Spatial early warning signals of social and epidemiological tipping points in a coupled behaviour-disease network

Brendon Phillips1, Madhur Anand2 & Chris T. Bauch1

The resurgence of infectious diseases due to vaccine refusal has highlighted the role of interactions between disease dynamics and the spread of vaccine opinion on social networks. Shifts between disease elimination and outbreak regimes often occur through tipping points. It is known that tipping points can be predicted by early warning signals (EWS) based on characteristic dynamics near the critical transition, but the study of EWS in coupled behaviour-disease networks has received little attention. Here, we test several EWS indicators measuring spatial coherence and autocorrelation for their ability to predict a critical transition corresponding to disease outbreaks and vaccine refusal in a multiplex network model. The model couples paediatric infectious disease spread through a contact network to binary opinion dynamics of vaccine opinion on a social network. Through change point detection, we find that mutual information and join count indicators provided the best EWS. We also show the paediatric infectious disease natural history generates a discrepancy between population-level vaccine opinions and vaccine immunity status, such that transitions in the social network may occur before epidemiological transitions. These results suggest that monitoring social media for EWS of paediatric infectious disease outbreaks using these spatial indicators could be successful.

Resurgences of vaccine-preventable diseases severely stress public health systems, interrupt tourism and public services, and disrupt economies through the huge costs of large-scale interventions1. These impacts motivate the study of factors that support vaccine uptake. Undervaccination may be attributed to vaccine refusal2, the spread of anti-vaccine opinion facilitated by media coverage and its sensationalisation of true adverse vaccine effects3, the expectation of adverse effects4, misstatement of the cause of illnesses5, the spread of other rumours and false information6, and the effect of social norms7.

These phenomena show how the social diffusion of information is heavily responsible for the trajectory of disease spread through its ability to alter individual behaviour. Much work has modelled opinion dynamics for different applications through the use of voter models8 and majority opinion models9, among other frameworks, and their combination with network structure has revealed much about the occurrence of opinion cascades10 and forecasting11. For instance, models coupling behavioural dynamics and spreading processes change the predicted dynamics of influenza transmission14 and climate change15 alike.

Opinion propagation can be represented by information diffusion through social networks14. Similarly, infection spread is often conceptualized as spreading through a physical contact network15. Many models explore the dynamics of n-layer multiplex networks, where each layer represents a different aspect of the dynamics of a single coupled system16–18. In these cases, the theory of phase transitions in spatially structured systems is important. For instance, epidemic regimes have previously been modelled as the outcome of phase transitions in physical systems and are featured widely in the epidemiology literature20–26. Generally, phase transitions occur when a physical system moves from one state to another after going through some critical point. Second-order transitions occur when a macroscopic variable varies continuously27 and are called critical transitions28.

Systems approaching these critical transitions sometimes display characteristic spatial or temporal behaviours called early warning signals (EWS) that can predict coming epidemic disease outbreaks and other events29. EWS are perhaps better defined as statistically significant, recognisable and characteristic behaviours known to precede
critical transition in dynamical systems\(^{30,31}\). For instance, critical slowing down can precede both first- and second-order transitions\(^{31,32}\) and is accompanied by the divergence of correlation length in a physical system\(^3\).

Many statistics have been used to study EWS in spatially extended systems; temporal\(^34,35\) and spatial correlation\(^36,37\) have been found to precede transitions in spreading processes. Other measurements have been applied to spin systems, where each site in a lattice may be in one of two possible states, possibly partially dependent on the state of the neighbouring sites. The spin model has also been applied to opinion dynamics; a simple voter model with binary opinion dynamics is analogous to a physical spin system, where particles represent agents and spins represent different opinions\(^38\). Consensus formation can be seen as a second-order phase transition to an ordered state (where all spins are aligned). In this regime, the opinion of a single agent predicts the opinion of all other agents in the system\(^39\). Since the transition in finite networks is smooth\(^40\), the distance across the network over which the opinions of connected agents are strongly correlated increases smoothly; this is analogous to divergence of the correlation length of a physical system\(^41\).

Above some critical temperature, disordered systems take the form of a spin glass. In a spatial opinion model, this describes a state where opinions between neighbours are generally uncorrelated\(^42\). On a static network, this state induces a larger number of edges between dissimilar neighbours as compared to that of consensus regimes. This is related to join count statistics, where the numbers of edges between like neighbours are compared to the number between dislike neighbours as a test of geographical distribution. This is arguably the most natural and well-defined measure for graphs presenting binary data and is used for spatial analysis\(^43\).

The necessity of disease surveillance and early warning signals for outbreaks has been discussed in multiple contexts, from epidemic mitigation to bioterrorism prevention\(^44–46\). Potential mitigation of unnecessary expense motivates us to find reliable EWS that remain easily computable on large high-resolution data sets. Furthermore, the study of EWS in coupled disease-behaviour multiplex networks has received relatively little attention, suggesting a significant gap in the literature. Our objective is to evaluate and compare the relative merits of the mutual information, Moran’s I, Geary’s C and join count statistics as EWS of the occurrence of epidemics and changes in aggregate opinion on a coupled disease-behaviour network model. We use three differently parametrised models (V1, V2 and V3) coupling a binary vaccination opinion dynamic to an SIRV epidemic process. The resulting trends in the EWS for model V2 will be explored in the Results and Discussion sections, with V1 and V3 presented in Supplementary Information S5.

The outline of this paper is as follows: the Methods section will present the EWS and their derivations and give the details of the model used. The Results section will analyse the trends in the warning signals and the Discussion section will present a review of the study and any shortcomings of our approach, with further results pertinent to the study presented in the Supplementary Information.

**Methods**

We assume an acute, self-limiting infection that confers lifelong natural immunity upon recovery, and for which a vaccine is readily available. Similar premises have been used to represent the natural history of many paediatric infectious diseases such as measles\(^47\). In particular, we assume an SIRV\(_p\) natural history consisting of four mutually exclusive disease states. Agents are initially susceptible to infection (S). Upon infection the agent enters the infected state (S → I), which we treat as a combination of both the latent, ill and infectious periods\(^48\). Upon clearing the infection, agents enter the recovered state of lifelong immunity (I → R); additionally, susceptible agents may be vaccinated and so enter the vaccinated state (S → V\(_p\))\(^49\).

We also include injunctive social norms (i.e. peer pressure) as well as a perceived cost of vaccination that captures both economic costs and the fear of perceived adverse vaccine effects\(^50\). As in some models\(^36\), we include a noise parameter \(\kappa\) to account for environmental and demographic stochasticity\(^51\) with the simplifying assumption of perfect vaccination\(^52\) (reversion from the recovered state to the susceptible only through agent death). During simulation, the susceptible time step represents a single week.

The model. We model a multilayer network where each layer is given an identical undirected Erdős-Rényi random graph with size \(N\) and mean node degree \(\langle q \rangle\). Each agent \(n\) can be described by a pair of states; for instance, each agent is assigned the joint state \((V_n, V')\) at the start of the simulation with probability \(\alpha\) (they are a pro-vaccine vaccinated agent), else they are initialised with joint state \((N, S)\) (an anti-vaccine susceptible agent) with probability \(1 - \alpha\).

The social process follows an \(NV\) dynamic (Fig. 1b), representing pro- \((V)\) and anti-vaccine \((N)\) opinion for each agent \(n\). \(\xi\) represents the probability of any agent switching opinion randomly in each week and \(P_n(V \rightarrow V')\) represents the probability of switching from anti-vaccine opinion to pro-vaccine opinion \((N \rightarrow V)\) upon interaction with a disagreeing neighbour. We introduce an imitation dynamic by having each agent \(n\) compare its opinion with a single randomly chosen social contact (a neighbouring agent on the social layer) each week; \(n\) then changes its vaccination opinion only if there is disagreement (the agent and the neighbour have different vaccine opinions). This change of opinion depends on the perceived risk of vaccine adverse effects \(\kappa\) (“vaccine risk”) and \(I_n\) (the number of infected physical neighbours of \(n\)) according to the rules

\[
\begin{align*}
P_n(N \rightarrow V) & = \frac{1}{1 + \exp(-\Delta U_{n}^{N \rightarrow V})}, \\
P_n(V \rightarrow N) & = \frac{1}{1 + \exp(-\Delta U_{n}^{V \rightarrow N})}, \\
\end{align*}
\]

(1)

where the indices \(\Delta U_{n}^{N \rightarrow V}\) and \(\Delta U_{n}^{V \rightarrow N}\) in Eq. (1) are utility functions defined as
\[ \Delta U_n^{N \rightarrow V_s} = -\sigma \frac{Q_n^N - Q_n^{V_s}}{Q_n} - (\kappa - I_n), \quad \Delta U_n^{V_s \rightarrow N} = -\sigma \frac{Q_n^{V_s} - Q_n^N}{Q_n} + (\kappa - I_n). \]

Here, \( \sigma \) represents the strength of an injunctive social norm to maintain the current opinion (referred to as "social norm"), while \( Q_n^a \) represents the number of neighbours of \( n \) with opinion \( a \) and \( Q_n \) represents the total number of social contacts of \( n \).

The epidemiological dynamics (which we term ‘physical dynamics’) follow an SIRV\( p \) process (Fig. 1a), in which an agent \( n \) can progress through each of four disease compartments: \( S \) (susceptible), \( I \) (infected), \( R \) (recovered) and \( V_p \) (vaccinated). Every week (i.e. in each time step), each susceptible agent \( n \) interacts with all its physical neighbours; each effective interaction carries the probability \( p \) of infection (\( S \rightarrow I \)). Upon death (with probability \( \mu \) per week), an agent is "rebirthed" with either vaccinated (probability \( \alpha \cdot \mu \)) or susceptible (probability \( (1 - \alpha) \cdot \mu \)) status. (b) Per time step, each agent switches between pro- \( V_s \) and anti-vaccine \( N \) opinion with probabilities \( P_n(N \rightarrow V_s) \) and \( P_n(V_s \rightarrow N) \) respectively upon interaction with a dissenting neighbour. \( \alpha \) gives the probability of being birthed with pro-vaccine opinion \( V_s \).

\[ P_n(N \rightarrow V_s) + \alpha \cdot \mu + \xi \]

\[ P_n(V_s \rightarrow N) + (1 - \alpha) \cdot \mu + \xi \]

Figure 1. Representation of the infection (a) and opinion (b) dynamics of the model occurring on the physical and social layers of the 2-layer network, respectively. (a) Effective contacts occur between susceptible \( S \) and infected \( I \) agents with probability \( p \) per time step (1 week). Upon deciding to vaccinate (with probability \( P_n(N \rightarrow V_s) \)), a susceptible agent \( n \) becomes physically vaccinated (\( S \rightarrow V_p \)). Infection lasts \( \ell = 2 \) weeks after which agents recover (\( I \rightarrow R \)). Upon death (with probability \( \mu \) per week), an agent is "rebirthed" with either vaccinated (probability \( \alpha \cdot \mu \)) or susceptible (probability \( (1 - \alpha) \cdot \mu \)) status. (b) Per time step, each agent switches between pro- \( V_s \) and anti-vaccine \( N \) opinion with probabilities \( P_n(N \rightarrow V_s) \) and \( P_n(V_s \rightarrow N) \) respectively upon interaction with a dissenting neighbour. \( \alpha \) gives the probability of being birthed with pro-vaccine opinion \( V_s \).
where $X$ and $Y$ are discrete random variables; $x$ takes value on the set $\mathcal{X} = \{x_1, x_2, x_3, \ldots\}$ and $y$ on set $\mathcal{Y} = \{y_1, y_2, y_3, \ldots\}$, with $\mathbb{P}$ a joint probability mass function of $X$ and $Y$. Mutual information is an entropy-based quantification of the "shared information" of two random variables quantifying how knowledge of one decreases the uncertainty of the other and vice versa. Mutual information peaks at the critical temperature of spin systems during second-order transitions and has been widely used in detecting phase transitions of spin systems; an advantage of this statistic is its ability to quantify non-linear dependence, unlike Moran’s I and covariance which only account for linear dependence.

**Join counts** quantify the degree of clustering by giving the number of adjacencies between agents of different types. We divide the population into two attributive classes, with $V$ the compartment of pro-vaccine agents and $N$ the compartment of anti-vaccine agents. Let $[\Psi, \Omega]$ be the number of social interactions between agents with vaccine opinions $\Psi$ and $\Omega$; then $[N, N]$ represents the number of nearest-neighbour interactions between anti-vaccine agents, $[V, V]$ the number of interactions between pro-vaccine agents and $[N, V]$ the number of interactions between pro- and anti-vaccine agents. These can be written as

$$
[N, V] = \frac{1}{2} \sum_{i,j,k} \omega_{i,j,k}(x_j - x_k)^2, \quad [N, N] = \frac{1}{2} \sum_{i,j,k} \omega_{i,j,k}(1 - x_j)(1 - x_k), \quad [V, V] = \frac{1}{2} \sum_{i,j,k} \omega_{i,j,k} x_j x_k.
$$

where $\omega_{i,j,k} = 1$ if agents $i$ and $k$ are social neighbours, with $\omega_{i,j,k} = 0$ otherwise ($\omega$ is the adjacency matrix of the social network); $x_n$ represents the opinion score of agent $n$, defined as

$$
x_n = \begin{cases} 1 & k \in V \quad (x_n \text{ has a pro-vaccine opinion}) \\ 0 & \text{else} \end{cases}
$$

In an opinion model, clustering manifests as agents consistently having a higher number of like-minded neighbours than expected based on the global prevalence of the opinion; join counts are then used to test the null hypothesis of positive correlation. Join counts are used in many fields as a categorical test of spatial autocorrelation, including ecology and geographical information systems. In all parameter realisations here, the number of joins are counted naïvely rather than calculated. Joins between like-minded agents (e.g. $[N, N]$ and $[V, V]$) will be called similar joins, and edges between disagreeing neighbours (e.g. $[N, V]$) will be called dissimilar joins.

The Moran’s I coefficient $I$ quantifies spatial correlation and is defined

$$
I = \frac{N}{W} \cdot \frac{\sum_i \omega_{i,j,k}(x_j - x)(x_k - x)}{\sum_j (x_j - x)^2},
$$

where $W = \sum_{i,j,k} \omega_{i,j,k}$ and $x = \frac{1}{N} \sum_j x_j$ represents the mean opinion score of the population. Used as a global statistic, Moran’s I gives the degree of correlation between the values of neighbouring patches (agents and their social contacts); here, the numerical value of the vaccine opinion is the same as described in Eq. (5). Algebraic manipulation of Eq. (6) using Eq. (5) gives

$$
I = \frac{N}{W} \cdot \frac{2 \cdot [V, V] - 2x \cdot (2 \cdot [V, V] + [N, V]) + W \cdot x^2}{(1 - 2x) \cdot [V] + N \cdot x^2},
$$

(full derivation given in Supplementary Information S4); we can then consider Moran’s I as a measure derived from the linear combination of join counts. Positive values signify spatial correlation, with negative values signifying anticorrelation.

The Geary’s C coefficient $C$ is yet another measure of spatial correlation based on the cross-product (like Moran’s I), but unlike Moran’s I it accounts for the difference in opinion between two neighbours. It is given as

$$
C = \frac{N - 1}{W} \cdot \frac{\sum_i \omega_{i,j,k}(x_j - x_k)^2}{\sum_i (x_j - x)^2}.
$$

Lower values show spatial correlation, and large values represent anticorrelation. Like Moran’s I (Eq. (7)), Geary’s C can also be expressed as a linear combination of join counts

$$
C = \frac{N - 1}{W} \cdot \frac{2 \cdot [N, V]}{(1 - 2x) \cdot [V] + N \cdot x^2};
$$

this expression is derived in Supplementary Information S4.

**Parametrisation.** The birth/death rate in the model was set at $\mu = 2.4 \times 10^{-4}$, giving each agent a mean life expectancy of 80 years. The network size $N = 40000$ was chosen to represent a small town where each agent $n$ has effective physical contact with $(Q_{un}) = 30$ neighbours per week, where an effective contact is defined as any interaction between agents that allows for infection and/or the communication of opinion. The case importation
proportion ratio $\epsilon = 2.5 \times 10^{-5}$ was added to provide periodic impulses of infection as a test of resilience in endemic disease regimes. Here, an ensemble of 100 simulations using parameters $\kappa = 0$, $\sigma = 0$ and $\alpha = 0.05$ returned the values $(S)$, $(R) < 0.05$ at equilibrium (defined in Supplementary Information S3), where $(\Psi)$ represents the mean number of agents with (social or physical) state $\Psi$, averaged over all realisations of that combination of parameter values.

The infectivity $p = 0.2$ was chosen to reflect the reproduction rate of a measles infection commonly estimated from empirical data$^4$; effective contacts occur in the simulation once per week during the period of infection. The probability of randomly switching opinion $\zeta = 1 \times 10^{-4}$ was included as a source of noise. We found that the parameter ranges for vaccine risk $\kappa \in [-1, 1]$ and social norm $\sigma \in [0, 3]$ were sufficiently broad to capture transitions in both social and physical dynamics (Fig. 2a,b), as well as the corresponding trends in the mutual information (Fig. 2c) and dissimilar join count (Fig. 2d).

The contours in each panel of Fig. 2 show the obvious correspondence between transitions in the social and physical dynamics of the model, and substantial changes in $\{\Psi\}$ and $\{N, V_n\}$ (Fig. 2d) and mutual information $\mathcal{I}$ (Fig. 2c) increase while the vaccine risk $\kappa$ and social norm $\sigma$ parameters increase towards their respective (pre-transition) threshold values. These trends are generally asymmetric about both transitions; this can be seen in Fig. 4, where post-transition trends do not exhibit similarly detectable warnings (if any). This Parametrisation applies to model V2. The corresponding parameterisations and contour plots of models V1 and V3, as well as their post-transition trends are presented through the Supplementary Information.

Results
Due to the low initial vaccine coverage $\alpha = 0.05$, all realisations demonstrated an initial epidemic spread (defined in Supplementary Information S3) over the first 7 weeks. After this period, the dynamics settled down to a quasi-equilibrium state characterized by fluctuations around a mean value that is the focus of our study – the following subsections are grouped by major findings of the model. The term model variables refer to the outputs $\{S\}, \{I\}, \{R\}, \{V_n\}, \{N\}$ and $\{V\}$.

Population vaccine immunity status can differ from aggregate vaccine opinion. Because only susceptible individuals are vaccinated and individuals cannot become ‘unvaccinated’ (but may change their opinion about the vaccine over their lifetime), the population-averaged vaccine opinion is not equal to the population-averaged vaccine immunity status, even at the quasi-equilibrium state. With no social pressure ($\sigma = 0$), a small increase in vaccine risk $\kappa \to 0.03125$ pushes the system to endemic infection and anti-vaccine consensus (Fig. 3a) despite a high vaccination rate (Fig. 3d). Towards an explanation, if an agent is newly birthed into this regime, the probability of having an infected neighbour vanishes $(I_n \to 0)$, so that

$$P_a(N \to V_i) = P_a(V_i \to N) \approx \frac{1}{2},$$

similar to Eq. (1), with the agents’ probability of being vaccinated over their lifetime as

$$0.05 + 0.95 \sum_{m=1}^{80} \frac{1}{2} \left(1 - \frac{1}{2}\right)^{m-1} \approx 1,$$

under the assumptions that the average agent with anti-vaccine opinion is almost certain to interact with a disagreeing contact. A similar calculation explains the phenomenon of high vaccination rate (Fig. 3e) despite mixed consensus (Fig. 3b) when vaccine risk becomes neutral ($\kappa \to 0$). A vaccine then perceived as beneficial ($\kappa \to -0.03125$) intuitively results in a high vaccination rate (Fig. 3f) and pro-vaccine consensus (Fig. 3c); in both these regimes, the disease survives only through case importation. In the absence of social norms and vaccine risk, the population’s aggregate vaccine opinion may not be a good indicator of its vaccine immunity profile (and vice versa); Eq. (2) shows that the probability of changing opinion depends only on $I_n$ when social norm $\sigma = 0$. In this region, the pattern of disease spread will be determined by the initial conditions of the physical dynamics; slight changes in vaccine risk $\kappa$ will push the network towards either of the consensuses, with minimal
effect on the high vaccination rate (Fig. 3d). This phenomenon is shared by models V1 and V3, as shown in Supplementary Information S5.1.

### EWS trends identify approaching transitions in both social and physical layers

The trends in both model dynamics and the proposed EWS are shown in Fig. 4, with social norm $\sigma = 0$ (column A) and $\sigma = 0.25$ (column B). The first vertical black line in all panels of Fig. 4 represents the transition in the social dynamics $K_s$, defined as the smallest $\kappa$ value at which $\langle V \rangle \approx \langle N \rangle$ (the mean number of pro-vaccine agents equals the number of anti-vaccine agents); the second vertical black line represents the transition in the physical dynamics $K_p$, similarly defined as the earliest $\kappa$ where $\langle R \rangle \approx \langle V \rangle$. Multiple physical and social transitions were found for some parameter combinations; these trends and the attendant behaviour of the EWS for models V1 and V3 can be seen in Supplementary Information S5.3. We also note that the equilibrium values of EWS and model variables were averaged over $15 \to 20$ realisations of all parameter combinations.

All EWS show recognisable trends preceding both transitions for both social norm values $\sigma = 0$ (column A) and $\sigma = 0.25$ (column B); for instance $(N, N)$ (Fig. 4c,d), $(N, V)$ (Fig. 4e,f), $(\langle \mathcal{A} \rangle)$ (Fig. 4i,j) and $(\langle \mathcal{E} \rangle)$ (Fig. 4k,l) increase sharply preceding $K_s$, with all but $(N, V)$ (Fig. 4e,f) approaching some maximum value preceding $K_p$, while $(V, V)$ (Fig. 4c,d) and $(\langle \mathcal{A} \rangle)$ (Fig. 4g,h) sharply decrease and approach some minimum value before $K_s$ and $K_p$ respectively. Mutual information $(\langle \mathcal{A} \rangle)$ (Fig. 4g,h) in particular shows clear changes in trend well before the social transition $K_s$ occurs. Though $(N, V)$, $\langle V \rangle$ (Fig. 4i,j) shows a similar rising-falling pattern for both $\sigma = 0$, 0.25, its maximum value with $\sigma = 0.25$ (Fig. 4f) is much lower than that for $\sigma = 0$ (Fig. 4e). For $\sigma = 0.25$, the mean of the Geary’s C $(\langle \mathcal{E} \rangle)$ (Fig. 4l) shows almost no change, though its envelope broadens post-transition; we see this as a failure of the EWS (no forewarning given). Similar observations hold for model V1 (Supplementary Figs. S15 and S16, where skewness $\gamma_1$ is used to quantify asymmetry of the trend of each EWS).

We can then say that all proposed EWS other than $(\langle \mathcal{E} \rangle)$ give appreciable signals approaching $K_s$ and $K_p$ when $\sigma = 0.25$ (Fig. 4B). $K_p$ precedes $K_s$ (Fig. 4a,b), showing that a shift in consensus will always precede a crisis in vaccination coverage in this model. Also shown is a marked decrease in $K_p - K_s$ (the gap between the two transitions $K_s$ and $K_p$, which we call the intertransition distance) as the social norm strengthens (for example, $\sigma \to 0.25$ in column (B) of Fig. 4). The generalisation of these trends to all tested values of $\sigma$ is confirmed in Fig. 5a, where $K_p - K_s$ is everywhere positive, though the distance between $K_s$ and $K_p$ vanishes with increasing $\sigma$; the inset of Fig. 5a shows the location of $K_s$ (blue) and $K_p$ (red) with respect to $\kappa$, so that $K_p - K_s$ (purple) gives the width of the area between the two curves in the inset graph at each $\sigma$. Other disparate models of the disease display largely similar concave decreases in the intertransition distance, suggesting that this behaviour arises generally from the model dynamics rather than in some specific subspace of the parameter space (see Supplementary Information S5.2).

### Stronger social norms result in decreased lead time for all EWS

The findings of the preceding subsection are intuitive, as vaccination depends more heavily on individual vaccination opinion than the number of infected neighbouring agents in Eq. (2), so that the opinion dynamics exert more influence than any feedback effect occurring in the physical dynamics. However, the vanishing intertransition distance $K_p - K_s$ presents a problem if we depend on predictions of $K_s$ to enact interventions avoiding the collapse of the system to a non-vaccinated regime (i.e. avoiding $K_p$). For social norms of increasing strength, we can therefore look at the
trend in the lead time $B_t = K_s - \Pi_q$, where $\Pi_q$ represents some $\kappa$ value at which we can assert that a signal occurs in some sequence $\Psi$ of $\kappa$ values; since we’ve established that $K_s$ precedes $K_p$ everywhere, then necessarily any warning of a social transition also warns of the following physical transition, so the quantity $-\Pi_\Psi K_p$ is not discussed here.

One way for us to quantify this lead time is to use a change point detection tool to find $\kappa$ values at which the two classes EWS (subscript WS) and model variables (subscript S/P) give signals (i.e. display statistically significant changes in trend/behaviour). Here, the Lanzante test$^{62}$ from the trend$^{63}$ package in R is applied to various sequences of equilibrium $\kappa$ values to find change points of EWS ($B_{WS}^{Lan}$, Fig. 5b) and model variables ($B_{S/P}^{Lan}$, Fig. 5c) respectively. (In other words, we computed the change test on the mean of all stochastic realisations at each parameter value combination, rather than computing the change test on each individual time series and then taking the average.) Further discussion of the method of application of this and other change point detection tests to series of $\kappa$ values can be found in Supplementary Information S5.2.

Failure of a warning signal or model variable $\Psi$ occurs when the warning given comes after the social transition, so that $\Psi > K_{B_s}$. Looking at the trend in the lead times $B_{WS}^{Lan}$ (Fig. 5b), the positivity of some curves shows that some of the proposed EWS do indeed give early warnings of coming transitions (largely for $\sigma < 1.875$). Failures of $(N', V')$ and $(N, V)$ occur in the range $1.875 < \sigma \leq 2.125$, while all other tests give valid warning signals everywhere $\sigma < 2.5$; model variables $(R)$ and $(V_p)$ fail in the range $1.65 \leq \sigma \leq 2.125$ (Fig. 5c). The failure of all the tests after $\sigma = 2.5$ likely results from insufficient length of the EWS’ $\kappa$-series; the inset of Fig. 5a shows that $K_s \to -1$ as $\sigma$ increases.

Figure 4. Trends of the EWS’ equilibrium values approaching the transitions of the social and physical dynamics $K_s$ and $K_p$ (marked in order by the first and second vertical black lines) respectively, demonstrating the signals given by each tool with respect to the perceived vaccine cost $\kappa$. The intervals in each panel represent one standard deviation of the mean equilibrium value in each stochastic realisation of the model. Social norm $\sigma = 0$ for column (A), and $\sigma = 0.25$ for column (B). (a,b) Social dynamics $(V_s)$ (green, solid), $(N)$ (red, solid) and physical dynamics $(R)$ (black, dashed), $(V_p)$ (blue, dashed). (c,d) Join counts: $(N, N)$ (blue), $(N, V_s)$ (red), $(V_s, V_p)$ (green). (e,f) Dissimilar join count $(N', V_s)$ alone (red). (g,h) Mutual information $(\mathcal{I})$ (red). (i,j) Moran’s $I$ $(\mathcal{I})$ (red). (k,l) Geary’s $C$ $(\mathcal{E})$ (red).
Figure 5c is largely similar, showing failure of all the signals around $\sigma = 2.5$ (as in Fig. 5b). In line with our focus on social dynamics as a predictor, $V_s$ appears to be the best performing signal of all the model variables; as was reasonably expected, $I$ appears to perform badly, since its role as a transitory compartment in the disease dynamics means that it never “gathers” sufficient agents over the course of each realisation to give a true indication of the state of the system (other than indicating the presence or absence of endemic infection).

Since the perceived cost of vaccination $\kappa \geq -1$ in this study, our method of detecting the change point will not accurately predict a change point $\Pi, \Psi$ close to $-1$. Since no one warning signal gives the highest lead time for any large contiguous range of $\sigma$ values, there is unfortunately no single objective way to choose a “strongest” signal; they are all suitable tools to predict coming crises in aggregate opinion and vaccination dynamics. However, it is worth noting that mutual information $M$ and the dissimilar join count $NV_s$, $s$ perform better than the other indicators; Fig. 5b shows that mutual information $M$ gave the largest lead time measured for 45% of tested social norm $\sigma$ values, and $V_s$, $V_s$ and $NV$, $s$ both give the largest lead time for 32% $\sigma$ values (multiple EWS showed an identical lead time for some values of $\sigma$). Lead times of all EWS for all three models under other various change point tests are discussed in Supplementary Information S5.4.

EWS can provide better forewarning than trends in model variables. One final question is whether the proposed EWS (mutual information, Moran’s $I$, Geary’s $C$, join counts) give earlier warnings than simply monitoring trends in model variables (such as the number of infections or pro-vaccine agents, using a change point test for prediction in both cases). There are many ways to quantify this, including maximin comparison (finding the larger of minimum values of classes WS and S/P) and maximax comparison (finding the larger of the maxima of each class) at each value of the social norm $\sigma$.

To compare the minima of the EWS and model variable lead times, we define $\chi_{\text{min}}$ as

$$\chi_{\text{min}} = \min(B_{\text{WS}}^{\text{Lan}}) - \min(B_{\text{S/P}}^{\text{Lan}}),$$

and we specify a tolerance $\varepsilon$ to be 1% of the total range of $\chi_{\text{min}}$

$$\varepsilon = \frac{\max(\chi_{\text{min}}) - \min(\chi_{\text{min}})}{100}.$$

Hence, if $\chi_{\text{min}} > \varepsilon_{\text{min}}$, then the EWS is outperforming simple monitoring of trends (model variables).

The blue curve in Fig. 6 compares the minima of the EWS (WS) and model variable (S/P) classes (maximin comparison), showing the $\sigma$ values for which the worst-performing (least lead time) EWS $\min(B_{\text{WS}}^{\text{Lan}})$ is either
better, equal or worse than the worst-performing model variable \( \min(B_{S/P}) \). Points in the green-shaded region represent \( \sigma \) values where the EWS' performance is at worst still better than that of the model variables.

The EWS outperformed simple monitoring of trends in variables for 45.4% of the tested \( \sigma \) values; maximin comparison shows that EWS are at worst still better than model variables for a large number of \( \sigma \) values, with the two classes performing equally badly in 18.2% of the \( \sigma \) values. Performance of the EWS and the model variables in this test was considered comparable or equal if the difference between the two minimum lead times fell under the tolerance \( \varepsilon_{\min} \), so that \( |\chi_{\min} - \chi_{\min}^{\max}| \), performance was equal for 18.2% of tested \( \sigma \) values, showing that the added computation of the EWS does not always yield a benefit. Otherwise, the points and portion of the blue curve falling in the red-shaded region of Fig. 6 represents values of \( \sigma \) where the model variables outperformed the EWS (that is, the minimum lead time of the model variables exceeded the minimum lead time of the EWS); this occurred for 36.4% of tested \( \sigma \) values.

The second part of the comparison (Fig. 6, red curve) is between the maxima of the lead times; as above, we define the comparison variable \( \chi_{\max}^{\max} \) and tolerance \( \varepsilon_{\max} \) as

\[
\chi_{\max}^{\max} = \max(B_{WS}) - \max(B_{S/P}) \quad \varepsilon_{\max} = \frac{\max(\chi_{\max}^{\max}) - \min(\chi_{\max}^{\max})}{100}.
\]

The green-shaded portion of Fig. 6 also shows the \( \sigma \) values where the EWS outperformed the model variables here, in that the maximum lead time given by the EWS exceed that given by the model variables \( \chi_{\max}^{\max} \); points falling within the red-shaded area of Fig. 6 show for which \( \sigma \) values the model variables outperform the EWS. From the second row of the inset table in Fig. 6, the two maxima are considered equal \( |\chi_{\max}^{\max} - \chi_{\max}^{\max}^{\max}| \) for 63.6% of tested \( \sigma \) values, while the EWS outperformed the model variables \( \chi_{\max}^{\max} > \chi_{\max}^{\max} \) for only 13.7% of \( \sigma \) values.

This shows that the EWS' lead times are at least equal to those of the model variables for around 63.6% of \( \sigma \) values and are absolutely larger for 77.3% of \( \sigma \) values, demonstrating that though monitoring the model variables (both social and physical) is itself valuable, the EWS offer better performance (using the Lanzante change point test). In both (blue and red) curves of Fig. 6, there is no apparent pattern to the positivity/negativity of \( \chi_{\max}^{\max} \) and \( \chi_{\max}^{\max} \). These comparisons are given for other tests and models in Supplementary Information S4.

Discussion

Here we studied a range of early warning signals for critical transitions in a two-layer coupled behaviour-disease model for paediatric infectious diseases. We compared the indicators to one another and the approach of simply monitoring trends in model variables. We found that the performance of the indicators was variable depending on model parameters, but the mutual information statistic and the dissimilar join count showed consistently high pre-transition lead times over various strengths of the social norm, many times giving the highest lead times of all the EWS. Through maximin and maximax comparisons, we found that using EWS provide more advance warning than simply monitoring trends in model variables in a clear majority of cases.

We note that join counts have the additional advantage of easy computability, since they require only counting pairs of a given type. This contrasts with other more computationally intensive indicators such as autocorrelation which require making decisions about whether to study lag-1 or higher order lags, as well as choosing parameter values governing computation of residuals. Moran's I was also shown to predict the approach of transitions, although perhaps this finding is trivial considering that it is a linear combination of similar and dissimilar join counts. Its predictive power was not as strong as many of the other indicators such as join count and mutual information, hence the added complexity of its calculation may not justify its use. Potential downfalls of the mutual information statistic include its computational complexity and the availability of a suitable data set pairing the personal health of each agent with their social activity.

We also showed that a population may have relatively high vaccine coverage despite a low pro-vaccine opinion. This discrepancy between social and physical dynamics is due to the paediatric infectious disease natural
history we assumed in the model. Unlike influenza, where revaccination must occur seasonally, an individual who receives a sufficient number of measles or chickenpox vaccine doses generally has lifelong immunity and therefore the opinion towards the vaccine can decline well before the level of vaccine immunity does. (Individuals can change their opinion but never become ‘unvaccinated.’) The implication of this asymmetry is that monitoring social media networks for changes in opinion using early warning signals like mutual information might provide advance warning of outbreak hot-spots.

The distance between the change point in the EWS indicators and the critical transition in the social dynamics decreases as the social norm grows stronger, as does the distance between the transitions in social and physical dynamics of the model. Given the relative scale of the social norm and vaccine risk parameter values used, stronger social norms decrease the time interval between birth and vaccination decision (the vaccination rate converges to its equilibrium value in fewer time steps than in other regimes); feedback between this and the disease incidence in the network (which affects the number of infected neighbours in each agent’s neighbourhood) alters the probability function controlling the vaccination decisions, effecting faster alignment of majority opinion and vaccination coverage.

This study only lays the foundation for the investigation of spatial EWS for such coupled-behaviour systems. There is much work to be done before they can be meaningfully applied to empirical data. For instance, we assumed that the network was static. This simplifying assumption could be relaxed in future work by using an evolving social dynamic in which agents are allowed to form or break links with new agents based on their node degree69 or vaccine opinion and associated social pressures59. The rate of interaction between agents was also assumed fixed in our model (relative to the speed of other dynamics, such as the birth/death interval and the length of illness). A valid extension of the model would be a variable rate of communication between agents, since the rate of communication has been shown to influence the rate and efficiency of opinion formation68. A further avenue of research would explore interventions to turn populations away from critical transitions. This could answer research questions such as: How far in advance must we act to prevent a social or physical critical transition, and does this change our interpretation of the EWS? If we assume that any of the EWS can be used for monitoring, how would this change in trend alter the reliability of the EWS?

Many researchers and public health bodies are drawing attention to global resurgences of vaccine-preventable illness, and speak to the vast efforts and multiple approaches taken to mitigating outbreaks. A few of these approaches have focused on human behaviour and opinion dynamics, either by directly tracking aggregate vaccine opinion, or monitoring alerts and media reports. Our work demonstrates the potential uses of early warning systems of critical transitions in preventative epidemiology. In particular, our work provides proof-of-concept for the idea of monitoring social networks for early warning signals of both social and epidemiological shifts, and also suggests several EWS indicators that might work well for this purpose.

Data availability

The datasets generated during and analysed during the current study are available from the corresponding author upon reasonable request.

Received: 19 November 2019; Accepted: 6 April 2020;
Published online: 06 May 2020

References

1. Western, U. N. D. G. U. & Africa, C. Socio-economic impact of ebola virus disease in west african countries: A call for national and regional containment, recovery and prevention. Tech. Rep., United Nations Development Group (UNDG) - Western and Central Africa (2015).
2. Phadke, V. K., Bednarzycz, R. A., Salmon, D. A. & Omer, S. B. Association between vaccine refusal and vaccine-preventable diseases in the united states: A review of measles and pertussis. JAMA 315, 1149–1158 (2016).
3. Hanley, S. J., Yoshioka, E., Ito, Y. & Kishi, R. HPV vaccination crisis in japan. The Lancet 385, 2571 (2015).
4. Smith, L. E., Weinman, J., Amlôt, R., Yiend, J. & Rubin, G. J. Parental expectation of side effects following vaccination is self-enhancing. JAMA Pediatr 170, 286–292 (2016).
5. Freed, L., Katz, S. L. & Clark, S. J. Safety of vaccinations: Miss America, the media, and public health. JAMA 276, 1869–1872 (1996).
6. Smith, T. C. Vaccine rejection and hesitancy: A review and call to action. Open Forum Infectious Diseases 4 (2017).
7. Sobo, E. J. Social cultivation of vaccine refusal and delay among Waldorf (Steiner) school parents. Medical Anthropology Quarterly 29, 381–399 (2015).
8. Das, A., Gollapudi, S. & Munagala, K. Modeling opinion dynamics in social networks. In Proceedings of the 7th ACM International Conference on Web Search and Data Mining, WSDM ’14, 403–412 (ACM, 2014).
9. Krapivsky, P. & Redner, S. Dynamics of majority rule in two-state interacting spin systems. Physical Review Letters 90, 238701 (2003).
10. Jalili, M. & Perc, M. Information cascades in complex networks. Journal of Complex Networks 5, 665–693 (2017).
11. De, A., Valera, I., Ganguly, N., Bhattacharya, S. & Gomez Rodriguez, M. Learning and forecasting opinion dynamics in social networks. In 30th Conference on Neural Information Processing Systems, 9 (2016).
12. Paul, M. I., Oredzo, M. & Broniatowski, D. Twitter improves influenza forecasting. PLOS Currents Outbreaks (2014).
13. Bury, T. M., Bauch, C. T. & Anand, M. Charting pathways to climate change mitigation in a coupled socio-climate model. PLOS Computational Biology 15, e1007000 (2019).
14. Collinson, S., Khan, K. & Heffernan, J. M. The effects of media reports on disease spread and important public health measurements. PLOS ONE 10, e0141423 (2015).
15. Keeling, M. J. & Eames, K. T. Networks and epidemic models. Journal of the Royal Society Interface 2, 295–307 (2005).
16. Bauch, C. T. & Galvani, A. P. Social factors in epidemiology. Science 342, 47–49 (2013).
17. Alvarez-Zuzek, L. G., La Rocca, C. E., Iglesias, J. R. & Braunstein, L. A. Epidemic spreading in multiplex networks influenced by opinion exchanges on vaccination. PLOS ONE 1–14 (2017).
18. Xia, C. et al. A new coupled disease-awareness spreading model with mass media on multiplex networks. Information Sciences 471, 185–200 (2019).
19. Zuzek, L. G. A., Buono, C. & Braunstein, L. A. Epidemic spreading and immunization strategy in multiplex networks. In Journal of Physics: Conference Series, vol. 640, 012007 (IOP Publishing, 2015).
20. Drake, J. M. et al. The statistics of epidemic transmission. PLOS Computational Biology 15, 1–14 (2019).
21. Scheffer, M. et al. Anticipating critical transitions. *Science* **338**, 344–348 (2012).
22. O’Regan, S. M. & Drake, J. M. Theory of early warning signals of disease emergence and leading indicators of elimination. *Theoretical Ecology* **6**, 333–357 (2013).
23. O’Regan, S. M., Lillie, J. W. & Drake, J. M. Leading indicators of mosquito-borne disease elimination. *Theoretical Ecology* **9**, 269–286 (2016).
24. E. A., Beccaria, M. & Macorini, G. Critical behavior in a stochastic model of vector mediated epidemics. *Scientific Reports* **6** (2016).
25. Doncaster, C. P. et al. Early warning of critical transitions in biodiversity from compositional disorder. *Ecological Society of America* **97** (2016).
26. Kuehn, C. A mathematical framework for critical transitions: Normal forms, variance and applications. *Journal of Nonlinear Science* **23**, 457–510 (2013).
27. Dunweg, B., Landau, D. P. & Milchev, A. I. *Computer simulations of surfaces and interfaces*, vol. 114 (Springer Science & Business Media, 2013).
28. O’Dea, E. B., Park, A. W. & Drake, J. M. Estimating the distance to an epidemic threshold. *Journal of the Royal Society Interface* **15**, 1–10 (2018).
29. Kuehn, C. A mathematical framework for critical transitions: Bifurcations, fast-slow systems and stochastic dynamics. *Physica D: Nonlinear Phenomena* **240**, 1620–1635 (2011).
30. Brett, T. S., Drake, J. M. & Rohani, P. Anticipating the emergence of infectious diseases. *Journal of the Royal Society Interface* **14**, 20170115 (2017).
31. Scheffer, M. et al. Early-warning signals for critical transitions. *Nature* **461** (2009).
32. Strogatz, S. H. *Nonlinear Dynamics and Chaos: With Application To Physics, Biology, Chemistry, And Engineering* 2 edn (Westview Press, 2014).
33. Das, M. & Green, J. R. Critical fluctuations and slowing down of chaos. *Nature communications* **10**, 2155 (2019).
34. Diks, C., Hommes, C. & Wang, J. Critical slowing down as an early warning signal for financial crises. *Empirical Economics* (2018).
35. Donangelo, P., Fort, H., Dallos, V., Scheffer, M. & van Nes, E. H. Early warnings for catastrophic shifts in ecosystems: Comparison between spatial and temporal indicators. *International Journal of Bifurcation and Chaos* **20**, 315–321 (2010).
36. Jentsch, P. C., Anand, M. & Bauch, C. T. Spatial autocorrelation as an early warning signal of regime shifts in a multiplex disease-model behaviour network. *Journal of Theoretical Biology* **448**, 17–25 (2018).
37. Orozco-Fuentes, S. et al. Early warning signals in plant disease outbreaks. *Ecological Modelling* **393**, 12–19 (2019).
38. Barabási, L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509–512 (1999).
39. Mukherjee, S. & Chatterjee, A. Disorder-induced phase transition in an opinion dynamics model: Results in two and three dimensions. *Physical Review E* **94**, 1–5 (2016).
40. Binder, K. Finite size effects on phase transitions. *Ferroelectrics* **73**, 43–67 (1987).
41. Zhang, W., Lim, C. C., Korniss, G. & Szymanski, B. K. Opinion dynamics and influencing on random geometric graphs. *Scientific Reports* **4**, 1–9 (2014).
42. Vincent, E. & Dupuis, V. *Spin glasses: experimental signatures and salient outcomes*, vol. 275 of *Springer Series in Materials Science* (Springer Science & Business Media, 2009).
43. Jentsch, P. C., Anand, M. & Bauch, C. T. Spatial autocorrelation as an early warning signal of regime shifts in a multiplex disease-model behaviour network. *Journal of Theoretical Biology* **448**, 17–25 (2018).
44. Orozco-Fuentes, S. et al. Early warning signals in plant disease outbreaks. *Ecological Modelling* **393**, 12–19 (2019).
45. Murray, J. & Cohen, A. L. *Infectious Disease Surveillance in China*, chap. 2, 23–33 (Elsevier, 2017).
46. Specco, A. & Timpka, T. Algorithms for detecting and predicting influenza outbreaks: metanarrative review of prospective evaluation. *BMJ Open* **6**, e010683 (2016).
47. Bauch, C. T. & Bhattacharyya, S. Evolutionary game theory and social learning can determine how vaccine scares unfold. *PLOS Computational Biology* **8**, 1–12 (2012).
48. Anderson, R. M. & May, R. M. *Infectious Diseases of Humans: Dynamics and Control* (Oxford University Press, 1992).
49. Vespignani, A. & Moreno, Y. Epidemic spreading in complex networks. *Physics Reports* **414**, 29–88 (2005).
50. Das, M. & Green, J. R. Critical fluctuations and slowing down of chaos. *Nature communications* **10**, 2155 (2019).
51. Donangelo, P., Fort, H., Dallos, V., Scheffer, M. & van Nes, E. H. Early warnings for catastrophic shifts in ecosystems: Comparison between spatial and temporal indicators. *International Journal of Bifurcation and Chaos* **20**, 315–321 (2010).
52. Jentsch, P. C., Anand, M. & Bauch, C. T. Spatial autocorrelation as an early warning signal of regime shifts in a multiplex disease-model behaviour network. *Journal of Theoretical Biology* **448**, 17–25 (2018).
53. Orozco-Fuentes, S. et al. Early warning signals in plant disease outbreaks. *Ecological Modelling* **393**, 12–19 (2019).
54. Murray, J. & Cohen, A. L. *Infectious Disease Surveillance in China*, chap. 2, 23–33 (Elsevier, 2017).
55. Specco, A. & Timpka, T. Algorithms for detecting and predicting influenza outbreaks: metanarrative review of prospective evaluation. *BMJ Open* **6**, e010683 (2016).
56. Bauch, C. T. & Bhattacharyya, S. Evolutionary game theory and social learning can determine how vaccine scares unfold. *PLOS Computational Biology* **8**, 1–12 (2012).
57. Anderson, R. M. & May, R. M. *Infectious Diseases of Humans: Dynamics and Control* (Oxford University Press, 1992).
58. Vespignani, A. & Moreno, Y. Epidemic spreading in complex networks. *Physics Reports* **414**, 29–88 (2005).
59. Das, M. & Green, J. R. Critical fluctuations and slowing down of chaos. *Nature communications* **10**, 2155 (2019).
60. Diks, C., Hommes, C. & Wang, J. Critical slowing down as an early warning signal for financial crises. *Empirical Economics* (2018).
61. Donangelo, P., Fort, H., Dallos, V., Scheffer, M. & van Nes, E. H. Early warnings for catastrophic shifts in ecosystems: Comparison between spatial and temporal indicators. *International Journal of Bifurcation and Chaos* **20**, 315–321 (2010).
62. Jentsch, P. C., Anand, M. & Bauch, C. T. Spatial autocorrelation as an early warning signal of regime shifts in a multiplex disease-model behaviour network. *Journal of Theoretical Biology* **448**, 17–25 (2018).
63. Orozco-Fuentes, S. et al. Early warning signals in plant disease outbreaks. *Ecological Modelling* **393**, 12–19 (2019).
64. Murray, J. & Cohen, A. L. *Infectious Disease Surveillance in China*, chap. 2, 23–33 (Elsevier, 2017).
65. Specco, A. & Timpka, T. Algorithms for detecting and predicting influenza outbreaks: metanarrative review of prospective evaluation. *BMJ Open* **6**, e010683 (2016).
66. Patterson, S. & Bamieh, B. Interaction-driven opinion dynamics on online social networks. In *Proceedings of the First Workshop on Social Media Analytics*, 98–105 (2010).

Acknowledgements

This research was funded by an NSERC Discovery Grant to C.T.B.
**Author contributions**
C.T.B. conceived the idea and supervised the project. B.P. designed the methodology and designed and performed all simulations and data analyses. All authors interpreted and discussed the results of the work. B.P. and C.T.B. wrote and edited the manuscript with input from M.A., C.T.B. and M.A. provided valuable critical feedback.

**Competing interests**
The authors declare no competing interests.

**Additional information**

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-63849-0.

Correspondence and requests for materials should be addressed to B.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2020