Full title: Limitations of non-mechanistic methods for characterizing pathogen-pathogen interactions: A simulation study

Short title: Limitations of non-mechanistic methods for pathogen interactions

Sarah C. Kramer[1], Sarah Pirikahu[1,2], Cana Kussmaul[1,3], Lulla Opatowski[4,5], Matthieu Domenech de Cellès[1]

1. Max Planck Institute for Infection Biology, Infectious Disease Epidemiology group, Charitéplatz 1, Campus Charité Mitte, 10117 Berlin, Germany
2. Department of Public Health, University of Otago, Wellington, New Zealand
3. Institute of Public Health, Charité – Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany
4. Institut Pasteur, Université Paris Cité, Epidemiology and Modelling of Antibiotic Evasion (EMAE) unit, Paris, France
5. Université Paris-Saclay, Université de Versailles Saint-Quentin-en-Yvelines, Inserm U1018, CESP, Infectious diseases, interactions and anti-infective evasion team, Montigny-Le-Bretonneux, France

Abstract

Pathogen-pathogen interactions occur when infection with one pathogen influences one's chance of infection or disease due to another. Increasingly, evidence suggests that interactions are a common feature of infectious disease epidemiology. However, due to both the nonlinearities and stochasticity inherent to infectious disease transmission, and the frequency of confounding (e.g., by shared seasonal forcing), purely statistical methods for characterizing interactions may be prone to failure. Here, we perform a simulation study to evaluate several non-mechanistic approaches for inferring causality from time series data: generalized additive models (GAMs), Granger causality, transfer entropy, and convergent cross-mapping (CCM). Specifically, we use a two-pathogen mechanistic transmission model, calibrated to produce dynamics resembling outbreaks of influenza and respiratory syncytial virus (RSV), to generate synthetic datasets with a range of values for interaction strength and duration. We then apply each method to all synthetic datasets. We find that Granger causality, transfer entropy, and CCM all fail to consistently infer whether data contain signal of an interaction; in particular, methods tend to incorrectly identify interactions where none are modeled (average sensitivity = 80.6%, 92.1%, 72.1%, respectively; average specificity = 31.0%, 33.3%, 33.1%). Furthermore, we find little to no association between point estimates from each method and true interaction strength. In contrast, GAMs infer the existence of interactions more accurately than the other methods (sensitivity = 85.2%, specificity = 72.5%), and consistently yield larger point estimates for stronger interactions. However, their practical utility is limited by an inability to evaluate interaction asymmetry (i.e., whether the effect of pathogen A on pathogen B is identical to that of B on A). We conclude that accurately and comprehensively characterizing pathogen-pathogen interactions based on outbreak data remains a significant challenge. For this reason, it is critical that any proposed methods be rigorously evaluated before being used to draw conclusions about interactions.

Author Summary

Pathogen-pathogen interactions occur when infection with one pathogen either increases or decreases a person's risk of infection or illness due to a second, distinct pathogen. Because interactions affect several common human pathogens, including influenza and SARS-CoV-2, a better understanding of interactions could improve epidemic control. However, past work has shown that simple methods commonly used to study interactions can lead to inaccurate conclusions. Here, we tested four methods frequently used in other fields, including ecology and neuroscience, to see whether they may also be useful for identifying interactions. Specifically, we tested each method using simulated outbreak data generated using a mathematical model. We found that most methods struggled to correctly determine whether an interaction effect was present; in particular, methods often falsely identified interactions when none occurred. Although one of the tested methods, generalized additive models, performed comparatively well at identifying interactions, it provided relatively little additional information about the interactions. Because pathogen-pathogen interactions are so challenging to study, it is important that researchers rigorously test methods before applying them to interactions, so as not to publish potentially misleading results. More broadly, a complete understanding of interactions will likely require a variety of approaches, including both laboratory and modeling studies.

Introduction

Although infectious diseases are most commonly studied in isolation, pathogen-pathogen interactions, in which infection with one pathogen modifies the risk of infection or disease due to another pathogen, are rapidly gaining more attention. Recent evidence suggests that a complex web of interactions exists between respiratory viruses, including influenza virus, respiratory syncytial virus (RSV), rhinovirus, human metapneumovirus (HMPV), and severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) (1–11). Many of these viruses may in turn affect infection with bacteria like *S. pneumoniae* and *Haemophilus influenzae* (12,13), while the impact of immunosuppression induced by viruses like HIV and measles on susceptibility to a wide range of pathogens has long been recognized (14,15). Indeed, it seems likely that interactions are the norm, rather than the exception, in infectious disease epidemiology. The mechanisms underlying interactions are varied, and include both direct (e.g., competition for resources) and indirect (e.g., modulation of host cells or the immune response) processes (8,16). The effects of these interactions, both for individual-level health and for outbreak dynamics at the population level, can be substantial (8,16). For this reason, characterizing these interactions is a necessary step in improving our understanding of the epidemiology of a wide range of infectious diseases.

Unfortunately, pathogen-pathogen interactions are difficult to characterize. In part, this is because interactions themselves are complex: they can vary in strength and duration, may be positive or negative (i.e., causing either an increase or a decrease in susceptibility to or severity of infection with a subsequent pathogen), and may be symmetric (i.e., having an equivalent impact in both directions) or asymmetric (8). To fully understand an interaction, each of these components must be elucidated. A further challenge is introduced when attempting to study interactions using surveillance data. Because of nonlinearities and stochasticity in the underlying transmission dynamics of the pathogens under study, patterns in the observed number of cases or deaths do not necessarily reflect patterns in the underlying drivers (17,18). Additionally, confounders, such as shared seasonal forcing, may create associations between observed time series that are not due to interaction effects. For these reasons, simple metrics calculated from incidence data are prone to failure: measures of coinfection prevalence often underestimate the true strength of an interaction (19), while phase differences are not a reliable indicator of whether an interaction is positive or negative (20). Although fitting mechanistic transmission models to empirical data is a promising alternative (1,8,16), these methods are

computationally intensive, and the development of appropriate models requires a good understanding of the natural history and transmission dynamics of the included pathogens.

A variety of non-mechanistic methods have been designed to infer causality between time series, including Granger causality (21,22), transfer entropy (23), and, most recently, convergent cross-mapping (CCM) (24). Meanwhile, methods like generalized additive models (GAMs) (25) can flexibly account for potential confounders while simultaneously inferring relationships between variables. Compared to a mechanistic modeling approach, these methods are relatively easy to implement, and require much less understanding of the underlying transmission dynamics. These methods are popular in many fields, such as econometrics (26,27), neuroscience (28,29) or ecology and environmental science (24,30–33). They have also been used in infectious disease epidemiology, specifically in elucidating the impact of climatic drivers (18,34) and transmission dynamics by age (35). However very few studies have evaluated them in the context of pathogen-pathogen interactions. Cobey and Baskerville (36) evaluated CCM as a tool for identifying lifelong cross-immunity between pathogen strains, but did not consider shorter-term interactions. Meanwhile, Randuineau (37) tested Granger causality and transfer entropy, but focused only on unidirectional interactions.

In this work, we conducted a simulation study to assess whether several commonly-used non-mechanistic methods can correctly classify the interaction effect between two pathogens when confronted with synthetic data generated by a mathematical model. Methods were tested for interactions with a range of strengths and durations, and, where possible, implementations controlling for the shared effect of seasonal forcing on both pathogens were also run. Specifically, methods were evaluated not only on their ability to infer whether or not an interaction was present, but also on whether they could provide information about the relative strength of an interaction. Here, we aim to determine whether these methods show promise in identifying and characterizing interactions based on real-world outbreak data.

Methods

Implementation

All analyses were conducted in R version 4.4.0 (38). All code used for this project has been published on GitHub (39).

Mathematical Model

We generated ten years of synthetic data for two cocirculating pathogens using a compartmental SEITRSxSEITRS model, where S stands for susceptible, E for exposed but not infectious, I for infected and able to transmit, T for a transient period post-infection, and R for recovered and immune. Here, individuals' infection status relative to both pathogens is tracked simultaneously (e.g., individuals in compartment X_{SI} are susceptible to pathogen A and infected with pathogen B). The interaction is modeled as a change in the force of infection for pathogen B among those currently infected (I) or recently infected (T) with pathogen A, and vice versa, consistent with evidence suggesting that virus-virus interactions can arise due to infection with one virus influencing susceptibility to another (2,3,7,40). Specifically, the extent of the change in the force of infection (θ_λ) represents the strength of the interaction effect, such that values above 1 indicate positive interactions (e.g., θ_λ = 2 implies a doubling of the force of infection), values below 1 indicate negative interactions (e.g., θ_λ = 0.5 represents a halving of the force of infection, while θ_λ **=** 0 represents complete inhibition), and a value of 1 indicates no interaction. Meanwhile, the time spent in compartment T ($7/\delta$, where δ is the rate at which individuals leave the T compartment) represents its duration. A model schematic can be seen in Figure 1, and the full model equations can be found in the Supplementary Materials (S1 Text). We have previously used a similar model to infer interactions between influenza and RSV (1).

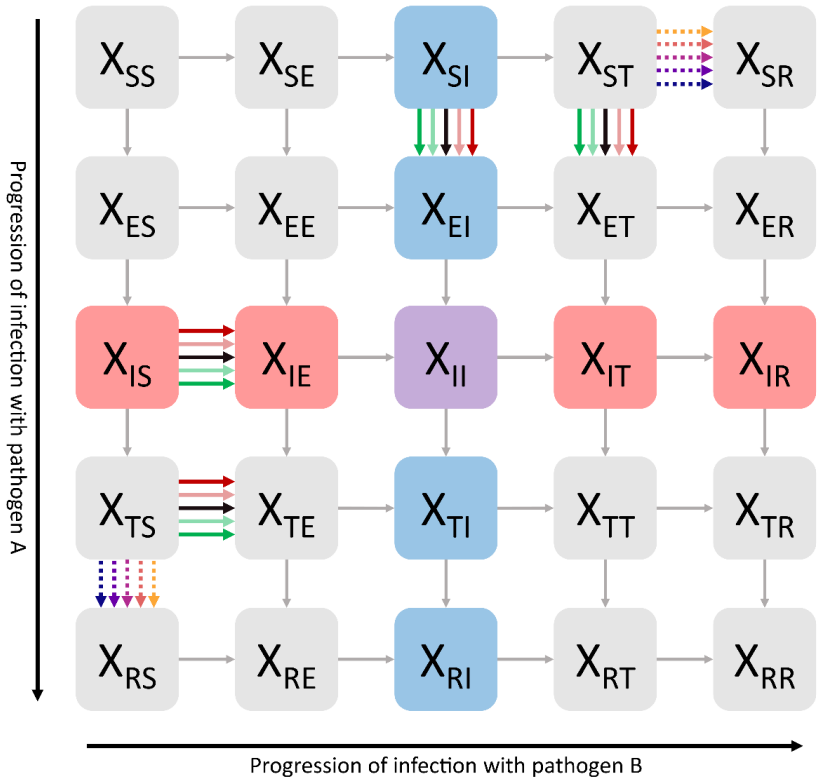


Figure 1. Schematic of the model used to generate synthetic data. Each box represents a distinct model state, with the first subscript indicating infection status with respect to pathogen A, and the second subscript indicating status with respect to pathogen B. Infection with pathogen A is indicated in red and progresses vertically, while infection with pathogen B is indicated in blue and progresses horizontally; coinfection is shown in purple. Arrows represent possible transitions between states. Transitions impacted by changes in interaction strength are shown in a gradient of colors from green to red, while transitions impacted by changes in interaction duration are shown as dotted arrows with a gradient of colors as in Figures 2, 4, and 5.

We parameterized the model such that pathogen A was based on influenza and pathogen B was based on RSV, two viruses between which a moderate-to-strong, negative interaction is likely to exist (1,2). Where information was available, parameter values for each of the two pathogens were based on past observational and modeling studies of both pathogens. Otherwise, parameter values were chosen such that, in the absence of an interaction effect, simulations reproduced patterns resembling real-world influenza and RSV outbreaks in temperate climates (i.e., annual periodicity with a single peak in winter). Specific parameter values used, as well as sources where available, can be found in Table 1.

Table 1. Parameters and parameter values used to generate synthetic data. Ranges indicate parameters that were allowed to vary in the main analysis. Values used for sensitivity analyses are shown in parentheses and italics.

Parameter	Description	Value(s) for Pathogen A (Influenza)	Value(s) for Pathogen B (RSV)	Source
R_{eff}	Effective reproductive number at the start of the simulation	1.05 – 1.4	1.6 – 2.0	(50,51)
$1 / \sigma$	Average latent period (days)	1	5	(52)
$1 / \gamma$	Average infectious period (days)	5	10	(53–55)
$1 / \omega$	Average duration of immunity (years)	1	0.6 – 1.0	(55–57)
α	Amplitude of seasonal forcing	0.2 *(0.05, 0.3, 0.4, 0.5)*	0.2 *(0.05, 0.3, 0.4, 0.5)*	(1,58,59)
ϕ	Week of maximum seasonal forcing (relative to beginning of season)	26	26	NA
ρ	Reporting probability	0.15	0.05	(60,61)
k	Dispersion parameter for the observation model	0.04 *(0, 0.1)*	0.02 *(0, 0.05)*	(62)
β_{sd}	Standard deviation of extra-demographic stochasticity	0.1 *(0.01)*	0.05 *(0.005)*	(63,64)
θ_{λ}	Interaction strength	0, 0.25, 0.5, 1, 2, 4	Same as for A	NA
δ	Inverse of interaction duration (i.e., rate of transition between T and R compartments)	1, 1/4, 1/13	Same as for A	NA

Influenza and RSV display similar seasonal patterns, with both viruses circulating primarily during the winter in temperate regions (41). Indeed, some work suggests that influenza viruses and RSV respond similarly to several environmental drivers, including temperature, humidity, and rainfall (18,42–47). To capture this, we introduced a common seasonal driver to the transmission rate of both pathogens, such that:

$$\beta_i(t) = \bar{\beta}_i \left(1 + \alpha \cos\left(\frac{2\pi}{52.25}(t - \phi)\right)\right) \quad (1)$$

where $\beta_i(t)$ is the transmissibility of pathogen *i* at time *t*, $\bar{\beta}_i$ is the yearly average transmissibility of pathogen *i*, α is the seasonal amplitude of transmissibility, and ϕ is the week during which transmissibility is highest. The division by 52.25 inside the cosine function imposes a yearly periodicity on the transmission rate, and accounts for the weekly timescale of our synthetic data. This shared

seasonal forcing represents a potential confounder when estimating the interaction effect between the two pathogens.

Unlike RSV, which is comparatively antigenically stable, influenza viruses undergo rapid “antigenic drift” (48). Furthermore, there is year-to-year variation in which influenza subtype(s) drive seasonal outbreaks. To capture the impact of this antigenic diversity on population-level susceptibility, we incorporated yearly “surges” in immunity loss into the model of influenza only, as in (49). Specifically, these surges occurred on average during the thirteenth week of each season, and could vary in size from 5-30% of the recovered population. The model population was set to 5 million, and the birth and death rates were both set to 2×10^{-4} per week (i.e., 10.45 births per 1000 people per year), such that the population size remained constant over time.

Generation of Synthetic Data

Synthetic data were generated using a wide range of interaction strengths (θ_λ = 0, 0.25, 0.5, 1, 2, 4) and durations ($7/\delta$ = 1, 4, or 13 weeks), including both positive and negative interactions, as well as simulations with no interaction between the two pathogens. All interactions were symmetric with regard to both strength and duration; in other words, the impact of pathogen A on pathogen B was of the same strength and duration as the impact of B on A.

For each strength-duration pair, we generated 100 synthetic datasets. The timing and magnitude of surges in immune loss (described above) were allowed to vary in each of these 100 simulations, as were values of R_{effA}, R_{effB}, ω_B (see S1 Text for details). This allowed for the generation of some datasets where pathogen A tended to peak before pathogen B, and others where pathogen B tended to peak before pathogen A, a factor we have found in past work to be important in determining whether an interaction can be accurately characterized based on data. The same 100 sets of values for R_{effA}, R_{effB}, ω_B, and the timing and magnitude of the surges in immune loss were used when generating datasets for each of the eighteen combinations of interaction strength and duration.

All datasets were generated using a stochastic model implementation. We incorporated both demographic and extra-demographic (or environmental) stochasticity, with extra-demographic stochasticity modeled by multiplying the transmission rate of each pathogen by gamma white noise

with standard deviation equal to β_{sd} at each timestep (see Table 1) (65,66). The observation model, which draws the weekly number of observed cases from a negative binomial distribution with mean, $\mu = \rho_n H_n(t)$ (where H_n is the total number infected with pathogen n at time t), and dispersion parameter k (i.e., variance $\sigma^2 = \mu + k\mu^2$), contributes additional stochasticity.

All simulations were first run for 20 years to allow the system to approach equilibrium. The model was then run for a further 10 years; these final 10 years of data were used in the analyses described below. To prevent the extinction of pathogen A, we added 100 individuals infected with pathogen A to the system from outside the model population each year during week 14. A representative dataset, where we have varied only the strength and duration of the interaction, can be seen in Figure 2A; further examples can be found in Figure S1.

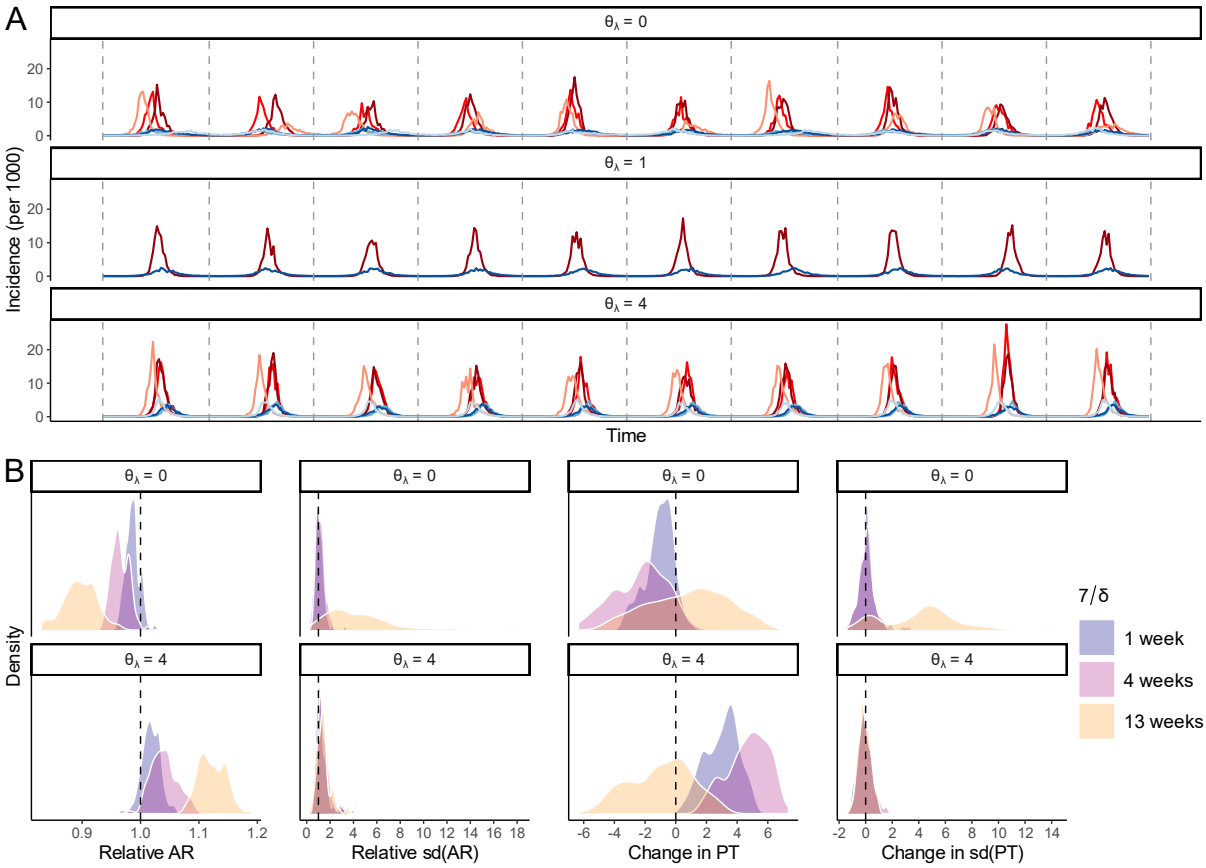


Figure 2. Impact of interactions on outbreak dynamics. (A) Example data consisting of ten years of simulated observations for pathogen A (red) and pathogen B (blue), generated with varying interaction strength (θ_λ) and duration ($7/\delta$) and all other parameters held constant at the values shown in Table 1 (here, Ri_1 = 1.38, Ri_2 = 1.89, $1/\omega_2$ = 49.5 weeks). Specifically, data are shown for simulations with a

strong negative interaction (top), no interaction (middle), and a strong positive interaction (bottom); line colors show the duration of the interaction, with the darkest colors indicating a duration of 1 week, medium colors indicating duration 4 weeks, and the lightest colors indicating duration 13 weeks. Dotted vertical lines indicate the beginning of each new epidemic season, with each season lasting one year. (B) The distribution of the change in attack rate (i.e., the total number of observed cases in a single year), year-to-year variation (i.e., standard deviation across all years) in attack rate, peak timing (i.e., the week during which observed incidence for a given year is maximal), and year-to-year variation (i.e., standard deviation across all years) in peak timing of pathogen B when either a strong negative (top) or strong positive (bottom) interaction is at play, relative to the dynamics when no interaction is present. Shading indicates the duration of the interaction; dotted vertical lines indicate no change.

In addition to the main analysis, we conducted several sensitivity analyses in which we varied 1) the number of years of data used for inference (20 vs. 10), 2) the amplitude of seasonal forcing (α = 0.05, 0.3, 0.4, 0.5), and 3) the amount of process (β_{sd_A} = 0.01, β_{sd_B} = 0.005) and observational (k_A = 0, 0.1; k_B = 0, 0.05) noise.

The transmission model was coded and run using the package pomp (version 6.1) (67).

Statistical Methods

We tested five non-mechanistic methods: (1) Pearson correlations, (2) generalized additive models (GAMs), (3) Granger causality, (4) transfer entropy, and (5) convergent cross-mapping (CCM). Characteristics of these methods are summarized in Table 2. Briefly, we applied each method to all 100 datasets for each of the 16 interaction parameter combinations listed above (see “Generation of Synthetic Data”). To fulfill the assumptions of the first three methods in particular, we log-transformed and centered the data prior to analysis (30); for consistency, this was done for all five methods tested. For methods accounting for shared seasonality, the magnitude of seasonal forcing at each timepoint was similarly transformed. Finally, we checked whether each synthetic dataset was stationary using both the Augmented Dickey–Fuller (68) and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) (69) tests. Nonstationary datasets (n=2 of 1800) were removed from consideration.

Table 2. Overview of all non-mechanistic methods evaluated.

Method	Basis for inferring causality	Notation	Identifies direction (symmetric/ asymmetric)?	Identifies sign (positive/ negative)?	Controls for confounding?	Software used
Pearson Correlation	Association	$r_{Pearson}$	No	Yes	No	R (package "stats")
GAMs	Association	r_{GAM}	No	Yes	Yes (User must specify period of confounder)	R (package "brms" (70)) Stan (71)
Granger Causality	Prediction	$G_{X\to Y}$	Yes	No	Yes (User must input data on confounder)	R (packages "stats," "vars" (72,73), "VARtests" (74))
Transfer Entropy	Prediction	$T_{X\to Y}^{(k,l)}(t)$	Yes	No	Yes (User must input data on confounder)	JIDT (75) R (package "rJava" (76))
CCM	State space reconstruction	ρ	Yes	No	Yes (Implicitly)	R (package "rEDM" (77))

A method was said to have identified an interaction for a given dataset if the results were statistically significant ($p < 0.05$). Meanwhile, the effect estimates were evaluated as indicators of the relative strength of the interactions; for Pearson correlations and GAMs, effect estimates also indicated sign (i.e., estimates could be positive or negative). We describe this process in detail for each method in the following subsections.

Pearson Correlation

The Pearson correlation coefficient measures the strength of the linear association between two datasets, and is defined as:

$$r_{Pearson} = \frac{cov(X,Y)}{s_X s_Y} \quad (2)$$

where $cov(X,Y)$ is the covariance of *X* and *Y*, s_X is the standard deviation of *X*, and s_Y is the standard deviation of *Y*. Values above 0 indicate a positive association, while values below 0 indicate a negative association; values of exactly 1 or –1 indicate a perfect linear relationship between the two datasets. Note that this method cannot be used to test directionality, i.e., $r_{XY} = r_{YX}$.

We calculated the Pearson correlation coefficient between the time series of our two modeled pathogens for all synthetic datasets. Point estimates and two-tailed p-values were obtained using the cor.test function in the package “stats” (version 4.4.0). Unlike many of the other methods tested here, correlation coefficients can be negative, allowing us to evaluate the extent to which interaction sign is correctly inferred. This approach is simple and does not account for potential confounding due to the shared seasonal driver. Furthermore, as we have emphasized in the introduction above, the complexity and nonlinearity inherent to infectious disease dynamics can lead to unintuitive relationships between observed case counts. However, although we do not necessarily expect this method to perform well, we have included it because correlation- and regression-based methods continue to be used to characterize pathogen-pathogen interactions, and, to our knowledge, the accuracy of this approach has yet to be formally evaluated.

Generalized Additive Models

Unlike in standard regression, where an outcome of interest is fit using the raw values of a set of predictors, generalized additive models (GAMs) fit the outcome data to smooth functions of the predictors (25,78). By using a penalized likelihood approach, GAMs ensure that these smooths contain enough “wiggliness” to fit the data, but not so much as to cause overfitting. This allows GAMs to flexibly describe nonlinear relationships between variables. It is this ability that makes GAMs a potentially attractive approach for identifying interactions: by fitting the seasonal cycles in the data, we can explore the association between the time series after removing the effects of shared seasonal forcing. Because our synthetic data represent case counts and not the underlying force of infection, this method does not perfectly control for confounding; nonetheless, it offers a potential advantage over the correlation coefficient method discussed above.

For each synthetic dataset, we fit a bivariate normal GAM to the time series of case counts for both pathogens, with a smooth on week of the year (i.e., 1 through 53; WOY) as the only predictor. When fitting GAMs, it is necessary to specify the basis dimension for any smooths, which sets an upper bound on the flexibility of the curve (25,78). For WOY, we set this value to 53. Furthermore, we used a cyclic cubic regression spline, which ensures that the smooth is continuous at the beginning and end of the year (25). Thus, the fitted model for each dataset was of the form:

$$\boldsymbol{X} \sim N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

$$\boldsymbol{\mu} = [\mu_1, \mu_2]^T$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$$

$$\mu_i = \alpha_i + s_i(WOY) \quad (3)$$

where $\boldsymbol{X} = [X_1, X_2]^T$ is a matrix containing case counts for the full study period for both pathogens, which is modeled as a bivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. The mean and standard deviation for each pathogen *i* are indicated by μ_i and σ_i, respectively, while ρ is the correlation coefficient between the two time series. In the final equation, "s" indicates a smooth term of the specified predictor; α_i represents the pathogen-specific intercept.

All GAMs were fit using a Bayesian approach as implemented by the "brms" package (version 2.22.0) (70), which interfaces with the probabilistic programming language Stan (71). For each pair of time series, we ran 2000 warmup iterations, followed by 1000 sampling iterations. In addition to the relationship between the predictor and outcome variables, the bivariate normal model also fits the residual correlation between predictors. We took the median of the posterior distribution of these correlation coefficients to be our point estimate of interaction strength. We constructed 95% confidence intervals by calculating the 2.5th and 97.5th percentiles of this distribution, and estimates were considered to be significant if this interval did not contain 0. As with correlation coefficients calculated from the raw data, this method is incapable of distinguishing between symmetric vs. asymmetric interactions. Datasets where fitting these models resulted in 1) any divergent transitions after warmup, or 2) lack of convergence, as indicated by an R-hat of greater than 1.05 for any model parameter, were removed from consideration before further analysis. Additional methodological details can be found in S1 Text.

Granger Causality

The previous two methods rely on correlations between the time series of two pathogens to determine whether an interaction is present, which, as discussed above, is potentially problematic. Wiener-Granger causality (Granger causality for short) (21,22) instead infers causation based on predictive ability. Specifically, Granger causality measures the extent to which knowledge of the past values of *X*

improves the prediction of *Y* above and beyond the predictive capability of past values of *Y* (and any confounders of interest) alone. Here, we calculate this quantity using vector autoregressive models (VARs), as in Granger's original implementation (21), although this framework can be extended to more complicated time series models (79,80). To account for shared seasonality, we calculated the true extent of seasonal forcing over time according to Equation 1, and included this time series as a predictor in all models. We therefore test for Granger causality by comparing the following two models:

$$Y_t = \sum_{i=1}^{p} a_i Y_{t-i} + \sum_{i=1}^{p} b_i X_{t-i} + S_t + \varepsilon_t, \qquad \varepsilon \sim N(0, \sigma_\varepsilon^2) \quad (4)$$

and

$$Y_t = \sum_{i=1}^{p} \alpha_i Y_{t-i} + S_t + \eta_t, \qquad \eta_t \sim N(0, \sigma_\eta^2) \quad (5)$$

where *p* is the order of the autoregressive models, and *S* is the seasonal forcing component of Equation 1 (i.e., everything except $\overline{\beta}_t$, which is pathogen-specific). The terms ε_t and η_t represent the remaining error in Y_t not explained by the models, which are assumed normally distributed with variance equal to σ_ε^2 and σ_η^2, respectively. If σ_ε^2 is significantly less than σ_η^2, it is said that *X* "Granger causes" *Y*. Despite this terminology, we emphasize that forecasting ability does not necessarily indicate that one variable has a causal effect on another. Thus, *X* may "Granger cause" *Y*, but may not actually have a causal effect on *Y* (81,82).

To apply this method to our synthetic datasets, we first identified the ideal order for each synthetic dataset as the value that minimized the BIC of the fitted VAR models. We allowed for a maximum order of 20, the sum of the upper bounds of the generation time for both pathogens (2 weeks for pathogen A and 5 weeks for pathogen B) and the maximum interaction duration tested (13 weeks) (see S1 Text for a derivation of these values). The magnitude of the estimated effect of one pathogen on the other was then calculated as initially suggested by (83):

$$G_{X \to Y} = ln\left(\frac{\sigma_\eta^2}{\sigma_\varepsilon^2}\right) \quad (6)$$

These values were deemed to be statistically significant if any elements of b_i were non-zero, as evaluated by an F-test.

All analyses were conducted using the “stats” (version 4.4.0), “vars” (version 1.6.1) (72,73), and “VARtests” (version 2.0.5) (74) packages; a link to the full code used can be found under “Implementation” below. Note that, unlike the above methods based on correlations, Granger causality allows for the effect of pathogen A on pathogen B to differ from the effect of pathogen B on pathogen A, meaning that asymmetric interactions can theoretically be identified. However, the effect estimate calculated in Equation 6 is always positive; thus, Granger causality is incapable of distinguishing between positive and negative interactions.

Transfer Entropy

Like Granger causality, transfer entropy (23,84) estimates the extent to which past information on *X* informs the current state of *Y*, above and beyond past information on *Y* itself. However, rather than VARs and other parametric time series models, transfer entropy instead relies on ideas from information theory (85). Specifically, the transfer entropy from *X* to *Y*, conditional on the shared seasonal forcing, is defined as:

$$T_{X\to Y|S}^{(k,l,m)}(t) = H\left(Y_t \middle| \boldsymbol{Y}_{t-1}^{(k)}, \boldsymbol{S}_{t-1}^{(m)}\right) - H\left(Y_t \middle| \boldsymbol{Y}_{t-1}^{(k)}, \boldsymbol{X}_{t-1}^{(l)}, \boldsymbol{S}_{t-1}^{(m)}\right) \quad (7)$$

Here, *k*, *l*, and *m* are the history lengths (i.e., the number of timepoints included in the analysis) for the target (*Y*), source (*X*), and seasonal forcing (*S*) data, respectively. Meanwhile, $H\left(Y_t | Y_{t-1}^{(k)}, S_{t-1}^{(m)}\right)$ is the conditional Shannon entropy (86) of Y_t given $Y_{t-1}^{(k)}$ and $S_{t-1}^{(m)}$, a measure of the amount of uncertainty in the value of *Y* at time *t*, after accounting for the previous history of *Y* up to *k* timepoints in the past and *S* up to *m* timepoints in the past. The term $H\left(Y_{t-1}^{(k)}, X_{t-1}^{(l)}, S_{t-1}^{(m)}\right)$ is similar, this time also conditioning on the past history of *X* up to timepoint *t – l*. Notably, Equation 7 is equivalent to the conditional mutual information between Y_t and $\boldsymbol{X}_{t-1}^{(l)}$ given $\boldsymbol{Y}_{t-1}^{(k)}$ and $\boldsymbol{S}_{t-1}^{(m)}$, $I\left(X_t : \boldsymbol{X}_{t-1}^{(l)} \middle| \boldsymbol{Y}_{t-1}^{(k)}, \boldsymbol{S}_{t-1}^{(m)}\right)$, a measure of the reduction in uncertainty about the value of *Y* at time *t* due to knowledge of the past history of *X*, after accounting for the past history of *Y* and *S* (85). Like with Granger causality, values significantly above zero are taken to imply that X causes Y.

Overall, transfer entropy may be viewed as a nonparametric alternative to Granger causality (84,85,87), and for Gaussian variables, the two are equivalent (88). Because VARs may not perform well when dynamics are highly nonlinear (89), and data transformations can greatly affect the resulting Granger causality estimate (90), transfer entropy may be a more appropriate approach in these cases. However, calculating transfer entropy is more computationally intensive than running the relatively simple VARs necessary to calculate Granger causality (85).

We calculated the transfer entropy for all datasets using the Kraskov-Stögbauer-Grassberger (KSG) technique with four nearest neighbors (91), as implemented by the Java Information Dynamics Toolkit (JIDT) (75). The history length for the target (i.e., outcome) data was fixed to 2 for pathogen A and 5 for pathogen B, based on the generation times for influenza and RSV, respectively; the history length for the source (i.e., predictor) data was set to 20 for both pathogens (see S1 Text). For the shared seasonal forcing, history length was set to 1. Due to the range of interaction durations modeled, we also allowed for lags greater than 1 week between *X* and *Y*; specifically, we tested lags of 1, 2, 4, and 13 weeks. Only results for the lag yielding the highest estimates of transfer entropy are shown (92). Significance was assessed by generating 500 permutations of the data under the null hypothesis, and comparing our point estimates to estimates calculated from these surrogate datasets (75).

Analyses were conducted using JIDT (version 1.6.1) (75) and the R package “rJava” (version 1.0.11) (76). Like Granger causality, point estimates for transfer entropy are always positive, and therefore do not allow for the characterization of positive vs. negative interactions.

Convergent Cross-Mapping

The above methods make sense for systems where *X* and *Y* contain unique information, and therefore the predictability of *Y* in the absence of information about *X* can be assessed simply by removing the time series of *X* from consideration. However, in nonlinear dynamical systems driven by an underlying deterministic skeleton, a time series *Y* will itself contain information on the past values of its cause *X*. In such situations, the utility of Granger causality and transfer entropy is controversial (24,30).

418 Convergent cross-mapping (CCM) is a relatively new approach that takes advantage of this inherent
419 lack of independence between time series (24). Specifically, to determine whether *X* causes *Y*, CCM
420 uses *L* lagged values of *Y* to construct a shadow manifold, M_Y, according to Takens' theorem (93).
421 Then, the method assesses whether the values of *X* can be estimated based on the topology of M_Y. In
422 other words, unlike Granger causality and transfer entropy, which infer that *X* causes *Y* if past values
423 of *X* improve predictions of some current value of *Y*, CCM infers that *X* causes *Y* if current values of *Y*
424 can be used to "predict" current or past values of *X*. Crucially, for causation to be demonstrated, the
425 quality of these predictions must increase with increasing *L*, or the amount of data available for
426 building the shadow manifold (i.e., there must be convergence).
427
428 To perform CCM, we first chose the ideal embedding dimensions (i.e., the number of lags used to
429 construct the shadow manifolds) for each synthetic dataset by determining which values yielded the
430 most accurate within-pathogen predictions for each pathogen. We allowed a maximum embedding
431 dimension of 2 for pathogen A and 5 for pathogen B (see S1 Text). We also allowed for a lagged
432 effect of one pathogen on the other by selecting the negative lag that yielded the highest cross-map
433 skill (94); a maximum lag of 20 weeks was permitted (see S1 Text). The selected values for the
434 embedding dimensions and lags were then used to run CCM with 100 samples for each value of *L*.
435 Point estimates were taken to be the median cross-map skill across all samples for the largest *L*. We
436 note that, since CCM is expected to yield null results for two variables that are not causally linked but
437 have a common environmental driver (24), there is no need to explicitly account for the shared
438 seasonal forcing between pathogens when applying this method.
439
440 We tested two methods of assessing significance. The first approach considers a result to be
441 significant so long as there is convergence, as in CCM's original implementation (24). Specifically, we
442 use a nonparametric bootstrap approach to determine whether the cross-map skill for the maximum
443 value of *L* is significantly greater than the cross-map skill for the minimum value of *L*, as in (36). The
444 second approach, originally suggested by (18), compares the cross-map skill obtained from the data
445 for the maximum value of *L* to a null distribution of cross-map skill calculated using seasonal
446 surrogates of the hypothesized causal pathogen's data. Specifically, we generate 500 surrogate time
447 series by maintaining the seasonal pattern in the data, but reshuffling the residuals. A p-value is then
448 calculated as:

$$p = \frac{\sum_{i=1}^{n}\left(\rho_{true} < \rho_{surr,\ i}\right) + 1}{n + 1} \quad (8)$$

where ρ_{true} is the cross-map skill calculated from the data, $\rho_{surr,i}$ is the cross-map skill calculated using surrogate time series *i*, and *n* is the number of seasonal surrogates considered (30,95). Compared to the first method, this approach may be better suited to systems where shared seasonality presents a challenge to causal inference (18).

Analyses were conducted using the R package "rEDM" (version 1.15.4) (77). As with Granger causality and transfer entropy, cross-map skill is always greater than zero. For this reason, CCM does not allow for positive and negative interactions to be differentiated (although similar methods exist that use scenario exploration to characterize the sign of the effect of one variable on another (18,96)).

Performance Assessment

We assessed the overall performance of each method by calculating the sensitivity (i.e., the proportion of datasets where the presence of an interaction is correctly inferred) and specificity (i.e., the proportion of datasets where the absence of an interaction is correctly inferred). Here, we considered results using Pearson correlation coefficients and GAMs to be accurate only if they correctly identified the sign of the interaction; for all other methods, which only yield positive point estimates, accuracy was assessed on the basis of significance alone. While this decision technically disadvantages correlation coefficients and GAMs, we feel that it is also important to assess the accuracy of the extra information these methods capture about interaction sign.

To determine whether the true underlying interaction parameters influenced accuracy, we also calculated the percentage of synthetic datasets correctly classified by each method for each combination of interaction parameters. Again, for Pearson correlation coefficients and GAMs, we considered both significance and sign in assessing accuracy.

If the tested methods are correctly identifying the relative strength of the modeled interactions, we expect that the methods will yield larger point estimates for datasets where the strength of the

underlying interaction is greater. To assess this, we fit a linear mixed-effects regression model on the point estimates obtained from each method, using true strength, true duration, and the interaction between them as predictors, as well as a random intercept on the parameter set. To account for the fact that point estimates from some methods are necessarily positive, we took the reciprocal of true interaction strength for negative interactions prior to fitting; interactions with strength 0 (i.e., complete inhibition) were assigned a value of 16. We also log2-transformed both the absolute value of the point estimates, and the true interaction strengths, such that the regression models captured the relative change in outcome resulting from a doubling of interaction strength. This was done because the range of point estimates produced varied by method.

Results

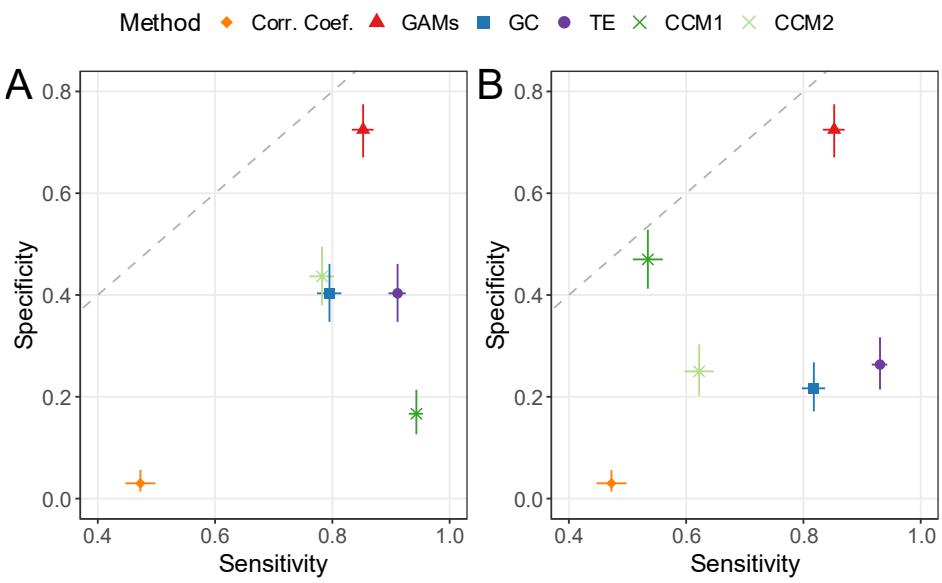
Symmetric interactions between pathogens led to observable differences in epidemic dynamics

Changing the strength and duration of the interaction between our two modeled pathogens led to observable differences in the dynamics of both pathogens, particularly for long-lasting interactions (Figure 2A, Figure S1). This suggests that outbreak data, which are collected at the population level, may nonetheless contain signal of individual-level interactions. When holding all other model parameters constant, negative interactions generally reduced both the average attack rate and the average peak week of pathogen B. In contrast, positive interactions had the opposite effect (Figure 2B). However, this was not always the case: long-lasting negative interactions could push outbreaks later, and vice versa. Furthermore, while the timing and size of outbreaks of pathogen B were typically highly consistent across seasons, long-lasting negative interactions greatly increased year-to-year variability. Similar patterns were observed for pathogen A (Figure S2). These findings illustrate a key point from the introduction: although interactions can impact infectious disease transmission, these effects can be subtle and their direction unintuitive.

Non-mechanistic methods applied to incidence data frequently fail in determining whether or not an interaction between two pathogens exists

The overall accuracy of all methods at identifying the presence or absence of an interaction effect is displayed in Figure 3. For Pearson correlation coefficients and GAMs, negative interactions were only considered to be correctly inferred if the resulting point estimate was also negative; for all other

510 methods, which can only yield positive point estimates, results were considered accurate for negative
511 interactions as long as point estimates were significantly greater than zero. For Granger causality and
512 transfer entropy, only the results of tests accounting for seasonality are shown (see Figure S3 for
513 results ignoring shared seasonality). Of the tested methods, GAMs performed best, with 85.2%
514 sensitivity and 72.5% specificity. All other methods showed much worse performance: while sensitivity
515 was often high, specificity was below 50% for all other methods. In other words, among simulations
516 where no interaction was modeled, all methods other than GAMs yielded false positives more often
517 than not. For most methods capable of distinguishing interaction direction (Granger causality, transfer
518 entropy, and CCM using seasonal surrogates), false positives were more common when trying to
519 classify the interaction effect of pathogen B on pathogen A. CCM using convergence as the criterion
520 for significance was the only exception, but improved specificity came at the cost of sensitivity. Overall
521 performance quality was similar for these three methods; as expected (20,97), Pearson correlation
522 coefficients were by far the least accurate.
523



524

525 **Figure 3. Sensitivity and specificity of all tested methods.** Results for the interaction effect of
526 pathogen A on pathogen B are shown in (A), while results for the effect of pathogen B on A are
527 shown in (B). Method is indicated by both point shape and color; CCM1 refers to the method
528 assessing significance based on convergence, while CCM2 refers to the method assessing
529 significance using seasonal surrogates. The crosshairs on each point indicate 95% confidence
530 intervals, obtained using a binomial test. The dashed diagonal line shows where sensitivity is equal to
531 specificity. Note that the x-axis begins at 0.4. For correlation coefficients and GAMs, negative

interactions were only considered to be correctly detected if the associated point estimate was also negative; all other methods cannot distinguish between positive and negative interactions, and so were considered to have correctly identified an interaction if the point estimate was significantly above 0. For Granger causality and transfer entropy, only results for implementations controlling for shared seasonality are shown.

Accuracy was dependent on the true strength and duration of the interaction

Accuracy for all methods by true strength and duration of the modeled interaction is plotted in Figure 4. In general, accuracy was higher for interactions with longer duration, perhaps due to increased signal of longer interactions in the data, especially when interaction sign was negative. Accuracy was typically similar for positive and negative interactions, although CCM performed better for positive interactions, at least when the effect of pathogen B on pathogen A was being tested (Figure 4H, J), and correlation coefficients performed particularly poorly for negative interactions (Figure 4A). Interaction strength had little effect on accuracy for all of the tested methods.

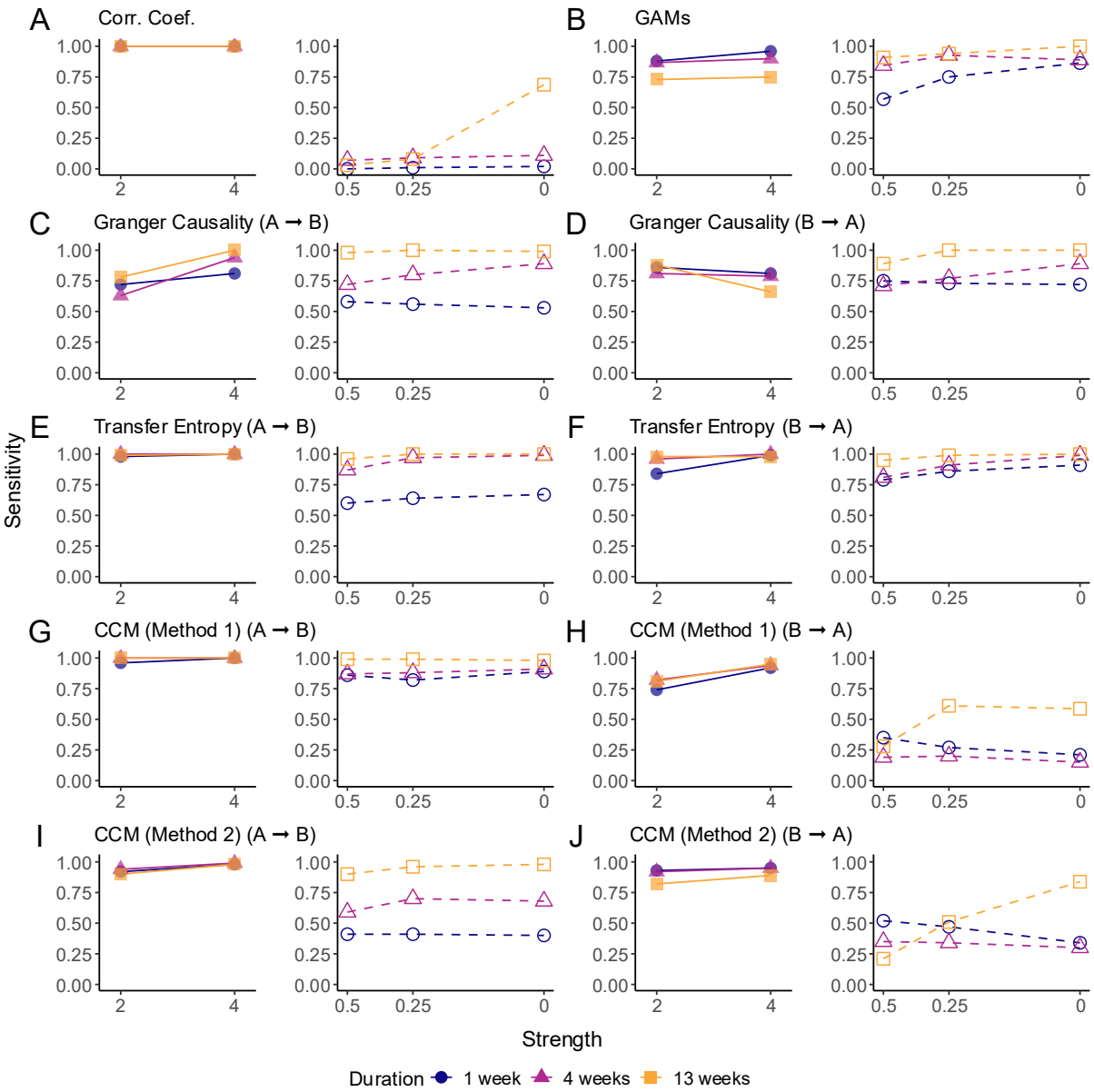


Figure 4. Sensitivity of all methods by true interaction strength and duration. Results are shown for correlation coefficients (A), GAMs (B), Granger causality (C and D), transfer entropy (E and F), CCM using convergence to assess significance (G and H), and CCM using seasonal surrogates to assess significance (I and J); for methods capable of distinguishing directionality, results for the effect of pathogen A on pathogen B are shown in (C), (E), (G), and (I), while results for the effect of B on A are shown in (D), (F), (H), and (J). For each method, results for positive interactions are shown in the left panel, and results for negative interactions are shown on the right. True interaction strength is shown on the x-axis, arranged from weaker interactions on the left to stronger interactions on the right. Results from simulations generated with positive interactions are displayed as filled points and connected with solid lines, while results from simulations with negative interactions are displayed as

hollow points and connected with dotted lines. Point and line color indicate the true interaction duration.

Magnitude of the point estimates was not consistently associated with true interaction strength

In addition to identifying whether an interaction between two pathogens exists, methods should ideally provide some information about the interaction. Figure 5 shows the extent to which a doubling of the true interaction strength used for data generation yields a concomitant increase in the point estimates from each method tested. For both GAMs and transfer entropy, we consistently find a significant, positive relationship between the true interaction strength and the inferred point estimates, suggesting that these methods can, to some extent, characterize the relative strength of an interaction on the basis of observed case data. As above, performance was better for longer-lasting interactions. However, it is important to note that, for transfer entropy, the variation in the magnitude of point estimates across all true interaction strengths is small. For GAMs, the differences in magnitude are more substantial, although a large amount of overlap still exists (Figure S4). Furthermore, a doubling in the true interaction strength consistently leads to a much smaller increase in the point estimates returned by these methods, particularly for short-lived interactions. Thus, while these methods may correctly infer the relative interaction strength among many datasets, they are unlikely to yield meaningful conclusions about absolute interaction strength based on real-world datasets. Encouragingly, the sign of the true interaction and of the estimate returned by the GAM approach overwhelmingly agree (Figure 4, Figure S4). For Granger causality, a significant association exists for stronger interactions, but the association for shorter interactions is null. Interestingly, for both correlation coefficients and CCM, we often find a significant negative relationship, indicating that, for these methods, larger point estimates are found for data generated using weaker interactions.

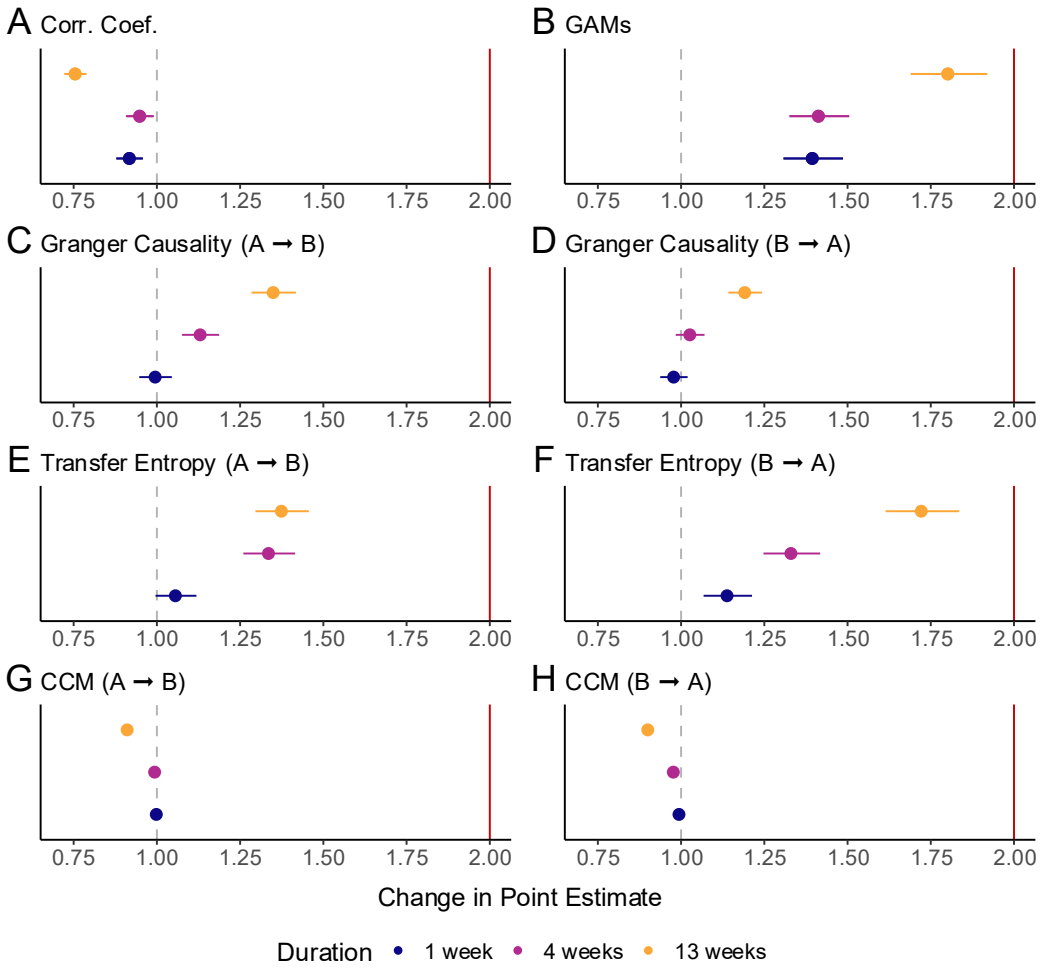


Figure 5. Multiplicative change in point estimates returned in response to a doubling of true interaction strength. Results are shown for all methods ((A) correlation coefficients, (B) GAMs, (C) Granger causality (impact of A on B), (D) Granger causality (impact of B on A), (E) transfer entropy (A on B) (F) transfer entropy (B on A), (G) CCM (A on B), (H) CCM (B on A)). Points show the effect estimate obtained from the linear regression model on point estimates from each method, and lines show the 95% confidence intervals. Point color indicates the true duration of the interaction. The red vertical line at 2.0 indicates the expected result if methods work perfectly (i.e., a doubling in the point estimates when the true strength is doubled), while the dotted gray vertical line at 1.0 indicates no change. Because these results are based only on the point estimates and not on whether the results are significant, results for both CCM methods are equivalent, and are therefore not shown separately.

Discussion

Although pathogen-pathogen interactions are common, they are notoriously difficult to study in human populations. Due to the nonlinear dynamics underlying infectious disease transmission, the

pervasiveness of confounding, and the inherent complexity of interactions themselves, many simple statistical methods fail when applied to questions about interactions (19,20). Here, we tested five non-mechanistic statistical methods and evaluated their ability to identify and characterize the interaction between two pathogens using synthetic data. We found that even these more sophisticated methods often failed to detect whether an interaction affecting the force of infection was present at all. In cases where an interaction was present, information about the relative strength of the interaction provided by these methods was ambiguous at best.

Surprisingly, out of all of the methods tested, GAMs achieved the highest performance: in addition to correctly identifying the presence or absence of an interaction across the full range of interaction parameters tested, GAMs could also provide information about the sign and relative strength of an interaction. This was true even though this is fundamentally a regression-based approach, which is not expected to perform well in causal inference even if shared seasonality is controlled for (17). It is unclear why GAMs exhibit such high performance. One possibility is that, because GAMs cannot consider the effect of A on B and of B on A separately, they are inherently observing the interaction effects of both pathogens simultaneously. This could give them an advantage in the case of a symmetric interaction. However, in a sensitivity analysis where the effect of pathogen B on pathogen A was set to 1.0 (i.e., no effect) for all simulations, we observed similar accuracy (Figure S5). Of course, the fact that GAMs are incapable of determining the direction of an interaction remains a significant weakness. Furthermore, our implementation of GAMs assumes a regular, seasonal cycle for both pathogens, and accuracy is highly dependent on the strength of the underlying seasonal forcing (Figure S7).

In contrast, although several of the tested methods (Granger causality, transfer entropy, and CCM) were developed to perform causal inference on time series data, they displayed notably worse performance. Although these results may seem counterintuitive, they are in line with past work. Cobey and Baskerville (36) found that CCM often incorrectly identified cross-immunity between strains where none was modeled. When applied to real-world data on several childhood pathogens in two cities, no common interaction effects were identified between locations. Likewise, Randuineau (37) found that neither Granger causality nor transfer entropy could correctly characterize the direction or strength of unidirectional interactions. It is interesting to note that, despite being theoretically better-suited to data

from systems with underlying deterministic dynamics (24), CCM performs similarly to Granger causality overall, and is slightly outperformed by transfer entropy. This echoes the results of Barraquand et al. (30), who found that Granger causality and CCM were often in agreement when identifying interactions in ecological data. Results were largely unchanged when we considered 20 (rather than 10) years of data (Figure S6), perhaps because, due to the cyclical nature of our data, little new information is gained by examining additional seasons. In the case of CCM using seasonal surrogates, the inclusion of an additional decade of data made it almost impossible to identify null interactions (Figure S6).

We emphasize that our results do not imply that these methods are inherently ineffective. Indeed, given that these methods are successfully applied in several other contexts (24,29,30,35,98), it is particularly important to consider why they perform so poorly when considering pathogen-pathogen interactions specifically. The existence of shared seasonal forcing between our two pathogens is likely a key factor. Past work has shown that both Granger causality and CCM have a high false-positive rate when a strongly autocorrelated shared driver is at play, even when the shared driver is controlled for (30). We tested two implementations of Granger causality and transfer entropy: one accounting for the strength of the shared driver and one ignoring it. While controlling for the strength of seasonal forcing greatly reduced the number of false positives identified by both Granger causality and transfer entropy (Figure S3), specificity remained low, particularly for tests of the effect of pathogen B on pathogen A, suggesting that fully accounting for the influence of a shared driver is challenging. CCM theoretically accounts for confounding implicitly (24). However, as mentioned above, CCM performed no better than Granger causality or transfer entropy. Notably, our sensitivity analyses revealed no consistent relationship between the extent of shared forcing and method performance (Figure S7), suggesting that even very slight confounding can be a substantial problem.

This is concerning, as shared seasonal drivers are incredibly common among infectious diseases, including among those hypothesized to interact (99,100). For example, evidence suggests that transmission of both influenza and RSV, on which the pathogens in our model are based, is higher when temperatures are lower, and when humidity is either low or very high (42–44,46). Temperature and humidity are also likely to play a role in the transmission of SARS-CoV-2 (101), as well as seasonal coronaviruses (102) and rhinoviruses (103). Because this is a simulation study, we were able

659 to control for the exact strength of seasonal forcing over time. However, in real-world outbreaks, the
660 exact relationship between potential drivers (e.g., temperature) and pathogen transmission rates is not
661 known, and seasonal patterns in pathogen incidence are not necessarily a proxy for the underlying
662 forcing variable (17,18).
663
664 The impact of observation noise on method performance should also be considered, as the utility of
665 these methods may be limited by data that are not perfectly observed (24,81). Unfortunately, for many
666 infectious diseases, real-world outbreak data are notoriously error-laden: surveillance systems relying
667 on healthcare utilization and testing will miss mild and asymptomatic cases, and sampling effort is
668 often much lower outside of the outbreak season (1). Interestingly, our sensitivity analyses suggest
669 that reducing observation noise to zero does not have a purely beneficial impact on method
670 performance. Rather, lower observation noise tended to reduce the number of false negatives, but
671 increased false positives, while increasing observation noise had the opposite effect (Figure S6).
672 Furthermore, this pattern was not consistent across all methods, suggesting that the exact impact of
673 noisy observations is difficult to predict. In contrast, modifying the amount of process noise had little
674 effect on performance (Figure S6), a result that is mostly in line with Cobey and Baskerville (36).
675
676 It is also worth discussing that, although all of the interactions we modeled were symmetric, most
677 methods were more likely to correctly classify the effect of pathogen A on pathogen B than the effect
678 of B on A. This could be, in part, due to the seasonal driver. CCM is known to be unreliable when a
679 downstream variable becomes synchronized to a forcing variable (24); in cases involving a shared
680 driver, this is because this leads to data (here, observed cases data) that closely resemble the driver,
681 potentially leading to false positives (36). In our simulated data, the strength of seasonal forcing is
682 much more highly correlated with observed cases of pathogen B than of pathogen A, perhaps
683 explaining the discrepancy in our results. Requiring the optimal cross-map lag to be negative as an
684 additional significance criterion can mitigate this issue (36,94), but, for systems like seasonal infectious
685 disease outbreaks, which have high periodicity, this approach is not valid (104). Because most
686 pathogen-pathogen interactions are not well-understood, it is often unclear whether the interaction is
687 uni- or bidirectional. For this reason, failure to accurately assess directionality represents a significant
688 drawback of these methods. Even GAMs, which performed very well otherwise, cannot help here, as
689 they offer no information about directionality.

690

691 Ultimately, inferences about pathogen-pathogen interactions on the basis of incidence data may be
692 particularly challenging simply because these data are related to the interaction of interest through a
693 long and indirect causal chain (see Vignette 1 in (17) for a related discussion). While the exact
694 mechanisms are rarely understood, many interactions are likely the result of processes such as direct
695 or indirect competition, modulation of the host's innate or adaptive immune system, or alteration of the
696 host's microbiome (8,16). These individual-level processes may impact how likely one is to become
697 infected with or to transmit a second pathogen (modeled here as a direct impact of the interaction on
698 force of infection). This in turn will impact the total number of people in a population who become
699 infected, which is itself observed with noise. This is a much less direct process than, say, predator-
700 prey interactions, for which methods like Granger causality and CCM can perform well (24,30).
701 Unfortunately, we only observe the final outcome of this causal chain: the observed number of cases
702 of each pathogen in a population at each time point. It is possible that the tested methods struggle to
703 identify any signal of the interaction from these downstream data.
704

705 Given these challenges, none of the tested methods appear to be promising tools for characterizing
706 pathogen-pathogen interactions on the basis of infectious disease outbreak data. Approaches utilizing
707 mechanistic models of infectious disease transmission, like the one used to generate the synthetic
708 data for this study, may represent an attractive alternative. By using equations to describe the
709 transmission of pathogens throughout a population, these models explicitly account for nonlinear
710 dynamics, as well as for potential confounders. Models can then be fit to surveillance data to gain
711 insight about mechanisms of interest, including interaction effects. Although this can be
712 computationally intensive, and improvements in the efficiency of methods for model fitting are
713 necessary to handle, for example, models of greater than two cocirculating pathogens (9), mechanistic
714 modeling methods show clear promise. In fact, this approach has already been used to better
715 understand interactions between influenza and RSV (1,105), influenza and *S. pneumoniae* (12,106),
716 influenza and SARS-CoV-2 (6), RSV and both hMPV and human parainfluenza virus (HPIV) (9,107),
717 and different serotypes of dengue (108) and cholera (66). A key advantage of these methods over the
718 ones tested in this work is that they can provide quantitative estimates and confidence intervals for
719 specific characteristics of the interaction, such as its strength and duration.
720

We acknowledge several limitations of the current work. In particular, we have focused here on interactions between two epidemic viruses with short infectious periods and with strong periodicity in their transmission dynamics. The methods tested here may perform better for, for example, interactions involving endemic colonizing bacteria, for which carriage may persist for several weeks and prevalence is consistently high (8,16). Likewise, a modeling study found that year-to-year variability in peak incidence was critical in allowing interaction effects to be inferred from population-level data (12); more promising results may therefore be expected for pathogens showing less regular outbreak patterns. Our model was also more likely to produce outbreaks where pathogen B peaked before pathogen A (81% of datasets where no interaction was modeled) than the other way around, which could have impacted our ability to infer in particular the impact of pathogen A on pathogen B (1). However, we found little influence of the difference in peak timing between pathogens on the ability of the tested methods to identify the impact of pathogen A on pathogen B (Figure S8). Only Granger causality performed better for datasets where pathogen A peaked first, and the effect was slight; interestingly, CCM using seasonal surrogates performed slightly worse for datasets where pathogen A peaked first. Finally, although we have tested several popular methods, there are others that we have not considered (109,110). Our results do not necessarily imply that there are no non-mechanistic methods available that could provide accurate information about interactions. Rather, they emphasize the importance of rigorously testing any existing or novel methods, using an approach similar to that described in this work, before these methods are used to draw conclusions about pathogen-pathogen interactions from real-world data.

We have shown that, when confronted with infectious disease outbreak data, even sophisticated nonmechanistic methods developed with causal inference in mind are incapable of reliably identifying and characterizing pathogen-pathogen interactions. We conclude that these methods are not appropriate tools for understanding interactions on the basis of outbreak data, at least for interactions between highly seasonal, acute viruses like the one modeled here, and that interaction effects inferred using these methods should be viewed with an abundance of caution. Rather, a combination of laboratory studies and mechanistic modeling approaches is likely to be necessary to fully characterize both the individual- and population-level impact of pathogen-pathogen interactions.

Acknowledgments

The authors gratefully acknowledge Anabelle Wong, Pietro Gemo, and Franziska Frederking for their helpful comments on the manuscript.

755 References

756 1. Kramer SC, Pirikahu S, Casalegno JS, Domenech de Cellès M. Characterizing the interactions
757 between influenza and respiratory syncytial viruses and their implications for epidemic control. Nat
758 Commun. 2024 Nov 20;15(1):10066.
759 2. Chan KF, Carolan LA, Korenkov D, Druce J, McCaw J, Reading PC, et al. Investigating Viral
760 Interference Between Influenza A Virus and Human Respiratory Syncytial Virus in a Ferret Model
761 of Infection. J Infect Dis. 2018 Jul 2;218(3):406–17.
762 3. Wu A, Mihaylova VT, Landry ML, Foxman EF. Interference between rhinovirus and influenza A
763 virus: a clinical data analysis and experimental infection study. Lancet Microbe. 2020
764 Oct;1(6):e254–62.
765 4. Gonzalez AJ, Ijezie EC, Balemba OB, Miura TA. Attenuation of influenza A virus disease severity
766 by viral coinfection in a mouse model. J Virol [Internet]. 2018 Dec 1;92(23). Available from:
767 http://dx.doi.org/10.1128/JVI.00881-18
768 5. Essaidi-Laziosi M, Geiser J, Huang S, Constant S, Kaiser L, Tapparel C. Interferon-dependent
769 and respiratory virus-specific interference in dual infections of airway epithelia. Sci Rep. 2020 Jun
770 24;10(1):10246.
771 6. Domenech de Cellès M, Casalegno JS, Lina B, Opatowski L. Estimating the impact of influenza
772 on the epidemiological dynamics of SARS-CoV-2. PeerJ. 2021 Dec 3;9:e12566.
773 7. Dee K, Goldfarb DM, Haney J, Amat JAR, Herder V, Stewart M, et al. Human Rhinovirus infection
774 blocks severe acute respiratory syndrome Coronavirus 2 replication within the respiratory
775 epithelium: Implications for COVID-19 epidemiology. J Infect Dis. 2021 Jul 2;224(1):31–8.
776 8. Wong A, Barrero Guevara LA, Goult E, Briga M, Kramer SC, Kovacevic A, et al. The interactions
777 of SARS-CoV-2 with cocirculating pathogens: Epidemiological implications and current knowledge
778 gaps. PLoS Pathog. 2023 Mar;19(3):e1011167.
779 9. Bhattacharyya S, Gesteland PH, Korgenski K, Bjørnstad ON, Adler FR. Cross-immunity between
780 strains explains the dynamical pattern of paramyxoviruses. Proc Natl Acad Sci U S A. 2015 Oct
781 27;112(43):13396–400.
782 10. Geiser J, Boivin G, Huang S, Constant S, Kaiser L, Tapparel C, et al. RSV and HMPV infections
783 in 3D tissue cultures: Mechanisms involved in virus-host and virus-virus interactions. Viruses.
784 2021 Jan 19;13(1):139.
785 11. Wen X, Suryadevara N, Kose N, Liu J, Zhan X, Handal LS, et al. Potent cross-neutralization of

786 respiratory syncytial virus and human metapneumovirus through a structurally conserved antibody
787 recognition mode. Cell Host Microbe. 2023 Aug 9;31(8):1288–300.e6.
788 12. Shrestha S, Foxman B, Berus J, van Panhuis WG, Steiner C, Viboud C, et al. The role of
789 influenza in the epidemiology of pneumonia. Sci Rep. 2015 Oct 21;5(1):15314.
790 13. Avadhanula V, Rodriguez CA, Devincenzo JP, Wang Y, Webby RJ, Ulett GC, et al. Respiratory
791 viruses augment the adhesion of bacterial pathogens to respiratory epithelium in a viral species-
792 and cell type-dependent manner. J Virol. 2006 Feb;80(4):1629–36.
793 14. Mina MJ, Metcalf CJE, de Swart RL, Osterhaus ADME, Grenfell BT. Long-term measles-induced
794 immunomodulation increases overall childhood infectious disease mortality. Science. 2015 May
795 8;348(6235):694–9.
796 15. Pawlowski A, Jansson M, Sköld M, Rottenberg ME, Källenius G. Tuberculosis and HIV co-
797 infection. PLoS Pathog. 2012 Feb;8(2):e1002464.
798 16. Opatowski L, Baguelin M, Eggo RM. Influenza interaction with cocirculating pathogens and its
799 impact on surveillance, pathogenesis, and epidemic profile: A key role for mathematical
800 modelling. PLoS Pathog. 2018 Feb;14(2):e1006770.
801 17. Barrero Guevara LA, Kramer SC, Kurth T, Domenech de Cellès M. Causal inference concepts
802 can guide research into the effects of climate on infectious diseases. Nat Ecol Evol. 2025
803 Feb;9(2):349–63.
804 18. Deyle ER, Maher MC, Hernandez RD, Basu S, Sugihara G. Global environmental drivers of
805 influenza. Proc Natl Acad Sci U S A. 2016 Nov 15;113(46):13081–6.
806 19. Domenech de Cellès M, Goult E, Casalegno JS, Kramer SC. The pitfalls of inferring virus–virus
807 interactions from co-detection prevalence data: application to influenza and SARS-CoV-2.
808 Proceedings of the Royal Society B: Biological Sciences. 2022 Jan 12;289(1966):20212358.
809 20. Shrestha S, King AA, Rohani P. Statistical inference for multi-pathogen systems. PLoS Comput
810 Biol. 2011 Aug;7(8):e1002135.
811 21. Granger CWJ. Investigating causal relations by econometric models and cross-spectral methods.
812 Econometrica. 1969 Aug;37(3):424.
813 22. Wiener N. The Theory of Prediction. In: F: Beckenback E, editor. Modern Mathematics for the
814 Engineer. New York, NY: McGraw-Hill Book Company; 1956.
815 23. Schreiber T. Measuring information transfer. Phys Rev Lett. 2000 Jul 10;85(2):461–4.
816 24. Sugihara G, May R, Ye H, Hsieh CH, Deyle E, Fogarty M, et al. Detecting causality in complex

ecosystems. Science. 2012 Oct 26;338(6106):496–500.

25. Wood SN. Generalized Additive Models: An Introduction with R, Second Edition. CRC Press; 2017. 496 p.
26. Hoover KD. Causality in Economics and Econometrics. In: The New Palgrave Dictionary of Economics. London: Palgrave Macmillan UK; 2008. p. 1–13.
27. Sims CA. Money, Income, and Causality. Am Econ Rev. 1972;62(4):540–52.
28. Seth AK, Barrett AB, Barnett L. Granger causality analysis in neuroscience and neuroimaging. J Neurosci. 2015 Feb 25;35(8):3293–7.
29. Timme NM, Lapish C. A tutorial for information theory in neuroscience. eNeuro [Internet]. 2018 May;5(3). Available from: http://dx.doi.org/10.1523/ENEURO.0052-18.2018
30. Barraquand F, Picoche C, Detto M, Hartig F. Inferring species interactions using Granger causality and convergent cross mapping. Theor Ecol. 2021 Mar;14(1):87–105.
31. Singh NK, Borrok DM. A Granger causality analysis of groundwater patterns over a half-century. Sci Rep. 2019 Sep 6;9(1):12828.
32. Sasaki T, Collins SL, Rudgers JA, Batdelger G, Baasandai E, Kinugasa T. Dryland sensitivity to climate change and variability using nonlinear dynamics. Proc Natl Acad Sci U S A. 2023 Aug 29;120(35):e2305050120.
33. Lee S, Lee B, Lee J, Song J, McCarty GW. Detecting causal relationship of non-floodplain wetland hydrologic connectivity using convergent cross mapping. Sci Rep. 2023 Oct 11;13(1):17220.
34. Nova N, Deyle ER, Shocket MS, MacDonald AJ, Childs ML, Rypdal M, et al. Susceptible host availability modulates climate effects on dengue dynamics. Ecol Lett. 2021 Mar;24(3):415–25.
35. Kissler SM, Viboud C, Grenfell BT, Gog JR. Symbolic transfer entropy reveals the age structure of pandemic influenza transmission from high-volume influenza-like illness data. J R Soc Interface. 2020 Mar;17(164):20190628.
36. Cobey S, Baskerville EB. Limits to causal inference with state-space reconstruction for infectious disease. PLoS One. 2016 Dec 28;11(12):e0169050.
37. Randuineau B. Interactions between pathogens: what are the impacts on public health? [Internet]. Université Pierre et Marie Curie - Paris VI; 2015. Available from: https://theses.hal.science/tel-01487918
38. R Core Team. R: A Language and Environment for Statistical Computing [Internet]. Vienna,

Austria: R Foundation for Statistical Computing; 2024. Available from: https://www.R-project.org/

39. Kramer SC. Statistical methods for inferring pathogen-pathogen interactions [Internet]. 2025. Available from: https://github.com/sarahckramer/Pathogen_interaction_simulations

40. Drori Y, Jacob-Hirsch J, Pando R, Glatman-Freedman A, Friedman N, Mendelson E, et al. Influenza A Virus Inhibits RSV Infection via a Two-Wave Expression of IFIT Proteins. Viruses [Internet]. 2020 Oct 16;12(10). Available from: http://dx.doi.org/10.3390/v12101171

41. Bloom-Feshbach K, Alonso WJ, Charu V, Tamerius J, Simonsen L, Miller MA, et al. Latitudinal variations in seasonal activity of influenza and respiratory syncytial virus (RSV): a global comparative review. PLoS One. 2013 Feb 14;8(2):e54445.

42. Lowen AC, Mubareka S, Steel J, Palese P. Influenza virus transmission is dependent on relative humidity and temperature. PLoS Pathog. 2007 Oct 19;3(10):1470–6.

43. Shaman J, Kohn M. Absolute humidity modulates influenza survival, transmission, and seasonality. Proc Natl Acad Sci U S A. 2009 Mar 3;106(9):3243–8.

44. Yuan H, Kramer SC, Lau EHY, Cowling BJ, Yang W. Modeling influenza seasonality in the tropics and subtropics. PLoS Comput Biol. 2021 Jun;17(6):e1009050.

45. Tamerius JD, Shaman J, Alonso WJ, Bloom-Feshbach K, Uejio CK, Comrie A, et al. Environmental predictors of seasonal influenza epidemics across temperate and tropical climates. PLoS Pathog. 2013 Mar;9(3):e1003194.

46. Baker RE, Mahmud AS, Wagner CE, Yang W, Pitzer VE, Viboud C, et al. Epidemic dynamics of respiratory syncytial virus in current and future climates. Nat Commun. 2019 Dec 4;10(1):5512.

47. Paynter S, Yakob L, Simões EAF, Lucero MG, Tallo V, Nohynek H, et al. Using mathematical transmission modelling to investigate drivers of respiratory syncytial virus seasonality in children in the Philippines. PLoS One. 2014 Feb 27;9(2):e90094.

48. Bedford T, Suchard MA, Lemey P, Dudas G, Gregory V, Hay AJ, et al. Integrating influenza antigenic dynamics with molecular evolution. Elife. 2014 Feb 4;3:e01914.

49. Yaari R, Katriel G, Huppert A, Axelsen JB, Stone L. Modelling seasonal influenza: the role of weather and punctuated antigenic drift. J R Soc Interface. 2013 Jul 6;10(84):20130298.

50. Biggerstaff M, Cauchemez S, Reed C, Gambhir M, Finelli L. Estimates of the reproduction number for seasonal, pandemic, and zoonotic influenza: a systematic review of the literature. BMC Infect Dis. 2014 Sep 4;14:480.

51. Reis J, Shaman J. Simulation of four respiratory viruses and inference of epidemiological

879 parameters. Infect Dis Model. 2018 Mar 19;3:23–34.
880 52. Lessler J, Reich NG, Brookmeyer R, Perl TM, Nelson KE, Cummings DAT. Incubation periods of
881 acute respiratory viral infections: a systematic review. Lancet Infect Dis. 2009 May;9(5):291–300.
882 53. Carrat F, Vergu E, Ferguson NM, Lemaitre M, Cauchemez S, Leach S, et al. Time lines of
883 infection and disease in human influenza: a review of volunteer challenge studies. Am J
884 Epidemiol. 2008 Apr 1;167(7):775–85.
885 54. Munywoki PK, Koech DC, Agoti CN, Kibirige N, Kipkoech J, Cane PA, et al. Influence of age,
886 severity of infection, and co-infection on the duration of respiratory syncytial virus (RSV) shedding.
887 Epidemiol Infect. 2015 Mar;143(4):804–12.
888 55. Moore HC, Jacoby P, Hogan AB, Blyth CC, Mercer GN. Modelling the seasonal epidemics of
889 respiratory syncytial virus in young children. PLoS One. 2014 Jun 26;9(6):e100422.
890 56. Hogan AB, Glass K, Moore HC, Anderssen RS. Exploring the dynamics of respiratory syncytial
891 virus (RSV) transmission in children. Theor Popul Biol. 2016 Aug;110:78–85.
892 57. Hodgson D, Pebody R, Panovska-Griffiths J, Baguelin M, Atkins KE. Evaluating the next
893 generation of RSV intervention strategies: a mathematical modelling study and cost-effectiveness
894 analysis. BMC Med. 2020 Nov 18;18(1):348.
895 58. Zheng Z, Weinberger DM, Pitzer VE. Predicted effectiveness of vaccines and extended half-life
896 monoclonal antibodies against RSV hospitalizations in children. NPJ Vaccines. 2022 Oct
897 27;7(1):127.
898 59. Krauer F, Guenther F, Treskova-Schwarzbach M, Schoenfeld V, Koltai M, Jit M, et al.
899 Effectiveness and efficiency of immunisation strategies to prevent RSV among infants and older
900 adults in Germany: a modelling study. BMC Med. 2024 Oct 17;22(1):478.
901 60. Galanti M, Comito D, Ligon C, Lane B, Matienzo N, Ibrahim S, et al. Active surveillance
902 documents rates of clinical care seeking due to respiratory illness. Influenza Other Respi Viruses.
903 2020 Sep;14(5):499–506.
904 61. Cohen C, Kleynhans J, Moyes J, McMorrow ML, Treurnicht FK, Hellferscee O, et al.
905 Asymptomatic transmission and high community burden of seasonal influenza in an urban and a
906 rural community in South Africa, 2017-18 (PHIRST): a population cohort study. Lancet Glob
907 Health. 2021 Jun;9(6):e863–74.
908 62. Brugger J, Althaus CL. Transmission of and susceptibility to seasonal influenza in Switzerland
909 from 2003 to 2015. Epidemics. 2020 Mar;30(100373):100373.

63. Roberts MG, Nishiura H. Early estimation of the reproduction number in the presence of imported cases: pandemic influenza H1N1-2009 in New Zealand. PLoS One. 2011 May 26;6(5):e17835.
64. Dorigatti I, Cauchemez S, Pugliese A, Ferguson NM. A new approach to characterising infectious disease transmission dynamics from sentinel surveillance: application to the Italian 2009-2010 A/H1N1 influenza pandemic. Epidemics. 2012 Mar;4(1):9–21.
65. He D, Ionides EL, King AA. Plug-and-play inference for disease dynamics: measles in large and small populations as a case study. J R Soc Interface. 2010 Feb 6;7(43):271–83.
66. Bretó C, He D, Ionides EL, King AA. Time series analysis via mechanistic models. Ann Appl Stat. 2009 Mar 1;3(1):319–48.
67. King AA, Nguyen D, Ionides EL. Statistical Inference for Partially Observed Markov Processes via the R Package pomp. J Stat Softw. 2016 Mar 29;69:1–43.
68. Said SE, Dickey DA. Testing for unit roots in autoregressive-moving average models of unknown order. Biometrika. 1984 Dec;71(3):599.
69. Kwiatkowski D, Phillips PCB, Schmidt P, Shin Y. Testing the null hypothesis of stationarity against the alternative of a unit root. J Econom. 1992 Oct;54(1-3):159–78.
70. Bürkner PC. Brms: An R package for Bayesian multilevel models using Stan. J Stat Softw [Internet]. 2017;80(1). Available from: http://dx.doi.org/10.18637/jss.v080.i01
71. Stan Development Team. Stan Reference Manual [Internet]. 2024. Available from: https://mc-stan.org
72. Pfaff B. VAR, SVAR and SVEC Models: Implementation Within R Package vars. J Stat Softw [Internet]. 2008;27(4). Available from: http://dx.doi.org/10.18637/jss.v027.i04
73. Pfaff B. Analysis of Integrated and Cointegrated Time Series with R. Second Edition. New York, NY: Springer; 2008.
74. Belfrage M, Catani P, Ahlgren N. VARtests: Tests for Error Autocorrelation, ARCH Errors, and Cointegration in Vector Autoregressive Models [Internet]. 2018. Available from: https://CRAN.R-project.org/package=VARtests
75. Lizier JT. JIDT: An information-theoretic toolkit for studying the dynamics of complex systems. Front Robot AI [Internet]. 2014 Dec 2;1. Available from: http://dx.doi.org/10.3389/frobt.2014.00011
76. Urbanek S. rJava: Low-Level R to Java Interface [Internet]. 2024. Available from: https://CRAN.R-project.org/package=rJava
77. Park J, Smith C, Sugihara G, Deyle E. rEDM: Empirical Dynamic Modeling (’EDM') [Internet].

2024. Available from: https://CRAN.R-project.org/package=rEDM

78. Simpson GL. Modelling Palaeoecological Time Series Using Generalised Additive Models. Frontiers in Ecology and Evolution [Internet]. 2018;6. Available from: https://www.frontiersin.org/articles/10.3389/fevo.2018.00149

79. Detto M, Molini A, Katul G, Stoy P, Palmroth S, Baldocchi D. Causality and persistence in ecological systems: a nonparametric spectral granger causality approach. Am Nat. 2012 Apr;179(4):524–35.

80. Barnett L, Seth AK. Granger causality for state-space models. Phys Rev E Stat Nonlin Soft Matter Phys. 2015 Apr;91(4):040101.

81. Granger CWJ. Testing for causality. J Econ Dyn Control. 1980 Jan;2:329–52.

82. Diebold FX. Elements of Forecasting. Department of Economics, University of Pennsylvania; 2006.

83. Geweke J. Measurement of linear dependence and feedback between multiple time series. J Am Stat Assoc. 1982 Jun;77(378):304–13.

84. Palus M, Komárek V, Hrncír Z, Sterbová K. Synchronization as adjustment of information rates: detection from bivariate time series. Phys Rev E Stat Nonlin Soft Matter Phys. 2001 Apr;63(4 Pt 2):046211.

85. Terry Bossomaier, Lionel Barnett, Michael Harré, Joseph T. Lizier. An Introduction to Transfer Entropy. Cham, Switzerland: Springer Nature; 2016.

86. Shannon CE. A mathematical theory of communication. Bell Syst Tech J. 1948 Oct;27(4):623–56.

87. Amblard PO, Michel O. The relation between Granger causality and directed information theory: A review. Entropy (Basel). 2012 Dec 28;15(1):113–43.

88. Barnett L, Barrett AB, Seth AK. Granger causality and transfer entropy are equivalent for Gaussian variables. Phys Rev Lett. 2009 Dec 4;103(23):238701.

89. Certain G, Barraquand F, Gårdmark A. How do MAR(1) models cope with hidden nonlinearities in ecological dynamics? Methods Ecol Evol. 2018 Sep;9(9):1975–95.

90. Roberts DL, Nord S. Causality tests and functional form sensitivity. Appl Econ. 1985 Feb 1;17(1):135–41.

91. Kraskov A, Stögbauer H, Grassberger P. Estimating mutual information. Phys Rev E Stat Nonlin Soft Matter Phys. 2004 Jun;69(6 Pt 2):066138.

92. Wibral M, Pampu N, Priesemann V, Siebenhühner F, Seiwert H, Lindner M, et al. Measuring

information-transfer delays. PLoS One. 2013 Feb 28;8(2):e55809.

93. Takens F. Detecting strange attractors in turbulence. In: Lecture Notes in Mathematics. Berlin, Heidelberg: Springer Berlin Heidelberg; 1981. p. 366–81. (Lecture notes in mathematics).

94. Ye H, Deyle ER, Gilarranz LJ, Sugihara G. Distinguishing time-delayed causal interactions using convergent cross mapping. Sci Rep. 2015 Oct 5;5(1):14750.

95. North BV, Curtis D, Sham PC. A note on the calculation of empirical P values from Monte Carlo procedures. Am J Hum Genet. 2002 Aug;71(2):439–41.

96. Sugihara G. Nonlinear forecasting for the classification of natural time series. Philos Trans Phys Sci Eng. 1994 Sep 15;348(1688):477–95.

97. Coenen AR, Weitz JS. Limitations of correlation-based inference in complex virus-microbe communities. mSystems [Internet]. 2018 Jul;3(4). Available from: http://dx.doi.org/10.1128/mSystems.00084-18

98. Hannisdal B, Haaga KA, Reitan T, Diego D, Liow LH. Common species link global ecosystems to climate change: dynamical evidence in the planktonic fossil record. Proc Biol Sci [Internet]. 2017 Jul 12;284(1858). Available from: http://dx.doi.org/10.1098/rspb.2017.0722

99. Moriyama M, Hugentobler WJ, Iwasaki A. Seasonality of respiratory viral infections. Annu Rev Virol. 2020 Sep 29;7(1):83–101.

100. Martinez ME. The calendar of epidemics: Seasonal cycles of infectious diseases. PLoS Pathog. 2018 Nov;14(11):e1007327.

101. Morris DH, Yinda KC, Gamble A, Rossine FW, Huang Q, Bushmaker T, et al. Mechanistic theory predicts the effects of temperature and humidity on inactivation of SARS-CoV-2 and other enveloped viruses. Elife [Internet]. 2021 Jul 13;10. Available from: http://dx.doi.org/10.7554/eLife.65902

102. Ijaz MK, Brunner AH, Sattar SA, Nair RC, Johnson-Lussenburg CM. Survival characteristics of airborne human coronavirus 229E. J Gen Virol. 1985 Dec;66 (Pt 12)(12):2743–8.

103. Karim YG, Ijaz MK, Sattar SA, Johnson-Lussenburg CM. Effect of relative humidity on the airborne survival of rhinovirus-14. Can J Microbiol. 1985 Nov;31(11):1058–61.

104. Sugihara G, Deyle ER, Ye H. Reply to Baskerville and Cobey: Misconceptions about causation with synchrony and seasonal drivers. Proc Natl Acad Sci U S A. 2017 Mar 21;114(12):E2272–4.

105. Waterlow NR, Toizumi M, van Leeuwen E, Thi Nguyen HA, Myint-Yoshida L, Eggo RM, et al. Evidence for influenza and RSV interaction from 10 years of enhanced surveillance in Nha Trang,

Vietnam, a modelling study. PLoS Comput Biol. 2022 Jun;18(6):e1010234.

106. Domenech de Cellès M, Arduin H, Lévy-Bruhl D, Georges S, Souty C, Guillemot D, et al. Unraveling the seasonal epidemiology of pneumococcus. Proc Natl Acad Sci U S A. 2019 Jan 29;116(5):1802–7.

107. Howerton E, Williams TC, Casalegno JS, Dominguez S, Gunson R, Messacar K, et al. Using COVID-19 pandemic perturbation to model RSV-hMPV interactions and potential implications under RSV interventions. Nat Commun. 2025 Aug 6;16(1):7261.

108. Reich NG, Shrestha S, King AA, Rohani P, Lessler J, Kalayanarooj S, et al. Interactions between serotypes of dengue highlight epidemiological impact of cross-immunity. J R Soc Interface. 2013 Sep 6;10(86):20130414.

109. Runge J, Gerhardus A, Varando G, Eyring V, Camps-Valls G. Causal inference for time series. Nat Rev Earth Environ. 2023 Jun 27;4(7):487–505.

110. Cliff OM, Bryant AG, Lizier JT, Tsuchiya N, Fulcher BD. Unifying pairwise interactions in complex dynamics. Nat Comput Sci. 2023 Oct;3(10):883–93.

1017 Supporting Information

1018 **S1 Text. Supplementary methods and results.**