medRxiv doi:10.1101/2020.05.01.20088047
based on data from June 25, 2020

The reproduction number of COVID-19 and its correlation with public health interventions

Kevin Linka · Mathias Peirlinck · Ellen Kuhl

Received: May 1, 2020

Abstract Throughout the past six months, no number has dominated the public media more persistently than the reproduction number of COVID-19. This powerful but simple concept is widely used by the public media, scientists, and political decision makers to explain and justify political strategies to control the COVID-19 pandemic. Here we explore the effectiveness of political interventions using the reproduction number of COVID-19 across Europe. We propose a dynamic SEIR epidemiology model with a time-varying reproduction number, which we identify using machine learning. During the early outbreak, the basic reproduction number was 4.22±1.69, with maximum values of 6.33 and 5.88 in Germany and the Netherlands. By May 10, 2020, it dropped to 0.67±0.18, with minimum values of 0.37 and 0.28 in Hungary and Slovakia. We found a strong correlation between passenger air travel, driving, walking, and transit mobility and the effective reproduction number with a time delay of 17.24±2.00 days. Our new dynamic SEIR model provides the flexibility to simulate various outbreak control and exit strategies to inform political decision making and identify safe solutions in the benefit of global health.

Keywords COVID-19 · epidemiology · SEIR model · reproduction number · machine learning

1 Motivation

Since the beginning of the new coronavirus pandemic in December 2020, no other number has been discussed more controversially than the reproduction number of COVID-19 [36]. Epidemiologists use the basic reproduction number R_0 to quantify how many new infections a single infectious individual creates in an otherwise completely susceptible population [13]. The public media, scientists, and political decision makers across the globe have started to adopted the basic reproduction number as an illustrative metric to explain and justify the need for community mitigation strategies and political interventions [21]: An outbreak will continue for $R_0 > 1$ and come to an end for $R_0 < 1$ [25]. While the concept of R_0 seems fairly simple, the reported basic reproduction number for COVID-19 varies hugely depending on country, culture, calculation, stage of the outbreak [36]. Knowing the precise number of R_0 is important, but challenging, because of limited data and incomplete reporting [12]. It is difficult–if not impossible–to measure R_0 directly [50]. The earliest COVID-19 study that followed the first 425 cases of the Wuhan outbreak via direct contact tracing reported a basic reproduction number of 2.2 [33]. However, especially during the early stages of the outbreak, information was limited because of insufficient testing, changes in case definitions, and overwhelmed healthcare systems [47]. Most basic reproduction numbers of COVID-19 we see in the public media today are estimates of mathematical models that depend critically on the choice of the model, the initial conditions, and numerous other modeling assumptions [12]. To no surprise, the mathematically predicted basic reproduction numbers cover a wide range, from 2.2–3.6 for exponential growth models to 4.1–6.5 for more sophisticated compartment models [36].

Compartment models are a popular approach to simulate the epidemiology of an infectious disease [29]. A prominent compartment model is the SEIR model that represents the timeline of a disease through the interplay of four compartments that contain the susceptible, exposed, infectious, and recovered populations [6]. The SEIR model has three characteristic parameters, the transition rates β from the susceptible to the exposed state, α from the exposed to the infectious state, and γ from the infectious to the recovered

Kevin Linka · Mathias Peirlinck · Ellen Kuhl
Department of Mechanical Engineering, Stanford University, Stanford, California, United States, klinka/mpeirlin/ekuhl@stanford.edu

state [25]. The latter two are disease specific parameters associated with the inverses of the latent period $A = 1/\alpha$ during which an individual is exposed but not yet infectious, and the infectious period $C = 1/\gamma$ during which an individual can infect others [32]. For COVID-19, depending on the way of reporting, these two times can vary anywhere between $A = 2$ to 6 days and $C = 3$ to 18 days [40,42,44]. The most critical feature of any epidemiology model is the transition from the susceptible to the exposed state. This transition typically scales with the size of the susceptible and infectious populations S and I, and with the contact rate β, the inverse of the contact period $B = 1/\beta$ between two individuals of these populations [25]. The product of the infectious period and the contact rate defines the reproduction number $R = C\beta$ [12]. Community mitigation strategies and political interventions seek to reduce the contact rate β, and with it the reproduction number R, to control the outbreak of a pandemic [44].

The first official case of COVID-19 in Europe was reported on January 24, 2020. Within only 45 days, the pandemic spread across all 27 countries of the European Union [15]. On March 17, for the first time in its history, the European Union closed all its external borders to prevent a further spreading of the disease [16]. Within the following two weeks, many local governments supplemented the European regulations with lockdowns and national travel restrictions. In response, passenger air travel within the European Union dropped by up to 95% [18]. These drastic measures have stimulated a wave of criticism, especially because initially, it was entirely unclear to which extent they would succeed in reducing the number of new infections [38].

In this study, as Europe begins to relax these constraints, we correlate the effect of Europe-wide travel restrictions to the outbreak dynamics of COVID-19. We introduce a dynamic SEIR model with a time-varying contact rate $\beta(t)$ that transitions smoothly from the initial contact rate β_0 at the beginning of the outbreak to the effective contact rate β_t under global travel restrictions and local lockdown. We express the time-varying contact rate $\beta(t) = R(t)/C$ as a function of the effective reproduction number $R(t)$ and use Bayesian inference to learn the evolution of the reproduction number for each country of the European Union from its individual outbreak history [15]. Our model allows us to precisely quantify the initial basic reproduction number R_0, the effective reproduction number R_t, and the adaptation time t^* to achieve this reduction, which are important quantitative metrics of the effectiveness of national public health intervention. Our model also specifies the exact time delay Δt between the implementation of political actions and their effects on the outbreak dynamics of COVID-19. This time delay is particularly important to plan exit strategies and estimate risks associated with gradually or radically relaxing current local lockdowns and global travel restrictions.

2 Methods

Epidemiology modeling. We model the epidemiology of the COVID-19 outbreak using an SEIR model with four compartments, the susceptible, exposed, infectious, and recovered populations, governed by a set of ordinary differential equations [34], see Appendix,

$$\begin{aligned}
\dot{S} &= -\beta SI/N \\
\dot{E} &= +\beta SI/N - \alpha E \\
\dot{I} &= \qquad\quad + \alpha E - \gamma I \\
\dot{R} &= \qquad\qquad\quad + \gamma I\,.
\end{aligned}$$

The transition rates between the four compartments, β, α, and γ, are inverses of the contact period $B = 1/\beta$, the latent period $A = 1/\alpha$, and the infectious period $C = 1/\gamma$, and $N = S + E + I + R$ is the total population. We interpret the latency rate α and the infectious rate γ as disease-specific for COVID-19, and assume that they are constant across all 27 countries of the European Union. We interpret the contact rate $\beta = \beta(t)$ as behavior specific, and assume that it is different for each country and can vary in time to reflect the effect of societal and political actions. For easier interpretation, we express the contact rate $\beta(t) = R(t)/C$ in terms of the time-varying effective reproduction number $R(t)$. For the effective reproduction number, we postulate a hyperbolic tangent type ansatz,

$$R(t) = R_0 - \tfrac{1}{2}[\,1 + \tanh([t - t^*]/T)\,][R_0 - R_\mathrm{t}]\,.$$

This ansatz ensures a smooth transition from the basic reproduction number R_0 at the beginning of the outbreak to the current reproduction number R_t under travel restrictions and lockdown, where t^* is the adaptation time and T is the transition time, see Appendix.

COVID-19 outbreak and mobility data. We draw the COVID-19 outbreak data for all 27 countries of the European Union [15]. From these data, we extract the newly confirmed cases as the difference between today's and yesterday's reported cases. We sample all European air traffic data from the Eurocontrol dashboard, a pan-European Organization dedicated to support European aviation [19]. In addition, we approximate car, walking, and transit mobility using a database generated from cell phone data [4]. These data represent the relative volume of location requests per city, subregion, region, and country, scaled by the baseline volume on January 13, 2020. We smoothen the weekday-weekend fluctuations in outbreak and mobility data by applying a moving averaging window of seven days.

Machine learning. To analyze the evolution of the effective reproduction number for each country, and predict possible exit scenarios, we identify the initial exposed and infectious populations E_0 and I_0 and the effective reproduction number $R(t)$ using the reported COVID-19 cases in all

27 countries of the European Union [15]. For each country, our simulation window begins on the day at which the number of reported cases surpasses 100 individuals and ends on May 10, 2020 for the initial simulation and on June 20, 2020 for the prediction. We fix the latency and infectious periods to $A = 2.5$ days and $C = 6.5$ days [31, 33, 47]. To account for uncertainties in the initial exposed and infectious populations E_0 and I_0 and in the effective reproduction number $R(t)$, we use Bayesian inference with Markov-Chain Monte-Carlo to estimate the following set of model parameters $\vartheta = \{E_0, I_0, \sigma, R_0, R_\mathrm{t}, t^*, T\}$. Here, σ represents the width of the likelihood $p(\hat{D}(t)\,|\,\vartheta)$ between the time-varying reported new cases $\hat{D}(t)$ and the simulated affected population $D(t,\vartheta)$. We adopt a Student's t-distribution for the likelihood between the data and the model predictions [11, 30] with a confirmed case number-dependent width,

$$p(\hat{D}(t)\,|\vartheta) \sim \mathrm{StudentT}_{\nu=4}(\ \mathrm{mean} = D(t,\vartheta),\ \mathrm{width} = \sigma\sqrt{D(t,\vartheta)}).$$

We apply Bayes' rule to obtain the posterior distribution of the parameters [41, 45] using the prior distributions in Table 1 and the reported case numbers [15],

$$p(\vartheta\,|\,\hat{D}(t)) = \frac{p\left(\hat{D}(t)\,|\,D(t,\vartheta)\right)\,p(\vartheta)}{p\left(\hat{D}(t)\right)}.$$

We solve this distribution numerically using the NO-U-Turn sampler [26] implementation of the python package PyMC3 [46]. We use two chains: The first 1000 samples are used to tune the sampler, and are later discarded; the subsequent 1000 samples are used to estimate the set of parameters ϑ. Chain convergence requires a geometric ergodicity between the Markov transition and the target distribution. In PyMC3 this is detected by split $\hat{R}$ statistics, which identifies convergence by comparing the variance between the chains. From the converged posterior distributions, we sample multiple combinations of parameters that describe the time evolution of reported cases. These posterior samples allow us to quantify the uncertainty on each parameter.

Table 1 Prior distributions for the initial exposed and infectious populations E_0 and I_0, width of likelihood σ, basic and effective reproduction numbers R_0 and R_t, adaptation time t^*, and transition time T.

Parameter	Distribution
E_0	LogNormal(log($D(t=A)$),1.5)
I_0	LogNormal(log($D(t=0)$),1.5)
σ	HalfCauchy($\beta = 1$)
R_0	Normal(2.5,2)
R_t	Normal(2.5,2)
t^*	Normal(10,10)
T	LogNormal(log(3),1.5)

To probe the effect of different exit strategies, we explore three possible projections of the effective reproduction number $R(t)$ for each posterior parameter sample set and predict the outbreak dynamics for a 40-day period after our initial manuscript submission, from May 10 until June 20, 2020. The first scenario assumes a constant effective reproduction number $R(t) = R_\mathrm{t}$, the second and third scenarios simulate the effect of a linear return from R_t to the country-specific basic reproduction number R_0, either rapidly within one month, or more gradually within three months. In the revision of our manuscript, we added the reported daily new cases from May 10 until June 20, 2020 to compare our model predictions against the real case data.

3 Results

Figure 1 illustrates the outbreak dynamics of COVID-19 for all 27 countries of the European Union. The dots represent daily new cases. The brown and red curves illustrate the fit of the SEIR model and the effective reproduction number for the time period until May 10, 2020. The gray shaded area highlights the model predictions for the 40-day period of gradual reopening, from May 10 until June 20, 2020. The dashed brown, orange, and red curves illustrate the projections for three possible exit strategies: a constant continuation at the effective reproduction number R_t from May 10, 2020, a gradual return to the basic reproduction number R_0 within three months, and a rapid to R_0 within one months.

Table 2 and Figures 2 and 3 summarize the basic reproduction number R_0 at the beginning of the COVID-19 outbreak and the effective reproduction number R_t as of May 10, 2020. The basic reproduction number R_0 has maximum values in Germany, the Netherlands, and Spain, with 6.33, 5.88, and 5.19 and minimum values in Bulgaria, Croatia, and Lithuania with 1.29, 0.93, and 0.91. The population weighted mean of the basic reproduction number across the European Union is $R_0 = 4.22 \pm 1.69$. The effective reproduction number R_t is significantly lower than the initial basic reproduction number R_0. In most countries, it is well below the critical value of $R_\mathrm{t} = 1.0$. It has maximum values in Sweden, Bulgaria, and Poland all with 1.01, 0.99, and 0.96 and minimum values in Lithuania, Hungary, and Slovakia with 0.41, 0.37, and 0.28. The population weighted mean of the basic reproduction number across the European Union is $R_\mathrm{t} = 0.67 \pm 0.18$.

Figure 4 provides a direct correlation between the reduction in mobility and the effective reproduction number of the COVID-19 outbreak across Europe. The purple, blue, grey, and black dots represent the reduction in air traffic, driving, walking, and transit mobility, the red curves show effective reproduction number with 95% confidence interval. The mean time delay Δt highlights the temporal delay between reduction in mobility and effective reproduction number. Spearman's rank correlation ρ, a measure of the statisti-

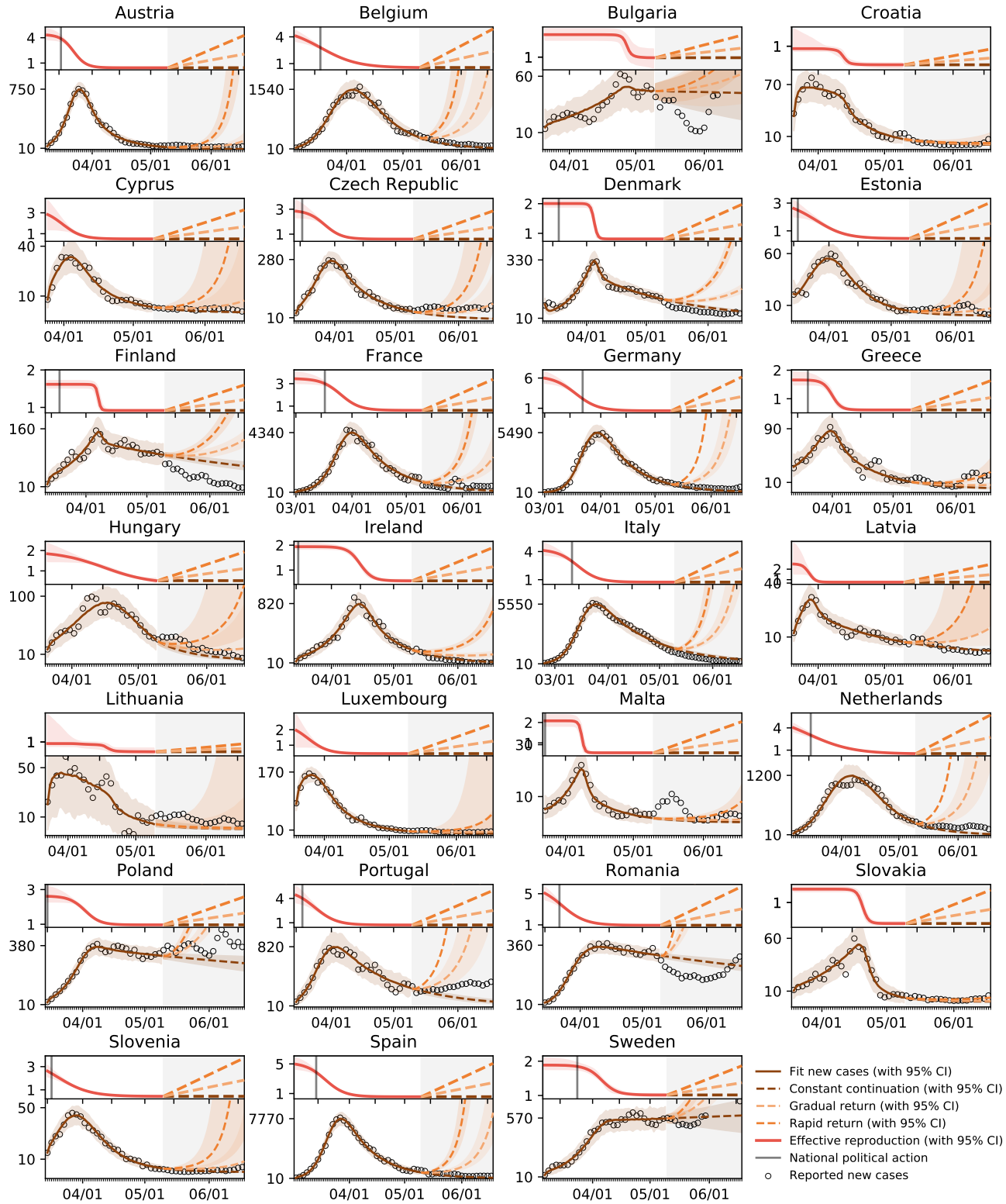

Fig. 1 Outbreak dynamics of COVID-19 across Europe and prediction of different exit strategies. The dots represent daily new cases. The brown and red curves illustrate the fit of the SEIR model and the effective reproduction number for the time period until May 10, 2020. The gray shaded area highlights the model predictions from May 10 until June 20, 2020. The dashed brown, orange, and red curves illustrate the projections for three possible exit strategies: a constant continuation at the effective reproduction number R_t from May 10, 2020, a gradual return to the basic reproduction number R_0 within three months, and a rapid to R_0 within one months.

Table 2 Parameters of the COVID-19 outbreak across Europe. Basic reproduction number R_0, effective reproduction number R_t, adaptation time t^*, adaptation speed T, and time delay Δt for fixed latency period $A = 2.5$ days and infectious period $C = 6.5$ days.

Country	Population	R_0	R_t	t^*	T	Δt
Austria	8.840.521	4.38±0.36	0.45±0.01	13.37±0.68	6.49±0.47	8.33±1.70
Belgium	11.433.256	5.00±0.73	0.54±0.03	13.31±2.84	19.30±1.57	4.00±2.35
Bulgaria	7.025.037	1.29±0.04	0.99±0.07	37.04±1.99	1.64±1.56	43.00±2.83
Croatia	4.087.843	0.93±0.22	0.49±0.03	22.36±2.90	2.46±3.61	27.33±2.36
Cyprus	1.189.265	3.35±1.14	0.50±0.02	6.60±2.87	8.02±1.47	14.00±0.00
Czech Republic	10.629.928	2.92±0.47	0.60±0.01	14.04±1.71	8.44±1.12	14.00±1.73
Denmark	5.793.636	2.00±0.05	0.81±0.01	24.74±0.29	1.72±0.45	24.00±2.92
Estonia	1.321.977	3.12±0.78	0.45±0.04	10.72±3.80	14.19±2.87	12.50±2.06
Finland	5.515.525	1.62±0.05	0.92±0.01	25.05±0.51	1.20±0.68	24.25±2.49
France	66.977.107	3.46±0.29	0.62±0.02	24.79±1.30	10.58±1.17	10.50±1.50
Germany	82.905.782	6.33±0.64	0.58±0.01	17.06±1.39	12.41±0.71	3.25±1.92
Greece	10.731.726	1.66±0.12	0.61±0.02	18.93±0.87	4.38±1.22	17.33±3.09
Hungary	9.775.564	1.97±0.55	0.37±0.15	25.62±6.55	20.23±7.33	31.67±1.89
Ireland	4.867.309	1.94±0.06	0.57±0.03	30.78±0.53	5.94±1.28	30.00±3.46
Italy	60.421.760	4.25±0.42	0.74±0.01	19.24±1.57	12.06±1.13	5.00±0.71
Latvia	1.927.174	2.50±0.89	0.76±0.01	6.99±1.32	2.70±0.90	14.67±1.89
Lithuania	2.801.543	0.91±0.88	0.41±0.09	26.23±9.88	2.25±6.25	34.67±1.89
Luxembourg	607.950	2.42±1.21	0.46±0.01	5.77±4.20	8.78±1.88	10.00±2.35
Malta	484.630	2.08±0.14	0.51±0.03	16.24±0.42	1.21±0.51	23.00±0.00
Netherlands	17.231.624	5.88±0.88	0.49±0.03	7.61±3.12	23.25±1.92	0.75±2.77
Poland	37.974.750	2.62±0.26	0.96±0.01	18.15±1.47	7.38±1.47	20.33±0.94
Portugal	10.283.822	5.10±0.86	0.73±0.02	8.93±1.86	10.40±1.22	8.67±2.62
Romania	19.466.145	6.06±0.84	0.95±0.01	8.12±1.60	11.55±0.70	8.33±2.36
Slovakia	5.446.771	1.46±0.04	0.28±0.03	31.80±0.47	2.65±0.70	40.25±0.43
Slovenia	2.073.894	3.83±0.96	0.44±0.03	4.65±3.50	15.33±1.96	6.33±2.36
Spain	46.796.540	5.19±0.50	0.57±0.01	15.90±1.17	10.70±0.69	5.50±2.60
Sweden	10.175.214	1.89±0.09	1.01±0.03	29.70±1.30	7.99±2.65	23.75±2.77
European Union	446.786.293	4.22±1.69	0.67±0.18	18.61±6.43	10.82±4.65	17.24±2.00

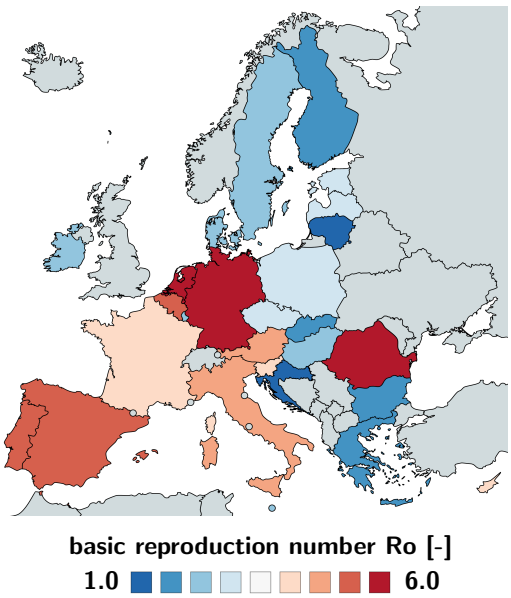

Fig. 2 Basic reproduction number R_0 of the COVID-19 outbreak across Europe. The basic reproduction number characterizes the initial number of new infectious created by one infectious individual. It has maximum values in Germany, the Netherlands, and Spain, with 6.33, 5.88, and 5.19 and minimum values in Bulgaria, Croatia, and Lithuania with 1.29, 0.93, and 0.91.

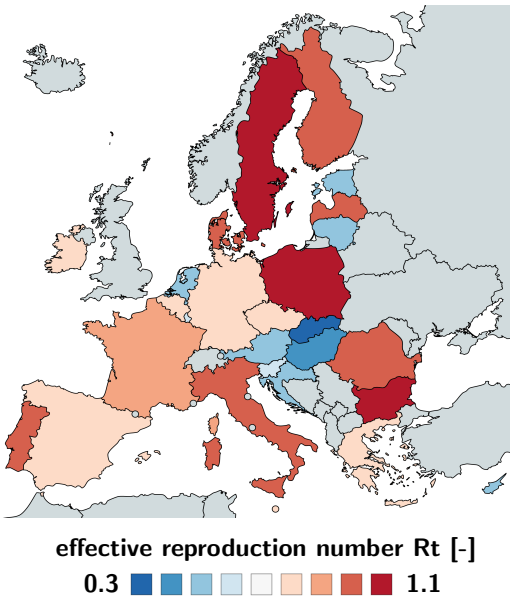

Fig. 3 Effective reproduction number R_t of the COVID-19 outbreak across Europe. The effective reproduction number characterizes the current number of new infectious created by one infectious individual. It has maximum values in Sweden, Bulgaria, and Poland all with 1.01, 0.99, and 0.96 and minimum values in Lithuania, Hungary, and Slovakia with 0.41, 0.37, and 0.28 as of May 10, 2020.

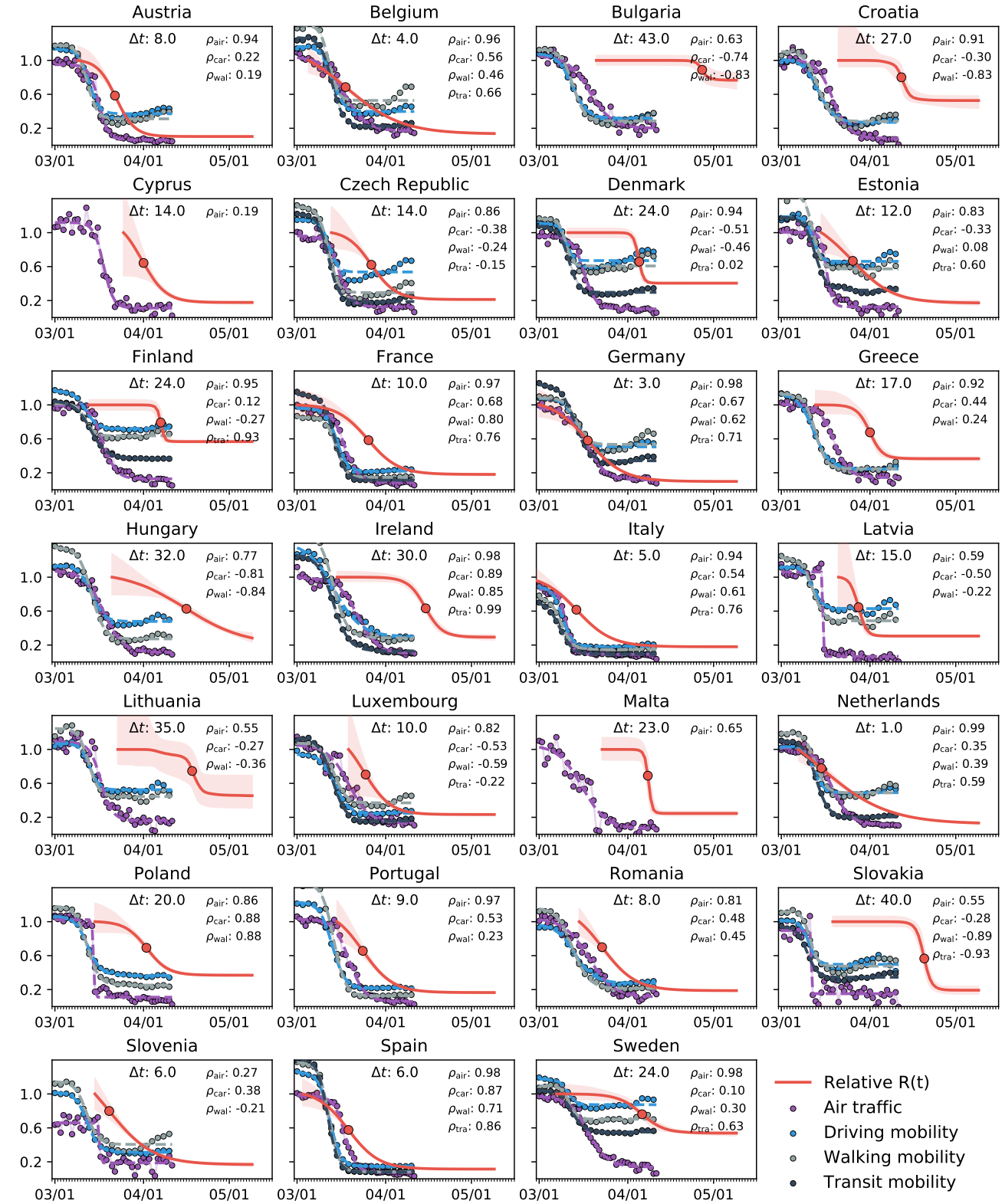

Fig. 4 Correlation between reduction in mobility and effective reproduction number of the COVID-19 outbreak across Europe. Purple, blue, grey, and black dots represent reduction in air traffic, driving, walking, and transit mobility; red curves show effective reproduction number $R(t)$ with 95% confidence interval. The mean time delay Δt highlights the temporal delay between reduction in mobility and effective reproduction number. Spearman's rank correlation ρ, measures of the statistical dependency between mobility and reproduction, and reveals the strongest correlation in the Netherlands, Germany, Ireland, Spain, and Sweden with 0.99 and 0.98.

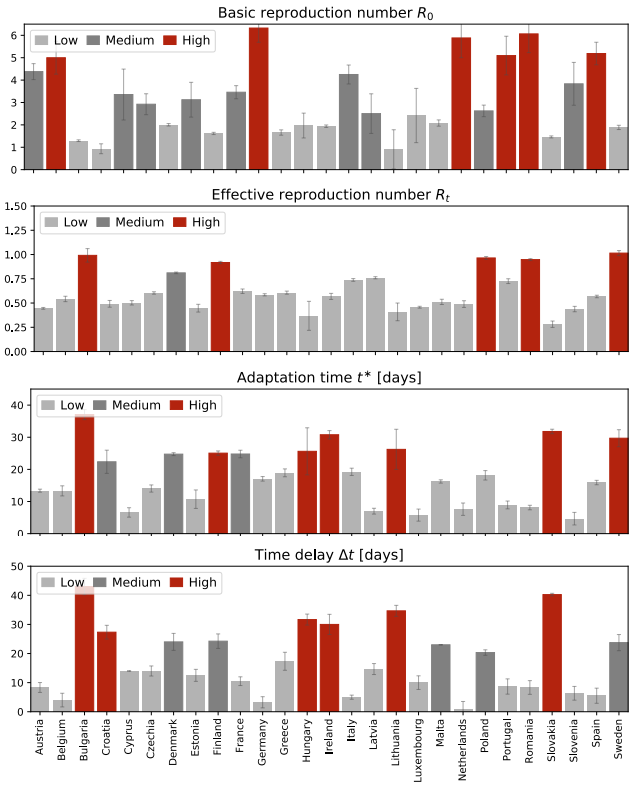

Fig. 5 Parameters of the COVID-19 outbreak across Europe. Basic reproduction number R_0, effective reproduction number R_t, adaptation time t^* and time delay Δt. The adaptation time t^* characterizes the time between the beginning of the outbreak and the reduction in the effective reproduction number; the time delay Δt characterizes the mean time between the reduction in air travel, driving, walking, and transit mobility and the reduction in the effective reproduction number.

cal dependency between both variables, reveals the strongest correlation in the Netherlands, Germany, Ireland, Spain, and Sweden with 0.99 and 0.98. Only in Slovakia, Slovenia and Lithuania, where the number of cases has not yet plateaued and the effective reproduction number does not show a clear smoothly decaying trend, there is no significant correlation between mobility and the effective reproduction number.

Figure 5 summarizes the learned basic reproduction number R_0, the effective reproduction number R_t, the adaptation time t^*, and the time delay Δt for all 27 countries of the European Union. The adaptation time t^* characterizes the time between the beginning of the outbreak at 100 confirmed cases and the reduction in the effective reproduction number and is a quantitative measure for the reaction time in the population. The time delay Δt characterizes the mean time between the reduction in air travel, driving, walking, and transit mobility and the reduction in the effective reproduction number and is a quantitative measure for the effect of mobility.

Table 2 and Figures 6 and 7 summarize the adaptation time t^* and the time delay Δt. The adaptation time t^* has

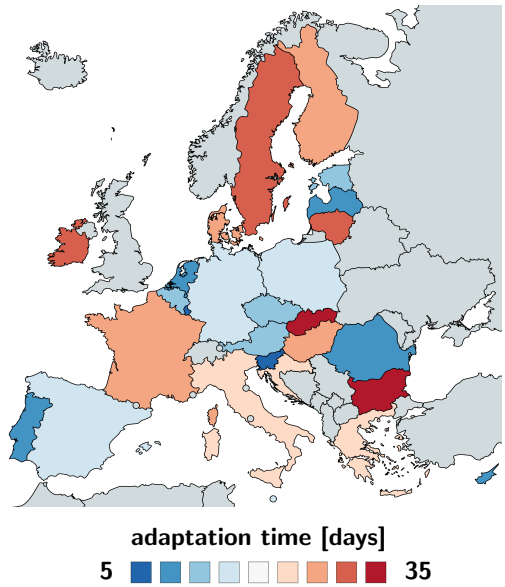

Fig. 6 Adaptation time t^* between beginning of the outbreak and reduction of the effective reproduction number across Europe. The adaptation time characterizes the time between the beginning of the outbreak at 100 confirmed cases and the reduction in the effective reproduction number. It has maximum values in Bulgaria and Slovakia with 37.04 and 31.80 days and minimum values in Luxembourg and Slovenia with 5.77 and 5.64 days.

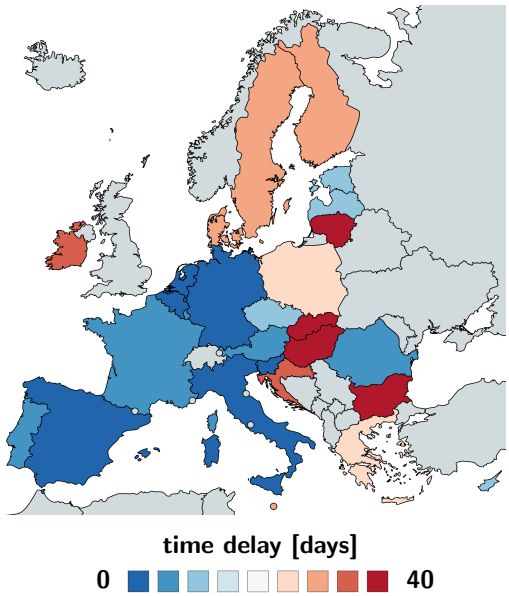

Fig. 7 Time delay Δt between reduction of air travel and reduction of the effective reproduction number across Europe. The time delay characterizes the mean time between the reduction in air travel, driving, walking, and transit mobility and the reduction in the effective reproduction number. It has maximum values in Bulgaria and Slovakia with 43.00 and 40.25 days and minimum values in Germany and the Netherlands both with 3.25 and 0.75 days.

maximum values in Bulgaria and Slovakia with 37.04 and 31.80 days and minimum values in Luxembourg and Slovenia with 5.77 and 5.64 days. The mean adaptation time

across the European Union is $t^* = 18.61 \pm 6.43$ days. The time delay Δt has maximum values in Bulgaria and Slovakia with 43.00 and 40.25 days and minimum values in Germany and the Netherlands both with 3.25 and 0.75 days. The mean time delay across the European Union is $\Delta t = 17.24 \pm 2.00$ days.

4 Discussion

Mathematical models can inform political interventions. As many countries begin to explore safe exit strategies from total lockdown, shelter in place, and national travel restrictions to manage the COVID-19 pandemic, political decision makers are turning to mathematical models for advise [10]. A powerful quantitative concept to characterize the contagiousness and transmissibility of the new coronavirus is the basic reproduction number R_0 [50]. This number explains–in simple terms–how many new infections are caused by a single one infectious individual in an otherwise completely susceptible population [13]. However, against many false claims, the *basic* reproduction number does not measure the effects of public health interventions [12]. Here, we quantify these effects, for every point in time, for every country, using the *effective* reproduction number $R(t)$, a time-dependent metric that changes dynamically in response to community mitigation strategies and political actions. We learn the effective reproduction number from case data of the COVID-19 outbreak across Europe using Bayesian inference and systematically correlate it to political interventions.

The classical SEIR model can predict a natural equilibrium and herd immunity. The SEIR model has advanced to the model of choice for the outbreak dynamics of COVID-19 [36]. It belongs to a class of infectious disease models that epidemiologists characterize as compartment models [14]. Compartment models represent the population via a sequence of compartments through which the population passes as the disease progresses. Out of the many different compartment models, the SEIR model seems best suited to mimic the epidemiology of COVID-19 via four compartments: the susceptible, exposed, infectious, and recovered populations. For more than three decades [6], epidemiologists have successfully applied the SEIR model to understand the outbreak dynamics of the measles, chickenpox, mumps, polio, rubella, pertussis, and smallpox [25]. For this class of diseases, the outbreak ends as the number of daily new cases, βSI, decreases. As such, the classical SEIR model is self-regulating: It naturally converges to an endemic equilibrium, at which either the susceptible group S, or the infectious group I, or both have become small enough to prevent new infections [32]. In epidemiology, this equilibrium is known as herd immunity [22]. In a homogeneous, well-mixed population, herd immunity occurs once a fraction of $(1 - 1/R_0)$ of the population has become immune, either through the disease itself or through vaccination, see Appendix. For the basic reproduction number of $R_0 = 4.22 \pm 1.69$ we found in this study, the herd immunity threshold would be 78%. This value is lower than 94% for the measles, 89% for chickenpox with, 86% for mumps and rubella, and 80% for polio [3], but significantly higher than the values of 16% to 27% for the seasonal flu [7]. The countries with the highest prevalence, Luxembourg with 0.72%, Sweden with 0.71%, and Spain with 0.64% [15], do currently not even come close to these values, not even when including asymptomatic cases that are believed to increase the prevalence by an order of magnitude [43], resulting in 7.2%, 7.1%, and 6.4%. Knowing the precise basic reproduction number of COVID-19 will be critical to estimate the conditions for herd immunity and predict the success of vaccination strategies.

The dynamic SEIR model can predict the effects of public health interventions. The classical SEIR model is a valuable tool to understand the interplay of the susceptible, exposed, infectious, and recovered populations under unconstrained conditions. However, for the current COVID-19 pandemic, similar to SARS, MERS, or Ebola, the dynamics of these four populations are tightly regulated by public health interventions [10] including isolation, quarantine, physical distancing, and community containment [9, 53]. This implies that model parameters like the contact rate β, the rate at which an infectious individual comes into contact and infects others, are not constant, but modulated by social behavior and political action [5]. Here we explicitly account for a dynamic contact rate $\beta(t)$ and express it as a function of the time-varying effective reproduction number $R(t)$ [55]. This allows us to "bend the curve" and predict temporary equilibrium states, far away from the equilibrium state of herd immunity, but stable under current conditions [32]. Yet, these states can quickly become unstable again once the current regulations change [53]. Our dynamic SEIR model allows us to study precisely these scenarios.

The time-varying effective reproduction number reflects the strength of public health interventions. To model temporal changes in the reproduction number, we propose a hyperbolic tangent type ansatz for the effective reproduction number $R(t)$. This functional form can naturally capture the basic reproduction number R_0, the converged reproduction number under the current constraints R_t, the adaptation time t^*, and the transition time T, see Appendix. Figure 11 illustrates how our hyperbolic tangent type model compares against a constant and a random walk type reproduction number. The constant reproduction number in Figure 11, left, nicely captures the exponential increase

during the early stages of the outbreak, but fails to "bend the curve" before herd immunity occurs. Nonetheless, several recent studies have successfully used an SEIR model with a constant reproduction number to model the outbreak dynamics of COVID-19 in China [42] and in Europe [34] by explicitly reducing the total population N to an affected population $N^* = \eta N$. The scaling coefficient $\eta = N^*/N$ is essentially a fitting parameter that indirectly quantifies the level of confinement [5]. For example, when averaged over 30 Chinese provinces, the mean affected population was $\eta = 5.19 \cdot 10^{-5} \pm 2.23 \pm 10^{-4}$, suggesting that the effect of COVID-19 was confined to only a very small fraction of the total population [42]. The Gaussian random walk in Figure 11, left, naturally captures the effects of public health interventions, however, in a daily varying, rather unpredictable way. It is a valuable method to analyze case data retrospectively, but since it does not allow for a closed functional form, it is not very useful to make informed predictions. We conclude that the hyperbolic tangent based ansatz in Figure 11, middle, with four physically meaningful parameters, is the most useful approach to represent the time-varying effective reproduction number $R(t)$ for our current purposes.

Bayesian inference identifies basic and effective reproduction numbers from reported cases. Unfortunately, we can neither measure the basic nor the effective reproduction number directly. However, throughout the past six months, the COVID-19 pandemic has probably generated more quantitative data than any infectious disease in history. Parametric Bayesian methods offers incredible opportunities to evaluate these data and learn correlations and trends [39]. Here we learn the effective reproduction number $R(t)$ directly from the reported COVID-19 cases in all 27 countries of the European Union, starting from the day of the first reported case on January 24, until May 10, 2020. This not only allows us to identify the model parameters and confidence intervals, but also to quantify correlations between travel restrictions and reduced effective reproduction numbers. Table 2 and Figures 2 and 3 summarize our basic reproduction numbers R_0 and effective reproduction numbers R_t for all 27 countries. Our mean basic reproduction number of $R_0 = 4.22 \pm 1.69$ exceeds the first estimates of 1.4 to 2.5 from the World Health Organization based on a tracing study that reported a value of 2.2 during the early outbreak in Wuhan [33]. However, our results agree well with the more recent values of 5.7 for the Wuhan outbreak [47] and with a recent review that suggested values from 4.1 to 6.5 calculated with SEIR models [36]. Our basic reproduction number of 4.22 is lower than the numbers of 18 for measles, 9 for chickenpox, 7 for mumps, 7 for rubella, and 5 for poliomyelitis [3]. Compared to the SARS coronavirus with a range from 2 to 5 [36], our values of SARS-CoV-2 in Table 2 are rather on the high end, suggesting that the new coronavirus would spread more rapidly than SARS [54]. Knowing the precise basic reproduction number is critical to estimate the number of contacts to trace, if we want to successfully control the dynamics of COVID-19 through contact trancing [24].

Political mitigation strategies reduce the effective reproduction number with a time delay of two weeks. Freedom of movement is the fundamental principle of the European Union. On March 13, 2020, the World Health Organization declared Europe the epicenter of the COVID-19 pandemic with more reported cases and deaths than the rest of the world combined [51]. To prevent a further spreading of the pandemic, four days later, for the first time in history, the European Union closed all its external borders [16]. In the following two weeks, the local governments augmented the European regulations with local lockdowns and national travel restrictions. Figure 4 shows that these measures had an enormous effect on the mobility within the European Union: By March 22, 2020, the average passenger air travel in Europe was cut in half, and as of May 10, it is reduced by 86% in Germany, 92% in France, 93% in Italy, and 95% in Spain [18]. These drastic actions have triggered an ongoing debate about the effectiveness of different outbreak strategies and the appropriate level of constraints [38]. Table 2 and Figures 4 to 7 summarize our time-varying effective reproduction number $R(t)$ and highlight the time delay of its reduction with respect to the European travel restrictions. An important socio-economical metric is mean time delay of $\Delta t = 17.24 \pm 2.00$ days between the reduction of air traffic, driving, walking, and transit mobility and the inflection point of the reproduction number curve. Figures 5 and 7 show that this time delay varies hugely across Europe with the fastest response of 0.75 days in the Netherlands, followed by Germany with 3.25 days, Belgium with 4.00 days, and Italy with 5.00 days. These fast response times naturally also reflect decisions on the national level. France had the first reported COVID-19 case in Europe on January 24, 2020 and acted rigorously and promptly by introducing the first national measures on March 16 [52]. Similarly, Italy, Spain, and Germany had introduced their national measures on March 9, March 9, and March 13, 2020 [48]. Figures 5 and 7 clearly highlight the special role of Sweden, where the government focusses efforts on encouraging the right behavior and creating social norms rather than mandatory restrictions: The time delay of 23.75 days is above the European Union average of 17.24 days, and Sweden is one of the few countries where the effective reproduction number has not yet decreased below one. Taken together, these results confirm that, especially during the early stages of an outbreak, controlling mobility can play a critical role in spreading a disease [8]. However,

these drastic political measures have stimulated an active ongoing debate when and how it would be safe to lift these restrictions.

Exit strategies will have different effects in individual countries. Political decision makers around the globe are currently trying to identify safe exit strategies from global travel restrictions and local lockdown. Mathematical models can provide guidelines and answer what-if scenarios. Our predictions in Figure 1 show projections of the number of total cases, for three possible exit strategies from lockdown: a continuation at a constant effective reproduction number R_t, a gradual return to the basic reproduction number R_0 within three months, and a rapid to R_0 within one months. Naturally, the case numbers increase in all three cases, with the steepest increase for the most rapid return. Interestingly, our method provides significantly different confidence intervals for different countries suggesting that a controlled return will be more predictable in some countries like Austria and less in others. Our projections suggest that in Sweden, were policy makers had encouraged each individual to take responsibility for their own health rather than enforcing political constraints, the projected case numbers will follow the current curve, without major deviations. Strikingly, in most countries, the newly reported case numbers upon gradual reopening, from May 10 to June 20, 2020, follow the dashed brown curves of the prediction with a constant effective reproduction number. This suggests that most countries have learnt how to successfully control the pandemic and manage new outbreaks.

Limitations. Just like any infectious disease model, our model inherently faces limitations associated with data uncertainties from differences in testing, inconsistent diagnostics, incomplete counting, and delayed reporting. For our specific study of COVID-19, we encounter a few additional limitations: First, although a massive amount of data are freely available through numerous well-documented public databases, the selection of the model naturally limits what we can predict and it remains challenging to map the available information into the format of the SEIR model. Second, the initial conditions for our exposed and infectious populations will always remain unknown and many new first cases have been reported throughout the past couple of weeks. To reduce the influence of unknown initial conditions, our parametric Bayesian inference algorithm learns these populations alongside the effective reproduction number. Third, in its current state, our model does not distinguish between community mitigation strategies, local public health recommendations, and global political actions [9]. We are currently integrating the current approach into a global network model that will provide more granularity to include other community mitigation strategies in addition to mobility. Fourth, our current model is not directly informed by mobility data. We have recently proposed a new method that uses a stochastic process to directly incorporate mobility as a latent variable into the present SEIR model framework [35]. Fifth, and probably most importantly, our current knowledge limits our ability to make firm predictions about the recovered group, which will be critical to estimate the return to normal. Recent studies have shown that the unreported asymptomatic population is huge, up to an order of magnitude larger than the reported symptomatic population traced in our study [43]. A related challenge is that the number of reported cases strongly depends on the testing strategy of each country. A possibility to eliminate testing bias could be to use death counts rather than case counts [23]; however, this would also require a consistent Europe-wide definition of death with versus death caused by COVID-19. In general, more targeted tests will be needed to identify the size of the asymptomatic population and explore whether it behaves differently in terms of contact rate and infectious period, which would both radically change the overall reproduction number. As more data become available, we are confident that we will learn from uncertainty quantification, become more confident in our model predictions, and learn how to quickly extract important trends.

5 Conclusion

We quantified the effectiveness of public health interventions using the effective reproduction number R, the time-varying reproduction number of the COVID-19 pandemic, across all 27 countries of the European Union. We adopted an SEIR epidemiology model with a dynamic effective reproduction number, which we learned for each country from its individual reported cases using Bayesian inference . We found that, during the early stages of the COVID-19 outbreak, the basic reproduction number across Europe was $R_0 = 4.22 \pm 1.69$. Massive public health interventions as well as social learning have successfully reduced the effective reproduction number to $R_t = 0.67 \pm 0.18$ by May 10, 2020. Strikingly, this reduction displays a strong correlation with mobility in the form of air travel, driving, walking and transit mobility with a mean time delay of 17.24 ± 2.00 days. This time delay is an important metric as we seek to identify safe exit strategies from current lockdown and travel restrictions. To highlight the predictive potential of our model, we simulated different exit strategies from lockdown that either maintain the current status quo, gradually return to normal, or rapidly return to the early exponential growth. Upon gradual reopening, from May 10 to June 20, 2020, the newly reported case numbers in most countries followed the prediction that maintained the current effective reproduction num-

ber suggesting that most countries were able to successfully manage the pandemic and control new outbreaks. Our dynamic epidemiology model provides the flexibility to simulate the effects and timelines of various outbreak control and exit strategies to inform political decision making and identify solutions that minimize the impact of COVID-19 on global health.

Appendix

The SEIR model. The SEIR model is a popular model in the epidemiology of infectious diseases [25]. It represents the timeline of a disease through four compartments that characterize the dynamics of the susceptible, exposed, infectious, and recovered populations [6]. The transition between these populations is governed by a set of ordinary differential equations [29],

$$\begin{aligned} \dot{S} &= -\beta SI/N \qquad\qquad - \mu S + \mu N \\ \dot{E} &= +\beta SI/N - \alpha E \qquad - \mu E \\ \dot{I} &= \qquad\qquad + \alpha E - \gamma I - \mu I \\ \dot{R} &= \qquad\qquad\qquad + \gamma I - \mu R . \end{aligned} \tag{1}$$

The transition rates between the four populations, the contact rate β, the latency rate α, and the infectious rate γ, are inverses of the contact period $B = 1/\beta$, the latent period $A = 1/\alpha$, and the infectious period $C = 1/\gamma$. The set of equations (1) includes vital dynamics at an equivalent birth and death rate μ, such that the sum of all four equations, (1.1) to (1.4), is equal to zero,

$$\dot{S} + \dot{E} + \dot{I} + \dot{R} = 0 . \tag{2}$$

This implies that the sum of the four populations is constant and equal to the total population N,

$$S + E + I + R = \text{const.} = N . \tag{3}$$

For the SEIR model with vital dynamics (1), the basic reproduction number R_0, the number of new infections caused by one infectious individual in a completely susceptible population [13], is

$$R_0 = \frac{\alpha}{\alpha+\mu}\frac{\beta}{\gamma+\mu} . \tag{4}$$

The magnitude of R_0 plays a critical role in the outbreak dynamics of an infectious disease [36]. Here we are interested in studying the outbreak dynamics of COVID-19, for which the time period is short, and we can neglect the effects of vital dynamics. This implies that the set of equations (1) reduces to the following system,

$$\begin{aligned} \dot{S} &= -\beta SI/N \\ \dot{E} &= +\beta SI/N - \alpha E \\ \dot{I} &= \qquad\qquad + \alpha E - \gamma I \\ \dot{R} &= \qquad\qquad\qquad + \gamma I , \end{aligned} \tag{5}$$

and the basic reproduction number (4) simplifies to the following expression,

$$R_0 = \beta/\gamma = C\beta = C/B . \tag{6}$$

Many infections diseases, including COVID-19, display a significant latent period during which individuals have been infected but are not yet infectious themselves. These individuals are represented through the exposed population E. A special case of the SEIR model is the SIR model, which follows from the set of equations (5) with $\alpha \to \infty$ as

$$\begin{aligned} \dot{S} &= -\beta SI/N \\ \dot{I} &= +\beta SI/N - \gamma I \\ \dot{R} &= \qquad\qquad + \gamma I . \end{aligned} \tag{7}$$

While the SIR model is conceptually simpler and lends itself to closed form solutions, for the outbreak dynamics of the COVID-19 pandemic, the invisible exposed, but not yet infectious population plays a critical role. Throughout this study, we therefore focus on the SEIR model. We reparameterize the absolute SEIR model (5) and scale it by the total population N, to obtain the fractions of the susceptible, exposed, infectious, and recovered populations,

$$s = S/N \qquad e = E/N \qquad i = I/N \qquad r = R/N . \tag{8}$$

This introduces the relative SEIR model,

$$\begin{aligned} \dot{s} &= -\beta si \\ \dot{e} &= +\beta si - \alpha e \\ \dot{i} &= \qquad\qquad + \alpha e - \gamma i \\ \dot{r} &= \qquad\qquad\qquad + \gamma i , \end{aligned} \tag{9}$$

parameterized in the fractional populations, s, e, i, and r, which sum up to one,

$$s + e + i + r = 1 . \tag{10}$$

Endemic equilibrium. The hallmark of a typical epidemic outbreak is that it begins with a small infectious population I_0. The infectious population $I(t)$ increases, reaches a peak, and then decays to zero [25]. Throughout the outbreak, the susceptible population $S(t)$ decreases, but the final susceptible population S_∞ always remains larger than zero. This final state is called the endemic equilibrium. To estimate the endemic equilibrium of the COVID-19 pandemic, we divide equation (5.1) by equation (5.4),

$$\frac{\dot{S}}{\dot{R}} = -\frac{\beta SI}{\gamma NI} \tag{11}$$

Separation of the variables and using the definition of the basic reproduction number $R_0 = \beta/\gamma$ yields the following equation,

$$\frac{\dot{S}}{S} = -\frac{R_0}{N}\dot{R} , \tag{12}$$

which we integrate in time,

$$\int \frac{1}{S}\frac{\mathrm{d}S}{\mathrm{d}t}\,\mathrm{d}t = -\int \frac{R_0}{N}\frac{\mathrm{d}R}{\mathrm{d}t}\,\mathrm{d}t\,, \tag{13}$$

to obtain the following expression,

$$\ln(S(t)) - \ln(S(0)) = -R_0[R(t) - R(0)]/N\,. \tag{14}$$

Here $S(0)$ and $R(0)$ are the initial susceptible and recovered populations and $S(t)$ and $R(t)$ are these populations at time t. Using $\ln(S(t)) - \ln(S(0)) = \ln(S(t)/S(0))$ and applying the exponential function on both sides of the equation introduces the following explicit representation for the susceptible population at time t,

$$S(t) = S(0)\,\exp(-R_0[R(t) - R(0)]/N)\,. \tag{15}$$

According to equation (8), we scale the populations with the total population N as $s_0 = S(0)/N$ and $r_0 = R(0)/N$, and evaluate equation (15) at the limit $t \to \infty$ with $s_\infty = S(\infty)/N$, $e_\infty = 0$, $i_\infty = 0$, and $r_\infty = R(\infty)/N = 1 - s_\infty$, to obtain the following expression for the susceptible population at endemic equilibrium,

$$s_\infty = s_0\,\exp(-R_0[r_\infty - r_0]) = 1 - r_\infty\,. \tag{16}$$

This transcendental equation has an explicit solution in terms of the Lambert function W,

$$\begin{aligned} s_\infty &= -W(-s_0 R_0 \exp(-R_0[1 - r_0]))/R_0 \\ e_\infty &= 0 \\ i_\infty &= 0 \\ r_\infty &= 1 + W(-s_0 R_0 \exp(-R_0[1 - r_0]))/R_0\,. \end{aligned} \tag{17}$$

The endemic equilibrium condition (17) confirms that, unless $S(0) = 0$, the final susceptible population will always be larger than zero, $S_\infty > 0$ [32].

Public health interventions. The classical SEIR model (1) assumes that the disease develops freely and that the contact rate β, latency rate α, and infectious rate γ are constant throughout the course of the outbreak. It is obvious that the contact rate β will change in response to community mitigation strategies and political actions, e.g., local lockdown and global travel restrictions [20]. Here, to account for the effects of public health interventions, we introduce a time-varying contact rate $\beta(t)$ and rewrite the system of equations (5),

$$\begin{aligned} \dot{S} &= -\beta(t)\,SI/N \\ \dot{E} &= +\beta(t)\,SI/N - \alpha E \\ \dot{I} &= \qquad\qquad + \alpha E - \gamma I \\ \dot{R} &= \qquad\qquad\qquad\quad + \gamma I\,. \end{aligned} \tag{18}$$

We make a hyperbolic tangent type ansatz for the contact rate $\beta(t)$,

$$\beta(t) = \beta_0 - \tfrac{1}{2}[1 + \tanh([t - t^*]/T)][\beta_0 - \beta_\mathrm{t}]\,, \tag{19}$$

where β_0 is the initial contact rate at the onset of the pandemic, β_t is the contact rate in response to public health interventions, t^* is the adaptation time, and T is the transition time. For easier interpretation, we reparameterize the system (18) in term of the time-dependent effective reproduction number $R(t) = \beta(t)/\gamma$,

$$\begin{aligned} \dot{S} &= -R(t)\,\gamma SI/N \\ \dot{E} &= +R(t)\,\gamma SI/N - \alpha E \\ \dot{I} &= \qquad\qquad + \alpha E - \gamma I \\ \dot{R} &= \qquad\qquad\qquad\quad + \gamma I\,. \end{aligned} \tag{20}$$

With equation (19), the effective reproduction number takes the following hyperbolic tangent type form,

$$R(t) = R_0 - \tfrac{1}{2}[1 + \tanh([t - t^*]/T)][R_0 - R_\mathrm{t}]\,. \tag{21}$$

This ansatz ensures a smooth transition from the initial basic reproduction number $R_0 = \beta_0/\gamma$ at the beginning of the outbreak to the effective reproduction number $R_\mathrm{t} = \beta_\mathrm{t}/\gamma$ in response to public health interventions, where t^* and T are the adaptation and transition times. From equation (16), we can estimate the constrained equilibrium in response to public health interventions,

$$\begin{aligned} s^*_\infty &= -W(-s_\mathrm{t} R_\mathrm{t} \exp(-R_\mathrm{t}[1 - r_\mathrm{t}]))/R_\mathrm{t} \\ e^*_\infty &= 0 \\ i^*_\infty &= 0 \\ r^*_\infty &= 1 + W(-s_\mathrm{t} R_0 \exp(-R_\mathrm{t}[1 - r_\mathrm{t}]))/R_\mathrm{t}\,, \end{aligned} \tag{22}$$

where $s_\mathrm{t} = s_{t^*+T/2}$ and $r_\mathrm{t} = r_{t^*+T/2}$ are the fractions of the susceptible and recovered populations at time $t = t^* + T/2$, the time at which the effective reproduction number has fully adopted the new value $R(t) = R_\mathrm{t}$. Importantly, this constrained equilibrium is not equivalent to the natural endemic equilibrium, $s_\infty \le s^*_\infty$ and $r^*_\infty \le r_\infty$, since $R_\mathrm{t} \le R_0$.

Time-varying effective reproduction number. Figures 8 to 10 illustrate the outbreak dynamics of our SEIR model with a time-varying effective reproduction number. The gray curves highlight the hyperbolic tangent type nature of the effective reproduction number $R(t)$, the dark red, red, orange, and blue curves illustrate the dynamics of the susceptible S, exposed E, infectious I, and recovered R populations. Unless stated otherwise, we use a latent period of $A = 2.5$ days, an infectious period of $C = 6.5$ days, a basic reproduction number of $R_0 = 4.5$, a reproduction number under public health interventions of $R_\mathrm{t} = 0.75$, and adaptation and transition times of $t^* = 20$ days and $T = 15$ days. In all simulations, the effective reproduction number $R(t)$ transitions gradually from the initial basic reproduction number R_0 at the beginning of the outbreak to the effective reproduction number R_t associated with the public health interventions. The adaptation time t^* marks the midpoint of the transition and the transition time T is its duration. The outbreak is more pronounced for larger basic reproduction numbers

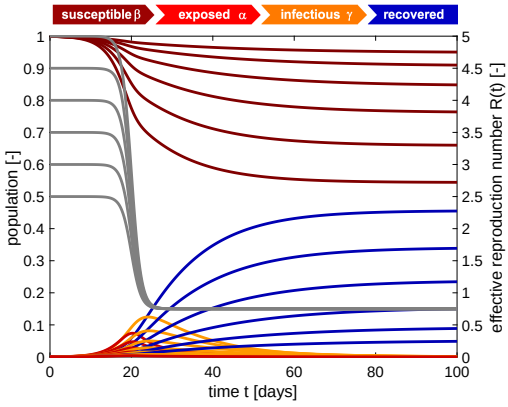

Fig. 8 SEIR model with time-varying effective reproduction number. Increasing the basic reproduction number R_0 increases the initial growth, and with it the number of cases. The temporary equilibrium for the smaller basic reproduction number of $R_0 = 2.5$ is $s^*_\infty = 0.948$ and $r^*_\infty = 0.052$ and for the larger basic reproduction number of $R_0 = 5.0$ is $s^*_\infty = 0.544$ and $r^*_\infty = 0.456$. Latent period $A = 2.5$ days, infectious period $C = 6.5$ days, basic reproduction number $R_0 = [5.0, 4.5, 4.0, 3.5, 3.0, 2.5]$, effective reproduction number $R_t = 0.75$, adaptation time $t^* = 20$ days, and transition time $T = 15$ days.

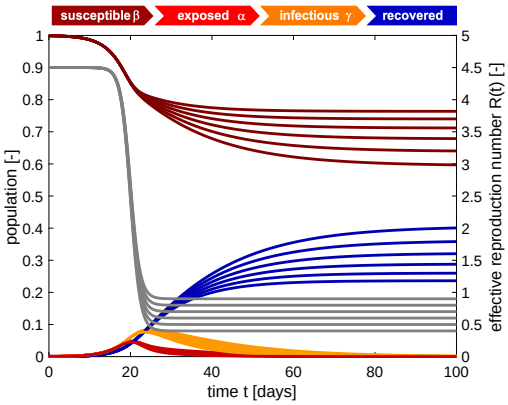

Fig. 9 SEIR model with time-varying effective reproduction number. Increasing the reproduction number R_t decreases the effect of interventions and increases the number of cases. The temporary equilibrium for the smaller effective reproduction number of $R_t = 0.4$ is $s^*_\infty = 0.764$ and $r^*_\infty = 0.236$ and for the larger effective reproduction number of $R_0 = 0.9$ is $s^*_\infty = 0.594$ and $r^*_\infty = 0.406$. Latent period $A = 2.5$ days, infectious period $C = 6.5$ days, basic reproduction number $R_0 = 4.5$, effective reproduction number $R_t = [0.4, 0.5, 0.6, 0.7, 0.8, 0.9]$, adaptation time $t^* = 20$ days, and transition time $T = 15$ days.

R_0 as we see in Figure 8, for larger intervention related reproduction numbers R_t as we see in Figure 9, and for larger adaptation times t^* as we see in Figure 10.

Constant, hyperbolic tangent, and random walk type effective reproduction numbers. To illustrate the effect of different time-varying effective reproduction numbers, we compare three different methods: a constant effective reproduction number, a smoothly decaying effective reproduction number of hyperbolic tangent type, and a daily varying effective reproduction number that follows a Gaussian random walk. The constant reproduction number has one parameter $R_t = R_0$. The hyperbolic tangent type reproduction number, $R_t = R_0 - \frac{1}{2}[1 + \tanh([t - t^*]/T)][R_0 - R_t]$, has four parameters, the basic and effective reproduction numbers R_0 and R_t, the adaptation time t^*, and the time delay Δt. The Gaussian random walk has three parameters, the drift μ, the daily stepwidth $\tau = \tau_1/[1.0 - s]$, and the smoothing parameter s. Table 3 summarizes the prior distributions for all three methods. Figure 11 compares the constant, hyperbolic tangent, and random walk type effective reproduction numbers for the example of Austria. The three graphs illustrate the number of reported cases as dots, the model fit as orange curves with 95% confidence interval, and the effective reproduction numbers as red curves with 95% confidence

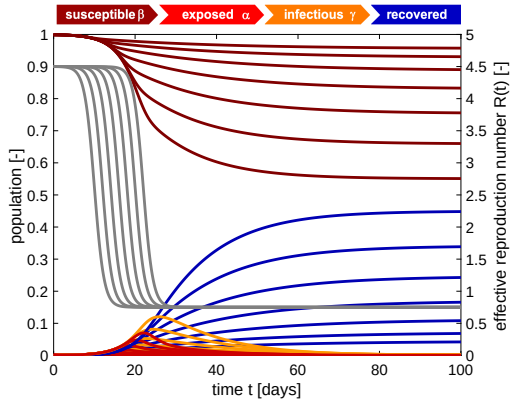

Fig. 10 SEIR model with time-varying effective reproduction number. Increasing the adaptation time t^* to interventions increases the time spent at a high reproduction number, and with it the number of cases. The temporary equilibrium for the faster adaptation of $t^* = 10$ days is $s^*_\infty = 0.956$ and $r^*_\infty = 0.044$ and for the slower adaptation of $t^* = 22$ days is $s^*_\infty = 0.550$ and $r^*_\infty = 0.450$. Latent period $A = 2.5$ days, infectious period $C = 6.5$ days, basic reproduction number $R_0 = 4.5$, effective reproduction number $R_t = R_0/6 = 0.75$, adaptation time $t^* = [10, 12, 14, 16, 18, 20, 22]$ days, and transition time $T = 15$ days.

Table 3 Prior distributions for time-varying effective reproduction number R(t) of constant, hyperbolic tangent, and Gaussian random walk type.

$R(t)$ ... constant	
Parameter	**Distribution**
R_0	Normal(2.5,2)

$R(t)$ ... hyperbolic tangent	
Parameter	**Distribution**
R_0	Normal(2.5,2)
R_t	Normal(2.5,2)
t^*	Normal(10,10)
T	LogNormal(log(3),1.5)

$R(t)$ ... Gaussian random walk	
Parameter	**Distribution**
$R(t)$	GRW($\mu, \tau_1/(1.0-s)$)
μ	Normal(0,2)
τ_1	Exponential(1/2)
s	Uniform(0,1)

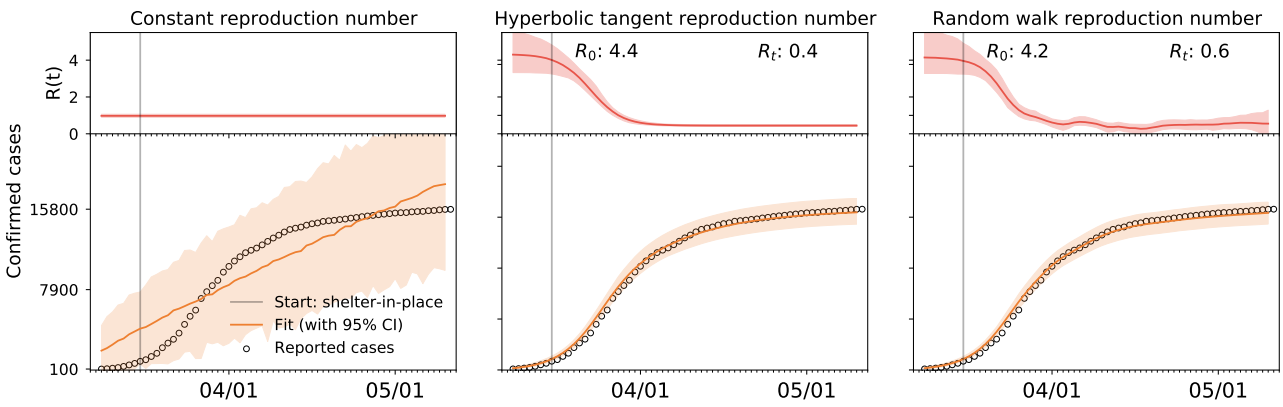

Fig. 11 Time-varying effective reproduction number R(t). Comparison of constant, hyperbolic tangent, and random walk type ansatz. The constant effective reproduction number predicts an exponential increase in the number of cases that fits the initial but not for the later stages of the COVID-19 outbreak, left. The hyperbolic tangent type reproduction number predicts a smooth early increase and later saturation of the number of cases, middle. The random walk type reproduction number predicts a daily varying, non-smooth early increase and later saturation of the number of cases, right. Dots represent reported cases; orange curves illustrate fit with 95% confidence interval; red curves shows effective reproduction number with 95% confidence interval; here illustrated for the case of Austria.

interval. Of all three methods, the constant ansatz can fit the early exponential increase of the COVID-19 outbreak, but not the later saturation. The random walk type ansatz can fit both the early exponential increase and the later saturation, but not with a closed form expression. Only the hyperbolic tangent type ansatz provides both a good fit and a closed functional form to compare the time lines of the outbreak in different countries and make informed predictions.

Herd immunity. An important consequence of the basic reproduction number R_0 is the condition for herd immunity [12]. Herd immunity occurs once the immune population, in our case the recovered population R, is large enough to protect susceptible individuals from infection [22]. We can express herd immunity in terms of the recovered fraction r, or in terms of the absolute recovered population R,

$$r > 1 - 1/R_0 \qquad \text{or} \qquad R > [1 - 1/R_0]\,N\,. \tag{23}$$

Importantly, upon relaxing public health interventions, the condition for herd immunity is not $R > 1 - 1/R_t$. Herd immunity is not a function of the reproduction number under public health interventions R_t–which is usually much smaller than the basic reproduction number R_0–but will depend on the natural basic reproduction number R_0 under unconstrained conditions.

Acknowledgments

We acknowledge stimulating discussions with Dr. Francisco Sahli Costabal. This work was supported by a DAAD Fellowship to Kevin Linka and a Stanford Bio-X IIP seed grant to Mathias Peirlinck and Ellen Kuhl and by the National Institutes of Health Grant U01 HL119578.

References

1. D. Ambrosi, M. BenAmar, C.J. Cyron, A. DeSimone, A. Goriely, J.D. Humphrey, E. Kuhl. Growth and remodelling of living tissues: Perspectives, challenges, and opportunities. J. Royal Soc. Interface 16 (2019) 20190233.
2. M. Alber, A. Buganza Tepole, W. Cannon, S. De, S. Dura-Bernal, K. Garikipati, G. Karniadakis, W.W. Lytton, P. Perdikaris, L. Petzold, E. Kuhl. Integrating machine learning and multiscale modeling: Perspectives, challenges, and opportunities in the biological, biomedical, and behavioral sciences. npj Digital Medicine 2 (2019) 115.
3. R. M. Anderson, R. M. May. Directly transmitted infectious diseases: control by vaccination. Science 215 (1982) 1053-1060.
4. Apple Mobility Trends. `https://www.apple.com/covid19/mobility`. accessed: June 25, 2020.
5. A. Arenas, W. Cota, J. Gomez-Gardenes, S. Gomez, C. Granell, J.T. Matamalas, D. Soriano-Panos, B. Steinegger. Derivation of the effective reproduction number R for COVID-19 in relation to mobility restrictions and confinement. medRxiv (2020) doi:10.1101/2020.04.06.20054320.
6. J. L. Aron, I. B. Schwartz. Seasonality and period-doubling bifurcation in an epidemic model. J. Theor. Bio 110 (1984) 665-679.
7. M. Biggerstaff, S. Cauchemez, C. Reed, M. Gambhir, L. Finelli. Estimates of the reproduction number for seasonal, pandemic, and zoonotic influenza: a systematic review of the literature. BMC Infectious Disease 14 (2014) 480.
8. M. Chinazzi, J. T Davis, M. Ajelli, C. Gioanni, ... A. Vespignani. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. Science (2020) doi:10.1126/science.aba9757.
9. D.K. Chu, E.A. Akl, S. Duda, K. Solo, S. Yaacoub, H.J. Schünemann. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: A systematic review and meta-analysis. Lancet 395 (2020) 1973-1987.
10. S. Cobey. Modeling infectious disease dynamics. Science (2020) doi:20.1126/science.aba5659.
11. J. Dehning, J. Zierenberg, F.P. Spitzner, M. Wibral, J. Pinheiro Neto, M. Wilczek, V. Priesemann, Inferring COVID-19 spreading rates and potential change points for case number forecasts. arXiv (2020) 2004.01105.

12. P. L. Delamater, E. J. Street, T. F. Leslie, Y. T. Yang, K. H. Jacobsen. Complexity of the basic reproduction number (R_0). Emerg. Infect. Disease 25 (2019) 1-4.
13. K. Dietz. The estimation of the basic reproduction number for infectious diseases. Stat. Meth. Med. Res. 2 (1993) 23-41.
14. P van den Diessche, J. Watmough. Reproduction numbers and sub-threshold endemic equilibria for compartment models of disease transmission. Math. Biosci. 180 (2002) 29-48.
15. European Centre for Disease Prevention and Control. Situation update worldwide. `https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases.` accessed: June 25, 2020.
16. European Commission. COVID-19: Temporary restriction on non-essential travel to the EU. Communication from the Commission to the European Parliament, the European Council and the Council. Brussels, March 16, 2020.
17. A. Erlich, D.E. Moulton, A. Goriely. Are homeostatic states stable? Bulletin Math. Bio. 81 (2019) 3219-3244.
18. Eurostat. Your key to European statistics. Air transport of passengers. `https://ec. europa.eu/eurostat`; accessed: June 25, 2020.
19. Eurocontrol. Flights 2020. Daily traffic variation. `http://eurocontrol.int`; accessed: June 25, 2020.
20. Y. Fang, Y. Nie, M. Penny. Transmission dynamics of the COVID-19 outbreak and effectiveness of government interventions: a data-driven analysis. J. Med. Virol. (2020) 1-15.
21. A.S. Fauci, H.C. Lane, R.R. Redfield. Covid-19–Navigating the uncharted. New Engl. J. Med. (2020) 382:1268-1269.
22. P. E. M. Fine. Herd immunity: history, theory, practice. Epidemiologic Reviews 15 (1993) 265-302.
23. S. Flaxman, S. Mishra, A. Gandy, H.J.T. Unwin, ..., S. Bhatt. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature (2020) doi:10.1038/s41586-020-2405-7.
24. J. Hellewell, S. Abbott, A. Gimma, N.I. Bosse, C.I. Jarvis, T.W. Russell, J.D. Munday, A.J. Kucharski, W.J. Edmunds. Feasibility of controlling COVID-19 outbreaks by isolation of cases and contacts. Lancet Global Health (2020) 8:e488-496.
25. H. W. Hethcote. The mathematics of infectious diseases. SIAM Review 42 (2000) 599-653.
26. M.D. Hoffman, A. Gelman. The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. J. Machine Learning Res. 15 (2014) 1593–1623.
27. J. Hsu. Here's how computer models simulate the future spread of new coronavirus. Scientific American (2020) February 23, 2020.
28. D. S. Hui, E. I. Azhar, T. A. Madani, F. Ntoumi, R. Kock, O. Dar, G. Ippolito, T. D. Mchogh, Z. A Memish, C. Drosten, A. Zumla, E. Petersen. The continuing 2019-nCoV epidemic thread of novel coronaviruses to global health - The latest 2019 novel coronavirus outbreak in Wuhan, China. Int. J. Infect. Diseases 91 (2020) 264-266.
29. W. O. Kermack, G. McKendrick. Contributions to the mathematical theory of epidemics, Part I. Proc. Roy. Soc. London Ser. A 115 (1927) 700-721.
30. K.L. Lange, R.J.A. Little, J.M.G. Taylor, Robust statistical modeling using the T distribution. J. Am. Stat. Ass. 84 (1989) 881-896.
31. S. A. Lauer, K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, A. S. Azman, N. G. Reich, J. Lessler. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. Ann. Int. Med. (2020) doi:10.7326/M20-0504.
32. M. Y. Li, J. S. Muldowney. Global stability for the SEIR model in epidemiology. Math. Biosci. 125 (1995) 155-164.
33. Q. Li, X. Guan, P. Wu, X. Wang, ... Z. Feng. Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. New Eng. J. Med. (2020) doi:10.1056/NEJMoa2001316.
34. K. Linka, M. Peirlinck, F. Sahli Costabal, E. Kuhl. Outbreak dynamics of COVID-19 in Europe and the effect of travel restrictions. Comp Meth Biomech Biomed Eng; in press. doi:10.1080/10255842.2020.1759560;
35. K. Linka, A. Goriely, E. Kuhl Global and local mobility as a barometer for COVID-19 dynamics medRxiv (2020) doi:10.1101/2020.06.13.20130658.
36. Y. Liu, A. A. Gayle, A. Wilder-Smith, J. Rocklöv. The reproductive number of COVID-19 is higher compared to SARS coronavirus. J. Travel Medicine (2020) doi:10.1093/jtm/taaa021.
37. B. F. Maier, D. Brockmann. Effective containment explains sub-exponential growth in confirmed cases of recent COVID-19 outbreak in mainland China. medRxiv (2020) doi:10.1101/2020.02.18.20024414.
38. B. Mason Meier, R. Habibi, Y. Tony Yang. Travel restrictions violate international law. Science 367 (2020) 1436.
39. G.C.Y. Pang, M. Alber, A. Buganza Tepole, W. Cannon, S. De, S. Dura-Bernal, K. Garikipati, G. Karniadakis, W.W. Lytton, P. Perdikaris, L. Petzold, E. Kuhl. Multiscale modeling meets machine learning: What can we learn? Arch. Comp. Meth. Eng. (2020) doi:10.1007/s11831-020-09405-5.
40. S. W. Park, B. M. Bolker, D. Champredon, D. J. D. Earn, M. Li, J. S. Weitz, B. T. Grenfell, J. Dushoff. Reconciling early-outbreak estimates of the basic reproductive number and its uncertainty: framework and applications to the novel coronavirus outbreak. medRxiv (2020) doi:10.1101/2020.01.30.20019877.
41. M. Peirlinck, F. Sahli Costabal, K.L. Sack, J.S. Choy, G.S. Kassab, J.M. Guccione, M. De Beule, P. Segers, E. Kuhl. Using machine learning to characterize heart failure across the scales. Biomech. Model. Mechanobio. 18 (2019) 1987-2001.
42. M. Peirlinck, K. Linka, F. Sahli Costabal, E. Kuhl. Outbreak dynamics of COVID-19 in China and the United States. Biomech. Model. Mechanobio. (2020) doi:10.1007/s10237-020-01332-5.
43. M. Peirlinck, K. Linka, F. Sahli Costabal, E. Bendavid, J. Bhattacharya, J.P.A. Ioannidis, E. Kuhl. Visualizing the invisible: The effect of asymptomatic transmission on the outbreak dynamics of COVID-19. medRxiv (2020) doi:10.1101/2020.05.23.20111419.
44. K. Prem, Y. Liu, A.J. Kucharski, R.M. Eggo, N. Davies. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modeling study. Lancet Public Health. doi:10.1016/S2468-2667(20)30073-6.
45. F. Sahli Costabal, K. Matsuno, J. Yao, P. Perdikaris, E. Kuhl. Machine learning in drug development: Characterizing the effect of 30 drugs on the QT interval using Gaussian process regression, sensitivity analysis, and uncertainty quantification. Comp. Meth. Appl. Mech. Eng. 348 (2019) 313-333.
46. J. Salvatier, T.V. Wiecki, C. Fonnesbeck. Probabilistic programming in Python using PyMC3. PeerJ Computer Science 2 (2016) e55.
47. S. Sanche, Y.T. Lin, C. Xu, E. Romero-Severson, N. Hengartner, R. Ke. High contagiousness and rapid spread of severe acute respiratory syndrome coronavirus 2. Emerg. Infect. Disease. (2020) doi:10.3201/eid2607.200282.
48. H. Sjödin, A. Wilder-Smith, S. Osman, Z. Farooq, J. Rocklöv. Only strict quarantine measures can curb the coronavirus disease (COVID-19) outbreak in Italy, 2020. Euro Surveill. 25 (2020) 2000280.
49. B. Tang, F. Xia, N.L. Bragazzi, Z. McCarthy, X. Wang, S. He, X. Sun, S. Tang, Y. Xiao, J. Wu. Lessons drawn from China and South Korea for managing COVID-19 epidemic: insights from a comparative modeling study. Bull World Health Organ (2020) doi:10.2471/blt.20.257238.
50. G. Viceconte, N. Petrosillo. COVID-19 R0: Magic number or conundrum? Infect. Dis. Rep. 12 (2020) 8516.
51. World Health Organization. WHO Virtual press conference on COVID-19. `https://www.who.int.docs/default-source/coronavirus/transcripts/who-audio-emergencies-corona-`

virus-press-conference-full-andfinal-11mar2020.pdf?sfvrsn=cb432bb3_2 recorded: March 11, 2020; accessed: June 25, 2020.

52. Wikipedia. 2020 Coronavirus pandemic in Europe. https://en.wikipedia.org/wiki/2020_corona-virus_pandemic_in_Europe; accessed: June 25, 2020.
53. A. Wilder-Smith, D.O. Freedman. Isolation, quarantine, social distancing and community containment: pivotal role for old-style public health measures in the novel coronavirus (2019-nCoV) outbreak. J. Travel Med. (2020) doi: 10.1093/jtm/taaa020.
54. A. Wilder-Smith, C.J. Chiew, V.J. Lee. Can we contain the COVID-19 outbreak with the same measures as for SARS? Lancet Infect. Dis. 20 (2020) e102-107.
55. J. Yuan, M. Li, G. Lv, Z.K. Lu. Monitoring transmissibility and mortality of COVID-19 in Europe. Int. J. Infectious Disease. doi:10.1016/j.ijid.2020.03.050.
56. S. Zhao, Q. Lin, J. Ran, S. S. Musa, G. Yang, W. Wang, Y. Lou, D. Gao, L. Yang, D. He, M. H. Wang, Maggie. Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. bioRxiv (2020) doi:10.1101/2020.01.23.916395.