SUPPLEMENTARY MATERIAL

Appendix A: Parameter schema in our model

The main aim of our model is to encourage users to customize parameters according to hypothetical scenarios and simulate the epidemic transmission cases number when infectious cases are imported. We deployed a web service to tune these parameters, as shown in Supplementary Table S1.

We grouped the parameters into three types. C1 and C2 are the number of imported infectious seeds and the total number of people in the region, respectively. R1–R4 characterize the transmissibility of epidemic disease by virus variant, vaccine, and environment. P1–P7 characterize government policy to COVID-19, such as social distancing or testing, and P8 is the days of the virus spreading freely before policies are carried out.

Appendix B: Policy effectiveness on the epidemic

Government policies are listed as P1–P7 in Supplementary Table S1. We computed policy effectiveness with different methods between P1–P5 and P6–P7.

(a) Data collection and data pre-processing

We analyzed OxCGRT[1] data compiled by the University of Oxford and extracted each country's daily policy levels. We performed one-hot encoding on the data to accurately calculate the impact of each policy and divided each policy into different levels. For example, we encoded "P1 = 1" as "P1_0 = 0, P1_1 = 1, and P1_2 = 0." According to the encoded policies, we can calculate the weight of each level of the policy in part b.

With the daily reports generated from Johns Hopkins University[2], we collected daily cases of each country and used the time-dependent method from the "R0" package[3] to calculate the daily Rt sequence for each country.

We also collected data on the population and gross domestic product (GDP) for each country from the United Nations Statistics Division[4].

(b) Data merging

Interventions, Rt sequences, population, and GDPs should be merged to form the training data of our machine learning model. For each country, we merged the data to observe the daily corresponding encoded policy levels, GDPs, population, and current Rt value.

(c) Training data generation

After merging the data, we only retained the days for countries that make changes on interventions. If a single country does not make policy changes on a single day, then data on that day will be deleted. For the reserved data, we also calculated the daily "Post_Rt" and "Rt_Ratio." "Post_Rt" is the current country's Rt value after 10 days, and "Rt_Ratio" is the value of "Post_Rt" over the current "Rt." We assume that after 10 days, the current interventions will be fully effective, and the impact could be expressed as "Rt_Ratio." We used the above methods to generate the training data for our model. The input of the model is the encoded policy levels, GDPs, and population for each day *X*. The predicted label of the model is the "Rt_Ratio" for day *X*.

(d) Time window selection

We used data in 2020 to independently calculate the effectiveness of interventions in Appendix A, which means considering the individual impact of variants and vaccines. During the whole year of 2020, the original strain was still the main source of transmission, and vaccines had not yet been released.

(e) Policy effectiveness of P1–P5

After mapping the essential policies to OxCGRT policies[5], we found that P1 (school closing), P2 (workplace closing), P3 (restrictions on gathering size), P4 (close public transport), and P5 (stay at home requirements) are the interventions that most countries have executed. Therefore, we used these policies in the simulation model.

With the training data mentioned in the above steps, we built a gradient regression tree model (Gradient Boost Regression Tree, *GBRT*)[6] to estimate the weight of each policy level.

In our analysis, we initialized the first tree for regression as follows:

$$F_0(x) = \underset{\gamma}{argmin} \sum_{i=1}^{n} L(y_i, \gamma),$$

Supplementary Table S1. Parameter settings in the model

ID	Name	Description	Measurement	Coding instructions	Facts of Xiamen outbreak
C1	Imported cases number	Number of imported cases	Integer number	-	1
C2	Population number	Population in the region	Integer number	-	5,163,970
R1	Virus variant	Type of virus variant	Discrete scale	- Original strain - Alpha strain - Beta strain - Gamma strain - Delta strain - Omicron strain	Delta(B.1.617.2)
R2	Diffusion environment	Type of diffusion environment	Discrete scale	- Hardest to spread the virus (E.g., Desert or grassland where is under-populated) - Hard to spread the virus (E.g., Village where people stay at home most time) - Normal environment (E.g., Normal living community) - Easy to spread the virus (Mass gathering) - Easiest to spread the virus (E.g., Confined, cold, humid environment)	Easiest
R3	Vaccine effectiveness	Effectiveness of the vaccine	Float value	-	30%
R4	Vaccination coverage	Percentage of people vaccinated	Float value	-	77%
P1	School closing	Closings of schools and universities	Ordinal scale	0 - No measures 1 - Require closing (only some levels or categories, e.g., just high school or just public schools) 2 - Require closing all levels	All universities, primary and secondary schools, and vocational schools have been adjusted to online teaching. Kindergartens have been suspended admission.
P2	Workplace closing	Closing of workplaces	Ordinal scale	0 - No measures 1 - require closing (or work from home) for some sectors or categories of workers 2 - require closing (or work from home) all-but-essential workplaces (e.g., grocery stores, doctors)	Government agencies and institutions at all levels implemented a flexible work system.
P3	Restrictions on gatherings	Cut-off size for bans on private gatherings	Ordinal scale	0 - No restrictions 1 - Restrictions on very large gatherings (the limit is above 1000 people) 2 - Restrictions on gatherings between 100–1000 people 3 - Restrictions on gatherings between 10–100 people 4 - Restrictions on gatherings of less than 10 people	All kinds of large-scale group dinner activities were suspended.
P4	Close public transport	Closing of public transport	Ordinal scale	0 - No measures 1 - Recommend closing (or significantly reduce volume/route/means of transport available) 2 - Require closing (or prohibit most citizens from using it)	Public transportation such as bus, BRT, and long-distance passenger transportation in the Tongan district was suspended.
P5	Stay at home requirements	Orders to “shelter-in-place” and otherwise confined to home	Ordinal scale	0 - No measures 1 - require not leaving the house with exceptions for daily exercise, grocery shopping, and “essential” trips 2 - Require not leaving the house with minimal exceptions (e.g., allowed to leave only once every few days, or only one person can leave at a time, etc.)	Residential communities implemented closed-loop management and reduced the flow of people.
P6	Quarantine high-risk people	Quarantine people with high-risk virus exposure	Ordinal scale	0 - No measures 1 - Quarantine close contacts and secondary close contacts of cases 2 - Quarantine visitors of high-risk regions 3 - Quarantine visitors of regions where cases have been found	Close contacts and secondary close contacts of cases were quarantined in designated places for 14 days.
P7	Time cost of testing for high-risk people	Time limit of testing for high-risk people	Ordinal scale	0 - No measures 1 - Finish testing for high-risk people in 5 days 2 - Finish testing for high-risk people in 3 days 3 - Finish testing for high-risk people in 1 day	The testing result was reported the next day of quarantine.
P8	Days of virus spread freely	Days of virus spread freely before policies carried out	Integer number	-	The first case was imported on September 6, 2021, and interventions were implemented approximately on September 14, 2021.

where $F_0(x)$ is the first weak learner of our tree model; y_i is the predicted value, which is the "*Rt_Ratio*" in our model; x is the input data; n indicates the total samples of our input data; L is the loss function to minimize the error of our predictions on "*Rt_Ratio,*"; and γ is the set of hyper parameters of the learner $F_0(x)$.

We then built the residual fitting of M trees for m = 1 to M and calculated the gradient direction of the loss function on the existing model as follows:

$$R_{im} = -\left[\frac{\partial L(y_i, f(x_i))}{\partial F(x_i)}\right]_{F(x)=F_{x-1}(x)}, \text{for } i = 1, \ldots, n.$$

Using the gradient vector and the previous sample points as the basis, we trained a weak classifier (basic GBRT) marked as $h_m(x_i)$ and added it to the entire model marked as $F_{m-1}(x_i) + \gamma \times h_m(x_i)$ to find the best weight. We used line search to find the weight of the classifier $h_m(x_i)$ and incorporated the new classifier into the loss function calculation to find the best weight γ_m:

$$\gamma_m = \underset{\gamma}{argmin} \sum_{i=1}^{n} L(y_i, F_{m-1}(x_i) + \gamma \times h_m(x_i))$$

Finally, we updated the entire model as $F_m(x) = F_{m-1}(x) + \gamma_m \times h_m(x)$.

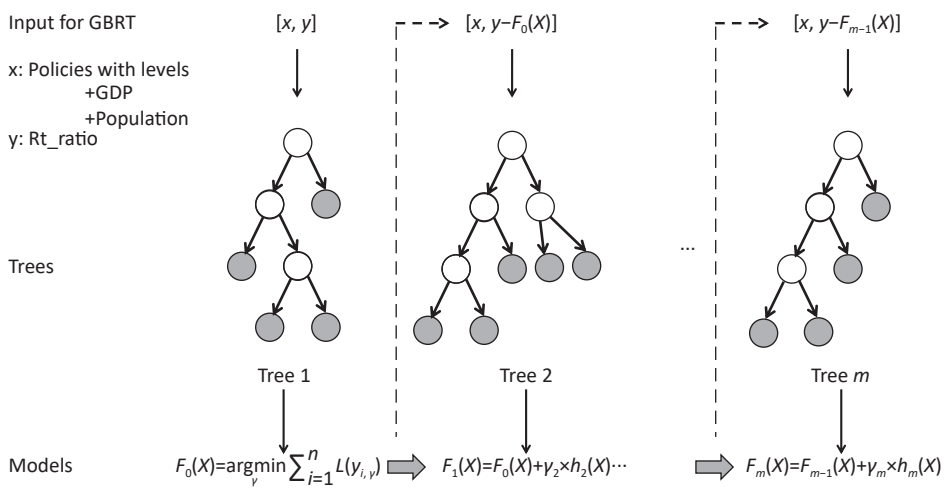
Supplementary Figure S1 provides a simplified explanation of the structure of the GBRT. The essence of GBRT is adding a set of tree models to build multiple weak learners. The purpose of the newly added tree is to fit the residuals, i.e., the data that cannot be properly fitted by the previous trees.

For the entire model $\mathrm{F_m(x)}$, we calculated the weights of the decrease in the purity of all non-leaf nodes when splitting. We then averaged the importance of each feature to all nodes to obtain the weight of each policy with a specific level marked as α.

For some nested policies, such as "Gathering limited to 10 people or less," the corresponding Oxford policy includes not only P3 = 4 but also P3 = 3 or P3 = 2.

(f) Policy effectiveness of P6–P7

P6 and P7 are the most effective policies in recent outbreaks in China. Thus we adopted these policies in our model in addition to the government policies from OxCGRT. We can tune the effective reproduction number when we empirically set P6 and P7. According to the Rt sequence obtained from the calculated effectiveness of P1 to P5, we can continuously simulate the effectiveness of P6 and P7 on the current Rt sequence. For the method details, please refer to Appendix C.



Supplementary Figure S1. The architecture of GBRT.

Appendix C: Simulation with parameter setting

The simulation includes the following steps.

(a) Calculate the basic reproduction number R_0 of the environment

The web service has two modes to obtain R_0. One is to enter a free float number between 1 and 5 based on the experience of COVID-19, and the other is to calculate R_0 from *R1–R4* in Supplementary Table S1.

Following a previous study[7], we obtained the distribution of R_0 values to abstract five transmission scenarios based on the sorted values of R_0. Specifically, our study distinguished the environment into five types and output the average R_0 of each type to represent the spread rate. The $R_0^{original}$ of original strain was presumed as 1.44 in hardest scenarios, 1.78 in hard scenarios, 2.04 in baseline scenarios, 2.50 in easy scenarios, and 3.39 in easiest scenarios as default values for parameter R2 in Supplementary Table S1.

In addition, we now know that different SARS-CoV-2 variants have different transmission rates. We also obtained the spread rate for each variant compared with the original strain. Assuming the original variant's strength is 1, we then assigned the corresponding coefficients to another variant. In reference to recent studies[8,9], we set the spread ratio of VOI (variant of interest) to the original strain M^{VOI} as 1.29 for Alpha strain, 1.25 for Beta strain, 1.38 for Gamma strain, 1.97 for Delta strain and 3.86 for Omicron strain as default values for parameter R1 in Supplementary Table S1.

The vaccine is an effective means to control the transmissibility of infectious diseases. Hence, we considered the coverage C^{VOI} and effectiveness V^{VOI} of vaccine in our model as parameters R3 and R4 in Supplementary Table S1.

We can calculate the final basic reproduction number as:

$$R_0 = R_0^{original} \times M^{VOI} \times (1 - C^{VOI} \times V^{VOI})$$

(b) Calculate effective reproduction number Rt after intervention

Referring to relevant research[5], we can set a functional relationship between policy levels and weight and the effective reproduction numbers (*Rt*). By combining with the initial R_0 value, we can obtain the deduction rate compared to R_0 under the influence of different policies. Thus, we can obtain R_t^1 value after policy execution as follows:

$$R_t = R_0 exp(-\gamma \times \sum_{k=1}^{m} \alpha_k I_k),$$

where I_k indicates whether policy k executes, α_k for the weights of a specific policy level, R_0 for the basic number of regeneration, andγ is the coefficient for tuning the deduction rate based on real-world scenarios.

To obtain the value of γ, we first performed Z-score normalization on α_k to ensure that $min(\sum_{k=1}^{m} \alpha_k I_k) = 0$ and $max(\sum_{k=1}^{m} \alpha_k I_k) = 1$. We assumed that the lockdown of Wuhan had the strongest impact on transmission, so its $\sum_{k=1}^{m} \alpha_k I_k = 1$. Thus, we can use R_0 and R_t before and after the lockdown of Wuhan to estimate γ. According to an article from JAMA[10] estimating the daily R_t values in Wuhan, we could obtain R_0 around 3.82 and R_t around 0.51. R_t after lockdown is difficult to precisely estimate; however, the article showed that R_t is always around 0.5 after lockdown comes into force. Thus, we chose a day from the middle and obtained R_t = 0.51.

We then inputted the value into our formula and obtained:

$$0.51 = 3.82 \times exp(-\gamma \times 1)$$

$\gamma = 2.015$

The original formula was changed as

$$R_t = R_0 exp(-2.015 \times \sum_{k=1}^{m} \alpha_k I_k)$$

In the regular calculation of feature importance on tree models, the training process only needs to be executed once. However, with limited data, the calculated feature importance, which is the policy weight in our paper, may not remain the same after rerunning the model. Thus, we conducted multiple rounds of training to generate policy weights and confidence intervals.

In Supplementary Figure S2, from 200 to 1,400 rounds, we recorded the weight of each policy level every 200 rounds. Except for P1 with unstable values, the weights of each policy level basically converged at 1,000 rounds. Thus, we trained the model 1,000 times to ensure the stability of our calculation.

For each iteration, we trained the model with the same training set mentioned above and saved the weights of each policy level.

We then applied the current policy level to the policy weight set to obtain the weighted sum of the coefficient. With the same policy level and 1,000 weight sets, we could generate 1,000 coefficient values and then calculate 1,000 R_t values. Using these R_t values as basis, we computed the 95% confidence interval and output the mean, upper, and lower boundaries of R_t.

The current model only calculates the effectiveness of P1–P5. For P6 and P7, we need the current output of R_t as input and perform additional calculations. As long as risk groups are found, they will be immediately isolated.

For P6, based on the analysis of epidemic cases in Guangzhou in mid-2021, we can calculate the proportion of different high-risk groups in the confirmed cases.

The total confirmed cases were 147. The close contacts and secondary close contacts of cases were 96, with a corresponding proportion of 65.3%.

The cases of visitors of high-risk regions and close contacts plus screening on main groups were 96 + 6 = 102 with the corresponding proportion of 69.3%.

The cases of visitors of regions with confirmed cases and the visitors of high-risk regions plus mass screening were 102 + 42 = 144 with the corresponding proportion of 97.9%.

Combined with the time cost of testing for high-risk people, assumed to be x days, the current R_t decreased steadily within n days. We set high-risk people as "1 − Quarantine close contacts and secondary close contacts of cases," accounting for 65.3% of the total number of confirmed cases. Hence, the final R_t decreased to:

$R^1_{1+n} = R^1_1 \times (1 - 65.3\%)$

R^1_1 is the current R_t before P6 and P7 executes (But after P1–P5);

R^1_{1+n} is the R_t after n days

Assuming that the test rate is stable every day, the R_t's deduction rate of each day in the n days is

$$R_t = \frac{(R^1_{1+n} - R^1_1)}{n}$$

After policies execute, R_t will temporarily be presumed the same value later until new policies are applied. Finally, we output the R_t sequence based on the given R_0 and policy levels I and policy weights α.

(c) Epidemic simulation

In this study, we assumed that imported cases were imported the day after their infection. We defined vaccination rates to remain constant from the first day of the simulation. We set the upper boundary of policy start day to 31, which may be larger than the incubation period for COVID-19.

	200 rounds	400 rounds	600 rounds	800 rounds	1,000 rounds	1,200 rounds	1,400 rounds
P1 as 1	0.044 (0.037, 0.045)	0.043 (0.037, 0.042)	0.043 (0.038, 0.042)	0.044 (0.039, 0.042)	0.044 (0.039, 0.042)	0.04 (0.034, 0.037)	0.039 (0.033, 0.035)
P1 as 2	0.071 (0.061, 0.071)	0.072 (0.066, 0.071)	0.072 (0.066, 0.071)	0.073 (0.067, 0.072)	0.073 (0.067, 0.072)	0.067 (0.06, 0.066)	0.065 (0.058, 0.061)
P2 as 1	0.122 (0.115, 0.126)	0.123 (0.117, 0.126)	0.122 (0.117, 0.123)	0.123 (0.119, 0.124)	0.123 (0.118, 0.123)	0.123 (0.119, 0.123)	0.123 (0.119, 0.123)
P2 as 2	0.074 (0.066, 0.073)	0.073 (0.066, 0.073)	0.073 (0.066, 0.072)	0.073 (0.067, 0.072)	0.073 (0.067, 0.071)	0.073 (0.068, 0.072)	0.073 (0.068, 0.071)
P3 as 1	0.08 (0.074, 0.083)	0.083 (0.077, 0.083)	0.084 (0.079, 0.083)	0.084 (0.08, 0.083)	0.084 (0.079, 0.082)	0.083 (0.079, 0.082)	0.083 (0.079, 0.082)
P3 as 2	0.075 (0.069, 0.075)	0.075 (0.07, 0.074)	0.075 (0.071, 0.074)	0.075 (0.071, 0.074)	0.074 (0.071, 0.074)	0.074 (0.071, 0.073)	0.074 (0.071, 0.073)
P3 as 3	0.071 (0.064, 0.074)	0.07 (0.064, 0.071)	0.071 (0.065, 0.072)	0.072 (0.065, 0.071)	0.071 (0.065, 0.07)	0.071 (0.065, 0.07)	0.071 (0.066, 0.07)
P3 as 4	0.033 (0.027, 0.034)	0.033 (0.027, 0.032)	0.033 (0.027, 0.032)	0.033 (0.027, 0.031)	0.033 (0.028, 0.031)	0.033 (0.028, 0.031)	0.033 (0.028, 0.03)
P4 as 1	0.082 (0.077, 0.087)	0.082 (0.076, 0.085)	0.082 (0.077, 0.084)	0.081 (0.076, 0.082)	0.082 (0.077, 0.082)	0.083 (0.08, 0.084)	0.083 (0.08, 0.083)
P4 as 2	0.067 (0.061, 0.067)	0.066 (0.059, 0.063)	0.066 (0.059, 0.063)	0.065 (0.06, 0.063)	0.066 (0.06, 0.063)	0.066 (0.061, 0.063)	0.067 (0.061, 0.064)
P5 as 1	0.054 (0.048, 0.054)	0.051 (0.045, 0.051)	0.051 (0.046, 0.049)	0.05 (0.046, 0.049)	0.05 (0.046, 0.049)	0.051 (0.047, 0.05)	0.051 (0.047, 0.049)
P5 as 2	0.086 (0.079, 0.086)	0.083 (0.076, 0.083)	0.084 (0.078, 0.083)	0.083 (0.078, 0.082)	0.083 (0.078, 0.082)	0.085 (0.079, 0.083)	0.084 (0.079, 0.082)

Supplementary Figure S2. Weights of each policy level expressed as mean + 95% *CI*.

After P8, C1, and C2 are entered, we have the number of input cases on day 1, the basic reproduction number (either from input or based on the calculation from variant and scenario), and the reproduction number sequence for the future.

We can also calculate generation time T using the statistics from clinical data[11]. Then,

$$\gamma(t) = \frac{1}{T}$$

According to the definition of R_0,

$R \propto b \times k \times T$,

where b is the infection rate of a single contact, k is the average contact rate of the infected and susceptible, and T is the generation time. We already calculated R_t sequences from the previous step. Therefore, we can calculate the daily transmission rate on the basis of the generated Rt sequence:

$$\beta(t) = k \times b = \frac{R_t}{T}$$

Based on the definition of group *I* in the SIR model:

$$\frac{dI}{dt} = \frac{\beta I \times S}{N} - \gamma i(t)$$

In generating *I*, *SIR* has considered the removed group *R.* Given that the patients before *T* (the generation time) days have been detected and isolated, we removed them from our *I* group. Hence, the new expression of group *I* is as follows:

$$\frac{\mathrm{dI}}{\mathrm{dt}} = \frac{\beta \mathrm{S}}{\mathrm{N}} \times \mathrm{I}(\text{with restricted time window}) = \frac{R_t}{T} \times \frac{\mathrm{S}}{\mathrm{N}} \times \mathrm{I}(\text{with restricted time window})$$

Starting from the second day, we used generation time *T* as the time window to restrict group *I*. When *T* > the time *t* when policy intervention is started, the size of the time window = *min (n, T)*. The onset on day *T* could be expressed as:

$$Onset(t) = \frac{\mathrm{dI}}{\mathrm{dt}} = \sum_{i=1}^{\min(t,T)} Infected_i \times \frac{R_t}{T} \times \frac{population\ of\ current\ scenario - total\ infected\ cases}{population\ of\ current\ scenario}$$

We can calculate future daily cases by combining with the variable "*start*," which is the time when policy executes as well as the variable *I(0) = base*.

After calculating $Onset(t)$ for each single day *t*, we established the process of mapping interventions, vaccines, and strains to the real-world epidemic curves.

REFERENCES

1. Hale T, Anania J, Angrist N, et al. Variation in government responses to COVID-19. https://www.bsg.ox.ac.uk/sites/default/files/2021-03/BSG-WP-2020-4032-v11.pdf. [2021-03-24].

2. Dong ES, Du HR, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis, 2020; 20, 533-4.

3. Wallinga J, Teunis P. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. Am J Epidemiol, 2004; 160, 509-16.

4. UNStats. New York: United Nations Publications, 2021.

5. Brauner JM, Mindermann S, Sharma M, et al. Inferring the effectiveness of government interventions against COVID-19. Science, 2021; 371, eabd9338.

6. Elith J, Leathwick JR, Hastie T. A working guide to boosted regression trees. J Anim Ecol, 2008; 77, 802-13.

7. Hilton J, Keeling MJ. Estimation of country-level basic reproductive ratios for novel coronavirus (SARS-CoV-2/COVID-19) using synthetic contact matrices. PLoS Comput Biol, 2020; 16, e1008031.

8. Campbell F, Archer B, Laurenson-Schafer H, et al. Increased transmissibility and global spread of SARS-CoV-2 variants of concern as at June 2021. Euro Surveill, 2021; 26, 2100509.

9. UK Health Security Agency. SARS-CoV-2 variants of concern and variants under investigation in England technical briefing 32. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1042688/RA_Technical_Briefing_32_DRAFT_17_December_2021_2021_12_17.pdf. [2022-03-01].

10. Inglesby TV. Public health measures and the reproduction number of SARS-CoV-2. JAMA, 2020; 323, 2186-7

11. Deng YH, You C, Liu YK, et al. Estimation of incubation period and generation time based on observed length-biased epidemic cohort with censoring for COVID-19 outbreak in China. Biometrics, 2021; 77, 929-41.