Estimating time-dependent infectious contact: a multi-strain epidemiological model of SARS-CoV-2 on the island of Ireland

Tsukushi Kamiya[1,2,*], Alberto Alvarez-Iglesias[1], John Ferguson[1], Shane Murphy[1], Mircea T. Sofonea[3], and Nicola Fitz-Simon[1,*]

[1]HRB Clinical Research Facility, National University of Ireland Galway, Ireland
[2]Center for Interdisciplinary Research in Biology (CIRB), Collège de France, CNRS, INSERM, Université PSL, Paris, France
[3]MIVEGEC, Univ. Montpellier, CNRS, IRD, Montpellier, France
[*]Corresponding author: tsukushi.kamiya@college-de-france.fr
[*]Corresponding author: nicola.fitz-simon@nuigalway.ie

May 3, 2022

Abstract

Mathematical modelling plays a key role in understanding and predicting the epidemiological dynamics of infectious diseases. We construct a flexible discrete-time model that incorporates multiple viral strains with different transmissibilities to estimate the changing infectious contact that generates new infections. Using a Bayesian approach, we fit the model to longitudinal data on hospitalisation with COVID-19 from the Republic of Ireland and Northern Ireland during the first year of the pandemic. We describe the estimated change in infectious contact in the context of government-mandated non-pharmaceutical interventions in the two jurisdictions on the island of Ireland. We take advantage of the fitted model to conduct counterfactual analyses exploring the impact of lockdown timing and introducing a novel, more transmissible variant. We found substantial differences in infectious contact between the two jurisdictions during periods of varied restriction easing and December holidays. Our counterfactual analyses reveal that implementing lockdowns earlier would have decreased subsequent hospitalisation substantially in most, but not all cases, and that an introduction of a more transmissible variant - without necessarily being more severe - can cause a large impact on the health care burden.

1 Introduction

During an epidemic, behavioural changes are encouraged, and sometimes mandated, to curtail infectious disease transmission. These changes aim to reduce the number of contacts between people indiscriminately (e.g., closure of schools, workplaces, commercial establishments, roads, and public transit; restriction of movement; cancellation of public events; maintenance of physical distances in public) and reduce the chance of infection upon contact (e.g., use of personal protective equipment). Furthermore, tracing and isolating known infectious cases can limit the contact between infectious and susceptible individuals. These actions are collectively referred to as non-pharmaceutical interventions (NPIs) and are often mandated by governments. Slowing the surge of infection (or "flattening the curve") affords an opportunity to reduce infection-induced mortality and mobility, alleviate health care burden and wait out an epidemic until pharmaceutical solutions (i.e., treatment and vaccines) become available. Implementation of mandated NPI in historical outbreaks, including during the 1918 influenza pandemic, was crucial for preventing excess death in the United States [1]. NPIs have also been mandated globally during the COVID-19 pandemic.

Mathematical modelling and quantitative analyses of empirical data plays a pivotal role in understanding and predicting epidemiological dynamics. Mechanistic epidemiological models have been widely applied to study the dynamics of SARS-CoV-2, and to make predictions of clinical outcomes under alternative scenarios (e.g., an assumed decrease in physical contact [2]). Despite their public health benefits, social distancing measures have been shown to incur high costs in several domains, including in economy [3], mental health [4], and civil liberty [5]. Thus, it is crucial to quantify infection contact, or its derivative quantities like the effective reproductive number, R - to monitor changes in infection burden, achieve desired public health outcomes and improve policy transparency and public engagement. While it is not possible to measure infectious contact directly, fitting a mathematical model to longitudinal data on observed processes such as reported cases and hospital admissions allows estimation of inter-individual infectious contact and its derivatives [6, 7].

A large contingency of epidemiological models follows a rich tradition of ordinary differential equation (ODE) models [8], which track the spread of infection and often immunity in a population. Specifically, ODE models assume that waiting time processes (such as infectious period and time to hospitalisation) are memoryless, that is to say, that the waiting time until an event (such as recovery and hospitalisation) does not depend on the elapsed time. Seen at the population level, this assumption deduces that times spent by individuals in each compartment are distributed exponentially, implying large individual variability. While mathematically convenient, the lack of memory is unsupported for certain epidemiological processes [9], and empirical evidence indicates other probability distributions with smaller individual variability and non-monotonic densities (e.g., gamma-, Weibull and log-normal distributions) are better equipped to describe those processes. Previous studies have also demonstrated that quantitative predictions of epidemiological outcomes depend on the assumed probability distribution in a variety of systems [10–13], including SARS-CoV-2 [7]. As such, it is pertinent to incorporate realistic waiting time distributions, particularly when one aims to obtain quantitative and short-term, rather than qualitative and long-term insights from

epidemiological models.

Here we develop a compartmental epidemiological model that accurately predicts inter-individual infectious contact over the first year of the epidemic in the Republic of Ireland (ROI) and Northern Ireland (NI). These neighbouring jurisdictions on the island of Ireland present a compelling contrast due to independent policymaking over a small geographical area. We use a discrete-time approach, which allows us to incorporate more accurate assumptions about times spent in each compartment [7].

The COVID-19 pandemic has been characterised by subsequent waves of novel variants with varying disease transmissibility and severity. To separate the effect of human behaviour from the difference in transmissibility of multiple strains, we explicitly model multiple SARS-CoV-2 strains, with differing transmissibilities, seeded in the population at differing times. Our multiple strain model allows consistent estimates of relative infectious contact between periods even when the dominant variant has changed.

Previous studies of SARS-CoV-2 have estimated the time-dependent contact ratio and its derived quantities using continuous (e.g., basis splines [14]), or piece-wise, discrete functions of time (e.g., consisting of the specified period corresponding to NPI mandates [7]). However, both approaches can be fraught with challenges. On the one hand, it is not obvious to choose the appropriate extent of smoothing of a continuous function, for example, by deciding the number of knots in a basis spline function. On the other hand, abrupt changes imposed by piece-wise functions are at odds with empirical data on human movement during the COVID-19 pandemic [15]. Furthermore, it is difficult to establish a precise definition of the level of an intervention over time as definitions changed over time [16, 17]. To address these concerns, we develop an intermediate approach, in which we introduce a prior that allows smoothness in infectious contact between neighbouring weeks in the absence of information otherwise from empirical data.

In the Irish context, compartmental models have been used by several other research groups to understand the dynamics of the virus, make forecasts of outcomes under various scenarios, and assess economic impacts of policy restrictions [18–23]. Our study complements these studies by providing a high-resolution description of the change in infectious contact over time, comparing the two jurisdictions on the island of Ireland. Leveraging the epidemiological model and estimated parameters, we also perform counterfactual analyses to explore the effects of alternative interventions on cumulative hospitalisation and assess the impact of a novel variant.

2 Methods

2.1 Epidemiological model and data

Our multi-strain discrete-time model consists of three types of host compartments (Fig. 1): a susceptible compartment (S) and two infectious compartments for viral strain s, ($J_{s,i}$ and $Y_{s,i}$ where i indicates the infection age, i.e., day since exposure). The compartments J and Y differ in their future clinical outcome: individuals in the components Y eventually get hospitalised while those in J remain out of hospitals. As our primary focus is the inference of NPI in the community, we did not consider within-hospital transmission, recurring hospital admissions of the same patients, or

demographic turnover. We ignored the dynamics of recovered hosts who were assumed to have minimal influence on the transmission during the period investigated.

Our model is parameterised by θ, the probability of hospitalisation, Δ_s, the daily probability of infection with strain s, and a discrete random variable H that characterises a set of probabilities governing daily transitions to hospital. Δ_s is informed by a discrete random variable, Z, that characterises a set of probabilities governing daily transitions into infected compartments, and is defined in Section 2.1.1. Z denotes the time in days from exposure of the infector to exposure of the infectee for a randomly chosen infectee-infector pair (i.e., generation interval), and can be viewed as the average relative contribution of each day to the individual reproduction number. The probability that infection occurs at infector age i, $\zeta_i = P(Z = i) = F_Z(i) - F_Z(i-1)$, where the cumulative distribution function $F_Z(i) = P(Z \leq i)$.

The event transmission of infection does not affect the individual's stay in the compartment; however, for transition out of the Y compartments, the event going to hospital at infection age i is conditional on still being in the compartment at infection age $i-1$. Given H, the random variable representing time from infection to hospital admission in hospitalised patients, we denote by η_i the probability of hospitalisation at infection age i given the individual was still not hospitalised at infection age $i-1$, that is $\eta_i = \frac{F_H(i)-F_H(i-1)}{1-F_H(i-1)}$. This is the discrete hazard of hospital admission at infection age i We use published estimates for Z and H, as described in section 2.1.3 below. Our model assumes that the proportion of infected people hospitalised and the processes governing hospitalisation and recovery over time are constant across strains.

2.1.1 Infection dynamics

We extend a discrete epidemiological modelling framework by Sofonea et al. [7] to accommodate multiple viral strains spreading simultaneously. First, we express the effective density of infectious host population on a given day d that contributes to transmission of the strain s as:

$$\bar{I}_s(d) = c(d) \sum_i \zeta_i \left(J_{s,i}(d) + Y_{s,i}(d) \right), \tag{1}$$

in which $J_{s,i}(d) + Y_{s,i}(d)$ is the number of individuals with strain s on day d with infection age i in the community. Multiplying by ζ_i and summing over infection ages, $\sum_i \zeta_i (J_{s,i}(d) + Y_{s,i}(d))$ can be regarded as the total 'potential for infection' in the community on day d. The effective infectious density, $\bar{I}_s(d)$, is this sum scaled by the infectious contact ratio, $c(d)$, on day d. The infectious contact ratio is the ratio of infectious contact rate on day d to infectious contact rate on day 0. The infectious density is thus a measure of the total amount of transmission in a completely susceptible population.

We allow susceptible hosts encounter the viral strain s with a probability $\Lambda_s(d)$, which as in [7] is assumed to follow the Michaelis-Menten function that saturates with the effective infectious density, $\bar{I}_s(d)$ and the contact rate; under assumptions about initial conditions as in [7] we derive the following:

$$\Lambda_s(d) = \frac{c(d)\, \tau_s\, \bar{I}_s(d)}{\frac{S_0}{\tau_s\, R_{0_1}} + c(d)\, \tau_s\, \bar{I}_s(d)}, \tag{2}$$

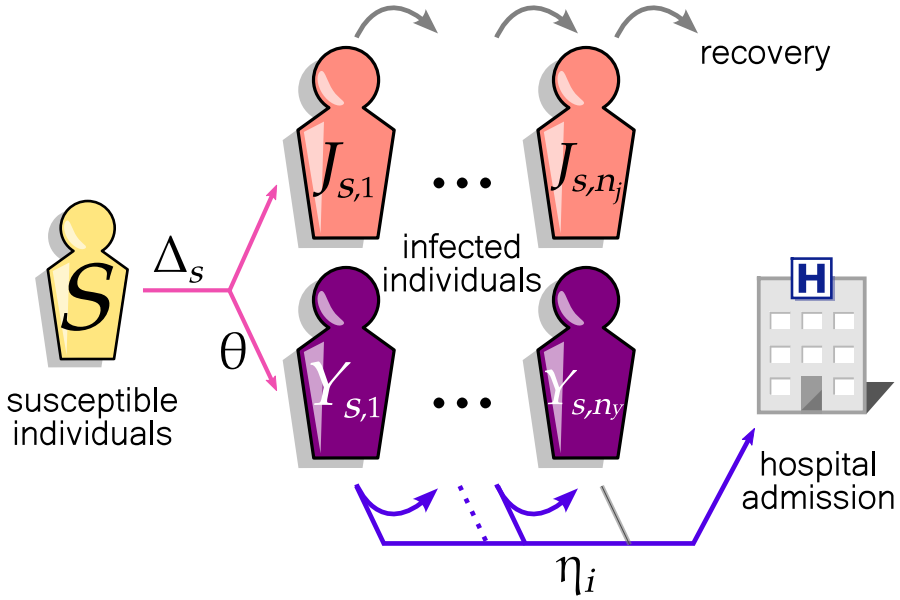

Figure 1: Discrete-time model of SARS-CoV-2 community transmission consists of susceptible S and infectious (J and Y) compartments of pathogen strain s. Each square represents a group of individuals with an identical contribution to the epidemiological dynamics. Infection with strain s occurs with probability Δ_s per day. Individuals in the components Y are infectious patients to be hospitalised. Once infected, individuals progress to the next square each day (J and Y), capturing the memory effect of the infection age. After spending n_j days, infectious hosts in J are no longer infectious. Alternatively, a fraction, θ of infectious hosts (in Y), is admitted to the hospital with a delay specified by the probabilities $\eta_1, ..., \eta_{n_y}$, where η_i is the probability that the individual is admitted to hospital on the day i, conditional on their being infectious for $i-1$ days. The grey arrows indicate the daily transition of individuals from one square to another that occurs with probability 1.

where τ_s is the relative transmission advantage of strain s, S_0 the population size, and R_{0_1} the basic reproductive number of the original strain. We define the τ_s as the ratio of the basic reproductive numbers of strain s to the original strain $R_{0_s} = \tau_s R_{0_1}$.

When a host encounters multiple strains, we model the interaction between strains assuming superinfection with priority determined by order of exposure: i.e., only the first strain that encounters a host establishes infection when the same host subsequently encounters multiple strains. Thus, in the case of two strains, the probability of getting infected with the strain $s, s \in \{1, 2\}$, $\Delta_s(d)$, is:

$$\begin{aligned} \Delta_s(d) = &\quad \Pr(\text{exposure to strain } s) - \Pr(\text{exposure to both strains and } s' \text{ first}) \\ = &\quad \Lambda_s(d) - \tfrac{\Lambda_s(d)\Lambda^2_{s' \neq s}(d)}{\Lambda_1(d)+\Lambda_2(d)} \end{aligned} \tag{3}$$

Similarly, expressions can be derived for more than two strains.

It follows then that the number of susceptibles on the next day is expressed as:

$$S(d+1) = (1 - \sum_s \Delta_s(d))\, S(d). \tag{4}$$

Of those exposed to either viral strain, the proportion θ will develop severe symptoms and eventually be admitted to the hospital (Fig. 1).

For less severe cases that do not result in hospitalisation, J, the infection progresses towards recovery until they are no longer infectious on the day n_j:

$$J_{s,1}(d+1) = (1-\theta)\, \Delta_s(d)\, S(d) \tag{5}$$

$$J_{s,i}(d+1) = J_{s,i-1}(d) \qquad 1 < i \leq n_j. \tag{6}$$

Those that develop severe symptoms, Y, are admitted to hospital with the probability η_i on the i-th day following exposure.

$$Y_{s,1}(d+1) = \theta\, \Delta_s(d)\, S(d), \tag{7}$$

$$Y_{s,i}(d+1) = (1-\eta_{i-1})\, Y_{s,i-1}(d), \qquad 1 < i \leq n_y. \tag{8}$$

It follows then that the number of hospital admissions on day $d+1$ equals

$$H(d+1) = \sum_s \sum_i \eta_i\, Y_{s,i}(d) \tag{9}$$

2.1.2 Observed longitudinal data

Epidemiological models are often fitted to data on infected cases - however, case data depends on levels of testing, which varied over time during the COVID-19 pandemic. It can also be problematic to rely on data on deaths - for instance, many deaths in the ROI occurred following outbreaks in care homes, and thus data on deaths may not reflect disease spread in the general community. Biases and uncertainty in estimating the reproductive number arising from such issues are discussed elsewhere [6]. Thus, hospital admission data is likely a better reflection of the community spread of SARS-CoV-2. We used daily COVID-19 hospital admissions in the ROI and NI, reported

respectively by the Central Statistics Office COVID data hub for the ROI [24], and the NI Department of Health [25].

Successive invasions of new variants have so far characterised the COVID-19 pandemic. Our study tracks two strains that circulated in the island of Ireland in the first 12 months of the pandemic: i.e., the original strain (initially detected in Wuhan, China) and the Alpha strain (also known as B.1.1.7., initially detected in Kent, UK). We used publicly accessible data on the frequency of the Alpha strain in the ROI [26], and NI [27], respectively.

2.1.3 Incorporating empirical estimates of waiting time distributions

Linking transitions within and between these components are two random variables, each describing a waiting time process. These are the infectious period (generation interval), Z, and the delay between infection exposure and hospitalisation, H. The probability distributions representing these random variables have been estimated elsewhere empirically for SARS-CoV-2 in a global and European context.

Generation interval

The generation interval refers to the time between infection events in a pair of infector and infectee, reflecting the incubation duration and recovery timing. Here, we used the distribution of this interval to model the relationship between the age of infection (i.e., time since exposure) and the infectiousness of the infector. We employed an estimate by Ferretti et al. [28] who found the variation in SARS-CoV-2 generation interval was best described by the Weibull distribution with the mean interval of 5.5 days (shape= 3.29 and scale= 6.12). We truncated the Weibull distribution at the upper-integer-rounded 99%-quantile — without this truncation, the discrete model would require infinite time-tracking sub-compartments due to a right-unbounded support $[0, \infty]$. We then discretised the distributions because the dynamics unfold in discrete-time intervals of one day in our model.

Exposure to hospital admission

The waiting time between exposure and hospital admission was estimated as the sum of the incubation period and the delay between symptom onset and hospitalisation. We assumed that the two waiting times were independent due to the absence of evidence otherwise. A meta-analysis of global, but predominately, Chinese data found that the SARS-CoV-2 incubation period was log-normally distributed with parameters $\mu = 1.63$ and $\sigma = 0.50$ [29], corresponding to a mean incubation time of 5.78 days (standard deviation of 3.97 days). The distribution of waiting time between symptom onset and hospitalisation was estimated assuming a gamma distribution by Public Health England with a mean of 5.14 days (standard deviation of 4.2 days) [30]. We fitted a gamma distribution to the simulated sum of the two distributions to represent the timing between exposure to infection and hospital admission (shape = 4.76 and rate = 0.435).

2.1.4 Weekly infectious contact ratio

Here, we defined the contact ratio c as the (potentially infectious) contact rate relative to the pre-pandemic, pre-intervention baseline (eq. 1 & 2) and estimated this quantity using a piece-wise function consisting of weekly intervals. Specifically, we estimated the ratio in each area a (NI and ROI), per week w (i.e., $c_{a,w}$) as a function of $\phi_{a,w}$, the log proportional change in the contact rate from the previous week. We index w from the date of the first public health intervention in either jurisdiction, which took place in ROI on 2020-03-12 (Supporting Information S1: Table S1 & S2); hence the preceding, pre-intervention contact ratios are defined as 1.0.

$$c_{a,w} = 1.0, \quad w = 0. \tag{10}$$

$$c_{a,w} = c_{a,(w-1)}\, e^{\phi_{a,w}}, \quad w \geq 1. \tag{11}$$

With this formulation, hierarchical Bayesian inference with priors on the $\phi_{a,w}$ allows us to estimate the time-varying weekly contact ratios with minimal prior information specific to the modelled system. Specifically, we used a prior $\sim \mathcal{N}(0, \epsilon)$, where ϵ is a hyperparameter specifying the standard deviation of ϕ, such that $c_{a,w}$ would equal $c_{a,(w-1)}$ in the absence of signals from epidemiological data (Table 1). A priori, this formulation avoids over-fitting random weekly variation at the potential risk of smoothing over valid signals of an abrupt change in the weekly contact ratio, for example, following an introduction of lockdown measures. To check for such bias, we examined the extent to which our smoothing approach affects the estimation of sudden changes in the infectious contact ratio, c. We showed that our formulation is unlikely to introduce substantial bias (Supporting Information S2).

2.1.5 Initial conditions

The first case of SARS-CoV-2 on the island of Ireland was identified in NI on 2020-02-27 from an individual travelling back from Northern Italy via Dublin Airport located in ROI (Table S2). Two days later, the first official case in the ROI was also confirmed from a traveller from Northern Italy (Table S1). Initially, most known cases are travel-related, and contact tracing may successfully contain infections. As our model solely tracks community transmission, we started our simulations on the first day that community transmission was detected on the island of Ireland: 2020-03-05 (Table S1). Coincidentally, the exponential growth of confirmed cases appears to have begun around 2020-03-05 in both ROI and NI [24, 31]. We account for the uncertainty of the beginning of community transmission by estimating the initial infectious density independently in the two jurisdictions (Table 1). We assume implicitly that the contribution of the travel-related cases is negligible once the infection starts growing exponentially in the community.

The first cases of the Alpha strain were reported in November and December 2020, respectively, in ROI and NI (Tables S1 & S2). Due to high connectivity with the island of Britain, the Alpha strain likely entered the island of Ireland soon after it emerged in England, where the strain was detected in mid-September [32]. By February 2021, the Alpha strain comprised the majority of infections in both ROI and NI. To estimate the date of introduction, we fitted a three-parameter logistic function to the longitudinal

data of the Alpha frequency and identified the date on which Alpha cases (frequency of Alpha × known new cases) intersects 1: the date of introduction was estimated as 2020-09-22. Again, we account for the sensitivity of the timing of introduction by estimating the founding infectious density of the Alpha strain, independently in the two areas (Table 1). In our model, viral strains differ only in their transmissibility, τ_s.

2.1.6 Fitting

We used a Bayesian approach to fit the above model to two types of longitudinal data from the ROI and NI: daily counts of hospital admissions and the Alpha strain frequency. Model parameters are detailed in Table 1. Hospital admissions per day were modelled as log-normally distributed with standard deviation parameters σ_h. The frequency of the Alpha strain was fitted assuming the beta proportion distribution with a standard deviation parameter, σ_f.

We fitted our model to the data from the first year of the pandemic from the first confirmed case of community transmission on the island, which was detected on 2020-03-05, in ROI, to the end of February 2021. Our modelling period precedes the widespread administration of the full course of vaccination in either jurisdiction: the proportion of fully (twice) vaccinated individuals in ROI and NI was less than 3% and 2% at the end of February 2021, respectively [24, 31].

Our model was written in Stan 2.21.2 and fitted through the RStan interface [33]. We fitted the model in parallel in four independent chains, each with 5000 sampled iterations and 1000 warmup iterations. For diagnostics, we confirmed over 400 effective samples and ensured convergence of independent chains using the $\hat{R}$ metric (values below 1.1 are considered an indication of multi-chain convergence) for all parameters [34]. We assessed the goodness of fit to data using standardised residuals (Supporting Information S3). We also quantified the posterior z-score and posterior contraction to examine the accuracy and precision of posterior distributions and the relative strength of data to prior information [35] (Supporting Information S4).

2.2 Counterfactual analyses

Estimating infectious contacts with a multi-strain model separates the effect of human behaviour from the difference in transmissibility of multiple strains. This separation allows us to leverage the epidemiological model and estimated parameters to simulate an epidemic based on data-generating processes consistent with the observed data. In turn, we can modify one part of the fitted model —while everything else is constant — to conduct counterfactual analyses, which allows us to explore the impact of different factors that affect disease transmission. Here, we explored two counterfactual scenarios: to examine the effect of lockdown timing; and to isolate the impact of the more transmissible Alpha strain on the hospitalisation outcome.

2.2.1 Effect of lockdown timing

We explored the impact of the timing of lockdown introduction by simulating an epidemic with parameters estimated from the model fitted to the observed data on hospitalisations and strain proportions, but the contact ratios counterfactually shifted earlier

Table 1: Description of model parameters and their fixed values, or prior distributions used in Bayesian statistical inference. We assigned an informed prior for R_0, τ_2 and a generic, weakly informative prior for $\bar{I}_{s,a}(0)$, ϵ and measurement error parameters.

Symbol	Description	Fixed value or prior	Source
Epidemiological parameters			
ζ	Infectivity	detailed in text	[36]
η	Exposure to hospital admission	detailed in text	[29, 30]
$S_{0,\mathrm{ROI}}$	Population size of the Republic of Ireland	4.92×10^6	[37]
$S_{0,\mathrm{NI}}$	Population size of Northern Ireland	1.89×10^6	[38]
$log(\bar{I}_{s,a}(0))$	Log initial effective infectious density (of strain s in area a)	$\mathcal{N}(0, 10)$	
R_{0_1}	Basic reproductive number of the original strain	$\mathcal{N}(2.79, 0.86)$	[39]
τ_2	Transmission advantage of the Alpha strain	$\mathcal{N}(1.31, 0.24)$	[32]
$\phi_{a,w}$	Log proportional change in the contract ratio from the previous week	$\mathcal{N}(0, \epsilon)$	
ϵ	Hyperprior specifying the standard deviation of ϕ	half-$\mathcal{N}(0, 1)$	
θ	Probability of hospital admission given infection	0.026	[40]
Measurement error			
σ_{h}	Standard deviation for hospital admission (log-normal distribution)	half-$\mathcal{N}(0, 1)$	
σ_{f}	Standard deviation for frequency of the Alpha strain (beta proportion distribution)	half-$\mathcal{N}(0, 1)$	

by seven and 14 days relative to the actual start of the three lockdowns imposed in the ROI and NI. We then compared counterfactual scenarios and reality by computing the ratio of the cumulative hospital admission numbers for the subsequent days under the counterfactual versus the observed scenarios. Our effect measure is the (natural) log of this quantity, denoted lnRR.

Suppose the intervention is to shift the first lockdown date earlier by seven days. Denoting the actual lockdown date d_l and the counterfactual contact ratio on day d as $c^*(d)$:

$$c^*(d) = \begin{cases} c(d) & \text{if } d < (d_l - 7) \\ c(d+7) & \text{if } d \geq (d_l - 7). \end{cases}$$

The counterfactual infectious density on day d follows from equation 1, and we denote this

$$\bar{I}_s^*(d) = c^*(d) \sum_i \zeta_i \left(J_{s,i}^*(d) + Y_{s,i}^*(d) \right.$$

where the J^*, Y^* denote the counterfactual numbers in these compartments on day d. From this follows the counterfactuals on day d $\Lambda_s^*(d)$, $\Delta_s^*(d)$, and $H^*(d+1)$ from equations 2 to 9. For the second of third lockdowns, we assume that the epidemic had proceeded as observed up to the second and third lockdown, respectively.

2.2.2 Impact of a more transmissible variant

We investigated the extent to which the introduction of the more transmissible Alpha strain contributed to public health burden by simulating an alternative epidemic with parameters estimated from the model fitted to the observed data on hospitalisations and strain proportions, but without introducing the Alpha strain in September 2020. We then used lnRR to compare the cumulative hospital admission numbers between the counterfactual and real scenarios over time until the end of the modelled period at the end of February 2021.

3 Results

3.1 Epidemiological model fit

Our discrete-time epidemiological model of SARS-CoV2 accurately described the time-course of hospital admissions and the frequency of the Alpha strain during the first year of the pandemic in the two jurisdictions on the island of Ireland, before the full course of vaccines were widely administered (Fig. 2; Supporting Information S3 & S4).

3.2 Estimated infectious contact ratios

We estimated a rapid decline in infectious contact ratios during the first month of the pandemic before a strict lockdown was implemented (Fig. 3; Tables S1 & S2). The first lockdown started on 2020-03-28 in both jurisdictions. By this date, the estimated infectious contact ratio was already down to about 60% of the pre-pandemic baseline

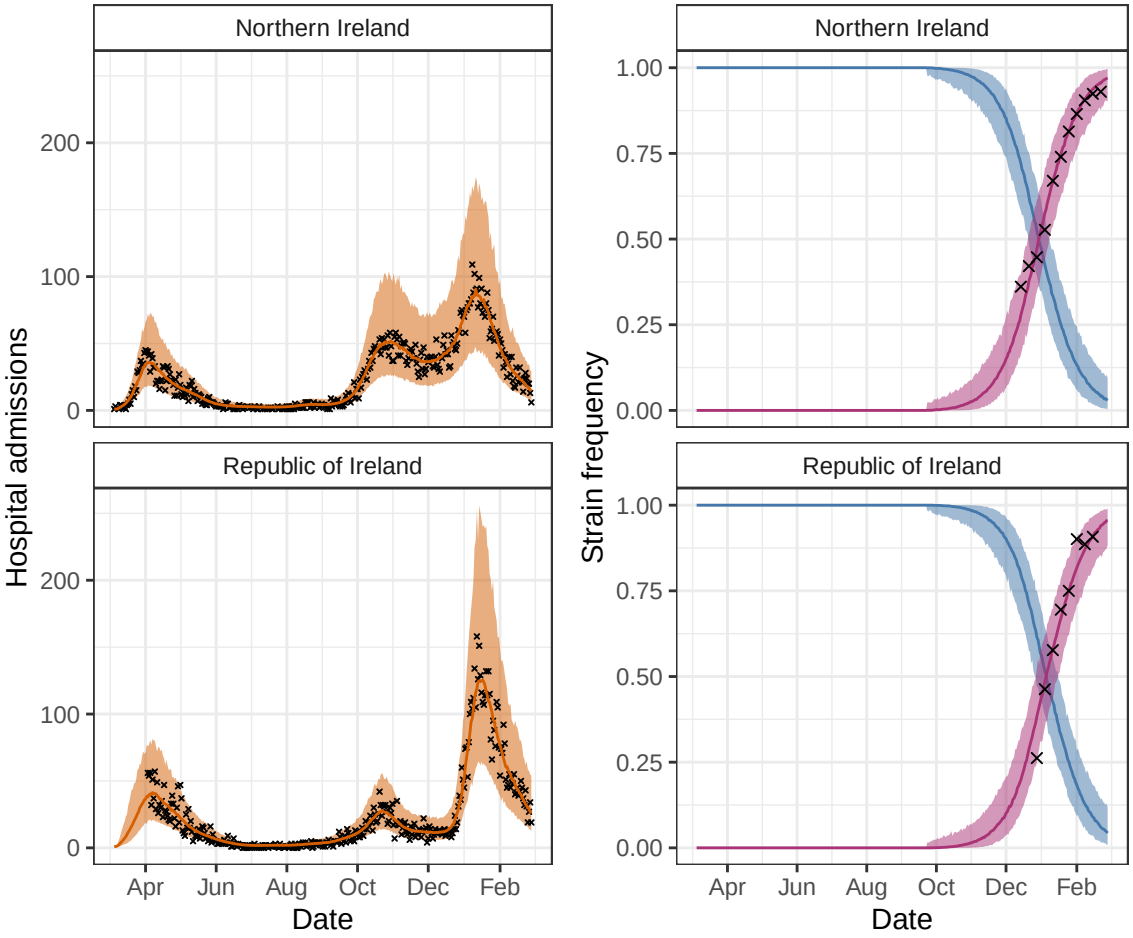

Figure 2: The fit of the epidemiological model to the longitudinal hospital admissions data and the frequency of the Alpha strain (the original strain in blue and Alpha strain in purple). The crosses indicate data and coloured bands correspond to 95% predictive intervals of the model, incorporating uncertainty in parameter estimation and sampling.

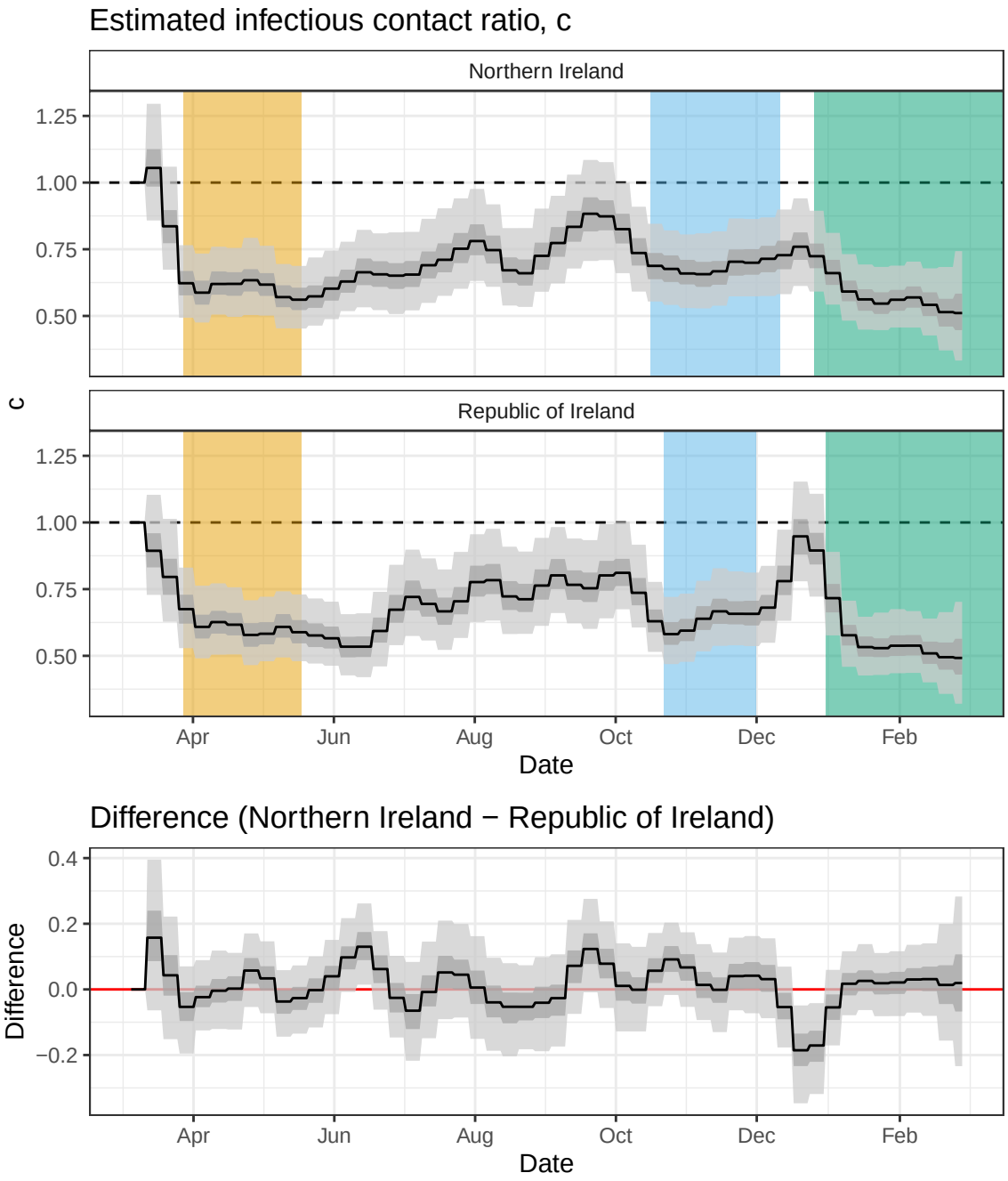

Figure 3: Estimated weekly infectious contact ratios in Northern Ireland (top) and the Republic of Ireland (middle) and differences in the contact ratio between the two jurisdictions (bottom). The three lockdown periods, corresponding to the most strict restrictions in each jurisdiction, are marked in yellow, blue and green, respectively. The black line, and grey bands correspond to the median, the 50% (dark) and 95% (light) predictive intervals.

in both jurisdictions. During this lockdown, estimated infectious contact fluctuated only slightly.

In both jurisdictions, the easing of the first lockdown began from 2020-05-18, and a long period of slow restriction easing took place during the summer months. In the ROI, the estimated infectious contact ratio increased from June and fluctuated between approximately 70-80% of the pre-pandemic baseline in July, August and September. In NI, the contact ratio rose to a peak around the end of July. We detected higher infectious contact in NI than the ROI in mid-June (indicated by 95% predictive intervals of the difference excluding zero; Fig. 3; bottom panel). Of potential relevance, we note that all non-essential retail outlets were allowed to reopen earlier in NI than the ROI during this period from 2020-06-12 and 2020-06-29, respectively (Tables S1 and S2). Infectious contact in NI decreased through August but elevated again to about 90% of baseline by the end of September with no parallel increase in the ROI. This period corresponds to the first time primary and secondary teaching resumed in person in both jurisdictions. The increased contact in NI mirrors a trend in detected England where the September schooling reopening led to increased cases, most notably among the teaching staff [41].

Ahead of the second lockdown, the estimated infectious contact ratio declined to about 60% of baseline in both jurisdictions during October. Unlike during the other two lockdowns, the contact ratios tended to increase during the lockdown period in both jurisdictions throughout November (Fig. 3; top and middle panels).

At the beginning of December in the ROI, several mitigation measures were lifted, allowing non-essential businesses, restaurants, cafes and gastro-pubs to open as well as relaxing household gathering restrictions (Table S1). This period coincides with an increasing trend in the estimated infectious contact ratio, which reached about 90% of the pre-pandemic baseline the week before Christmas. In NI, on the other hand, the lockdown remained in place almost two weeks longer, and the estimated infectious contact ratio reached a maximum of about 75% of the baseline value before Christmas. Our estimates indicate that infectious contact was substantially higher in the ROI than NI for two weeks over the Christmas period (indicated by 95% predictive intervals of the difference excluding zero; Fig. 3; bottom panel): the ROI experienced the highest per capita infection rate in the world during this period. [42]. In both NI and the ROI, the third lockdown introduced in late-December 2020 coincided with the lowest contact ratio (Fig. 3; green) followed by the first lockdown in late March (Fig. 3; yellow) and second lockdown in November (Fig. 3; green).

3.3 Counterfactual scenarios

3.3.1 Effect of lockdown timing

Lockdown measures have been shown effective in reducing the infection burden of SARS-CoV-2, and the timing of introduction is the most significant factor in determining their effectiveness [43, 44]. We found that bringing forward the lockdown dates by either seven or 14 days would have substantially reduced the cumulative hospitalisation over the subsequent 50 days from the date of lockdown in most scenarios (indicated by the 95% predictive interval of the lnRR excluding zero; Fig. 4). Of note,

we found that a counterfactual simulation to bring forward the second lockdown date by seven days showed a non-conclusive impact on the cumulative hospitalisation in the subsequent 50-day period in either jurisdiction (judged by the 95% predictive interval of the lnRR containing zero; Fig. 4). The second lockdown was preceded by a declining trend in contact ratios while the contact during the lockdown remained relatively higher than the first or third lockdown (Fig. 3).

3.3.2 Impact of a more transmissible variant

Our model estimated that the Alpha strain was approximately 19% more transmissible than the original strain (95% prediction intervals $[16.0, 21.8]$). It is worthwhile noting that our model does not consider continuous inputs of infection into the island of Ireland, despite the high connectivity among the British Isles. Thus, our estimate of the Alpha transmissibility may be confounded by repeated introductions, for example, from England, where the Alpha strain was first detected. Nonetheless, our estimate is consistent with those from England [32].

To assess the the impact of the Alpha strain, which arrived later and is more transmissible than the original strain, we compared the fitted model (Fig. 3; orange) to a counterfactual simulation without the Alpha strain, in which we assumed the same estimated contact ratio (Fig. 3; blue). We detected a statistically distinguishable impact of the Alpha strain on the cumulative hospital admissions by earlier January in both jurisdictions - approximately 3.5 months after the initial introduction (indicated by the 95% predictive interval of lnRR excluding zero; Fig. 5). By the end of February 2021, we show that the Alpha strain was responsible for a 38 and 55% increase in cumulative hospitalisation, in NI (median lnRR = 0.323) and the ROI (median lnRR = 0.437), respectively (Fig. 5). Our findings demonstrate that an introduction of a more transmissible variant - without necessarily being more severe - can cause a large impact on the health care burden.

4 Conclusion

We developed a multi-strain model of SARS-CoV-2 and estimated time-dependent infectious contact over the first 12 months of the pandemic on the island of Ireland. Unlike many earlier COVID-19 modelling studies that estimate the effective reproductive number of a single strain, our model explicitly incorporates multiple viral strains and focus on estimating infectious contact ratios. An important difference between the infectious contact ratio and the effective reproductive number is that the former is unaffected by changes in virus transmissibility, which is modelled independently. As such, our approach separates the effect of human behaviour from that of the difference in transmissibilities between multiple, co-circulating strains.

Examining the longitudinal patterns and geographical differences in the estimated infectious contact ratios allowed us to identify corresponding policies and events. In addition, we leveraged estimated parameters to conduct counterfactual analyses, in which we examined the role of lockdown timing and a novel variant on cumulative hospitalisation. In a companion paper, we extended the application of the estimated infectious contact ratios to causal inference [45]. Specifically, we used mobility and

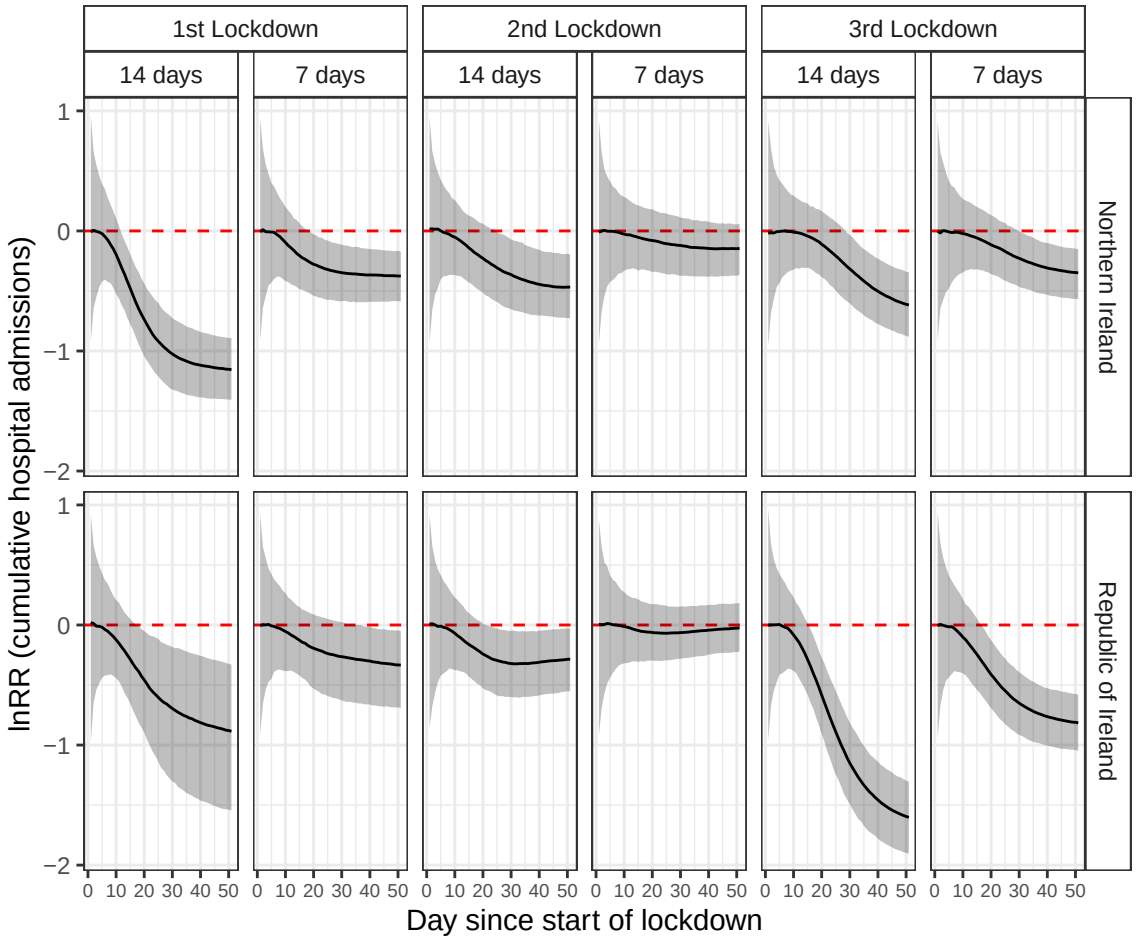

Figure 4: Counterfactual analysis demonstrates the effect of lockdown timing on epidemiological outcomes. The relative difference in cumulative hospital admissions between the counter-factual and factual scenarios is estimated as the log response ratio (lnRR). Shown are the lnRR following the introduction of three lockdowns in Northern Ireland and the Republic of Ireland, assuming that they would have started seven days and 14 days earlier. The black line and grey band indicates the median and 95% predictive interval, respectively.

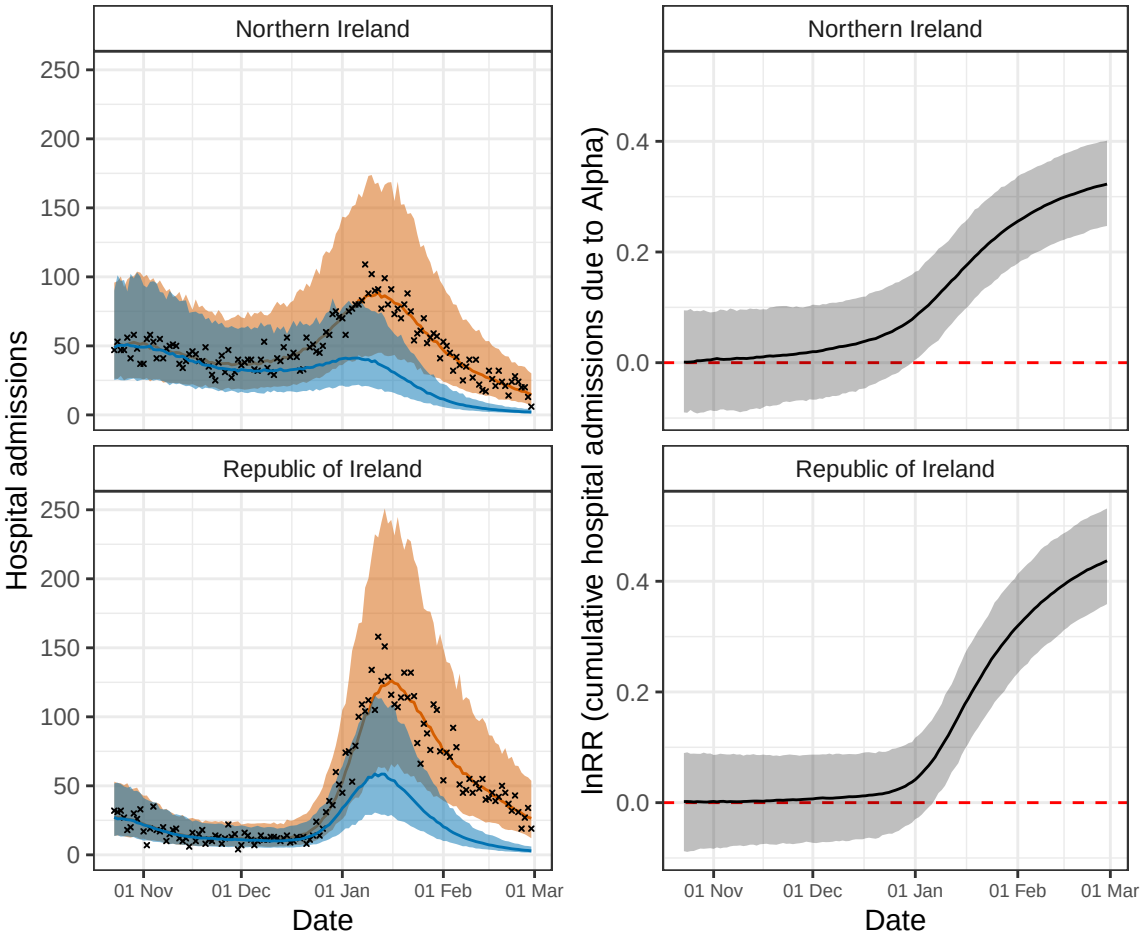

Figure 5: Counterfactual analysis shows the extent to which the Alpha strain elevated the burden of hospitalisation. To compute the impact of the Alpha strain, the counterfactual simulation (blue) assumes that the Alpha strain never invaded either jurisdiction. The crosses indicate data, and the coloured bands correspond to 95% predictive intervals of the fitted model (orange) and counterfactual scenario (blue), respectively, incorporating uncertainty in parameter estimation and sampling (left panels). The relative difference in cumulative hospital admissions between the two scenarios is estimated as the log response ratio (lnRR). The black line and grey band indicates the median and 95% predictive interval, respectively (right panels).

mask-wearing data to independently predict the infectious contact ratios estimated from our epidemiological model described in the current paper and subsequently compared observed hospitalisations with predicted hospitalisations under a counterfactual mask-wearing scenario.

We presented a generic, epidemic model parameterised for SARS-CoV-2 to fit longitudinal hospitalisation data, one of the most reliable and available data types [6]. Of most relevance to COVID-19 in 2022, our model lacks human age structure and vaccination: these omissions give rise to certain limitations. For example, hospitalisation risks increase with age while older individuals adjust their behaviour differently from young counterparts [46]. Thus ignoring the age structure may bias our estimate of infectious contact estimated from hospitalisation data. In addition, the lack of vaccination and associated immunity in our model restricted our scope to the first 12 months of the COVID-19 pandemic. Technically, our model can be extended modularly to relax these assumptions about age structure and vaccination. However, these extensions were outside the scope of this study due to challenges in parameterising these processes reliably. For instance, the output of age-structured models are highly sensitive to assumptions of age-specific contact patterns [47], which likely changed during the epidemic, yet empirical data for time-dependent contact matrices are scarcely available. Behavioural adjustment in response to the pandemic is further complicated by the interaction between age- and sex-specific effects [48]. Furthermore, it is difficult to track and parameterise the state of immunity generated by natural infections from multiple viral strains and multiple vaccine doses using compartmental models.

Finally, our work contributes to the growing COVID-19 modelling literature by providing a transparent Bayesian workflow for fitting a multi-strain epidemic model to longitudinal epidemiological data, which may be readily adapted to modelling SARS-CoV-2 in other jurisdictions and other infectious diseases.

Funding

This research was supported by a Science Foundation Ireland COVID-19 Rapid Response Grant ID: 20/COV/0264

References

[1] Markel H, Lipman HB, Navarro JA, Sloan A, Michalsen JR, Stern AM, et al. Nonpharmaceutical interventions implemented by US cities during the 1918-1919 influenza pandemic. JAMA. 2007;298(6):644–654.

[2] Davies NG, Kucharski AJ, Eggo RM, Gimma A, Edmunds WJ, Jombart T, et al. Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a modelling study. The Lancet Public Health. 2020;5(7):e375–e385.

[3] Ke X, Hsiao C. Economic impact of the most drastic lockdown during COVID-19 pandemic—The experience of Hubei, China. Journal of Applied Econometrics. 2021;10.1002/jae.2871.

[4] Marroquín B, Vine V, Morgan R. Mental health during the COVID-19 pandemic: Effects of stay-at-home policies, social distancing behavior, and social resources. Psychiatry Research. 2020;293:113419.

[5] Orzechowski M, Schochow M, Steger F. Balancing public health and civil liberties in times of pandemic. Journal of Public Health Policy. 2021;p. 1–9.

[6] Gostic KM, McGough L, Baskerville EB, Abbott S, Joshi K, Tedijanto C, et al. Practical considerations for measuring the effective reproductive number, R_t. PLOS Computational Biology. 2020;16(12):e1008409.

[7] Sofonea MT, Reyné B, Elie B, Djidjou-Demasse R, Selinger C, Michalakis Y, et al. Memory is key in capturing COVID-19 epidemiological dynamics. Epidemics. 2021;35:100459.

[8] Keeling MJ, Rohani P. Modeling infectious diseases in humans and animals. Princeton University Press; 2008.

[9] Sallet G. Mathematical Epidemiology. Springer Science & Business Media; 2018.

[10] Lloyd-Smith JO, Schreiber SJ, Kopp PE, Getz WM. Superspreading and the effect of individual variation on disease emergence. Nature. 2005;438(7066):355–359.

[11] Brand SP, Rock KS, Keeling MJ. The interaction between vector life history and short vector life in vector-borne disease transmission and control. PLOS Computational Biology. 2016;12(4):e1004837.

[12] Kamiya T, Greischar MA, Mideo N. Epidemiological consequences of immune sensitisation by pre-exposure to vector saliva. PLOS Neglected Tropical Diseases. 2017;11(10):e0005956.

[13] Kamiya T, Greischar MA, Wadhawan K, Gilbert B, Paaijmans K, Mideo N. Temperature-dependent variation in the extrinsic incubation period elevates the risk of vector-borne disease emergence. Epidemics. 2020;30:100382.

[14] Hong HG, Li Y. Estimation of time-varying reproduction numbers underlying epidemiological processes: A new statistical tool for the COVID-19 pandemic. PLOS One. 2020;15(7):e0236464.

[15] Google LLC. Google COVID-19 Community Mobility Reports; 2020. `https://www.google.com/covid19/mobility/` Accessed: 2020.

[16] RTE. What's in the government's medium-term plan for living with Covid?; 2020. Available from: https://www.rte.ie/news/2020/0915/1165322-living-with-covid-summary/.

[17] Department of the Taoiseach. Roadmap for reopening society and business; 2020. https://www.gov.ie/en/news/58bc8b-taoiseach-announces-roadmap-for-reopening-society-and-business-and-u/?referrer=/roadmap/ Accessed: December 2021.

[18] Irish Epidemiological Modelling Advisory Group. A population-level SEIR model for COVID-19 scenarios; 2020. https://www.gov.ie/en/publication/dc5711-irish-epidemiology-modelling-advisory-group-to-nphet-technical-notes/ Accessed: December 2021.

[19] Náraigh LÓ, Byrne Á. Piecewise-constant optimal control strategies for controlling the outbreak of COVID-19 in the Irish population. Mathematical Biosciences. 2020;330:108496.

[20] Cazelles B, Nguyen-Van-Yen B, Champagne C, Comiskey C. Dynamics of the COVID-19 epidemic in Ireland under mitigation. BMC Infectious Diseases. 2021;21(1):735.

[21] Jaouimaa FZ, Dempsey D, Van Osch S, Kinsella S, Burke K, Wyse J, et al. An age-structured SEIR model for COVID-19 incidence in Dublin, Ireland with framework for evaluating health intervention cost. PLOS One. 2021 12;16(12):1–25.

[22] Northern Ireland Department of Health. Modelling the COVID epidemic; 2020. https://www.health-ni.gov.uk/sites/default/files/publications/health/modelling%20the%20COVID%20epidemic.pdf Accessed: December 2021.

[23] Abbott S, Hellewell J, Thompson R, et al. Estimating the time-varying reproduction number of SARS-CoV-2 using national and subnational case counts. Wellcome Open Res. 2020;5(112). Available from: https://doi.org/10.12688/wellcomeopenres.16006.2.

[24] Department of Health. Ireland's COVID19 Data Hub; 2021. Available from: https://covid19ireland-geohive.hub.arcgis.com/.

[25] Northern Ireland Department of Health. COVID-19 Daily Dashboard Updates; 2021. Available from: https://www.health-ni.gov.uk/articles/covid-19-daily-dashboard-updates.

[26] NPHET. NPHET COVID-19 update – 25 February 2021; 2021. Available from: https://assets.gov.ie/125107/a6254f8f-4957-41bf-ba5f-26a6af3a954c.pdf.

[27] Department of Health, Northern Ireland. COVID-19 Infection Survey; 2021. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/articles/coronaviruscovid19infectionsurveyantibodydatafortheuk/16march2021/pdf.

[28] Ferretti L, Ledda A, Wymant C, Zhao L, Ledda V, Abeler-Dörner L, et al. The timing of COVID-19 transmission;Available from: http://dx.doi.org/10.2139/ssrn.3716879.

[29] McAloon C, Collins Á, Hunt K, Barber A, Byrne AW, Butler F, et al. Incubation period of COVID-19: a rapid systematic review and meta-analysis of observational research. BMJ Open. 2020;10(8):e039652.

[30] Pellis L, Scarabel F, Stage HB, Overton CE, Chappell LH, Lythgoe KA, et al. Challenges in control of COVID-19: short doubling time and long delay to effect of interventions. arXiv preprint arXiv:200400117. 2020;.

[31] GOV UK. Coronavirus (COVID-19) in the UK; 2021. Available from: https://coronavirus.data.gov.uk/.

[32] Public Health England. Investigation of novel SARS-CoV-2 variant, Variant of Concern 202012/01, Technical briefing 5; 2020. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/957504/Variant_of_Concern_VOC_202012_01_Technical_Briefing_5_England.pdf.

[33] Stan Development Team. RStan: the R interface to Stan, Version 2.18.2.; 2019. http://mc-stan.org.

[34] Stan Development Team. Stan Modeling Language User's Guide and Reference Manual, Version 2.18.0; 2018. Available from: http://mc-stan.org/.

[35] Betancourt M. Towards a principled Bayesian workflow; 2020.

[36] Nishiura H, Linton NM, Akhmetzhanov AR. Serial interval of novel coronavirus (COVID-19) infections. International journal of infectious diseases. 2020;93:284–286.

[37] Central Statistics Office. Population and migration estimates;. Available from: https://www.cso.ie/en/releasesandpublications/er/pme/populationandmigrationestimatesapril2020/.

[38] Office for National Statistics. Northern Ireland population mid-year estimate;. Available from: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates.

[39] Liu Y, Gayle AA, Wilder-Smith A, Rocklöv J. The reproductive number of COVID-19 is higher compared to SARS coronavirus. Journal of Travel Medicine. 2020;10.1093/jtm/taaa021.

[40] Irish Epidemiological Modelling Advisory Group. COVID-19 Hospital Utilisation Planning model; 2020. Available from: https://www.esri.ie/publications/covid-19-hospital-utilisation-planning-model-description-and-parameters.

[41] Aiano F, Mensah AA, McOwat K, Obi C, Vusirikala A, Powell AA, et al. COVID-19 outbreaks following full reopening of primary and secondary schools in England: cross-sectional national surveillance, November 2020. The Lancet Regional Health-Europe. 2021;6:100120.

[42] The Guardian. 'Reckless' Christmas easing of rules blamed for Ireland Covid surge; 2021. https://www.theguardian.com/world/2021/jan/11/reckless-christmas-rule-relaxation-blamed-for-irelands-dire-covid-surge.

[43] Born B, Dietrich AM, Müller GJ. The lockdown effect: A counterfactual for Sweden. PLOS One. 2021;16(4):e0249732.

[44] Huang X, Shao X, Xing L, Hu Y, Sin DD, Zhang X. The impact of lockdown timing on COVID-19 transmission across US counties. EClinicalMedicine. 2021;38:101035.

[45] Fitz-Simon N, Ferguson J, Alvarez-Iglesias A, Sofonea MT, Kamiya T. Understanding the role of mask-wearing during COVID-19 on the island of Ireland. medRxiv. 2022;.

[46] Elvén M, Kerstis B, Stier J, Hellström C, von Heideken Wågert P, Dahlen M, et al. Changes in physical activity and sedentary behavior before and during the COVID-19 pandemic: A Swedish population study. International Journal of Environmental Research and Public Health. 2022;19(5):2558.

[47] Bastien Reyné QR, Noûs C, Selinger C, Sofonea MT, Djidjou-Demasse R, Alizon S. Non-Markovian modelling highlights the importance of age structure on Covid-19 epidemiological dynamics. Mathematical Modelling of Natural Phenomena. 2022;.

[48] Barber SJ, Kim H. COVID-19 worries and behavior changes in older and younger men and women. The Journals of Gerontology: Series B. 2021;76(2):e17–e23.