model replaced by a Gaussian model (Supplementary Figure S8), such that the inference is not dominated by the early period with low prevalence, and for this case again the MAP estimates of parameters are more in accordance with other sources and with the results of §3.3.

4. Summary and discussion

We have explored various simple epidemic models, particularly investigating the importance of accurately characterising the infectiousness profile and the infection-to-death time distribution. These distributions are well characterised in the clinical literature and easy to incorporate into an SIR-type model structured by infected age. The results show that it is essential to incorporate them in models used for inferring the underlying epidemic dynamics from death data. Basic SIR models implicitly mis-specify the infectiousness and infection-to-death distributions and lead to highly misleading inferences about the impact of NPIs, the time and magnitude of peak infections, and the basic reproduction number.

In the particular context of the spring 2020 SARS-CoV-2 outbreak in England and Wales, we used the infected-age-structured model, SI_tD , to compare the hypothesis that the epidemic decline was on account of NPIs versus the hypothesis, which was entertained at the time, that the decline was owing to ‘herd immunity’. When using distributions for the infectious profile and infection-to-death distribution that are drawn from clinical literature, but fitting all the other model parameters, we found very strong evidence in favour of efficacious NPIs, rather than herd immunity, as the explanation for the epidemic decline (Figure 3). This conclusion was in spite of accommodating the possibility of an implausibly low IFR, δ , and high initial number infected, I_0 , both of which are to the benefit of the herd immunity hypothesis. This indicates that the

death data, combined with infectiousness profile and infection-to-death distributions available from clinical data, enable dismissing of the ‘herd immunity’ hypothesis, without needing to know the IFR very accurately.

The limitations of very basic SIR-type models in mis-specifying the infectiousness profile and the infection-to-death has been pointed out by Keeling and Rohani [34], Wearing et al. [53]. Our work extends this to consider the impact of mis-specification when the model is used to infer epidemiological parameters and dynamics from data, and particularly when there is an abrupt change in population mixing intensity. Our findings show major errors in the inferred dynamics—for example, timing of changes in population mixing can be wrong by weeks—using basic SIR-type models, suggesting that untangling the impact of NPIs adopted in quick succession [e.g. 8] may be especially error-prone if using a simple SIR-like model. Models such as $SE^2I^2U^2RD$ that incorporate multiple infected states (in this case pre-infectious, infectious, post-infectious) and multiple sub-compartments per state can potentially perform well if the number of states, the number of sub-compartments, and the values of rate parameters are suitably chosen [see also 31, 32], but in our view this entails more complexity and no advantage compared using the SI_tD model.

Some models for SARS-CoV-2 transmission were calibrated using multiple data streams, particularly in the UK where data on cases, hospitalisations, prevalence and seroprevalence were also available [47]. Doing so diminishes the influence of the deaths data, and hence diminishes the influence of how the infection-to-death distribution is specified, on the inferred quantities. However, other data streams bring their own modelling challenges — for example, how to connect the underlying epidemic model to data on cases when testing capacity is severely limited but rapidly growing — and entail added complexity that we have sought to avoid by basing inference upon the deaths data.

Our estimates for \mathcal{R}_0 are significantly larger when we take a fully Bayesian approach and fit parameters describing the infectiousness profile and infection-to-death distribution. Much of the variation between early estimates from case data has been attributed to assumptions for the generation time distribution [45], which is closely linked to the infectiousness profile. The need to quantify uncertainty in estimates of \mathcal{R}_0 for use in policy-making has been highlighted, for example by Royal Society SET-C [47]. Uncertainty analysis of a spatial agent-based model suggests that parameters that largely determine the infectiousness profile (exposure to onset delays, onset to isolation delays) remain significant sources of overall uncertainty [15]. We found that incorporating uncertainty by placing priors with relatively large variance on the four parameters characterising the infectiousness profile and infection-to-death distribution — as opposed to conditioning on them — yielded low-variance posteriors concentrated in implausible regions of parameter space, contrary to our initial expectation that the additional prior uncertainty would largely just inflate the posterior variance. We understand this to be the impact of model error, with the negative binomial observation model leading to inference being dominated by a period in which the deterministic dynamics are least reliable. A tendency of deterministic models to yield overconfident and error-prone estimates of \mathcal{R}_i from early incidence data

has also been documented [35].

We have argued in favour of an SI_tD model structured by infected age over a basic SIR model, especially when the goal is inference from death data, but this model is itself only a coarse population-level model with many limitations. The model assumes homogeneous mixing, and neglects stochastic effects which might be appreciable when prevalence is low. To keep the model simple, we accommodated only a step change in the population mixing intensity, α_i . For the observation model for deaths, we assumed a negative binomial model and conditional independence between different days, which are strong parametric assumptions upon which the inference can be sensitive. We did not attempt to model subtleties such as time-dependence in the IFR due to, for e.g., changes in hospital fatality depending on case loads within hospitals [13], or the time-dependence of generation intervals [24]. Hence there are many possible sources of model error. In the context of inference with this model, we have found that affording too much freedom through high prior variance on parameters enables overfitting of the deterministic part and amplifies the impact of model error. Erring on the side of choosing priors with high variance is therefore not necessarily a conservative strategy, at least when model error is non-negligible.

Software Availability

Open source software to reproduce the results in this paper is available at https://github.com/DGWhittaker/nottingham_covid_modelling.

Funding

This work was supported by the Wellcome Trust [grant number 212203/Z/18/Z]. DGW, ADH-R and GRM acknowledge support from the Wellcome Trust via a Senior Research Fellowship to GRM. KJB acknowledges support from a University of Nottingham Anne McLaren Fellowship. This research was funded in whole, or in part, by the Wellcome Trust [212203/Z/18/Z]. For the purpose of open access, the author has applied a CC-BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Acknowledgements

We thank Frank Ball and Theodore Kypraios for helpful discussions.

References

- [1] Ashcroft, P., Huisman, J. S., Lehtinen, S., Bouman, J. A., Althaus, C. L., Regoes, R. R. and Bonhoeffer, S. (2020) COVID-19 infectivity profile correction. *arXiv preprint arXiv:2007.06602*.

- [2] Backer, J. A., Klinkenberg, D. and Wallinga, J. (2020) Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020. *Eurosurveillance*, **25**. URL: <https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2020.25.5.2000062>.
- [3] Billah, M. A., Miah, M. M. and Khan, M. N. (2020) Reproductive number of coronavirus: A systematic review and meta-analysis based on global level evidence. *PloS one*, **15**, e0242128.
- [4] Britton, T. (2020) Basic estimation-prediction techniques for covid-19, and a prediction for stockholm. *medRxiv*. URL: <https://www.medrxiv.org/content/early/2020/05/14/2020.04.15.20066050>.
- [5] Cauchemez, S. and Ferguson, N. M. (2008) Likelihood-based estimation of continuous-time epidemic models from time-series data: application to measles transmission in London. *Journal of the Royal Society Interface*, **5**, 885–897.
- [6] Challen, R., Brooks-Pollock, E., Tsaneva-Atanasova, K. and Danon, L. (2021) Meta-analysis of the severe acute respiratory syndrome coronavirus 2 serial intervals and the impact of parameter uncertainty on the coronavirus disease 2019 reproduction number. *Statistical Methods in Medical Research*, 09622802211065159.
- [7] Clerx, M., Robinson, M., Lambert, B., Lei, C. L., Ghosh, S., Mirams, G. R. and Gavaghan, D. J. (2019) Probabilistic Inference on Noisy Time Series (PINTS). *Journal of Open Research Software*, **7**, 23.
- [8] Dehning, J., Zierenberg, J., Spitzner, F. P., Wibral, M., Neto, J. P., Wilczek, M. and Priesemann, V. (2020) Inferring change points in the spread of covid-19 reveals the effectiveness of interventions. *Science*, **369**. URL: <https://science.sciencemag.org/content/369/6500/eabb9789>.
- [9] Department for Business, Energy & Industrial Strategy and Department for Digital, Culture, Media & Sport (2020) *Working safely during coronavirus (COVID-19)*. URL: <https://www.gov.uk/guidance/working-safely-during-coronavirus-covid-19>. Accessed November 30, 2020.
- [10] DHSC (2020) *Daily tests processed and testing capacity (UK): 20 March to 22 September 2020*. URL: <https://www.gov.uk/government/publications/daily-tests-processed-and-testing-capacity-uk-20-march-to-22-september-2020>. Accessed December 7, 2020.
- [11] Diekmann, O. and Heesterbeek, J. A. P. (2000) *Mathematical Epidemiology of Infectious Diseases: Model Building, Analysis and Interpretation*. Wiley Series in Mathematical & Computational Biology. Wiley.
- [12] Diekmann, O., Othmer, H. G., Planque, R. and Bootsma, M. C. (2021) On discrete time epidemic models in Kermack-McKendrick form. *medRxiv*.

- [13] Docherty, A. B., Mulholland, R. H., Lone, N. I., Cheyne, C. P., De Angelis, D., Diaz-Ordaz, K., Donoghue, C., Drake, T. M., Dunning, J., Funk, S., García-Fiñana, M., Girvan, M., Hardwick, H. E., Harrison, J., Ho, A., Hughes, D. M., Keogh, R. H., Kirwan, P. D., Leeming, G., Nguyen-Van-Tam, J. S., Pius, R., Russell, C. D., Spencer, R., Tom, B. D., Turtle, L., Openshaw, P. J., Baillie, J. K., Harrison, E. M. and Semple, M. G. (2020) Changes in uk hospital mortality in the first wave of covid-19: the isaric who clinical characterisation protocol prospective multicentre observational cohort study. *medRxiv*. URL: <https://www.medrxiv.org/content/early/2020/12/22/2020.12.19.20248559>.
- [14] Dunn, P., Allen, L., Cameron, G. and Alderwick, H. (2020) *COVID-19 policy tracker*. URL: <https://www.health.org.uk/news-and-comment/charts-and-infographics/covid-19-policy-tracker>. Accessed December 7, 2020.
- [15] Edeling, W., Arabnejad, H., Sinclair, R., Suleimenova, D., Gopalakrishnan, K., Bosak, B., Groen, D., Mahmood, I., Crommelin, D. and Coveney, P. V. (2021) The impact of uncertainty on predictions of the covidsim epidemiological code. *Nature Computational Science*, **1**, 128–135.
- [16] Ferretti, L., Wymant, C., Kendall, M., Zhao, L., Nurtay, A., Abeler-Dörner, L., Parker, M., Bonsall, D. and Fraser, C. (2020) Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science*, **368**. URL: <https://science.sciencemag.org/content/368/6491/eabb6936>.
- [17] Fraser, C. (2007) Estimating individual and household reproduction numbers in an emerging epidemic. *PloS ONE*, **2**, e758.
- [18] Ganyani, T., Kremer, C., Chen, D., Torneri, A., Faes, C., Wallinga, J. and Hens, N. (2020) Estimating the generation interval for coronavirus disease (COVID-19) based on symptom onset data, March 2020. *Euro-surveillance*, **25**. URL: <https://www.eurosurveillance.org/content/10.2807/1560-7917.ES.2020.25.17.2000257>.
- [19] Google (2020) *COVID-19 Community Mobility Reports*. URL: <https://www.google.com/covid19/mobility/>. Accessed November 30, 2020.
- [20] Haario, H., Saksman, E. and Tamminen, J. (2001) An Adaptive Metropolis Algorithm. *Bernoulli*, **7**, 223–242. URL: <https://www.jstor.org/stable/3318737>. Publisher: International Statistical Institute (ISI) and Bernoulli Society for Mathematical Statistics and Probability.
- [21] Hansen, N., Müller, S. D. and Koumoutsakos, P. (2003) Reducing the Time Complexity of the Derandomized Evolution Strategy with Covariance Matrix Adaptation (CMA-ES). *Evolutionary Computation*, **11**, 1–18.

- [22] Harrison, E., Docherty, A. and Semple, C. (2020) *COVID-19: time from symptom onset until death in UK hospitalised patients*. URL: <https://assets.publishing.service.gov.uk>. Accessed November 30, 2020.
- [23] Hart, W., Maini, P., Yates, C. and Thompson, R. (2020) A theoretical framework for transitioning from patient-level to population-scale epidemiological dynamics: influenza a as a case study. *Journal of the Royal Society Interface*, **17**, 20200230.
- [24] Hart, W. S., Abbott, S., Endo, A., Hellewell, J., Miller, E., Andrews, N., Maini, P. K., Funk, S. and Thompson, R. N. (2022) Inference of the SARS-CoV-2 generation time using UK household data. *eLife*, **11**, e70767.
- [25] Hart, W. S., Maini, P. K. and Thompson, R. N. (2021) High infectiousness immediately before COVID-19 symptom onset highlights the importance of continued contact tracing. *eLife*, **10**, e65534.
- [26] Hawryluk, I., Mellan, T. A., Hoeltgebaum, H., Mishra, S., Schnekenberg, R. P., Whittaker, C., Zhu, H., Gandy, A., Donnelly, C. A., Flaxman, S. and Bhatt, S. (2020) Inference of covid-19 epidemiological distributions from brazilian hospital data. *Journal of The Royal Society Interface*, **17**, 20200596. URL: <https://royalsocietypublishing.org/doi/abs/10.1098/rsif.2020.0596>.
- [27] He, X., Lau, E. H., Wu, P., Deng, X., Wang, J., Hao, X., Lau, Y. C., Wong, J. Y., Guan, Y., Tan, X. et al. (2020) Temporal dynamics in viral shedding and transmissibility of covid-19. *Nature medicine*, **26**, 672–675.
- [28] Heffernan, J. M., Smith, R. J. and Wahl, L. M. (2005) Perspectives on the basic reproductive ratio. *Journal of the Royal Society Interface*, **2**, 281–293.
- [29] Held, L., Hens, N., Wallinga, J. and O’Neill, P. (2019) *Handbook of Infectious Disease Data Analysis*. CRC Press LLC, Milton. Available from: ProQuest Ebook Central, accessed 18 December 2020.
- [30] Huang, A. and Garcia-Carreras, B.ans Hitchings, M. e. a. (2020) A systematic review of antibody mediated immunity to coronaviruses: kinetics, correlates of protection, and association with severity. *Nat Commun*, **11**, 4704. URL: <https://doi.org/10.1038/s41467-020-18450-4>.
- [31] Hurtado, P. and Kiro Singh, A. (2019) Generalizations of the ‘linear chain trick’: incorporating more flexible dwell time distributions into mean field ode models. *J. Math. Biol.*, **79**.
- [32] İşlier, Z. G., Güllü, R. and Hörmann, W. (2020) An exact and implementable computation of the final outbreak size distribution under erlang distributed infectious period. *Mathematical Biosciences*, **325**, 108363.

- [33] Jarvis, C. I., Van Zandvoort, K., Gimma, A., Prem, K., Klepac, P., Rubin, G. J. and Edmunds, W. J. (2020) Quantifying the impact of physical distance measures on the transmission of covid-19 in the uk. *BMC medicine*, **18**, 1–10.
- [34] Keeling, M. J. and Rohani, P. (2008) *Modeling Infectious Diseases in Humans and Animals*. Princeton University Press. URL: <http://www.jstor.org/stable/j.ctvc4gk0>.
- [35] King, A. A., Domenech de Cellès, M., Magpantay, F. M. and Rohani, P. (2015) Avoidable errors in the modelling of outbreaks of emerging pathogens, with special reference to ebola. *Proceedings of the Royal Society B: Biological Sciences*, **282**, 20150347.
- [36] Kissler, S. M., Fauver, J. R., Mack, C., Olesen, S. W., Tai, C., Shiue, K. Y., Kalinich, C. C., Jednak, S., Ott, I. M., Vogels, C. B., Wohlgenuth, J., Weisberger, J., DiFiori, J., Anderson, D. J., Mancell, J., Ho, D. D., Grubaugh, N. D. and Grad, Y. H. (2020) Sars-cov-2 viral dynamics in acute infections. *medRxiv*. URL: <https://www.medrxiv.org/content/early/2020/12/01/2020.10.21.20217042>.
- [37] Kucharski, A. J., Russell, T. W., Diamond, C., Liu, Y., Edmunds, J., Funk, S., Eggo, R. M., Sun, F., Jit, M., Munday, J. D. et al. (2020) Early dynamics of transmission and control of covid-19: a mathematical modelling study. *The lancet infectious diseases*, **20**, 553–558.
- [38] Lehtinen, S., Ashcroft, P. and Bonhoeffer, S. (2020) On the relationship between serial interval, infectiousness profile and generation time. *medRxiv*. URL: <https://www.medrxiv.org/content/early/2020/09/18/2020.09.18.20197210>.
- [39] Lonergan, M. and Chalmers, J. D. (2020) Estimates of the ongoing need for social distancing and control measures post-“lockdown” from trajectories of covid-19 cases and mortality. *European Respiratory Journal*, **56**.
- [40] Lourenço, J., Pinotti, F., Thompson, C. and Gupta, S. (2020) The impact of host resistance on cumulative mortality and the threshold of herd immunity for sars-cov-2. *medRxiv*. URL: <https://www.medrxiv.org/content/early/2020/10/01/2020.07.15.20154294>.
- [41] Mathews, J. D., McCaw, C. T., McVernon, J., McBryde, E. S. and McCaw, J. M. (2007) A biological model for influenza transmission: Pandemic planning implications of asymptomatic infection and immunity. *PLOS ONE*, **2**, 1–6. URL: <https://doi.org/10.1371/journal.pone.0001220>.
- [42] ONS (2020) *Deaths registered weekly in England and Wales, provisional: week ending 13 November 2020*. URL: <https://www.ons.gov.uk/>. Accessed November 30, 2020.

- [43] ONS (2020) *Population estimates*. URL: <https://www.ons.gov.uk/>. Accessed November 30, 2020.
- [44] O’Driscoll, M., Dos Santos, G. R., Wang, L., Cummings, D. A., Azman, A. S., Paireau, J., Fontanet, A., Cauchemez, S. and Salje, H. (2020) Age-specific mortality and immunity patterns of SARS-CoV-2. *Nature*, 1–9.
- [45] Park, S. W., Bolker, B. M., Champredon, D., Earn, D. J., Li, M., Weitz, J. S., Grenfell, B. T. and Dushoff, J. (2020) Reconciling early-outbreak estimates of the basic reproductive number and its uncertainty: framework and applications to the novel coronavirus (SARS-CoV-2) outbreak. *MedRxiv*.
- [46] Robinson, M. D. and Smyth, G. K. (2008) Small-sample estimation of negative binomial dispersion, with applications to SAGE data. *Biostatistics*, **9**, 321–332. URL: <https://academic.oup.com/biostatistics/article/9/2/321/353777>. Publisher: Oxford Academic.
- [47] Royal Society SET-C (2020) *Reproduction number (R) and growth rate (r) of the COVID-19 epidemic in the UK: methods of estimation, data sources, causes of heterogeneity, and use as a guide in policy formulation*. URL: <https://royalsociety.org>. Members of the Royal Society SET-C and members of the SET-C sub-group.
- [48] Sherratt, K., Abbott, S., Meakin, S. R., Hellewell, J., Munday, J. D., Bosse, N., working group, C. C.-., Jit, M. and Funk, S. (2021) Exploring surveillance data biases when estimating the reproduction number: with insights into subpopulation transmission of covid-19 in england. *Philosophical Transactions of the Royal Society B*, **376**, 20200283.
- [49] Sun, K., Wang, W., Gao, L., Wang, Y., Luo, K., Ren, L., Zhan, Z., Chen, X., Zhao, S., Huang, Y., Sun, Q., Liu, Z., Litvinova, M., Vespignani, A., Ajelli, M., Viboud, C. and Yu, H. (2020) Transmission heterogeneities, kinetics, and controllability of SARS-CoV-2. *Science*. URL: <https://science.sciencemag.org/content/early/2020/11/23/science.abe2424>.
- [50] UK Government (2020) *The Health Protection (Coronavirus, Restrictions) (England) Regulations 2020*. URL: <https://www.legislation.gov.uk/uksi/2020/350/regulation/6/2020-03-26>. Accessed November 30, 2020.
- [51] — (2020) *The Health Protection (Coronavirus, Wearing of Face Coverings in a Relevant Place) (England) Regulations 2020*. URL: <https://www.legislation.gov.uk/uksi/2020/791/schedule>. Accessed November 30, 2020.
- [52] Verity, R., Okell, L. C., Dorigatti, I., Winskill, P., Whittaker, C., Imai, N., Cuomo-Dannenburg, G., Thompson, H., Walker, P. G. T., Fu, H., Dighe, A., Griffin, J. T., Baguelin, M., Bhatia, S., Boonyasiri, A., Cori,

A., Cucunubá, Z., FitzJohn, R., Gaythorpe, K., Green, W., Hamlet, A., Hinsley, W., Laydon, D., Nedjati-Gilani, G., Riley, S., van Elsland, S., Volz, E., Wang, H., Wang, Y., Xi, X., Donnelly, C. A., Ghani, A. C. and Ferguson, N. M. (2020) Estimates of the severity of coronavirus disease 2019: a model-based analysis. *The Lancet. Infectious diseases*, **20**, 669–677.

[53] Wearing, H. J., Rohani, P. and Keeling, M. J. (2005) Appropriate models for the management of infectious diseases. *PLoS medicine*, **2**, e174.

Appendix A. Expression for hazards, h_j , in the SI_tD model in terms of infection-to-death distribution

Denote by T the number of days between infection and subsequent death due to infection. The goal is to determine hazards, h_j in (1), such that the proportion with $T = 1$, $T = 2$, etc., is consistent with a prescribed infection-to-death distribution $\{\zeta_j\}$. Denoting by $p(T = j)$ the proportion of infected individuals who die at precisely time j after infection then we require h_j such that $p(T = j) = \delta\zeta_j$.

The hazard, h_j , is the proportion who, having survived up until j , then die on day j , hence $h_j = p(T = j|T \geq j)$. By standard axioms we have

$$h_j = p(T = j|T \geq j) = \frac{p(T = j \cap T \geq j)}{p(T \geq j)} = \frac{p(T = j)}{p(T \geq j)} = \frac{\delta\zeta_j}{1 - \delta P_{j-1}^\zeta},$$

which is expression (4). In the limit $j \rightarrow \infty$, then $h_j \rightarrow 0$ such that the proportion of infected individuals that ultimately die as a consequent of infection equals the prescribed IFR, δ .

Appendix B. Infectiousness profiles and time-to-death distributions for $SIRD$, $SIRD_{\Delta D}$ and $SE^2I^2U^2RD$ models

For our models without a latent state defined by equations (5) & (8), the proportion of newly infected hosts remaining in an infectious state (I) at infected-age j is simply those who are not removed by the previous time-step $A_j = (1 - \theta)^{j-1}$, and $\beta_j = \theta(1 - \theta)^{j-1}$. The probability mass function for death (as before conditioning on death as the outcome of the infection) can be calculated by summing over the probabilities of Markov chains that arrive in D at infected-age j . For the $SIRD$ model:

$$\zeta_j = \theta \begin{cases} 0, & j \leq 1 \\ 1, & j = 2 \\ (1 - \theta)^{j-2}, & j \geq 3 \end{cases}$$

The distribution of arrival times in D is shifted for the $SIRD_{\Delta D}$ model,

$$\zeta_j = \theta \begin{cases} 0, & j \leq \Delta D + 1 \\ 1, & j = \Delta D + 2 \\ (1 - \theta)^{j-2}, & j \geq \Delta D + 3 \end{cases}.$$

For the *SIURD* model:

$$\zeta_j = \theta \xi \begin{cases} 0, & j \leq 2 \\ 1, & j = 3 \\ \sum_{k=0}^{j-3} (1 - \theta)^k (1 - \xi)^{j-3-k}, & j \geq 4 \end{cases}.$$

For the slightly more complicated *SE²I²U²RD* model described in equation (9) we construct a matrix describing the Markov transitions following infection [12]

$$T = \begin{pmatrix} 1 - 2\eta & 2\eta & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 - 2\eta & 2\eta & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 - 2\theta & 2\theta & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 - 2\theta & 2\theta & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 - 2\xi & 2\xi & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 - 2\xi & 2\xi(1 - \delta) & 2\delta\xi \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix},$$

where $T_{n,k}$ is the proportion transitioning from state n to state k . Describing a newly infected individual in E_1 by a state vector $i_0 = (1, 0, 0, 0, 0, 0, 0, 0)$, the proportion of new infections in I_1 at infected-age j is $f_j^{I_1} = \sum_{k=1}^8 i_0 (T^{j-1})_{k,3} = (T^{j-1})_{1,3}$, where T^j indicates the j th power of T . Similarly $f_j^{I_2} = (T^{j-1})_{1,4}$. Hence $A_j = (T^{j-1})_{1,3} + (T^{j-1})_{1,4}$. If $f_{j'}^{I_i}$ denotes the proportion of infections in I_i at time-step j' after arriving in I_i , then $\sum (f_j^{I_i}) = \sum (f_{j'}^{I_i}) = \sum_{j'=1}^{\infty} (1 - 2\theta)^{j'-1} = 1/(2\theta)$, and thus $\beta_j = \theta((T^{j-1})_{1,3} + (T^{j-1})_{1,4})$. The proportion of deaths occurring at infected-age j is $\zeta_j = \frac{1}{\delta}((T^{j-1})_{1,8} - (T^{j-2})_{1,8})$.

Appendix C. Pseudo-steady distribution of infecteds by infected age

During a phase of epidemic growth with $S_i \approx N$ and $\lambda_{ij} = \lambda_j$ constant with respect to i , the distribution of I s by infected age reaches a dynamic equilibrium,

$$\frac{(I_{i,1}, \dots, I_{i,p})}{\sum_{j=1}^p I_{i,j}} = \frac{(I_{i+1,1}, \dots, I_{i+1,p})}{\sum_{j=1}^p I_{i+1,j}},$$

and the number of infecteds grows exponentially,

$$\sum_{j=1}^p I_{i+1,j} = c \sum_{j=1}^p I_{i,j} \text{ for some } c > 1.$$

Together these imply $I_{i+1,j} = cI_{i,j}$. Then under model (1), we get,

$$cI_{i,1} = N^{-1}S_i \sum_j \lambda_j I_{i,j},$$
$$cI_{i,j+i} = I_{i,j}(1 - \zeta_j),$$

and the solution of these equations for $\{I_{i,j}\}$, after eliminating c , is the equilibrium distribution.

Supplementary tables

Journal Pre-proof

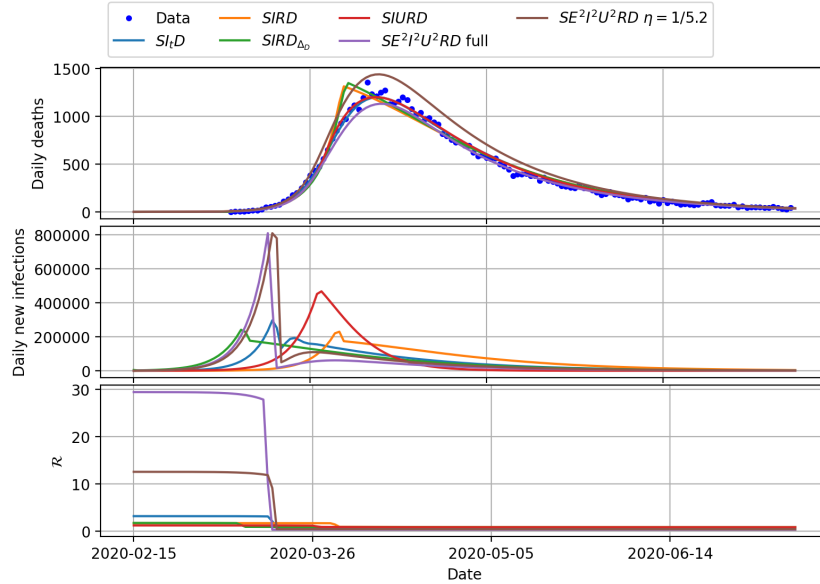
Table S1: MAP estimates of the model parameters and derived epidemiological parameters from fitting the various models to the real data, similar to the simulation study described in section §3.2. The (–) indicate parameters not relevant to the particular model. The derived parameters are: the basic reproduction number \mathcal{R}_0 ; the minimum value of \mathcal{R} , $\min_i\{\mathcal{R}_i\}$; the index of the day at which \mathcal{R}_i first decreases below 1, $\operatorname{argmin}_i\{\mathcal{R}_i < 1\}$; the maximum number of new infections on any day, $\max_i\{S_i - S_{i+1}\}$; the mean, β_μ , of the infectiousness profile; and the mean of the infection-to-death distribution, ζ_μ . The (A) and (B) tables differ in that (B) involves extra parameters being fixed when the models are fitted, as described in §3.2.

(A)							
Model parameters	Estimates					$SE^2I^2U^2RD$	
	SI_tD	$SIRD$	$SIRD_{\Delta D}$	$SIURD$	η free	$\eta = 1/5.2$	
I_0	844.5	34.9	4429.3	68.6	2980.6	4165.8	
ρ	3.2119	0.559	0.5222	1.244	2.548	0.8081	
t^*	31.55	45.66	24.51	41.37	30.37	31.77	
α_b	0.2806	0.5658	0.5506	0.7907	0.0135	0.0404	
θ	—	0.3215	0.2954	0.9992	0.0865	0.0642	
η	—	—	—	—	0.0715	—	
ξ	—	—	—	0.0405	0.1555	0.1212	
ΔD	—	—	21.7779	—	—	—	
ϕ	0.002	0.012	0.0096	0.003	0.0009	0.0015	
Derived parameters							
\mathcal{R}_0	3.21	1.74	1.77	1.24	29.44	12.58	
$\min_i\{\mathcal{R}_i\}$	0.79	0.86	0.85	0.86	0.35	0.44	
$\operatorname{argmin}_i\{\mathcal{R}_i < 1\}$	32	46	25	42	31	32	
$\max_i\{S_i - S_{i+1}\}$, ($\times 1000$)	296.7	230.0	241.2	467.5	809.6	809.5	
β_μ	5.2	3.11	3.38	1.001	22.65	16.88	
ζ_μ	18.69	4.11	25.38	26.67	32.97	30.02	
(B)							
Model parameters	Estimates				$SE^2I^2U^2RD$		
	SI_tD	$SIRD$	$SIRD_{\Delta D}$	$SIURD$			
I_0	844.5	68.4	998.3	109.6	532.1		
ρ	3.2119	0.4241	0.4192	0.4633	1.05		
t^*	31.5	46.6	34.5	39.3	34.5		
α_b	0.2806	0.4094	0.4137	0.3551	0.2869		
ϕ	0.002	0.0128	0.0123	0.0098	0.0054		
Fixed parameters							
θ	—	1/5.2	1/5.2	1/5.2	1/2.974653		
η	—	—	—	—	1/2.974653		
ξ	—	—	—	1/12.49	1/11.744308		
ΔD	—	—	12.49	—	—		
Derived parameters							
\mathcal{R}_0	3.21	2.21	2.18	2.41	3.12		
$\min_i\{\mathcal{R}_i\}$	0.79	0.79	0.79	0.75	0.79		
$\operatorname{argmin}_i\{\mathcal{R}_i < 1\}$	32	47	35	40	35		
$\max_i\{S_i - S_{i+1}\}$, ($\times 1000$)	296.7	324.1	333.7	430.6	306.3		

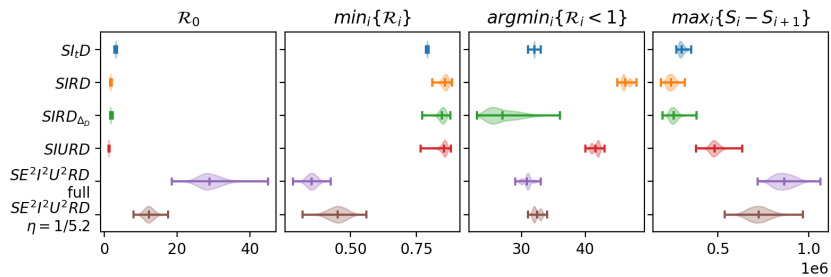
Supplementary Figures

Journal Pre-proof

A



B



C

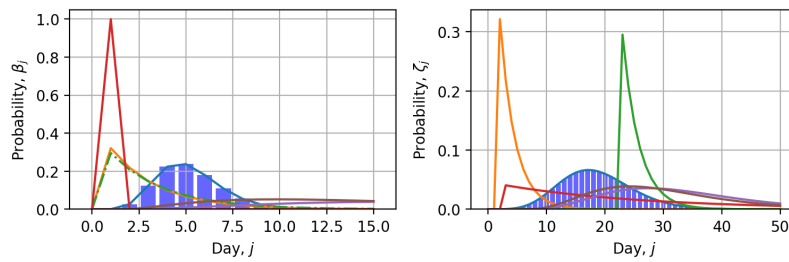
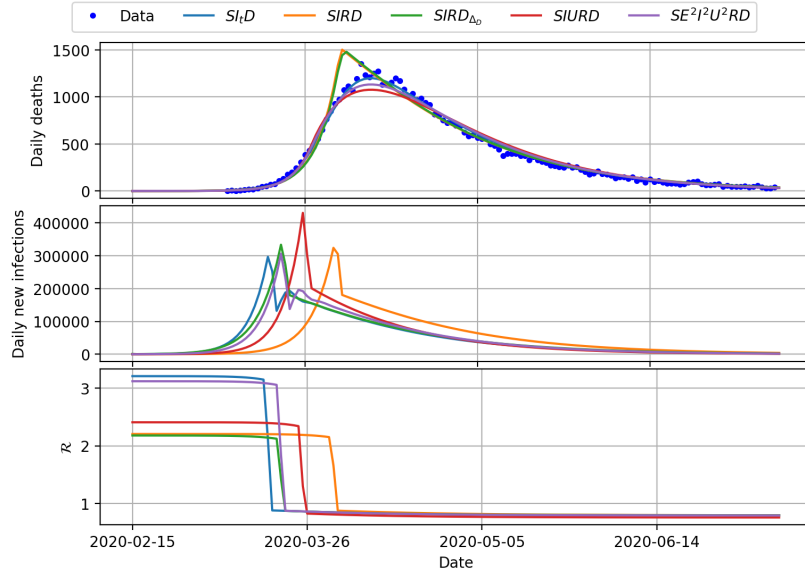
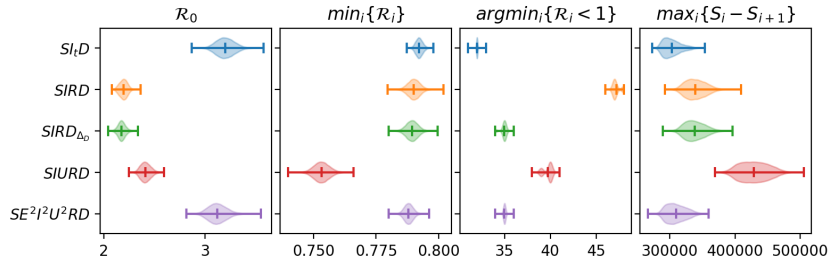


Figure S1: (A) Real data and dynamics of fitted SI_tD , $SIRD$, $SIRD_{\Delta D}$, $SIURD$ and $SE^2I^2U^2RD$ models, corresponding to MAP estimates of parameters per Table S1A. (B) Violin plots of the posterior marginals for derived epidemiological parameters under the different models (C) Infectiousness profiles and infection-to-death distributions for the fitted models.

A



B



C

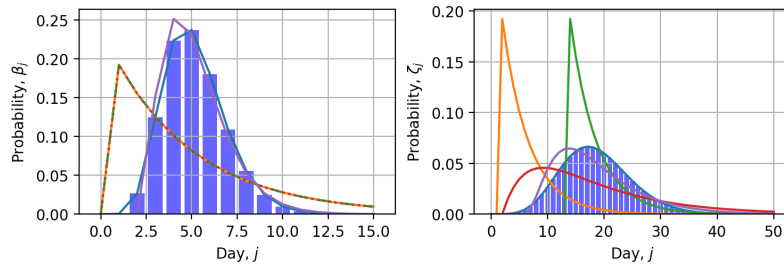


Figure S2: (A) Real data and dynamics of fitted SI_tD , $SIRD$, $SIRD_{\Delta D}$, $SIURD$ and $SE^2I^2U^2RD$ models for the parameters indicated in Table S1B; this version has additional fixed parameters, whose values are chosen *a priori* such that for each model the infectiousness profile and infection-to-death distribution, shown in (C), matches as closely as possible to the data-generating ones. (B) Violin plots of the derived epidemiological parameter uncertainties for each model.

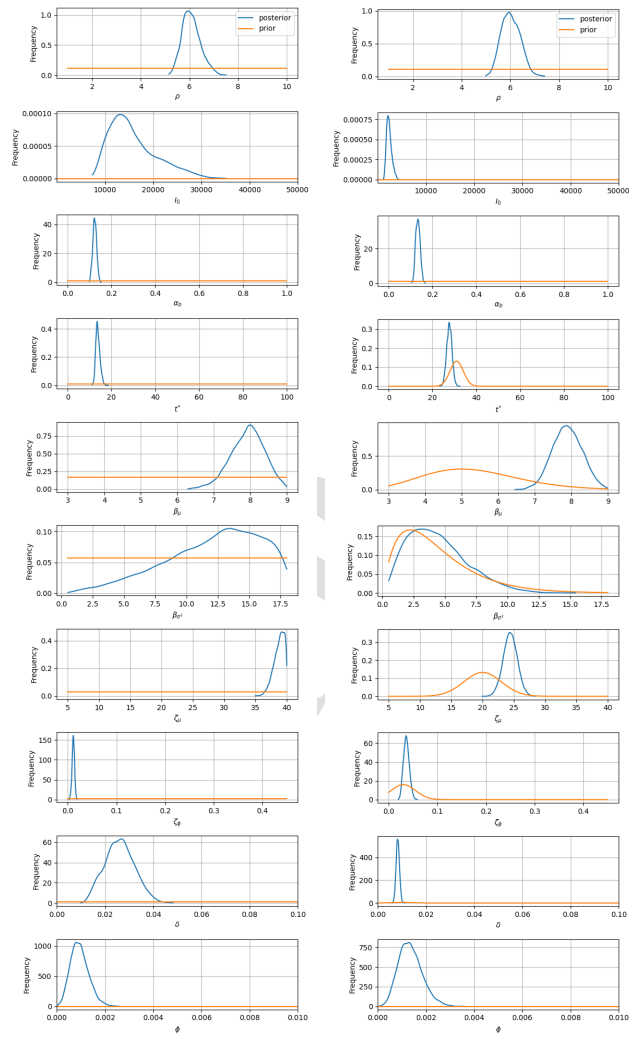


Figure S3: Comparison of posterior distributions from the ‘fully’ Bayesian inference in which the priors are diffuse (left) and informative (right). In each case, the posteriors (blue) are overlaid on the priors (orange). The right panel corresponds to results shown in Figure 7.

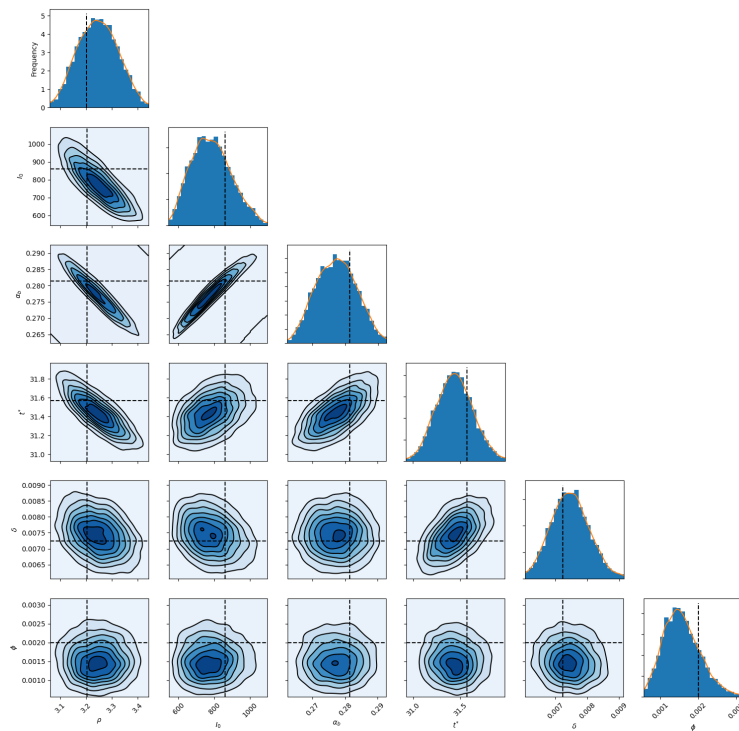


Figure S4: Bayesian inference of synthetic data from Figure 4 using parameters described in §3.3. Uni- and bivariate marginal posterior distributions for the model parameters are shown, with the true, data-generating values shown as black, dotted lines.

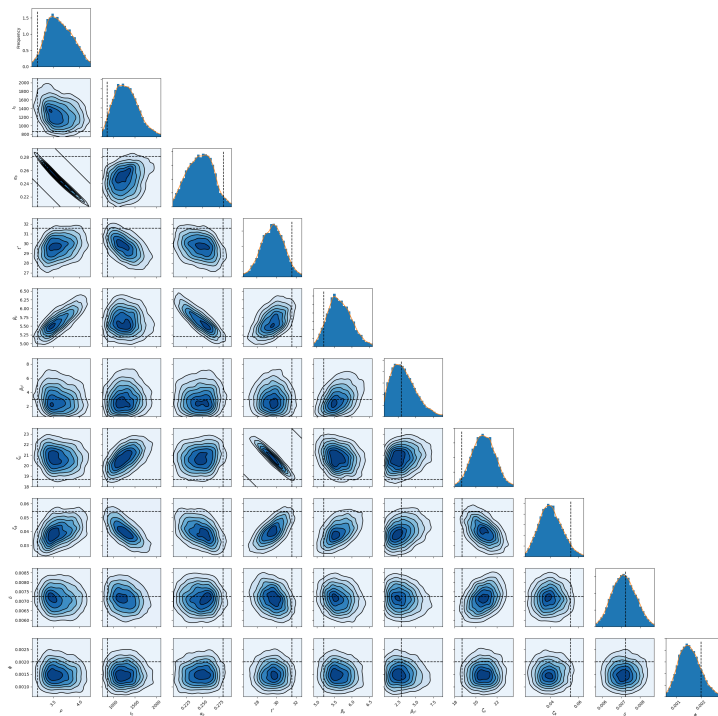


Figure S5: Bayesian inference of synthetic data from Figure 4 using parameters described in §3.4. Uni- and bivariate marginal posterior distributions for the model parameters are shown, with the true, data-generating values shown as black, dotted lines. These results are analogous to those in Fig S4, except here the analysis involved placing priors on—rather than conditioning on—the parameters β_μ , β_{σ^2} , ζ_μ , ζ_ϕ of the infectiousness profile and infection-to-death distribution.

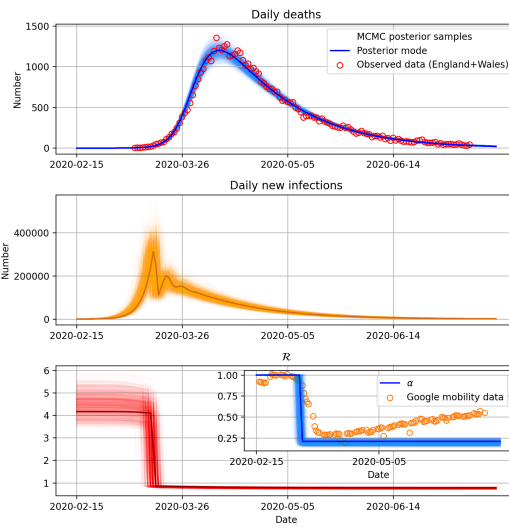
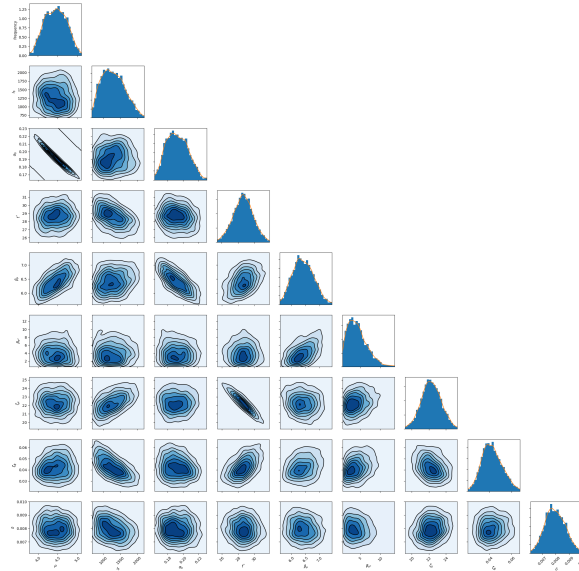


Figure S6: Bayesian inference of real data using the SI_tD model with uncertainty on $\beta_\mu, \beta_{\sigma^2}, \zeta_\mu$ and ζ_ϕ , with ϕ fixed to 0.01. (Upper panel) Uni- and bivariate marginal posterior distributions for the model parameters. (Lower panel) The interpretation is the same as that of panels (B-D) in Figure 6.

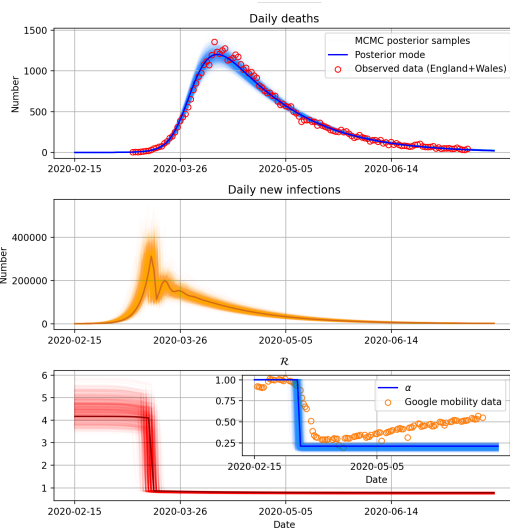
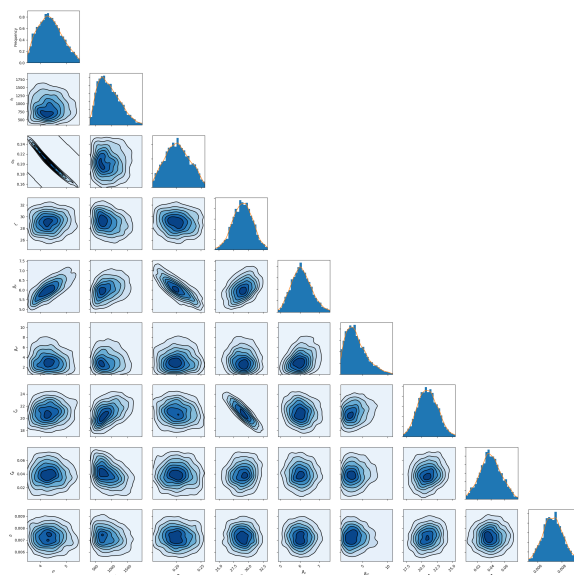


Figure S7: Bayesian inference of real data using the SI_tD model with β_{μ} , β_{σ^2} , ζ_{μ} and ζ_{ϕ} , with ϕ fixed to 0.1. (Upper panel) Uni- and bivariate marginal posterior distributions for the model parameters. (Lower panel) The interpretation is the same as that of panels (B-D) in Figure 6.

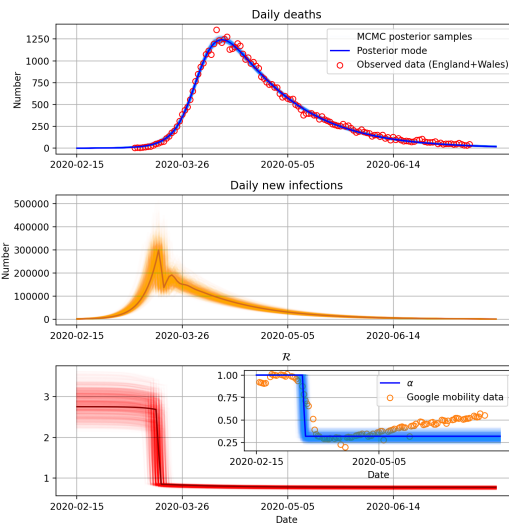
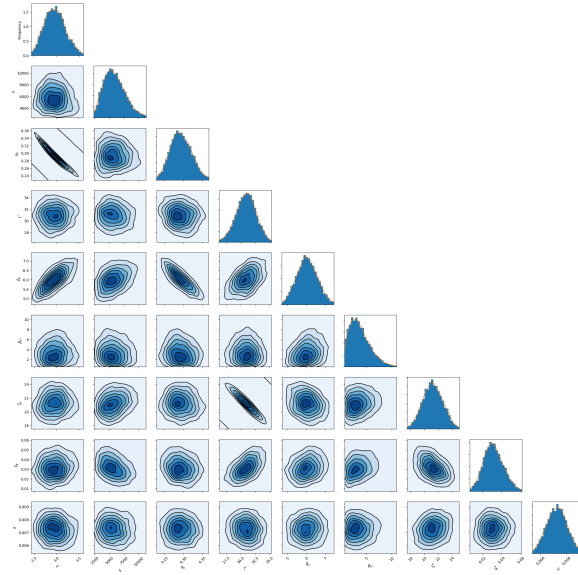


Figure S8: Bayesian inference of real data using the SI_tD model with uncertainty on $\beta_\mu, \beta_{\sigma^2}, \zeta_\mu$ and ζ_ϕ , with Gaussian noise σ fixed to 60. (Upper panel) Uni- and bivariate marginal posterior distributions for the model parameters. (Lower panel) The interpretation is the same as that of panels (B-D) in Figure 6.

- Simple models of SARS-CoV-2 transmission can describe death data in early 2020
- A drop in mixing intensity, and not herd immunity, explains the epidemic decline
- Mis-specifying delays to host outcomes yields misleading estimates for R
- Models structured by infected-age can capture realistic delays and their uncertainty
- Bayesian uncertainty quantification can be misleading if there is model error

Journal Pre-proof

Dominic G. Whittaker: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing - Original Draft, Writing - Review & Editing, Visualisation, Data curation

Alejandra D. Herrera-Reyes:

Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Writing - Original Draft, Writing - Review & Editing, Visualisation, Data curation

Maurice Hendrix: Software, Data curation, Writing - review & editing

Markus R. Owen: Conceptualization, methodology, Writing - original draft, Writing - review & editing

Leah R. Band: Conceptualization, methodology, Writing - original draft, Writing - review & editing

Gary R. Mirams: Conceptualization, Methodology, Writing - Original Draft, Writing - Review & Editing, Supervision, Funding Acquisition

Kirsty J. Bolton: Conceptualization, Methodology, Writing - Original Draft, Writing - Review & Editing

Simon P. Preston: Conceptualization, Methodology, software, Formal Analysis, Writing - Original Draft, Writing - Review & Editing, Supervision

Journal Pre-proof

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Dominic Whittaker is now an employee of GlaxoSmithKline Plc.