PNAS RESEARCH ARTICLE | COMPUTER SCIENCES POPULATION BIOLOGY OPEN ACCESS

How to avoid a local epidemic becoming a global pandemic

Nils Chr. Stenseth[a,b,1,2], Rudolf Schlatte[c,2], Xiaoli Liu[d,2], Roger Pielke Jr.[b,e], Ruiyun Li[b], Bin Chen[f,g,h], Ottar N. Bjørnstad[b,i], Dimitri Kusnezov[j], George F. Gao[k,l], Christophe Fraser[m,n], Jason D. Whittington[a,b], Yuqi Bai[o,p], Ke Deng[q,r], Peng Gong[s,t], Dabo Guan[o,t], Yixiong Xiao[u], Bing Xu[o,p], and Einar Broch Johnsen[c,1,2]

Contributed by Nils Chr. Stenseth; received November 28, 2022; accepted January 10, 2023; reviewed by Henning S. Mortveit and Stig W. Omholt

Here, we combine international air travel passenger data with a standard epidemiological model of the initial 3 mo of the COVID-19 pandemic (January through March 2020; toward the end of which the entire world locked down). Using the information available during this initial phase of the pandemic, our model accurately describes the main features of the actual global development of the pandemic demonstrated by the high degree of coherence between the model and global data. The validated model allows for an exploration of alternative policy efficacies (reducing air travel and/or introducing different degrees of compulsory immigration quarantine upon arrival to a country) in delaying the global spread of SARS-CoV-2 and thus is suggestive of similar efficacy in anticipating the spread of future global disease outbreaks. We show that a lesson from the recent pandemic is that reducing air travel globally is more effective in reducing the global spread than adopting immigration quarantine. Reducing air travel out of a source country has the most important effect regarding the spreading of the disease to the rest of the world. Based upon our results, we propose a digital twin as a further developed tool to inform future pandemic decision-making to inform measures intended to control the spread of disease agents of potential future pandemics. We discuss the design criteria for such a digital twin model as well as the feasibility of obtaining access to the necessary online data on international air travel.

Disease X | epidemiology | data science | coupled simulation model | digital twin model

Significance

We contribute a proof-of-concept model to inform decision-making on how to avoid a local epidemic developing into a global pandemic by reducing international air travel worldwide, coupled with a compulsory immigration quarantine when traveling between countries. The work highlights a major innovation: replacing the historical air travel data and fixed parameter values of our case study with a digital-twin model that continuously incorporates a live feed of air travel data and improved model parameter estimates for any novel infection. This may facilitate the rapid analysis of effects of intervention measures as a local epidemic may escalate into a global pandemic, and thus slow or even stop the spread.

Cases of pneumonia of unknown etiology (PUE) were observed in Wuhan, China (1, 2), and reported to the WHO China Country Office on December 31, 2019 (3). On January 7, 2020, Chinese health officials confirmed that the PUE outbreak was caused by a novel coronavirus, later named Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2). On January 10, 2020, the whole-genome sequence was reported to the WHO and shared through the Global Initiative on Sharing All Influenza Data (Accession numbers EPI_ISL_402119 and EPI_ISL_402121) (2). At this point, it became clear to the scientific community that the Wuhan COVID-19 epidemic was potentially very dangerous (4, 5) and might develop into a global pandemic—as we now know it did.

Despite past experiences with outbreaks of SARS-COV-1, influenza A/H1N1/2009pwd, and West African EBOV 2014/15, a complete lockdown of Wuhan city on January 23 and Hubei province on January 24, 2020, and the announcement of COVID-19 as a global pandemic by the WHO Director-General on March 12 (6), the rest of the world had not begun to realize the threat (7). This slow global response highlights the current limitations in effective coordination between the scientific community and policy makers as mediated by scientific advisory mechanisms (4, 8, 9).

By January 29, 2020, basic epidemiological calculations using the Wuhan data indicated that the virus had a plausible supercritical reproduction number of 1.5 to 2 and a probable doubling time of 6 d, clearly showing early warning signs that COVID-19 might develop into a global pandemic (10). Finding ways to anticipate and clearly communicate to decision makers when and how a local endemic epidemic might develop into a future global pandemic is an important aspect of outbreak response. Here we present a framework to support decision-making to prevent future local epizootics from becoming global pandemics—or more specifically, delaying the global spread of a disease agent and thus facilitating the early termination of an emergent, potential global pandemic. For this purpose, we use the early period of the COVID-19 pandemic as a case study. However, our perspective is more general, going beyond the COVID-19 pandemic for better preparedness for the next "Disease X."

Emerging respiratory infectious disease agents like SARS-CoV-2 often disseminate geographically on varying length and timescales, from diffusion via local transit and mobility networks to large steps through international air travel. Here we focus on the latter to explore how well our methods capture broad features of the global impacts observed during periods of interest. In this context, one may think of three major clusters

Author contributions: N.C.S. and E.B.J. designed research; B.C. performed research; R.S., X.L., and R.L. analyzed data; and N.C.S., R.S., X.L., R.P., R.L., B.C., O.N.B., D.K., G.F.G., C.F., J.D.W., Y.B., K.D., P.G., D.G., Y.X., B.X., and E.B.J. wrote the paper.

Reviewers: H.S.M., University of Virginia; and S.W.O., Norges teknisk-naturvitenskapelige universitet.

The authors declare no competing interest.

[1]To whom correspondence may be addressed. Email: n.c.stenseth@ibv.uio.no or einarj@ifi.uio.no.

[2]N.C.S., R.S., X.L., and E.B.J. contributed equally to this work.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2220080120/-/DCSupplemental.

Published February 27, 2023.

of modeling-based research. The first integrates travel patterns to explain asynchrony in epidemic development. A number of studies have modeled pandemic dynamics (from influenza to COVID-19) in this way using a metapopulation approach with realistic travel patterns in multiple forms of transport (11–17). The second involves modeling nonpharmaceutical interventions (NPIs) to control or delay an epidemic. In the absence of effective prophylactic and therapeutic countermeasures, early studies of the COVID-19 pandemic have focused on how different NPIs, such as travel restrictions, social distancing, lockdown measures, and border controls, help suppress COVID-19 transmission in diverse geopolitical settings, such as China, UK, USA (18–22). The third involves agent- or metapopulation-based simulations for scenario exploration using real-time data and calibration. For example, agent-based models were used in the UK to characterize COVID-19 dynamics and different NPIs (23, 24). The National Institute of Public Health in Norway used both a metapopulation SEIR (Susceptible, Exposed, Infectious, and Recovered) model and an agent-based model to interpret early scenarios of COVID-19 development in Norway (25, 26). Such efforts may help to understand, interpret, and forecast pandemic dynamics and prioritize NPIs and collective strategies of transmission control. However, the development of a new comprehensive approach to investigate the *global* transmission dynamic and effectiveness of mitigation strategies for future emerging pandemics during the early stage to support decision-making when stakes are high and knowledge is uncertain, has received little attention.

A Cosimulation Model Captures the Early Phase of the Pandemic

In this study, we propose computational methods that may improve the ability to inform decision-making seeking to prevent future transmission events at the animal–human interface from developing into global pandemic regimes (27). We use COVID-19 as a case study and develop a coupled simulation model (or, for short, a cosimulation model) (28). Cosimulation is the joint simulation of loosely coupled stand-alone subsimulators. A cosimulation algorithm takes care of time synchronization and interactions across the subsimulators. The interactions between these subsimulators are only synchronized at discrete communication points. The cosimulation here couples an air travel model based on actual air travel passenger numbers (29) with a standard epidemiological model (30–32). The model is a metapopulation model (16, 18, 19, 33), in which the exchange of people between communities is based on real-time travel data. Parameter estimates are taken from the beginning of January 2020 and the cosimulation platform is used to project hypothetical "what-if" scenarios capturing the effect of global air travel regulations and use of border quarantine upon arrival. The cosimulation model uses daily intercountry air travel data,* combined with epidemiological models for the countries of origin and destination (Fig. 1). For a brief summary of some earlier relevant studies, see the *SI Appendix, Earlier studies of pandemic spread*.

The cosimulation model combines a standard epidemiological model for each country worldwide with a model of daily air travel between these countries. We demonstrate that this cosimulation model can replicate the dynamics of the COVID-19 pandemic well by comparing the simulated number of infections to the number of documented cases when we include the air travel information for January to March 2020.

*Passengers traveling from one country to any other country via third countries are currently not tracked, posing a limitation to studying the effects of infected passengers on third countries.

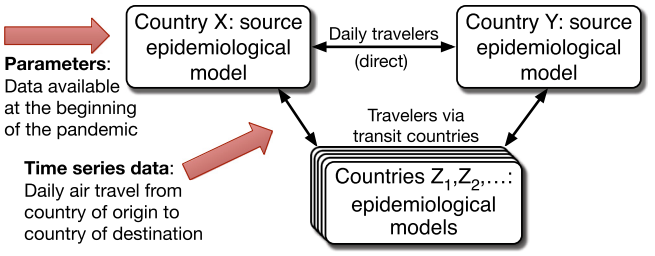

Fig. 1. Structure of the coupled-simulation model of the global spread of a disease agent.

All scenario simulations were started with an initial exposed population (E_i) of 0.000036 percent in China (corresponding to the recorded infection numbers 2 wk later), and no infections were assumed in the rest of the world. The coupled simulation model was projected for 91 d akin to the early COVID-19 spread from January 1 to March 31, 2020. In each country connected through air travel, the exposed population for each day was assumed affected by incoming exposed passengers from all over the world (weighted by country population size) with subsequent SEIR-like local daily spread.

Epidemiological Model. We designed a multicompartmental SEIR model (34) to describe endemic spread per country in the form

$$\begin{aligned}\frac{ds_j}{dt} &= b_j - \mu_j s_j - \beta_j s_j i_j,\\ \frac{de_j}{dt} &= \beta_j s_j i_j - (\mu_j + \sigma_j) e_j,\\ \frac{di_j}{dt} &= \sigma_j e_j - (\mu_j + \gamma_j) i_j,\\ \frac{dr_j}{dt} &= \gamma_j i_j - \mu_j r_j,\end{aligned} \qquad [1]$$

where s_j, e_j, i_j and r_j denote the fraction of susceptible, exposed, infected, and recovered individuals in a country j, respectively. The total fraction of population in a country j is $s_j + e_j + i_j + r_j = 1$. Eq. **1** describes the evolution of the fraction set of susceptible individuals, exposed, infected, and recovered individuals over time, respectively. In the equations, b_j is the population birth rate, μ_j is the natural death rate, β_j is the rate of transmission, $\frac{1}{\sigma_j}$ is the average duration of latent or exposed period, and $\frac{1}{\gamma_j}$ represents the average duration of infection in a country j. In this study, we assume that the local epidemics in all countries follow the same dynamics, so $\beta_j = \beta$, $\sigma_j = \sigma$, and $\gamma_j = \gamma$.

Model Parameter Selection. The parameters for the SEIR model were chosen to roughly correspond to the observed characteristics of the original COVID-19 strain globally. Guan et al. (35) has pointed out that the median SARS-CoV-2 incubation period is 4 d with interquartile range 2.0 to 7.0. Exposed period is shorter compared to incubation period. Hence, we selected the exposed days that fall within the interval [1, 6]. He et al. (36) estimated the basic reproductive number R_0 of SARS-CoV-2 to be 3.15. Liu et al. (37) carried out a review on the R_0 of the COVID-19 virus in 12 studiespublished from January 1 to February 7, 2020, and found that estimated R_0 ranged from 1.5 to 6.68 with a final mean value of 3.28. Ke et al. (38) reported that the mean value of R_0 was 5.8 in the US and that R_0 ranged from 3.6 to 6.1 in eight countries in the EU during the earlier stage of COVID-19. Therefore, we

selected the parameter values that made R_0 fall within the interval [2, 7]. In line with our focus on the early stage of the pandemic (from January 1, 2020 to the end of March 2020), we assumed the transmission rate to be 0.5, and the average duration of exposed and infection periods to be 2.5 and 9 d, respectively (39–41). Because the dynamic behavior of the birth and natural death processes are slower than those of the dynamics of the epidemic process, we set $b_j = 0$ and $\mu_j = 0$. Thus, together with normalized local population sizes for computational convenience, we used a transmission rate (β) of 0.5, an average duration of latency ($1/\sigma$) of 2.5 d and an average duration of infection ($1/\gamma$) of 9 d, corresponding to a reproduction number $R_0 = 4.5$ using the formula $R_0 = \frac{\sigma}{\sigma+\mu} \cdot \frac{\beta}{\gamma+\mu}$ (42).

We applied the model to all countries in the air travel passenger dataset using uniform model parameters. We calculated the correlation between the simulated levels and the daily country-level confirmed COVID-19 cases [from the data repository of the Center for Systems Science and Engineering at Johns Hopkins University (43)] to carry out the analysis.

The Air Travel Data and Model. We collected the monthly global flight and passenger data from January to March 2020, from the International Air Transportation Association (IATA) database (https://www.iata.org). The IATA database contains information on flights between 4,418 commercial airports worldwide and can be expected to have a 100% coverage of the global airline market. To simulate the spatiotemporal dynamics of early COVID-19 transmission on a global scale, we first aggregated the airport-level flight data into country-level origin–destination passenger flows. To minimize the number of assumptions and parameters, we evenly distributed monthly passengers over each month instead of accounting for weekend travel.

Given the number of passengers traveling from country k to country l on day t, represented as $p(k, l, t)$, and the current fraction of exposed individuals ($e_k(t)$) in each origin country k on day t, the travel data model calculates the changes in the number of exposed passengers entering country l as:

$$n_l^e(t) = \sum_{k \neq l} p(k, l, t) e_k(t). \quad [2]$$

This model assumes that the passengers from a country k form a representative sample of the whole population of country k and that symptomatic people do not travel. Note that the global spread is sensitive to how the air travel model handles very small numbers. We assumed that there cannot be less than one exposed passenger disembarking from an airplane. A consequence of this assumption is that if the representative sample from the SEIR model of country k suggests that 0.2 passengers traveling from country k to country i on a given day would be exposed, we reduce the number of exposed passengers to 0.

Coupled-Simulation Model. We combined the epidemiological model and the air travel model into a cosimulation model to conduct experiments that capture the effects of air travel on the virus spread, based on the worldwide time series travel data and the per country epidemiological models. The epidemiological model and travel data for each country l were integrated as follows. First, we used the epidemiological model to predict the local disease spread (i.e., the change in the fraction of susceptible, exposed, infected and recovered individuals) within country l in 1-d intervals to update the proportion of individuals in each epidemic state $s_l(t)$, $e_l(t)$, $i_l(t)$, $r_l(t)$. Then after each day, we computed the change in the number of passengers in each epidemic state caused by traveling for country l using the air travel model and then divided the population size to update the fraction of passengers in each epidemic state for each country l as

$$e_l(t+1) = e_l(t) + \alpha_l \frac{n_l^e(t)}{N_l(t)}, \quad [3]$$

where α_l describes to what extend the exposed passengers traveling to country l could finally enter country l and α is affected by different countries' policies and $N_l(t)$ represents the total population in country l at day t. For example, if country l adopts 10% quarantine measure and no other measures are adopted in other countries, then in this situation, the value of α_l is 0.9. In order to limit the parametric complexity, we did not track the total population changes in each country, which means that $N_l(t)$ remains fixed and does not change according to time. Meanwhile, we assumed that the fraction changes of infected and recovered from day t to day $t+1$ in a country l was only affected by local epidemic models described by Eq. **1**. We tracked the fraction changes in the number of susceptible and exposed individuals, and to keep the whole population equal to 1, the fraction of susceptible individuals was updated as:

$$s_l(t+1) = s_l(t) - \alpha_l \frac{n_l^e(t)}{N_l(t)}. \quad [4]$$

This coupling of the epidemiological and air travel models assumed that the increase and decrease in population size due to plane travel was irrelevant; hence, the numbers of susceptible, exposed, infected, and recovered individuals could be given as fractions of the whole population such that $s_l(t) + e_l(t) + i_l(t) + r_l(t) = 1$. To calculate the next day's disease spread, we started with $s_l(t+1)$, $e_l(t+1)$, $i_l(t+1)$, $r_l(t+1)$ and repeated the process by firstly using the per-country epidemiological models with updated numbers of susceptible, exposed, infected, and recovered individuals and then the air travel passengers as described above. This kind of setting would affect the results. There was difference between the proposed methods compared with continuously updating $N_l(t)$ owning to travel, but it did not greatly affect the final analysis.

The overall model was implemented in the Julia programming language and the differential equations for each day were solved using the solver for ordinary differential equations provided by the Differential Equations package. The simulation results were stored in a local SQLite database, and further analysis and plots were carried out in the Python programming language. Our implementation as well as the source data for our cosimulation case study are available as an open source, online artefact.[†]

Model Fit. The first comparison is between the cosimulation model and the reported cases of COVID-19 infections from January 22 to March 31, 2020. For 150 countries with simulated cases, the mean correlation was 0.92 (SD 0.09, 25th quartile 0.89, 75th quartile 0.98). Fig. 2 shows the correlation per country (a ranked list of these countries is presented in *SI Appendix*, Figs. S1 and S2). The model projection did not predict any cases for the 16 countries shown in grey color in Fig. 2. This is due to

[†]Zenodo model is available at https://doi.org/10.5281/zenodo.7472836.

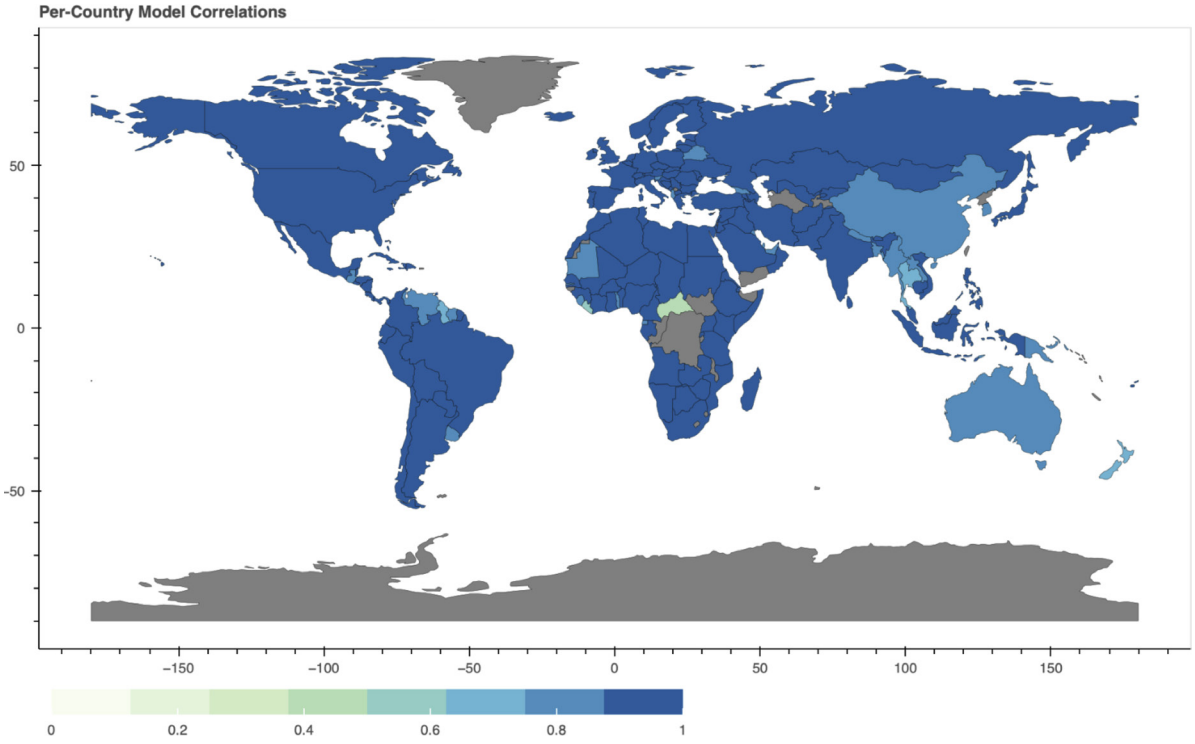

Fig. 2. Similarity between simulated and reported real cases per country during the period January 21, 2020 to March 31, 2020. The color gray denotes countries for which the model recorded no cases. See *SI Appendix*, Fig. S1 for an alternative plot, and *SI Appendix*, Fig. S2 for another similarity measure, both based on the same underlying data as this figure.

the low number of travelers and the expected incoming exposed travelers below one among the travelers from source countries on a given day. A notable case of low correlation between model predictions and the observed number of cases is South Korea, where an early cluster (stemming from the so-called Patient 31) led to early growth in the observed infections (44).

Across countries, the similarity between projected and documented cases varied over time (Fig. 3). The model/data similarity increased through the initial phase due to more data availability reducing uncertainty. The baseline similarity decreases as interventions are introduced in different countries, such as earlier quarantine in Australia and Vietnam and the start of lockdowns in some countries such as Norway (45), making the uniform SEIR (i.e., with no implemented interventions) parameters used by the model less reflective of reality.

The overall high degree of similarity between simulations and recorded cases demonstrates that the main assumption of spread by air travel and local propagation is captured by the cosimulation model, and thus may plausibly reflect the spread of infections with similar parameters in future outbreaks. Fig. 3 shows that the cosimulation model with fixed parameters correlates well with the reported cases of the COVID-19 pandemic until interventions were introduced in various countries.

The Cosimulation Model Can Be Used to Anticipate the Combined Effects of Early Interventions and thus Serve as a Tool for the Exploration and Appraisal of Policy Alternatives. Having first assessed the correlation between the model predictions and the observed infections, as reported by Johns Hopkins University (43), we then explored a range of possible what-if scenarios in the cosimulation model by reducing different fixed percentages for the number of passengers and the number of incoming exposed passengers per country, corresponding to interventions that reduce air travel and immigration quarantine rates. The effect of each hypothetical scenario is measured as the mean difference in the onset date of the national epidemics between the original predictions (without interventions) and the modified model (with the specified interventions). We define a country's onset date as the first day when the simulation reaches an infection fraction $I \geq 0.0001$ in that country. Our results show that use of such a cosimulation model has the potential to inform real-world policy to quickly evaluate alternative mitigation interventions and understand the potential efficacy of interventions on the global spread of a disease agent.

The cosimulation results demonstrated that 10 d after the outbreak, the model realistically reflects infection numbers as compared to official data. This makes it possible to perform hypothetical computational simulations. As an example, we consider the effects of reducing international air travel as well as implementing immigration quarantine-interventions for various countries (46–48).

In the first (and main) scenarios, we consider reduced air travel and introduced immigration quarantine restrictions (details in the *SI Appendix*) as follows:

- starting on January 21 we reduce air traffic globally by 10 to 100%, in increments of 10%; and
- starting on January 21 we reduce globally the number of cases coming into a country through quarantining by 10 to 100%, in increments of 10%.

In these projections, the interventions are introduced 10 d after January 11 as this was the day the scientific community realized that SARS-CoV-2 was potentially a very dangerous disease agent, and Wuhan's lockdown was initiated (18).

Fig. 4 shows the difference between the predictions without interventions and the modified model with the two sets of what-if interventions measured as the mean expected delay between the original and the experimentally modified simulations. The scenario analysis shows that, in general, reduction of air travel is more effective than immigration quarantine.

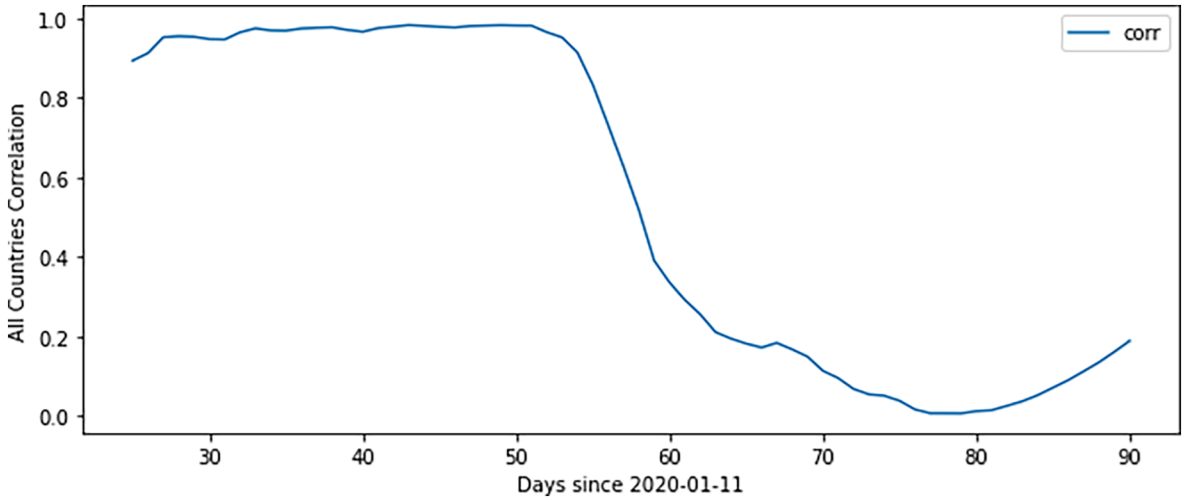

Fig. 3. Global correlations between measured and simulated cases for each day from January 25; see *SI Appendix*, Fig. S3 for a wider date range.

Table 1 presents the global effect of the reduction in international air travel and the increased use of immigration quarantine. The reduction in air travel (starting on January 21, 2020) is significantly more effective in reducing the spread of the disease agent than intensifying the use of immigration quarantine (also starting on January 21) as the absolute value of the coefficient of variable reduction in air travel (14.30) is larger than the absolute value of the coefficient of variable increased use of quarantine (7.68). This result is consistent with Fig. 4 where the gradient in the direction of reduction owning to air travel is sharper than that owning to quarantine.

A second set of analyses show the modeled effect of completely stopping air travel out of the country of origin (in our proof-of-concept case study, China) starting on January 2, 2020, through January 10, 2020, assuming no other interventions were carried out elsewhere in the world. It turns out this analysis is highly sensitive to initial conditions and is unrealistic in that the model assumes that all cases are localized in the initial country at the start of the experiment. Nevertheless, the results indicated that completely stopping air travel out of the country of origin immediately after the time when an outbreak is discovered would have an effect in delaying a local epidemic becoming a global pandemic (Fig. 5).

We also investigated the effect of combining a total air travel cessation from the country of origin with the scenario of the first experiment (global travel and quarantine restrictions of varying degrees). Fig. 6 shows the effect of hypothetically stopping all air

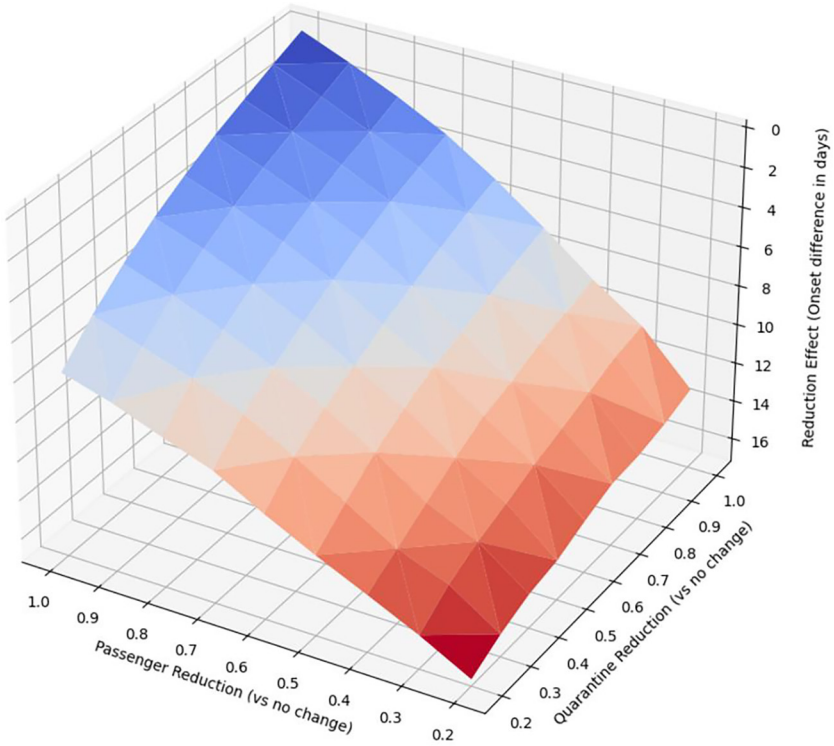

Fig. 4. Effect of reduction in air-travel and immigration quarantine on the Covid-19 epidemic, measured as mean value of the difference in epidemic onset between the simulation of the original model and that of the modified model corresponding to the what-if scenario across all countries. For the passenger reduction axis, 0.2 means air travel is reduced to 20%, whereas 1 means no reduction at all. For the quarantine reduction axis, 0.2 means 20% of exposed travelers arriving at a given country enter that country, whereas 1 means that there is no immigration quarantine intervention in any county. The effect axis is a measure of the degree to which the interventions (and combinations) can delay spread of the disease agent.

Table 1. Regression of the effect of the interventions (effect)—on the reduction in air travel globally and increased quarantine globally

					CI	
	Coef	SE	T	P > |t|	[0.025	0.975]
Intercept	21.49	0.26	82.3	0.000	20.97	22.01
Reduction in air travel	−14.305	0.29	−48.61	0.000	−14.89	−13.72
Increased use of quarantine	−7.682	0.29	−26.11	0.000	−8.27	−7.10

The intercept represents the maximum difference in a comparison of the original model (without interventions) with the modified model representing the what-if scenarios (with interventions). The overall R^2 is 0.975 (the adjusted R^2 is 0.974).

travel out of the country of origin on January 21, 2020, when combined with global air travel restrictions, again measured as the mean difference in epidemic onset. The z-axis value with ($X = 1$, $Y = 1$), which represents the effect of a 100% air travel stoppage out of China on January 21, 2020, without taking any other measures, shows that completely stopping air travel cannot prevent a global pandemic (when compared with Fig. 5), but it may delay global pandemic emergence; the greater the delay in air travel restrictions, the less the effect. The difference between Figs. 4 and 6 shows that the effect of completely stopping air travel from the country of origin decreases with generally more stringent air travel and quarantine restrictions worldwide.

Table 2 shows the global effect of a reduction in international air travel and the increased use of immigration quarantine plus completely stopping outgoing air travel from the country of origin on January 21, 2020. By comparing the results with Table 1, we project that completely stopping outgoing travel from the country of origin would slow down the effect of a reduction in air travel and increased use of quarantine. In Fig. 6, the Z-values at ($X = 0.2$, $Y = 0.2$), based on the regression models in Tables 1 and 2, are 17.09 and 18.13, respectively, which means that completely stopping outgoing travel from the country of origin has a negligible effect in such a situation.

In a third set of analyses, we reduced the transmission rate (β) between people within the country of origin of the pandemics (in this case China) by 20% (making R_0 equal to 4 rather than 4.5) for different combinations of immigration quarantine interventions and air travel reductions. As can be seen in Fig. 7, such reduction in human-to-human transmission has a major effect when combined with air travel restrictions and immigration quarantine.

Summary of the Findings. Using a simple cosimulation model that accurately replicated the observed emergence of the COVID-19 pandemic, we considered three forms of travel interventions in what-if scenario analyses: reducing the number of exposed people entering a country through i) reducing air travel; ii) implementing strict immigration quarantine interventions; and iii) completely stopping outgoing travel from the country of origin.

We have demonstrated that:

1. The cosimulation model well captured the observed global dynamics in the spread of a disease agent.
2. Reducing the number of flights globally was significantly more effective than global immigration quarantining in flattening the curve of infected people globally.
3. Increasing the effect of physical distancing and other human-to-human transmissions in the source country (only), combined with a reduction in air travel and immigration quarantining, had a larger effect in delaying the infection curve.

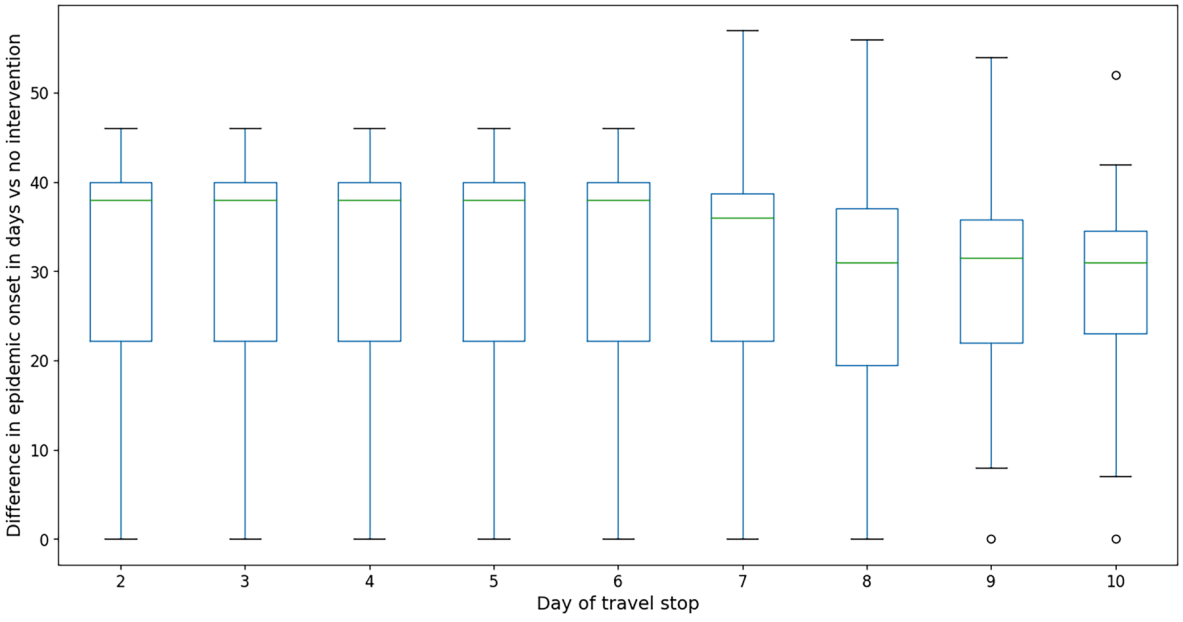

Fig. 5. Effect of 100% stop of air travel out of China on day 2 to 10 after the initial outbreak with no global interventions, measured in difference of onset days versus unmodified simulation.

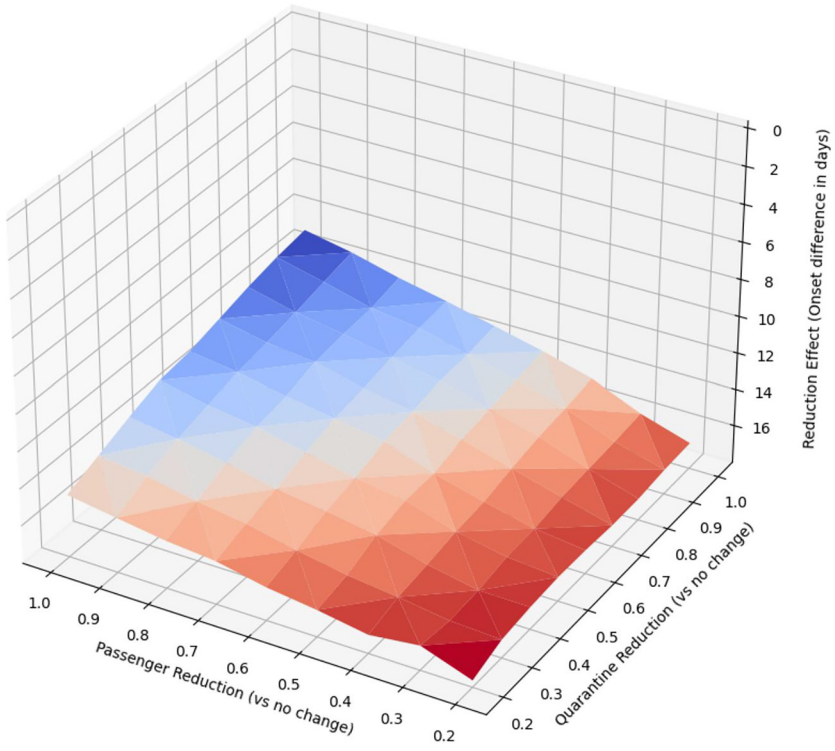

Fig. 6. Combined effect of completely stopping air-travel on January 21 from the country of origin combined with reduction in air-travel and immigration quarantine on the Covid-19 epidemic, measured as mean value of the difference in epidemic onset between the simulation of the original model and that of the modified model corresponding to the what-if scenario across all countries. Parameters on the X and Y axis have the same meaning as in Fig. 4.

It should be noted that immigration quarantine interventions are locally controlled, as is air travel out of (and into) a given country; however, global air travel is under international control.

The cosimulation model makes several simplifying assumptions and could thus be refined to provide even more precise estimates. For example, the SEIR model could be improved by accounting for age groups and relating those to air travel passenger statistics. The SEIR model could be made significantly more precise by estimating dynamic parameter changes for parameters in the local SEIR models; for example, machine learning techniques could be used to estimate the effect of introducing local interventions in given countries by combining both measured and simulated data.[‡] The air travel model could be improved by distinguishing weekend and holiday travel from business travel, using actual daily travel numbers, or by attempting to capture travel at the daily passenger level. (In contrast, the air travel model used above distributes monthly air travel data equally across all days in a month.) Interventions could further be regulated by the availability of local resources such as immigration quarantine capacities. Either or both the SEIR and air travel models could be replaced by stochastic models to simulate the likelihood of slowing or stopping a pandemic outbreak using travel restriction measures; the presented model is deterministic and hence will always model a global outbreak.

[‡]Our cosimulation model does not incorporate local interventions within any countries and our ability to predict further outbreak development could be significantly improved by introducing local dynamic parameters instead of using fixed global parameters. This is not important for our simulations but would certainly be critical for developing mechanisms for preventing (or delaying) a local epidemic developing into a global pandemic.

How Should a Digital Twin of a Global Epidemic Process Be Designed?

From a Cosimulation Model to a Digital Twin. A further developed cosimulation model for COVID-19 could be used as a computational framework to evaluate alternative intervention strategies for dealing with subsequent epidemic waves in the current COVID-19 pandemic caused by newly emerging SARS-CoV-2 variants. Such elaborations of the framework could highlight intervention strategies to inform efforts seeking to prevent future local epidemics from developing into a global pandemic. For this purpose, we need to generalize the cosimulation model from a model specific to the initial outbreak of COVID-19, which was based on specific parameter values for the epidemiological model and historical air travel data for the period January 22 to March 31, 2020. We now discuss how the cosimulation model can be turned into what is referred to as a *digital twin* (49).

A digital twin is a digital model of an underlying physical system, often called the physical twin. The digital twin is directly connected to the physical twin through streams of observations; this turns the digital twin into a live replica of the physical twin, able to provide insights into the dynamics of the physical twin in near real time (50). By replacing historical travel data and fixed parameter values of the current cosimulation model of the COVID-19 case study with continuous data assimilation, a digital twin will enable the real-time investigation of the potential effects of various alternative policy interventions.

To support the development of intervention strategies, the digital twin would require a more advanced model to further increase its precision and to expand the range of experiments that

Table 2. Regression of the effect of the interventions (effect) on the reduction in air travel globally and increased quarantine globally plus completely stopping outgoing air travel from the country of origin on January 21, 2020

					CI	
	Coef	SE	T	$P > \|t\|$	[0.025	0.975]
Intercept	19.900	0.14	138.26	0.000	19.61	20.19
Reduction in air travel	−5.598	0.16	−34.50	0.000	−5.92	−5.28
Increased use of quarantine	−3.232	0.16	−19.92	0.000	−3.56	−2.91

The overall R^2 is 0.95 (the adjusted R^2 is 0.95).

can be carried out, as well as to account for air travel in a general way, at any time when an epidemic occurs somewhere in the world. This can be done by continuously collecting and assimilating information from diverse data sources into the cosimulation model, such as worldwide air travel data at the present time. Furthermore, the parameter values of the epidemiological model would not be well-known at the time of an outbreak and would need to be refined continuously as our understanding of the outbreak improves. The following are the most important further developments needed:

1. Link the model with near real-time air travel information (such as daily passenger numbers reflecting travel at the geographical resolution of the model), including a digital twin architecture able to dynamically handle the inevitable instabilities and changes in the available data streams, integrate their diverse data representation formats and assess their integrity with respect to manipulation;
2. Increase the geographical resolution of the cosimulation model beyond the country level (e.g., to the level of large airports) and support multiscale cosimulation models;
3. Increase the detail of the epidemiological model (e.g., by including the demographic structure of populations and age-structured social interactions);
4. Allow a hierarchical model structure with a within-country metapopulation model with domestic travels within the country, combined with a global metapopulation model of international travel;
5. Update the parameters estimates of the local epidemiological models using feedback loops on a daily basis, to take maximal advantage of the increased amount of data pertaining to the epidemiological dynamics (both relating to the different geographic locations as well as relating to the air travel); and
6. Support what-if scenarios based on the cosimulation of heterogeneous models such as experiments that combine refined, stochastic models with the basic deterministic model to calculate the likelihoods of various outcomes.

The next emerging pandemic might have a very different infection pattern compared to SARS-COV-2; e.g., only infecting particular population groups. It is therefore essential to construct a digital twin that can flexibly reflect diverse infection patterns. Although the cosimulation architecture implemented in this paper combines SEIR compartment models for each country in the world with real passenger data from air travel between these countries, the architecture we consider does not require a particular compartment model. In a further developed digital twin model, different epidemiological models might be integrated on

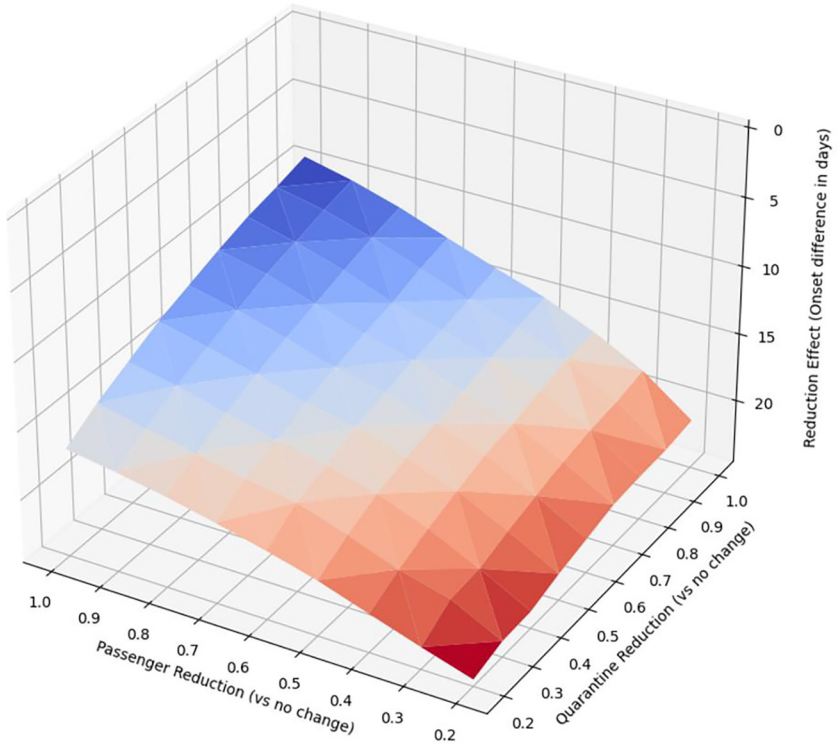

Fig. 7. Additional effect of reducing human-to-human transmission (β) within the source country (in this case China) by 20%. The effect axis is a measure of the degree to which the interventions (and combinations) can delay the spread of the virus. Note that the scale on the z-axis is different from that in Fig. 4; the meaning of the parameters in the X and Y axes is the same as in Fig. 4.

a per-country basis, depending on the availability of information (or preexisting models) and on the infection patterns of the disease agent. It would be worth considering the integration of local metapopulation compartment models that cover different regions within a country and how people travel between these areas [in particular, travel by car or train can be tracked by, e.g., mobile phone mobility data (51, 52)] as well as different age groups within a country. It would also be worth integrating agent-based models with compartment models. Both would be possible within a digital-twin architecture, provided that the models permit an interface that allows us to extract predictions regarding the exposure of the population to the disease agent. In particular, a family of different epidemiological models per country may be explored within the digital twin architecture to compare outcomes and thereby determine the models with the best fit for an emerging epidemic. Nevertheless, it seems essential for a global digital twin to use as its baseline epidemiological model per country a very simple default model that is transparent and easy to maintain, such as the SEIR model considered in this paper.

The resulting digital twin would enable real-time investigation regarding the effect of travel measures, as an epidemic develops. Such a digital twin would have the potential to accurately anticipate how a local epidemic might develop into a global pandemic, thereby enlarging the toolbox available to policy makers seeking to mitigate the likelihood of future pandemics. The digital twin could be used to explore what-if scenarios to evaluate alternative interventions with respect to different control objectives (which again would vary between countries, depending on different biological, societal and economic factors). To refine such a digital twin towards increased accuracy, parameters must be evaluated and tuned in real time; challenges in estimating reproduction numbers in real-time are discussed by Pellis et al. (53, 54). The results of the cosimulation model proposed in the present paper demonstrate that such a project is possible and important. A detailed model of this type could be augmented on the back end to understand additional consequences, for example, the impacts of decisions on the ability to respond internationally, the economic and trade impacts that could arise in the near term, or potential national security or geopolitical implications. Whereas these types of questions live outside of the model, they are often central to the science–policy interface and the issues that are faced by key decision makers.

Demonstrating the capability to develop a digital twin for pandemic forecasting and policy intervention evaluation is just one step in the development of an operational forecasting capability (55–58). Experience in the development of numerical weather forecasting models and the well-established relationships between research and operational agencies can prove useful and may be drawn upon in the development of forecasting capabilities related to pandemics (59, 60). Open-source software projects and international standardization boards can provide guidance on the efficient governance and logistics for a long-term international digital twin development project, integrating contributions from a wide range of competences. The development of a successful prediction enterprise in support of decision-making requires a dual focus—on prediction as a product and also as an integrated process that includes research, communication, and utilization (61). Turning research potential into tools for decision makers will require attention to each element of this process. In spring 2022, the US Centers for Disease Control and Prevention announced the creation of a new Center for Forecasting and Outbreak Analytics (62) modeled on the US National Weather Service, illustrating the great potential for better connecting pandemic forecasting with policy.

Challenges to Obtain Online Global Access to Air Travel and Epidemiological Data. Even with a fully developed digital-twin framework, many difficulties remain. The continuous collection, integration and storage of travel and epidemiological data at a global scale introduces several technological challenges, both with respect to the persistence and changing formats of data, their integration and their integrity; e.g., the collected data or data streams may be subject to systematic distortion to influence decision-making. Furthermore, accessing individual travel data will require international agreements. To discuss such difficulties in detail is beyond the scope of this contribution. However, if we are to prevent local epidemics from developing into global pandemics in the future, the technological challenges need to be addressed and such international agreements need to be obtained.

Discussion

Even with the best preparations, it is not possible to completely eliminate the risk of a local epidemic spreading to become a global pandemic. However, the approach here of a cosimulation coupled to a digital twin can be used to quickly and effectively identify potential outbreaks and take measures to prevent them from spreading. This approach can also be framed to facilitate implementing quarantine measures, providing medical treatment to those who are infected, and implementing public health measures to prevent the further spread of the disease. Additionally, a global digital twin can be used to monitor the progress of a pandemic and to inform decisions about how to respond to it. This can be especially important in situations where a vaccine is not immediately available, as we may have to live with the pathogen for a long time. A global digital twin can help us to navigate our way through a global pandemic and reduce its impact on society. Developing such a modeling framework now is an important advance step toward being well prepared for the next Disease X pandemic.

Data, Materials, and Software Availability. All study data are included in the article and/or *SI Appendix*.

ACKNOWLEDGMENTS. This work is funded by the Research Council of Norway through the COVID-19 Seasonality Project 312740 (N.C.S.) and the SIRIUS Center for Research-driven Innovation 237898 (R.S.), and the Penn State University Seed-Funded COVID-19 Project (O.N.B.). This work was strengthened thanks to feedback from reviewers Henning S. Mortveit and Stig William Omholt, and also from Joakim Sundnes, and Sasu A.O. Tarkoma.

Author affiliations: [a]Center for Pandemics and One Health Research, Sustainable Health Unit (SUSTAINIT), Faculty of Medicine, Oslo 0316, Norway; [b]Centre for Ecological and Evolutionary Synthesis, Department of Biosciences, University of Oslo, Oslo 0316, Norway; [c]Department of Informatics, University of Oslo, Oslo 0316, Norway; [d]Department of Computer Science, University of Helsinki, 00560 Helsinki, Finland; [e]Department of Environmental Studies, University of Colorado Boulder, Boulder, CO 80309; [f]Future Urbanity & Sustainable Environment (FUSE) Lab, Division of Landscape Architecture, Faculty of Architecture, University of Hong Kong, Hong Kong 999077, China; [g]Department of Geography, Urban Systems Institute, University of Hong Kong, Hong Kong 999077, China; [h]HKU Musketeers Foundation Institute of Data Science, The University of Hong Kong, Hong Kong 999077, China; [i]Department of Biology, Center for Infectious Disease Dynamics, Pennsylvania State University, University Park, PA 16802; [j]Deputy Under Secretary, Artificial Intelligence & Technology Office, US Department of Energy, Washington, DC 20585; [k]Chinese Academy of Sciences Key Laboratory of Pathogen Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, China; [l]Chinese Center for Disease Control and Prevention, Beijing 102206, China; [m]Pandemic Sciences Institute, University of Oxford, Oxford OX3 7DQ, UK; [n]Big Data Institute, Nuffield Department of Medicine, University of Oxford, Oxford 0X3 7LF UK; [o]Department of Earth System Science, Tsinghua University, Beijing 100084, China; [p]Ministry of Education Ecological Field Station for East Asia Migratory Birds, Tsinghua University, Beijing 100084, China; [q]Center for Statistical Science, Tsinghua University, Beijing 100084, China; [r]Department of Industrial Engineering, Tsinghua University, Beijing 100084, China; [s]Department of Earth Sciences, University of Hong Kong, Hong Kong 999077, China; [t]The Bartlett School of Sustainable Construction, University College London, London WC1E 6BT, UK; and [u]Business Intelligence Lab, Baidu Research, Beijing 100193, China

1. W. Tan *et al.*, A novel Coronavirus genome identified in a cluster of pneumonia cases–Wuhan, China 2019–2020. *China CDC Wkly.* **2**, 61–62 (2020).
2. N. Zhu *et al.*, A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* **382**, 727–733 (2020).
3. World Health Organisation (WHO), Pneumonia of unknown cause–China. https://www.who.int/emergencies/disease-outbreak-news/item/2020-DON229 (Accessed 23 November 2022).
4. S. Camporesi, F. Angeli, G. D. Fabbro, Mobilization of expert knowledge and advice for the management of the Covid-19 emergency in Italy in 2020. *Humanit. Soc. Sci. Commun.* **9**, 1–14 (2022).
5. C. Wang, P. W. Horby, F. G. Hayden, G. F. Gao, A novel coronavirus outbreak of global health concern [Review of A novel coronavirus outbreak of global health concern]. *The Lancet* **395**, 470–473 (2020), https://doi.org/10.1016/S0140-6736(20)30185-9.
6. D. Cucinotta, M. Vanelli, WHO declares COVID-19 a pandemic. *Acta Biomed.* **91**, 157–160 (2020).
7. N. C. Stenseth *et al.*, Lessons learnt from the COVID-19 pandemic. *Front. Public Health* **9**, 694705 (2021).
8. B. M. Vallejo, R. A. C. Ong, OCTA as an independent science advice provider for COVID-19 in the Philippines. *Humanit. Soc. Sci. Commun.* **9**, 1–9 (2022).
9. N. Brusselaers *et al.*, Evaluation of science advice during the COVID-19 pandemic in Sweden. *Humanit. Soc. Sci. Commun.* **9**, 1–17 (2022).
10. Q. Li *et al.*, Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia. *N. Engl. J. Med.* **382**, 1199–1207 (2020).
11. L. A. Rvachev, I. M. Longini, A mathematical model for the global spread of influenza. *Math. Biosci.* **75**, 3–22 (1985).
12. A. Flahault *et al.*, Modelling the 1985 influenza epidemic in France. *Stat. Med.* **7**, 1147–1155 (1988).
13. C. E. Walters, M. M. I. Meslé, I. M. Hall, Modelling the global spread of diseases: A review of current practice and capability. *Epidemics* **25**, 1–8 (2018).
14. T. F. Menkir *et al.*, Estimating internationally imported cases during the early COVID-19 pandemic. *Nat. Commun.* **12**, 311 (2021).
15. F. Pinotti *et al.*, Tracing and analysis of 288 early SARS-CoV-2 infections outside China: A modeling study. *PLoS Med.* **17**, e1003193 (2020).
16. R. Li *et al.*, Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* **368**, 489–493 (2020).
17. I. I. Bogoch *et al.*, Pneumonia of unknown aetiology in Wuhan, China: Potential for international spread via commercial air travel. *J. Travel Med.* **27**, taaa008 (2020).
18. H. Tian *et al.*, An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science* **368**, 638–642 (2020).
19. M. U. G. Kraemer *et al.*, The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* **368**, 493–497 (2020).
20. N. G. Davies *et al.*, Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: A modelling study. *Lancet Public Health* **5**, e375–e385 (2020).
21. S. Lai *et al.*, Assessing the effect of global travel and contact restrictions on mitigating the COVID-19 pandemic. *Engineering* **7**, 914–923 (2021).
22. S. Flaxman *et al.*, Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* **584**, 257–261 (2020).
23. C. C. Kerr *et al.*, Covasim: An agent-based model of COVID-19 dynamics and interventions. *PLoS Comput. Biol.* **17**, e1009149 (2021).
24. R. Hinch *et al.*, OpenABM-Covid19-An agent-based model for non-pharmaceutical interventions against COVID-19 including contact tracing. *PLoS Comput. Biol.* **17**, e1009146 (2021).
25. S. Engebretsen *et al.*, A real-time regional model for COVID-19: Probabilistic situational awareness and forecasting. *PLOS Computational Biology* **19**, e1010860 (2023).
26. G. Storvik, *et al.*, A sequential Monte Carlo approach to estimate a time varying reproduction number in infectious disease models: the Covid-19 case. arXiv [q-bio.PE] (2022). https://doi.org/10.1101/2021.10.25.21265166v1.
27. J. O. Lloyd-Smith *et al.*, Epidemic dynamics at the human-animal interface. *Science* **326**, 1362–1367 (2009).
28. C. Gomes, C. Thule, D. Broman, P. G. Larsen, H. Vangheluwe, Co-simulation: A survey. *ACM Comput. Surv.* **51**, 1–33 (2018).
29. Global flight and passenger ticket database.The International Air Transport Association. https://www.iata.org/. Accessed 23 September 2022.
30. O. N. Bjørnstad, *Epidemics: Models and Data Using R (Use R!),* 2nd Ed. (Springer, 2022).
31. O. Diekmann, H. Heesterbeek, T. Britton, *Mathematical Tools for Understanding Infectious Disease Dynamics* (Princeton University Press, 2013).
32. M. J. Keeling, P. Rohani, *Modeling Infectious Diseases in Humans and Animals* (Princeton University Press, 2008).
33. J. T. Wu, K. Leung, G. M. Leung, Nowcasting and forecasting the potential domestic and international spread of the 2019-nCoV outbreak originating in Wuhan, China: A modelling study. *Lancet* **395**, 689–697 (2020).
34. J. L. Aron, I. B. Schwartz, Seasonality and period-doubling bifurcations in an epidemic model. *J. Theor. Biol.* **110**, 665–679 (1984).
35. W.-J. Guan *et al.*, Clinical characteristics of coronavirus disease 2019 in China. *N. Engl. J. Med.* **382**, 1708–1720 (2020).
36. W. He, G. Y. Yi, Y. Zhu, Estimation of the basic reproduction number, average incubation time, asymptomatic infection rate, and case fatality rate for COVID-19: Meta-analysis and sensitivity analysis. *J. Med. Virol.* **92**, 2543–2550 (2020).
37. Y. Liu, A. A. Gayle, A. Wilder-Smith, J. Rocklöv, The reproductive number of COVID-19 is higher compared to SARS coronavirus. *J. Travel Med.* **27**, taab124 (2020).
38. R. Ke, E. Romero-Severson, S. Sanche, N. Hengartner, Estimating the reproductive number R0 of SARS-CoV-2 in the United States and eight European countries and implications for vaccination. *J. Theor. Biol.* **517**, 110621 (2021).
39. K. Prem *et al.*, The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: A modelling study. *Lancet Public Health* **5**, e261–e270 (2020), 10.1016/S2468-2667(20)30073-6.
40. J. A. Backer, D. Klinkenberg, J. Wallinga, Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020. *Euro Surveill.* **25**, 2000062 (2020).
41. S. A. Lauer *et al.*, The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: Estimation and application. *Ann. Intern. Med.* **172**, 577–582 (2020).
42. O. Diekmann, J. A. Heesterbeek, J. A. Metz, On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* **28**, 365–382 (1990).
43. E. Dong, H. Du, L. Gardner, An interactive web-based dashboard to track COVID-19 in real time. *Lancet Infect. Dis.* **20**, 533–534 (2020).
44. E. Jeong, M. Hagose, H. Jung, M. Ki, A. Flahault, Understanding South Korea's response to the COVID-19 outbreak: A real-time analysis. *Int. J. Environ. Res. Public Health* **17**, 9571 (2020).
45. J. M. Brauner *et al.*, Inferring the effectiveness of government interventions against COVID-19. *Science* **371**, eabd9338 (2021).
46. N. Eichler *et al.*, Transmission of severe acute respiratory syndrome coronavirus 2 during border quarantine and air travel, New Zealand (Aotearoa). *Emerg. Infect. Dis.* **27**, 1274–1278 (2021).
47. M. P. Hossain *et al.*, The effects of border control and quarantine measures on the spread of COVID-19. *Epidemics* **32**, 100397 (2020).
48. N. Steyn *et al.*, Managing the risk of a COVID-19 outbreak from border arrivals. *J. R. Soc. Interface* **18**, 20210063 (2021).
49. B. R. Barricelli, E. Casiraghi, D. Fogli, A survey on digital twin: Definitions, characteristics, applications, and design implications. *IEEE Access* **7**, 167653–167671 (2019).
50. F. Tao, H. Zhang, A. Liu, A. Y. C. Nee, Digital twin in industry: State-of-the-art. *IEEE Trans. Ind. Inf.* **15**, 2405–2415 (2019).
51. S. Engebretsen *et al.*, Time-aggregated mobile phone mobility data are sufficient for modelling influenza spread: The case of Bangladesh. *J. R. Soc. Interface* **17**, 20190809 (2020).
52. S. Engebretsen *et al.*, Regional probabilistic situational awareness and forecasting of COVID-19. medRxiv [Preprint] (2021) https://doi.org/10.1101/2021.10.25.21265166 (Accessed 15 December 2022).
53. L. Pellis *et al.*, Estimation of reproduction numbers in real time: Conceptual and statistical challenges. *J. R. Stat. Soc. Ser. A Stat. Soc.* **185**, 112–130 (2022).
54. G. S. Blair, Digital twins of the natural environment. *Patterns (N. Y.)* **2**, 100359 (2021).
55. C. Rivers *et al.*, Using "outbreak science" to strengthen the use of models during epidemics. *Nat. Commun.* **10**, 3102 (2019).
56. D. B. George *et al.*, Technology to advance infectious disease forecasting for outbreak management. *Nat. Commun.* **10**, 3932 (2019).
57. S. Pollett *et al.*, Identification and evaluation of epidemic prediction and forecasting reporting guidelines: A systematic review and a call for action. *Epidemics* **33**, 100400 (2020).
58. S. Pollett *et al.*, Recommended reporting items for epidemic forecasting and prediction research: The EPIFORGE 2020 guidelines. *PLoS Med.* **18**, e1003793 (2021).
59. C. Rivers, D. George, How to forecast outbreaks and pandemics. *Foreign Affairs* (2020) (23 September 2022).
60. R. Pielke, R. E. Carbone, Weather impacts, forecasts, and policy: An Integrated Perspective. *Bull. Am. Meteorol. Soc.* **83**, 393–406 (2002).
61. R. A. Pielke Jr., D. Sarewitz, R. Byerly, "Decision making and the future of nature: Understanding, using, and producing predictions" in *Prediction: Science, Decision Making, and the Future of Nature*, D. Sarewitz, R. A. Pielke, R. Byerly, Eds. (Island Press, 2000), pp. 361–387.
62. CDC, Center for forecasting and outbreak analytics. https://www.cdc.gov/forecast-outbreak-analytics/index.html (Accessed 23 September 2022).