

RESEARCH ARTICLE

The limitations of non-mechanistic methods for characterizing pathogen-pathogen interactions: A simulation study

Sarah C. Kramer^{1*}, Sarah Pirikahu^{1,2}, Cana Kussmaul^{1,3}, Lulla Opatowski^{4,5}, Matthieu Domenech de Cellès¹

1 Max Planck Institute for Infection Biology, Infectious Disease Epidemiology Group, Berlin, Germany, **2** Department of Public Health, University of Otago, Wellington, New Zealand, **3** Institute of Public Health, Charité – Universitätsmedizin Berlin, 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

* kramer@mpiib-berlin.mpg.de



OPEN ACCESS

Citation: Kramer SC, Pirikahu S, Kussmaul C, Opatowski L, Domenech de Cellès M (2026) The limitations of non-mechanistic methods for characterizing pathogen-pathogen interactions: A simulation study. PLoS Comput Biol 22(8): e1013859. <https://doi.org/10.1371/journal.pcbi.1013859>

Editor: Burcu Tepekule, Princeton University Department of Ecology and Evolutionary Biology, UNITED STATES OF AMERICA

Received: December 21, 2025

Accepted: July 22, 2026

Published: August 17, 2026

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pcbi.1013859>

Copyright: © 2026 Kramer et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/).

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), simple, correlative methods for characterizing interactions may be prone to failure. Here, we perform a simulation study to evaluate several more complex 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). Overall performance patterns were similar when methods were applied to two real-world datasets from Hong Kong and Canada.

which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data availability statement: All code necessary to reproduce the findings reported here is available in our Git repository: https://github.com/sarahckramer/Pathogen_interaction_simulations (DOI: [10.5281/zenodo.17791160](https://doi.org/10.5281/zenodo.17791160)).

The synthetic data used for the majority of the analyses can be reproduced using the code linked above. Data on influenza and RSV from Hong Kong and Canada used for the analysis of real-world data can be found at <https://www.chp.gov.hk/en/statistics/data/10/641/642/2274.html> and https://search.open.canada.ca/opendata/?sort=metadata_modified+desc&-search_text=fluwatch&page=1, respectively. Climate data from the US National Centers for Environmental Information can be downloaded using the R package GSODR.

Funding: The author(s) received no specific funding for this work.

Competing interests: I have read the journal's policy and the authors of this manuscript have the following competing interests: MddC received consulting fees from MSD, GSK, Moderna, and Vaxcyte for work unrelated to this project. LO received a research grant from Sanofi for work on pathogen-pathogen interactions. All other authors declare no competing interests.

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 from 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. Throughout the literature, the term “interaction” is applied broadly and sometimes inconsistently: population-level phenomena (e.g., reductions in cases of subsequent pathogens due to people remaining at home during convalescence) and even purely correlative relationships are sometimes included. In this work, we use the term to refer exclusively to interactions arising due to biological, within-host mechanisms. Even within this more specific definition, interactions are highly diverse. The biological mechanisms at play can include both direct (e.g., competition for resources) and indirect (e.g., modulation of host cells or the immune response) processes [1, 2]. Likewise, depending on the specific pathogens and mechanisms involved, interactions may influence susceptibility to infection, risk of onward transmission upon

infection, the likelihood or severity of symptoms among those infected, or a combination of these outcomes [1]. Recent evidence suggests that a complex web of such 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, 3–12]. Many of these viruses may in turn affect infection with bacteria like *S. pneumoniae* and *Haemophilus influenzae* [13, 14], while the impact of immunosuppression induced by viruses like HIV and measles on susceptibility to a wide range of pathogens has long been recognized [15, 16]. Indeed, it seems likely that interactions are the norm, rather than the exception, in infectious disease epidemiology. Although interactions as defined here occur at the level of the individual, their effects on outbreak dynamics at the population level can be substantial [1, 2]. 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 [1]. To fully understand an interaction, each of these components must be described. 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, statistical 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 statistical associations identified using observed data are rarely informative about the presence or nature of an underlying, biological interaction: measures of coinfection prevalence systematically 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]. Despite these findings, these simple approaches remain common in the interactions literature. Although fitting mechanistic transmission models to empirical data is a promising alternative [1–3], 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], and 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.

We note preemptively that the above methods were not developed specifically with pathogen-pathogen interactions in mind, and that observed case data commonly available for epidemiologic research may not always meet every assumption of these methods (see Methods and Discussion for more details). With that said, some of these methods have proven surprisingly effective in scenarios where their assumptions don't necessarily hold [30]. Furthermore, it is unfortunately common in infectious disease epidemiology research to see methods applied in situations where their assumptions are violated. Indeed, Barrero Guevara et al. found that time series regression was commonly used in studies evaluating the influence of weather on infectious disease transmission, despite the fact that associations between weather variables and downstream disease incidence do not necessarily reflect the association between weather and

transmission rates [17, 38]. Similarly, despite the work discussed above showing that prevalence ratios and phase differences are not reliable as indicators of pathogen-pathogen interactions, these methods remain popular. Given that recent research has already attempted to apply some of the methods tested here to pathogen-pathogen interactions [39, 40], we consider a formal test of these methods to be of critical importance, even if they may not be perfectly applicable.

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. More specifically, we evaluate the extent to which a method returning a statistically significant result reliably corresponds to a true, biological interaction between the modeled pathogens (and, conversely, the extent to which null results are indicative of a lack of such an interaction). Furthermore, we assessed whether the magnitude of estimates produced by each method could provide information about the relative strength of the underlying interaction. Here we focused specifically on bidirectional, symmetric interactions, in which infection with one pathogen modulates the individual-level risk of infection with another pathogen upon exposure. 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. Overall, we aimed to determine whether the tested 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 [41]. All code used for this project has been published on GitHub [42].

Mathematical model

Transmission 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 pathogen-pathogen interactions can arise due to infection with one pathogen influencing susceptibility to another [4, 5, 9, 43]. 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., $\theta_\lambda = 2$ implies a doubling of the force of infection), values below 1 indicate negative interactions (e.g., $\theta_\lambda = 0.5$ represents a halving of the force of infection, while $\theta_\lambda = 0$ represents complete inhibition), and a value of 1 indicates no interaction. In other words, θ_{λ_1} acts as a multiplier on the rate at which modeled individuals move from state X_{IS} to state X_{IE} , and from X_{TS} to X_{TE} , while θ_{λ_2} modifies the rate of transition between X_{SI} and X_{EI} and between X_{ST} and X_{ET} (see Fig 1). Meanwhile, the time spent in compartment T ($7/\delta$, where δ is the rate at which individuals leave the T compartment to enter the R compartment) represents the duration of the interaction effect. Thus, δ_1 is the rate of transition between X_{TS} and X_{RS} , and δ_2 is the rate of transition between X_{ST} and X_{SR} ; these parameters control the amount of time for which an increase or decrease in the force of infection of one pathogen persists after recovery from the initial pathogen. A model schematic can be seen in Fig 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 [3].

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 [3, 4]. Where information was

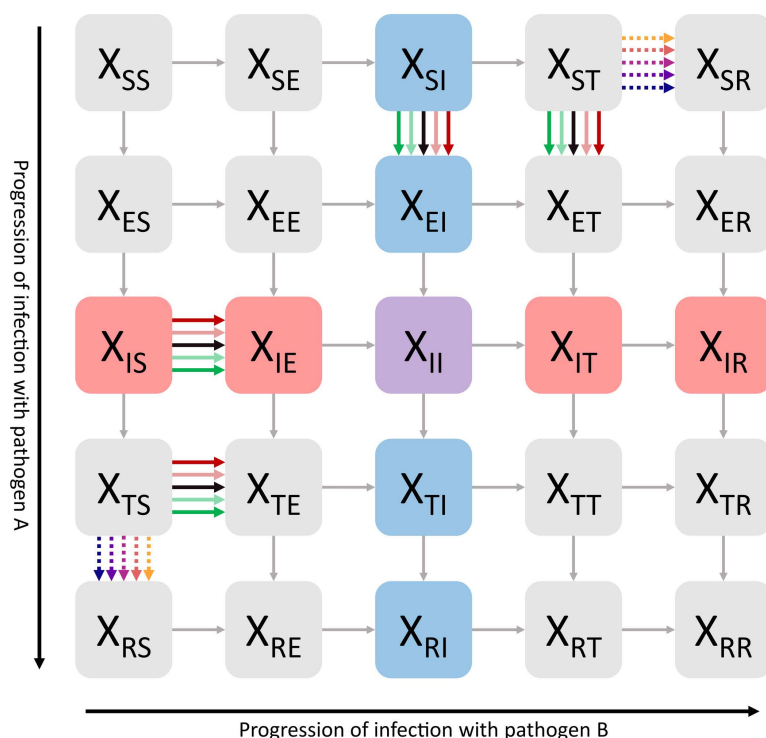


Fig 1. Schematic of the model used to generate synthetic data. Each box represents a distinct model state, with the first subscript indicating infection status with respect to pathogen A, and the second subscript indicating status with respect to pathogen B. Infection with pathogen A is indicated in 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 (θ_λ) are shown in a gradient of colors from green to red, while transitions impacted by changes in interaction duration ($7/\delta$) are shown as dotted arrows with a gradient of colors as in Figs 2, 4, and 5.

<https://doi.org/10.1371/journal.pcbi.1013859.g001>

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.

Influenza and RSV display similar seasonal patterns, with both viruses circulating primarily during the winter in temperate regions [59]. Indeed, some work suggests that influenza viruses and RSV respond similarly to several environmental drivers, including temperature, humidity, and rainfall [18, 60–65]. 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.

Table 1. Parameters and parameter values used to generate synthetic data. Parameters controlling the interaction effect are shown in bold; 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	[44, 45]
$1/\sigma$	Average latent period (days)	1	5	[46]
$1/\gamma$	Average infectious period (days)	5	10	[47–49]
$1/\omega$	Average duration of immunity (years)	1	0.6 – 1.0	[49–51]
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)	[3, 52, 53]
ϕ	Week of maximum seasonal forcing (relative to beginning of season)	26	26	NA
ρ	Reporting probability	0.15	0.05	[54, 55]
k	Dispersion parameter for the observation model	0.04 (0, 0.1)	0.02 (0, 0.05)	[56]
β_{sd}	Standard deviation of extra-demographic stochasticity	0.1 (0.01)	0.05 (0.005)	[57, 58]
θ_λ	Interaction strength (i.e., relative rate of transition between S and E compartments)	0, 0.25, 0.5, 1, 2, 4	Same as for A	NA
δ	Inverse of interaction duration (weeks) (i.e., rate of transition between T and R compartments)	1, 1/4, 1/13	Same as for A	NA

<https://doi.org/10.1371/journal.pcbi.1013859.t001>

Unlike RSV, which is comparatively antigenically stable, influenza viruses undergo rapid “antigenic drift” [66]. 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 [67]. 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 (see [S1 Text](#) for additional details). 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.

Observation model. The SEIRSxSEIRS model describes the true number of individuals in each model state at each timepoint. However, real-world surveillance systems cannot perfectly capture the true number of cases of a disease at any given time. Influenza and RSV in particular often cause mild or asymptomatic illness, such that many infected individuals do not seek healthcare and therefore are not included in surveillance data. For this reason, all model states in the transmission model ([Fig 1](#)) are latent, or unobserved. In order to produce synthetic data that are representative of real-world surveillance data, an observation model is needed to describe how observed data arise from the underlying latent process.

Here, we generate synthetic data by drawing observed cases from a negative binomial distribution each week. Specifically, the distribution at time t has mean $\mu = \rho_i H_i(t)$, where ρ_i is the probability that a case of pathogen i is reported, and H_i is the number of cases of pathogen i who have recovered between timepoints $t-1$ and t (i.e., we assume that cases report near the end of the infectious period), such that dH_A/dt is equal to $(X_{IS} + X_{IE} + X_{II} + X_{IT} + X_{IR}) \gamma_A$ and dH_B/dt is equal to $(X_{SI} + X_{EI} + X_{II} + X_{TI} + X_{RI}) \gamma_B$ (see also Eq 1 in [S1 Text](#)). The observation model has dispersion parameter k (i.e., the distribution has variance $\sigma^2 = \mu + k\mu^2$), such that higher k yields higher variability in reporting.

Generation of synthetic data

Synthetic data were generated using a wide range of interaction strengths ($\theta_\lambda = 0, 0.25, 0.5, 1, 2, 4$) and durations ($7/\delta = 1, 4, \text{ or } 13$ 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.

Here, we focus specifically on short-term interactions, as these are more realistic for unrelated pathogens where adaptive cross-immunity is unlikely, and our model is parameterized to produce outbreaks similar to those of influenza and RSV. While long-term interactions due to adaptive immune mechanisms can occur between related pathogens [10, 68, 69], understanding these interactions is a somewhat distinct problem, where knowledge of the underlying mechanisms (e.g., cross-immunity between paramyxoviruses [10], antibody-dependent enhancement between dengue serotypes [69]) may provide some prior information about the interaction's sign and duration. However, a test of CCM for long-term interactions can be found in [36].

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) [70, 71]. The observation model (see above) 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 Fig 2A; further examples can be found in S1 Fig.

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 ($\alpha = 0.05, 0.3, 0.4, 0.5$), and 3) the amount of process ($\beta_{sdA} = 0.01, \beta_{sdB} = 0.005$) 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) [72].

Real-world data

In addition to synthetic data, we applied all methods to observed influenza and RSV positivity data from Hong Kong [73] and Canada [74]. These data were chosen because we have previously fit a mechanistic transmission model to data from these two locations, and found evidence of a negative interaction in both datasets [3], allowing us to assess whether each method tested here yielded similar results.

The data are more comprehensively described and visualized in [3]. Briefly, data consist of the weekly number of tests positive for influenza and for RSV, as well as the total weekly number of tests conducted, over the course of four to six seasons. Samples were primarily obtained from inpatients and emergency department patients. From this data we calculated the weekly proportion of tests positive for influenza and for RSV, which we used as our case data when applying the statistical methods detailed in the following section. Note that, for Canada, we combined all (sub)types of influenza (H1N1, H3N2, and B); for Hong Kong, however, where H3N2 outbreaks often lead to multiple yearly peaks, we included only H1N1 and B, as in [3]. For methods that allowed it, we controlled for weekly mean absolute humidity, calculated from the US National Centers for Environmental Information's Global Surface Summary of the Day (GSOD) data [75], which we obtained using the R package "GSODR" [76], using the Clausius-Clapeyron relation [77]. Since the Canadian surveillance data cover the

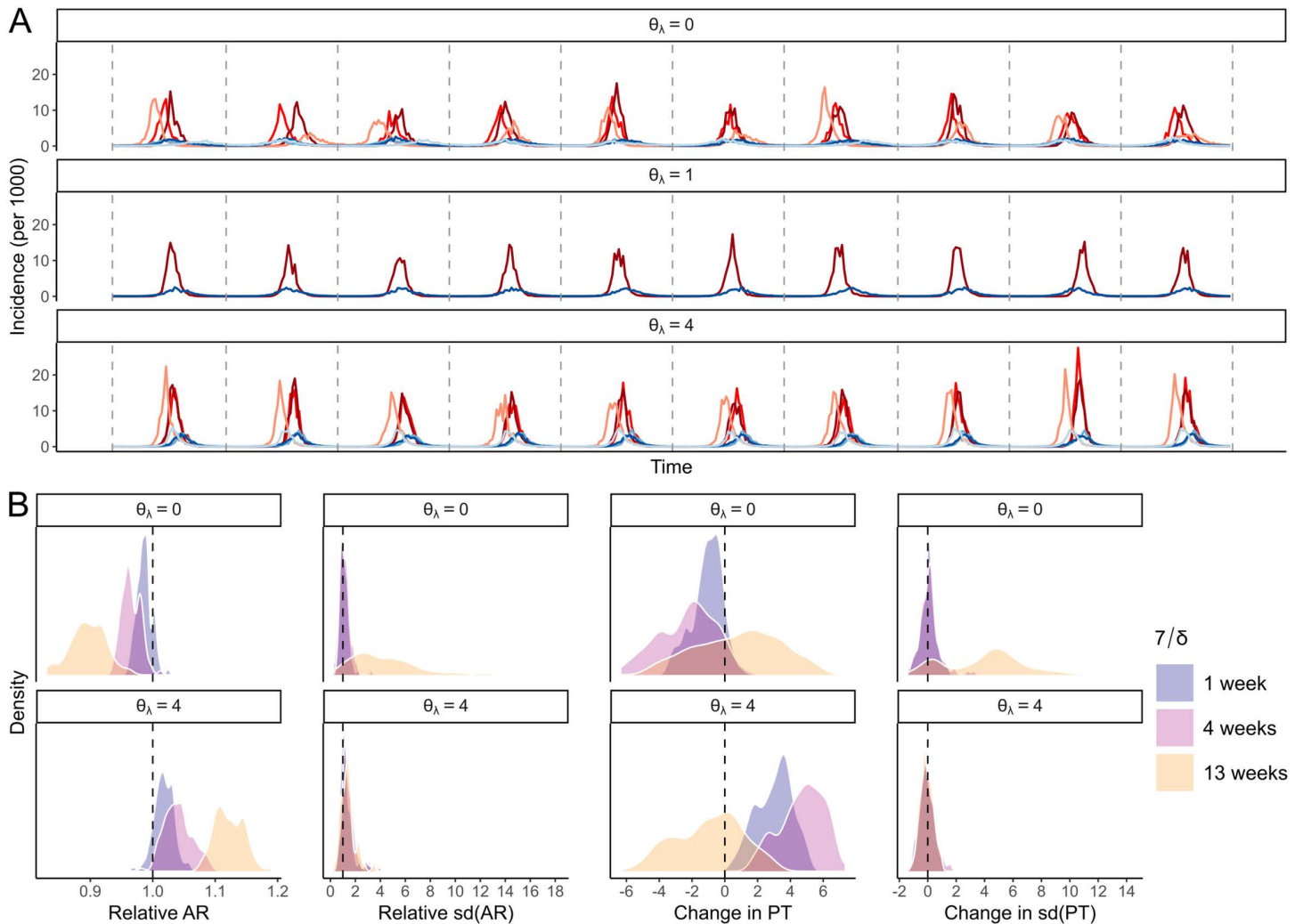


Fig 2. Impact of interactions on outbreak dynamics. (A) Example data consisting of ten years of simulated observations for pathogen A (red) and pathogen B (blue), generated with varying interaction strength (θ_λ) and duration ($7/\delta$) and all other parameters held constant at the values shown in Table 1 (here, $R_1 = 1.38$, $R_2 = 1.89$, $1/\omega_2 = 49.5$ 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.

<https://doi.org/10.1371/journal.pcbi.1013859.g002>

entire country, we used the median absolute humidity across all available stations each week. Absolute humidity data were chosen due to the likely influence of humidity on the transmission of both influenza and RSV [18, 61, 64].

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”), as well as to the surveillance data from Hong Kong and Canada. 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 [78] and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) [79] tests. Nonstationary datasets ($n = 2$ of 1800) were removed from consideration.

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:

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

where $\text{cov}(X, Y)$ 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 (as detailed in the Introduction), 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

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” [80]) Stan [81]
Granger Causality	Prediction	$G_{X \rightarrow Y}$	Yes	No	Yes (User must input data on confounder)	R (packages “stats,” “vars” [82, 83], “VARtests” [84])
Transfer Entropy	Prediction	$T_{X \rightarrow Y}^{(k, \ell)}(t)$	Yes	No	Yes (User must input data on confounder)	JIDT [85] R (package “rJava” [86])
CCM	State space reconstruction	ρ	Yes	No	Yes (Implicitly)	R (package “rEDM” [87])

<https://doi.org/10.1371/journal.pcbi.1013859.t002>

[25, 88]. By using a penalized likelihood approach, GAMs ensure that these smooths contain enough “wiggleness” 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., 1–53; WOY) as the only predictor. This smooth represents the seasonal periodicity in the data, which we assume here to be the same in all years. 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, 88]. 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:

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

$$\boldsymbol{\mu}_t = [\mu_{1,t}, \mu_{2,t}]^T$$

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

$$\mu_{i,t} = \alpha_i + s_i(\text{WOY}_t, k = 53) \quad (3)$$

where $\mathbf{X}_t = [X_{1,t}, X_{2,t}]^T$ is a matrix containing case counts at time t for both pathogens, which is modeled as a bivariate normal distribution with mean $\boldsymbol{\mu}_t$ and covariance matrix $\boldsymbol{\Sigma}$. The mean and standard deviation for each pathogen i at time t are indicated by $\mu_{i,t}$ 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) [80], which interfaces with the probabilistic programming language Stan [81]. 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, including formulations designed specifically for use with state-space models [89, 90]. To account for shared seasonality, we calculated the true extent of seasonal forcing over time according to Eq 1, and included this time series as a predictor in all models. We therefore test for Granger causality by comparing the following two models:

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

and

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

where p is the order of the autoregressive models, and S is the seasonal forcing component of Eq 1 (i.e., everything except $\bar{\beta}_i$, 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 [91, 92].

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). We also conducted a sensitivity analysis using a larger value (see S1 Text). The magnitude of the estimated effect of one pathogen on the other was then calculated as initially suggested by [93]:

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

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) [82, 83], and “VARtests” (version 2.0.5) [84] packages; a link to the full code used can be found under “Implementation” above. 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 Eq 6 is always positive; thus, Granger causality is incapable of distinguishing between positive and negative interactions.

Transfer entropy. Like Granger causality, transfer entropy [23, 94] 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 [95]. Specifically, the transfer entropy from X to Y , conditional on the shared seasonal forcing, is defined as:

$$T_{X \rightarrow Y|S}^{(k,l,m)}(t) = H(Y_t | Y_{t-1}^{(k)}, S_{t-1}^{(m)}) - H(Y_t | Y_{t-1}^{(k)}, X_{t-1}^{(l)}, S_{t-1}^{(m)}) \quad (7)$$

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, $H(Y_t | Y_{t-1}^{(k)}, S_{t-1}^{(m)})$ is the conditional Shannon entropy [96] 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 $H(Y_t | Y_{t-1}^{(k)}, X_{t-1}^{(l)}, S_{t-1}^{(m)})$ is similar, this time also conditioning on the past history of X up to timepoint $t - l$. More specifically, Shannon entropy measures uncertainty by accounting for both the range of possible values taken by Y_t , as well as the probability of each value, such that variables with a narrow possible range of values will have low entropy (i.e., low uncertainty), whereas variables with an equal probability of taking a wide range of values will have high entropy. Notably, Eq 7 is equivalent to the conditional mutual information between Y_t and $X_{t-1}^{(l)}$ given $Y_{t-1}^{(k)}$ and $S_{t-1}^{(m)}$, $I(Y_t : X_{t-1}^{(l)} | Y_{t-1}^{(k)}, S_{t-1}^{(m)})$, 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 [95]. Like with Granger causality, values above zero (i.e., where the uncertainty in Y at time t when accounting for the past histories of X , Y , S is lower than the uncertainty in Y at time t when accounting only for the past histories of Y and S alone) are taken to imply that X causes Y .

Overall, transfer entropy may be viewed as a nonparametric alternative to Granger causality [94, 95, 97], and for Gaussian variables, the two are equivalent [98]. Because VARs may not perform well when dynamics are highly nonlinear [99], and data transformations can greatly affect the resulting Granger causality estimate [100], 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 [95].

We calculated the transfer entropy for all datasets using the Kraskov-Stögbauer-Grassberger (KSG) technique with four nearest neighbors [101], as implemented by the Java Information Dynamics Toolkit (JIDT) [85]. 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). As with Granger causality, we also performed a sensitivity analysis where larger values were used (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 [102]. 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 [85].

Analyses were conducted using JIDT (version 1.6.1) [85] and the R package “rJava” (version 1.0.11) [86]. 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 by simply 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].

Convergent cross-mapping (CCM) is a relatively new approach that takes advantage of this inherent lack of independence between time series [24]. Specifically, to determine whether X causes Y , CCM uses L lagged values of Y to construct a shadow manifold, M_Y ; this is done according to Takens’ theorem [103], which states that information about all states of a dynamical system can be reconstructed based on the lagged values of a single state of that system. Then, the method assesses whether the values of X can be estimated based on the topology of M_Y . In other words, unlike Granger

causality and transfer entropy, which infer that X causes Y if past values of X improve predictions of some current value of Y , CCM infers that X causes Y if current values of Y can be used to “predict” current or past values of X . Crucially, for causation to be demonstrated, the quality of these predictions must increase with increasing L , or the amount of data available for building the shadow manifold (i.e., there must be convergence).

To perform CCM, we first chose the ideal embedding dimensions (i.e., the number of lags used to construct the shadow manifolds) for each synthetic dataset by determining which values yielded the most accurate within-pathogen predictions for each pathogen. We allowed a maximum embedding dimension of 2 for pathogen A and 5 for pathogen B (see [S1 Text](#)). We also allowed for a lagged effect of one pathogen on the other by selecting the negative lag that yielded the highest cross-map skill [104]; a maximum lag of 20 weeks was permitted (see [S1 Text](#)). As for the previous two methods, we also ran a sensitivity analysis using larger values for the maximum embedding dimensions and lags ([S1 Text](#)). The selected values for the embedding dimensions and lags were then used to run CCM with 100 samples for each value of L . Point estimates were taken to be the median cross-map skill across all samples for the largest L . We note that, since CCM is expected to yield null results for two variables that are not causally linked but have a common environmental driver [24], there is no need to explicitly account for the shared seasonal forcing between pathogens when applying this method.

We tested two methods of assessing significance. The first approach considers a result to be significant so long as there is convergence, as in CCM’s original implementation [24]. Specifically, we use a nonparametric bootstrap approach to determine whether the cross-map skill for the maximum value of L is significantly greater than the cross-map skill for the minimum value of L , as in [36]. The second approach, originally suggested by [18], compares the cross-map skill obtained from the data for the maximum value of L to a null distribution of cross-map skill calculated using seasonal surrogates of the hypothesized causal pathogen’s data. Specifically, we generate 500 surrogate time series by maintaining the seasonal pattern in the data, but reshuffling the residuals. A p-value is then calculated as:

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

where ρ_{true} is the cross-map skill calculated from the data, $\rho_{surrogate, i}$ is the cross-map skill calculated using surrogate time series i , and n is the number of seasonal surrogates considered [30, 105]. 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) [87]. 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, 106]).

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), for which the reciprocal is undefined, were assigned a value of 16; results remained qualitatively similar when different values were chosen. 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 (Figs 2A and 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 (Fig 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 (S2 Fig). 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 Fig 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 methods, which can only yield positive point estimates, results were considered accurate for negative interactions as long as point estimates were significantly greater than zero. For Granger causality and transfer entropy, only the results of tests accounting for seasonality are shown (see S3 Fig for results ignoring shared seasonality). Of the tested methods, GAMs performed best, with 85.2% sensitivity and 72.5% specificity. All other methods showed much worse performance: while sensitivity was often high, specificity was consistently below 50%. In other words, among simulations where no interaction was modeled, all methods other than GAMs yielded false positives more often than not. Results were no better when larger values were chosen for various method parameters (e.g., lags, history lengths, and embedding dimensions), and for CCM in particular, using larger embedding dimensions and lags often led to even lower specificity (S4 Fig). For most methods capable of distinguishing interaction direction (Granger causality, transfer entropy, and CCM using seasonal surrogates), false positives were more common when trying to classify the interaction effect of pathogen B on pathogen A. CCM using convergence as the criterion for significance was the only exception, but improved specificity came at the cost of sensitivity. Overall performance quality was similar for these three methods; as expected [20, 107], Pearson correlation coefficients were by far the least accurate.

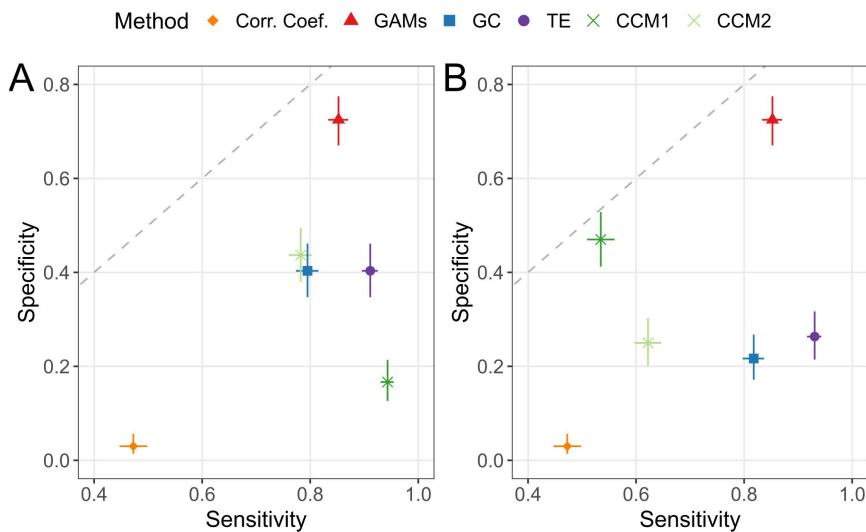


Fig 3. Sensitivity and specificity of all tested methods. Results for the interaction effect of pathogen A on pathogen B are shown in (A), while results for the effect of pathogen B on A are shown in (B). Method is indicated by both point shape and color; CCM1 refers to the method assessing significance based on convergence, while CCM2 refers to the method assessing significance using seasonal surrogates. The crosshairs on each point indicate 95% confidence intervals, obtained using a binomial test. The dashed diagonal line shows where sensitivity is equal to 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.

<https://doi.org/10.1371/journal.pcbi.1013859.g003>

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 Fig 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 (Fig 4H and 4J), and correlation coefficients performed particularly poorly for negative interactions (Fig 4A). Interaction strength had little effect on accuracy for all of the tested methods.

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. Fig 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 (S5 Fig). 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 any given real-world dataset. Encouragingly, the sign of the true interaction and of

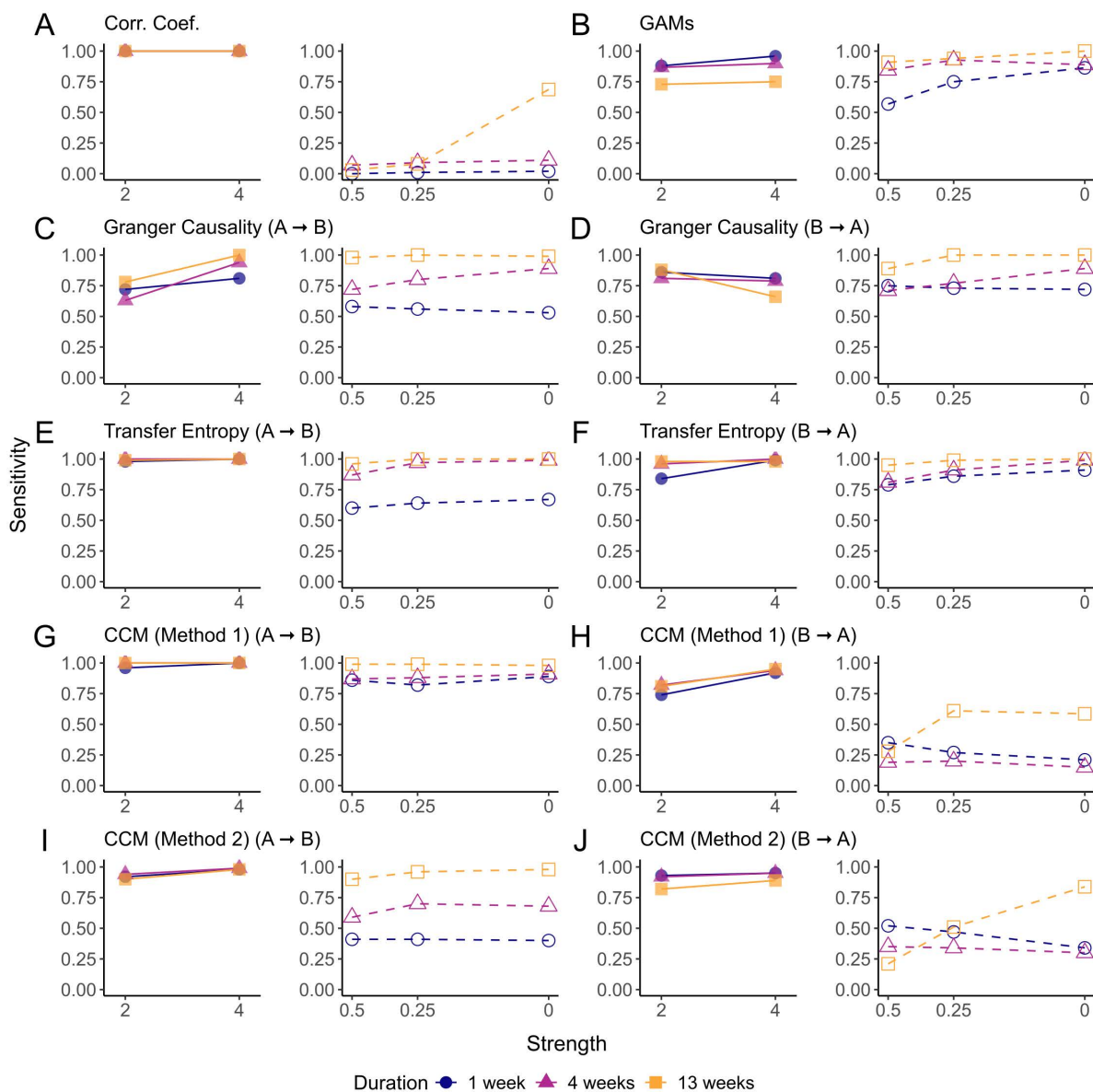


Fig 4. Sensitivity of all methods by true interaction strength and duration. Results are shown for correlation coefficients (A), GAMs (B), Granger causality (C and D), transfer entropy (E and F), CCM using convergence to assess significance (G and H), and CCM using seasonal surrogates to assess significance (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.

<https://doi.org/10.1371/journal.pcbi.1013859.g004>

the estimate returned by the GAM approach overwhelmingly agree (Figs 4 and S5). 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.

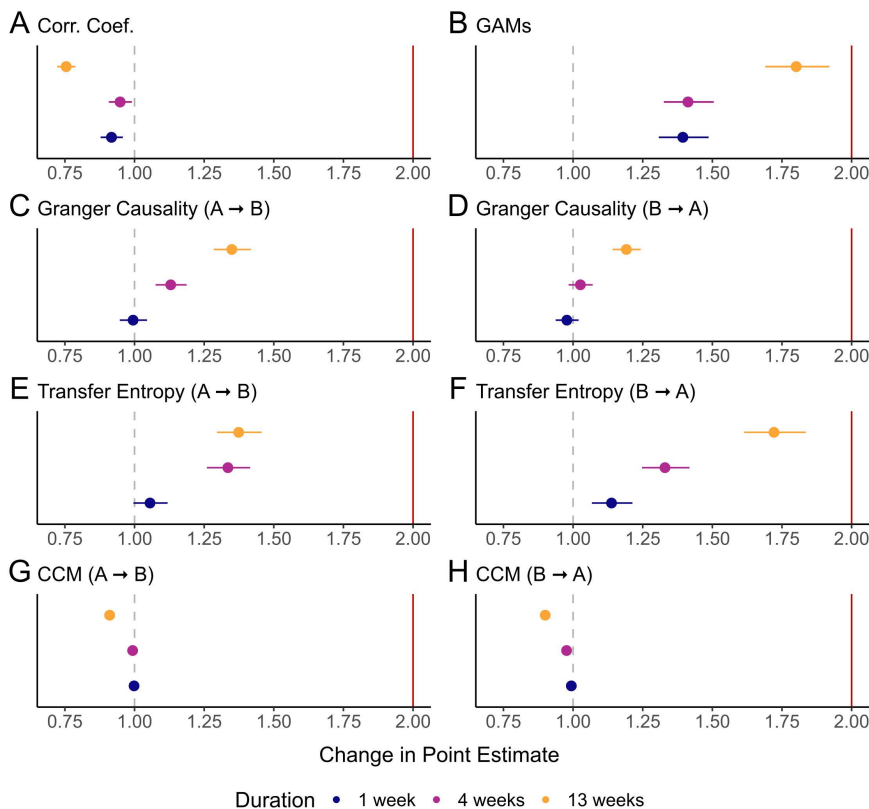


Fig 5. Multiplicative change in point estimates returned in response to a doubling of true interaction strength. Results are shown for all methods ((A) correlation coefficients, (B) GAMs, (C) Granger causality (impact of A on B), (D) Granger causality (impact of B on A), (E) transfer entropy (A on B) (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.

<https://doi.org/10.1371/journal.pcbi.1013859.g005>

Methods displayed inconsistent results concerning the presence and strength of the interaction between influenza and RSV when applied to real-world data

In a previous study, we fit a mechanistic model of influenza and RSV cocirculation to surveillance data from Hong Kong and Canada, and found evidence of a negative interaction in both locations. When confronted with these same real-world data, we find that all methods tested in the current work were able to detect evidence of an interaction effect of influenza on RSV in Hong Kong. For the effect of RSV on influenza in Hong Kong, and for both interaction directions in Canada, on the other hand, conclusions were not always consistent between methods (Fig 6). As in our simulation study, GAMs performed well, identifying a significant negative interaction in both Hong Kong and Canada (Fig 6B). Both Granger causality and transfer entropy were able to identify the effect of influenza on RSV in both locations, but typically failed to identify an impact of RSV on influenza (Fig 6C–6D). However, methods disagreed regarding the relative strength of these interactions across locations: Granger causality suggests a stronger interaction effect in Canada, transfer entropy suggests the opposite, and GAM results are consistent with a similar interaction strength in both Hong Kong and Canada. Finally, while CCM typically identified an interaction effect in Hong Kong, both implementations struggled to find evidence of an interaction in Canada, at least for the effect of influenza on RSV (Fig 6E).

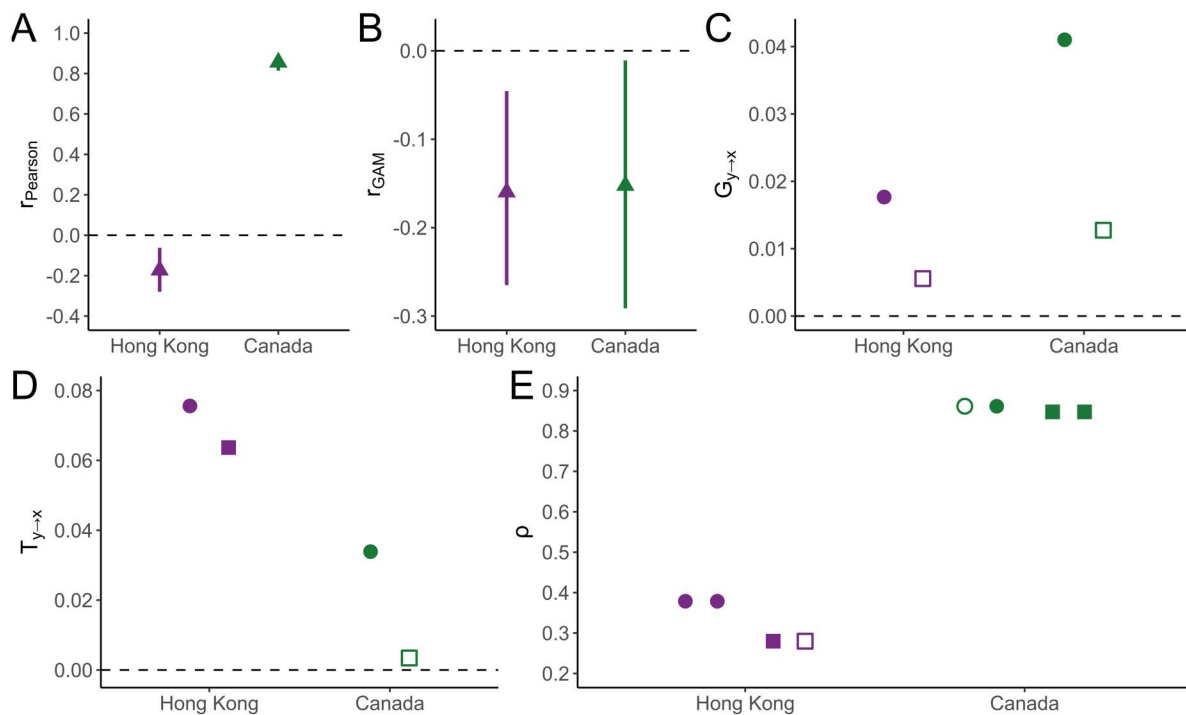


Fig 6. Point estimates returned by each method when confronted with influenza and RSV data from Hong Kong and Canada. Results for Pearson correlation coefficients are shown in (A), GAMs in (B), Granger causality in (C), transfer entropy in (D), and CCM in (E). Filled points represent statistically significant results ($p < 0.05$; for GAMs only, 95% confidence intervals not including 0), while open points represent results that were not significant. Circles show results for the interaction effect of influenza on RSV, and squares show the effect of RSV on influenza; for methods incapable of distinguishing direction (correlation coefficients and GAMs), triangles are shown. Results from Hong Kong are shown in purple, while results from Canada are shown in green. Horizontal dotted lines are plotted at values of 0. For correlation coefficients and GAMs, 95% confidence intervals are shown as vertical lines; other methods do not return confidence intervals. For CCM, results for the implementation using convergence to indicate significance are shown on the left, and results using seasonal surrogates are shown on the right; since these approaches only differ in how they evaluate significance, point estimates do not change depending on the method.

<https://doi.org/10.1371/journal.pcbi.1013859.g006>

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]. Furthermore, although we allowed the GAMs to flexibly control for the seasonality in observed case data, seasonality in observations does not necessarily correlate with seasonality in the underlying force of infection [18]. 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 (S6 Fig). 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 (S7 Fig).

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. While 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 (S8 Fig), 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 (S8 Fig).

We observed similar results when methods were applied to real-world data. Here again, GAMs identified a significant, negative interaction between influenza and RSV in both locations, while other methods were less able to consistently detect relationships between the two viruses. Methods were generally more likely to detect an effect of influenza on RSV than of RSV on influenza. Interestingly, CCM, which is arguably the method best rooted in causal theory out of those we tested, struggled the most: neither implementation was able to identify the interaction effect of influenza on RSV in Canada, while both Granger causality and transfer entropy could.

We emphasize that our results do not imply that these methods are inherently ineffective, and our work is not intended to devalue them. Indeed, these methods are successfully applied in several other contexts [24, 29, 30, 35, 108]. Rather, we find that these methods perform poorly in the specific context of pathogen-pathogen interactions, and it is important to consider why this is the case. 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 (S3 Fig), 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 (S7 Fig), 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 [109, 110]. 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 [60–62, 64]. Temperature and humidity are also likely to play a role in the transmission of SARS-CoV-2

[111], as well as seasonal coronaviruses [112] and rhinoviruses [113]. Because this is a simulation study, we were able to control for the exact strength of seasonal forcing over time. However, in real-world outbreaks, the exact relationship between potential drivers (e.g., temperature) and pathogen transmission rates is not known, and, as explained above, seasonal patterns in forcing cannot necessarily be inferred from seasonality in pathogen incidence [17, 18]. We had reasonable success applying Granger causality and transfer entropy to real-world influenza and RSV data while controlling for absolute humidity, at least when it came to detecting the effect of influenza and RSV. However, RSV and especially influenza are two pathogens for which the underlying climatic drivers are relatively well-studied, which is not the case for all pathogens.

The role of confounding in general should also be discussed. Both Granger causality and transfer entropy, as well as regression-based methods like GAMs, assume that all relevant confounders are observed and accounted for [17, 91, 95]. We are able to control for climate forcing, as weather data are widely collected and often freely available. However, we are unable to control for the proportion of the population susceptible to infection with each pathogen, a quantity that has a critical impact on transmission dynamics, and which is typically unobserved in reality [17, 38]. Furthermore, while we can be sure that weather is the only extrinsic variable driving transmission in our simulation study, this is not the case in real life. For example, given the apparent prevalence of pathogen-pathogen interactions, failure to account for other pathogens that interact with both pathogens of interest could lead to invalid results. Unlike Granger causality and transfer entropy, where confounding variables must be explicitly defined, GAMs control for patterns in the data resulting from confounding, while remaining mostly naïve to the nature of these confounders. While it could be argued that this gives GAMs an advantage, and may partially explain their success here, we emphasize again that there is no guarantee that patterns in observed data will be correlated with patterns in the underlying forcing functions [18]. Theoretically, approaches that control for confounding directly are expected to yield more accurate results.

The impact of observation noise on method performance should also be considered, as the utility of these methods may be limited by data that are not perfectly observed [24, 91]. Unfortunately, for many infectious diseases, real-world outbreak data are notoriously error-laden: surveillance systems relying on healthcare utilization and testing will miss mild and asymptomatic cases, and sampling effort is often much lower outside of the outbreak season [3]. Interestingly, our sensitivity analyses suggest that reducing observation noise to zero does not have a purely beneficial impact on method performance. Rather, lower observation noise tended to reduce the number of false negatives, but increased false positives, while increasing observation noise had the opposite effect (S8 Fig). Furthermore, this pattern was not consistent across all methods, suggesting that the exact impact of noisy observations is difficult to predict. In contrast, modifying the amount of process noise had little effect on performance (S8 Fig), a result that is mostly in line with Cobey and Baskerville [36].

In addition to observation noise, a more fundamental issue with the use of observed incidence data for these analyses is the potential loss of information when values corresponding to the latent state variables of the transmission model are processed through the observation model. For the interaction effect of pathogen A on pathogen B, for example, all individuals in compartments X_{IS} and X_{TS} at a given time may be impacted by the interaction. However, 1) H_A only includes cases at the time of reporting, and not for the full duration of their infection; 2) the total prevalence of pathogen A ($X_{IS} + X_{IE} + X_{II} + X_{IT} + X_{IR}$) includes four compartments beyond X_{IS} , where the interaction has no effect; 3) X_{TS} is not included in prevalence, as these individuals are no longer infected. With that said, we do not believe this to be a convincing explanation for the underperformance of the tested methods. First, we find that modeled incidence (H_A) is highly correlated with prevalence ($X_{IS} + X_{IE} + X_{II} + X_{IT} + X_{IR}$) (S9A Fig), suggesting a lack of information degradation. Furthermore, although X_{IS} itself makes up only about 40% of cases of pathogen A at a given time, on average (S9B Fig), all compartments making up the total prevalence of pathogen A contribute to its force of infection, which will have downstream effects on X_{IS} and therefore on the extent to which the interaction impacts overall dynamics. Finally, we find that values of X_{TS} are highest compared to H_A for the longest-duration interactions (S9C Fig), indicating that, in these scenarios, there may be the greatest mismatch between observed incidence and the number of individuals directly contributing to the interaction effect. However, these

are also the scenarios for which the tested methods perform best, suggesting that this is not inherently a problem. This makes sense, given that individuals in X_{TS} were included in H_A at previous timepoints, and many of the methods used here account for lags in the relationship between time series.

It is also worth discussing that, although all of the interactions we modeled were symmetric, most methods were more likely to correctly classify the effect of pathogen A on pathogen B than the effect of B on A. This result held for both simulated and real-world data. This could be, in part, due to the seasonal driver. CCM is known to be unreliable when a downstream variable becomes synchronized to a forcing variable [24]; in cases involving a shared driver, this is because this leads to data (here, observed cases data) that closely resemble the driver, potentially leading to false positives [36]. In our simulated data, the strength of seasonal forcing is much more highly correlated with observed cases of pathogen B than of pathogen A, perhaps explaining the discrepancy in our results. This was not the case in our real-world data, where RSV outbreaks were only slightly but not significantly more strongly correlated with absolute humidity data. However, we note that influenza outbreaks peaked before RSV outbreaks in almost every season contained in our observed data, which could have made finding an effect of RSV on influenza more difficult. Again, we emphasize that this issue is not unique to the methods tested here, and that we experienced similar difficulties when fitting a mechanistic model to these data [3]. Requiring the optimal cross-map lag to be negative as an additional significance criterion can mitigate this issue [36, 104], but, for systems like seasonal infectious disease outbreaks, which have high periodicity, this approach is not valid [114]. Because most pathogen-pathogen interactions are not well-understood, it is often unclear whether the interaction is uni- or bidirectional. For this reason, failure to accurately assess directionality represents a significant drawback of these methods. Even GAMs, which performed very well otherwise, cannot help here, as they offer no information about directionality.

Ultimately, the inherent complexity of pathogen-pathogen interactions may explain why many of the tested methods perform poorly here. Methods like Granger causality, transfer entropy, and CCM were originally developed with direct interactions between individual state variables in mind. In our transmission model, however, the interaction effect depends simultaneously on several model states: the total number of new infections with one pathogen is modulated by the total number of current and recent infections with the other pathogen. The issue is potentially amplified by applying the methods to outputs from the observation model, which aggregates over several states and returns underreported and error-laden case counts (although, as noted above, neither observation error nor the use of incidence data specifically appear to have any substantial impact on method performance). On a similar note, these observed data are related to the interaction of interest through a long and indirect causal chain (see Vignette 1 in [17] for a related discussion): individual-level processes impact how likely one is to become infected with or to transmit a second pathogen (modeled here as a change in the force of infection), which in turn will impact the total number of people in a population who become infected, which is itself observed with noise. For these reasons, it is perhaps unsurprising that many of the tested methods struggle to infer pathogen-pathogen interactions. Indeed, methods like Granger causality and CCM perform well for simpler, more direct processes like, for example, predator-prey interactions [24, 30]. That said, these methods are already being applied to infer interactions from real-world data [39, 40]. We therefore reemphasize our point from the Introduction that studies explicitly evaluating these methods in this context remain an urgent necessity.

Given these challenges, none of the tested methods appear to be promising tools for characterizing pathogen-pathogen interactions on the basis of infectious disease outbreak data. Approaches utilizing mechanistic models of infectious disease transmission, like the one used to generate the synthetic data for this study, may represent an attractive alternative. By using equations to describe the transmission of pathogens throughout a population, these models explicitly account for nonlinear dynamics, as well as for potential confounders, including unobserved state variables like the proportion susceptible. Models can then be fit to surveillance data to gain insight about mechanisms of interest, including interaction effects. Although this can be computationally intensive, and improvements in the efficiency of methods for model fitting are necessary to handle, for example, models of greater than two cocirculating pathogens [10], mechanistic modeling methods show clear promise. In fact, this approach has already been used to better understand interactions between

influenza and RSV [3, 115], influenza and *S. pneumoniae* [13, 116], influenza and SARS-CoV-2 [8], RSV and both hMPV and human parainfluenza virus (HPIV) [10, 117], and different serotypes of dengue [118] and cholera [71]. A key advantage of these methods over the ones tested in this work is that they can provide quantitative estimates and confidence intervals for specific characteristics of the interaction, such as its strength and duration.

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 [1, 2]. 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 [13]; 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 [3]. 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 (S10 Fig). 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 [119, 120]. 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 non-mechanistic 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.

Supporting information

S1 Text. Supplementary methods and results.

(PDF)

S1 Fig. Four representative data sets generated from the model. Figure layout and colors are as in Fig 2. In each panel, interaction strength (θ_λ) and duration ($7/\delta$) were varied, while all other parameters were held constant for that data-set. Specifically, parameter values were: (A) $Ri_1 = 1.36$, $Ri_2 = 1.88$, $1/\omega_2 = 32.5$ weeks; (B) $Ri_1 = 1.07$, $Ri_2 = 1.95$, $1/\omega_2 = 36.8$ weeks; (C) $Ri_1 = 1.11$, $Ri_2 = 1.80$, $1/\omega_2 = 44.7$ weeks; (D) $Ri_1 = 1.38$, $Ri_2 = 1.62$, $1/\omega_2 = 45.4$ weeks; all other parameters were set to the values in Table 1.

(SVG)

S2 Fig. Change in dynamics of pathogen A due to interaction. Similar to Fig 2B, the distribution of the change in attack rate, year-to-year variation in attack rate, peak timing, and year-to-year variation in peak timing of pathogen A 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.

(SVG)

S3 Fig. Sensitivity and specificity for implementations of Granger causality and transfer entropy accounting for vs. ignoring shared seasonal forcing. Method is indicated by both point shape and color, as in Fig 3; darker colors show the results from the main analysis, where shared seasonal forcing is accounted for, while the faded points show results from an analysis where shared seasonal forcing is not included. The crosshairs superimposed on each point represent 95% confidence intervals. As in Fig 3, the dashed diagonal lines show where sensitivity is equal to specificity. Results for the effect of pathogen A on pathogen B are shown on the left, and results for the effect of B on A are shown on the right.

(SVG)

S4 Fig. Change in sensitivity and specificity, relative to main results (Fig 3), when higher values are used as maximum orders for VAR models (Granger causality), history lengths (transfer entropy), and maximum embedding dimensions and lags (CCM). Method is indicated by both point shape and color, as in Fig 3. The crosshairs on each point represent 95% confidence intervals. The dashed vertical and horizontal lines indicate no change in sensitivity and specificity, respectively. Results for the interaction effect of pathogen A on pathogen B are shown on the left, while results for the effect of pathogen B on A are shown on the right.

(SVG)

S5 Fig. Point estimates returned by each method, according to true interaction strength and duration. Results are shown for all methods and directions (correlation coefficients (A), GAMs (B), Granger causality (C-D), transfer entropy (E-F), and CCM (G-H); results for effect of pathogen A on pathogen B shown in (C), (E), and (G), and results for effect of B on A in (D), (F), and (H)). True interaction strength is shown on the x-axis, while true direction is indicated by color. The black points and lines show results for simulations where no interaction was included; the gray dotted lines at zero indicate a null result (i.e., no interaction identified).

(SVG)

S6 Fig. Accuracy of GAMs for symmetric vs. asymmetric (unidirectional) interactions. Results for the asymmetric interaction, where pathogen B has no effect on susceptibility to pathogen A, are shown as dotted lines; the main text results, for a symmetric interaction, are shown as solid lines. True interaction strength is shown on the x-axis, while true interaction duration is shown by panel; colors and shapes are as in main text Fig 4.

(SVG)

S7 Fig. Sensitivity and specificity for all tested methods as the amplitude of shared seasonal forcing is varied. For all methods capable of distinguishing direction, the tested direction is indicated for each panel. Point color indicates the amplitude of seasonal forcing; shapes are as in Fig 3. The dotted diagonal line shows where sensitivity is equal to specificity. Dotted vertical and horizontal lines show the sensitivity and specificity, respectively, obtained in the main analysis ($\alpha=0.20$, also shown here by the blue point).

(SVG)

S8 Fig. Sensitivity and specificity for all tested methods for a range of sensitivity analyses varying process noise, observation noise, and the number of years of available data. For all methods capable of distinguishing direction, the tested direction is indicated for each panel. The specific sensitivity analysis is indicated by color, shapes are as in Fig 3, and crosshairs indicate 95% confidence intervals. The dotted diagonal line shows where sensitivity is equal to specificity. Dotted vertical and horizontal lines show the sensitivity and specificity, respectively, obtained in the main analysis.

(SVG)

S9 Fig. Relationship between latent state variables and incidence of pathogen A as returned by the observation model, across different values of interaction strength and duration. (A) The association between prevalence

($X_{IS} + X_{IE} + X_{II} + X_{IT} + X_{IR}$) and incidence (H_A) of pathogen A. Points represent individual timepoints, and the lines connecting them show the evolution of the system over time. The dotted gray lines indicate where prevalence and incidence are equal. (B) The proportion of prevalent cases of pathogen A over time belonging to each of the five model states containing individuals infected with pathogen A. (C) A comparison of the number of total prevalent cases of pathogen A (black) and the number of individuals recently infected with pathogen A and susceptible to pathogen B (green) over time. All subplots show results for a single synthetic dataset (see “Generation of Synthetic Data” in the main text), with the same dataset chosen for panels (A), (B), and (C).

(SVG)

S10 Fig. Impact of the difference in peak timing between pathogens A and B on method accuracy. Lines show the marginal effect of peak timing difference on the chance of correctly identifying the impact of pathogen A on pathogen B, for all methods capable of distinguishing directionality, after controlling for parameter set. Shaded areas represent 95% confidence intervals. Colors for each method are as in Fig 3. Positive values of peak timing difference indicate simulations where pathogen B peaked earlier in simulations with no modeled interaction, while negative values indicate simulations where pathogen A peaked first; the vertical dotted line at 0 indicates simulations where both pathogens peaked simultaneously. Density plots along the bottom of each panel show the relative frequency with which different values of peak timing difference were observed in our simulated dataset.

(SVG)

Acknowledgments

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

Author contributions

Conceptualization: Sarah C. Kramer, Sarah Pirikahu, Matthieu Domenech de Cellès.

Formal analysis: Sarah C. Kramer, Sarah Pirikahu.

Methodology: Sarah C. Kramer, Sarah Pirikahu.

Software: Sarah C. Kramer, Sarah Pirikahu, Cana Kussmaul.

Supervision: Matthieu Domenech de Cellès.

Validation: Cana Kussmaul.

Visualization: Sarah C. Kramer.

Writing – original draft: Sarah C. Kramer.

Writing – review & editing: Sarah C. Kramer, Sarah Pirikahu, Cana Kussmaul, Lulla Opatowski, Matthieu Domenech de Cellès.

References

1. Wong A, Barrero Guevara LA, Goult E, Briga M, Kramer SC, Kovacevic A, et al. The interactions of SARS-CoV-2 with cocirculating pathogens: Epidemiological implications and current knowledge gaps. *PLoS Pathog.* 2023;19(3):e1011167. <https://doi.org/10.1371/journal.ppat.1011167> PMID: [36888684](https://pubmed.ncbi.nlm.nih.gov/36888684/)
2. Opatowski L, Baguelin M, Eggo RM. Influenza interaction with cocirculating pathogens and its impact on surveillance, pathogenesis, and epidemic profile: a key role for mathematical modelling. *PLoS Pathog.* 2018;14(2):e1006770. <https://doi.org/10.1371/journal.ppat.1006770> PMID: [29447284](https://pubmed.ncbi.nlm.nih.gov/29447284/)
3. Kramer SC, Pirikahu S, Casalegno J-S, Domenech de Cellès M. Characterizing the interactions between influenza and respiratory syncytial viruses and their implications for epidemic control. *Nat Commun.* 2024;15(1):10066. <https://doi.org/10.1038/s41467-024-53872-4> PMID: [39567519](https://pubmed.ncbi.nlm.nih.gov/39567519/)

4. Chan KF, Carolan LA, Korenkov D, Druce J, McCaw J, Reading PC, et al. Investigating viral interference between influenza A virus and human respiratory syncytial virus in a ferret model of infection. *J Infect Dis*. 2018;218(3):406–17. <https://doi.org/10.1093/infdis/jiy184> PMID: 29746640
5. Wu A, Mihaylova VT, Landry ML, Foxman EF. Interference between rhinovirus and influenza A virus: a clinical data analysis and experimental infection study. *Lancet Microbe*. 2020;1(6):e254–62. [https://doi.org/10.1016/s2666-5247\(20\)30114-2](https://doi.org/10.1016/s2666-5247(20)30114-2) PMID: 33103132
6. Gonzalez AJ, Ijezie EC, Balemba OB, Miura TA. Attenuation of influenza A virus disease severity by viral coinfection in a mouse model. *J Virol*. 2018;92(23).
7. Essaidi-Laziosi M, Geiser J, Huang S, Constant S, Kaiser L, Tapparel C. Interferon-dependent and respiratory virus-specific interference in dual infections of airway epithelia. *Sci Rep*. 2020;10(1):10246. <https://doi.org/10.1038/s41598-020-67262-5>
8. Domenech de Cellès M, Casalegno JS, Lina B, Opatowski L. Estimating the impact of influenza on the epidemiological dynamics of SARS-CoV-2. *PeerJ*. 2021;9:e12566. <https://doi.org/10.7717/peerj.12566>
9. Dee K, Goldfarb DM, Haney J, Amat JAR, Herder V, Stewart M, et al. Human rhinovirus infection blocks severe acute respiratory syndrome coronavirus 2 replication within the respiratory epithelium: implications for COVID-19 epidemiology. *J Infect Dis*. 2021;224(1):31–8. <https://doi.org/10.1093/infdis/jiab147> PMID: 33754149
10. Bhattacharyya S, Gesteland PH, Korgenski K, Bjørnstad ON, Adler FR. Cross-immunity between strains explains the dynamical pattern of paramyxoviruses. *Proc Natl Acad Sci U S A*. 2015;112(43):13396–400. <https://doi.org/10.1073/pnas.1516698112> PMID: 26460003
11. Geiser J, Boivin G, Huang S, Constant S, Kaiser L, Tapparel C, et al. RSV and HMPV infections in 3D tissue cultures: mechanisms involved in virus-host and virus-virus interactions. *Viruses*. 2021;13(1):139. <https://doi.org/10.3390/v13010139> PMID: 33478119
12. Wen X, Suryadevara N, Kose N, Liu J, Zhan X, Handal LS, et al. Potent cross-neutralization of respiratory syncytial virus and human metapneumovirus through a structurally conserved antibody recognition mode. *Cell Host Microbe*. 2023;31(8):1288–1300.e6. <https://doi.org/10.1016/j.chom.2023.07.002> PMID: 37516111
13. Shrestha S, Foxman B, Berus J, van Panhuis WG, Steiner C, Viboud C. The role of influenza in the epidemiology of pneumonia. *Sci Rep*. 2015;5(1):15314. <https://doi.org/10.1038/srep15314>
14. Avadhanula V, Rodriguez CA, Devincenzo JP, Wang Y, Webby RJ, Ulett GC, et al. Respiratory viruses augment the adhesion of bacterial pathogens to respiratory epithelium in a viral species- and cell type-dependent manner. *J Virol*. 2006;80(4):1629–36. <https://doi.org/10.1128/JVI.80.4.1629-1636.2006> PMID: 16439519
15. Mina MJ, Metcalf CJE, de Swart RL, Osterhaus ADME, Grenfell BT. Long-term measles-induced immunomodulation increases overall childhood infectious disease mortality. *Science*. 2015;348(6235):694–9. <https://doi.org/10.1126/science.aaa3662> PMID: 25954009
16. Pawlowski A, Jansson M, Sköld M, Rottenberg ME, Källénus G. Tuberculosis and HIV co-infection. *PLoS Pathog*. 2012;8(2):e1002464. <https://doi.org/10.1371/journal.ppat.1002464> PMID: 22363214
17. Barrero Guevara LA, Kramer SC, Kurth T, Domenech de Cellès M. Causal inference concepts can guide research into the effects of climate on infectious diseases. *Nat Ecol Evol*. 2025;9(2):349–63. <https://doi.org/10.1038/s41559-024-02594-3> PMID: 39587221
18. Deyle ER, Maher MC, Hernandez RD, Basu S, Sugihara G. Global environmental drivers of influenza. *Proceedings of the National Academy of Sciences of the United States of America*. 2016;113(46):13081–6.
19. Domenech de Cellès M, Goult E, Casalegno J-S, Kramer SC. The pitfalls of inferring virus-virus interactions from co-detection prevalence data: application to influenza and SARS-CoV-2. *Proc Biol Sci*. 2022;289(1966):20212358. <https://doi.org/10.1098/rspb.2021.2358> PMID: 35016540
20. Shrestha S, King AA, Rohani P. Statistical inference for multi-pathogen systems. *PLoS Comput Biol*. 2011;7(8):e1002135. <https://doi.org/10.1371/journal.pcbi.1002135> PMID: 21876665
21. Granger CWJ. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*. 1969;37(3):424. <https://doi.org/10.2307/1912791>
22. Wiener N. The theory of prediction. In: Beckenback E, editor. *Modern mathematics for the engineer*. New York, NY: McGraw-Hill Book Company; 1956.
23. Schreiber T. Measuring information transfer. *Phys Rev Lett*. 2000;85(2):461–4.
24. Sugihara G, May R, Ye H, Hsieh C, Deyle E, Fogarty M, et al. Detecting causality in complex ecosystems. *Science*. 2012;338(6106):496–500. <https://doi.org/10.1126/science.1227079>
25. Wood SN. *Generalized additive models: an introduction with R*. Second ed. CRC Press; 2017.
26. Hoover KD. Causality in Economics and Econometrics. *The New Palgrave Dictionary of Economics*. Palgrave Macmillan UK; 2008. p. 1–13. https://doi.org/10.1057/978-1-349-95121-5_2227-1
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;35(8):3293–7.
29. Timme NM, Lapish C. A tutorial for information theory in neuroscience. *eNeuro*. 2018;5(3):ENEURO.0052-18.2018. <https://doi.org/10.1523/ENEURO.0052-18.2018> PMID: 30211307
30. Barraquand F, Picoche C, Detto M, Hartig F. Inferring species interactions using Granger causality and convergent cross mapping. *Theor Ecol*. 2021;14(1):87–105.

31. Singh NK, Borrok DM. A Granger causality analysis of groundwater patterns over a half-century. *Sci Rep*. 2019;9(1):12828. <https://doi.org/10.1038/s41598-019-49320-0>
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;120(35):e2305050120. <https://doi.org/10.1073/pnas.2305050120> PMID: 37603760
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;13(1):17220. <https://doi.org/10.1038/s41598-023-44071-0> PMID: 37821495
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;24(3):415–25. <https://doi.org/10.1111/ele.13652> PMID: 33300663
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;17(164):20190628. <https://doi.org/10.1098/rsif.2019.0628> PMID: 32183640
36. Cobey S, Baskerville EB. Limits to Causal Inference with State-Space Reconstruction for Infectious Disease. *PLoS One*. 2016;11(12):e0169050. <https://doi.org/10.1371/journal.pone.0169050> PMID: 28030639
37. Randuineau B. Interactions between pathogens: what are the impacts on public health?. Université Pierre et Marie Curie - Paris VI. 2015. <https://theses.hal.science/tel-01487918>
38. Gemo P, Barrero Guevara LA, Kussmaul C, Kramer SC, de Cellès MD. The pitfalls of incidence-based time series regression for inferring the effects of weather on infectious diseases. *medRxiv*. 2026.
39. Barth N, Carstens G, Kozanli E, Han W, Hermans L, Paolotti D, et al. Population-level associations in the spread of co-circulating respiratory viruses: a multi-method statistical investigation using incidence data. *Epidemics*. 2026. <https://doi.org/10.1016/j.epidem.2026.100925>
40. Chen Y, Tang F, Cao Z, Zeng J, Qiu Z, Zhang C, et al. Global pattern and determinant for interaction of seasonal influenza viruses. *J Infect Public Health*. 2024;17(6):1086–94. <https://doi.org/10.1016/j.jiph.2024.04.024> PMID: 38705061
41. R Core Team. R: A Language and Environment for Statistical Computing. 2024. <https://www.R-project.org>
42. Kramer SC. Statistical methods for inferring pathogen-pathogen interactions. 2025. Accessed 2023 October 1. https://github.com/sarahckramer/Pathogen_interaction_simulations
43. Drori Y, Jacob-Hirsch J, Pando R, Glatman-Freedman A, Friedman N, Mendelson E. Influenza A virus inhibits RSV infection via a two-wave expression of IFIT proteins. *Viruses*. 2020;12(10).
44. 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;14:480. <https://doi.org/10.1186/1471-2334-14-480> PMID: 25186370
45. Reis J, Shaman J. Simulation of four respiratory viruses and inference of epidemiological parameters. *Infect Dis Model*. 2018;3:23–34. <https://doi.org/10.1016/j.idm.2018.03.006> PMID: 30839912
46. Lessler J, Reich NG, Brookmeyer R, Perl TM, Nelson KE, Cummings DAT. Incubation periods of acute respiratory viral infections: a systematic review. *Lancet Infect Dis*. 2009;9(5):291–300. [https://doi.org/10.1016/S1473-3099\(09\)70069-6](https://doi.org/10.1016/S1473-3099(09)70069-6) PMID: 19393959
47. Carrat F, Vergu E, Ferguson NM, Lemaître M, Cauchemez S, Leach S, et al. Time lines of infection and disease in human influenza: a review of volunteer challenge studies. *Am J Epidemiol*. 2008;167(7):775–85. <https://doi.org/10.1093/aje/kwm375> PMID: 18230677
48. Munywoki PK, Koech DC, Agoti CN, Kibirige N, Kipkoech J, Cane PA, et al. Influence of age, severity of infection, and co-infection on the duration of respiratory syncytial virus (RSV) shedding. *Epidemiol Infect*. 2015;143(4):804–12. <https://doi.org/10.1017/S0950268814001393> PMID: 24901443
49. Moore HC, Jacoby P, Hogan AB, Blyth CC, Mercer GN. Modelling the seasonal epidemics of respiratory syncytial virus in young children. *PLoS One*. 2014;9(6):e100422. <https://doi.org/10.1371/journal.pone.0100422> PMID: 24968133
50. Hogan AB, Glass K, Moore HC, Anderssen RS. Exploring the dynamics of respiratory syncytial virus (RSV) transmission in children. *Theor Popul Biol*. 2016;110:78–85. <https://doi.org/10.1016/j.tpb.2016.04.003> PMID: 27155294
51. Hodgson D, Pebody R, Panovska-Griffiths J, Baguelin M, Atkins KE. Evaluating the next generation of RSV intervention strategies: a mathematical modelling study and cost-effectiveness analysis. *BMC Med*. 2020;18(1):348. <https://doi.org/10.1186/s12916-020-01802-8> PMID: 33203423
52. Zheng Z, Weinberger DM, Pitzer VE. Predicted effectiveness of vaccines and extended half-life monoclonal antibodies against RSV hospitalizations in children. *NPJ Vaccines*. 2022;7(1):127. <https://doi.org/10.1038/s41541-022-00550-5> PMID: 36302926
53. Krauer F, Guenther F, Treskova-Schwarzbach M, Schoenfeld V, Koltai M, Jit M, et al. Effectiveness and efficiency of immunisation strategies to prevent RSV among infants and older adults in Germany: a modelling study. *BMC Med*. 2024;22(1):478. <https://doi.org/10.1186/s12916-024-03687-3> PMID: 39420374
54. Galanti M, Comito D, Ligon C, Lane B, Matienzo N, Ibrahim S, et al. Active surveillance documents rates of clinical care seeking due to respiratory illness. *Influenza Other Respir Viruses*. 2020;14(5):499–506. <https://doi.org/10.1111/irv.12753> PMID: 32415751
55. Cohen C, Kleynhans J, Moyes J, McMorrow ML, Treurnicht FK, Helfferssee O, et al. Asymptomatic transmission and high community burden of seasonal influenza in an urban and a rural community in South Africa, 2017–18 (PHIRST): a population cohort study. *Lancet Glob Health*. 2021;9(6):e863–74.
56. Brugger J, Althaus CL. Transmission of and susceptibility to seasonal influenza in Switzerland from 2003 to 2015. *Epidemics*. 2020;30(100373):100373. <https://doi.org/10.1016/j.epidem.2019.100373>

57. 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;6(5):e17835. <https://doi.org/10.1371/journal.pone.0017835> PMID: 21637342
58. 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;4(1):9–21. <https://doi.org/10.1016/j.epidem.2011.11.001> PMID: 22325010
59. 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;8(2):e54445. <https://doi.org/10.1371/journal.pone.0054445> PMID: 23457451
60. Lowen AC, Mubareka S, Steel J, Palese P. Influenza virus transmission is dependent on relative humidity and temperature. *PLoS Pathog*. 2007;3(10):1470–6. <https://doi.org/10.1371/journal.ppat.0030151> PMID: 17953482
61. Shaman J, Kohn M. Absolute humidity modulates influenza survival, transmission, and seasonality. *Proc Natl Acad Sci U S A*. 2009;106(9):3243–8. <https://doi.org/10.1073/pnas.0806852106>
62. Yuan H, Kramer SC, Lau EHY, Cowling BJ, Yang W. Modeling influenza seasonality in the tropics and subtropics. *PLoS Comput Biol*. 2021;17(6):e1009050. <https://doi.org/10.1371/journal.pcbi.1009050> PMID: 34106917
63. 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;9(3):e1003194. <https://doi.org/10.1371/journal.ppat.1003194> PMID: 23505366
64. 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;10(1):5512. <https://doi.org/10.1038/s41467-019-13562-y> PMID: 31797866
65. 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;9(2):e90094. <https://doi.org/10.1371/journal.pone.0090094> PMID: 24587222
66. Bedford T, Suchard MA, Lemey P, Dudas G, Gregory V, Hay AJ, et al. Integrating influenza antigenic dynamics with molecular evolution. *Elife*. 2014;3:e01914. <https://doi.org/10.7554/eLife.01914> PMID: 24497547
67. 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;10(84):20130298. <https://doi.org/10.1098/rsif.2013.0298> PMID: 23676899
68. Gostic KM, Bridge R, Brady S, Viboud C, Worobey M, Lloyd-Smith JO. Childhood immune imprinting to influenza A shapes birth year-specific risk during seasonal H1N1 and H3N2 epidemics. *PLoS Pathog*. 2019;15(12):e1008109. <https://doi.org/10.1371/journal.ppat.1008109> PMID: 31856206
69. Katzelnick LC, Gresh L, Halloran ME, Mercado JC, Kuan G, Gordon A, et al. Antibody-dependent enhancement of severe dengue disease in humans. *Science*. 2017;358(6365):929–32.
70. 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;7(43):271–83. <https://doi.org/10.1098/rsif.2009.0151> PMID: 19535416
71. Bretó C, He D, Ionides EL, King AA. Time series analysis via mechanistic models. *Ann Appl Stat*. 2009;3(1):319–48.
72. King AA, Nguyen D, Ionides EL. Statistical inference for partially observed Markov processes via the R package pomp. *J Stat Softw*. 2016;69:1–43.
73. Centre for Health Protection. Detection of pathogens from respiratory specimens. 2022. Accessed 2022 August 1. <https://www.chp.gov.hk/en/statistics/data/10/641/642/2274.html>
74. Public Health Agency of Canada. Overview of influenza monitoring in Canada. 2023. Accessed 2024 March 19. <https://www.canada.ca/en/public-health/services/diseases/flu-influenza/influenza-surveillance/about-fluwatch.html>
75. National Centers for Environmental Information. Global Surface Summary of the Day - GSOD. 2022. Accessed 2022 August 1. <https://www.ncei.noaa.gov/access/metadata/landing-page/bin/iso?id=gov.noaa.ncdc:C00516>
76. Sparks AH, Hengl T, Nelson A. GSODR: global summary daily weather data in R. *J Open Source Softw*. 2017;2(10):177.
77. Wallace JM, Hobbs PV. Atmospheric science: an introductory survey. Elsevier; 2006.
78. Said SE, Dickey DA. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*. 1984;71(3):599–607. <https://doi.org/10.1093/biomet/71.3.599>
79. Kwiatkowski D, Phillips PCB, Schmidt P, Shin Y. Testing the null hypothesis of stationarity against the alternative of a unit root. *J Econom*. 1992;54(1–3):159–78.
80. Bürkner PC. Brms: An R package for Bayesian multilevel models using Stan. *J Stat Softw*. 2017;80(1).
81. Stan Development Team. Stan Reference Manual. 2024. <https://mc-stan.org>
82. Pfaff B. VAR, SVAR and SVEC models: implementation within R Package vars. *J Stat Softw*. 2008;27(4).
83. Pfaff B. Analysis of integrated and cointegrated time series with R. Second edition. New York, NY: Springer; 2008.
84. Belfrage M, Catani P, Ahlgren N. VARtests: Tests for Error Autocorrelation, ARCH Errors, and Cointegration in Vector Autoregressive Models. 2018. <https://CRAN.R-project.org/package=VARtests>
85. Lizier JT. JIDT: An information-theoretic toolkit for studying the dynamics of complex systems. *Front Robot AI*. 2014.
86. Urbanek S. rJava: Low-Level R to Java Interface. 2024. <https://CRAN.R-project.org/package=rJava>

87. Park J, Smith C, Sugihara G, Deyle E. rEDM: Empirical Dynamic Modeling ('EDM'). 2024. <https://CRAN.R-project.org/package=rEDM>
88. Simpson GL. Modelling palaeoecological time series using generalised additive models. *Frontiers in Ecology and Evolution*. 2018;6.
89. 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;179(4):524–35. <https://doi.org/10.1086/664628> PMID: 22437181
90. Barnett L, Seth AK. Granger causality for state-space models. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2015;91(4):040101. <https://doi.org/10.1103/PhysRevE.91.040101> PMID: 25974424
91. Granger CWJ. Testing for causality. *Journal of Economic Dynamics and Control*. 1980;2:329–52. [https://doi.org/10.1016/0165-1889\(80\)90069-x](https://doi.org/10.1016/0165-1889(80)90069-x)
92. Diebold FX. *Elements of forecasting*. Department of Economics, University of Pennsylvania; 2006.
93. Geweke J. Measurement of linear dependence and feedback between multiple time series. *J Am Stat Assoc*. 1982;77(378):304–13.
94. Palus M, Komárková V, Hrnčíř Z, Sterbová K. Synchronization as adjustment of information rates: detection from bivariate time series. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2001;63(4 Pt 2):046211. <https://doi.org/10.1103/PhysRevE.63.046211> PMID: 11308934
95. Bossomaier T, Barnett L, Harré M, Lizier JT. *An introduction to transfer entropy*. Cham, Switzerland: Springer Nature; 2016.
96. Shannon CE. A mathematical theory of communication. *Bell System Technical Journal*. 1948;27(3):379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
97. Amblard P-O, Michel O. The relation between granger causality and directed information theory: a review. *Entropy*. 2012;15(1):113–43. <https://doi.org/10.3390/e15010113>
98. Barnett L, Barrett AB, Seth AK. Granger causality and transfer entropy are equivalent for Gaussian variables. *Phys Rev Lett*. 2009;103(23):238701. <https://doi.org/10.1103/PhysRevLett.103.238701> PMID: 20366183
99. Certain G, Barraquand F, Gårdmark A. How do MAR(1) models cope with hidden nonlinearities in ecological dynamics?. *Methods in Ecology and Evolution*. 2018;9(9):1975–95.
100. Roberts DL, Nord S. Causality tests and functional form sensitivity. *Appl Econ*. 1985;17(1):135–41.
101. Kraskov A, Stögbauer H, Grassberger P. Estimating mutual information. *Phys Rev E Stat Nonlin Soft Matter Phys*. 2004;69(6 Pt 2):066138. <https://doi.org/10.1103/PhysRevE.69.066138> PMID: 15244698
102. Wibral M, Pampu N, Priesemann V, Siebenhühner F, Seiwert H, Lindner M, et al. Measuring information-transfer delays. *PLoS One*. 2013;8(2):e55809. <https://doi.org/10.1371/journal.pone.0055809> PMID: 23468850
103. Takens F. Detecting strange attractors in turbulence. *Lecture Notes in Mathematics*. Berlin Heidelberg: Springer; 1981. p. 366–81. <https://doi.org/10.1007/bfb0091924>
104. Ye H, Deyle ER, Gilarranz LJ, Sugihara G. Distinguishing time-delayed causal interactions using convergent cross mapping. *Sci Rep*. 2015;5:14750. <https://doi.org/10.1038/srep14750> PMID: 26435402
105. North BV, Curtis D, Sham PC. A note on the calculation of empirical P values from Monte Carlo procedures. *Am J Hum Genet*. 2002;71(2):439–41. <https://doi.org/10.1086/341527> PMID: 12111669
106. Sugihara G. Nonlinear forecasting for the classification of natural time series. *Philos Trans Phys Sci Eng*. 1994;348(1688):477–95.
107. Coenen AR, Weitz JS. Limitations of correlation-based inference in complex virus-microbe communities. *mSystems*. 2018;3(4):e00084-18. <https://doi.org/10.1128/mSystems.00084-18> PMID: 30175237
108. 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*. 2017;284(1858):20170722. <https://doi.org/10.1098/rspb.2017.0722> PMID: 28701561
109. Moriyama M, Hugentobler WJ, Iwasaki A. Seasonality of respiratory viral infections. *Annu Rev Virol*. 2020;7(1):83–101.
110. Martinez ME. The calendar of epidemics: seasonal cycles of infectious diseases. *PLoS Pathog*. 2018;14(11):e1007327. <https://doi.org/10.1371/journal.ppat.1007327> PMID: 30408114
111. 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*. 2021. <https://doi.org/10.7554/eLife.65902>
112. Ijaz MK, Brunner AH, Sattar SA, Nair RC, Johnson-Lussemburg CM. Survival characteristics of airborne human coronavirus 229E. *J Gen Virol*. 1985;66(12):2743–8.
113. Karim YG, Ijaz MK, Sattar SA, Johnson-Lussemburg CM. Effect of relative humidity on the airborne survival of rhinovirus-14. *Can J Microbiol*. 1985;31(11):1058–61. <https://doi.org/10.1139/m85-199> PMID: 3004682
114. 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;114(12):E2272–4.
115. Waterlow NR, Toizumi M, van Leeuwen E, Thi Nguyen H-A, 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;18(6):e1010234. <https://doi.org/10.1371/journal.pcbi.1010234> PMID: 35749561
116. Domenech de Cellès M, Arduin H, Lévy-Bruhl D, Georges S, Souty C, Guillemot D. Unraveling the seasonal epidemiology of pneumococcus. *Proc Natl Acad Sci U S A*. 2019;116(5):1802–7.

117. Howerton E, Williams TC, Casalegno J-S, 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;16(1):7261. <https://doi.org/10.1038/s41467-025-62358-w> PMID: [40770182](https://pubmed.ncbi.nlm.nih.gov/40770182/)
118. 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;10(86):20130414. <https://doi.org/10.1098/rsif.2013.0414> PMID: [23825116](https://pubmed.ncbi.nlm.nih.gov/23825116/)
119. Runge J, Gerhardus A, Varando G, Eyring V, Camps-Valls G. *Nature Reviews Earth & Environment.* 2023;4(7):487–505.
120. Cliff OM, Bryant AG, Lizier JT, Tsuchiya N, Fulcher BD. Unifying pairwise interactions in complex dynamics. *Nat Comput Sci.* 2023;3(10):883–93. <https://doi.org/10.1038/s43588-023-00519-x> PMID: [38177751](https://pubmed.ncbi.nlm.nih.gov/38177751/)