Supplementary material: Transient indicators of tipping points in infectious diseases

Suzanne M. O'Regan Eamon B. O'Dea Pejman Rohani John M. Drake

7 July 2020

S1 Simulation models

Our simulation models are partially observed Markov processes. The state variables of the models are the number of susceptible, infected, and removed individuals. To generate realizations of the process, these state variables were updated according to the transition probabilities in table S1 using the implementation of Gillespie's algorithm in the *pomp* package (King et al., 2016). This stochastic model has a deterministic skeleton nearly equivalent to the deterministic model, equation 1 in the main text, that we use to analyze transient dynamics. The only difference is that the deterministic skeleton of our stochastic model (table S1) includes a term modeling the importation of infection. We keep this term small enough that it will have little effect on the expected trajectory of transient dynamics and include it only to eliminate complicating effects of stochastic extinction in our simulation results.

Observed data are calculated as the number of transitions from the infected to the removed class over the course of a week. We refer to these transitions as cases, and they are meant to model the number of reported visits to healthcare providers, which are commonly available for notifiable diseases. The two simulated diseases are measles and pertussis (whooping cough), with respective parameters specified in table S2. We simulated 1000 replicates of each parameter set, calculated indicators, and then calculated areas under the receiver operator characteristic curve (AUC) to quantify the classification ability of indicators. A value close to 1 indicates little overlap in the distributions of the indicators for different levels of immunity and thus a strong classification ability. A value close to 0.5 indicates high overlap and thus poor classification ability. Having described the common features of simulations in our two scenarios, we now turn to the specifics of each.

In the emergence scenario, we suppose data from two outbreaks of a disease that is well-controlled (*i.e.*, for which $R_0(1-p) < 1$) are observed and we would like to know if the disease was closer to the point of emergence in one of the outbreaks. The outbreaks are initialized with 1 individual in the infected class, a fraction of the population in the removed class that is equal to the vaccine uptake p, and the balance of the population in the susceptible class. The values of p used were 0.95 to 0.99, in steps of 0.01. Simulations were

Table S1: Transition probabilities for the simulation models

Event	$(\Delta S, \Delta I, \Delta R)$	Transition probability
Birth of susceptible	$(1, 0, 0)$	$\mu N_0(1-p)\Delta t$
Death of susceptible	$(-1, 0, 0)$	$\mu S\Delta t$
Death of infectious	$(0, -1, 0)$	$\mu I\Delta t$
Death of recovered	$(0, 0, -1)$	$\mu R\Delta t$
Transmission	$(-1, 1, 0)$	$\beta SI/(S+I+R)\Delta t$
Importation of infectious	$(-1, 1, 0)$	$\eta S\Delta t$
Recovery of infectious	$(0, -1, 1)$	$\gamma I\Delta t$
Birth of immune	$(0, 0, 1)$	$\mu N_0 p\Delta t$

Table S2: Parameters for deterministic and stochastic SIR models.

Symbol	Definition	Value
μ	per capita birth/death rate	0.02 yr^{-1}
N_0	population size	10^7
R_0	basic reproduction number	17
γ (pertussis)	per capita recovery rate	$365/22$ yr^{-1}
γ (measles)	per capita recovery rate	$365/13$ yr^{-1}
β	per capita transmission rate	$(\gamma+\mu)R_0$ yr^{-1}
η	per capita importation rate	varies
p^*	critical vaccination threshold	$1-1/R_0$
p	vaccine uptake	varies

run until the outbreak ended. The importation rate η was set to zero so that the data corresponded to the observation of a single outbreak. Some simulated trajectories are displayed in figure S1.

In the elimination scenario, we suppose that 14 years of weekly case counts are available and we would like evidence that the disease is closer to the point of elimination in the second half of the time series. As we shall explain when we describe our method of analysis for this data, 14 years is roughly the minimum time span over which two independent calculations of our indicators can be done for a pertussis time series. The vaccine uptake p jumps from zero to some fixed value in the middle of the time series, which models the beginning of a mass vaccination program. The simulations are initialized at the endemic equilibrium and run for a burn-in period of 13 years prior to producing output. Consequently, the distribution of states at the beginning of the available data is close to the stationary distribution. The jumps in p used were 0.2, 0.4, and 0.6. The importation rate η was set to $365/N_0$. Figure S2 displays some of our simulated data sets for this scenario.

S2 Analysis of reactivity

S2.1 Reactivity of the disease-free equilibrium

Letting $x = R_0(1-p) < 1$ and differentiating the reactivity of the disease-free equilibrium $\nu_d = (-\mu - \Gamma(1-x))/2 + \sqrt{(-\mu-\Gamma(1-x))^2 - 4\mu\Gamma(1-x) + (\Gamma x)^2}/2$ yields

$$\frac{d\nu_d}{dx} = \frac{\Gamma}{2}\left(1 + \frac{\Gamma(2x-1)+\mu}{\sqrt{\Gamma^2x^2 - 4\Gamma\mu(1-x) + (-\mu-\Gamma(1-x))^2}}\right).$$

There is a stationary point at $x = 0$. The second derivative of ν_d is

$$\frac{d^2\nu_d}{dx^2} = \frac{\Gamma^2(\Gamma-\mu)^2}{2(\Gamma^2x^2 - 4\Gamma\mu(1-x) + (-\mu-\Gamma(1-x))^2)^{3/2}},$$

which is strictly positive for $0 \le x \le 1$. Therefore, reactivity of the disease-free equilibrium has a global minimum at $x = 0$ and consequently is strictly increasing for $0 < x < 1$.

S2.2 Reactivity of the endemic equilibrium

Letting $x = R_0(1-p)$, then $x > 1$ since $p < p^*$, and the derivative of endemic reactivity $\nu_e = -\frac{\mu x}{2} + \frac{1}{2}\sqrt{(-\mu x)^2 + (\Gamma - \mu(x-1))^2}$ is

$$\frac{d\nu_e}{dx} = \frac{\mu}{2}\left[\frac{2\mu x - (\Gamma+\mu)}{\sqrt{\mu^2x^2 + (\Gamma-\mu(x-1))^2}} - 1\right].$$

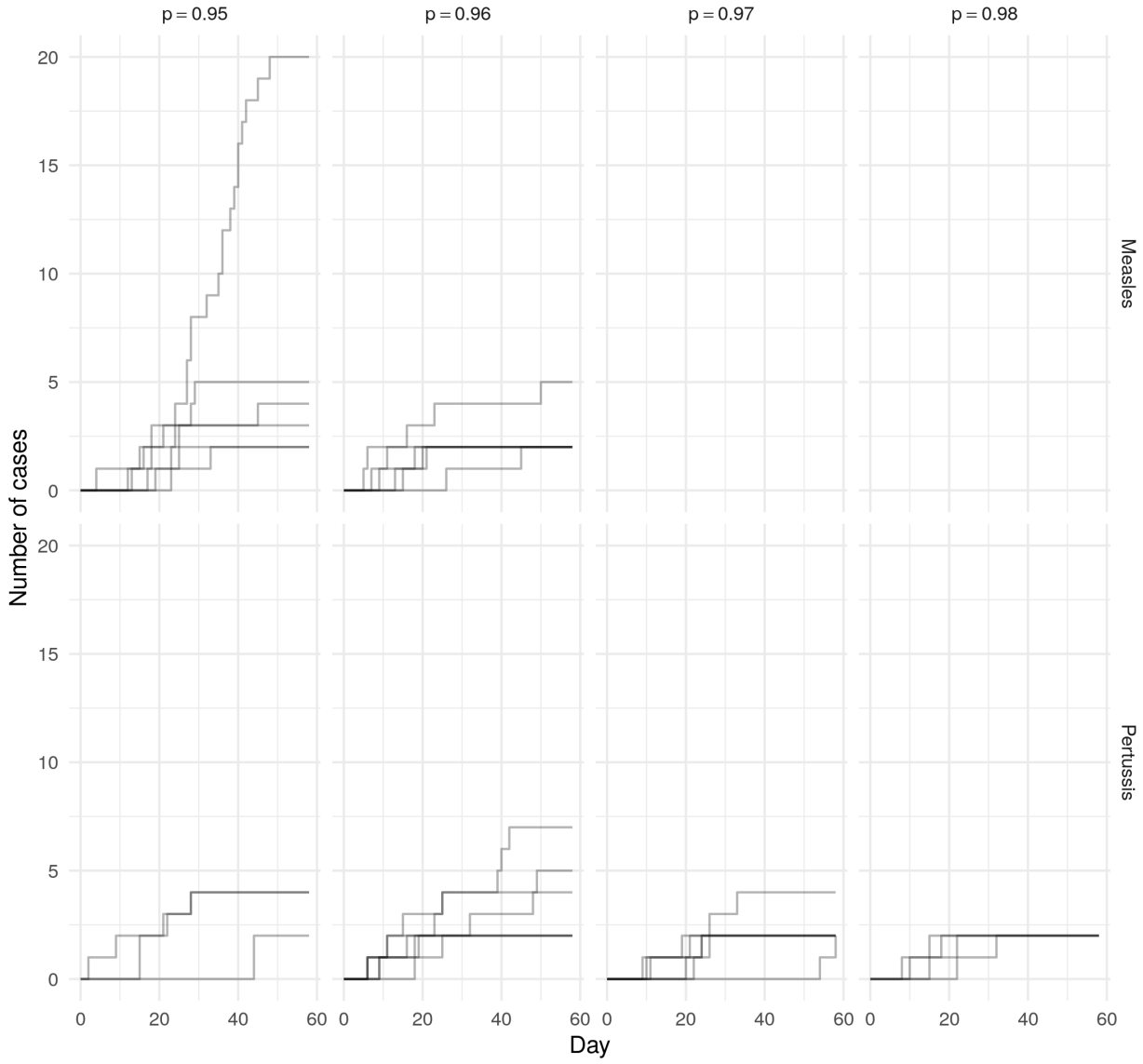

Figure S1: Trajectories of simulated outbreaks for scenario in our emergence simulation study. Simulation replicates in which the initially infected patient failed to transmit are not plotted because they could not be used to calculate some indicators. These outbreaks of size 1 were frequent when vaccine uptake p was higher. The trajectories resemble those of the historical outbreaks in figure 1 in that they appear roughly linear overall.

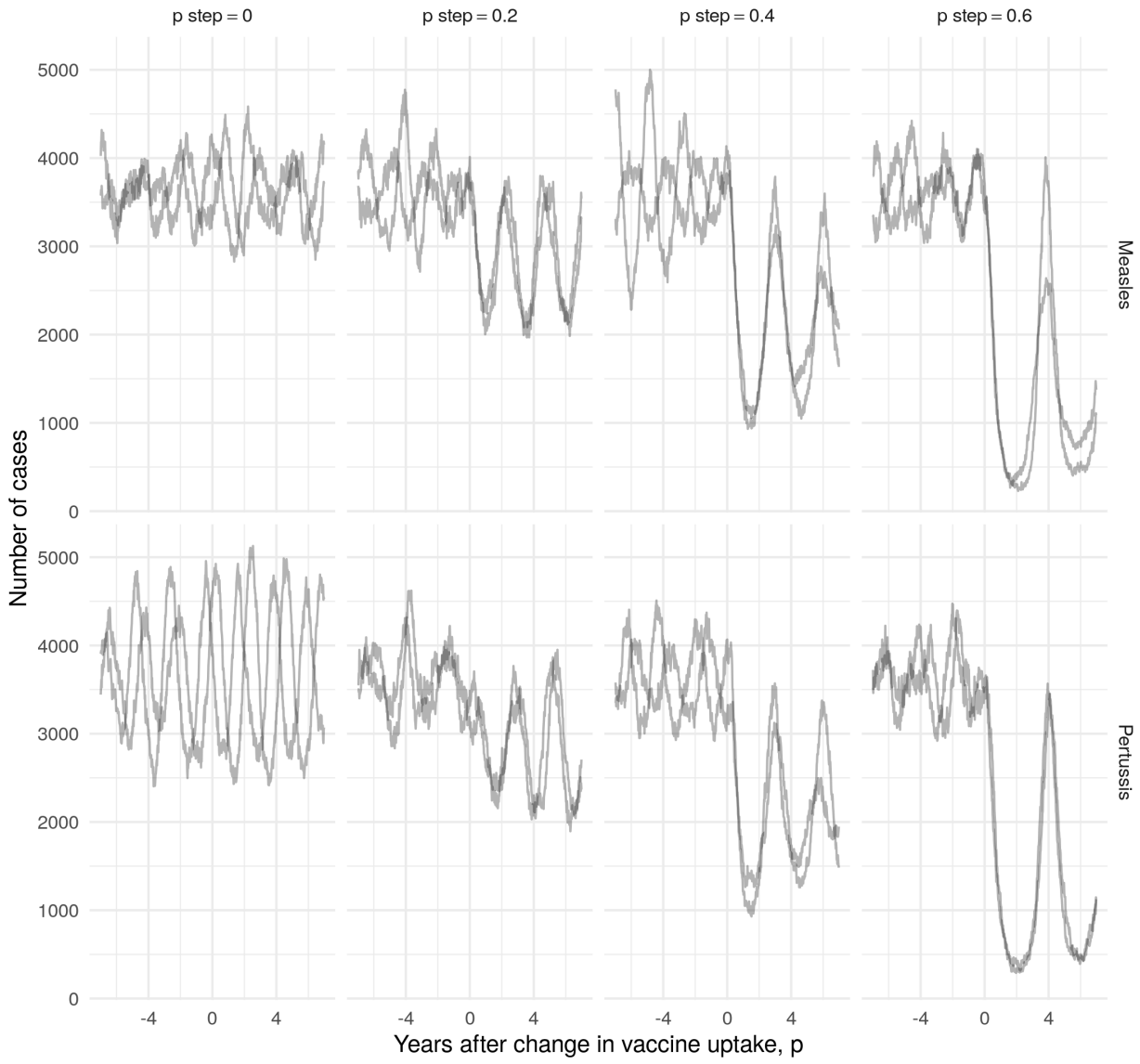

Figure S2: Two sample trajectories of simulated case reports for each of our scenario in our elimination simulation study. "p step" in the labels for columns of panels refers to the change in vaccine uptake p. The pertussis simulations with a p step of 0.2 most resemble the real data in figure 2A

.

Assuming $\gamma > 0$, $\mu > 0$ and $x > 1$, the endemic reactivity ν_e has a single stationary point, at $x_c = \Gamma/\mu + 1$. The second derivative of ν_e is

$$\frac{d^2\nu_e}{dx^2} = \frac{\mu^2(\Gamma+\mu)^2}{2(\mu^2x^2 + (\Gamma - \mu(x-1))^2)^{3/2}},$$

which is positive at x_c and additionally positive for all $x > 1$. By the second derivative test, the reactivity function has a global minimum at x_c, and therefore, reactivity is a decreasing function of x if $1 < x < x_c$. Consequently, as p approaches p^* from the left, reactivity increases with p.

S3 Derivation of transient indicators for emergence scenarios

We show in section S2.1 that the reactivity of the disease-free equilibrium increases with x. Figure S3 illustrates that, for the disease-free equilibrium, the eigenvector corresponding to the eigenvalue $-\mu$ is always aligned with the axis of the S variable, whereas the other eigenvector is aligned with the axis of the I variable only when $x = 0$. As x moves from zero to one, this other eigenvector sweeps toward the S axis, thus closing the angle between the two eigenvectors from the initial value of $\pi/2$. Because only one of the eigenvectors projects onto the I axis, and its eigenvalue is negative, the I variable cannot contribute to transient growth of the norm. The main idea of our practical indicators is to characterize the behavior of the initial growth rate and maximum of the norm of the S variable for a given perturbation of I. Although improved performance may result from additionally including the behavior of the I variable, we use a simpler approach for the benefit of clarity in this work.

As a practical indicator of reactivity, we use an estimate of the rate at which S decreases during an outbreak. Although this indicator does not account for the I component of the norm of $(S - \bar{S}, I - \bar{I})$, it is clear from the structure of equation 3 that this indicator behaves in the same way as true reactivity. The element in row 1, column 2 of the matrix $\mathbf{J_d}$ in table 1 is multiplied by I to determine how quickly S decreases due to infection, and this element grows with x.

As a practical indicator of maximum amplification, we use an estimate of the total decrease in S. In the previous paragraph, we have noted that S decreases more rapidly for any given value of I as x increases. Further, from the second row of the matrix $\mathbf{J_d}$ in table 1, I evidently decreases less rapidly for a given value of I as x increases. Thus by the time I becomes close to zero (*i.e.*, the outbreak ends), the total decrease in S must increase with x. This behavior is the same as for theoretical value of maximum amplification (figure 4), which we shall report on more fully in *Results* of the main text.

References

King AA, Nguyen D, Ionides EL. Statistical Inference for Partially Observed Markov Processes via the R Package pomp. Journal of Statistical Software 2016;69(12):1–43. doi:10.18637/jss.v069.i12.

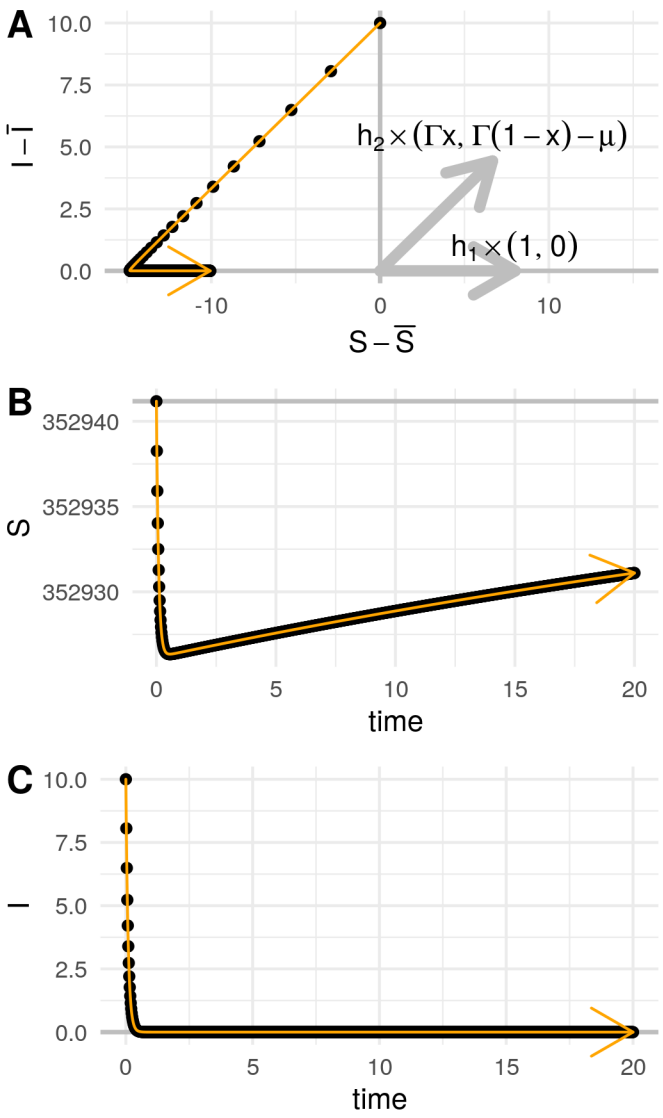

Figure S3: Near a stable disease-free equilibrium $(\bar{S}, \bar{I})$, transient growth of a perturbation is caused by growth in $S - \bar{S}$. (**A**) The grey arrows plot the eigenvectors of the Jacobian of the disease-free equilibrium, $\mathbf{J_d}$ in table 1, for the values of the measles parameters in table S2 and $p = 82/85$. The arrowheads of the eigenvectors are labeled with formulas for their S and I components, in which h_1 and h_2 are normalization constants. Only one vector has an I component, and the shared S components permit transient growth of $S - \bar{S}$. The series of points on the path indicated by the orange arrow are an example of a trajectory of the system that displays transient growth. The points give the position of the system each week in state space and the direction of the arrow indicates the time-ordering of the points. The parameters used to calculate this trajectory are the same as those used for the eigenvectors. (**B**, **C**) Plots of the dynamics of individual variables. The horizontal grey reference lines indicate the equilibrium values $\bar{S}$ and $\bar{I}$.

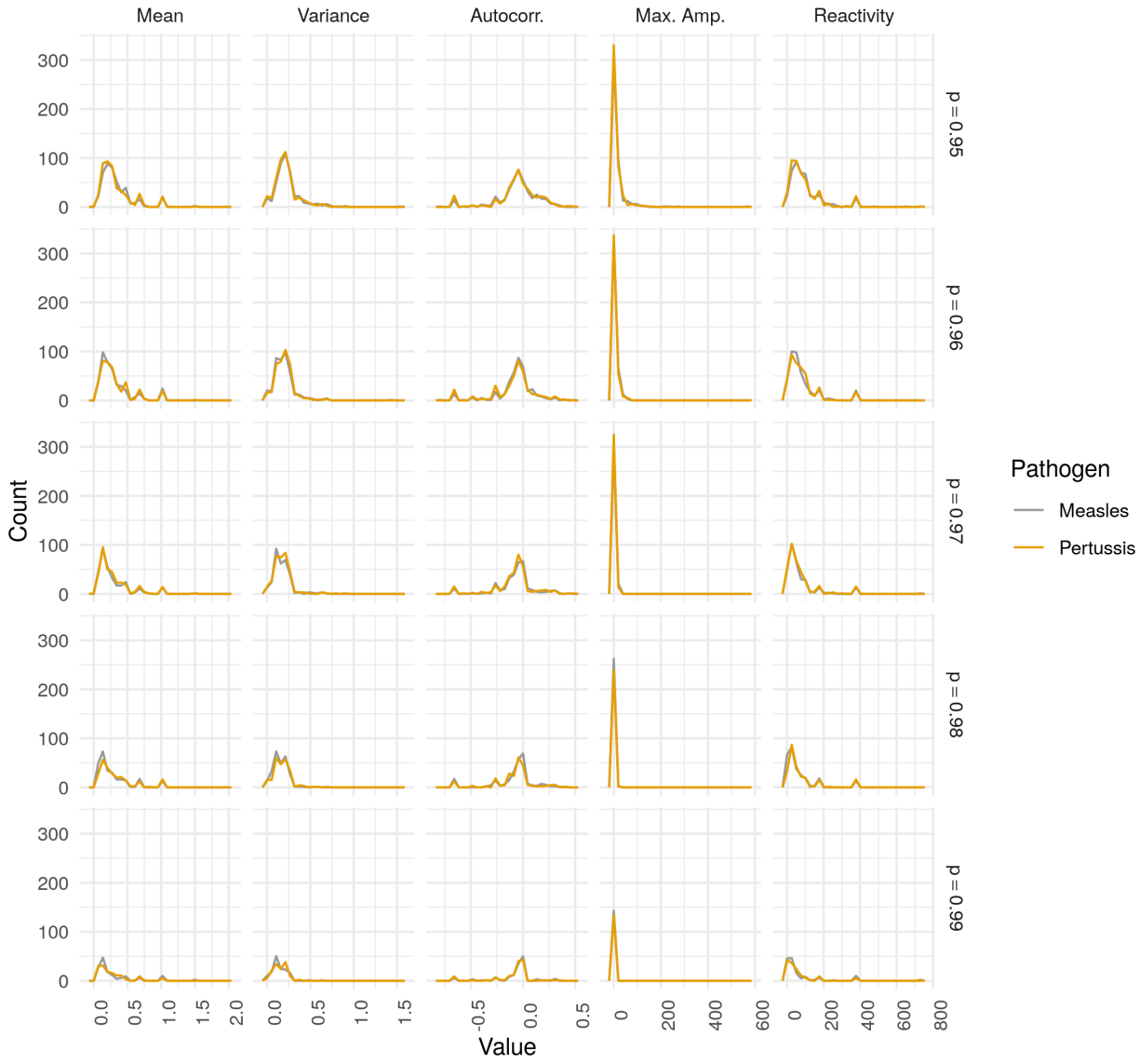

Figure S4: The distributions of all indicators calculated from simulated outbreaks shifted to the right as vaccine uptake decreased. The measles and pertussis simulations with the same vaccine uptake had statistically indistinguishable distributions, illustrating the insensitivity of these indicators of emergence to the infectious period of the pathogen given a fixed reproduction number.

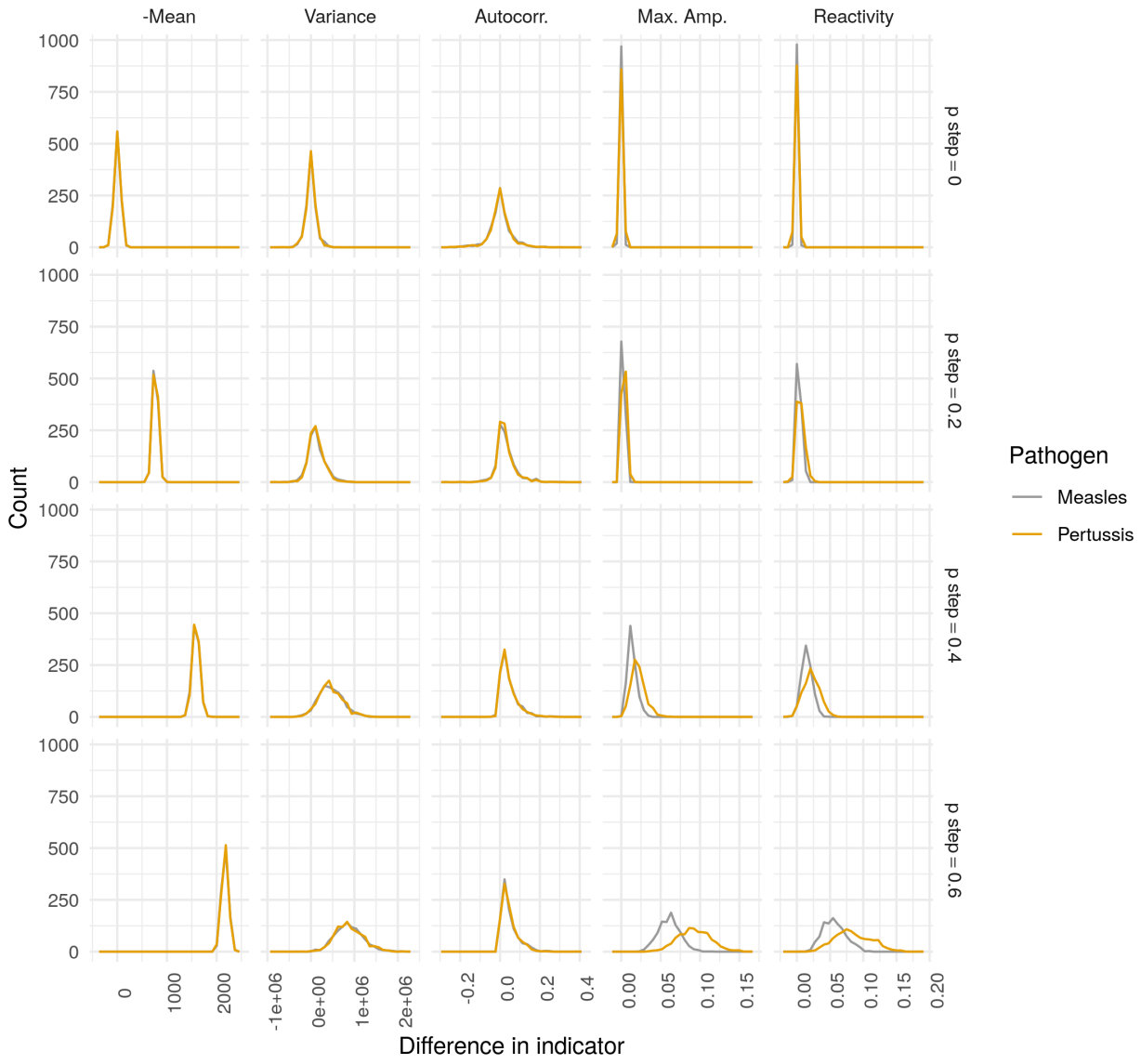

Figure S5: The distributions of the difference in indicators from two windows of a time series of an endemic disease shifted to the right as the difference in vaccine uptake p between the windows increased. This difference is referred to as "p step" in the panel labels on the right of the plot.