scientific reports

OPEN

Estimating the course of the COVID-19 pandemic in Germany via spline-based hierarchical modelling of death counts

Tobias Wistuba, Andreas Mayr & Christian Staerk✉

We consider a retrospective modelling approach for estimating effective reproduction numbers based on death counts during the first year of the COVID-19 pandemic in Germany. The proposed Bayesian hierarchical model incorporates splines to estimate reproduction numbers flexibly over time while adjusting for varying effective infection fatality rates. The approach also provides estimates of dark figures regarding undetected infections. Results for Germany illustrate that our estimates based on death counts are often similar to classical estimates based on confirmed cases; however, considering death counts allows to disentangle effects of adapted testing policies from transmission dynamics. In particular, during the second wave of infections, classical estimates suggest a flattening infection curve following the "lockdown light" in November 2020, while our results indicate that infections continued to rise until the "second lockdown" in December 2020. This observation is associated with more stringent testing criteria introduced concurrently with the "lockdown light", which is reflected in subsequently increasing dark figures of infections estimated by our model. In light of progressive vaccinations, shifting the focus from modelling confirmed cases to reported deaths with the possibility to incorporate effective infection fatality rates might be of increasing relevance for the future surveillance of the pandemic.

The COVID-19 pandemic continues to have severe impacts on public health in many parts of the world. During the course of the pandemic, different non-pharmaceutical interventions (NPIs) have been implemented to mitigate the spread of the virus, including closures of schools and nurseries, cancellation of public events, regulations regarding social distancing, closures of non-essential shops and further measures[1]. Since many of these interventions impose a large burden on society and economy, it is crucial to understand which measures are effective in reducing the spread of the virus.

An important figure in the context of the pandemic is the basic reproduction number R_0, which is defined as the expected number of secondary infected individuals by an infectious individual in a completely susceptible population. The basic reproduction number R_0 is a population-specific measure of the contagiousness of the virus, which can be expressed as the product of the duration of infectiousness, the probability of infection for a contact between infected and susceptible individuals and the average rate of contacts between infected and susceptible individuals[2]. A systematic review and meta-analysis estimated $\hat{R}_0 = 3.3$ (95% confidence interval: 2.8 to 3.8) for the initial spread of SARS-CoV-2 in China, 2020[3], while Ahmad et al.[4] estimated $\hat{R}_0 = 2.3$ for Pakistan and Locatelli et al.[5] estimated $\hat{R}_0 = 2.2$ (95% CI 1.9 to 2.6) for Western Europe (all estimates based on early variants of SARS-CoV-2 in 2020).

In contrast to the basic reproduction number R_0, the effective reproduction number R_t is the time-dependent counterpart, defined as the average number of secondary infections resulting from an infectious individual in a specific region at a certain time t. In particular, the effective reproduction number R_t accounts for changes in the rate of contacts over time, which can be affected by multiple factors including the number of susceptible individuals, the implemented mitigation measures and the adaptive behaviour of the population. Therefore, the effective reproduction number R_t will typically be lower than the basic reproduction number R_0. In practice, the

Department of Medical Biometry, Informatics and Epidemiology, University Hospital Bonn, Bonn, Germany. ✉email: staerk@imbie.uni-bonn.de
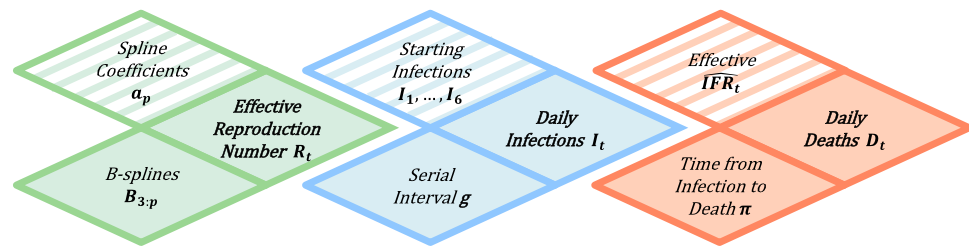
Figure 1. Schematic overview of the adapted Bayesian hierarchical model (cf. Flaxman et al.[11]).

effective reproduction number R_t is usually estimated based on currently available surveillance data, particularly based on daily numbers of newly confirmed cases[6,7]. While numbers of confirmed cases are crucial for nowcasting the current development of the effective reproduction number[8,9], they are largely influenced by the implemented testing policies. In particular, short-term changes in numbers of conducted SARS-CoV-2 tests complicate the accurate estimation of the effective reproduction number, as such changes can lead to time-varying numbers of undetected infections (so-called "dark figures" of infections[10]) in relation to numbers of actually confirmed cases.

Data on COVID-19 related deaths can provide an additional, retrospective viewpoint on the course of the pandemic and the assessment of NPIs. Important and influential approaches hence also focused on modelling the spread of SARS-CoV-2 based on numbers of reported deaths[11,12]. In particular, the Imperial College COVID-19 Response Team[11] developed a Bayesian hierarchical model to estimate the impact of NPIs on the effective reproduction number in different European countries during the first wave of infections in spring 2020. In the further course of the pandemic, several interventions (e.g., closures of schools and non-essential shops) have been adapted, relaxed or restricted to particular regions, while others (e.g., cancellations of public events and face mask regulations) have largely kept in place. To account for this development, the Bayesian model of Flaxman et al.[11] has been updated and extended to estimate the effectiveness of NPIs during the second infection wave in Europe[13].

Other authors have argued that resulting effect estimates of NPIs are non-robust and highly model-dependent[14,15]. In particular, the selection of NPIs to be included in the model predetermine the potential change points for the effective reproduction number and thus can have large effects on the estimates attributed to individual prevention measures. Furthermore, not only the implemented NPIs change over time, but also the adherence and awareness of the population, which may not be adequately described by categorical variables for the implemented prevention measures (cf. studies[16,17] of mobility patterns during the first phase of the pandemic in Germany). Another limitation of the original model of Flaxman et al.[11] is that the infection fatality rate (IFR) is assumed to be constant over time. However, the IFR of COVID-19 increases largely with increasing age[18,19]. As the age distribution of infections changes substantially during the course of the pandemic[20], the effective IFR should not be regarded as constant over time when modelling the number of infections based on deaths data.

In this study we adapt the Bayesian hierarchical model of Flaxman et al.[11] to estimate the course of the effective reproduction number in Germany by continuous smoothing splines, without the need for additional information regarding the timings of specific interventions. While our model is primarily driven by the numbers of reported deaths similar as in previous modelling approaches[11,15], a main contribution of this study is the additional incorporation of the changing age distribution of confirmed infections to account for changes in effective IFR. We compare our model estimates of the effective reproduction number with classical estimates derived solely from the numbers of confirmed cases in combination with nowcasting[6,21], showing that estimates of our model based on death counts tend to be more robust to changes in testing policies. Furthermore, by considering the total number of estimated infections per confirmed case (IPCC) as a varying factor, we are able to discuss dark figures of infections over the course of the pandemic in Germany.

The paper is structured as follows: the proposed Bayesian hierarchical model is described in the "Methods" section, while a more detailed comparison with the model of Flaxman et al.[11] is provided in the Supplement to this paper (see Section S.1). Model estimates for the first year of the pandemic in Germany are presented in the "Results" section; complimentary results for the individual 16 German federal states are provided in the Supplement (see Section S.4). The paper concludes with a discussion of the merits and limitations of the proposed hierarchical model in the context of related approaches.

Methods

To estimate the course of the pandemic in Germany based on death counts, we extend the Bayesian hierarchical model from Flaxman et al.[11] by adjusting for time-dependent effective infection fatality rates and by considering splines for modelling the effective reproduction number over time. Figure 1 provides a schematic representation of the adapted hierarchical model. The main idea of the model is to estimate the effective reproduction number R_t retrospectively from daily numbers of reported deaths D_t associated with COVID-19. For this, we use death counts for Germany (and also for the 16 German federal states, see Section S.4 of the Supplement) based on daily situation reports published by the German federal agency for disease control and prevention (Robert Koch Institute, RKI[22]). All methods were carried out in accordance with relevant guidelines and regulations.

On the last level of the hierarchical model (see Fig. 1), the reported deaths D_t are modelled with a negative binomial distribution

$$D_t \sim \text{NB}\Big(d_t,\, d_t + \frac{d_t^2}{\psi}\Big), \tag{1}$$

with mean d_t and variance $d_t + \frac{d_t^2}{\psi}$. A half-normal distribution is considered as the prior for the dispersion parameter $\psi \sim \mathcal{N}^+(0, 5)$. Numbers of daily reported deaths D_t are linked to expected numbers of daily infections $I_\tau, \tau < t$, by taking into account estimates for the effective infection fatality rate (IFR) and the time distribution between infection and reported death. More specifically, expected numbers of daily reported deaths d_t are given by

$$d_t = \sum_{\tau=1}^{t-1} \pi_{t-\tau} \cdot \widehat{\text{IFR}}_\tau \cdot \nu \cdot I_\tau, \tag{2}$$

where π denotes the (discretized) distribution for the time between infection and reported death for lethal infections, $\widehat{\text{IFR}}_\tau$ is the estimated effective infection fatality rate and I_τ denotes the expected number of infections for day τ. Additionally, the parameter ν reflects the uncertainty regarding the estimated $\widehat{\text{IFR}}_\tau$; similarly to the model of Flaxman et al.[11], a normal distribution with mean 1 and standard deviation 0.1 is considered as the prior for ν. The distribution π for the time between infection and reported death is obtained as the sum of two components: the incubation period and the time between symptom onset and reported death. Based on results from Lauer et al.[23], for the incubation period we use a log-normal distribution with mean 5.52 (days) and standard deviation 2.43 (days). The distribution for the time between symptom onset and reported death is adopted from Flaxman et al.[11] and given by a gamma distribution with mean 17.8 (days) and standard deviation 8.01 (days). Since our model is based on daily data, we consider a discretized version π of the distribution for the sum of the two periods (see Section S.1 of the Supplement for details on the discretization).

The IFR is an important link when trying to infer total numbers of infections from death counts. As the IFR of COVID-19 largely depends on the age of the infected individuals[18,19,24], it is important to take the age structure of infections into account, which has been shown to vary over the course of the pandemic in Germany[20]. Thus, in contrast to the original model of Flaxman et al.[11] using time-constant averaged IFRs, we consider time-dependent effective IFRs which reflect the changing age distribution of infections. Here, we estimate weekly effective IFRs based on the assumption that the age distributions of true infections can be approximated by the age distributions of confirmed cases[20]. In particular, based on data from the RKI[25], let $C_{a,w}$ denote the number of confirmed cases in calendar week w for age-group $a \in \{0\text{–}9, 10\text{–}19, ..., 70\text{–}79, 80+\}$. Furthermore, let $\widehat{\text{IFR}}_a$ denote the estimated infection fatality rate for age group a based on Brazeau et al.[24] (see Section S.3 of the Supplement for sensitivity analyses using alternative age-specific IFR estimates from O'Driscoll et al.[18] and Levin et al.[19]). The effective IFR for day τ in calendar week w is estimated as a weighted average of age-specific IFR estimates:

$$\widehat{\text{IFR}}_\tau = \frac{1}{C_w} \sum_{a \in A} C_{a,w} \cdot \widehat{\text{IFR}}_a. \tag{3}$$

Finally, IFR estimates are shifted backwards by ten days to adjust for the delay between the reporting date of cases C_t and the date of infections I_t. This delay can be expressed as the sum of the incubation time with an expected value of 5.52 days[23] and the time between onset of symptoms and reporting date (median delay of four days for German data[26]).

On the previous level of the hierarchical model (see Fig. 1), expected numbers of daily infections I_t are modelled based on estimated infections I_τ of the preceding days $\tau < t$ via

$$I_t = R_t \cdot \sum_{\tau=1}^{t-1} I_\tau \cdot g_{t-\tau}, \tag{4}$$

where R_t denotes the effective reproduction number and g the (discretized) generation time distribution (i.e. the time between two infections in a transmission pair). As infection times are generally unknown, direct estimation of the generation time is rather difficult and commonly approximated by the serial interval. Based on results from Nishiura et al.[27], the generation time g is modelled using a discretized log-normal distribution with mean 4.7 (days) and standard deviation 2.9 (days). An important and idealistic assumption of the considered semi-mechanistic model is that the country is viewed as a closed environment[11], so that all infections are assumed to occur within the German population. Consequently, numbers of infections for the first days have to be initialized. Similarly to Flaxman et al.[11], for the first six modelling days $t = 1, \ldots, 6$, expected numbers of infections $I_t \sim \text{Exp}(1/\lambda)$ are exponentially distributed with mean $\lambda \sim \text{Exp}(10)$. The starting date for modelling ($t = 1$) is considered to be the 15th of January 2020, which is defined as the earliest date so that 60 days later, at least 10 cumulative deaths associated with COVID-19 have been reported in Germany.

On the first level of the hierarchical model (see Fig. 1), the effective reproduction number is modelled with a smoothing spline via

$$R_t = \max\Big(\sum_p a_p B_{p,3}(t),\, 0\Big), \tag{5}$$

where $B_{p,3}$ are B-splines of third degree between equidistant knots (each with distance of two weeks) and a_p denote the corresponding spline coefficients[28]. The maximum in Eq. (5) is taken to ensure that the effective reproduction number is non-negative. For a smoothing effect, the priors for the spline coefficients $a_p, p > 1$, are

considered to be normal distributions $a_p \sim \mathcal{N}(a_{p-1}, \theta)$ with the previous spline coefficients a_{p-1} as expected values and common variance $\theta \sim \mathcal{N}^+(0, 1)$, while the prior for the first spline coefficient $a_1 \sim \mathcal{N}^+(0, 1)$ is considered to be a half-normal distribution.

By modelling the true infections, our approach also allows for the analysis of dark figures of undetected infections during the pandemic. The total number of infections $I = C + U$ is composed of the number of confirmed (detected) cases C and the number of undetected infections U. The total number of infections per confirmed case (IPCC) is given by the ratio I/C, which we estimate as a time-varying factor of dark figures. Since there is an average delay of ten days between infections and reporting dates (see Methods), the IPCC factor for day t is estimated as the ratio of infections at day $t - 10$ estimated by our model and the 7-day-mean of confirmed cases at day t.

While the general hierarchical structure of our model is the same as in Flaxman et al.[11], there are two main differences: First, the original model of Flaxman et al. is based on the assumption of time-constant IFRs (for different countries), whereas our model incorporates time-varying estimates of the effective IFR (cf. Fig. S1 of the Supplement). Second, the model of Flaxman et al. yields piecewise constant estimates for the effective reproduction number which can only change at prespecified time points where non-pharmaceutical interventions were adapted; on the other hand, our model based on splines provides a smooth and data-driven way of estimating the effective reproduction number over time, without requiring to prespecify potential change points. A complete and compact formulation of the proposed Bayesian model including all assumed prior distributions is presented in Section S.1 of the Supplement, comparing it also to the original model of Flaxman et al.[11].

The Bayesian hierarchical model is implemented in R[29] via the add-on package `rstan` for Stan[30]. The implementation of smoothing splines is based on Kharratzadeh[28]. Posterior samples are obtained by the No-U-Turn Sampler (NUTS) for adaptive Hamiltonian Monte Carlo, which takes multiple steps based on first-order gradients and automatically tunes the number of steps L as well as the step size ϵ without requiring any manual tuning by the data analyst[31]. Compared to classical random-walk Metropolis or Gibbs samplers, NUTS is particularly less prone to slow mixing in cases of highly correlated parameters induced by hierarchical model structures[32]. For algorithmic details on NUTS we refer to Hoffman and Gelman[31]. Using eight independent NUTS chains with 2000 iterations each (considering burn-in periods of 1000 iterations), convergence diagnostics indicate that the algorithm provides a representative sample from the posterior distribution of our model (see Section S.5 of the Supplement).
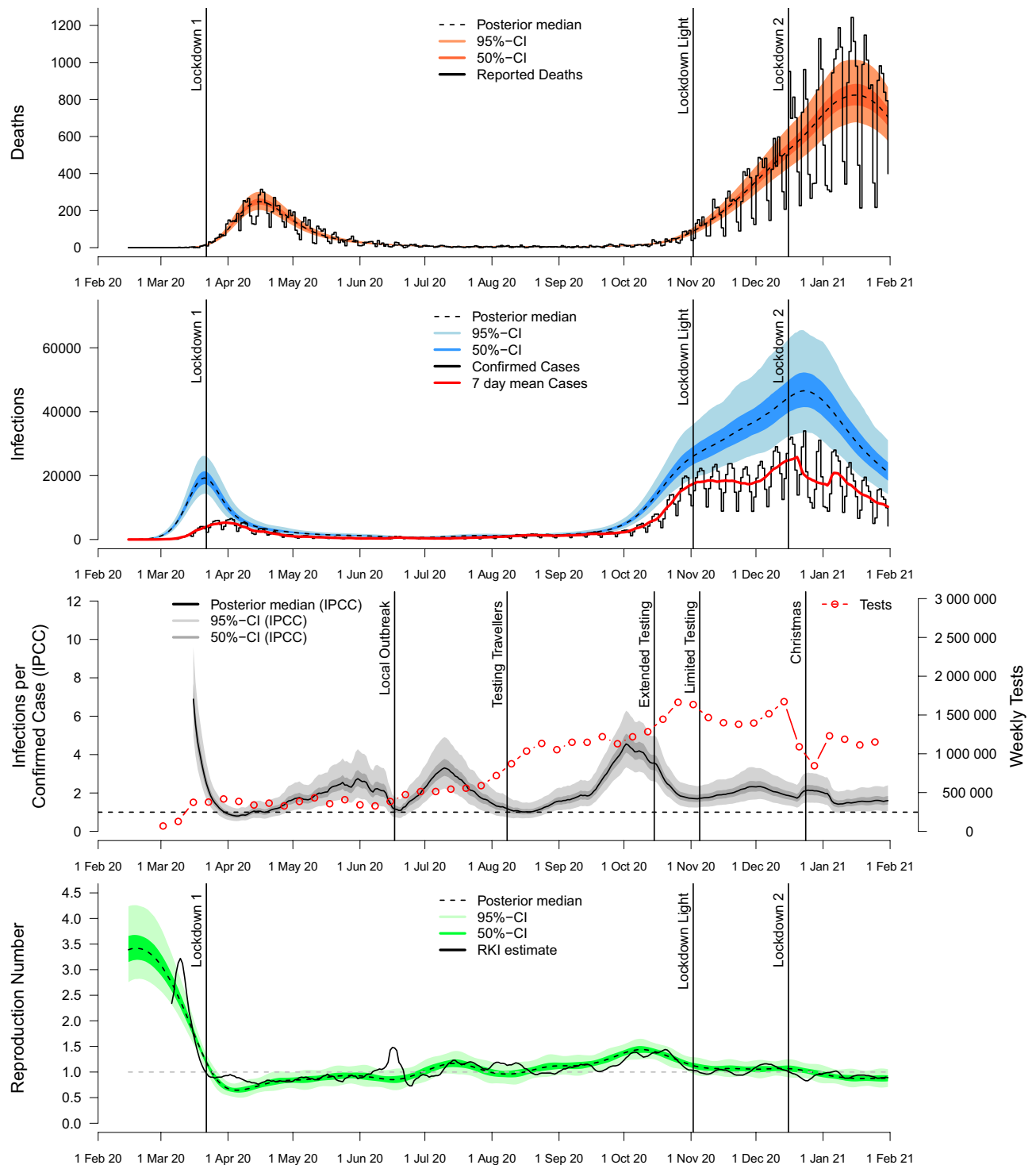
Results

Using German surveillance data[22,25], we model the course of the pandemic for Germany as a whole country as well as separately for each of the 16 German federal states during the first year of the pandemic. Here we present detailed national results, while further results for the individual federal states can be found in Section S.4 of the Supplement.

Results of the Bayesian hierarchical model (cf. Fig. 1) for Germany are depicted in Fig. 2. During the first year of the pandemic, death counts show two pronounced waves, with a shorter first wave in spring 2020 with fewer deaths compared to the second wave in autumn and winter 2020/2021. Figure 2 (first graph) indicates that the hierarchical model yields a good fit to the course of reported numbers of deaths. Note that daily reported deaths and confirmed cases show a characteristic weekly oscillating pattern. Estimates of our model, however, capture the average tendency and are robust to such reporting artefacts due to the use of smoothing splines with biweekly equidistant knots for modelling the effective reproduction number.

Figure 2 additionally illustrates the introduction of major non-pharmaceutical interventions (NPIs) in Germany. Note that, due to the federal structure of Germany, there have been local differences regarding the implementation of NPIs (not shown here). Around the 23rd of March 2020, the "first lockdown" (Lockdown 1) was introduced in Germany, including strict contact restrictions and closures of non-essential shops. At this point, schools and nurseries had already been closed for one week and public events had been banned for two weeks in most parts of the country. Approximately three to four weeks after the introduction of the "first lockdown", the first wave reached its maximum regarding the number of reported deaths. During summer 2020, cases and death counts were at relatively low levels. After a rapid increase of cases in October 2020, the German government introduced the so-called "lockdown light" to mitigate the spread of the virus. During this period, restaurants and leisure facilities were closed, while schools and shops remained opened with hygiene concepts in place. However, no noticeable decline of reported deaths was observed and confirmed cases remained at relatively high levels, which led to the introduction of the "second lockdown" in December 2020 (Lockdown 2), including closures of schools and non-essential shops amongst further measures. Three to four weeks after the introduction of the "second lockdown", a decline of daily deaths was reported similarly as during the first wave in spring 2020. Note that, as expected, the general course of confirmed cases precedes the course of reported deaths.

Instead of modelling directly the numbers of confirmed cases, which are largely influenced by the adopted testing regime, the Bayesian hierarchical model provides estimates of the course of true infections. Estimated infections in turn precede the course of confirmed cases due to the incubation period and reporting delays (see second graph of Fig. 2). However, results show a remarkable difference between the course of estimated infections and confirmed cases during the second wave of infections in autumn/winter 2020/21: following the "lockdown light" introduced at the 2nd of November 2020, the course of confirmed cases suggests a flattening or even decreasing trend, while results of our hierarchical model indicate that true numbers of estimated infections continued to rise and reached a maximum around one week after the "second lockdown" in December 2020.

Such effects should also be viewed in light of changing testing policies, which can result in varying dark figures of undetected infections. During the first year of the pandemic (i.e. between the 1st of February 2020 and the 31st of January 2021), there were in total 2, 221, 838 confirmed SARS-CoV-2 cases in Germany; for

Figure 2. Results for the course of the COVID-19 pandemic in Germany via spline-based hierarchical modelling of death counts (cf. Fig. 1) using age-specific IFR estimates from Brazeau et al.[24]. Our model provides estimates of total numbers of infections over time (second graph) and of infections per confirmed case (IPCC) as a time-varying factor for dark figures of infections (third graph). By disentangling effects of changes in testing from transmission dynamics, estimates of our model for the effective reproduction number based on death counts tend to be more robust compared to RKI estimates based on confirmed cases (fourth graph). Model estimates are based on posterior medians, together with 50% and 95% credible intervals (CIs).

the same time period, the model estimates 4.449 million (95% credible interval: [3.245m; 6.177m]) infections based on age-specific IFR estimates from Brazeau et al.[24], yielding an overall estimated infections per confirmed

case (IPCC) factor of 2.002 [1.460; 2.780]. Using alternative age-specific IFR assumptions based on O'Driscoll et al.[18] and Levin et al.[19] (cf. Fig. S1 of the Supplement), the model estimates 5.577m [3.966m; 7.931m] and 3.008m [2.155m; 4.275m] infections for the first year of the pandemic, resulting in overall estimated IPCC factors of 2.510 [1.785; 3.569] and 1.354 [0.970; 1.924], respectively. Although absolute numbers of estimated infections and levels of dark figures vary considerably for the different age-specific IFRs, general temporal trends of estimated infections and dark figures remain quite stable, while effective reproduction numbers estimated by our model are very robust regarding different age-specific IFR assumptions (see Section S.3 of the Supplement for detailed results of sensitivity analyses).

The third graph of Fig. 2 depicts the estimated course of the infections per confirmed case (IPCC) as a factor for dark figures of undetected infections. Results show that estimated dark figures are much larger before the "first lockdown" in spring 2020 than in the following course, reflecting limited testing capacities at the beginning of the pandemic. It can be observed that changes in estimated dark figures are often associated with shifts in testing policies and practice (here, vertical lines in the third graph of Fig. 2 indicate dates of important shifts). In particular, estimated dark figures declined sharply in June 2020 in the context of a local super-spreading event in a slaughterhouse in North Rhine-Westphalia, resulting in temporarily increased targeted testing of factory employees. During August 2020, when incidence rates were low and targeted testing of travellers returning from summer holidays was intensified, the IPCC factor is estimated to be close to one, indicating that most infections had been identified by testing in that period. Following increasing dark figures of infections in September and October 2020, on the 15th of October the Robert Koch Institute (RKI) changed its recommendations towards less stringent indications for SARS-CoV-2 tests, leading to a further increase in numbers of conducted tests and a considerable decrease in estimated dark figures of infections. However, with increasing incidences towards the end of October, on the 5th of November the recommendations were again updated towards more restricted testing, which again results in increasing dark figures estimated by our model. Finally, temporarily increased dark figures are estimated following Christmas, reflecting the decrease of conducted tests during this holiday period. Overall, our results indicate that the hierarchical modelling approach can reliably detect shifts in testing policies, even though model estimates of dark figures are solely based on numbers of confirmed cases and reported deaths, without incorporating any direct information regarding the numbers of conducted tests.

Finally, the fourth graph of Fig. 2 shows the resulting course of the effective reproduction number estimated by our hierarchical model based on death counts, in comparison with official estimates from the Robert Koch Institute (RKI), which are based on the evolution of confirmed cases[6,21]. It can be observed that estimates based on death counts are often similar to classical estimates based on confirmed cases. However, model estimates based on death counts prove to be more robust against shifts in testing policies. In particular, confirmed cases indicate a short-term spike in the effective reproduction number linked to a local super-spreading event in June 2020, whereas our model does not estimate a spike during this period; instead, it estimates reduced dark figures of infections, suggesting that the spike in the effective reproduction number was mainly related to targeted testing of contact persons. Furthermore, after the "lockdown light" at the 2nd of November 2020, classical estimates of the effective reproduction number tend to be smaller than model estimates based on death counts. Although the differences do not seem large, they imply considerably different interpretations regarding the course of the pandemic: while classical estimates (with values smaller or equal to one) suggest a flattening or even decreasing trend of infections following the "lockdown light", estimates of our model (with values larger or equal to one) suggest that true numbers of infections continued to rise (cf. second graph of Fig. 2). Note that the "lockdown light" was introduced more or less at the same time when the RKI recommendations for testing were adapted (see third graph of Fig. 2), indicating that our model is able to disentangle overlying effects of reduced testing (resulting in larger dark figures) and the adaptation of NPIs on numbers of confirmed cases.

Discussion

We have extended and adapted the Bayesian hierarchical model of Flaxman et al.[11] for modelling the course of the COVID-19 pandemic in Germany based on death counts. A main feature of the proposed approach is the smooth and data-driven way of estimating the effective reproduction number. As a result, there is no need to prespecify discrete change points for timings of non-pharmaceutical interventions (NPIs) as in the original model of Flaxman et al.[11], diminishing the chance that potential implicit biases are incorporated into the model[14,15]. While our approach shows parallels with the Bayesian model developed in Wood[15], which uses smoothing splines to estimate effective reproduction numbers in the United Kingdom, an important additional characteristic of our work is the adjustment for time-varying effective infection fatality rates (IFRs), which are estimated to change substantially over the course of the pandemic as a result of varying age distributions of infections[20].

Results for German surveillance data illustrate that the proposed retrospective model can provide additional valuable insights regarding the course of effective reproduction numbers and dark figures of true infections. While estimated reproduction numbers of our model are often similar to classical estimates from the Robert Koch Institute (RKI) based on confirmed cases[6,21], the proposed modelling approach based on death counts proves to be more robust against shifts in testing policies. In contrast to classical estimation methods relying solely on confirmed cases, our approach has the potential to disentangle overlying effects of shifts in testing policies and actual changes in the effective reproduction number, as illustrated for the second wave of infections in Germany in November 2020, where the "lockdown light" was introduced concurrently with the adaptation of testing recommendations. Further results presented in the Supplement illustrate that our Bayesian modelling approach yields robust estimates for the different developments of the pandemic in the 16 individual German federal states (using the same model and priors as for the full country). Here one should note that death counts were at relatively low levels during summer 2020, so that estimating the effective reproduction number and dark figures of infections for individual federal states comes with larger uncertainty, which is also reflected in the

wider credible intervals of model estimates (particularly for federal states with smaller populations, see e.g. the results for Mecklenburg-Western Pomerania and Saarland in Figs. S11 and S15).

An important methodological conclusion of this study is that the assumption of a constant effective IFR as in the original model of Flaxman et al.[11] from the beginning of the pandemic is not suitable for modelling the full first year of the pandemic in Germany. In particular during summer 2020 with relatively low incidences in Germany, there was a considerably younger age distribution of infections compared to periods with higher incidences in spring 2020 and winter 2020/2021, implying lower effective IFRs in the summer period (see Fig. S1 of the Supplement). Assuming a constant effective IFR, the hierarchical model would, for example, estimate unreasonable small numbers of infections during summer 2020 (even lower than numbers of confirmed cases). This observation illustrates that the effective IFR is a key parameter of the hierarchical model to retrospectively infer numbers of true infections from reported death counts. Only by combining the adjustment for time-dependent effective IFRs with the smooth modelling via splines, our model estimates of effective reproduction numbers largely resembled the official RKI estimates based on confirmed cases[21]. At the same time, by modelling death counts instead of inferring only from confirmed cases, our hierarchical model is more robust to changes in testing policies and provides valuable information regarding dark figures of undetected infections over time.

Our study is also related to the recent work of Schneble et al.[10], which estimates relative changes in the case detection ratio (CDR) over time for different age groups (the CDR is the reciprocal of infections per confirmed case, IPCC). The authors employ a smooth generalized linear mixed model for confirmed cases and variables indicating whether the infections resulted in COVID-19 related deaths (without modelling the actual dates of deaths). While this classical mixed modelling approach provides age group specific estimates of *relative* changes in the CDR, our proposed Bayesian hierarchical modelling approach also yields estimates of *absolute* numbers of dark figures of infections as well as estimates of the effective reproduction number, by considering age-specific IFR estimates and dates of reported deaths. Our numerical results regarding trends in dark figures of infections generally support the results of Schneble et al.[10]: dark figures in Germany are estimated to be largest at the beginning of the pandemic and, after a period of relatively low estimated dark figures (i.e. large CDR) during summer 2020, numbers of undetected cases are estimated to increase sharply in September 2020.

An obvious limitation of our modelling approach relying on death counts is that it can only reflect the course of the pandemic in retrospect. While this is partly true for all modelling approaches, including the traditional ones based on confirmed cases (due to the incubation period and reporting delays), one clearly has to acknowledge that a timely reporting and analysis is essential for estimating the effects of political decisions (e.g., lockdown measures or other NPIs). In this context, our model based on death counts will always yield results several weeks later than day-by-day estimates relying on reported cases. In particular, our retrospective model is not designed for "nowcasting" the current development of the pandemic in real time[8,9,33]. A limitation of the presented analysis of German surveillance data is that dates of daily reported deaths may deviate from actual dates of deaths[9]. Similar to the Bayesian model of Flaxman et al.[11], our approach relies on parametric assumptions, particularly regarding the distribution between infections and reported deaths. While a comparison of our results with official RKI estimates indicates that the considered distribution based on previous studies[11,23] is appropriate for German data, the specific parametric assumptions may not be generally transferable to other countries with different reporting characteristics.

The Bayesian hierarchical modelling approach also relies on various other assumptions[11], among them statistical ones including the implemented prior distributions for model parameters (see "Methods" section and Section S.1 of the Supplement). From a more practical perspective, the model also relies on the assumption of a closed environment (no new infections imported from outside of the population) and on the assumption that cases are insusceptible for another (second) infection with COVID-19. For the estimation of the effective infection fatality rate (IFR) we assumed that the evolving age distribution of infections can be approximated by the corresponding age distribution of confirmed cases[20]; furthermore, it relies on the assumption that age-specific IFR estimates from Brazeau et al.[24] are applicable to Germany (see Section S.3 of the Supplement for sensitivity analyses with alternative IFR estimates). In the current modelling approach we do not account for vaccinations, which does not pose an important limitation for the considered time period with low total numbers of administered vaccinations until the end of January 2021.

In this context, we have reason to believe that our hierarchical approach is particularly flexible regarding extensions for current and future challenges of COVID-19 modelling, as it shifts the focus from modelling confirmed cases towards reported deaths. Since we have included time-varying IFRs to account for changing age distributions of infections, future research could make use of this flexibility and incorporate also time-varying vaccination effects into our model, as well as potentially altered intrinsic severity of emerging SARS-CoV-2 variants[34]. Adjustment for vaccinations could be achieved via a general population-wise factor or age group specific parameters representing the rates of vaccinated individuals in the respective groups (see also related studies[35–37] for different modelling approaches of vaccination effects). Particularly in light of progressive vaccination programs in many countries, it can be expected that there will be additional changes in implemented testing regimes during the further course of the pandemic. For example, in the second year of the pandemic in Germany, SARS-CoV-2 rapid antigen tests were generally provided for free except for the short time period between 11th of October 2021 and 12th of November 2021, where free tests were restricted to those which fulfil certain eligibility criteria, including children and pregnant women. While this study focused on the first year of the pandemic in Germany, our modelling approach based on death counts is expected to be more robust to such sharp changes in testing policies compared to approaches based only on numbers of confirmed cases. In particular, Contreras et al.[38] have illustrated that limited test-trace-and-isolate (TTI) capacities can lead to a "metastable regime with the risk of sudden explosive growth" in (undetected) infections. Furthermore, vaccinated individuals may generally be less likely to be tested due to asymptomatic or mild symptomatic infection[39], which may induce larger dark figures of infections in the vaccinated part of the population. Such vaccination effects are

expected to further complicate the reliable estimation of the effective reproduction number based only on the numbers of confirmed cases. While our retrospective modelling approach is primarily based on the development of reported death counts, numbers of hospitalizations and intensive care unit cases are important additional and more timely indicators, which could also be integrated into the estimation of the course of the pandemic[40,41].

In conclusion, the presented retrospective spline-based modelling approach for estimating effective reproduction numbers and dark figures of infections provides additional insights regarding the course of the pandemic. In particular, by incorporating effective infection fatality rates for modelling the link between infections and deaths, the hierarchical model can disentangle overlying effects of changes in testing and mitigation measures. Future research should be targeted at integrating various pieces of information for modelling the further course of the pandemic, including data on vaccinations, confirmed cases, hospitalizations, intensive care unit cases and death counts.

Data availability

All data used for this study are publicly available from the Robert Koch Institute (RKI). RKI estimates of the effective reproduction number are available online[21]. Data on daily reported death counts are available online[22]. Data on weekly confirmed cases for different age groups are available online[25]. Data on individual confirmed cases are available online[26].

Code availability

Source code for reproducing all results of the paper is publicly available on Github: https://github.com/wistubaT/SplineModelGermany.

References

1. Ebrahim, S. H. *et al.* All hands on deck: A synchronized whole-of-world approach for COVID-19 mitigation. *Int. J. Infect. Dis.* **98**, 208–215. https://doi.org/10.1016/j.ijid.2020.06.049 (2020).
2. Dietz, K. The estimation of the basic reproduction number for infectious diseases. *Stat. Methods Med. Res.* **2**(1), 23–41. https://doi.org/10.1177/096228029300200103 (1993).
3. Alimohamadi, Y., Taghdir, M. & Sepandi, M. Estimate of the basic reproduction number for COVID-19: A systematic review and meta-analysis. *J. Prev. Med. Public Health* **53**(3), 151–157. https://doi.org/10.3961/jpmph.20.076 (2020).
4. Ahmad, Z., Arif, M., Ali, F., Khan, I. & Nisar, K. S. A report on COVID-19 epidemic in Pakistan using SEIR fractional model. *Sci. Rep.* **10**, 22268. https://doi.org/10.1038/s41598-020-79405-9 (2020).
5. Locatelli, I., Trächsel, B. & Rousson, V. Estimating the basic reproduction number for COVID-19 in Western Europe. *PLoS ONE* **16**(3), e0248731. https://doi.org/10.1371/journal.pone.0248731 (2021).
6. Hamouda, O. Schätzung der aktuellen Entwicklung der SARS-CoV-2-Epidemie in Deutschland-Nowcasting. *Epidemiol. Bull.* **17**, 10–15. https://doi.org/10.25646/6692.4 (2020).
7. Demongeot, J. *et al.* Estimation of daily reproduction numbers during the COVID-19 outbreak. *Computation* **9**(10), 109. https://doi.org/10.3390/computation9100109 (2021).
8. Günther, F., Bender, A., Katz, K., Küchenhoff, H. & Höhle, M. Nowcasting the COVID-19 pandemic in Bavaria. *Biometrical J.* **63**(3), 490–502. https://doi.org/10.1002/bimj.202000112 (2021).
9. Schneble, M., De Nicola, G., Kauermann, G. & Berger, U. Nowcasting fatal COVID-19 infections on a regional level in Germany. *Biometrical J.* **63**(3), 471–489. https://doi.org/10.1002/bimj.202000143 (2020).
10. Schneble, M., De Nicola, G., Kauermann, G. & Berger, U. A statistical model for the dynamics of COVID-19 infections and their case detection ratio in 2020. *Biometrical J.* **63**(8), 1623–1632. https://doi.org/10.1002/bimj.202100125 (2021).
11. Flaxman, S. *et al.* Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* **584**(7820), 257–261. https://doi.org/10.1038/s41586-020-2405-7 (2020).
12. Brauner, J. M. *et al.* Inferring the effectiveness of government interventions against COVID-19. *Science* **371**, 6531. https://doi.org/10.1126/science.abd9338 (2021).
13. Sharma, M. *et al.* Understanding the effectiveness of government interventions against the resurgence of COVID-19 in Europe. *Nat. Commun.* **12**, 5820. https://doi.org/10.1038/s41467-021-26013-4 (2021).
14. Chin, V., Ioannidis, J. P., Tanner, M. A. & Cripps, S. Effect estimates of COVID-19 non-pharmaceutical interventions are non-robust and highly model-dependent. *J. Clin. Epidemiol.* **136**, 96–132. https://doi.org/10.1016/j.jclinepi.2021.03.014 (2021).
15. Wood, S. N. Inferring UK COVID-19 fatal infection trajectories from daily mortality data: Were infections already in decline before the UK lockdowns? *Biometrics*. https://doi.org/10.1111/biom.13462 (2021).
16. Bönisch, S. *et al.* Effects of coronavirus disease (COVID-19) related contact restrictions in Germany, March to May 2020, on the mobility and relation to infection patterns. *Front. Public Health* **8**, 619. https://doi.org/10.3389/fpubh.2020.568287 (2020).
17. Schlosser, F. *et al.* COVID-19 lockdown induces disease-mitigating structural changes in mobility networks. *Proc. Natl. Acad. Sci.* **117**(52), 32883–32890. https://doi.org/10.1073/pnas.2012326117 (2020).
18. O'Driscoll, M. *et al.* Age-specific mortality and immunity patterns of SARS-CoV-2. *Nature* **590**(7844), 140–145. https://doi.org/10.1038/s41586-020-2918-0 (2021).
19. Levin, A. T. *et al.* Assessing the age specificity of infection fatality rates for COVID-19: Systematic review, meta-analysis, and public policy implications. *Eur. J. Epidemiol.* **35**, 1123–1138. https://doi.org/10.1007/s10654-020-00698-1 (2020).
20. Staerk, C., Wistuba, T. & Mayr, A. Estimating effective infection fatality rates during the course of the COVID-19 pandemic in Germany. *BMC Public Health* **21**, 1073. https://doi.org/10.1186/s12889-021-11127-7 (2021).
21. Robert Koch-Institut. Nowcasting und R-Schätzung: Schätzung der aktuellen Entwicklung der SARS-CoV-2-Epidemie in Deutschland. https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Projekte_RKI/Nowcasting.html (2021).
22. Robert Koch-Institut. Aktueller Lage-/Situationsbericht des RKI zu COVID-19. https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Situationsberichte/Gesamt.html (2021).
23. Lauer, S. A. *et al.* The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: Estimation and application. *Ann. Intern. Med.* **172**(9), 577–582. https://doi.org/10.7326/M20-0504 (2020).
24. Brazeau, N. *et al.* Covid-19 Infection Fatality Ratio: Estimates from Seroprevalence Vol. 34 (Imperial College London, 2020). https://doi.org/10.25561/83545.

25. Robert Koch-Institut. COVID-19-Fälle nach Altersgruppe und Meldewoche. https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Daten/Altersverteilung.html;jsessionid=A1ED0937C08D73850C426A1EF89EAFAF.internet102?nn=2386228 (2021).
26. Bundesamt für Kartographie und Geodäsie, Robert Koch-Institut (RKI). Fallzahlen in Deutschland. https://npgeo-corona-npgeo-de.hub.arcgis.com (2021).
27. Nishiura, H., Linton, N. M. & Akhmetzhanov, A. R. Serial interval of novel coronavirus (COVID-19) infections. *Int. J. Infect. Dis.* **93**, 284–286. https://doi.org/10.1016/j.ijid.2020.02.060 (2020).
28. Kharratzadeh, M. Splines In Stan. https://mc-stan.org/users/documentation/case-studies/splines_in_stan.html (2017).
29. R Core Team. *R: A Language and Environment for Statistical Computing*. https://www.R-project.org/ (R Foundation for Statistical Computing, 2021).
30. Stan Development Team. *Stan Modeling Language Users Guide and Reference Manual, Version 2.27*. https://mc-stan.org (2021).
31. Hoffman, M. D. & Gelman, A. The No-U-Turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *J. Mach. Learn. Res.* **15**(1), 1593–1623 (2014).
32. Betancourt, M. J. & Girolami, M. Hamiltonian Monte Carlo for hierarchical models. *arXiv* (2013). Preprint at: arXiv:1312.0906.
33. Li, T. & White, L. F. Bayesian back-calculation and nowcasting for line list data during the COVID-19 pandemic. *PLoS Comput. Biol.* **17**(7), e1009210. https://doi.org/10.1371/journal.pcbi.1009210 (2021).
34. Bhattacharyya, R. P. & Hanage, W. P. Challenges in inferring intrinsic severity of the SARS-CoV-2 Omicron variant. *N. Engl. J. Med.* **386**(7), e14. https://doi.org/10.1056/NEJMp2119682 (2022).
35. Moore, S., Hill, E. M., Tildesley, M. J., Dyson, L. & Keeling, M. J. Vaccination and non-pharmaceutical interventions for COVID-19: A mathematical modelling study. *Lancet Infect. Dis.* **21**(6), 793–802. https://doi.org/10.1016/S1473-3099(21)00143-2 (2021).
36. Sonabend, R. *et al.* Non-pharmaceutical interventions, vaccination, and the SARS-CoV-2 delta variant in England: A mathematical modelling study. *Lancet* **398**(10313), 1825–1835. https://doi.org/10.1016/S0140-6736(21)02276-5 (2021).
37. Malik, A. *et al.* Sensitivity analysis of COVID-19 with quarantine and vaccination: A fractal-fractional model. *Alex. Eng. J.* **61**(11), 8859–8874. https://doi.org/10.1016/j.aej.2022.02.024 (2022).
38. Contreras, S. *et al.* The challenges of containing SARS-CoV-2 via test-trace-and-isolate. *Nat. Commun.* **12**, 378. https://doi.org/10.1038/s41467-020-20699-8 (2021).
39. Rovida, F. *et al.* SARS-CoV-2 vaccine breakthrough infections are asymptomatic or mildly symptomatic and are infrequently transmitted. *medRxiv*https://doi.org/10.1101/2021.06.29.21259500 *(2021).*
40. Oshinubi, K., Ibrahim, F., Rachdi, M. & Demongeot, J. Functional data analysis: Application to daily observation of COVID-19 prevalence in France. *AIMS Math.* **7**(4), 5347–5385. https://doi.org/10.3934/math.2022298 (2022).
41. Fritz, C. *et al.* Statistical modelling of COVID-19 data: Putting generalised additive models to work. *arXiv* (2022). Preprint at: arXiv:2201.02182.

Author contributions

T.W. developed and implemented the hierarchical model, supported by C.S.; T.W., A.M. and C.S. contributed to the conception and design of the project, the interpretation of the results and the writing of the manuscript. All authors have read and approved the final version.

Funding

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-13723-y.

Correspondence and requests for materials should be addressed to C.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.