

Homophily impacts the success of vaccine roll-outs

Giulio Burgio¹, Benjamin Steinegger ¹ & Alex Arenas ¹✉

Physical contacts do not occur randomly, rather, individuals with similar socio-demographic and behavioral characteristics are more likely to interact among them, a phenomenon known as homophily. Concurrently, the same characteristics correlate with the adoption of prophylactic tools. As a result, the latter do not unfold homogeneously in a population, affecting their ability to control the spread of infectious diseases. Focusing on the case of vaccines, we reveal that, provided an imperfect vaccine efficacy, three different dynamical regimes exist as a function of the mixing rate between vaccinated and not vaccinated individuals. Specifically, depending on the epidemic pressure, vaccine coverage and efficacy, we find the final attack rate to decrease, increase or vary non monotonously with respect to the mixing rate. We corroborate the phenomenology through Monte Carlo simulations on a temporal real-world contact network. Besides vaccines, our findings hold for any prophylactic tool that reduces but not suppress the probability of transmission, indicating a universal mechanism in spreading dynamics.

¹Departament d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, Tarragona, Catalonia, Spain. ✉email: alexandre.arenas@urv.cat

Vaccines have been crucial in humanity's struggle to protect itself from infectious diseases¹. In the 20th century, vaccines enabled to control a series of diseases² as well as the eradication of smallpox³. Nevertheless, vaccines are not uniformly adopted in the population. While on a world-wide scale a lack of access impedes equitable adoption of vaccines^{4,5}, among high income countries vaccine hesitancy is the primary barrier^{6–8}.

Vaccine hesitancy widely correlates with age, socio-economic status, education level, or ethnicity^{9–12}. Concurrently, these same factors shape the interaction patterns in society, leading to what is commonly referred to as homophily¹³, i.e., similarity of social contacts. As a consequence, social interactions are relatively homogeneous with regard to many socio-demographic or behavioral characteristics¹³, and vaccination status is not an exception^{6,7,14–19}. This non-uniform, clustered vaccine adoption strongly determines how, and whether, the virus spreads in the population. A great example of this effect is provided by the recurrent measles outbreaks in high-income countries caused by clusters of vaccine hesitant individuals^{6,7,14–16,18}.

These recurring outbreaks sparked modeling studies that analyzed the impact of homophilic vaccine adoption on the disease dynamics^{20–24}. Due to the high quality of vaccines against measles, these models assumed vaccine efficacy of almost 100%, and showed that clustered adoption always increases the final attack rate. In contrast, vaccines as the ones against influenza or variants of concern of SARS-CoV-2 have relatively low efficacy, between 20% and 80%^{25,26}.

In this work, through the joint exploration of imperfect immunization and vaccine uptake, we offer a wider picture in which vaccination clustering is not always detrimental. Adjustable interaction rates allow us to model different contact structures with respect to the vaccination status of the interacting individuals. In such mean-field approximation, we reveal the existence of three dynamical regimes in relation to how the vaccination assortativity affects the epidemic size. This phenomenology is then corroborated by simulations on a real-world temporal contact network. Overall, integrating recent findings on the effect of homophilic adoption of digital proximity tracing apps²⁷, we point out a general mechanism in spreading dynamics, implied by any prophylactic measure that reduces but not suppress the probability of transmission.

Results and discussion

Modeling the contact structure and the spreading dynamics.

We consider the standard susceptible-infected-recovered (SIR) model, with transmission probability β , recovery rate μ , and contact rate k . Accordingly, in the absence of protected individuals, and assuming homogeneous mixing, the basic reproduction number of the disease is $R_0 = \beta k / \mu$ ²⁸. The fraction of people who received a vaccine is fixed as $V \in [0, 1]$. Upon encounter, an infected individual transmits the infection to a vaccinee at a reduced probability $\beta(1 - \varepsilon)$, where $\varepsilon \in [0, 1]$ represents the vaccine efficacy.

Under random mixing, the effective reproduction number would be given by $R = R_0(1 - \varepsilon V)$. In order to include homophilic interactions, we now parametrize the mixing relation between vaccinated and not vaccinated people, i.e., the contact matrix \mathbf{K} , with a parameter $\alpha \in [0, 1]$. We denote the entries of \mathbf{K} as k_{ij} with $i, j \in \{V, N\}$, where 'V' and 'N' stand for 'vaccinated' and 'not vaccinated', respectively. We introduce α through the relation $k_{NV} = \alpha V k$, hence interpolating from complete homophily ($\alpha = 0$) to random mixing ($\alpha = 1$). The remaining contact rates follow from the balance equation $(1 - V)k_{NV} = V k_{VN}$ and the constraint $k = k_{NN} + k_{NV} = k_{VV} + k_{VN}$. Accordingly, \mathbf{K} has the

following entries:

$$k_{NV} = \alpha V k \quad (1)$$

$$k_{NN} = [1 - \alpha V] k \quad (2)$$

$$k_{VN} = \alpha(1 - V) k \quad (3)$$

$$k_{VV} = [1 - \alpha(1 - V)] k. \quad (4)$$

The degree of homophily regarding vaccine uptake, h , i.e., the probability that during a contact both individuals are either vaccinated or not, thus reads

$$h = \frac{1}{k} [(1 - V)k_{NN} + V k_{VV}] = 1 - 2\alpha V(1 - V). \quad (5)$$

Note that α may take values larger than one (with an upper bound depending on V), indicating disassortative mixing. Broad sociological evidence^{13,29,30}, however, largely excludes this possibility. Recently, the adoption of contact tracing apps has been shown to strongly correlate with age, income and nationality^{31–33}, effectively leading to homophilic (assortative) patterns.

Similarly, we tried to estimate assortativity in vaccine uptake indirectly through age-stratification. We leverage the correlations in the contact patterns among age groups and the levels of vaccine coverage within each age group (see the "Methods" section for details). Specifically, for each region/country considered, we combined the age-stratified contact matrices³⁴ with the data on vaccine adoption against SARS-CoV-2 (counting full vaccinations only), from January 2021 –when first full vaccinations appeared– to September 2021^{35–38}. As shown in Fig. 1, the estimated mixing rate stays well below 1 for almost all the time, sometimes fluctuating around it only for very low levels of vaccine coverage ($V \lesssim 0.02$ for Catalonia).

It must be said, however, that it would be naïve to take the temporal trends in Fig. 1 as ready-to-use data. Indeed, we neglect other important features beyond age such as socio-economic classes or spatial patterns^{9,12}. Including these additional features would reinforce the homophilic structure and thus decrease the mixing rate. Therefore, the results deriving from our simple analysis must be solely understood as a qualitative indication in line with existing literature reporting homophily in health behavior^{17,30,33}, hence with the choice $\alpha \in [0, 1]$. The latter is always kept fixed here, leaving the study of its implicit time-dependence for future work.

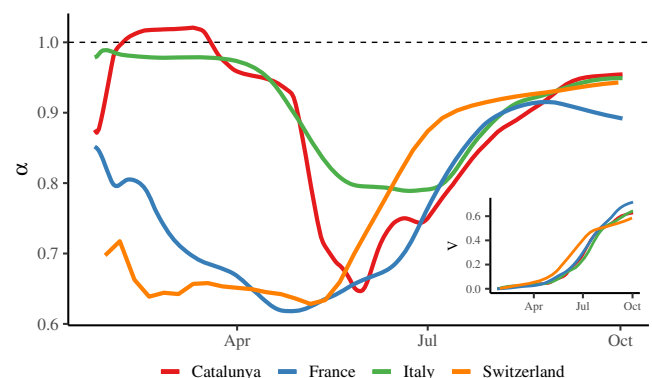


Fig. 1 Mixing parameter from age-stratified data on vaccine uptake.

Weekly moving average of the mixing parameter, α , as inferred for different countries/regions starting from the contact matrix among age groups and the SARS-CoV-2 vaccines uptake of each group (whose aggregate, V , is shown in the inset plot) from January to September 2021.

We denote with $X_Y \equiv X_Y(t)$ the fraction of vaccinated ($Y = V$) and not vaccinated ($Y = N$) people in compartment $X \in \{S, I\}$ at time t . Then, the differential equations governing the dynamics read as

$$\dot{I}_N(t) = \beta [k_{NN}I_N(t) + k_{NV}I_V(t)]S_N(t) - \mu I_N(t) \quad (6)$$

$$\dot{I}_V(t) = \beta [k_{VN}I_N(t) + k_{VV}I_V(t)](1 - \varepsilon)S_V(t) - \mu I_V(t) \quad (7)$$

$$\dot{S}_N(t) = -\beta [k_{NN}I_N(t) + k_{NV}I_V(t)]S_N(t) \quad (8)$$

$$\dot{S}_V(t) = -\beta [k_{VN}I_N(t) + k_{VV}I_V(t)](1 - \varepsilon)S_V(t). \quad (9)$$

Calculation of the reproduction number. Linearizing Eqs. (6)–(9) around the disease-free equilibrium $(I_N, I_V, S_N, S_V) \approx (0, 0, 1, 1)$, we note the Jacobian matrix \mathbf{J} to act as the null matrix in the (S_N, S_V) two-dimensional subspace, having only zero entries in the respective columns. This subspace thus provides two zero eigenvalues. The other two are provided by the 2×2 block \mathbf{J}_I related to the I_N and I_V compartments, taking the form

$$\mathbf{J}_I = \begin{pmatrix} \beta k_{NN} - \mu & \beta k_{NV} \\ \beta(1 - \varepsilon)k_{VN} & \beta(1 - \varepsilon)k_{VV} - \mu \end{pmatrix}. \quad (10)$$

The effective reproduction number, R , is found from the spectral radius $\rho(\mathbf{J})$ as $R = \rho(\mathbf{J}) + 1$ ³⁹. Since $\rho(\mathbf{J}) = \max\{0, \rho(\mathbf{J}_I)\}$, we can just look at $\rho(\mathbf{J}_I)$. If $\rho(\mathbf{J}_I) > 0$ ($R > 1$) the disease-free equilibrium is unstable and a finite fraction of the population gets infected. By computing the spectral radius and inserting the explicit expressions of the \mathbf{K} matrix entries, one gets

$$R = \frac{R_0}{2} \left[2 - \alpha + \alpha\varepsilon(1 - V) - \varepsilon + \sqrt{[\alpha - \alpha\varepsilon(1 - V) + \varepsilon]^2 - 4\alpha\varepsilon V} \right]. \quad (11)$$

One can easily check that R/R_0 decreases monotonously with respect to all the figuring parameters. The monotonous dependence on α was expected, as the higher is the latter, the lower is the number of effectively susceptible people an infected individual can come in contact with. The monotonic shape of R distinguishes the adoption of vaccines from contact tracing apps, for which R can show instead a non monotonous dependence on the mixing rate²⁷.

As anticipated, Eq. (11) reduces to $R = R_0(1 - \varepsilon V)$ for $\alpha = 1$, i.e., random mixing. Interestingly, the symmetry between vaccine efficacy, ε , and coverage, V , holding for random mixing, breaks for homophilic adoption. As proved in the ‘‘Methods’’ section and illustrated in Fig. 2, when keeping the product εV fixed, a higher coverage lowers the reproduction number more than a higher efficacy.

By imposing $R = 1$ and solving for α , we find the critical value of mixing, α_c , above which the disease cannot thrive, to be

$$\alpha_c = \left(1 - \frac{1}{R_0} \right) \frac{1 - R_0(1 - \varepsilon)}{1 - R_0(1 - \varepsilon) - \varepsilon(1 - V)}, \quad (12)$$

given that $\alpha_c \geq [2(1 - 1/R_0) - \varepsilon]/[1 - \varepsilon(1 - V)]$, with $R_0 > 1$, is satisfied. In particular, the condition is always met for $\varepsilon = 1$ and never met for $\varepsilon = 0$ (meaning eradication is not possible in this case, i.e., $R > 1$). The critical values of vaccine uptake, V_c , and efficacy, ε_c , are found from Eq. (12) by solving for the respective variable.

Three dynamical regimes. The final attack rate, understood as the final fraction of individuals that got infected (and eventually transitioned to the recovered/removed compartment) throughout all the duration of an outbreak, exhibits three different dynamical

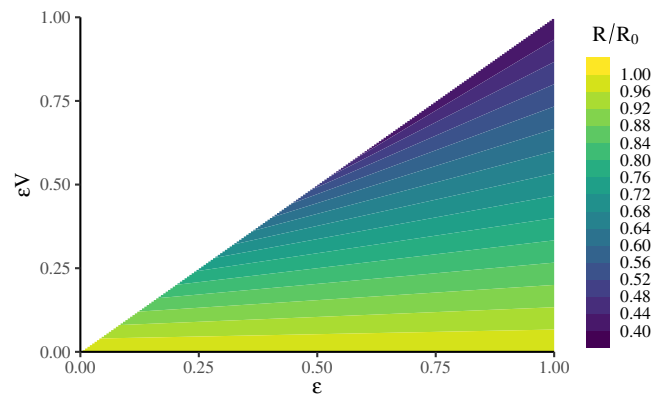


Fig. 2 Analytical mean field prediction for the reproduction number as a function of vaccine efficacy and coverage. Normalized reproduction number, R/R_0 , respect to the the basic reproduction number, R_0 , as a function of the vaccine efficacy, ε , and of εV , the product between vaccine efficacy and vaccine coverage, V . The normalized reproduction number is computed from Eq. (11). For any given value of εV (a horizontal slice of the heatmap), ε can vary from that value to 1 (respectively, V varies from 1 to that value), for $\varepsilon V \leq \varepsilon$ (respectively, $\leq V$).

regimes with respect to its dependence on the mixing rate, α . This is shown in Fig. 3a, where, by increasing the basic reproduction number, R_0 , the dependence of the final attack rate on α is first monotonously decreasing, then concave and finally monotonously increasing. Following previous work²⁷, we refer to these three regimes as critical, intermediate, and saturated, respectively.

These overall regimes result from the competition between two nearly complementary regimes, as illustrated in Fig. 3b, c. With no mixing at all ($\alpha = 0$), vaccinated and not vaccinated form two disconnected components. Accordingly, the disease is free to spread in the not vaccinated cluster whenever $R = R_0 > 1$; on the other hand, $R = R_0(1 - \varepsilon) < R_0$ in the vaccinated one, so the spread does not occur if ε is high enough to ensure $R < 1$. Increasing the mixing (i.e., α), each cluster is ‘diluted’ with nodes from the other. As a consequence, the reciprocal protection that vaccinated people provide among them to keep infection chains localized is partially lost. Instead, not vaccinated individuals profit from vaccinated individuals in their vicinity and are thus subject to a lower infection probability. Mixing is therefore beneficial for not vaccinated people and detrimental for vaccinated ones. Then, depending on the basic reproduction number, R_0 , the vaccine coverage, V , and the vaccine efficacy, ε , the overall system falls in one of the three dynamical regimes.

The same qualitative picture holds for the peak of prevalence (see Fig. 3d–f), i.e., the maximum number of simultaneously infectious people during the epidemic. Note that, for each set of fixed parameters, the respective regimes of the peak of prevalence and of the final attack rate may not coincide. For example, at $R_0 = 3$, the final attack rate is in the intermediate regime while the peak of prevalence is in the critical one (Fig. 3a). In Fig. 4, we show how the three regimes can be also accessed by varying the vaccine coverage, V , or the vaccine efficacy, ε , or both. Specifically, increasing V and/or ε makes the dynamics pass from the saturated to the critical regime, going through the intermediate one. Accordingly, a sufficiently high efficacy makes random mixing always beneficial.

In the case of a perfect vaccine, i.e., $\varepsilon = 1$, the system is always in the critical regime. Indeed, in this case, we see from Eqs. (6)–(9) that the dynamics reduces to a standard SIR model within the not vaccinated sub-population (S_N, I_N). Such dynamics is then

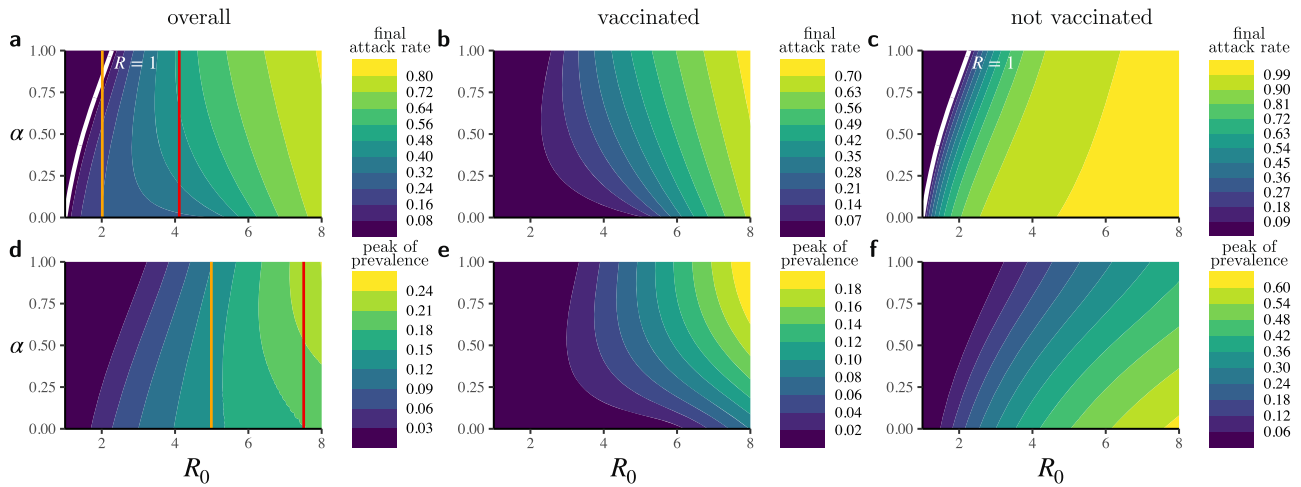


Fig. 3 The three dynamical regimes as predicted by the mean field model. Final attack rate **a-c** and peak of prevalence **d-f** as functions of the basic reproduction number, R_0 , and the mixing parameter, α , as resulting from the numerical iteration of the differential equations governing the system dynamics (Eqs. (6)–(9)). Here, the vaccine coverage is $V = 0.7$ and the vaccine efficacy $\epsilon = 0.8$. In order to highlight the competing processes at the base of the dynamics we show, besides the results for the population overall (**a, d**), those for the vaccinated individuals (**b, e**) and for the not vaccinated ones (**c, f**), separately. The white solid line indicates the critical curve, $\alpha_c \equiv \alpha_c(R_0)$, at which $R = 1$, as computed from Eq. (12). The orange and red vertical lines, demarcate the boundaries between, respectively, the critical and the intermediate regime, and the intermediate and the saturated one, for both the final attack rate (**a**) and the peak of prevalence (**d**).

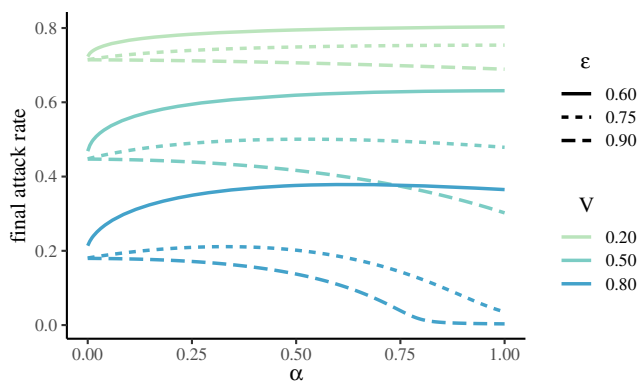


Fig. 4 Final attack rate for different values of vaccine efficacy and coverage. Final attack rate as a function of the mixing parameter, α , for different combinations of vaccine efficacy, ϵ , and coverage, V , given a basic reproduction number $R_0 = 2.5$. As a complement to Fig. 3, this plot illustrates how all three dynamical regimes can be also explored by varying V or ϵ , or both.

solely driven by the contact rate k_{NN} , which is a linearly decreasing function of the mixing rate, α .

As mentioned in the section “Introduction”, it has been recently²⁷ revealed that the same three dynamical regimes found here for vaccination, are also induced by the homophilic adoption of digital proximity tracing apps for epidemic containment. Indeed, the existence of these regimes can be understood within a common underlying mechanism, enclosing a broad class of prophylactic measures. The essential ingredient is the lack of sufficient individual, independent protection offered by the prophylactic tool to the adopter. This is evident for contact tracing apps, as adoption does not directly protect the adopter; instead, individuals only indirectly benefit if adoption is sufficiently widespread such that transmission chains between adopters can be stopped. Similarly, if the vaccine is imperfect, isolated adoption in a high prevalence environment may not substantially reduce the individual infection probability.

Widespread adoption among an individual’s contacts is thus necessary to provide mutual protection.

Accordingly, for both vaccines and contact tracing apps, if coverage is low in comparison to the epidemic pressure, only an assortative/clustered adoption can provide mutual protection and thus non adopters cannot be protected (saturated regime). On the contrary, if coverage is high, protection is sufficiently strong such that nonadopters can be protected and thus mixing between adopters and nonadopters is beneficial (critical regime). In between we then find the intermediate regime. In all three cases, this phenomenology holds for any prophylactic measure that reduces the transmission probability since it is mathematically equivalent to the model presented here. In this sense, equivalent results could be found for the use of face masks or the adoption of social distancing.

Monte Carlo simulations on a real-world contact network. To corroborate our theory, we performed numerical simulations upon a temporal contact network estimated via Bluetooth signal exchanges in the Copenhagen Networks Study⁴⁰. The data was collected during a month with almost 700 participants. In order to study different levels of assortativity in vaccine adoption, we distribute vaccines algorithmically until a preset value of α is reached (see the “Methods” section for details)⁴¹.

The results for the final attack rate, and the peak of prevalence, respectively, reported in Fig. 5a, b for $\epsilon = 1.0$, in Fig. 5c, d for $\epsilon = 0.8$, and in Fig. 5e, f for $\epsilon = 0.6$, confirm the existence of the three dynamical regimes identified by our model. Accordingly, the beneficial effect of random mixing vanishes when the vaccine does not provide sufficient protection. In the saturated regime, mixing is slightly detrimental. Furthermore, as in the mean-field case, we observe that final attack rate and peak of prevalence can be in different dynamical regimes. Particular structural constraints of this aggregated network do not allow for arbitrarily small values of α , making the saturated regime less evident. Nonetheless, higher homophilic levels (even $\alpha \approx 0$) can be reached by artificially controlling the mixing, in which case the decrease in the final attack rate would be better appreciated.

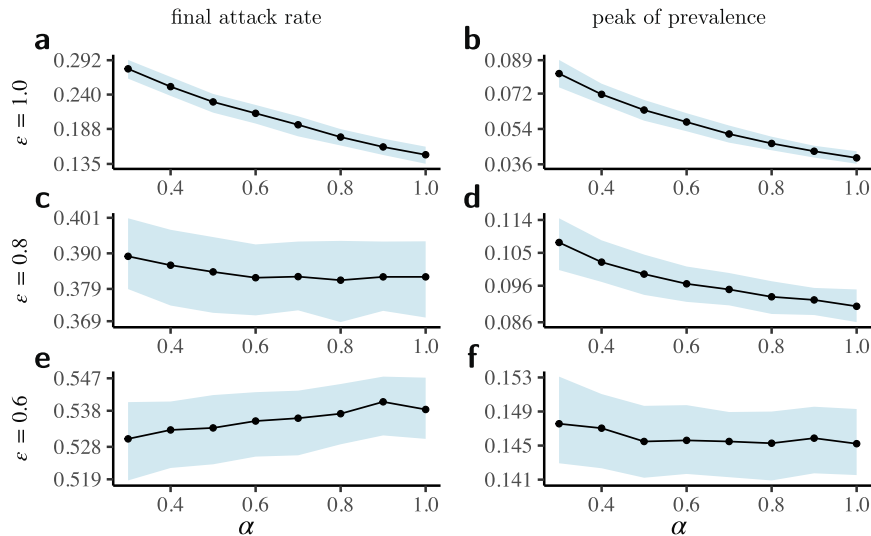


Fig. 5 The three dynamical regimes confirmed for a real-world temporal contact network. Final attack rate and peak of prevalence, respectively, for vaccine efficacy, $\epsilon = 1.0$ (a, b), $\epsilon = 0.8$ (c, d), and $\epsilon = 0.6$ (e, f) as functions of the mixing parameter, α , resulting from the numerical simulations performed on top of a real-world temporal contact network from the Copenhagen Networks Study⁴⁰. This consists of $N = 672$ nodes and 374,884 edges (pairwise interactions) spread over 8064 timestamps, binning 4 weeks of recording time into 5-min intervals. Dots indicate the median value, whereas the shaded area indicates first and third quartiles. Each point is obtained by averaging over 2×10^4 runs. We fixed the vaccine coverage to $V = 0.5$, the recovery rate to $\mu = 4.6 \times 10^{-4}$ (corresponding to a mean infection time of 7.5 days) and the transmission probability to $\beta = 1.142 \times 10^{-1}$, yielding $R_0 = 6$ from the estimation $R_0 = \beta\kappa/\mu$, where $\kappa = \overline{s^2}/\overline{s}$, being \overline{s} ($\overline{s^2}$) the network average of the (squared) number of contacts per timestamp.

Conclusions

Depending on the need and the available information, the model can be applied to different specific populations and at different scales. Currently, many epidemic models only implicitly consider the mixing between not vaccinated and vaccinated individuals through stratification according to age (Fig. 1)^{42,43} or socio-economic status⁴⁴. Our findings urge also to account for the impact that subgroup-specific levels of mixing, can have on the system as a whole. Such approaches may provide a further tool to interpret epidemiological data. Additionally, more realistic models can act as useful guides for policy makers to design non-pharmaceutical interventions.

For instance, with respect to the current COVID-19 pandemic, various countries require the green pass to enter restaurants or the workplace. The intention behind the green pass is to nudge individuals to get vaccinated or that non vaccinated individuals reduce their number of physical contacts⁴⁵. However, at the same time, the green pass reshapes the social fabric⁴⁶ effectively reducing the mixing rate between vaccinated and non vaccinated individuals. In light of our results, due to the associated reduction in the mixing rate, models that do not include homophily in vaccine adoption may actually overestimate the impact of the green pass. Overall, our results show that more detailed information on the correlations between social interactions and health behavior (vaccination, face mask, social distancing, contact tracing apps) would lead to a more comprehensive analysis.

Shortly after having finished this analysis, we became aware that Hiraoka et al.⁴⁷ and Watanabe et al.⁴⁸ made an analysis very similar to ours. The phenomenology they uncover is equivalent to the one we found here. In this sense, the findings are very robust regarding model assumptions, as Hiraoka et al.⁴⁷ consider random networks and an all-or-nothing vaccine⁴⁹, while we focus on the mean-field case and a temporal real-world network assuming a leaky vaccine⁴⁹. Additionally, Fefferman et al.^{50,51} showed that the presence of homophily adds interesting phenomenology regarding the temporal trajectory of the epidemics beyond the three dynamical regimes in the final attack rate.

From a more theoretical standpoint, we showed that the presence of homophily in vaccine adoption leads to three different dynamical regimes (critical, intermediate, saturated). Furthermore, the phenomenology presented here also extend to any prophylactic measure that reduces the transmission probability such as the use face masks⁴⁸, social distancing⁵², as well as digital contact tracing^{27,53}. Accordingly, the phenomenology induced by the presence of homophily is robust with respect to the adoption of different health behaviors. This robustness of our findings across prophylactic tools hints to a general feature of spreading dynamics. Eventually, the results highlight how correlated meta-data such as vaccination status can add rich phenomenology beyond the network structure itself^{54–57}.

Methods

Homophily level of SARS-Cov-2 vaccine adoption. We infer the mixing parameter α of SARS-CoV-2 vaccine adoption through an age-stratified approach. We exclusively focus on the heterogeneous adoption and contact structure with respect to age strata. Due to limited data availability, we neglect further correlations such as socio-economic factors. In this sense, the inferred value of α likely represents an upper bound.

To calculate α we make use of the empirical contact matrix C^{34} and the vaccination coverage $v_i(t)$ in age-strata $i = 1, 2, \dots, M$ at time $t^{35–38}$. The entry C_{ij} of the contact matrix represents the average number of daily contacts of an individual in age-strata i with an individual in age-strata j . Vaccination data is stratified into intervals of 10 years (0–9, 10–19, ..., 70–79, 80+), while the contact matrices use 5 year bins. Therefore, as a first step, we average the empirical contact matrices to transform them into 10 year bins according to the census data.

Generally, the empirical contact matrices do not respect the balance equation $C_{ij}N_i = C_{ji}N_j$, where N_i represents the number of individuals in age-strata i . There are various approaches to fix this issue⁵⁸. Here, we defined the matrix C as

$$C_{ij} = \frac{N_i \tilde{C}_{ij} + N_j \tilde{C}_{ji}}{N_i} \tag{13}$$

such that the balance equation is met, being \tilde{C} the empirical matrix. To infer $\alpha(t)$ we leverage its relation with $h(t)$, i.e., Eq. (5). Namely, we first calculate h as

$$h(t) = \frac{\sum_{i=1}^M \sum_{j=1}^M C_{ij} N_i [v_i(t)v_j(t) + (1 - v_i(t))(1 - v_j(t))]}{\sum_{i=1}^M \sum_{j=1}^M C_{ij} N_i} \tag{14}$$

From Eq. (5), we find the mixing rate as $\alpha(t) = (1-h)/[2V(1-V)]$, where V corresponds to the fraction of vaccinated individuals in the overall population.

Asymmetry between vaccine efficacy and coverage in lowering the reproduction number. Here we prove that, when keeping fixed the product ϵV between vaccine efficacy, ϵ , and coverage, V , a higher coverage is more effective than a higher efficacy in lowering the reproduction number, R . To this end, it is convenient to rewrite Eq. (11) in the form

$$R = \frac{R_0}{2} \left[2 - \alpha(1 + \epsilon V) - \epsilon(1 - \alpha) + \sqrt{\epsilon^2(1 - \alpha)^2 + \epsilon[2\alpha(1 - \alpha)(1 + \epsilon V)] + [\alpha(1 + \epsilon V)]^2 - 4\alpha\epsilon V} \right]. \quad (15)$$

Differentiating Eq. (15) with respect to ϵ while keeping the product ϵV fixed, one finds

$$\left. \frac{\partial R}{\partial \epsilon} \right|_{\epsilon V = \text{const.}} = \frac{R_0}{2} (1 - \alpha) \left[-1 + \frac{\epsilon + \alpha(1 + \epsilon V - \epsilon)}{\sqrt{[\epsilon(1 - \alpha)]^2 + \epsilon[2\alpha(1 - \alpha)(1 + \epsilon V)] + c_2(\epsilon V)}} \right]. \quad (16)$$

This is zero for $\alpha = 1$. Supposing $\alpha \neq 1$ and imposing $\partial R/\partial \epsilon|_{\epsilon V = \text{const.}} \geq 0$, after some algebra, one gets the condition $4\alpha\epsilon V \geq 0$, which is always satisfied. Therefore, it holds $(\partial R/\partial \epsilon)|_{\epsilon V = \text{const.}} \geq 0$ and, in a complementary manner, $(\partial R/\partial V)|_{\epsilon V = \text{const.}} = -(\epsilon/V)(\partial R/\partial \epsilon)|_{\epsilon V = \text{const.}} \leq 0$, which completes the proof.

Assortative seeding of vaccines on a network. We consider the temporal contact network from the Copenhagen Networks Study⁴⁰ that was inferred through Bluetooth signals. We retained those interactions with an associated received signal strength indication (RSSI) not lower than -74 dBm, corresponding to physical distances approximately up to 2 m⁵⁹. The resulting temporal network involves $N = 672$ individuals and $374,884$ pairwise interactions spread over 8064 time-stamps, binning 4 weeks of recording time into 5 -min intervals.

To distribute the vaccine in the population, we provisionally aggregate the temporal network to get a weighted, static one, where the weight of an edge, w_{ij} , is the duration of time its end nodes interacted, understood as the number of 5 -min bins this occurred. Accordingly, the homophily, h , is computed as the sum of the weights (duration of contacts) over the homophilic edges, normalized by the sum over all the edges. In line with Eq. (5), given the vaccination status, v_i ($v_i = 1$ if vaccinated and 0 otherwise), of an individual i , we can express h as

$$h = \frac{\sum_{i=1}^N \sum_{j=1}^N w_{ij} [v_i v_j + (1 - v_i)(1 - v_j)]}{\sum_{i=1}^N \sum_{j=1}^N w_{ij}}. \quad (17)$$

Then α is obtained from Eq. (5) as $\alpha = (1-h)/[2V(1-V)]$.

To reach a desired mixing level α , we initially distribute the vaccine at random ($\alpha = 1$) and then iteratively swap the vaccination status of two randomly selected neighboring nodes whenever this leads to a decrease in α , until the latter attains the preset value (± 0.01). Additionally, if the average strength (number of contacts) of vaccinated individuals is higher (lower) than that of not vaccinated ones, a swap is only allowed if it increases (decreases) the strength of not vaccinated ones. Without this condition, the algorithm induces a spurious correlation between vaccination status and the strength of a node. Given the role that frequently interacting nodes play in driving the spreading dynamics, such correlation can importantly affect the results. Therefore, as done in Burgio et al.²⁷—but not in a previous, related work¹⁷—only innocuous, not correlating swaps are carried out. The code for the seeding of the vaccine is publicly available⁴¹.

Data availability

All the data used in this study is publicly available and accordingly referenced.

Code availability

The code to reproduce all the results presented here is available at github and archived at zenodo.org⁴¹.

Received: 21 December 2021; Accepted: 25 February 2022;

Published online: 28 March 2022

References

- McNeill, W. H. *Plagues and Peoples*. (Anchor Books, New York, USA, 1976).
- Pollard, A. J. & Bijker, E. M. A guide to vaccinology: from basic principles to new developments. *Nat. Rev. Immunol.* **21**, 83–100 (2021).

- Smith, G. L. & McFadden, G. Smallpox: anything to declare? *Nat. Rev. Immunol.* **2**, 521–527 (2002).
- Smith, J., Lipsitch, M. & Almond, J. W. Vaccine production, distribution, access, and uptake. *Lancet (London, England)* **378**, 428–438 (2011).
- Wouters, O. J. et al. Challenges in ensuring global access to COVID-19 vaccines: production, affordability, allocation, and deployment. *Lancet (London, England)* **397**, 1023–1034 (2021).
- May, T. & Silverman, R. D. ‘Clustering of exemptions’ as a collective action threat to herd immunity. *Vaccine* **21**, 1048–1051 (2003).
- Parker, A. A. et al. Implications of a 2005 measles outbreak in Indiana for sustained elimination of measles in the United States. *N. Engl. J. Med.* **355**, 447–455 (2006).
- Editorial. Vaccine hesitancy: a generation at risk. *Lancet Child Adolesc. Health* **3**, 281 (2019).
- Robertson, E. et al. Predictors of COVID-19 vaccine hesitancy in the UK household longitudinal study. *Brain Behav. Immun.* **94**, 41–50 (2021).
- Larson, H. J., Jarrett, C., Eckersberger, E., Smith, D. M. D. & Paterson, P. Understanding vaccine hesitancy around vaccines and vaccination from a global perspective: a systematic review of published literature, 2007–2012. *Vaccine* **32**, 2150–2159 (2014).
- MacDonald, N. E. & on Vaccine Hesitancy, S. W. G. Vaccine hesitancy: definition, scope and determinants. *Vaccine* **33**, 4161–4164 (2015).
- Casini, F., Pantovic, A., Al-Ajlouni, Y., Failla, G. & Ricciardi, W. Attitudes, acceptance and hesitancy among the general population worldwide to receive the COVID-19 vaccines and their contributing factors: a systematic review. *EclinicalMedicine* **40**, [https://www.thelancet.com/journals/eclinm/article/PIIS2589-5370\(21\)00393-X/fulltext](https://www.thelancet.com/journals/eclinm/article/PIIS2589-5370(21)00393-X/fulltext) (2021).
- McPherson, M., Smith-Lovin, L. & Cook, J. M. Birds of a feather: homophily in social networks. *Annu. Rev. Sociol.* **27**, 415–444 (2001).
- Richard, J. L., Masserey-Spicher, V., Santibanez, S. & Mankertz, A. Measles outbreak in Switzerland—an update relevant for the European football championship (EURO 2008). *Euro Surveill.* **13**, 8043 (2008).
- Omer, S. B. et al. Geographic clustering of nonmedical exemptions to school immunization requirements and associations with geographic clustering of pertussis. *Am. J. Epidemiol.* **168**, 1389–1396 (2008).
- Atwell, J. E. et al. Nonmedical vaccine exemptions and pertussis in California, 2010. *Pediatrics* **132**, 624–630 (2013).
- Barclay, V. C. et al. Positive network assortativity of influenza vaccination at a high school: implications for outbreak risk and herd immunity. *PLoS ONE* **9**, e87042 (2014).
- Lieu, T. A., Ray, G. T., Klein, N. P., Chung, C. & Kulldorff, M. Geographic clusters in underimmunization and vaccine refusal. *Pediatrics* **135**, 280–289 (2015).
- Edge, R., Keegan, T., Isba, R. & Diggle, P. Observational study to assess the effects of social networks on the seasonal influenza vaccine uptake by early career doctors. *BMJ Open* **9**, e026997 (2019).
- Mbah, M. L. N. et al. The impact of imitation on vaccination behavior in social contact networks. *PLoS Comput. Biol.* **8**, e1002469 (2012).
- Salathé, M. & Bonhoeffer, S. The effect of opinion clustering on disease outbreaks. *J. R. Soc. Interface* **5**, 1505–1508 (2008).
- Liu, F. et al. The role of vaccination coverage, individual behaviors, and the public health response in the control of measles epidemics: an agent-based simulation for California. *BMC Public Health* **15**, 447 (2015).
- Glasser, J. W., Feng, Z., Omer, S. B., Smith, P. J. & Rodewald, L. E. The effect of heterogeneity in uptake of the measles, mumps, and rubella vaccine on the potential for outbreaks of measles: a modelling study. *Lancet Infect. Dis.* **16**, 599–605 (2016).
- Kuylen, E., Willem, L., Broeckhove, J., Beutels, P. & Hens, N. Clustering of susceptible individuals within households can drive measles outbreaks: an individual-based model exploration. *Sci. Rep.* **10**, 19645 (2020).
- Centers for Disease Control and Prevention. *CDC Seasonal Flu Vaccine Effectiveness Studies*. <https://www.cdc.gov/flu/vaccines-work/effectiveness-studies.htm> (2021).
- Pouwels, K. B. et al. Effect of Delta variant on viral burden and vaccine effectiveness against new SARS-CoV-2 infections in the UK. *Nat. Med.* **27**, 2127–2135 (2021).
- Burgio, G., Steinegger, B., Rapisardi, G. & Arenas, A. Homophily in the adoption of digital proximity tracing apps shapes the evolution of epidemics. *Phys. Rev. Res.* **3**, 033128 (2021).
- Anderson, R. M. & May, R. M. *Infectious Diseases of Humans: Dynamics and Control* (Oxford University Press, 1992).
- Centola, D. The spread of behavior in an online social network experiment. *Science* **329**, 1194–1197 (2010).
- Centola, D. An experimental study of homophily in the adoption of health behavior. *Science* **334**, 1269–1272 (2011).
- Salathé, M. et al. Early evidence of effectiveness of digital contact tracing for SARS-CoV-2 in Switzerland. Preprint at *medRxiv* <https://www.medrxiv.org/content/early/2020/10/04/2020.09.07.20189274> (2020)

32. Munzert, S., Selb, P., Gohdes, A., Stoetzer, L. F. & Lowe, W. Tracking and promoting the usage of a COVID-19 contact tracing app. *Nat. Hum. Behav.* **5**, 247–255 (2021).
33. Moreno López, J. A. et al. Anatomy of digital contact tracing: role of age, transmission setting, adoption and case detection. *Sci. Adv.* **7**, eabd8750 (2021).
34. Prem, K., Cook, A. R. & Jit, M. Projecting social contact matrices in 152 countries using contact surveys and demographic data. *PLoS Comput. Biol.* **13**, e1005697 (2017).
35. Generalitat de Catalunya. Catalunya's COVID-19 vaccine uptake data. <https://dadescovid.cat/descarregues> (2021).
36. Etalab, Direction interministérielle du numérique (DINUM) - Gouvernement Français. France's COVID-19 vaccine uptake data. <https://www.data.gouv.fr/fr/datasets> (2021).
37. Commissario straordinario per l'emergenza Covid-19 - Presidenza del Consiglio dei Ministri - Governo Italiano. Italy's COVID-19 vaccine uptake data. <https://github.com/italia/covid19-opendata-vaccini/tree/master/dati> (2021).
38. Federal Office of Public Health FOPH - Swiss Confederation. Switzerland's COVID-19 vaccine uptake data. <https://www.covid19.admin.ch/en/vaccination/persons> (2021).
39. Van den Driessche, P. Reproduction numbers of infectious disease models. *Infect. Dis. Model.* **2**, 288–303 (2017).
40. Sapiezynski, P., Stopczynski, A., Lassen, D. D. & Lehmann, S. Interaction data from the Copenhagen Networks Study. *Sci. Data* **6**, 315 (2019).
41. Burgio, G. & Steinegger, B. Code to reproduce all the results presented in this paper. <https://doi.org/10.5281/zenodo.6108284> (2022).
42. Bubar, K. M. et al. Model-informed COVID-19 vaccine prioritization strategies by age and serostatus. *Science (New York, NY)* **371**, 916–921 (2021).
43. Sonabend, R. et al. Non-pharmaceutical interventions, vaccination, and the SARS-CoV-2 delta variant in England: a mathematical modelling study. *Lancet (London, England)* **398**, 1825–1835 (2021).
44. Brand, S. P. C. et al. COVID-19 transmission dynamics underlying epidemic waves in Kenya. *Science (New York, NY)* **374**, 989–994 (2021).
45. Wilf-Miron, R., Myers, V. & Saban, M. Incentivizing vaccination uptake: the “green pass” proposal in Israel. *JAMA* **325**, 1503–1504 (2021).
46. Dada, S. et al. Learning from the past and present: social science implications for COVID-19 immunity-based documentation. *Humanit. Soc. Sci. Commun.* **8**, 219 (2021).
47. Hiraoaka, T., Rizi, A. K., Kivelä, M. & Saramäki, J. Herd immunity and epidemic size in networks with vaccination homophily. Preprint at [arXiv:2112.07538](https://arxiv.org/abs/2112.07538) (2021).
48. Watanabe, H. & Hasegawa, T. Impact of assortative mixing by mask-wearing on the propagation of epidemics in networks. Preprint at [arXiv:2112.06589](https://arxiv.org/abs/2112.06589) (2021).
49. Shim, E. & Galvani, A. P. Distinguishing vaccine efficacy and effectiveness. *Vaccine* **30**, 6700–6705 (2012).
50. Fefferman, N. H., Silk, M. J., Pasquale, D. K. & Moody, J. Homophily in risk and behavior complicate understanding the covid-19 epidemic curve. Preprint at [medRxiv https://www.medrxiv.org/content/early/2021/03/20/2021.03.16.21253708](https://www.medrxiv.org/content/early/2021/03/20/2021.03.16.21253708) (2021).
51. Young, M. J., Silk, M. J., Pritchard, A. J. & Fefferman, N. H. Diversity in valuing social contact and risk tolerance lead to the emergence of homophily in populations facing infectious threats. Preprint at [arXiv:2111.11362](https://arxiv.org/abs/2111.11362) (2021).
52. Kadelka, C. & McCombs, A. Effect of homophily and correlation of beliefs on covid-19 and general infectious disease outbreaks. *PLoS ONE* **16**, 1–20 (2021).
53. Rizi, A. K., Faqeh, A., Badie-Modiri, A. & Kivelä, M. Epidemic spreading and digital contact tracing: effects of heterogeneous mixing and quarantine failures. [arXiv preprint arXiv:2103.12634](https://arxiv.org/abs/2103.12634) (2021).
54. Gómez-Gardeñes, J., Gómez, S., Arenas, A. & Moreno, Y. Explosive synchronization transitions in scale-free networks. *Phys. Rev. Lett.* **106**, 128701 (2011).
55. Newman, M. E. J. & Clauset, A. Structure and inference in annotated networks. *Nat. Commun.* **7**, 11863 (2016).
56. Peel, L., Larremore, D. B. & Clauset, A. The ground truth about metadata and community detection in networks. *Sci. Adv.* **3**, e1602548 (2017).
57. Artime, O. & De Domenico, M. Percolation on feature-enriched interconnected systems. *Nat. Commun.* **12**, 2478 (2021).
58. Arregui, S., Aleta, A., Sanz, J. & Moreno, Y. Projecting social contact matrices to different demographic structures. *PLoS Comput. Biol.* **14**, e1006638 (2018).
59. Sekara, V. & Lehmann, S. The strength of friendship ties in proximity sensor data. *PLoS ONE* **9**, e100915 (2014).

Acknowledgements

G.B. acknowledges financial support from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No. 945413 and from the Universitat Rovira i Virgili (URV). B.S. acknowledges financial support from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant Agreement No. 713679 and from the Universitat Rovira i Virgili (URV). A.A. acknowledges support by Ministerio de Economía y Competitividad (Grants No. PGC2018-094754-B-C21 and No. FIS2015-71582-C2-1), Generalitat de Catalunya (Grant No. 2017SGR-896), Universitat Rovira i Virgili (Grant No. 2019PFR-URV-B2-41), ICREA Academia, and the James S. McDonnell Foundation (Grant No. 220020325). We thank L. Arola-Fernández for helpful comments and suggestions.

Author contributions

A.A. conceived the study. A.A., G.B., and B.S. designed the approach. G.B. performed the analysis. G.B. and B.S. drafted the manuscript. All authors edited the manuscript and approved its final version.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Alex Arenas.

Peer review information *Communications Physics* thanks Claus Kadelka and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at <http://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) 2022