The pitfalls of inferring virus-virus interactions from co-detection prevalence data: Application to influenza and SARS-CoV-2

Matthieu Domenech de Cellès[1,*], Elizabeth Goult[1],
Jean-Sébastien Casalegno[2,3], Sarah Kramer[1]

1. Max Planck Institute for Infection Biology, Infectious Disease Epidemiology group, Charitéplatz 1, Campus Charité Mitte, 10117 Berlin, Germany

2. Laboratoire de Virologie des HCL, IAI, CNR des virus à transmission respiratoire (dont la grippe) Hôpital de la Croix-Rousse F-69317 Lyon cedex 04, France

3. Virpath, Centre International de Recherche en Infectiologie (CIRI), Université de Lyon Inserm U1111, CNRS UMR 5308, ENS de Lyon, UCBL F-69372 Lyon cedex 08, France

***Corresponding author:** Dr. Matthieu Domenech de Cellès, Max Planck Institute for Infection Biology, Charitéplatz 1, Campus Charité Mitte, 10117 Berlin, Germany. E-mail address: domenech@mpiib-berlin.mpg.de

Keywords SARS-CoV-2; COVID-19; influenza; virus-virus interaction; mathematical modeling

Abstract

There is growing experimental evidence that many respiratory viruses—including influenza and SARS-CoV-2—can interact, such that their epidemiological dynamics may not be independent. To assess these interactions, standard statistical tests of independence suggest that the prevalence ratio—defined as the ratio of co-infection prevalence to the product of single-infection prevalences—should equal unity for non-interacting pathogens. As a result, earlier epidemiological studies aimed to estimate the prevalence ratio from co-detection prevalence data, under the assumption that deviations from unity implied interaction. To examine the validity of this assumption, we designed a simulation study that built on a broadly applicable epidemiological model of co-circulation of two respiratory viruses causing seasonal epidemics. By focusing on the pair influenza–SARS-CoV-2, we first demonstrate that the prevalence ratio systematically under-estimates the strength of interaction, and can even misclassify antagonistic or synergistic interactions that persist after clearance of infection. In a global sensitivity analysis, we further identify properties of viral infection—such as a high reproduction number or a short infectious period—that blur the interaction inferred from the prevalence ratio. Altogether, our results suggest that epidemiological studies based on co-detection prevalence data provide a poor guide to assess interactions among respiratory viruses.

Main Text

Introduction

The pandemic of coronavirus disease 2019 (COVID-19), caused by the novel severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), has emphasized the persistent threat posed by respiratory viruses. In addition to SARS-CoV-2, other major respiratory viruses like influenza and the respiratory syncytial virus (RSV) cause a substantial burden every year, estimated at 78 million cases of lower respiratory infections and 130,000 associated deaths worldwide in 2016 [1]. As evidenced by the current and past pandemics, the large host range of respiratory viruses—and the correspondingly high risk of spillover from animals into humans—also makes them prime candidates for emergence of currently unknown "diseases X" [2]. Interaction—here broadly defined as the ability of one pathogen to affect infection or disease caused by another pathogen—is an intriguing yet under-studied aspect of respiratory viruses' biology [3]. Although different nomenclatures have been proposed [4], such interactions can be classified according to their sign, either positive (synonymously, synergistic or facilitatory) or negative (synonymously, antagonistic or competitive). According to experimental evidence, various biological mechanisms exist which make either sign *a priori* plausible [4]. Examples include, in the case of positive interactions, up-regulation of viral target receptors [5] or cell fusion [6]; and, in the case of negative interactions, blocking of viral replication caused by the interferon response [7, 8]. Intriguingly, different respiratory viruses may have opposing effects on COVID-19, *e.g.*, rhinoviruses can inhibit SARS-CoV-2 infection via the interferon response [8], while influenza A viruses can facilitate it via up-regulation of ACE2, the cognate receptor of SARS-CoV-2 in human cells [5, 9, 10]. SARS-CoV-2 interactions may have far-reaching implications for predicting not only the future course of the COVID-19 pandemic, but also the indirect effects of non-COVID-19 vaccines on COVID-19 [11]. Indeed, vaccines that directly target a pathogen may also indirectly affect non-target pathogens that interact with this target pathogen [12, 13, 14].

Because of their relevance to epidemiology and public health, a natural question is how best to identify and estimate interactions between respiratory viruses. Arguably, challenge studies in animals or humans provide the strongest form of evidence, because they can pinpoint the within-host mechanisms of interaction in a controlled experimental setting. However, such studies remain scarce and, more generally, it is not easy to predict their consequences at the scale of human populations [15]. Hence, epidemiological studies—ideally informed by experimental evidence to narrow the search range of interacting pathogens—remain indispensable to assess interactions, but it is unclear whether methods commonly used in such studies are well-suited to this task.

In particular, recent studies of SARS-CoV-2 interactions used a test-negative design [16] to compare

the risk of SARS-CoV-2 infection among those infected with another respiratory virus (*e.g.*, influenza) to that among those uninfected [17, 18, 19]. The underlying idea is conceptually simple: if two (or more) viruses do not interact and circulate independently, then the frequency of co-detection estimated from cross-sectional data should be approximately equal to the product of each virus's detection frequency—conversely, any significant deviation from equality should indicate interaction. However, earlier epidemiological and ecological modeling studies have cautioned against seemingly intuitive metrics of interaction [15, 20, 21]. In fact, to our knowledge the validity of this study design has not yet been systematically tested for respiratory viruses that cause seasonal epidemics.

In this study, we aimed to determine if epidemiological studies based on co-detection prevalence data enabled reliable estimation of interactions between respiratory viruses. To do so, we designed a simulation study that built on a general epidemiological model of co-circulation of two respiratory viruses. We show that cross-sectional estimates of co-infection prevalence—interpreted either alone or in combination with estimates of single-infection prevalences—provide a poor guide to assess interaction. Hence, we argue that earlier epidemiological studies based on this design should be interpreted with caution and that further longitudinal studies will be needed to elucidate the epidemiological interactions of SARS-CoV-2 with other respiratory viruses.

Methods

Transmission model of viral co-circulation We developed a deterministic model of circulation of two respiratory viruses, assumed to interact during the infectious period (*i.e.*, the period of transmissible viral infection, denoted by I) or during a transient period following clearance of infection (denoted by T). According to experimental evidence, such interactions can result from an antiviral state caused by non-specific innate immune responses (such as the interferon response), which develop early during infection and can persist for a short period after clearance of infection [7]. In contrast, we did not model long-term interactions (effected, for example, by adaptive cross-immunity), which are less likely for different species of respiratory viruses [7]. The model was similar to that originally proposed by Shrestha *et al.* [15], with the addition of a latent period (denoted by E) and of a realistic distribution for the infectious period, modeled as a Gamma distribution with shape parameter 2 [22]. The transmission dynamic of each virus was therefore represented by an SEITR model [23], where S represents susceptible individuals and R recovered individuals. Following Shrestha *et al.*, we used a double index notation to indicate the infection status with respect to each virus, *e.g.*, X_{SE} represents the proportion of individuals susceptible to virus 1 and exposed to virus 2. As we primarily focused on respiratory viruses that cause seasonal epidemics lasting a few months, we made

the reasonable assumption of a constant, closed population.

The model was defined by a set of $6 \times 6 = 36$ ordinary differential equations, represented schematically in Fig. 1. The force of infection of each virus $i = \{1, 2\}$ was given by:

$$\begin{aligned} \lambda_1(t) &= R_1\gamma_1 p_1(t) \\ p_1(t) &= \sum_{x\in\Xi}[X_{I_a x}(t) + X_{I_b x}(t)] \\ \lambda_2(t) &= R_2\gamma_2 p_2(t) \\ p_2(t) &= \sum_{x\in\Xi}[X_{x I_a}(t) + X_{x I_b}(t)] \end{aligned}$$

where $\Xi = \{S, E, I_a, I_b, T, R\}$ is the set of state variables, R_i is the reproduction number of virus i, $1/\gamma_i$ the average infectious period of virus i, and $p_i(t)$ the prevalence of infection with virus i. Importantly, as in [24] we assumed that R_i captured pre-existing population immunity. Hence, this parameter is best interpreted here as the initial reproduction number in a partially immune population, as opposed to the basic reproduction number in a fully susceptible population [24]. We also defined the prevalence of individuals co-infected (purple compartments in Fig. 1):

$$p_{12}(t) = X_{I_a I_a}(t) + X_{I_a I_b}(t) + X_{I_b I_a}(t) + X_{I_b I_b}(t)$$

Metric to infer interaction from co-detection prevalence data Standard statistical tests of independence suggest that the following prevalence ratio (PR):

$$\mathrm{PR}(t) = \frac{p_{12}(t)}{p_1(t) \times p_2(t)}$$

could be used to infer interaction [18, 17, 19]. Intuitively, a prevalence ratio above unity indicates that the frequency of co-detection is higher than that expected by chance, suggesting that co-infection is facilitated—that is, that the interaction is positive, *i.e.*, synergistic [20, 21]. Correspondingly, a prevalence ratio below unity would suggest a negative, or antagonistic interaction. In numerical applications, we calculated the prevalence ratio at the time of peak co-infection prevalence, $t_{\max} = \arg\max_t p_{12}(t)$ (cf. Fig. 2), as we reasoned that empirical studies would have maximal statistical power to detect co-infection at that time point. Nevertheless, this choice is arbitrary, and we considered an alternative calculation in a sensitivity analysis, described below. In the following, we drop the time argument ($\mathrm{PR} = \mathrm{PR}(t_{\max})$) and we simply refer to the prevalence ratio calculated at that time point.

Of note, other definitions of the prevalence ratio are possible and have been used in previous studies.

For example, earlier studies of the association between SARS-CoV-2 and influenza compared the fraction of individuals infected with virus 2 among those infected with virus 1 to the fraction infected with virus 2 among those uninfected with virus 1—that is, a test negative design [18, 17, 19]. Using the above notations and after some algebra, the corresponding prevalence ratio $\mathrm{PR}^{'}$ equals:

$$\mathrm{PR}^{'} = \frac{\frac{p_{12}}{p_1}}{\frac{p_2 - p_{12}}{1-p_1}} = \mathrm{PR}\frac{1-p_1}{1-\mathrm{PR}\times p_1}$$

However, this alternative prevalence ratio is no longer symmetric in virus 1 and 2, which implies an arbitrary choice of virus 1. We therefore prefer our formulation, but we point out that the two prevalence ratios are approximately equal for low prevalence of infection with virus 1. Furthermore, it can be shown that $\mathrm{PR}^{'} \geq 1 \iff \mathrm{PR} \geq 1$, such that the sign of the interaction inferred from either ratio is identical.

Model parametrization In numerical applications, we considered the pair influenza (virus 1)–SARS-CoV-2 (virus 2) and we fixed the parameters accordingly (Table 1). Specifically, for influenza we assumed an average latent period of 1 day and an average infectious period of 4 days, resulting in an average generation time of 3 days [25, 26]. For SARS-CoV-2, we assumed an average latent period of 4 days and an average infectious period of 5 days (average generation time of 6.5 days) [27, 28, 29]. The reproduction number of influenza was fixed to 1.3 [24] and that of SARS-CoV-2 to 2.5 [28, 30]. To initialize the model, we assumed that a small fraction $X_{ES}(0) = E_{0,1} = 10^{-3}$ had been exposed to influenza and $X_{SE}(0) = E_{0,2} = 10^{-5}$ to SARS-CoV-2. These initial conditions were chosen to reflect the epidemiological situation in early 2020 in Europe, where influenza was already circulating before the emergence of SARS-CoV-2 [31]. Other individuals were assumed fully susceptible ($X_{SS}(0) = 1 - E_{0,1} - E_{0,2}$), and all other compartments were initialized to 0. As explained above, the initial reproduction number R_i was supposed to capture pre-existing population immunity [24]. For simplicity we considered only symmetric interactions, that is, the effect of virus 1 on virus 2 was assumed equal to that of virus 2 on virus 1. Furthermore, we assumed that interaction could not change sign over the course of infection, and we therefore tested negative ($0 \leq \theta^{(T)}, \theta^{(I)} \leq 1$) and positive ($1 \leq \theta^{(T)}, \theta^{(I)} \leq 5$) interactions separately.

Simulation protocol In all scenarios, the model was integrated numerically for a period of 400 days, with state variable values recorded every 5×10^{-2} days.

Sensitivity analyses for influenza and SARS-CoV-2 To verify the robustness of our results, we conducted two sensitivity analyses. First, we considered an alternative prevalence ratio, similarly defined but averaged ±14 days around the time of peak co-infection prevalence. Second, although earlier experimental

studies found that influenza can affect SARS-CoV-2 infection [9, 10, 5], the effect of SARS-CoV-2 on influenza infection, if any, is currently unknown. Previous experimental studies—*e.g.*, of influenza and RSV [7]—demonstrated the possibility of non-symmetric interactions, where one virus affects the other, but not the other way around. We therefore tested an alternative hypothesis of non-symmetric interactions, for which influenza affected SARS-CoV-2 infection, while SARS-CoV-2 did not affect influenza infection ($\theta_2^{(I)} = \theta_2^{(T)} = 1$).

Global sensitivity analyses To examine more generally the properties of viral infection and interaction that affected the prevalence ratio, we conducted a global sensitivity analysis for a broad range of respiratory viruses [32]. For simplicity, we assumed a fully symmetric model with identical characteristics of the two viruses, and we then proceeded in three steps. First, we used a Latin hypercube design to sample 10^3 values (over the ranges indicated in Table 1, [33]) of the following five parameters: average latent period ($1/\sigma$), average infectious period ($1/\gamma$), average post-infectious period ($1/\delta$), degree of interaction during the infectious period ($\theta^{(I)}$), and degree of interaction during the post-infectious period ($\theta^{(T)}$). Second, we simulated the model and calculated the prevalence ratio for every parameter set. Finally, we used a Normal generalized additive regression model (GAM) to simultaneously estimate the association between the prevalence ratio and every input parameter [34]. For every parameter, the association was modeled using a basis of cubic splines, with a maximum basis dimension of 10. Preliminary analyses indicated that the prevalence ratio was sensitive to the reproduction number, in isolation and in interaction with other parameters. To simplify the regression model, we therefore ran the global sensitivity analysis for three different values of the reproduction number (1.5, 2.0, and 2.5).

Numerical implementation We implemented and simulated all the models using the `pomp` package [35] in R version 3.6.3 [36]. For the global sensitivity analysis, we used the `mgcv` package [34] to fit the GAMs and the `ggeffects` package [37] to plot the marginal effect of each input parameter. Finally, we used the `renv` package to keep track of all packages' version and to increase the results' reproducibility [38].

Results

The prevalence ratio correctly identifies the sign, but not the degree, of uniform interactions We first considered interactions of equal strength during the infectious and post-infectious periods ($\theta = \theta^{(I)} = \theta^{(T)}$)—henceforth referred to as *uniform* interactions. Example simulations of negative, neutral, and positive interactions between influenza and SARS-CoV-2 are plotted in Fig. 2. Compared with the

no-interaction scenario (peak co-infection prevalence: 0.4%), the peak amplitude of co-infection was lower for negative interaction (0.1%) and higher for positive interaction (2.8%). In all scenarios, however, the peak time was approximately identical. Next, we examined the general relationship between the strength of interaction and the prevalence ratio for different values of the post-infectious period in the range 1–14 days (Fig. 3). We found that the prevalence ratio equalled 1 for non-interacting viruses and thus permitted correct identification of neutral interactions ($\theta = 1$). For interacting viruses ($\theta \neq 1$), the prevalence ratio also correctly estimated the sign of the interaction, but systematically under-estimated its strength. Because of a concave association, under-estimation became more severe as the strength of interaction increased. The degree of under-estimation also increased with the duration of the post-infectious period. Hence, we found evidence that the prevalence ratio enabled qualitative, but not quantitative, estimation of uniform interactions.

Higher interaction post-infection can cause the prevalence ratio to misidentify non-uniform interactions Next, we considered the more general case of interactions that differed during the infectious and the post-infectious periods, or *non-uniform* interactions ($\theta^{(I)} \neq \theta^{(T)}$). For these experiments, we assumed an average post-infectious period of 7 days and we tested negative ($0 \leq \theta^{(I)}, \theta^{(T)} \leq 1$) and positive ($1 \leq \theta^{(I)}, \theta^{(T)} \leq 5$) interactions separately. Because higher values of θ actually resulted in lower interaction when the true interaction was assumed negative, in the following we define the strength of interaction as $1 - \theta$ for negative interactions and as θ for positive interactions during either the infectious or the post-infectious period. As shown in Fig. 4, we found that the prevalence ratio was a monotonic function of the strength of interaction during the infectious period, either decreasing for negative interactions or increasing for positive interactions. Hence, in either case stronger interaction during the infectious period helped the prevalence ratio identify the true interaction. In contrast, higher interaction during the post-infectious period blurred the interaction inferred from the prevalence ratio. For weak interaction during infection ($0.8 \leq \theta^{(I)} \leq 1.75$), these two opposing effects combined caused the prevalence ratio to misidentify the sign of interaction in scenarios with strong interaction post-infection. In the other scenarios, the prevalence ratio correctly identified the sign of the interaction, but substantially under-estimated its strength (*e.g.*, prevalence ratio of 0.44 for $\theta^{(I)} = 0$ and $\theta^{(T)} = 1$, of 1.82 for $\theta^{(I)} = 5$ and $\theta^{(T)} = 1$). These experiments demonstrate that the prevalence ratio is an unreliable measure of interaction between influenza and SARS-CoV-2.

Sensitivity analyses demonstrate the results' robustness for influenza and SARS-CoV-2 In sensitivity analyses, we first verified that our results were robust to an alternative calculation of the prevalence ratio (Fig. S1). Second, we repeated our analyses for non-symmetric interactions with no effect of SARS-CoV-

2 on influenza infection (Fig. S2). The results were broadly comparable to those for symmetric interactions (Fig. 4), except that fewer parameter combinations caused the prevalence ratio to mis-identify the sign of interaction. However, the strength of interaction was also more severely under-estimated in this scenario (prevalence ratio range: 0.80–1.41, compared with 0.44–1.82 for symmetric interactions).

Global sensitivity analysis highlights properties of viral infection that obscure or facilitate estimation of interaction In a global sensitivity analysis of positive interactions ($\theta^{(I)}, \theta^{(T)} \geq 1$), we assessed how different properties of viral infection and interaction affected the prevalence ratio. As shown in Fig. 5, the prevalence ratio decreased with the average latent period, the average post-infectious period, and the strength of interaction post-infection. Hence, these three parameters blurred the interaction inferred from the prevalence ratio. Conversely, the average length of, and the strength of interaction during, the infectious period increased with the prevalence ratio and therefore facilitated estimation of the interaction. Of note, higher values of the reproduction number dampened all these variations. To understand the effect of each parameter on the prevalence ratio, we propose that some insights can be gained by examining how and when each parameter affects the prevalences of single infections and co-infection. For example, it is likely that parameters that enhance interaction after infection (*i.e.*, higher $\theta^{(T)}$ and $1/\delta$) affect single-infection prevalences more rapidly and strongly than co-infection prevalence, thereby decreasing the prevalence ratio. In sum, these results confirm our earlier experiments on influenza and SARS-CoV-2 and highlight additional factors that make it difficult to interpret the prevalence ratio as a measure of interactions between respiratory viruses.

Discussion

In this study, we aimed to determine if the prevalence ratio—defined as the ratio of the prevalence of co-infection to the product of individual infection prevalences—enabled reliable estimation of interactions between respiratory viruses. To do so we designed a simulation study that built on a broadly applicable epidemiological model of co-circulation of two respiratory viruses. By focusing on the pair influenza–SARS-CoV-2, we first demonstrated that the prevalence ratio systematically under-estimated the strength of interaction, and could even mistake the sign of interactions that persisted after clearance of infection. In a global sensitivity analysis, we further identified properties of viral infection—such as a high reproduction number, a long latent period, or a short infectious period—that blurred the interaction inferred from the prevalence ratio. Our results show that, in the absence of precise information about the timing of interaction, epidemiological studies designed to estimate the prevalence ratio, or variations thereof, may be unreliable.

With the likely prospect of COVID-19 becoming endemic, there is a pressing need to elucidate the potential interactions of SARS-CoV-2 with other pathogens, in particular respiratory viruses. Thus far, most epidemiological studies of SARS-CoV-2 interaction used simple statistics of co-circulation, such as the prevalence of co-infection, the prevalence ratio, or some variation thereof [18, 17, 19, 39, 40]. As we showed here, however, such studies—even those carefully designed to control for various sources of bias like age or co-morbidities—are likely uninformative. Besides the prevalence ratio, we found that the prevalence of co-infection was also an unreliable measure of interaction, as low prevalences could be consistent with strong, positive interactions (Fig. 2, bottom panel). As suggested by our global sensitivity analysis, these deficiencies may be even more severe for SARS-CoV-2 infection, characterized by a relatively long latent period and a high reproduction number [28]. In sum, we submit that further epidemiological studies will be needed to elucidate the interactions of SARS-CoV-2 with other respiratory viruses.

More generally, our study adds to the growing body of evidence demonstrating the shortcomings of seemingly intuitive measures of interaction. Using the same model, Shrestha *et al.* demonstrated the unreliability of phase as an indicator of interaction [15]. Using a SIS-like model of multiple pathogens causing chronic infection, Hamelin *et al.* showed that the prevalence ratio (as defined in this study) exceeded unity for non-interacting pathogens [21]. In contrast, we found that the prevalence ratio equalled unity for non-interacting pathogens (Fig. 3). This discrepancy, explained by the different pathogens and modeling assumptions considered in [21], highlights the sensitivity of the prevalence ratio to the characteristics of infection. More generally, it suggests that our results cannot be extrapolated to pathogens not well described by the SIR-like model used here. Using a series of SIS and SIRS models, Man *et al.* examined the properties of the odds ratio, defined as the ratio of the odds of one type in the presence of the other type, relative to the odds of this type in the absence of the other type—a quantity closely related to the alternative prevalence ratio PR' defined above [20]. They proved that odds ratio exceeding unity could mask negative interactions. Despite differences in the scope of, and the models used in, this study, our results replicate this finding (Fig. 4). Furthermore, the association between the prevalence ratio and the interaction parameter in our study (Fig. 3) is comparable to that in [20] (Figure 2A, SIRS direct model). Finally, in a field study to assess interactions between an intestinal pathogen and nematodes in mice (where the true sign of interaction was known from previous experimental evidence), Fenton *et al.* reported that statistical methods based on cross-sectional data performed poorly and typically estimated the wrong sign of interaction [41]. Our results align with these findings, and we second Fenton *et al.*'s caution against the use of such methods to study pathogen-pathogen interactions. In sum, our study broadly agrees with previous evidence, and provides new evidence specific to the epidemiology of respiratory viruses.

The shortcomings of the prevalence ratio demonstrated here might suggest the need for new statistical

methods to estimate interaction from co-detection prevalence data. However, seconding Fenton *et al.* [41], we propose that methods based on longitudinal data—collected at an appropriately fine time scale—offer a more promising avenue of research. Among those methods, mathematical models of transmission provide a powerful tool to formulate and test biologically explicit mechanisms of interaction, while capturing the underlying, non-linear dynamics of infection of each pathogen [42]. Robust statistical inference techniques now facilitate fitting these models to epidemiological time series [43, 44], as demonstrated by earlier successful applications in the field of pathogen interactions [45, 46, 47, 48, 49, 50]. Alternatively, advanced regression models have been developed to assess interactions between respiratory viruses [51], but such models may be limited because they lack a mechanistic formulation of interaction. Altogether we propose that empirical or mechanistic models of longitudinal data will be required to study the interactions of SARS-CoV-2 with other respiratory viruses, and more generally the interactions between respiratory pathogens [52].

Our study has four important limitations. First, because we used a deterministic model expressed in proportions, we sidestepped the important issue of statistical uncertainty, caused for example by finite sample size or imperfect measurement of infection prevalences. As the prevalence ratio was found to systematically under-estimate the strength of interaction, such uncertainty—inevitable in practice—may further limit the ability of the prevalence ratio to correctly identify interactions. Second, for simplicity we did not include confounding variables (*e.g.*, age) that may also affect estimation of the prevalence ratio. Third, we considered only short-term interactions that rapidly waned after clearance of infection. Although such interactions appear to be the most biologically plausible for different species of respiratory viruses [7, 8], long-term interactions resulting from adaptive cross-immunity have been documented and could be relevant to other systems, such as the multiple types or subtypes of influenza [53, 54]. Fourth, for simplicity we only modeled interactions that affected susceptibility to infection, because experimental evidence suggests this mechanism predominates among respiratory viruses [7, 8]. However, other mechanisms—like changes in the transmissibility or the duration of infection—are biologically likely and could be tested for other classes of pathogens. Acknowledging all these limitations, our simple model could serve as a building block for further research on epidemiological interactions.

In conclusion, our results show that the inherently complex, non-linear dynamic of respiratory viruses makes the interpretation of seemingly intuitive measures of interaction difficult, if not impossible. Despite these pitfalls, other statistical or mathematical methods based on longitudinal data should enable epidemiological research on pathogen interactions. Indeed, with increasing evidence that SARS-CoV-2 and other pathogens do not circulate in isolation but within polymicrobial systems, such research should remain a priority.

References

[1] GBD 2016 Lower Respiratory Infections Collaborators. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory infections in 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Infect Dis. 2018 11;18(11):1191–1210.

[2] WHO Research and Development Blueprint. 2018 Annual review of diseases prioritized under the Research and Development Blueprint. World Health Organization; 2018. Available from: `https://s.gwdg.de/fUFmoO`.

[3] Griffiths EC, Pedersen AB, Fenton A, Petchey OL. The nature and consequences of coinfection in humans. J Infect. 2011 Sep;63(3):200–6.

[4] DaPalma T, Doonan BP, Trager NM, Kasman LM. A systematic approach to virus-virus interactions. Virus Res. 2010 Apr;149(1):1–9.

[5] Bai L, Zhao Y, Dong J, Liang S, Guo M, Liu X, et al. Coinfection with influenza A virus enhances SARS-CoV-2 infectivity. Cell Res. 2021 04;31(4):395–403.

[6] Goto H, Ihira H, Morishita K, Tsuchiya M, Ohta K, Yumine N, et al. Enhanced growth of influenza A virus by coinfection with human parainfluenza virus type 2. Med Microbiol Immunol. 2016 Jun;205(3):209–18.

[7] 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 07;218(3):406–417.

[8] Dee K, Goldfarb DM, Haney J, Amat JAR, Herder V, Stewart M, et al. Human rhinovirus infection blocks SARS-CoV-2 replication within the respiratory epithelium: implications for COVID-19 epidemiology. J Infect Dis. 2021 Mar;.

[9] Smith JC, Sausville EL, Girish V, Yuan ML, Vasudevan A, John KM, et al. Cigarette Smoke Exposure and Inflammatory Signaling Increase the Expression of the SARS-CoV-2 Receptor ACE2 in the Respiratory Tract. Dev Cell. 2020 06;53(5):514–529.e3.

[10] Ziegler CGK, Allon SJ, Nyquist SK, Mbano IM, Miao VN, Tzouanas CN, et al. SARS-CoV-2 Receptor ACE2 Is an Interferon-Stimulated Gene in Human Airway Epithelial Cells and Is Detected in Specific Cell Subsets across Tissues. Cell. 2020 05;181(5):1016–1035.e19.

[11] Thindwa D, Garcia Quesada M, Liu Y, Bennett J, Cohen C, Knoll MD, et al. Use of seasonal influenza and pneumococcal polysaccharide vaccines in older adults to reduce COVID-19 mortality. Vaccine. 2020 07;38(34):5398–5401.

[12] Brogden KA, Guthmiller JM, Taylor CE. Human polymicrobial infections. Lancet. 2005;365(9455):253–5.

[13] Shann F. The non-specific effects of vaccines. Arch Dis Child. 2010 Sep;95(9):662–7.

[14] Cowling BJ, Fang VJ, Nishiura H, Chan KH, Ng S, Ip DKM, et al. Increased risk of noninfluenza respiratory virus infections associated with receipt of inactivated influenza vaccine. Clin Infect Dis. 2012 Jun;54(12):1778–83.

[15] Shrestha S, King AA, Rohani P. Statistical inference for multi-pathogen systems. PLoS Comput Biol. 2011 Aug;7(8):e1002135.

[16] Vandenbroucke JP, Pearce N. Test-Negative Designs: Differences and Commonalities with Other Case-Control Studies with "Other Patient" Controls. Epidemiology. 2019 11;30(6):838–844.

[17] Stowe J, Tessier E, Zhao H, Guy R, Muller-Pebody B, Zambon M, et al. Interactions between SARS-CoV-2 and influenza, and the impact of coinfection on disease severity: a test-negative design. Int J Epidemiol. 2021 May;.

[18] Kim D, Quinn J, Pinsky B, Shah NH, Brown I. Rates of Co-infection Between SARS-CoV-2 and Other Respiratory Pathogens. JAMA. 2020 May;323(20):2085–2086.

[19] Nowak MD, Sordillo EM, Gitman MR, Paniz Mondolfi AE. Coinfection in SARS-CoV-2 infected patients: Where are influenza virus and rhinovirus/enterovirus? J Med Virol. 2020 10;92(10):1699–1700.

[20] Man I, Wallinga J, Bogaards JA. Inferring Pathogen Type Interactions Using Cross-sectional Prevalence Data: Opportunities and Pitfalls for Predicting Type Replacement. Epidemiology. 2018 09;29(5):666–674.

[21] Hamelin FM, Allen LJS, Bokil VA, Gross LJ, Hilker FM, Jeger MJ, et al. Coinfections by noninteracting pathogens are not independent and require new tests of interaction. PLoS Biol. 2019 12;17(12):e3000551.

[22] Lloyd AL. Destabilization of epidemic models with the inclusion of realistic distributions of infectious periods. Proc Biol Sci. 2001 May;268(1470):985–93.

[23] Keeling MJ, Rohani P. Modeling infectious diseases in humans and animals. Princeton: Princeton University Press; 2008. Available from: `http://www.loc.gov/catdir/toc/fy0805/2006939548.html`.

[24] Chowell G, Miller MA, Viboud C. Seasonal influenza in the United States, France, and Australia: transmission and prospects for control. Epidemiol Infect. 2008 Jun;136(6):852–64.

[25] Carrat F, Vergu E, Ferguson NM, Lemaitre 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 Apr;167(7):775–85.

[26] Vink MA, Bootsma MCJ, Wallinga J. Serial intervals of respiratory infectious diseases: a systematic review and analysis. Am J Epidemiol. 2014 Nov;180(9):865–75.

[27] Bi Q, Wu Y, Mei S, Ye C, Zou X, Zhang Z, et al. Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. Lancet Infect Dis. 2020 08;20(8):911–919.

[28] Li R, Pei S, Chen B, Song Y, Zhang T, Yang W, et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). Science. 2020;368(6490):489–493.

[29] Flaxman S, Mishra S, Gandy A, Unwin HJT, Mellan TA, Coupland H, et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. Nature. 2020 08;584(7820):257–261.

[30] Kucharski AJ, Russell TW, Diamond C, Liu Y, Edmunds J, Funk S, et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. Lancet Infect Dis. 2020 05;20(5):553–558.

[31] Melidou A, Pereyaslov D, Hungnes O, Prosenc K, Alm E, Adlhoch C, et al. Virological surveillance of influenza viruses in the WHO European Region in 2019/20 - impact of the COVID-19 pandemic. Euro Surveill. 2020 11;25(46).

[32] Wu J, Dhingra R, Gambhir M, Remais JV. Sensitivity analysis of infectious disease models: methods, advances and their application. J R Soc Interface. 2013 Sep;10(86):20121018.

[33] Spencer JA, Shutt DP, Moser SK, Clegg H, Wearing HJ, Mukundan H, et al. Distinguishing Viruses Responsible for Influenza-Like Illness. medRxiv. 2021;p. 2020–02.

[34] Wood SN. Generalized additive models: an introduction with R. 2nd ed. CRC press; 2017.

[35] Aaron A King and Edward L Ionides and Carles Martinez Bretó and Stephen P Ellner and Matthew J Ferrari and Bruce E Kendall and Michael Lavine and Dao Nguyen and Daniel C Reuman and Helen Wearing and Simon N Wood. pomp: Statistical Inference for Partially Observed Markov Processes; 2020. R package, version 2.7. Available from: `https://kingaa.github.io/pomp/`.

[36] R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria; 2020. Https://www.R-project.org/. Available from: `https://www.R-project.org/`.

[37] Daniel Lüdecke. ggeffects: Tidy Data Frames of Marginal Effects from Regression Models. Journal of Open Source Software. 2018;3(26):772.

[38] Kevin Ushey. renv: Project Environments; 2021. R package version 0.12.5. Available from: `https://CRAN.R-project.org/package=renv`.

[39] Lansbury L, Lim B, Baskaran V, Lim WS. Co-infections in people with COVID-19: a systematic review and meta-analysis. J Infect. 2020 08;81(2):266–275.

[40] Musuuza JS, Watson L, Parmasad V, Putman-Buehler N, Christensen L, Safdar N. Prevalence and outcomes of co-infection and superinfection with SARS-CoV-2 and other pathogens: A systematic review and meta-analysis. PLoS One. 2021;16(5):e0251170.

[41] Fenton A, Knowles SCL, Petchey OL, Pedersen AB. The reliability of observational approaches for detecting interspecific parasite interactions: comparison with experimental results. Int J Parasitol. 2014 Jun;44(7):437–45.

[42] Metcalf CJE, Walter KS, Wesolowski A, Buckee CO, Shevliakova E, Tatem AJ, et al. Identifying climate drivers of infectious disease dynamics: recent advances and challenges ahead. Proc Biol Sci. 2017 Aug;284(1860).

[43] Ionides EL, Nguyen D, Atchadé Y, Stoev S, King AA. Inference for dynamic and latent variable models via iterated, perturbed Bayes maps. Proc Natl Acad Sci U S A. 2015 Jan;112(3):719–24.

[44] King AA, Nguyen D, Ionides EL. Statistical Inference for Partially Observed Markov Processes via the R Package pomp. Journal of Statistical Software. 2016;69(1):1–43.

[45] Shrestha S, Foxman B, Weinberger DM, Steiner C, Viboud C, Rohani P. Identifying the interaction between influenza and pneumococcal pneumonia using incidence data. Sci Transl Med. 2013 Jun;5(191):191ra84.

[46] Opatowski L, Varon E, Dupont C, Temime L, van der Werf S, Gutmann L, et al. Assessing pneumococcal meningitis association with viral respiratory infections and antibiotics: insights from statistical and mathematical models. Proc Biol Sci. 2013 Aug;280(1764):20130519.

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

[48] Shrestha S, Foxman B, Berus J, van Panhuis WG, Steiner C, Viboud C, et al. The role of influenza in the epidemiology of pneumonia. Sci Rep. 2015 Oct;5:15314.

[49] Domenech de Cellès M, Arduin H, Lévy-Bruhl D, Georges S, Souty C, Guillemot D, et al. Unraveling the seasonal epidemiology of pneumococcus. Proc Natl Acad Sci U S A. 2019 01;116(5):1802–1807.

[50] Noori N, Rohani P. Quantifying the consequences of measles-induced immune modulation for whooping cough epidemiology. Philos Trans R Soc Lond B Biol Sci. 2019 06;374(1775):20180270.

[51] Nickbakhsh S, Mair C, Matthews L, Reeve R, Johnson PCD, Thorburn F, et al. Virus-virus interactions impact the population dynamics of influenza and the common cold. Proc Natl Acad Sci U S A. 2019 Dec;.

[52] 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 02;14(2):e1006770.

[53] Terajima M, Babon JAB, Ennis FA, et al. Cross-reactive human B cell and T cell epitopes between influenza A and B viruses. Virology journal. 2013;10(1):1–10.

[54] Laurie KL, Guarnaccia TA, Carolan LA, Yan AWC, Aban M, Petrie S, et al. Interval Between Infections and Viral Hierarchy Are Determinants of Viral Interference Following Influenza Virus Infection in a Ferret Model. J Infect Dis. 2015 Dec;212(11):1701–10.

Declarations

Availability of data and materials

All R programming codes can be found at `https://transfer.mpiib-berlin.mpg.de/s/mKpySLNBcZoQt8D`.

Competing interests

Competing interests MDdC received postdoctoral funding (2017–2019) from Pfizer and consulting fees from GSK. All other authors declare no competing interests.

Funding

No specific funding was used for this study.

Authors Contributions

MDdC conceived of the study design and performed the analysis. EG, JSC, SK conducted bibliographical research and provided content expertise. All authors helped draft, and approved the final version of, the manuscript.

Acknowledgements

We thank Laura Barrero and Michael Briga for helpful comments on the manuscript.

List of abbreviations

- COVID-19: coronavirus disease 2019
- SARS-CoV-2: severe acute respiratory syndrome coronavirus 2
- RSV: respiratory syncytial virus

List of supplementary materials

- Figures S1–S2

Figures and tables

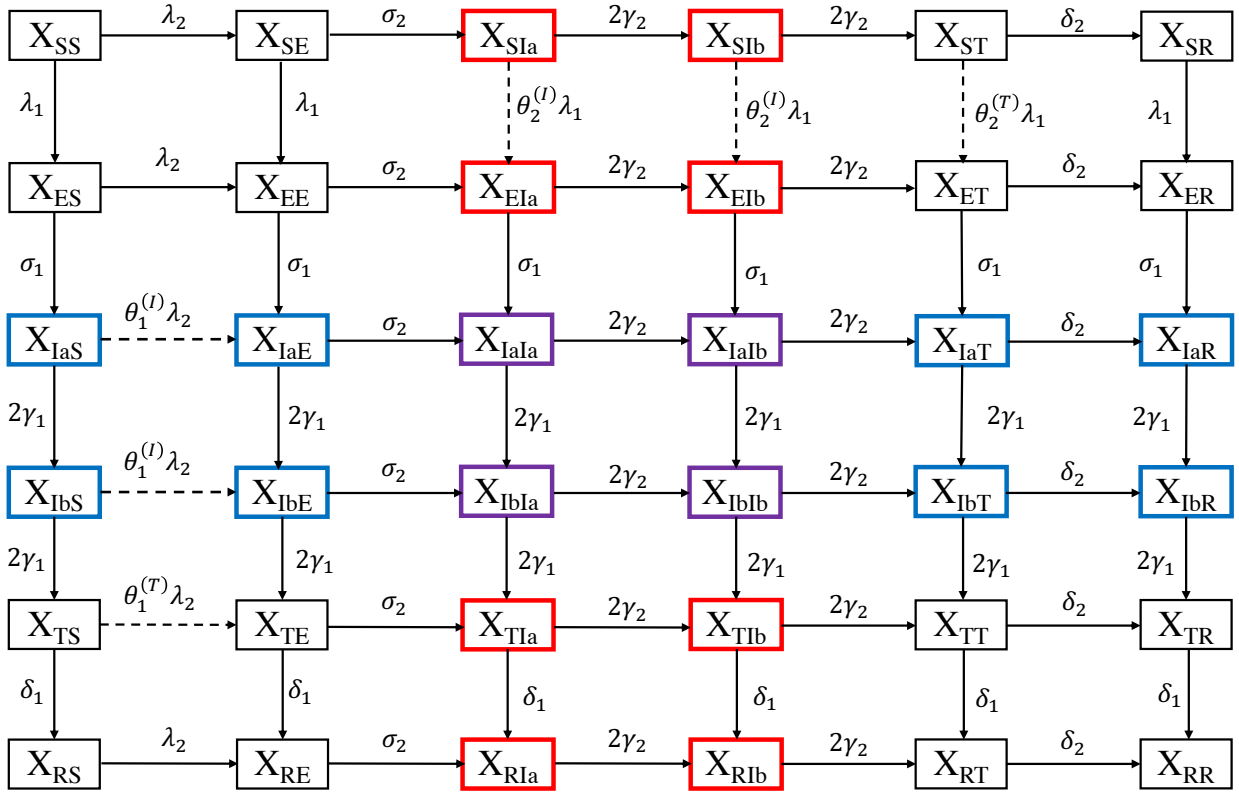

Figure 1: **Schematic of epidemiological model of viral co-circulation.** Individuals infectious with virus 1 are highlighted in blue, with virus 2 in red, and with both viruses in purple. Dashed lines indicate epidemiological transitions affected by interactions.

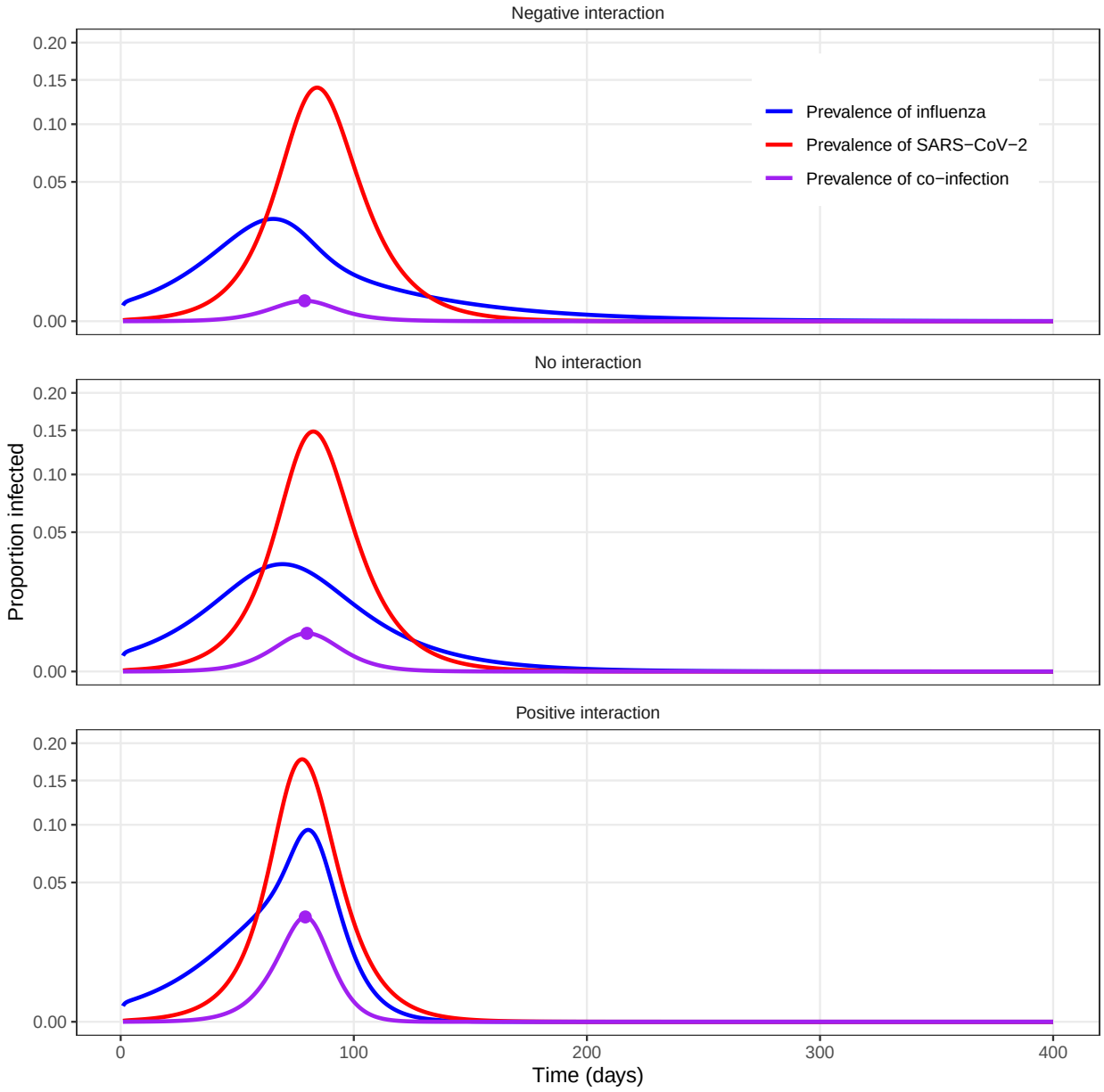

Figure 2: **Example model simulations for different types of uniform interactions between influenza and SARS-CoV-2**. The simulations depicted correspond to uniform interactions ($\theta = \theta_1^{(I)} = \theta_1^{(T)} = \theta_2^{(T)} = \theta_2^{(I)}$), either negative ($\theta = 0$, upper panel), neutral ($\theta = 1$, middle panel), or positive ($\theta = 5$, lower panel). Here the average post-infectious period was fixed to $\frac{1}{\delta} = \frac{1}{\delta_1} = \frac{1}{\delta_2} = 1$ day; other parameters were fixed to model the coupled dynamics of influenza and SARS-CoV-2 (cf. Table 1). The purple points indicate the peak of co-infection prevalence, when the prevalence ratio was calculated (numerical values of the prevalence ratio: 0.45, 1.00, and 1.72 from top to bottom). In every panel, the y-axis values are square-root transformed to highlight the peaks.

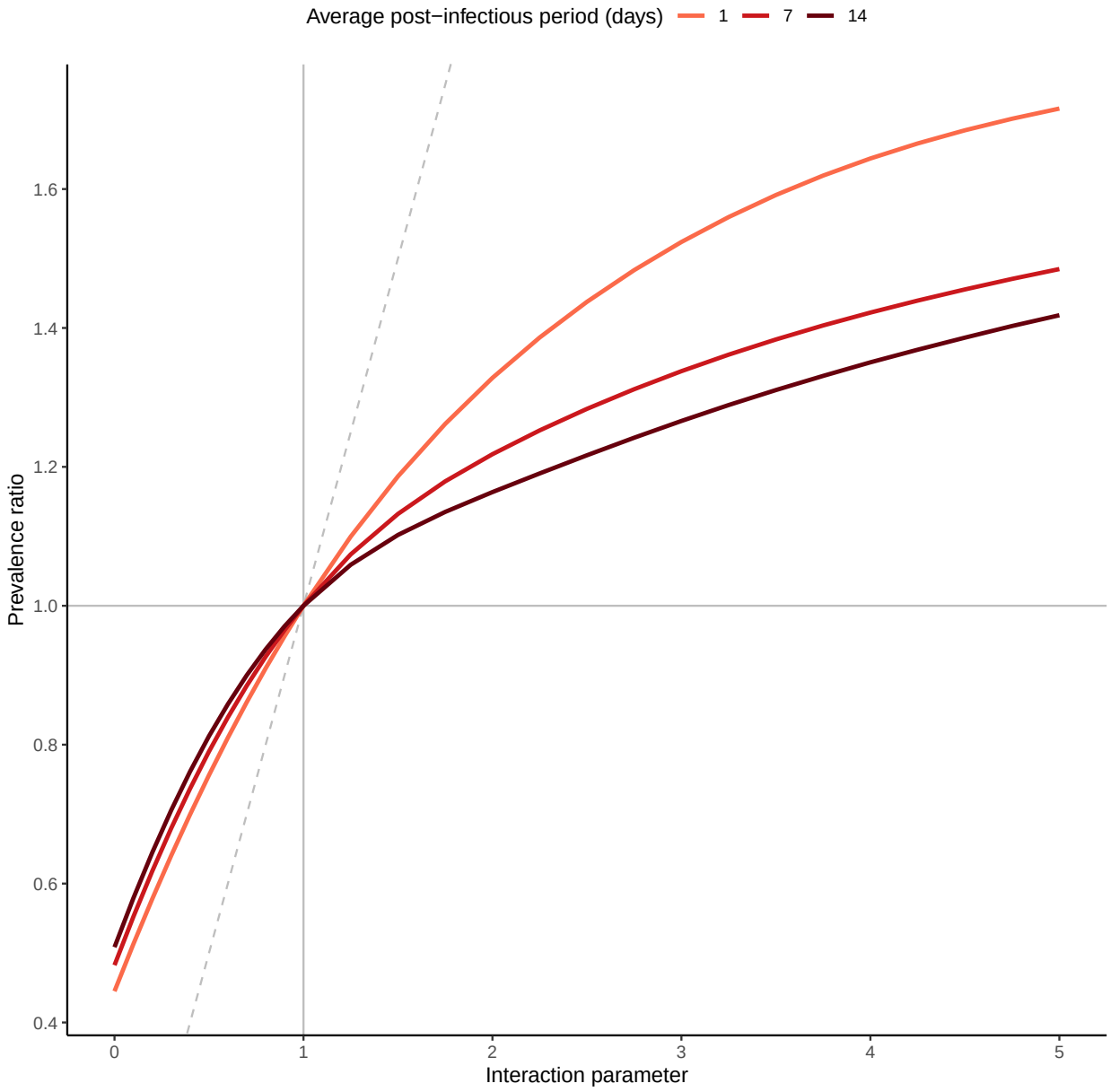

Figure 3: **Relationship between strength of interaction and prevalence ratio for uniform interactions between influenza and SARS-CoV-2**. The scenarios tested correspond to $\theta_1^{(I)} = \theta_1^{(T)} = \theta_2^{(T)} = \theta_2^{(I)}$ (x-axis), for three different values of the average post-infectious period ($\frac{1}{\delta} = \frac{1}{\delta_1} = \frac{1}{\delta_2}$); other parameters were fixed to model the coupled dynamics of influenza and SARS-CoV-2 (cf. Table 1). The dashed grey identity line depicts equality between the prevalence ratio and the true strength of interaction.

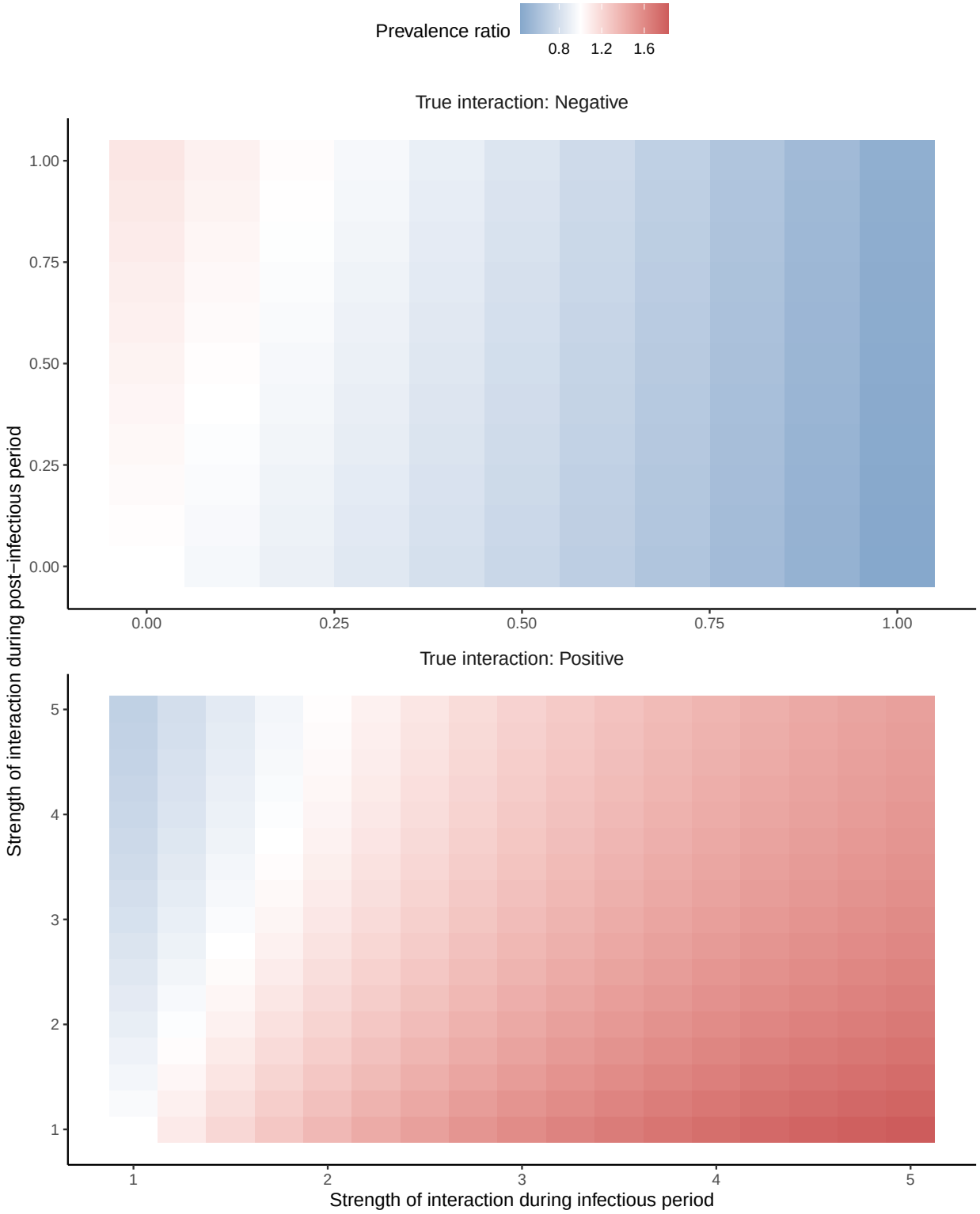

Figure 4: **Relationship between strength of interaction and prevalence ratio for non-uniform interactions between influenza and SARS-CoV-2**. The scenarios tested correspond to $\theta_1^{(I)} = \theta_2^{(I)} = \theta^{(I)}$ and $\theta_1^{(T)} = \theta_2^{(T)} = \theta^{(T)}$; other parameters were fixed to model the coupled dynamics of influenza and SARS-CoV-2(cf. Table 1). For negative interactions (top panel), the x-axis represents $1-\theta^{(I)}$ and the y-axis $1-\theta^{(T)}$; for positive interactions (bottom panel) $\theta^{(I)}$ and $\theta^{(T)}$. Hence, in either panel the true strength of interaction increases from left to right and from bottom to top.

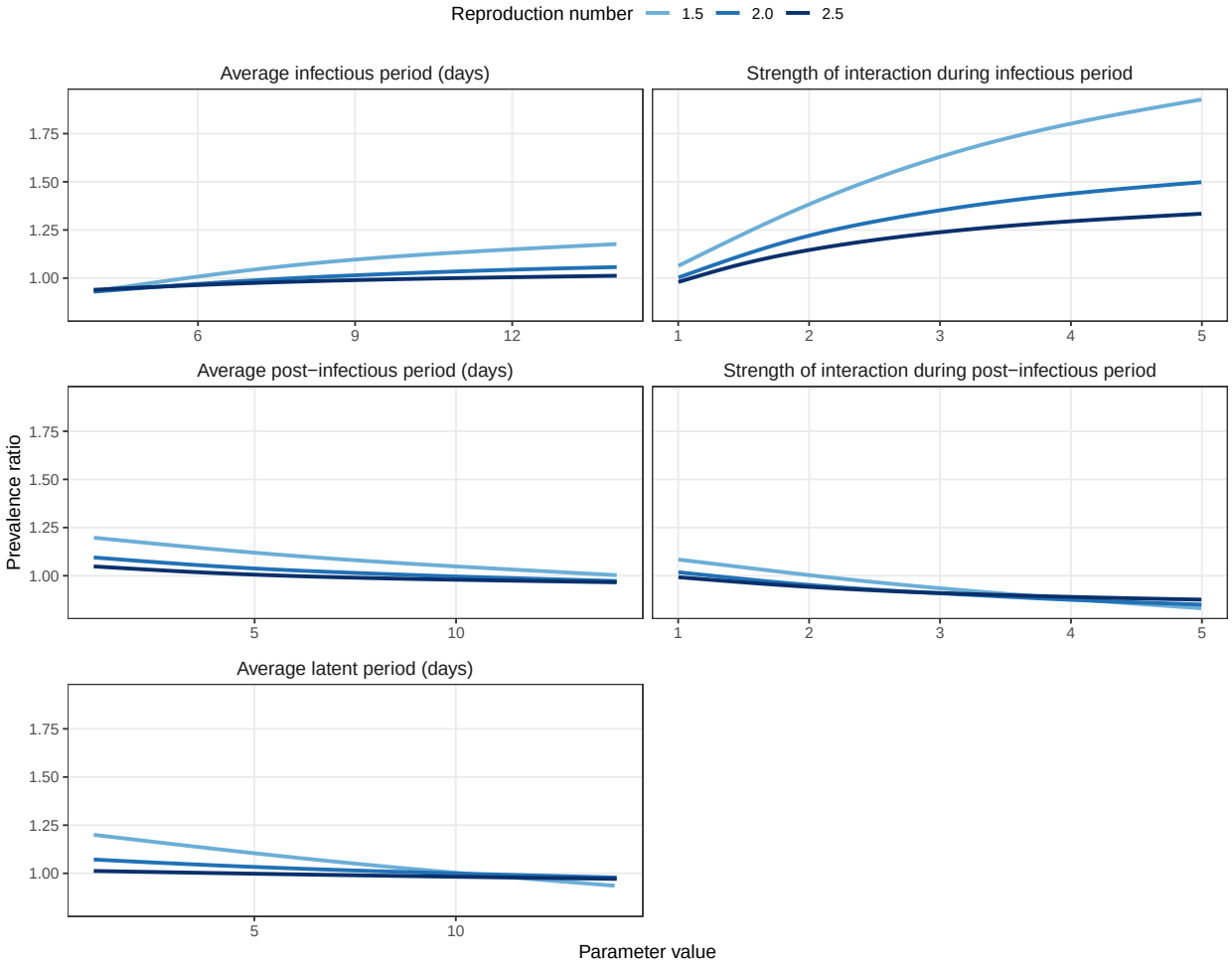

Figure 5: **Global sensitivity analysis for positive virus-virus interactions**. The association between the prevalence ratio and each input parameter was estimated using a GAM with cubic splines (sample size $n = 10^3$), for three different values of the reproduction number (1.5, 2.0, and 2.5). The corresponding adjusted R-squared equalled 96.3%, 96.6%, and 98.4%. For visual clarity, the x-axis values differ between panels.

Parameter	Meaning	Fixed value or interval (influenza–SARS-CoV-2 analysis)	Fixed value or interval (global sensitivity analysis)	Source/Comment
σ_1^{-1}	Average latent period of influenza	1 day	$\sigma_1^{-1} = \sigma_2^{-1} = \sigma^{-1}$ $\sigma^{-1} \in [1,14]$ days	[26]
σ_2^{-1}	Average latent period of SARS-CoV-2	4 days		[29]
γ_1^{-1}	Average infectious period of influenza	4 days	$\gamma_1^{-1} = \gamma_2^{-1} = \gamma^{-1}$ $\gamma^{-1} \in [4,14]$ days	[27]
γ_2^{-1}	Average infectious period of SARS-CoV-2	5 days		[28]
R_1	Reproduction number of influenza	1.3	$R_1 = R_2 = R$ $R \in \{1.5, 2.0, 2.5\}$	[25]
R_2	Reproduction number of SARS-CoV-2	2.5		[29, 31]
$E_{0,1}$	Initial fraction exposed to influenza	10^{-3}	$E_{0,1} = E_{0,2} = 10^{-5}$	Assumption: Influenza circulated before SARS-CoV-2 [32]
$E_{0,2}$	Initial fraction exposed to SARS-CoV-2	10^{-5}		
$\delta^{-1} = \delta_1^{-1} = \delta_2^{-1}$	Average post-infectious period	1–14 days	1–14 days	[34]
$\theta^{(I)} = \theta_1^{(I)} = \theta_2^{(I)}$	Strength of interaction during infectious period	0–5	1–5	Assumption
$\theta^{(T)} = \theta_1^{(T)} = \theta_2^{(T)}$	Strength of interaction during post-infectious period	0–5	1–5	Assumption

Table 1: **List of model parameters**.

Supplementary Materials

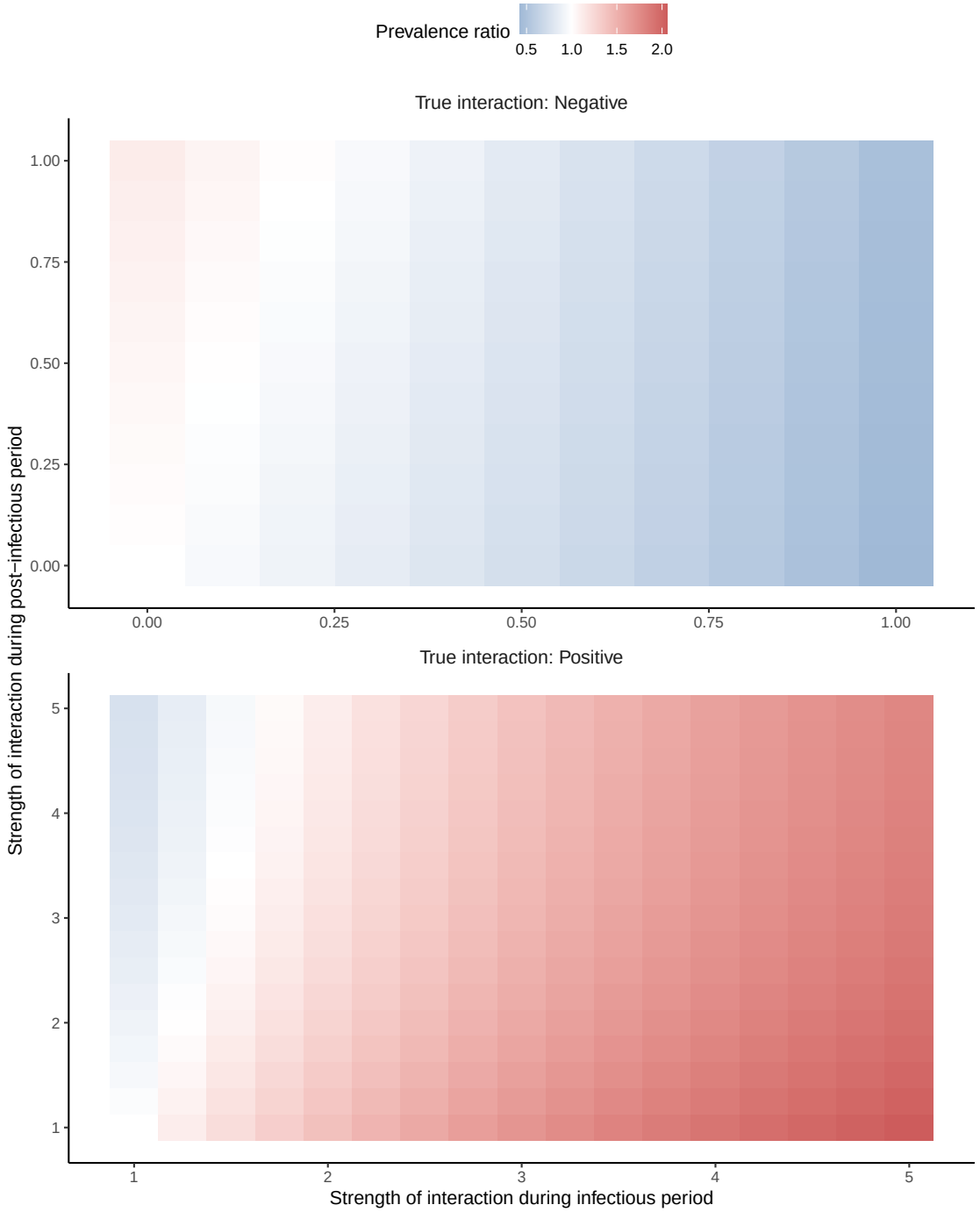

Figure S1: **Relationship between strength of interaction and prevalence ratio for non-uniform interactions between influenza and SARS-CoV-2 (prevalence ratio time-averaged ± 14 days around the peak time of co-infection)**. The scenarios tested correspond to $\theta_1^{(I)} = \theta_2^{(I)} = \theta^{(I)}$ and $\theta_1^{(T)} = \theta_2^{(T)} = \theta^{(T)}$; other parameters were fixed to model the coupled dynamics of influenza and SARS-CoV-2(cf. Table 1). For negative interactions (top panel), the x-axis represents $1 - \theta^{(I)}$ and the y-axis $1 - \theta^{(T)}$; for positive interactions (bottom panel) $\theta^{(I)}$ and $\theta^{(T)}$. Hence, in either panel the true strength of interaction increases from left to right and from bottom to top.

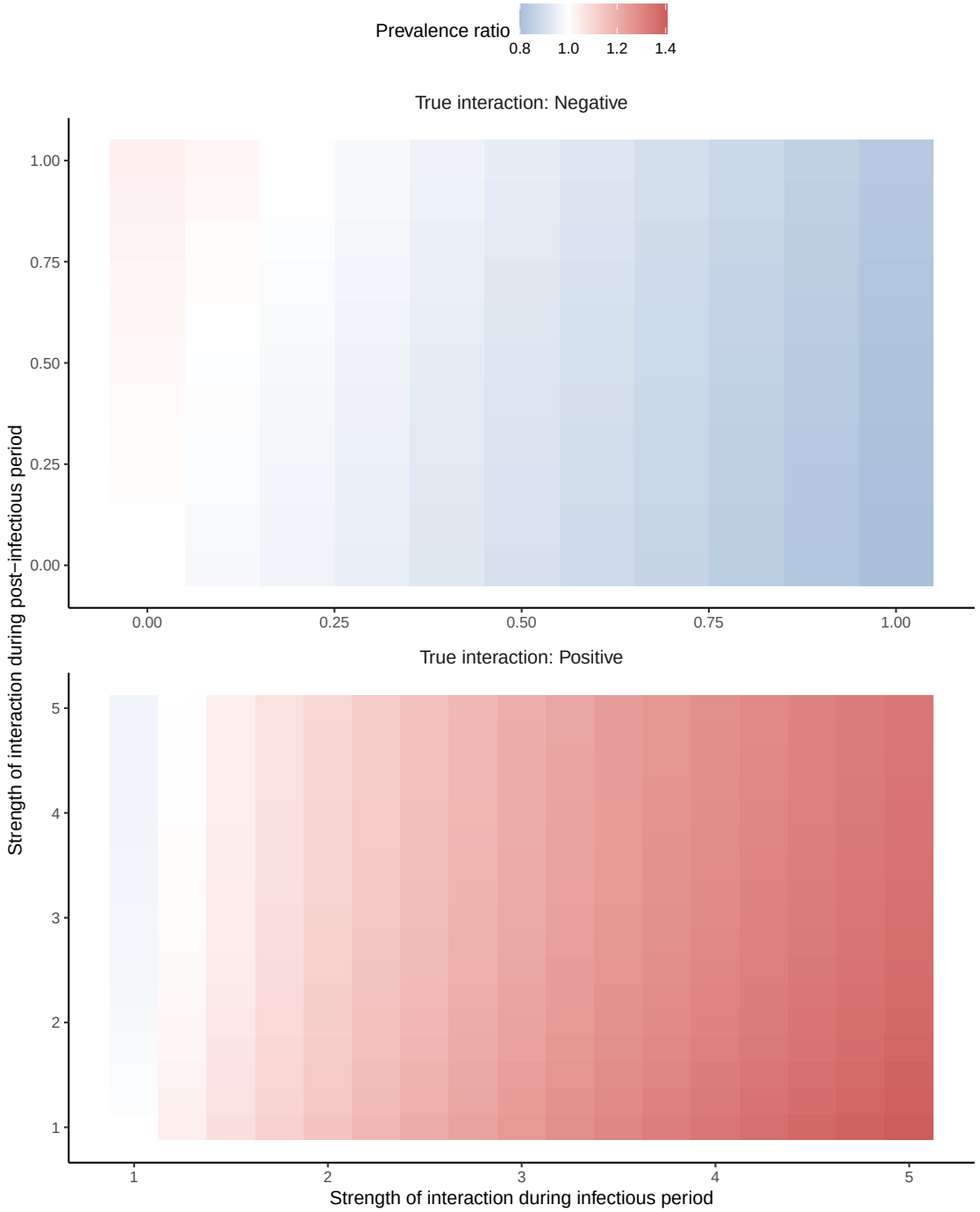

Figure S2: **Relationship between strength of interaction and prevalence ratio for non-symmetric, non-uniform interactions between influenza and SARS-CoV-2**. The scenarios tested correspond to $\theta_1^{(I)} = \theta^{(I)}, \theta_2^{(I)} = 1$ and $\theta_1^{(T)} = \theta^{(T)}, \theta_2^{(T)} = 1$ (that is, no effect of SARS-CoV-2 on influenza); other parameters were fixed to model the coupled dynamics of influenza and SARS-CoV-2 (cf. Table 1). For negative interactions (top panel), the x-axis represents $1-\theta_1^{(I)}$ and the y-axis $1-\theta_1^{(T)}$; for positive interactions (bottom panel) $\theta_1^{(I)}$ and $\theta_1^{(T)}$. Hence, in either panel the true strength of interaction increases from left to right and from bottom to top.