Statistics
in Medicine

Research Article

Received 31 August 2012 Accepted 25 October 2013 Published online 25 November 2013 in Wiley Online Library

(wileyonlinelibrary.com) DOI: 10.1002/sim.6041

Bayesian parameter inference for dynamic infectious disease modelling: rotavirus in Germany

Felix Weidemann,[a*†] Manuel Dehnert,[a,c] Judith Koch,[a] Ole Wichmann[a] and Michael Höhle[a,b]

Understanding infectious disease dynamics using epidemic models based on ordinary differential equations requires the calibration of model parameters from data. A commonly used approach in practice to simplify this task is to fix many parameters on the basis of expert or literature information. However, this not only leaves the corresponding uncertainty unexamined but often also leads to biased inference for the remaining parameters because of dependence structures inherent in any given model. In the present work, we develop a Bayesian inference framework that lessens the reliance on such external parameter quantifications by pursuing a more data-driven calibration approach. This includes a novel focus on residual autocorrelation combined with model averaging techniques in order to reduce these estimates' dependence on the underlying model structure. We applied our methods to the modelling of age-stratified weekly rotavirus incidence data in Germany from 2001 to 2008 using a complex susceptible–infectious–susceptible-type model complemented by the stochastic reporting of new cases. As a result, we found the detection rate in the eastern federal states to be more than four times higher compared with that of the western federal states (19.0% vs 4.3%), and also the infectiousness of symptomatically infected individuals was estimated to be more than 10 times higher than that of asymptomatically infected individuals (95% credibility interval: 8.1–19.6). Not only do these findings give valuable epidemiological insight into the transmission processes, we were also able to examine the considerable impact on the model-predicted transmission dynamics when fixing parameters beforehand. Copyright © 2013 John Wiley & Sons, Ltd.

Keywords: Bayesian inference; disease transmission; ordinary differential equations; model averaging; residual autocorrelation; rotavirus

1. Introduction

In applied infectious diseases epidemiology, there has been a substantial increase in the utilization of mathematical modelling, especially to support decision-making in public health policy [1]. Although the focus most often lies on the correct representation of epidemiological aspects during model building, an accurate assessment of parameter inference for such dynamic models is in many cases neglected.

The aim of the present work was not only to develop a mathematical model based on ordinary differential equations (ODE) for the dynamic transmission and stochastic observation of rotavirus infections in Germany but also to adequately address uncertainty in such an approach. We therefore applied a Bayesian framework for parameter inference and model evaluation with an additional focus on residual autocorrelation. Understanding the transmission dynamics of rotavirus in Germany is of particular importance because of a possible implementation of routine childhood immunization against rotavirus, which is currently being discussed by the Standing Committee on Vaccination in Germany.

[a]*Department for Infectious Disease Epidemiology, Robert Koch Institute, Berlin, Germany*
[b]*Department of Mathematics, Stockholm University, Stockholm, Sweden*
[c]*Faculty of Biotechnology and Bioinformatics, University of Applied Sciences Weihenstephan-Triesdorf, Freising, Germany*
**Correspondence to: Felix Weidemann, Department for Infectious Disease Epidemiology, Robert Koch Institute, Berlin, Germany.*
[†]*E-mail: weidemannf@rki.de*

With more than 110 million cases yearly and 400 000 deaths worldwide, rotavirus is the most common cause of acute gastroenteritis [2]. To analyse the potential impact of a vaccination programme against rotavirus transmission, susceptible–infectious–recovered (SIR)-type modelling has already been applied for the USA [3], England and Wales [4, 5], Western Europe [6], Australia [7] and Kyrgyzstan [8]. A modelling approach based on partial differential equations was used to investigate the vaccination impact on incidence and healthcare cost in the USA and Mexico [9, 10]. Mathematical modelling has also been used to investigate the role of single demographic or epidemiological aspects, such as birth rates [11].

In order to conduct inference for such models, likelihood-based approaches [12, 13] and Bayesian approaches [14–16] including Bayesian melding [17] as well as plug-and-play approaches [18, 19] based on frequentist or Bayesian methods have been applied. However, many of the models applied to rotavirus transmission, and also other infectious diseases, have a high complexity level. In these cases, it has become a habit to base many parameter estimates on values from previous epidemiological studies or on expert opinion alone [3, 5, 8]. This means to fix the majority of parameters and to perform data-driven model calibration for only a few model parameters, not accounting for the full epidemiological information contained in any actual incidence data available. However, in the setting of SIR-type models, such a procedure may yield misleading results due to the high parameter sensitivity of such models combined with possible parameter collinearity.

Within our work, we want to present a more comprehensive view on uncertainty compared with previous ODE-based approaches in epidemic modelling by using a Bayesian approach. We formulated a complex age-structured model capturing the underlying disease transmission dynamics via a system of ODEs, where we considered in detail all relevant aspects regarding the transmission of the virus, for example, infectiousness, seasonality and waning immunity. Moreover, an additional stochastic observational component describes the reporting process based on the previously computed transmission dynamics. In doing so, we also provide a more mechanistic framework for describing routinely collected surveillance data compared with more time series-orientated works [20, 21]. Within a Bayesian framework, we performed model inference combining available incidence data with knowledge from focused epidemiological studies. To additionally address uncertainty concerning model selection, we applied Bayesian model averaging techniques, which allowed us to compute posterior estimates and incidence predictions unconditioned on a particular model, but on a set of models instead. Moreover, by adjusting the likelihood impact on the basis of the cumulative autocorrelation (CA) of the residuals, we were able to obtain a more realistic variance estimation accounting for potential mismatches between our model and the real underlying processes.

We organized the present work as follows. In Section 2, we present the German rotavirus incidence data used to motivate our work. Furthermore, we will explain the dynamic transmission model together with the stochastic observational component. In Section 3, we will present the Bayesian two-step estimation and model averaging procedure, yielding results for the posterior distribution of the model parameters. In Section 4, we apply our methods to the rotavirus incidence data and provide an interpretation of the epidemiological and statistical insights. We discuss the gained knowledge on rotavirus transmission, possible extensions of our model and further analyses in Section 5.

2. Data, dynamic transmission and stochastic observation

This section explains the rotavirus data and our proposed model consisting of dynamic transmission and stochastic observation. It is important to distinguish the function of the two model components: the deterministic transmission model describes the disease's dynamic prevalence within the population at any time point and determines the mechanisms of infection spreading, whereas the stochastic observational model explains how underlying transmission dynamics lead to reported cases.

2.1. Epidemiological data basis

Since 2001, acute rotavirus infection has been notifiable in Germany [22]. Laboratory-confirmed cases are routinely reported to the local health offices and forwarded electronically via the state health authority to the Robert Koch Institute. Each data set includes information on age, sex, federal state of residency, onset of symptoms, hospitalization and fatal outcome.

We examined data on the weekly number of rotavirus infections from 2001 onwards. After the licensure of rotavirus vaccines in 2006, the estimated rotavirus vaccination coverage reached 20% in 2008, averaged over all federal states [23]. Therefore, we expected a significant impact on the rotavirus incidence in 2009 due to the increased coverage such that data from 2009 and the following years were excluded from the analysis.

The data were separated by the two regions of the eastern federal states (EFS) and western federal states (WFS) of Germany (the state of Berlin is counted as part of the WFS). It is a common phenomenon among notifiable gastroenteritis in Germany that the reported incidence in the EFS is significantly higher than in the WFS [24]. These differences are not necessarily a result of a higher disease burden in the EFS but are assumed to arise from different healthcare-seeking behaviours of parents with sick children and different diagnostic test-ordering behaviours of doctors.

From 2001 till 2008, a total of 441 508 rotavirus cases were reported to the Robert Koch Institute. This translates into an average annual incidence of 1097 and 20 cases per 100 000 population in the age groups <5 and $\geqslant 5$ years of age, respectively, as well as 47 and 175 cases per 100 000 in the WFS and EFS, respectively. The data suggest a strong seasonal pattern of rotavirus incidence with its maximum during March with an average peak of more than 3000 weekly cases and its minimum in August with just above 200 cases per week. One can observe a distinct increase in the peak number of reported cases beginning in 2005. We expect the 2004 change in reimbursement for a hospitalized rotavirus case to be a major reason for this increase (where a gastroenteritis caused by a specific pathogen was reimbursed higher than that without a specific pathogen being identified) [22].

Additional demographic data on monthly birth rates, annual age-stratified mortality rates, age-stratified migration data from 1990 to 2008 and age-stratified population counts in the EFS and WFS for the years 2001–2008 were obtained from the GENESIS database at the Federal Statistical Office [25]. Because the rotavirus-associated case fatality in Germany is below 0.01%, there was no need to account for disease-induced mortality.

2.2. *The dynamic transmission model*

We constructed a detailed susceptible–infectious–susceptible-type model based on ODEs to mimic the underlying rotavirus transmission dynamics in Germany, where we were mainly inspired by the models presented in [3,5]. Because our long-term goal is to utilize the model to investigate the potential impact of a routine vaccination against rotavirus, we aimed at covering all relevant epidemiological and demographic aspects affecting the transmission within the underlying population. Thus, the resulting model might look overly complicated compared with the data available, but this complexity is necessary in order to implement vaccination mechanisms into the model at a later stage.

To take the heterogeneity of succeeding rotavirus infections in the individual into account, we differentiated between several states of susceptibility. Therefore, the uninfected population was split into three states of susceptibility S_k $(k = 1, 2, 3)$, corresponding to the number of infections already suffered (none, one and at least two, respectively). The states S_2 and S_3 yield a specific relative risk of infection α_k $(k = 2, 3)$ compared with the first susceptibility state S_1. Moreover, each state S_i corresponds to a specific risk of developing symptoms θ_k $(k = 1, 2, 3)$. A structural overview of the model is given in Figure 1.

The starting point of the model is the state M containing all newly born infants protected by maternal antibodies. With loss of maternal protection, these infants move to the first susceptibility state S_1. Following each susceptibility state S_k is a pair of states corresponding to symptomatic and asymptomatic infections I_k, A_k $(k = 1, 2, 3)$. In case of infection, the probability θ_k of developing symptoms and moving to I_k, before passing through a period of asymptomatic infection (A_k), depends on the corresponding susceptibility level k, whereas infected individuals not developing any symptoms move to A_k directly. Recovery from symptoms happens at a rate μ, whereas individuals recover from asymptomatic infection at a rate ω, moving to the next susceptibility state S_{k+1}. The only exception is the recovery from the third infection, where individuals move back to the third susceptibility state S_3. The state variables $M(t)$, $S_k(t)$, $I_k(t)$ and $A_k(t)$ represent the absolute numbers of individuals belonging to these states at time t.

There exists evidence for immunity waning in the absence of boosters through subsequent rotavirus infections [26,27]. To incorporate this waning protection, we assumed that individuals move from state S_k to S_{k-1} $(k = 2, 3)$ at a constant rate β. For a more realistic representation, we introduced two intermediate states of susceptibility S_{kA} and S_{kB} $(k = 2, 3)$, in order to model a temporal delay.

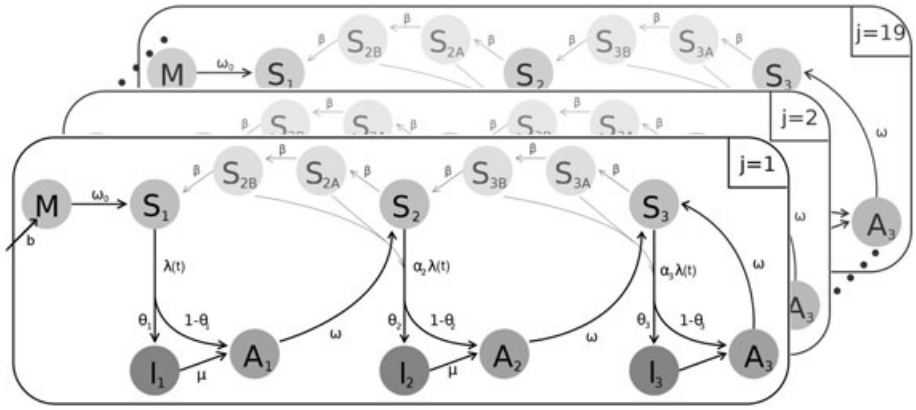

Figure 1. Layer structure of the age-stratified transmission model, in which each age group is represented by one layer.

Thus, the time to loss of immunity after infection corresponds to a $\Gamma(3, \beta)$-distribution instead of an $\exp(\beta)$-distribution.

Rotavirus incidence is the highest among children under 5 years of age with its peak at 1 year of age as the data suggest. To account for this heterogeneity, the data were stratified according to age by distinguishing between $n_D = 10$ age groups: one group for each of the first 5 years of age, one group from 5 to 19 years of age and four subsequent age groups with a range of 20 years each. In order to later be able to include fixed vaccination ages into the model at 2, 4 and 6 months of age and for a more realistic ageing process, the model population was decomposed even further. Thus, the first year of age was split into six equally broad age groups, whereas each group for the second till fifth years of age was divided into two age groups, which yields a total of $n_A = 19$ model age groups and 266 age-specific model states. To denote a specific age group of a model state, the state's name is marked with an upper index (j), $(j = 1, \ldots, n_A)$. Ageing from one group to the next happens at a rate δ_j determined by the length of the age groups.

In order to model the dynamic evolution of the underlying population size as accurately as possible, we fixed time-specific and age-specific death rates $d_t^{(j)}$ as well as migration rates $m_t^{(j)}$ for each age group $(j = 1, \ldots, n_A)$ and also a time-specific birth rate b_t, each from 1990 till 2008.

2.3. *Disease transmission*

In transmission modelling, the rate at which susceptibles become infected is typically determined by the force of infection. In this case, the force of infection is a composition of several components affecting the transmission of rotavirus, that is, infectiousness, susceptibility, contact pattern and seasonality. Regarding infectiousness, we assumed that both symptomatically and asymptomatically infected individuals contribute to the force of infection, but with symptomatically infected individuals having a different relative intensity $p > 0$. Thus, the age-specific force of infection $\lambda_{\text{inf}}^{(j)}$ induced by the j-th age group takes the form

$$\lambda_{inf}^{(j)}(t) = \sum_{k=1}^{3} \left(p I_k^{(j)}(t) + A_k^{(j)}(t) \right).$$

The contact pattern within transmission is modelled through a who-acquires-infection-from-whom matrix $\boldsymbol{C} \in \mathbb{R}_{\geq 0}^{n_A \times n_A}$. The component $c_{i,j}$ of the matrix refers to the average number of weekly contacts of an individual from age group i with the j-th age group and also reflects the relative susceptibility and infectiousness of the respective age groups.

The seasonality component mimics the varying transmission rates over the course of a year because of climatic or other environmental factors, which is assumed to be equal for all age groups. Therefore, we assumed a positive periodic form by using

$$\lambda_{\text{seas}}(t) = \exp\left\{a_1 \cos 2\pi\left(\frac{t}{52} - b_1\right) + a_2 \cos 2\pi\left(\frac{t}{26} - b_2\right)\right\}.$$

In contrast to previous modelling approaches [3,5], we chose such an exponential form to allow the peak in transmission to be more seasonally localized rather than harmonic. Note that, although λ_{seas} is called the seasonal term, $\boldsymbol{\lambda}_{\text{inf}} = (\lambda_{\text{inf}}^{(1)}, \ldots, \lambda_{\text{inf}}^{(n_A)})$ also varies over time, as it depends on the time-varying variables $I_k^{(j)}(t)$ and $A_k^{(j)}(t)$. As this is only an indirect time dependence, it seems more suitable to refer to $\boldsymbol{\lambda}_{\text{inf}}(t)$ as a state-dependent function and to $\lambda_{\text{seas}}(t)$ as a purely time-dependent function.

Combining these parts, we obtain the force of infection

$$\boldsymbol{\lambda}(t) = \frac{1}{N(t)}\lambda_{\text{seas}}(t)\boldsymbol{C}\boldsymbol{\lambda}_{\text{inf}}(t),$$

with $N(t)$ denoting the total population size at time t. Note that $\boldsymbol{\lambda}(t)$ is a time-dependent n_A-dimensional vector providing the force of infection for each of the $n_A = 19$ age groups. For individuals in the states S_2 and S_3, who already suffered from a previous infection, the force of infection $\boldsymbol{\lambda}(t)$ is multiplied by the corresponding relative risk α_k $(k = 2, 3)$ to mimic the higher immunity level.

Altogether, the disease status of the population is represented through 266 possible states, and the corresponding transmission dynamics are governed by a system of ODEs given in Appendix 5.

2.4. *Stochastic observation of new cases*

The preceding ODE system models the absolute number of individuals in each state over time, which yields the actual number of currently infected individuals in each age group. Through an observational component, we now link this deterministic quantity with the actual data, which consist of weekly numbers of newly reported cases and are subject to stochasticity. Similar models for stochastic case reporting were also introduced in [15, 16].

To obtain the number of newly infected cases from the model, we have to consider the derivatives of the states $I_i^{(j)}$ given in the ODE system (7) in Appendix 5. The instantaneous rate of new symptomatic infections $H^{(j)}(t)$ is given by the corresponding derivatives in (7):

$$H^{(j)}(t) = \theta_1\lambda^{(j)}(t)S_1^{(j)}(t) + \alpha_2\theta_2\lambda^{(j)}(t)\left(S_2^{(j)}(t) + S_{2A}^{(j)}(t) + S_{2B}^{(j)}(t)\right)$$
$$+ \alpha_3\theta_3\lambda^{(j)}(t)\left(S_3^{(j)}(t) + S_{3A}^{(j)}(t) + S_{3B}^{(j)}(t)\right).$$

Therefore, the expected number of newly symptomatically infected individuals $Y^{(j)}(t)$ in age group j in week t is

$$Y^{(j)}(t) = \int_t^{t+1} H^{(j)}(u)\,\mathrm{d}u.$$

We differentiated between cases from the EFS and WFS; that is, the expected number of new infections in each age group is divided regionally

$$Y^{(j)}(t) = Y^{(j,w)}(t) + Y^{(j,e)}(t).$$

There is no evidence for differences in the underlying transmission dynamics between the two regions, for example, different contact behaviours or hygienic measures. Thus, we assumed the actual incidences (per 100 000 population) to be equal in both regions for each age group. This implies that the ratio of expected numbers of new cases in age group j in both ESF and WFS is equal to the ratio of populations $\text{Pop}^{(j,i)}(t)$ $(i = e, w)$ of age group j in these regions

$$Y^{(j,e)}(t) = w^{(j,e)}(t)Y^{(j)}(t), \quad Y^{(j,w)}(t) = \left(1 - w^{(j,e)}(t)\right)Y^{(j)}(t), \tag{1}$$

with

$$w^{(j,e)}(t) = \frac{\mathrm{Pop}^{(j,e)}(t)}{\mathrm{Pop}^{(j,e)}(t) + \mathrm{Pop}^{(j,w)}(t)}.$$

However, there is evidence that the differences in reported disease incidence arise from different habits of doctors (related to the ordering of diagnostic tests for patients with diarrhoea) and also different healthcare-seeking behaviours of parents with sick children in the two regions. Thus, we introduced a time and region specific detection rate $h^{(i)}(t)$ $(i = e, w)$, which gives the probability of a symptomatic case to be reported. To avoid overparametrization and because no auxiliary information is available, we assume this detection rate to be age independent, although an age-dependent detection rate might be more realistic. With the change in reimbursement happening at the end of 2004, the detection rate $h^{(i)}(t)$ was assumed to be constant in each region for the time before 2005 and after the first half year of 2005 with linear growth in between. Hence, the time-dependent detection rates $h^{(i)}(t)$ were defined as follows:

$$h^{(i)}(t) = \begin{cases} h^{(i)} & \text{if } t \leqslant t_{2005} \quad \text{(until 2005)} \\ h^{(i)} + \left(q^{(i)}h^{(i)} - h^{(i)}\right)\frac{t-t_{2005}}{26} & \text{if } t_{2005} < t \leqslant t_{2005} + 26 \quad \text{(first half year 2005)} \\ q^{(i)}h^{(i)} & \text{if } t > t_{2005} + 26 \quad \text{(after first half year 2005)} \end{cases}$$

with $h^{(i)}$ and $q^{(i)}$ $(i \in \{e, w\})$ representing the detection rates before the year of 2005 and their relative increase afterwards, respectively, and t_{2005} denoting the first week of year 2005.

Finally, the weekly number of reported cases $X^{(j,i)}(t)$ in age group j and region i was assumed to be negative-binomially distributed

$$X^{(j,i)}(t) \sim \mathrm{NegBin}\left(h^{(i)}(t) \cdot Y^{(j,i)}(t), d\right),$$

with expectation $h^{(i)}(t) \cdot Y^{(j,i)}(t)$ and dispersion coefficient d. A negative binomial distribution was chosen over a Poisson distribution to allow for a higher variance as this was suggested by the data. To keep the number of parameters low, the same dispersion parameter d was applied to all age groups and both regions.

3. Bayesian inference and model averaging

In the past, calibration of dynamic transmission models has been heavily based on results from external epidemiological and clinical studies or on expert opinion. Transmission models based on ODEs are known to behave very sensitively to their input parameters, which is why we think that a data-driven inference, not solely based on detached previous results, is more appropriate for these models. A Bayesian framework with prior elicitation based on this external knowledge seems especially suited for this. Furthermore, we addressed the aspects of residual autocorrelation, which is an inherent issue in infectious disease modelling and may lead to bias in parameter inference, especially underestimation of the parameter variance. To account for this bias, we applied an approach based on readjusting the likelihood function given the residual autocorrelation in the data, which is a particularly suitable method for our complex model, as it does not require the estimation of additional parameters. The accuracy of this procedure was investigated analytically and by simulation as described in the supporting information.

3.1. Prior elicitation

Altogether, a 17-dimensional parameter vector ϑ from Section 2 is to be estimated

$$\vartheta = \Big(\mu, \omega, \omega_0, \beta \in \mathbb{R}_{>0}, \mathrm{p} \in \mathbb{R}_{>0}, \mathrm{a} \in \mathbb{R}^2, \mathrm{b} \in (0,1)^2, \mathrm{c} \in \mathbb{R}^3_{>0},$$
$$h^{(e)}, h^{(w)} \in (0,1), q^{(e)}, q^{(w)} \in \mathbb{R}_{>0}, \mathrm{d} \in \mathbb{R}_{>0}\Big).$$

This includes many parameters with an intuitive epidemiological interpretation, for instance the recovery rates from symptomatic or asymptomatic infection μ and ω, respectively. In order to construct prior distributions, we searched the epidemiological literature for estimated values and corresponding uncertainty given through 95% confidence intervals. For some parameters, in particular those concerning transmission, that is, the components of the contact matrix $\boldsymbol{C}$ or the relative infectiousness with symptomatic infection p, reliable evidence was not available in the literature. Hence, we assumed vague priors yielding a wide range for these parameters. On the natural space, most parameters are restricted to the positive real axis or to the interval $(0, 1)$. To take these parameter restrictions into account, we first applied an appropriate log or logit transformation if necessary. This allowed us to quantify priors using the same distributional family for the whole real axis, regardless of each parameter's natural space. We chose to construct the individual prior densities on the unrestricted transformed parameter space via a skew normal distribution ($\mathcal{SN}$) [28], as it offers three distribution parameters to be calibrated as opposed to the beta or gamma distribution, for example, making it a suitable choice in the case of prior information available as point estimators and a confidence interval. The $\mathcal{SN}$-density to the parameters m, σ and s is given by

$$\psi_{m,\sigma,s}(x) = \frac{2}{\sigma}\phi\left(\frac{x-m}{\sigma}\right)\Phi\left(s\left(\frac{x-m}{\sigma}\right)\right),$$

where ϕ and Φ are the standard normal distribution density and distribution function, respectively.

We transformed the estimated values and borders of the reported 95% confidence interval of each parameter via the log or logit transformation and interpreted the resulting values as the 0.025, 0.5 and 0.975 quantiles of the corresponding prior distribution on the transformed parameter space. This is a suitable interpretation, as the quantiles are invariant with respect to monotone transformation, whereas expected values are not. Then, for each parameter, the parameters m, σ and s were computed such that the corresponding skew normal distribution fits the three desired quantiles on the transformed space. To obtain the according distribution on the natural parameter space, we applied density transformation, which led to the following density functions:

$$f_X(x) = \frac{\psi_{m,\sigma,s}\left(\text{logit}\,(x)\right)}{x(1-x)}, \quad \text{if logit}\,(X) \sim \mathcal{SN}(m,\sigma,s),$$

or

$$f_X(x) = \frac{\psi_{m,\sigma,s}\left(\log(x)\right)}{x}, \quad \text{if } \log(X) \sim \mathcal{SN}(m,\sigma,s).$$

For most parameter configurations, these density functions yield shapes comparable with those of a beta or gamma distribution on their respective scale but offer more flexibility owing to the three parameters available. The joint prior distribution π of all parameters was then constructed as the product of the individual prior densities assuming independence between the parameters. See Table I for the prior distribution of each parameter and the corresponding knowledge we found from literature.

In order to avoid overparametrization of the model, we exclusively fixed the parameters on infection-induced immunity, which we found to have the highest collinearity to other model parameters. From epidemiological studies [26, 37, 38], the relative risk for a subsequent infection is defined at a factor of $\alpha_2 = 0.6$ and $\alpha_3 = 0.4$ after one or two infections, respectively. The probability of developing symptoms is $\theta_1 = 0.5$, $\theta_2 = 0.25$ and $\theta_3 = 0$ with the first, second and third infections, respectively.

3.2. Approximate posterior distributions

The likelihood of a single parameter vector ϑ was computed by assuming that each data point $D_t^{(j,i)}$ corresponding to age group j, region i and week t is independently negative-binomially distributed with expectation $h_\vartheta^{(i)}(t)Y_\vartheta^{(j,i)}(t)$ and dispersion parameter d_ϑ as described in Section 2.4. Hereby, subscript ϑ indicates the dependence of the corresponding variables on the model input parameter. Thus, the log-likelihood corresponding to the underlying independent negative binomial distribution of the observed case numbers D is given by

$$LL(\vartheta|D) = \sum_{t\in T}\sum_{j=1}^{n_D}\sum_{i=e,w}\log\left\{\text{f}\left(D_t^{(j,i)}\,\middle|\,h_\vartheta^{(i)}(t)\bar{Y}_\vartheta^{(j,i)}(t), d_\vartheta\right)\right\},$$

Table I. Prior distributions of all epidemiological parameters are given as skew normal distribution on an appropriate transformed scale.

Parameter	Interpretation	Prior distribution	Literature source
μ	Rate of recovery from symptomatic infection (1/duration in weeks)	$\log \mu \sim \mathcal{SN}(0.69, 0.31, 83)$	[29]
ω	Rate of recovery from asymptomatic infection (1/duration in weeks)	$\log \omega \sim \mathcal{SN}(0.08, 0.34, -1.38)$	[30, 31]
ω_0	Rate of waning maternal protection (1/duration in weeks)	$\log \omega_0 \sim \mathcal{SN}(-2.11, 0.35, 0)$	[32–34]
p	Relative infectiousness of symptomatically infected individuals	$\text{logit}\, p \sim \mathcal{SN}\,(0, 1.12, 0)$	(Assumption)
β	Rate of waning natural immunity	$\log \beta \sim \mathcal{SN}\,(-3.30, 0.43, -0.13)$	[26, 29, 35]
$c_i, i = 1, 2, 3$	Contact rates	$\log c_i \sim \mathcal{SN}\,(0, 0.82, 0)$	(Assumption)
$a_i, i = 1, 2$	Amplitude of seasonality	$a_i \sim \mathcal{SN}\,(0, 1.02, 0)$	(Assumption)
$b_i, i = 1, 2$	Phase shift of seasonality	$\text{logit}\, b_i \sim \mathcal{SN}\,(0, 1.87, 0)$	(Assumption)
$h^{(i)}, i = e, w$	Proportions of detected cases among symptomatic cases in eastern and western federal states before 2005	$\text{logit}\, h^{(i)} \sim \mathcal{SN}(-1.87, 0.47, -2.94)$	[36]
$q^{(i)}, i = e, w$	Relative increase of detection rate in eastern and western federal states	$\log m^{(i)} \sim \mathcal{SN}(0, 0.36, 0)$	(Assumption)
d	Dispersion parameter of the observations' negative binomial distribution	$\log d \sim \mathcal{SN}(-0.01, 0.82, 0)$	(Assumption)

Literature evidence where available is given in the last column.

where $\mathrm{f}(\cdot\,|\mu, d)$ is the probability mass function of the negative binomial distribution with expectation μ and dispersion d. Furthermore, $\bar{Y}$ denotes the model predicted incidences, which may be an aggregation of some of the n_A age strata defined in the model, such that they correspond to the n_D age groups available in the data, if necessary.

As a first step, we then computed the unnormalized posterior log-density $\log \mathscr{P}(\vartheta)$ of a given parameter vector ϑ as the sum of the prior log-density $\log \pi(\vartheta)$ and the earlier log-likelihood $LL(\vartheta\,|D)$:

$$\log \mathscr{P}(\vartheta) = LL\,(\vartheta|D) + \log \pi\,(\vartheta)\,, \tag{2}$$

which is the classical definition of the posterior distribution in Bayesian inference once you account for the normalization constant [39]. Unfortunately, the computational effort required for computing the normalized posterior distribution in this complex model setting is quite large. Typically, this computation would be performed by Markov Chain Monte Carlo techniques, which require a few hundred thousand evaluations of the posterior to guarantee convergence of the sample distribution because of the large number of parameters to be estimated within this multivariable setting. Because a single evaluation of our model and the likelihood is computationally quite expensive, we instead relied on the posterior being asymptotically normal around the posterior mode. Because of the size of our data set consisting of more than 8000 observations, the assumption of a nearly asymptotic posterior appears to be fair, as the asymptotic properties already hold for smaller sample sizes as shown in the simulation study in Section 3.5 of the supporting information. Therefore, on the basis of the unnormalized posterior log-density (2), we numerically computed the posterior mode using an alternating combination of gradient and Nelder–Mead simplex methods [40]. With every start of the Nelder–Mead method, the algorithm allowed the optimum to escape from a potential local maximum while intermediate gradient steps improved the overall performance of the algorithm, which eventually yields the posterior mode ϑ^*.

However, considering the potentially high residual autocorrelation within the data, we were concerned about overemphasizing the strength of the data subject to the independence assumption given the model. The remaining autocorrelation becomes particularly apparent when looking at the residuals from such modelling, that is, the deviation of the data to the expected case numbers coming from the model, which still exhibit a strong autocorrelation. Holding on to the assumption of independent observations would result in an overrating of the likelihood and consequently in an underestimation of the posterior variance.

Therefore, in a second step, we adjusted the likelihood for the actual information content of the data as induced by the autocorrelation of the residuals. This was carried out by estimating the CA, $\widehat{\mathrm{CA}}^{(j,i)}(\vartheta^*)$, for each given time series, that is, each age group and both regions, on the basis of the optimal calibration of the model after the first step using the posterior mode ϑ^*. The CA may be interpreted as a measure for the effective sample size as defined in [41]. Assuming that the log-likelihood $LL(\vartheta|D)$ of a parameter vector ϑ is not based on T independent observations for each time series, but on a number of independent observations given by $\widehat{\mathrm{CA}}(\vartheta^*)^{-1}T$, we discounted the likelihood by that number as suggested similarly in [42]. This might seem to be an ad hoc adjustment method, but it has a theoretical motivation (Section 2 of the supporting information) and allows a valid inference without having to reflect dependent observations through the model directly, which is a much more complicated task. The estimation approach of the CA, $\widehat{\mathrm{CA}}^{(j,i)}(\vartheta^*)$, for a certain age group j and region i, given the first step posterior mode ϑ^*, is explained in detail in Appendix 5. Altogether, this leads to an adjusted step 2 log-likelihood given by

$$LL_{\mathrm{CA}}(\vartheta|D) = \sum_{t\in T}\sum_{j=1}^{n_D}\sum_{i=e,w} \widehat{\mathrm{CA}}^{(j,i)}(\vartheta^*)^{-1} \log\left\{\mathrm{f}\left(D_t^{(j,i)}\,\middle|\,h_\vartheta^{(i)}(t)\bar{Y}_\vartheta^{(j,i)}(t), d_\vartheta\right)\right\}.$$

During this second step, the final unnormalized posterior log-density $\log \mathscr{P}_{\mathrm{CA}}$ incorporating the log-likelihood LL_{CA} is subject to optimization

$$\log \mathscr{P}_{\mathrm{CA}}(\vartheta) = LL_{\mathrm{CA}}(\vartheta|D) + \log \pi(\vartheta), \tag{3}$$

leading to a second posterior mode $\vartheta^*_{\mathrm{CA}}$. Computation of the second step posterior mode was performed analogously to the first step. On the basis of the asymptotic properties, our final approximate posterior distribution is defined as normally distributed with expectation $\vartheta^*_{\mathrm{CA}}$ and variance $-H^{-1}$, where H is the Hessian matrix of $\log \mathscr{P}_{\mathrm{CA}}$ evaluated at $\vartheta^*_{\mathrm{CA}}$, that is,

$$\vartheta_{\mathrm{CA}} \sim \mathcal{N}\left(\vartheta^*_{\mathrm{CA}}, -\left(\frac{d^2}{d\vartheta^2}\log \mathscr{P}_{\mathrm{CA}}(\vartheta^*_{\mathrm{CA}})\right)^{-1}\right).$$

Note that as one convergence criterion of the optimization process, we wanted the Hessian H at the final optimum to be negative definite such that the resulting posterior covariance is well defined.

3.3. Model averaging based on posterior distributions

In addition to the uncertainty in parameter estimation, we also wanted to address uncertainty in model selection by applying Bayesian model averaging techniques. The set of different models results from the choice between six different contact structures, represented through different contact matrices $\boldsymbol{C}_1, \ldots, \boldsymbol{C}_6 \in R_{>0}^{19\times 19}$, and three possible parameter spaces: $\mathscr{Z} = (Z_{(\mu,\omega)}, Z_{(\mu)}, Z_{()})$, where the lower index indicates the parameters that were fixed during estimation. The contact matrices we investigated are an extended pool of contact structures already used in [3, 5, 8]. A simplified representation of the six different types of contact matrix is shown later, where each matrix is given as a composition of submatrices regarding either the first 14, the subsequent three or the last two rows and columns, which correspond to the age groups young (0–4 years), middle (5–59 years) and old (60+ years), respectively. All components within one submatrix were assumed to be equal as determined by the corresponding parameter c_i, while certain submatrices were also modelled to share the same component value, according to the contact patterns shown in the following. Thus, each matrix calibration except for $\boldsymbol{C}_1$ depends on three parameters $c_1, c_2, c_3 > 0$.

$$\boldsymbol{C}_1 = \begin{pmatrix} c_1 & c_1 & c_1 \\ c_1 & c_1 & c_1 \\ c_1 & c_1 & c_1 \end{pmatrix}, \quad \boldsymbol{C}_2 = \begin{pmatrix} c_1 & 0 & 0 \\ 0 & c_2 & 0 \\ 0 & 0 & c_3 \end{pmatrix}, \quad \boldsymbol{C}_3 = \begin{pmatrix} c_1 & c_1 & c_1 \\ c_2 & c_2 & c_2 \\ c_3 & c_3 & c_3 \end{pmatrix},$$

$$\boldsymbol{C}_4 = \begin{pmatrix} c_1 & c_1 & c_3 \\ c_1 & c_2 & c_3 \\ c_3 & c_3 & c_3 \end{pmatrix}, \quad \boldsymbol{C}_5 = \begin{pmatrix} c_1 & c_2 & c_3 \\ c_2 & c_2 & c_3 \\ c_3 & c_3 & c_3 \end{pmatrix}, \quad \boldsymbol{C}_6 = \begin{pmatrix} c_1 & c_3 & c_3 \\ c_3 & c_2 & c_3 \\ c_3 & c_3 & c_3 \end{pmatrix}.$$

Regarding the parameter space, we decided to examine how fixing parameters beforehand affects the model fit and the estimation of other relevant model parameters. For that investigation, we selected the parameters determining the duration of illness, as the corresponding evidence has a high degree of uncertainty and we expected those parameters to significantly correlate with other model parameters. In the cases of parameters being fixed, they were fixed at their respective point estimates as shown in Table I.

Given any set of models $\{M_i\}_{i\in I}$, the Bayes factor $B_{i,0}$ for a model M_i was computed as

$$B_{i,0} = \frac{\mathrm{pr}\,(D|M_i)}{\mathrm{pr}\,(D|M_0)},$$

with M_0 being any reference model and $\mathrm{pr}\,(D|M_i)$ being the marginal likelihood of model M_i. As the posterior in our case is approximated by a normal distribution, the marginal likelihood can be computed exactly by

$$\begin{aligned}\mathrm{pr}\,(D|M_i) &= \int_\Theta \exp\left\{LL_{\mathrm{CA}}\left(D\,|\vartheta, M_i\right)\pi_i\left(\vartheta\,|M_i\right)\right\}\,\mathrm{d}\vartheta \\ &= (2\pi)^{p_i/2}\,|-H_i|^{\frac{1}{2}}\,LL_{\mathrm{CA}}\left(D\,\middle|\vartheta^*_{\mathrm{CA},i}, M_i\right)\pi_i\left(\vartheta^*_{\mathrm{CA},i}\,|M_i\right),\end{aligned} \tag{4}$$

where p_i is the dimension of the input parameter ϑ_i belonging to model M_i and H_i is the Hessian matrix of $\log \mathscr{P}_{\mathrm{CA}}\left(\cdot\,|D, M_i\right)$ evaluated at $\vartheta^*_{\mathrm{CA},i}$ as in the preceding equation. In the case of a nonnormal posterior, the preceding term corresponds to the marginal likelihood approximation based on Taylor expansion as suggested in [43].

To guarantee comparability of the models, we had to make sure that the final posterior distribution of each model was adjusted by the same CA factor. Otherwise, high autocorrelations measured in the first step would lead to higher posterior modes in the second step optimization, which is counterintuitive. Therefore, we had to choose one set of CAs $\{\widehat{\mathrm{CA}}^{(i,j)}\}$ to be used for every model. Hence, we computed the mean CA $\{\overline{\mathrm{CA}}^{(i,j)}\}$ over all $m = 18$ models and proceeded with those mean CAs in the second step of our optimization instead. Because the variation of the estimated CAs $\mathrm{CA}^{(i,j)}$ for a specific time series (i, j) given the $m = 18$ models is small, taking the mean is a convenient choice.

Finally, we computed the individual model weights w_i from the Bayes factors assuming equal prior model probability, resulting in

$$w_i = \frac{\mathrm{pr}\,(D|M_i)}{\sum_{k=1}^{m}\mathrm{pr}\,(D|M_k)}.$$

Note the possibility of using only a subset of models defined earlier for the averaging procedure. Two classes of subsets are of particular interest: those models that share a common contact structure and those models that share the same parameter space, which means that the parameters μ and ω are either fixed or subject to estimation for each model. We will refer to those model subsets as horizontal and vertical averaging, respectively.

3.4. Computing the averaged model predictions

Knowing the weights $\{w_i\}$ corresponding to an ensemble of models $\{M_i\}_{i\in I}$, we computed the model predictions on the age-stratified incidence $X_t^{(i,j)}$ and its expectation $h_t^{(i)}\cdot Y_t^{(i,j)}$ for any time t by sampling a model and then sampling from its respective posterior distribution. To obtain the desired quantities, we applied the following algorithm for our ensemble of models:

- For k in 1 to K, repeat the following:
  (1) Draw a model $M^{(k)}$ from $\{M_i\}_{i\in I}$ according to the probabilities $\{w_i\}_{i\in I}$
  (2) Draw a parameter vector $\vartheta^{(k)}$ according to the posterior distribution $\mathscr{P}_{\mathrm{CA}}^{M^{(k)}}$ of $\vartheta^{(k)}$ in model $M^{(k)}$
  (3) Compute the expected incidences $h_t^{(i)}\cdot Y_t^{(i,j)}$ resulting from model $M^{(k)}$ with parameter $\vartheta^{(k)}$ and draw a sample for the observed incidence $X_t^{(i,j)} \sim \mathrm{NegBin}\left(h_t^{(i)}\cdot Y_t^{(j,i)}, d_{\vartheta^{(k)}}\right)$ for $t = 1,\ldots,T$, $i = e, w$ and $j = 1,\ldots,n_D$

The sample consisting of the K arrays for each of $X = \left(X_t^{(i,j)}\right)$ and $\mathbb{E}[X] = \left(\mathbb{E}\left[X_t^{(i,j)}\right]\right) = \left(h_t^{(i)} \cdot Y_t^{(i,j)}\right)$ has the desired distribution given by

$$\mathbb{P}\left(X_t^{(i,j)} = x\right) = \sum_{k \in I} w_i \cdot \mathbb{P}\left(X_t^{(i,j)} = x \,|M_k\right) \quad \text{and} \quad \phi\,(x) = \sum_{k \in I} w_i \cdot \phi\,(x\,|M_k)\,, \tag{5}$$

where $\mathbb{P}$ and ϕ denote the probability mass function of X and density function of $\mathbb{E}\,[X]$, respectively.

3.5. The averaged posterior distribution

Given a set of weighted models $\{M_i, w_i\}_{i \in I}$, we can compute the averaged posterior distribution as a weighted sum of the corresponding cumulative distribution functions:

$$\mathscr{P}\,(\vartheta \leqslant \vartheta_0) = \sum_{i \in I} w_i \cdot \mathscr{P}_{\mathrm{CA}}^{M_i}\,(\vartheta \leqslant \vartheta_0)\,. \tag{6}$$

By interpreting fixed parameters as a probability distribution with its entire mass concentrated in one point, we can therefore compute an averaged cumulative posterior distribution function, even if not all considered models share the same parameter space such as in the case of vertical averaging, where an averaging of the posterior densities is not possible.

In the case of horizontal averaging, because all models share a common parameter space, we obtain an averaged probability density computed equivalently:

$$\phi\,(\vartheta) = \sum_{i \in I} w_i \cdot \phi^{M_i}\,(\vartheta)\,,$$

where $\phi^{M_{(i)}}$ denotes the probability density belonging to model M_i. Such a representation is also possible for the marginal posterior densities of individual parameters, if the parameter posterior distribution was subject to estimation.

4. Application to German rotavirus incidence data

This section contains the results of modelling the reported rotavirus incidence data from Germany presented in Section 2.1. We computed the first-step and second-step marginal likelihood as described earlier for all 18 models originating from the six different contact patterns and the three possible parameter configurations in $\mathscr{Z} = \left\{Z_{(\mu,\omega)}, Z_{(\mu)}, Z_{()}\right\}$ (the index indicates the parameters that were fixed in advance). On the basis of the marginal likelihoods, we applied Bayesian averaging to a selected set of models, as described in the following.

In our base case analysis, we selected all 18 models to be subject to averaging. The corresponding model weights are displayed in the left panel of Table II. When comparing the model probabilities, it becomes clear that the models using contact pattern $\boldsymbol{C}_6$ reached the highest marginal likelihoods. This coincides with the results from the European POLYMOD study on social mixing patterns [44], with contact matrix $\boldsymbol{C}_6$ yielding the closest fit to the German contact pattern from this study. The contact and parameter combination $(\boldsymbol{C}_6, Z_{()})$ resulted in a marginal likelihood with a log difference of more than 4 compared with any other model, yielding a probability of $w(\boldsymbol{C}_6, Z_{()}) = 0.977$ and hence dominating all other models. The model weights computed in this base case analysis were used later for computing predictions regarding the German rotavirus incidence.

Table II. Model weights resulting from the model-specific marginal likelihoods.

Total weights	$\boldsymbol{C}_1$	$\boldsymbol{C}_2$	$\boldsymbol{C}_3$	$\boldsymbol{C}_4$	$\boldsymbol{C}_5$	$\boldsymbol{C}_6$
$Z_{(\mu,\omega)}$	0.000	0.000	0.000	0.000	0.000	0.023
$Z_{(\mu)}$	0.000	0.000	0.000	0.000	0.000	0.000
$Z_{()}$	0.000	0.000	0.000	0.000	0.000	0.977

Vertical weights	$\boldsymbol{C}_1$	$\boldsymbol{C}_2$	$\boldsymbol{C}_3$	$\boldsymbol{C}_4$	$\boldsymbol{C}_5$	$\boldsymbol{C}_6$
$Z_{(\mu,\omega)}$	0.034	0.000	0.999	0.275	0.888	0.023
$Z_{(\mu)}$	0.966	0.000	0.000	0.177	0.111	0.000
$Z_{()}$	0.000	1.000	0.001	0.548	0.001	0.977

The left panel shows weights resulting from all 18 constructed models being subject of averaging, whereas the right panel shows the vertical weights resulting from averaging among models sharing the same contact structure.

In a second analysis, we applied horizontal averaging; that is, we averaged among the three sets of models sharing the same parameter configuration given through either $Z_{(\mu,\omega)}$, $Z_{(\mu)}$ or $Z_{()}$. Each of these sets contained six models with pairwise different contact patterns. All three sets of models yielded the same result of contact structure $\boldsymbol{C}_6$ giving the best model fit. This could also be derived from the model probabilities computed in our first analysis by calculating the probabilities $w(\boldsymbol{C}_j, z\,|z)$ conditioned on the parameter configuration $z \in \mathscr{Z}$.

In an analogous analysis, we calculated the vertical model weights; that is, we averaged within each of the six sets of models that share the same contact structure but differ with respect to their parameter configuration $z \in \mathscr{Z}$. These vertical weights cannot easily be obtained by looking at the total weights displayed in the left panel of Table II, owing to the limited number of digits used. However, the recalculated vertical weights for each contact pattern $\boldsymbol{C}_i, i = 1,\ldots,6$, are given in the right panel of Table II. Surprisingly, there is generally no dominant parameter configuration $z \in \mathscr{Z}$ as the parameter configuration yielding the highest model probability varies for different underlying contact structures.

To obtain predictions for the German reported rotavirus incidence for each age group and region, we computed a total of $K = 5000$ samples from the 18 models according to the base case model weights as shown in the left panel of Table II and each model's individual posterior distribution. Using this sample, we calculated the expected number of reported cases $\mathbb{E}\left[X_t^{(j,i)}\right]$ for each age group $j = 1,\ldots,n_D$, region $i = e, w$ and week $t = 1,\ldots,T$ as well as the 95% prediction interval for the expectation $\mathbb{E}\left[X_t^{(j,i)}\right] = h_t^{(i)} \cdot Y_t^{(j,i)}$ with respect to the uncertainty regarding model choice and the respective input parameter. We also computed the 95% prediction interval for the observation $X_t^{(j,i)} \sim \text{NegBin}(h_t^{(i)} \cdot Y_t^{(j,i)}, d)$ including also uncertainty regarding the stochastic observation of events. These results are plotted in Figure 2 for the three age groups 0–4, 5–59 and 60–99 years of age for both EFS and WFS. We see that the vast majority of observed data points are included

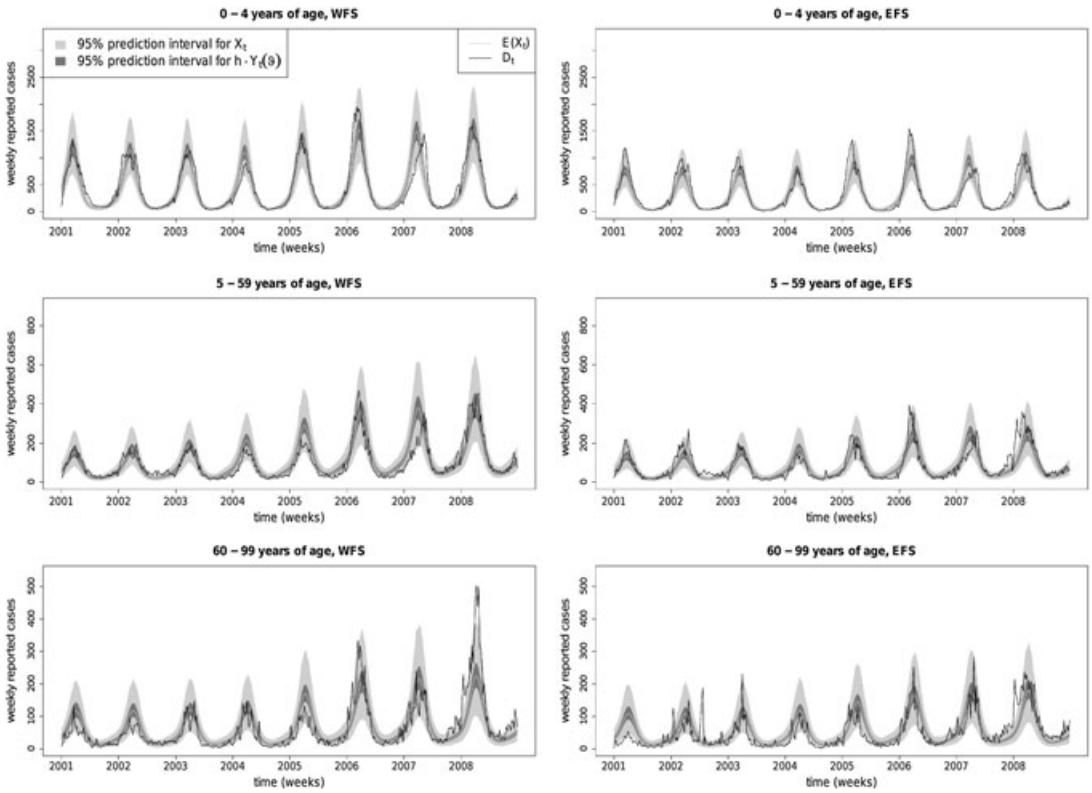

Figure 2. Weekly number of reported rotavirus cases D_t among the three aggregated age groups of 0–4, 5–59 and 60–99 years of age in the western (left) and eastern (right) federal states. Also shown are the aggregated model-averaged predictive distributions of X_t and $\mathbb{E}[X_t]$. Note the different scaling of the y-axis for each age group.

within the 95% prediction intervals, with no clear sign of overestimation or underestimation by the model-predicted expectation in any group. Considering the age group 60–99 years, which exhibits a higher degree of fluctuation due to a smaller base incidence, the model is still able to catch the average incidence. The model-predicted seasonality matched the observed rotavirus seasonality very well, with the only exception of the 2007 season. The reasons for the observed delay in incidence are still unknown to the epidemiologists; possibilities are the prevalence of a specific serotype or special environmental conditions in that season. Although our model was primarily developed to mimic the long-term dynamics of the expected German rotavirus incidence, it also captures the observational uncertainty very well for any age group and region, which makes it very suitable for future predictions.

4.1. Epidemiological insights

Beside the ability to compute a predictive distribution for the future rotavirus incidence, we also obtained posterior distributions for each model parameter. For epidemiological quantities, which directly find a representation as a parameter in the model, we can interpret these posterior distributions as updated knowledge when combining prior knowledge with the available time series data. Especially for those parameters, that come with a high degree of prior uncertainty due to difficulties in assessing them in studies in real life, the estimates are of particular interest. Examples of such parameters are the detection ratio or the relative infectiousness of symptomatically infected individuals, which includes aspects such as higher excretion of virus and also a lower number of potentially infectious contacts due to bed rest. From the weighted posterior distributions, we can compute point estimates by taking, for example, the posterior median as well as 95% equi-tailed credibility intervals (CI).

Especially the difference in underdetection between the EFS and WFS is of interest, because significant incidence differences between EFS and WFS have been found among other notifiable diseases in Germany as well [24]. However, despite the lack of detailed prior information on the reporting behaviour in Germany, we were able to obtain sharp estimates for the parameters $h^{(w)}$ and $h^{(e)}$. For this analysis, we considered only the three models using contact structure $\boldsymbol{C}_6$ and computed their marginal posterior distributions for the parameters $h^{(w)}$ and $h^{(e)}$. In Figure 3, we see that the posterior median for the detection ratio in the WFS $h^{(w)}$ was computed at 4.5% (95% CI 4.1–4.9%), 4.1% (3.7–4.5%) and 4.3% (4.0–4.7%) for the model using parameter configuration $Z_{(\mu,\omega)}$, $Z_{(\mu)}$ and $Z_{()}$, respectively. The averaged posterior density using the weights from the vertical averaging regarding contact pattern $\boldsymbol{C}_6$ shows a median estimate of 4.3% (3.9–4.7%) for parameter $h^{(w)}$.

Regarding the estimation of the detection ratio in the EFS, the same models computed posterior medians of 19.6% (18.1–21.1%), 18.3% (16.9–19.7%) and 19.0% (17.6–20.4%). The averaged posterior distribution for this parameter suggests a median estimate of 19.0% (17.6–20.5%). These results not only imply that the detection ratio in the EFS is more than four times higher compared with that in the WFS but also yield concrete estimates for the disease's underdetection, which is otherwise difficult to assess. Note that these ratios describe only the period from 2001 till 2004. Our model allowed a temporal break of these parameters at the end of 2004, such that the detection ratios from 2004 onwards are estimated at 6.3% (5.7–6.9%) and 24.1% (22.3–25.9%) in the WFS and EFS, respectively. This reflects

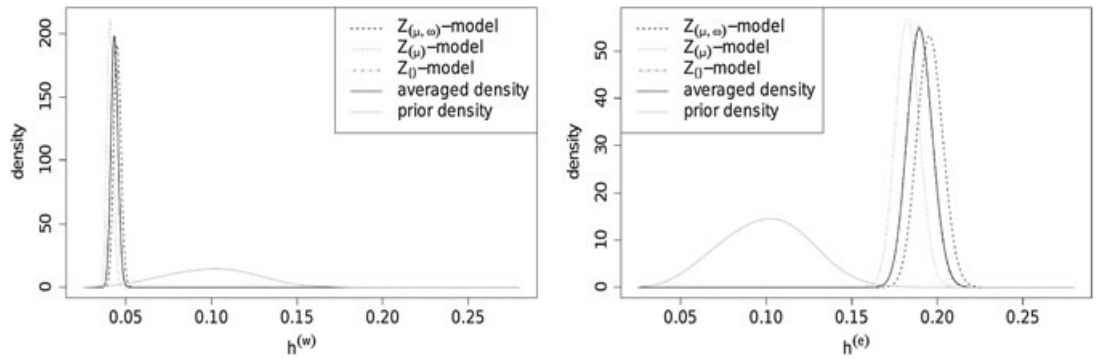

Figure 3. Single and vertically averaged posterior distribution of detection rates $h^{(w)}$ (left) and $h^{(e)}$ (right) in the western and eastern federal states coming from the models using contact pattern $\boldsymbol{C}_6$.

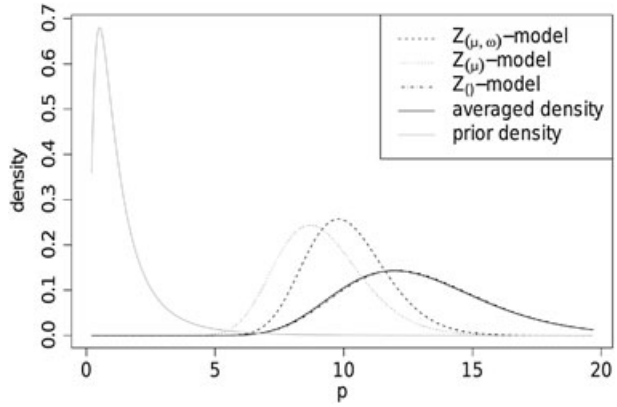

Figure 4. Single and vertically averaged posterior distribution of relative infectiousness of symptomatically infected individuals p according to the models using contact pattern $\boldsymbol{C}_6$.

the potential effect of the change in the reimbursement such that laboratory confirmation was requested for more cases after 2004. Note that the relative increase was higher in the WFS, which may also imply a higher degree of underdetection in that region beforehand.

We conducted the same analysis for the parameter p, which represents the relative infectiousness of symptomatically infected individuals, using again the three models with contact pattern $\boldsymbol{C}_6$ (Figure 4). According to the choice of parameter space from $Z_{(\mu,\omega)}$, $Z_{(\mu)}$ and $Z_{()}$, we obtained posterior medians of 10.0 (7.4–13.6), 9.0 (6.3–12.9) and 12.7 (8.2–19.6), respectively. Calculating the averaged posterior distribution using the same weights as before, we obtained an averaged posterior median of 12.6 with CI (8.1–19.6). This estimate suggests that a symptomatically infected person causes more than 10 times as many secondary rotavirus infections as an infected person without symptoms. This difference may be easily explained by the increased excretion of virus and only few viruses being necessary for infection. However, it was not entirely clear whether these aspects would outweigh the dampening effects, such as lesser physical contacts and a potentially higher attention for hygienic countermeasures, that might actually decrease the number of secondary cases induced by a symptomatic infection. Therefore, with the goal of taking targeted intervention measures to reduce disease burden, it is important to know which group has the highest impact on the force of infection, where this new insight provides an explicit answer to that problem.

Another interesting aspect of the employed model is the interplay of immunity gain through infection versus immunity loss over time. Considering the high incidence among children and the rising incidence among elderly, it appears that the typical age-specific immunity states within the model change over the life course (Figure 1). Although children rapidly move through the first two infectivity states I_k and A_k ($k = 1, 2$), once they arrive at S_3, they frequently go through asymptomatic infection A_3, maintaining their high immunity level. With high age, the contact rate decreases and so does the force of infection such that a loss of immunity ($S_3 \rightarrow S_2$) becomes more common. Therefore, according to the model, infections among the elderly happen less frequently than in the middle-age group but with a higher chance of proceeding symptomatically, which explains the higher reported incidences in the elderly.

4.2. Comparison with different transmission models

We re-implemented the transmission models presented in [3, 5] and applied them to the German rotavirus data. Both model structures were taken as described by the ODE systems in the respective works for the USA [3] and England and Wales [5] and supplemented with adopted versions of our observational component and our definition of the force of infection including the proposed contact pattern $\boldsymbol{C}_6$. Using the same inference approach, both models yielded significantly lower marginal log-likelihoods. The reference models had in particular problems to mimic the incidences in the age groups 60+, with especially the model from [3] heavily underestimating the corresponding number of reported cases.

4.3. Statistical insights

The main intention of our Bayesian approach was to account for uncertainty originating from both parameter quantification and model selection. Because the calculation of the model-specific posterior distribution was conducted independently for each model, we were able to examine how different model assumptions yield different results for the parameter vector ϑ. This is of special interest considering the consequences such estimates may have on health policymaking.

4.3.1. Parameter spaces and collinearities. By looking at the median estimates for the relative infectiousness of symptomatically infected individuals p as shown in Figure 4, it becomes clear that certain parameter estimates react very sensitively to changes in model structure and parameter space. Of particular interest is the impact of fixing parameter values in advance as often performed in the literature [3,5,8]. In our case, the models using configurations $(\boldsymbol{C}_6, Z_{(\mu,\omega)})$ and $(\boldsymbol{C}_6, Z_{(\mu)})$, where one or both of the parameters $\mu = 2.333$ and $\omega = 0.875$ were fixed, had a posterior median of variable p at 10.0 and 9.0, respectively. In contrast, the model using the full space $(\boldsymbol{C}_6, Z_{()})$ yielded a median estimate for p at 12.6.

Adding only ω to the parameter vector ϑ did yield a slight decrease in the estimated value for p, whereas the posterior median for ω itself at 1.93 was higher than its corresponding fixed value. Adding also μ, we obtained notably different estimates. In the corresponding model $(\boldsymbol{C}_6, Z_{()})$, the posterior medians for ω and μ were given at 1.01 and 1.02, significantly lower than the estimates from model $(\boldsymbol{C}_6, Z_{(\mu)})$, whereas the median for p at 12.6 increased. With a closer look at the corresponding posterior correlations, it seems that the model prefers parameter vectors yielding the same μ/ω ratio and compensates variations in these mean durations of infection by adjusting the contact rates c_1, c_2, c_3 and p.

The preceding insights are also useful when thinking about potential extensions of the model structure. One modification could be the introduction of an infectious period preceding the symptomatic phase, which was left out in the original model because epidemiological studies [29,30] have shown this period to be rather short. However, although the estimated durations of symptomatic and asymptomatic phase might change under such a model variation, the relative infectiousness p is likely to be robust. This can be seen from the duration parameters μ and ω varying significantly among the compared models, while the estimate for parameter p remains relatively stable, as it is primarily affected by the mean infectiousness of a childhood case (many symptomatic infections) versus an adult case (almost no symptomatic infections), but only partially by the duration of symptomatic and asymptomatic periods.

4.3.2. Two-step optimization. Also of interest is the impact of the second step optimization on the posterior, in which we adjusted for the estimated CA in the data as given in equation (3). Using the modified likelihood LL_{CA} had significant effects on the final estimation results, as we can see by analysing the posterior estimation for parameter $h^{(e)}$.

As shown in Figure 3 in model $(\boldsymbol{C}_6, Z_{(\mu,\omega)})$, the second step posterior median for $h^{(e)}$ was 19.0% with 95% CI (17.6–20.5%). In contrast, after the first optimization step using only the posterior density as defined in equation (2), the median was estimated at 19.6% (19.3–19.9%). By reducing the impact of the likelihood by a factor determined by the CA, the posterior density flattened around the posterior mode. As a consequence, the posterior median estimates moved closer to the prior medians, but more importantly, we obtained higher estimates for the posterior variance and hence also wider credibility regions.

Furthermore, by downscaling the likelihood, the marginal likelihoods of each of the selected models, as computed via equation (4), moved closer together. This can also be observed in the case of vertical averaging with respect to contact structure $\boldsymbol{C}_5$, where the largest difference of the marginal likelihoods was 7.2 on the log scale after the second step procedure while the minimal difference after only one estimation step was 25.4, which would lead to nearly degenerated model weights.

5. Discussion

In this paper, we presented a realistic model for the transmission and reporting of rotavirus in Germany representing all epidemiologically relevant aspects. Such a detailed mechanistic model is a necessary precondition for investigating potential intervention measures such as a nationwide routine vaccination.

Although our model building was heavily focused on describing the rotavirus dynamics in Germany, one can easily point out those aspects of model building that need to be adapted to other pathogens sharing similar transmission characteristics.

We used a Bayesian framework to put focus on the analysis of uncertainty regarding model parameters and model selection. To address the strong remaining autocorrelation in the residuals of our ODE-based modelling, we applied a pragmatic and robust correction method by scaling the likelihood function. The Bayesian approach allowed us to gain knowledge on the underlying epidemiology of disease transmission as was performed similarly in [15, 16]. Among other epidemiological insights, we obtained estimates for the region-specific detection rates in Germany and also found the force of infection to be primarily driven by symptomatically as opposed to asymptomatically infected individuals.

As the underlying transmission process is exceedingly complex, it is not feasible to construct a mathematical model embodying the 'true dynamics'. Hence, the observed residual autocorrelation is an indicator of this imperfection. Thus, acknowledging this model dependence is important when interpreting the corresponding inference results. To lessen the reliance on certain model assumptions, besides accounting for the inherent residual autocorrelation, we additionally applied methods to integrate several model structures into the estimation process, using Bayesian averaging tools to generate synthesized results. Altogether, our approach is novel in the field of disease transmission modelling based on ODEs as uncertainty regarding both model structure and parameter is treated in a more comprehensive way, and we are able to incorporate this into the future analysis of intervention strategies.

Considering the strong dependence between the parameter and model, and also the multivariate correlation among the input parameters, we found that one should be very careful with fixing crucial input parameters as it is sometimes carried out in other studies. Because the fixing of certain model parameters affects the inference of each other input parameter, the resulting transmission dynamics governed by the final model calibration might be biased. Because of an erroneous predictive behaviour of the model, this might have serious consequences for public health decisions. It is important to note that such effects cannot be revealed by conducting univariate sensitivity analysis on parameters in question, as is often performed to quantify predictive uncertainty in the context of infectious disease modelling.

We also applied adapted versions of the models that were developed previously [3, 5] to the German rotavirus incidence data. On the basis of marginal likelihood criteria, our model yielded a significantly better fit, especially regarding the elder age groups. However, those reference models were aimed at the rotavirus epidemiology in the USA and England and Wales and also primarily at the target group of children younger than 5 years of age, which requires a different focus in model building. This makes a comparison of the models difficult but again shows how dependent such models are on the underlying data and processes used for developing them.

Although our model captures very well the long-term trends of rotavirus incidence in Germany among all age groups, including its seasonal pattern, it would have also been favourable if the model could account for possible interseasonal variation regarding the onset and magnitude of the incidence peak. To improve performance in this matter, it might be worthwhile to introduce more stochasticity into the transmission component. Another possible extension is a further stratification of the observational component to investigate differences in the disease underdetection and reporting over time and with respect to age, for example, by including age and time-dependent detection rates and overdispersion. Furthermore, it might be desirable to adopt suitable Markov chain Monte Carlo techniques suitable for such complex model settings, which would simultaneously enable corresponding methods for model averaging as in [45]. However, in the presented work, we already took a good step towards a more reasonable treatment of uncertainty in epidemiological model building.

Appendix A

A.1 The system of differential equations

The following system of differential equations describes the unobserved transmission dynamics according to our model. These equations hold for all age groups $j = 1, \ldots, n_A$ and use the indicator function $\mathbb{I}_A(x)$, which is 1 for $x \in A$ and 0 otherwise. $N^{(j)}$ denotes the total size of age group j.

$$
\begin{aligned}
\frac{dM^{(j)}}{dt} &= \mathbb{I}_{\{1\}}(j)b_t + \mathbb{I}_{\{2,\ldots,19\}}(j)\delta^{(j-1)}M^{(j-1)} - \left(\omega_0 + \delta^{(j)} + d^{(j)} - \frac{m^{(j)}}{N^{(j)}}\right)M^{(j)} \\
\frac{dS_1^{(j)}}{dt} &= \mathbb{I}_{\{2,\ldots,19\}}(j)\delta^{(j-1)}S_1^{(j-1)} + \omega_0 M^{(j)} + \beta S_{2B}^{(j)} - \left(\lambda^{(j)}(t) + \delta^{(j)} + d^{(j)} - \frac{m^{(j)}}{N^{(j)}}\right)S_1^{(j)} \\
\frac{dI_1^{(j)}}{dt} &= \mathbb{I}_{\{2,\ldots,19\}}(j)\delta^{(j-1)}I_1^{(j-1)} + \theta_1\lambda^{(j)}(t)S_1^{(j)} - \left(\mu + \delta^{(j)} + d^{(j)} - \frac{m^{(j)}}{N^{(j)}}\right)I_1^{(j)} \\
\frac{dA_1^{(j)}}{dt} &= \mathbb{I}_{\{2,\ldots,19\}}(j)\delta^{(j-1)}A_1^{(j-1)} + (1-\theta_1)\lambda^{(j)}(t)S_1^{(j)} - \left(\omega + \delta^{(j)} + d^{(j)} - \frac{m^{(j)}}{N^{(j)}}\right)A_1^{(j)} \\
\frac{dS_2^{(j)}}{dt} &= \mathbb{I}_{\{2,\ldots,19\}}(j)\delta^{(j-1)}S_2^{(j-1)} + \omega A_1^{(j)} + \beta S_{3B}^{(j)} - \left(\alpha_2\lambda^{(j)}(t) + \beta + \delta^{(j)} + d^{(j)} - \frac{m^{(j)}}{N^{(j)}}\right)S_2^{(j)} \\
\frac{dS_{2A}^{(j)}}{dt} &= \mathbb{I}_{\{2,\ldots,19\}}(j)\delta^{(j-1)}S_{2A}^{(j-1)} + \beta S_2^{(j)} - \left(\alpha_2\lambda^{(j)}(t) + \beta + \delta^{(j)} + d^{(j)} - \frac{m^{(j)}}{N^{(j)}}\right)S_{2A}^{(j)} \\
\frac{dS_{2B}^{(j)}}{dt} &= \mathbb{I}_{\{2,\ldots,19\}}(j)\delta^{(j-1)}S_{2B}^{(j-1)} + \beta S_{2A}^{(j)} - \left(\alpha_2\lambda^{(j)}(t) + \beta + \delta^{(j)} + d^{(j)} - \frac{m^{(j)}}{N^{(j)}}\right)S_{2B}^{(j)} \\
\frac{dI_2^{(j)}}{dt} &= \mathbb{I}_{\{2,\ldots,19\}}(j)\delta^{(j-1)}I_2^{(j-1)} + \theta_2\alpha_2\lambda^{(j)}(t)\left(S_2^{(j)} + S_{2A}^{(j)} + S_{2B}^{(j)}\right) - \left(\mu + \delta^{(j)} + d^{(j)} - \frac{m^{(j)}}{N^{(j)}}\right)I_2^{(j)} \\
\frac{dA_2^{(j)}}{dt} &= \mathbb{I}_{\{2,\ldots,19\}}(j)\delta^{(j-1)}A_2^{(j-1)} + (1-\theta_2)\alpha_2\lambda^{(j)}(t)\left(S_2^{(j)} + S_{2A}^{(j)} + S_{2B}^{(j)}\right) - \left(\omega + \delta^{(j)} + d^{(j)} - \frac{m^{(j)}}{N^{(j)}}\right)A_2^{(j)} \\
\frac{dS_3^{(j)}}{dt} &= \mathbb{I}_{\{2,\ldots,19\}}(j)\delta^{(j-1)}S_3^{(j-1)} + \omega\left(A_2^{(j)} + A_j^3\right) - \left(\alpha_3\lambda^{(j)}(t) + \beta + \delta^{(j)} + d^{(j)} - \frac{m^{(j)}}{N^{(j)}}\right)S_3^{(j)} \\
\frac{dS_{3A}^{(j)}}{dt} &= \mathbb{I}_{\{2,\ldots,19\}}(j)\delta^{(j-1)}S_{3A}^{(j-1)} + \beta S_3^{(j)} - \left(\alpha_3\lambda^{(j)}(t) + \beta + \delta^{(j)} + d^{(j)} - \frac{m^{(j)}}{N^{(j)}}\right)S_{3A}^{(j)} \\
\frac{dS_{3B}^{(j)}}{dt} &= \mathbb{I}_{\{2,\ldots,19\}}(j)\delta^{(j-1)}S_{3B}^{(j-1)} + \beta S_{3A}^{(j)} - \left(\alpha_3\lambda^{(j)}(t) + \beta + \delta^{(j)} + d^{(j)} - \frac{m^{(j)}}{N^{(j)}}\right)S_{3B}^{(j)} \\
\frac{dI_3^{(j)}}{dt} &= \mathbb{I}_{\{2,\ldots,19\}}(j)\delta^{(j-1)}I_3^{(j-1)} + \theta_3\alpha_3\lambda^{(j)}(t)\left(S_3^{(j)} + S_{3A}^{(j)} + S_{3B}^{(j)}\right) - \left(\mu + \delta^{(j)} + d^{(j)} - \frac{m^{(j)}}{N^{(j)}}\right)I_3^{(j)} \\
\frac{dA_3^{(j)}}{dt} &= \mathbb{I}_{\{2,\ldots,19\}}(j)\delta^{(j-1)}A_3^{(j-1)} + (1-\theta_3)\alpha_3\lambda^{(j)}(t)\left(S_3^{(j)} + S_{3A}^{(j)} + S_{3B}^{(j)}\right) - \left(\omega + \delta^{(j)} + d^{(j)} - \frac{m^{(j)}}{N^{(j)}}\right)A_3^{(j)}
\end{aligned}
\tag{7}
$$

A.2 Computing the cumulative autocorrelation for the rotavirus transmission model

The CA is a measure for the number of effectively independent data points within a time series similar to the effective sample size presented in [41]. The CA was used as an adjustment factor of the log-likelihood function to account for the reduced information content within dependent samples as described in Section 3.2. For the rational behind adjusting the likelihood on the basis of the CA, see also Section 2 of the supporting information providing an analytical assessment. The CA of a time series $X = (X_t)_{t=1,\ldots,T}$ is defined by

$$
\mathrm{CA} = \left(\sum_{\tau=-T}^{+T}\hat{\rho}(\tau)\right), \tag{8}
$$

where $\hat{\rho}(\tau)$ are estimates of the symmetric autocorrelation of lag τ within the time series X.

In our application, there are $2 \cdot n_D$ time series to analyse, resulting from the n_D age groups and the two regions supported by the data. For each of the time series, we carried out an Anscombe transformation of the residuals [46] with the aim of making the residuals to be distributed 'as normal as possible'. The reasoning behind this transformation is that all dependence within normally distributed random vectors are contained in their correlation. The series of Anscombe residuals $(r_t^A)_{t=1,\ldots,T}$ for a time series $X = (X_t)_{t=1,\ldots,T}$ is defined as

$$
r_t^A = \frac{A(X_t) - A(\hat{\mu}_t)}{V(\hat{\mu}_t)^{\frac{1}{6}}},
$$

with $(\hat{\mu}_t)_{t=1,\ldots,T}$ being the series of estimated expectations at each time point and the function A being

$$A(x) = \int_{-\infty}^{x} V(\mu)^{-\frac{1}{3}} \mathrm{d}\mu, \tag{9}$$

where $V(\mu)$ denotes the expectation-dependent variance function. In the case of a negative binomial response, this function takes the form $V(\mu) = \mu + \mu^2/d$. Hence, (9) leads to

$$r_t^A = \frac{\frac{3}{2} X_t^{\frac{2}{3}} {}_2F_1\left(\frac{1}{3}, \frac{2}{3}, \frac{5}{3}, -\frac{X_t}{\hat{d}}\right) - \frac{3}{2}\hat{\mu}_t^{\frac{2}{3}} {}_2F_1\left(\frac{1}{3}, \frac{2}{3}, \frac{5}{3}, -\frac{\hat{\mu}_t}{\hat{d}}\right)}{\left(\hat{\mu}_t + \frac{\hat{\mu}_t^2}{\hat{d}}\right)^{\frac{1}{6}}}.$$

Here, ${}_2F_1$ denotes the Gaussian hypergeometric function [47]. We applied this transformation to each of the 20 time series $(D_t^{(j,i)})_{t=1,\ldots,T}$ with the corresponding expectations

$$\hat{\mu}_t^{(j,i)} = h_{\vartheta^*}^{(i)}(t) Y_{\vartheta^*}^{(j,i)}(t)$$

computed through our model and dispersion $\hat{d} = d_{\vartheta^*}$ given by the posterior mode ϑ^* from the first optimization step.

To obtain an estimation of the autocorrelations within the newly constructed time series $(r_t^A)_{t=1,\ldots,T}^{(j,i)}$, we fitted an ARMA$(p, q)$-process to each time series separately. The orders p^* and q^* of the process

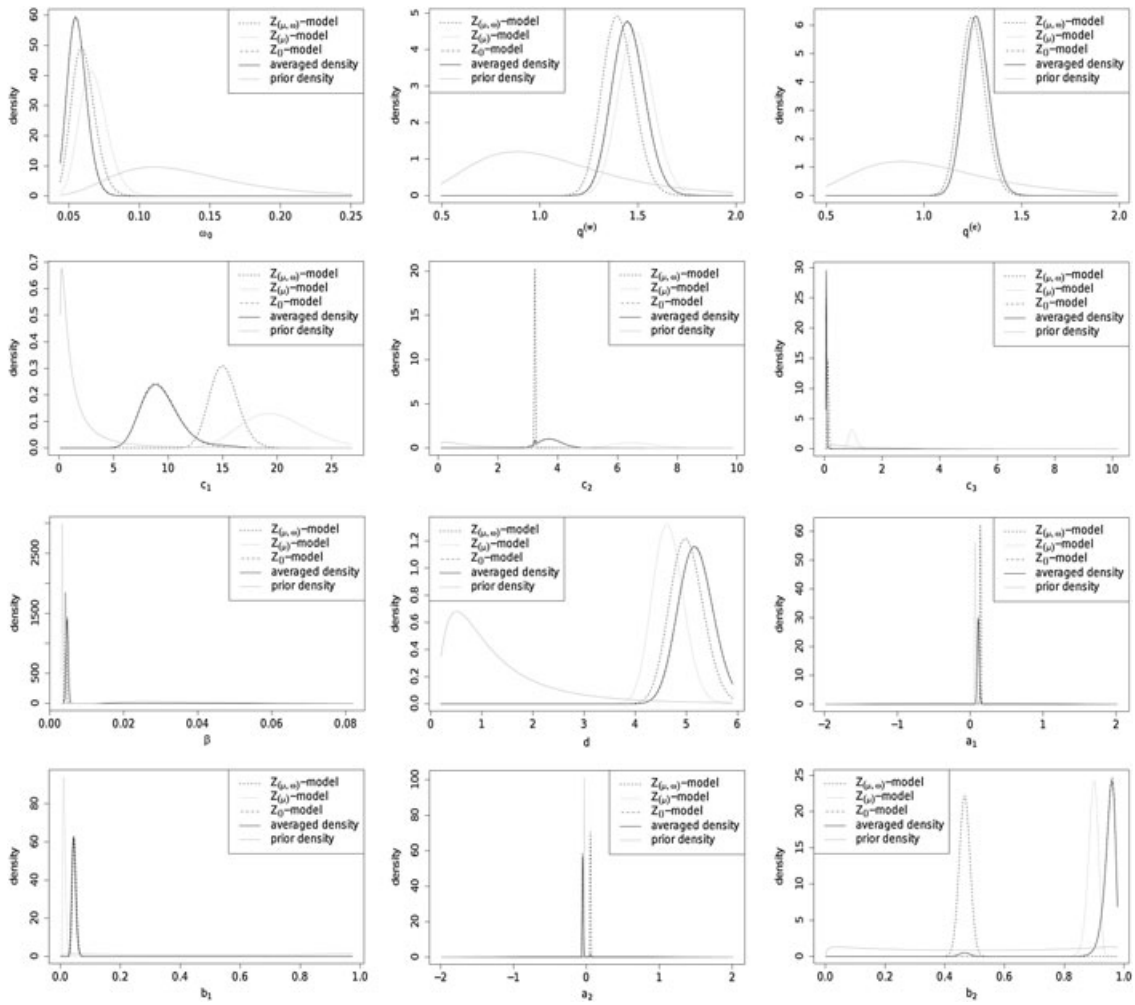

Figure A1. Averaged posterior densities of all parameters to be estimated coming from the models using contact pattern $\boldsymbol{C}_6$.

were chosen through the Akaike information criteria and may differ for each time series. From the coefficient estimates for the ARMA(p^*, q^*) process, we computed the autocorrelation function (ACF) $\rho(\tau)$ as described in [48]. We preferred this parametric ARMA(p, q)-based estimation over the empirical ACF estimates as the latter appeared to suffer from high variance.

Using this estimated ACF $\hat{\rho}$ for each time series $(D_t^{(j,i)})_{t=1,\ldots,T}$ ($j = 1, \ldots, 10$, $i = e, w$), we computed the estimated CA, $\widehat{\mathrm{CA}}^{(j,i)}(\vartheta^*)$, on the basis of the posterior mode ϑ^* computed in the first step according to definition (8).

A.3 Posterior distributions of all model parameters

Posterior distributions from vertical averaging using contact pattern $\boldsymbol{C}_6$ for all parameters to be estimated in each model are shown in Figure A1.

Acknowledgements

We are grateful to Bernhard Renard, Robert Koch Institute, for granting us access to the Linux cluster of the RKI bioinformatics group to perform the computer-intensive computations. We also thank the associate editor and two reviewers for their valuable comments and suggestions.

References

1. O'Neill PD. Introduction and snapshot review: relating infectious disease transmission models to data. *Statistics in Medicine* 2010; **29**(20):2069–2077.
2. Tate JE, Burton AH, Boschi-Pinto C, Steele AD, Duque J, Parashar UD. 2008 estimate of worldwide rotavirus-associated mortality in children younger than 5 years before the introduction of universal rotavirus vaccination programmes: a systematic review and meta-analysis. *The Lancet Infectious Diseases* 2012; **12**(2):136–141.
3. Pitzer VE, Viboud C, Simonsen L, Steiner C, Panozzo CA, Alonso WJ, Miller MA, Glass RI, Glasser JW, Parashar UD, Grenfell BT. Demographic variability, vaccination, and the spatiotemporal dynamics of rotavirus epidemics. *Science* 2009; **325**(5938):290–294. DOI: 10.1126/science.1172330.
4. Atkins KE, Shim E, Pitzer VE, Galvani AP. Impact of rotavirus vaccination on epidemiological dynamics in England and Wales. *Vaccine* 2012; **30**(3):552–564. DOI: 10.1016/j.vaccine.2011.11.064.
5. Atchison C, Lopman B, Edmunds WJ. Modelling the seasonality of rotavirus disease and the impact of vaccination in England and Wales. *Vaccine* 2010; **28**(18):3118–3126. DOI: 10.1016/j.vaccine.2010.02.060.
6. Van Effelterre T, M. Soriano-Gabarró M, Debrus S, Claire Newbern E, Gray J. A mathematical model of the indirect effects of rotavirus vaccination. *Epidemiology and Infection* 2010; **138**(6):884–897. DOI: 10.1017/S0950268809991245.
7. Shim E, Banks C, Castillo-Chavez HT. Seasonality of rotavirus infection with its vaccination. *AMS Contemporary Mathematics Book Series* 2006; **410**:327–348.
8. Freiesleben de Blasio B, Kasymbekova K, Flem E. Dynamic model of rotavirus transmission and the impact of rotavirus vaccination in Kyrgyzstan. *Vaccine* 2010; **28**(50):7923–7932. DOI: 10.1016/j.vaccine.2010.09.070.
9. Shim E, Castillo-Chavez C. The epidemiological impact of rotavirus vaccination in the United States and Mexico. *Mathematical and Statistical Estimation Approaches in Epidemiology* 2009:303–323.
10. Shim E, Galvani AP. Impact of transmission dynamics on the cost-effectiveness of rotavirus vaccination. *Vaccine* 2009; **27**(30):4025–4030. DOI: 10.1016/j.vaccine.2009.04.030.
11. Pitzer VE, Viboud C, Lopman BA, Patel MM, Parashar UD, Grenfell BT. Influence of birth rates and transmission rates on the global seasonality of rotavirus incidence. *Journal of the Royal Society Interface* 2011; **8**(64):1584–1593. DOI: 10.1098/rsif.2011.0062.
12. Brookhart MA, Hubbard AE, van der Laan MJ, Colford JM, Eisenberg JNS. Statistical estimation of parameters in a disease transmission model: analysis of a Cryptosporidium outbreak. *Statistics in Medicine* 2002; **21**(23):3627–3638.
13. Whitaker HJ, Farrington CP. Estimation of infectious disease parameters from serological survey data: the impact of regular epidemics. *Statistics in Medicine* 2004; **23**(15):2429–2443.
14. Elderd BD, Dukic VM, Dwyer G. Uncertainty in predictions of disease spread and public health responses to bioterrorism and emerging diseases. *PNAS* 2006; **103**(42):15693–15697. DOI: 10.1073/pnas.0600816103.
15. Birrell PJ, Ketsetzis G, Gay NJ, Cooper BS, Presanis AM, Harris RJ, Charlett A, Zhang XS, White PJ, Pebody RG, De Angelis D. Bayesian modeling to unmask and predict influenza A/H1N1pdm dynamics in London. *PNAS* 2011; **108**(45):18238–18243. DOI: 10.1073/pnas.1103002108.
16. Dorigatti I, Cauchemez S, Pugliese A, Ferguson NM. A new approach to characterising infectious disease transmission dynamics from sentinel surveillance: application to the Italian 2009–2010 A/H1N1 influenza pandemic. *Epidemics* 2012; **4**(1):9–21. DOI: 10.1016/j.epidem.2011.11.001.
17. Coelho FC, Codeço CT, Gomes MGM. A Bayesian framework for parameter estimation in dynamical models. *PLoS ONE* 05 2011; **6**(5):e19616.
18. Toni T, Welch D, Strelkowa N, Ipsen A, Stumpf MP. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society Interface* 2009; **6**(31):187–202. DOI: 10.1098/rsif.2008.0172.
19. He D, Ionides EL, King AA. Plug-and-play inference for disease dynamics: measles in large and small populations as a case study. *Journal of the Royal Society Interface* 2010; **7**(43):271–283. DOI: 10.1098/rsif.2009.0151.

20. Held L, Höhle M, Hofmann M. A statistical framework for the analysis of multivariate infectious disease surveillance counts. *Statistical Modelling* 2005; **5**(3):187–199. DOI: 10.1191/1471082X05st098oa.
21. Paul M, Held L, Toschke AM. Multivariate modelling of infectious disease surveillance data. *Statistics in Medicine* 2008; **27**(29):6250–6267.
22. Koch J, Wiese-Posselt M. Epidemiology of rotavirus infections in children <5 years of age: Germany, 2001–2008. *Pediatric Infectious Disease Journal* 2011; **30**(2):112–117.
23. Dudareva S, Koch J, An der Heiden M, Oberle D, Keller-Stanislawski B, Wichmann O. Impact of rotavirus vaccination in regions with low and moderate vaccine uptake in Germany. *Human Vaccines & Immunotherapeutics* 2012; **8**(10):1407–1415. DOI: 10.4161/hv.21593.
24. Rosner B, Stark K, Werber D. Epidemiology of reported *Yersinia enterocolitica* infections in Germany, 2001–2008. *BMC Public Health* 2010; **10**(1):337–344.
25. Federal Bureau of Statistics, GENESIS Online Database. https://www-genesis.destatis.de/genesis/online, Accessed: 2013-09-01.
26. Fischer TK, Valentiner-Branth P, Steinsland H, Perch M, Santos G, Aaby P, Mølbak K, Sommerfelt H. Protective immunity after natural rotavirus infection: a community cohort study of newborn children in Guinea-Bissau, West Africa. *The Journal of Infectious Disease* 2002; **186**(5):593–597. DOI: 10.1086/342294.
27. Moulton LH, Staat MA, Santosham M, Ward RL. The protective effectiveness of natural rotavirus infection in an American Indian population. *The Journal of Infectious Disease* 1998; **178**(6):1562–1566. DOI: 10.1086/314504.
28. Azzalini A. A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics* 1985; **12**:171–178.
29. Anderson EJ, Weber SG. Rotavirus infection in adults. *The Lancet Infectious Diseases* 2004; **4**(2):91–99. DOI: 10.1016/S1473-3099(04)00928-4.
30. Pickering LK, Bartlett AV III, Reves RR, Morrow A. Asymptomatic excretion of rotavirus before and after rotavirus diarrhea in children in day care centers. *The Journal of Pediatrics* 1988; **112**(3):361–365. DOI: 10.1016/S0022-3476(88)80313-5.
31. Wilde J, Yolken R, Willoughby R, Eiden J. Improved detection of rotavirus shedding by polymerase chain reaction. *The Lancet* 1991; **337**(8737):323–326. DOI: 10.1016/0140-6736(91)90945-L.
32. Clark HF, Marcello A, Lawley D, Reilly M, DiNubile M. Unexpectedly high burden of rotavirus gastroenteritis in very young infants. *BMC Pediatrics* 2010; **10**(1):40–46.
33. Clemens J, Rao M, Ahmed F, Ward R, Huda S, Chakraborty J, Yunus M, Khan MR, Ali M, Kay B, van Loon F, Sack D. Breast-feeding and the risk of life-threatening rotavirus diarrhea: prevention or postponement? *Pediatrics* 1993; **92**(5):680–685.
34. Pickering LK, Morrow AL, Herrera I, O'Ryan M, Estes MK, Guilliams SE, Jackson L, Carter-Campbell S, Matson DO. Effect of maternal rotavirus immunization on milk and serum antibody titers. *Journal of Infectious Diseases* 1995; **172**(3):723–728. DOI: 10.1093/infdis/172.3.723.
35. Ward RL, Bernstein DI, Shukla R, McNeal MM, Sherwood JR, Young EC, Schiff GM. Protection of adults rechallenged with a human rotavirus. *Journal of Infectious Diseases* 1990; **161**(3):440–445. DOI: 10.1093/infdis/161.3.440.
36. Soriano-Gabarró M, Mrukowicz J, Vesikari T, Verstraeten T. Burden of rotavirus disease in european union countries. *The Pediatric Infectious Disease Journal* 2006; **25**(1):7–11.
37. F. Raúl Velásquez DOM, Ruiz-Palacios GM. Rotavirus infection in infants as protection against subsequent infections. *Vaccine* 1996; **335**(14):1022–1028.
38. Mrukowicz JZ, Thompson J, Reed GW, Tollefson SJ, Kobayashi M, Araki K, Wright PF. Epidemiology of rotavirus in infants and protection against symptomatic illness afforded by primary infection and vaccination. *Vaccine* 1999; **17**(7–8):745–753.
39. Bernardo JM, Bayarri MJ, Berger JO, Dawid AP, Heckerman D, Smith AFM, West M. *Bayesian Statistics 8*. Oxford University Press: Oxford, 2007.
40. Press WH, Teukolsky SA, Vetterling WT, Flannery BP. *Numerical Recipes 3rd Edition*. Cambridge University Press: New York, 2007.
41. Thiébaux HJ, Zwiers FW. The interpretation and estimation of effective sample size. *Journal of Applied Meteorology* May 1984; **23**:800–811. DOI: 10.1175/1520-0450(1984)0232.0.CO;2.
42. McMillan H, Freer J, Pappenberger F, Krueger T, Clark M. Impacts of uncertain river flow data on rainfall–runoff model calibration and discharge predictions. *Hydrological Processes* 2010; **24**(10):1270–1284.
43. Raftery AE. Approximate Bayes factors and accounting for model uncertainty in generalised linear models. *Biometrika* 1996; **83**(2):251–266. DOI: 10.1093/biomet/83.2.251.
44. Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, Massari M, Salmaso S, Tomba GS, Wallinga J, Heijne J, Sadkowska-Todys M, Rosinska M, Edmunds WJ. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Medicine* 2008; **5**:e74.
45. Friel N, Pettitt AN. Marginal likelihood estimation via power posteriors. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 2008; **70**(3):589–607.
46. Hardin JW, Hilbe JM. *Generalized Linear Models and Extensions*. Stata Press: College Station, Texas, 2007.
47. Abramowitz M, Stegun I. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover Publications: New York, 1964.
48. Brockwell PJ, Davis RA. *Time Series: Theory and Methods*. Springer: New York, 1991.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site.