Social Science & Medicine 312 (2022) 115350

Contents lists available at ScienceDirect

Social Science & Medicine

journal homepage: www.elsevier.com/locate/socscimed

# Health behavior homophily can mitigate the spread of infectious diseases in small-world networks

Hendrik Nunner [a,b,*], Vincent Buskens [a,b], Alexandra Teslya [c], Mirjam Kretzschmar [b,c]

[a] Department of Sociology/ICS, Utrecht University, Padualaan 14, 3584 CH Utrecht, Netherlands
[b] Centre for Complex System Studies (CCSS), Utrecht University, Leuvenlaan 4, 3584 CW Utrecht, Netherlands
[c] Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Universiteitsweg 100, 3584 CX Utrecht, the Netherlands

ARTICLE INFO

*Keywords:*
Small-world networks
Adaptive networks
Homophily
Health behavior
Risk perception
Infectious diseases
Epidemics
Agent-based simulations

ABSTRACT

Research has repeatedly shown that the spread of infectious diseases is influenced by properties of our social networks. Small-world like structures with densely connected clusters bridged by only a few connections, for example, are not only known to diminish disease spread, but also to increase the chance for a disease to spread to any part of the network. Clusters composed of individuals who show similar reactions to avoid infections (health behavior homophily), however, might change the effect of such clusters on disease spread. To study the combined effect of health behavior homophily and small-world network properties on disease spread, we extend a previously developed ego-centered network formation model and agent-based simulation. Based on more than 80,000 simulated epidemics on generated networks varying in clustering and homophily, as well as diseases varying in severity and infectivity, we predict that the existence of health behavior homophilous clusters reduce the number of infections, lower peak size, and flatten the curve of active cases. That is because agents perceiving higher risks of infections can protect their cluster from infections comparatively quickly by severing only a few bridging ties. A comparison with epidemics in static network structures shows that the incapability to act upon risk perceptions and the low connectivity between clusters in static networks lead to diametrically opposed effects with comparatively large epidemics and prolonged epidemics. These finding suggest that micro-level behavioral adaptation to health risks mitigate macro-level disease spread to an extent that is not captured by static network models of disease spread. Furthermore, this mechanism can be used to design information campaigns targeting proxies for groups with lower risk perception.

## 1. Introduction

The COVID-19 pandemic has revealed the vulnerability of our globalized world to the spread of infectious diseases. As a reaction to this, many countries implemented physical distancing measures to lower the number of infections for a disease mainly transmitted through human-to-human contact in social networks (Li et al., 2020; Shen et al., 2020; WHO, 2021). In other words, reducing social contacts should reduce the transmission of the virus.

However, authority-initiated interventions are not the only factors affecting social network structure to lower the risk of infections. Avoidant health behaviors, such as keeping away from large crowds, avoiding symptomatic coworkers, or avoiding sexual contact because a partner has a sexually transmitted disease, are common self-imposed measures known to mitigate disease spread (Bish and Michie, 2010; Funk et al., 2009). These health behaviors depend on heterogeneous risk perceptions (Bish and Michie, 2010; Leppin and Aro, 2009) that, depending on the composition of a group, may facilitate or hinder the spread of infectious diseases on the macro-level, such as COVID-19 (d'Andrea et al., 2022) or HIV (Koku and Felsher, 2020). Furthermore, risk perceptions depend on personal properties, such as age, gender, ethnicity, educational level, or marital status (Bish and Michie, 2010); properties known to facilitate *homophily* (McPherson et al., 2001), a tendency of people to be connected to others with similar characteristics.

Additionally, certain network properties are known to affect the spread of diseases in networks. On the one hand, small average path length reduces the probability of a disease to die out before reaching

* Corresponding author. Department of Sociology/ICS, Utrecht University, Padualaan 14, 3584 CH Utrecht, the Netherlands.
*E-mail addresses:* h.nunner@uu.nl (H. Nunner), v.buskens@uu.nl (V. Buskens), a.i.teslya@umcutrecht.nl (A. Teslya), m.e.e.kretzschmar@umcutrecht.nl (M. Kretzschmar).

https://doi.org/10.1016/j.socscimed.2022.115350
Received 14 March 2022; Received in revised form 14 June 2022; Accepted 1 September 2022
Available online 18 September 2022

distant points of the network (Watts and Strogatz, 1998). On the other hand, the presence of clusters, densely connected areas within a network that have only few ties bridging to nodes outside this area, can reduce the size of an epidemic (Badham and Stocker, 2010; Keeling, 1999; Miller, 2009). These two features are of particular interest for the spread of infectious diseases, as they constitute the typical structure of our *small-world*-like social networks (Milgram, 1967; Watts and Strogatz, 1998).

Although we know quite well how small-world properties and homophily affect disease spread in networks separately, it is less clear how the two shape epidemics in combination. Furthermore, network models of disease spread typically neglect adaptive changes in the network structure, despite their potentially large effect on predictions (Bansal et al., 2010). We therefore ask: *What are the effects of health behavior homophily on epidemics in adaptive small-world networks?* and *How do epidemics differ between adaptive and static networks?*

### 1.1. Network models of disease spread

For almost 40 years, scientists use network theory to study the spread of infectious diseases (e.g., Klovdahl, 1985; May and Anderson, 1987). A typical approach is to study how the structural properties of relationships (edges) between individuals (nodes) affect the spread of infectious diseases.

Many network models of disease spread, however, assume that relationships do not change over time (Bansal et al., 2010). In their review on epidemic processes in complex networks, Pastor-Satorras et al. (2015) argue that static networks are good approximations, when the network evolves at a much slower pace than the dynamic process working within it. Examples for processes that change network topology slowly are demographic changes, such as births and deaths, migration, or social or economic changes. Short-term changes, however, can be the result of behavioral changes to adapt to health risks thus requiring the consideration of the interdependent dynamics between social networks and infectious diseases (Bansal et al., 2010). An established approach to account for such adaptations as well as heterogeneity on the host level are individual-based models for infectious disease transmission (Bedson et al., 2021; Willem et al., 2017; Verelst et al., 2016). Simulation studies on adaptive networks suggest, for example, that epidemic size is smaller in adaptive networks compared to their static counterparts, no matter whether ties to infectious nodes are simply cut (Gross et al., 2006) or whether infectious nodes rewire randomly (Leung et al., 2018; Risau-Gusman and Zanette, 2009). Another typical application of individual-based models is to test the effectivity of network-based interventions (Verelst et al., 2016), such as vaccination (Breban, 2011; Nunner et al., 2022) or information campaigns (Mao and Yang, 2012; Nyabadza et al., 2010).

### 1.2. Small-world networks

"My it's a small world." is according to Milgram (1967, p. 61) a typical statement we make upon realizing that a stranger knows the same person we do. In a seminal study, he discovered that it takes only about five intermediaries to send a letter from distant places in the US (i.e., Wichita, Kansas and Omaha, Nebraska) to a stockbroker in Boston, Massachusetts. This surprisingly low number is caused by the typical properties of our social networks: high clustering and short average path length (Watts and Strogatz, 1998). That is, small-world networks typically have densely connected areas that are connected by just a few bridges. Translated into our social lives that means we form clusters with people who are also connected among each other (e.g., friends, family), while we typically know few people who are not part of that inner social circle (e.g., a former fellow student). The tendency for high clustering or *triadic closure* (Simmel, 1950) is often studied in the social sciences (Granovetter, 1977), also because social cohesive structures have advantages in terms of cooperation and trust for the people in these structures (Coleman, 1994).

Despite the advantages of social cohesion, we typically do not find fully connected networks. In addition to obvious reasons (spatial distance, too many people, personal sympathies), Jackson (2008) presents a process of consideration preceding the formation of contacts. He argues that while social ties provide benefits (e.g., emotional support, sense of belonging), there are also costs to maintain the relationships (e.g., time, effort). Thus, the answer to whether a social tie is formed or kept is a trade-off between the benefits from having the tie and the costs to maintain it. On the one hand, this trade-off creates tightly-knit clusters within our social networks that are limited in size. On the other hand, clusters are not disconnected from each other, but bridged by a few ties. Burt (1992) argues that bridging ties provide their own advantages such as access to and control of resources and information. Consequently, network properties, such as the number of relations, the degree of clustering, or average path length, depend on costs and benefits shaped by social context.

Previous research on disease spread in social networks has shown repeatedly that increased clustering leads to smaller epidemics (Badham and Stocker, 2010; Keeling, 1999; Miller, 2009). Small-world networks that combine high clustering with short average path lengths, however, facilitate the spread of infectious diseases (Watts and Strogatz, 1998). That is because transmission events are subject to probabilities, and the more transmissions are required to reach distant parts of the network, the more likely the transmission chain breaks. Ties bridging two clusters lower the average path length by providing shortcuts to access distant nodes with only a few steps. It follows that while clusters hamper the spread of infections, increasing numbers of bridges between clusters facilitate the disease to quickly spread through the entire network.

### 1.3. Health behavior homophily

In the context of infectious diseases, Kasl and Cobb (1966) define health behavior as "[…] any activity undertaken by a person believing himself to be healthy, for the purpose of preventing disease or detecting it in an asymptomatic stage" (p. 531). A systematic review of 26 studies on a variety of airborne diseases (SARS, avian influenza, H5N1, swine flu, H1N1) reveals avoidant behavior as one of the most commonly adapted behaviors to prevent disease (Bish and Michie, 2010). Typical avoidant behaviors are removing social ties (Funk et al., 2010), voluntary quarantining (Tracy et al., 2009), or avoiding public places (Jones and Salathé, 2009).

A systematic review of 28 studies on SARS and avian influenza (Leppin and Aro, 2009) showed that independent of how risk is conceptualized, the perception of risks is the most important determinant of health decisions. The two main drivers of risk perception are (i) how probable a person believes it is to get infected, and (ii) how severe that person thinks the disease is (Bish and Michie, 2010; Leppin and Aro, 2009). Risk perceptions, however, are highly subjective in nature (Bults et al., 2015), meaning that the more people perceive themselves to be susceptible to a disease and the more severe the consequences of the disease are perceived, the more likely they are to change behavior to save themselves from a potential infection. It follows, that a person who perceives a high risk of catching a disease with presumably severe effects on personal health is more likely to avoid potentially infectious contacts than a person who does not.

Although risk perception is an individual characteristic and the resulting behavioral changes occur on the individual level, homophily regarding risk perception exerts effects beyond that. Homophily indicates the extent to which it is more likely for a social tie to exist between similar people than between dissimilar people. Causes for homophily can be either the demographic structure of a group (*baseline* homophily) or explicit preferences for similarity (*inbreeding* homophily; McPherson et al., 2001). Independent of what causes homophily to emerge, risk perceptions and health behavior often correlate with other properties facilitating homophily, such as age, gender, ethnicity,

educational level, or marital status (Bish and Michie, 2010). McPherson et al. (2001) argue that as a result of homophily transmission processes tend to be socially localized. Kitchovitch and Lì (2010), for example, showed that higher levels of risk avoidance in scale-free networks limits the disease to the area around highly connected nodes, while nodes with lower contact numbers remain unaffected.

## 2. The model

Although numerous studies were concerned with the effects of small-world properties (Badham and Stocker, 2010; Keeling, 1999; Miller, 2009; Watts and Strogatz, 1998) or homophily (d'Andrea et al., 2022; Kitchovitch and Lì, 2010; Koku and Felsher, 2020) on epidemics, there is, to our knowledge, no study that looks into their interplay. Furthermore, simulation studies typically consider static networks (Badham and Stocker, 2010; d'Andrea et al., 2022; Keeling, 1999; Kitchovitch and Lì, 2010; Miller, 2009; Watts and Strogatz, 1998). To study the combined effects of small-world properties and homophily on epidemics in adaptive social networks, we created the so-called *Small-Worlds Infectious Disease Model* (*SWIDM*), a specific model case of the *Networking during Infectious Diseases Model* (*NIDM*; Nunner et al., 2021).

To facilitate comparability, reproducibility, and understanding of our model we are guided by the updated version of the *ODD protocol* ("Overview", "Design concepts", "Details") of individual- and agent-based models (Grimm et al., 2010). Here, we offer mostly top-level descriptions of the model components relevant for this paper. In *Online appendix A – Additional model definitions*, we offer more detailed and formal model definitions.

### 2.1. Purpose

The SWIDM was developed to investigate how (i) social mixing of heterogeneous risk perception (random, homophilous), (ii) variations in small-world network properties (clustering, path length), and (iii) the possibility of network changes (static, adaptive) affect the transmission of infectious diseases in social networks.

### 2.2. Basic principles

The SWIDM is a network formation and disease spread model for a closed population of autonomous agents (nodes) that can create infection-relevant relationships (edges). Networking decisions are based on the trade-off between the social benefits (e.g., sense of belonging, affection), costs (e.g., time, effort, money), and (potential) harm of infections (e.g., symptoms, absence from work, hospitalization) that a tie has for an agent. Agents seek to myopically maximize personal utility by (i) forming ties to other agents (requires the consent of the opposite agent), (ii) maintaining existing ties to other agents, and (iii) dissolving ties to other agents (unilateral decision). Decisions on these actions are based on an agent's position in the network, individual risk perception related to the disease, and the distribution of disease states among agents in the network (see section Utility).

In their choice to create or maintain ties, agents have a preference for triadic closure versus openness in their networks. This preference balances two things: first, a preference for closed triads (Coleman, 1994), and second, the strategic advantage to bridge structural holes (Burt, 1992). Finally, the underlying agent selection process captures baseline homophily (McPherson et al., 2001) by implementing larger probabilities for changing ties with agents who have similar risk perceptions. This process does not integrate explicit preferences regarding risk perception of alters (inbreeding homophily) and is in line with findings that correlations between personal properties, such as risk perceptions and health behavior, are driven by higher likelihoods to meet similar others (e.g., regarding age, gender, ethnicity, educational level, marital status) (Bish and Michie, 2010; McPherson et al., 2001).

### 2.3. Dynamics of disease transmission and network formation

Disease dynamics and adaptive network dynamics are simulated in discrete time steps (for detailed pseudocode refer to *Online appendix A – Additional model definitions*, *Process overview and scheduling*). At the start of each time step, the simulation updates the epidemiological state of the population. That is, for each agent in random order, we first determine its disease state. If an agent is susceptible, the likelihood of the agents to become infected is:

$$\pi_i = 1 - (1-\gamma)^{t_{i_I}}, \tag{1}$$

with $\gamma$ being the probability to get infected per single contact and $t_{i_I}$ the number of infected neighbors. If an agent is infected and has been infected for a fixed number of time steps ($\tau$), the agent recovers. If an agent is recovered, the agent cannot get re-infected until the simulation run ends. Note that the random order ensures that all nodes get an equal chance to be processed first, as later processed agents may have an increased number of infectious neighbors than if they had been processed earlier.

Following the update of the population's epidemiological state, social network dynamics are computed. That is, for each agent $i$ in random order a set $J$ consisting of a fixed proportion ($\varphi$) of agents from the entire population ($N$) is selected. To realize the concept of *focused* interaction in an agent's social vicinity (Feld, 1981), $J$ is composed of neighbors at distance 1 and distance 2, as well as agents randomly selected from the entire network. Furthermore, to account for homophily, each single agent $j$ is selected either to be the most similar agent regarding risk perception ($r$) or an agent selected without regard to risk perception. More formally, we draw a random number $u$ from the uniform distribution $U(0,1)$ and compare this to the homogeneity parameter $\omega$ such that:

$$j = k, \text{with } k \begin{cases} \text{such that} |r_i - r_k| = \min \forall k' \in N^* |r_i - r_{k'}|, & \text{if } u \le \omega \\ \text{is a randomly selected element of } N^*, & \text{if } u > \omega \end{cases} \tag{2}$$

For each agent $j$ in $J$, $i$ dissolves an existing tie, if $i$'s utility (see section Utility) without tie $ij$ exceeds the utility with tie $ij$; or $i$ proposes a new tie to $j$, if $i$'s utility with tie $ij$ exceeds the utility without $ij$. Note that only if $j$ accepts the proposal ($j$'s utility with tie $ij$ also exceeds the utility without $ij$), a new tie $ij$ is formed.

### 2.4. Utility

Utility for an agent $i$ is defined in the SWIDM as the trade-off between the social benefits ($B_i$), social maintenance costs ($C_i$), and (potential) harm ($D_i$) of infections based on the network connections held by $i$:

$$U_i = B_i - C_i - D_i. \tag{3}$$

Social benefits are defined as the weighted sum of the benefits for ties (left summand) and the benefits for the proportion of closed triads (right summand):

$$B_i = b_1 \cdot t_i + b_2 \cdot \left(1 - 2 \cdot \frac{|x_i - \alpha|}{\max(\alpha, 1-\alpha)}\right), \tag{4}$$

with $x_i$ denoting the actual proportion of closed triads $i$ belongs to, $\alpha$ the preferred proportion of closed triads, and $t_i$ the number of ties agent $i$ possesses. Social maintenance costs are assumed to be quadratic in the number of ties $t_i$ to model increasing marginal costs of additional ties:

$$C_i = c_1 \cdot t_i + c_2 \cdot t_i^2. \tag{5}$$

The combination of benefits and costs for social ties $([b_1 \cdot t_i] - [c_1 \cdot t_i + c_2 \cdot t_i^2])$ allows us to control the number of ties the agents seek to establish apart from clustering considerations. This largely controls the degrees of the agents and thus the average degree of the network. Further, we control the degree of clustering in the network by

defining the proportion of closed triads agents seek to be a part of ($\alpha$).

(Potential) harm of infections is a combination of *perceived* probability to get infected ($p_i$) and *perceived* severity of the disease ($s_i$):

$$D_i = p_i \cdot s_i. \tag{6}$$

That is, agents transform actual probability to get infected $\pi_i$ (Equation (1)) into a subjective version of the same, depending on their disease state:

$$p_i = \begin{cases} \pi_i^{2-r}, & \text{if } i \text{ is susceptible,} \\ 1, & \text{if } i \text{ is infected,} \\ 0, & \text{if } i \text{ is recovered.} \end{cases} \tag{7}$$

Furthermore, agents transform actual severity of the disease ($\sigma$) into a subjective version of the same, depending on their disease state:

$$s_i = \begin{cases} \sigma^r, & \text{if } i \text{ is susceptible,} \\ \sigma, & \text{if } i \text{ is infected,} \\ 0, & \text{if } i \text{ is recovered.} \end{cases} \tag{8}$$

Consequently, agents with risk perception values ($r$) below 1 underestimate, while agents with risk perception values above 1 overestimate the probability of infections and disease severity. We can therefore control the severity ($\sigma$) and infectivity ($\gamma$, operationalized as transmission probability per contact and time step) of the disease, as well as the agents' perception of these parameters ($r$).

## 3. Simulation

### 3.1. Simulation procedure & output data

A simulation run is initialized by setting the model parameters (see Table 1). Thereupon, we simulate the social network dynamics in a disease free population until a *pairwise stable* network (Jackson and Wolinsky, 1996) emerges. That is a notion of equilibrium where no agent benefits from unilaterally breaking an existing tie and no pair of agents benefits from jointly creating a non-existing tie. This ensures that changes in network structure at the onset of an epidemic depend solely on the presence of an infectious disease. If this initialization phase did not lead to a pairwise stable and connected network within 50 time steps, we restarted the initialization until we had over 80,000 networks suitable for the next phase (80,305 to be precise).

The observation phase begins when a randomly selected agent gets infected (*index case*). The observation phase ends when no infected agents are left (every agent is either susceptible or recovered). Each observation phase is executed twice for each network and parameter setting. That is, one simulation without network dynamics and one simulation with network dynamics. Thus, both conditions start with the same network structure and index case. The only difference is that agents in the static condition cannot change their network ties, while agents in the adaptive condition can.

During the observation phase we keep track of all varied parameters and a range of outcome variables (see Table 1). On the agent level, we record the number of social ties ($t_i$), the proportion of closed triads ($x_i$), and the disease state ($d_i$). On the network level, we record the number of susceptible ($|\boldsymbol{S}|$), infected ($|\boldsymbol{I}|$), and recovered agents ($|\boldsymbol{R}|$), the number of broken ($t_G^-$) and created ties ($t_G^+$), network clustering ($\mathcal{C}_G$), average path length ($\mathcal{L}_G$), and homophily ($\mathcal{H}_G$). Furthermore, we keep track of average degree ($\mathcal{D}_G$), as it is known to be a significant factor for diffusion processes even for small variations.

Network clustering is defined as the average of all agents' local clustering coefficients, while average path length consists of the average of all shortest paths between all pairs of agents in the network (Watts and Strogatz, 1998). Following the concept of quantifying degree-based assortative mixing by Newman (2002), we quantify homophily with the Pearson correlation coefficient of risk perception between all pairs of tied agents.

**Table 1**
State variables, scales, and settings for model parameters.

| State variable | Scale | Setting |
|---|---|---|
| **I.I. Agent, parameters** | | |
| Benefit per social tie[a] | $b_1 \in \mathbb{R}_0^+$ | $b_1 = 1.0$ |
| Benefit for triadic closure | $b_2 \in \mathbb{R}_0^+$ | $b_2 = 0.5$ |
| Preferred proportion of closed triads[b] | $0 \leq \alpha \leq 1$ | $\alpha \sim U[0, 1]$ |
| Simple cost per tie[a] | $c_1 \in \mathbb{R}_0^+$ | $c_1 = 0.2$ |
| Marginal cost per tie[a] | $c_2 \in \mathbb{R}_0^+$ | $c_2 = 0.05$ |
| Risk perception | $0 \leq r \leq 2$ | $r \sim U[r_{\min}, r_{\max}]$ |
| **I.II. Agent, outcomes** | | |
| Number of social ties | $t_i \in \mathbb{N}_0$ | |
| Proportion closed triads | $0 \leq x \leq 1$ | |
| Disease state | $d \in \{S, I, R\}$ | |
| **II.I. Network, model parameters** | | |
| Number of agents | $N \in \mathbb{N}_0$ | $N = 80$ |
| Minimum risk perception | $0 \leq r_{\min} < 2$ | $r_{\min} \sim U[0, 1]$ |
| Maximum risk perception | $0 < r_{\max} \leq 2$ | $r_{\max} \sim U[1, 2]$ |
| Likelihood of ties similar in risk perception[c] | $0 \leq \omega \leq 1$ | $\omega \sim U[0, 1]$ |
| **II.II. Network, outcomes** | | |
| Number of susceptible agents | $\lvert\boldsymbol{S}\rvert \in \mathbb{N}_0$ | |
| Number of infected agents | $\lvert\boldsymbol{I}\rvert \in \mathbb{N}_0$ | |
| Number of recovered agents | $\lvert\boldsymbol{R}\rvert \in \mathbb{N}_0$ | |
| Number of broken ties | $t_G^- \in \mathbb{N}_0$ | |
| Number of created ties | $t_G^+ \in \mathbb{N}_0$ | |
| Clustering | $0 \leq \mathcal{C}_G \leq 1$ | |
| Average path length | $\mathcal{L}_G \in \mathbb{R}_0^+$ | |
| Homophily | $-1 \leq \mathcal{H}_G \leq 1$ | |
| Average degree | $\mathcal{D}_G \in \mathbb{R}_0^+$ | |
| **III. Infectious disease, model parameters** | | |
| Disease severity | $\sigma > 1$ | $\sigma \sim U(1, 100]$ |
| Infectivity[d] | $0 \leq \gamma \leq 1$ | $\gamma \sim U[0.01, 0.20]$ |
| Recovery time in time steps | $\tau > 0$ | $\tau = 5$ |

[a] The combination of $b_1 = 1.0$, $c_1 = 0.2$, and $c_2 = 0.05$ sets the preferred number of ties to 8.

[b] The effect of $\alpha$ on network clustering and average path length is depicted in Figure *B.1* in *Online appendix B – Additional analyses*.

[c] The effect of $\omega$ on homophily is depicted in Figure *B.1* in *Online appendix B – Additional analyses*.

[d] Infectivity is operationalized as transmission probability per contact and time step.

### 3.2. Parameter settings

We fixed the population size to $N = 80$, the preferred number of ties per agent to 8 (realized through the combination of $b_1 = 1.0$, $c_1 = 0.2$, and $c_2 = 0.05$; see section Utility for details), the number of agents to be evaluated per time step to 16 ($\varphi = 0.2$), and the recovery time to $\tau = 5$ time steps. These parameters constitute a relevant and interesting framework for our further studies. That is, keeping average degree constant, and thus creating a homogeneous degree distribution, enables us to minimize the otherwise strong effect of degree on epidemic dynamics (e.g., Danon et al., 2011; Nunner et al., 2021), and thus to study the network properties of interest (clustering, average path length) in isolation. Additionally, empirical studies support an average of 8 contacts per day being a relevant magnitude for networks of respiratory disease spread (Danon et al., 2013; Leung et al., 2017), while 5 days recovery time is within the average range for respiratory diseases, like influenza (Longini et al., 2005), and people typically have between 5 and 30 contacts relevant for respiratory disease transmissions per day (Mossong et al., 2008). Finally, we opted for a constant homogeneous setting to minimize sources of variation, while test simulations have shown that these settings produce interesting dynamics.

Submodels are realized through randomized initial settings of the remaining model parameters. This allows us to generate epidemics on a whole variety of different scenarios and to study how sensitive the outcomes are to changes in parameter settings. Variations in clustering and average path length are realized through uniform random samples

of preferred proportion of closed triads ($\alpha \sim U[0, 1]$). Note that the joint variation of clustering and path length is key in the collective dynamics of small-world networks (Watts and Strogatz, 1998). Uniform random samples of likelihood of ties between agents that are similar in risk perception ($\omega \sim U[0, 1]$) are used to realize variations in homophily. Further, we set randomized bounds for risk perception on the network level ($r_{\min} \sim U[0, 1]$, $r_{\max} \sim U[1, 2]$) and assign randomized values within the previously set bounds to each agent ($r_i \sim U[r_{\min}, r_{\max}]$) to realize variations of risk perception on the network and individual level. Finally, we vary disease severity ($\sigma \sim U(1, 100]$) and infectivity ($\gamma \sim U[0.01, 0.20]$).

Note that, although we choose plausible settings supported by empirical examples, we do not study a specific disease in a specific context but consider parameter variations that contribute to answering our specific research questions. Please refer to Nunner et al. (2021) for detailed elaborations on the model's theoretical foundations, its functional form, and additional sensitivity analyses.

### 3.3. Data and source code availability

The data, the Java 8 source code to generate the data (including an executable program and an easy to use graphical user interface), and the R scripts to analyze the data during the current study are available under the GPLv3 license in the GitHub repository, https://github.com/hnunner/nidm-simulation.

## 4. Analysis

To investigate the effect of small-world properties and health behavior homophily on epidemics, we divide the variables into two categories. First, independent variables describing network and disease properties: (i) network clustering, (ii) average path length, (iii) homophily, (iv) average degree, (v) disease severity, (vi) infectivity, and (vii) average risk perception. Note that the listed network properties are outcome variables resulting from the network formation process, but independent variables in the analysis of epidemics. Second, dependent variables characterizing epidemics and network dynamics: (i) the percentage of infected agents throughout the entire epidemic (*final size*), (ii) the number of time steps from first infection until all agents are either susceptible or recovered (*duration*), (iii) the maximum number of simultaneously infected agents (*epidemic peak size*), and (iv) the number of network changes during the epidemic (number of broken and created ties).

We use box-and-whisker plots to compare dependent variables between adaptive and static networks. Box-and-whisker show the median, interquartile range (*IQR*: $[Q_1, Q_3]$), minimum (at most $Q1 - 1.5 \cdot IQR$), maximum (at most $Q3 + 1.5 \cdot IQR$), and outliers. We use Kendall rank correlation coefficients to describe the relationship between non-normally distributed variables (i.e., final size - duration, final size - peak size).

Bivariate effects are analyzed visually using scatter plots with independent variables on the x-axis and dependent variables on the y-axis. Due to the large number of simulations, we sort the independent variables and bin them into 100 points on the x-axis, thus, showing the mean of $\frac{80{,}305}{100} \approx 803$ simulations per point. Standard errors and 95% confidence intervals are omitted, as they hardly exceed the size of the dots in every plot.

Finally, we perform regression analyses to study model behavior in a multivariate manner. We realize that significance tests do not have the conventional interpretation because we do not have a usual sample. Tests should be interpreted more descriptively and are used to find the best fitting models given the sample of model parameters drawn (cf. Buskens and Yamaguchi, 1999; Buskens and Snijders, 2016). We use models to linearly approximate the logit of final size by combinations of the independent variables, since final size describes the percentage of a binary response (infected/not infected) (Long, 1997). We use linear models for duration, peak size, and number of network changes. We first create models for each dependent variable and the main effects for all standardized independent variables. Second, we create interaction effects starting from all possible interactions. We then remove all effects that are not significant at $p < 0.001$ or do not contribute to reduce unexplained variance at $R^2 \geq 0.001$. For highly collinear parameters (e.g., clustering and path length) we select parameters to minimize collinearity, while maximizing $R^2$. We use an inductive approach rather than testing theoretically informed interaction effects to get a more complete picture of the model behavior and not to miss potentially counterintuitive interactions.

## 5. Results & discussion

In the following, we discuss the results of our simulations starting with a comparison of epidemics in adaptive and static networks. Thereafter, we discuss how health behavior homophily affects epidemics in adaptive and small-world networks and conclude with how these dynamics differ in static networks. Additional analyses can be found in *Online appendix B – Additional analyses*, where we show, among others, that the effect of degree on final size is negligible (*Effect of degree on final size*) and can therefore be omitted for further analyses.

### 5.1. Epidemics in adaptive vs. static networks

While in the majority of simulated epidemics almost all agents get infected in the static networks ($Mdn = 97.50\%$), only about one fifth of the agents get infected in the adaptive networks ($Mdn = 21.25\%$). Because duration and peak size are correlated with final size (final size - duration: $r_\tau = 0.61$, $p < 0.05$; final size - peak size: $r_\tau = 0.89$, $p < 0.05$), epidemics in static networks are on average longer (static networks: $Mdn = 17$ time steps; adaptive networks: $Mdn = 15$ time steps) and have higher maximum numbers of simultaneously infected agents (static networks: $Mdn = 53.75\%$; adaptive networks: $Mdn = 11.25\%$). That is, few infected agents need on average less time to recover than many infected agents, given they do not get infected at the same time step. This relation, however, reverses for large final sizes in static networks (Fig. 1**b**, orange box). Here, epidemics take significantly less time than for medium final sizes (Fig. 1**b**, green box). In adaptive networks, on the other hand, the duration is largely independent of final size when final size is larger than 10%. Consequently, many infected agents increase the rate of transmission events and thus the speed of disease spread if agents cannot actively avoid infections. Adaptive agents in networks with large final sizes, on the other hand, manage to delay infections, thus slowing down the course of epidemic. As a result, static networks show not only higher epidemic peaks, but must also have more simultaneously infected nodes on average.

### 5.2. Effects of health behavior homophily on epidemics in adaptive small-world networks

Fig. 2 shows bivariate effects of network properties on epidemics in adaptive networks. The data reveal negative effects on final size for homophily, network clustering, and average path length. Furthermore,

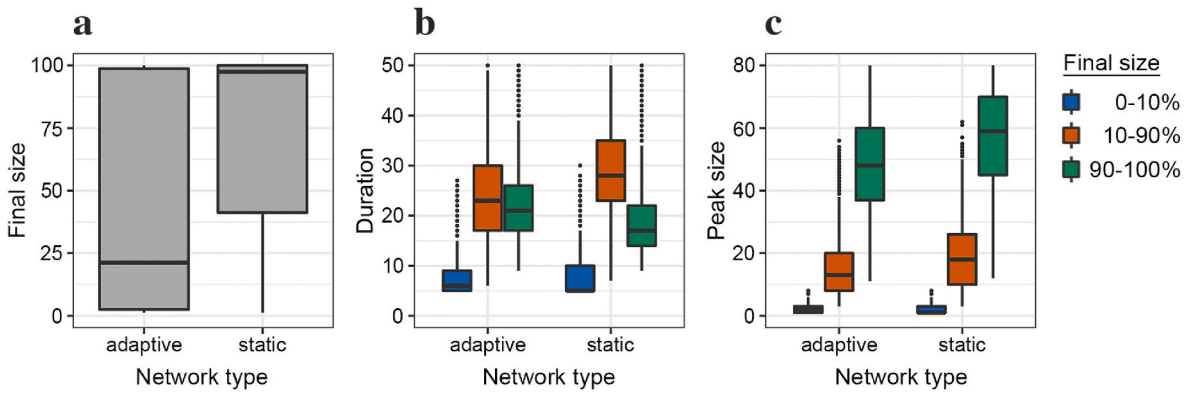


**Fig. 1. Comparison of epidemics in adaptive and static networks.** Each panel shows box-and-whisker plots divided by network type (left: adaptive, right: static). Panels show the average outcome for final size (a), duration (b), and peak size (c). Plots for duration and peak size are further divided into three groups depending on final size: low (0–10%; blue), medium (10–90%; orange), and high (90–100%; green). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

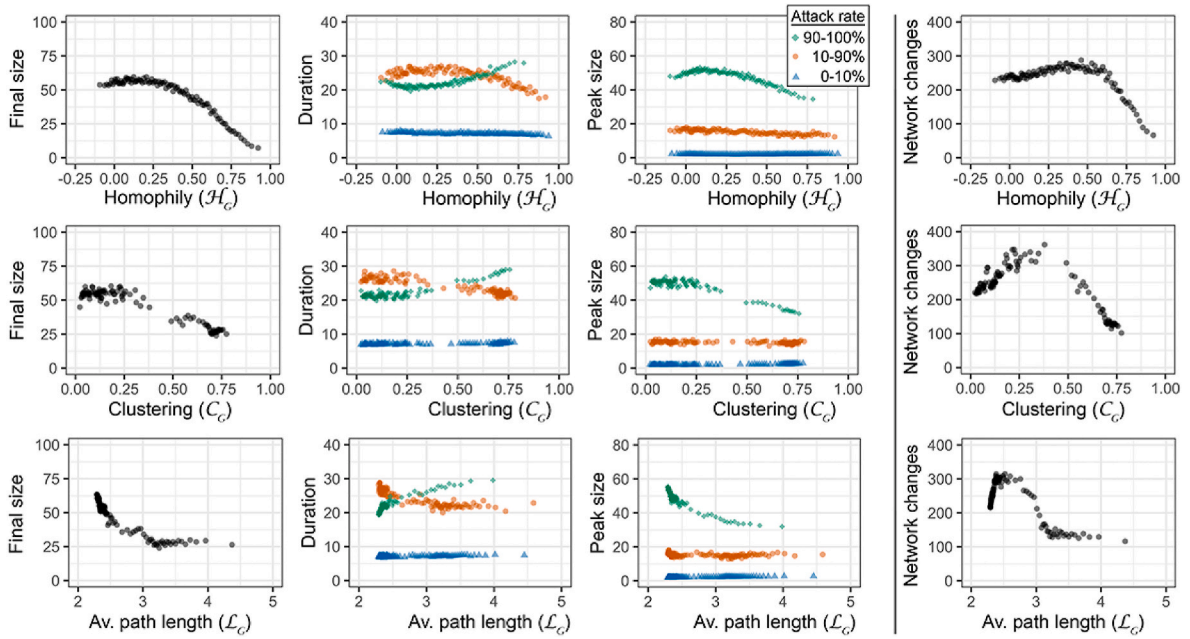


**Fig. 2. Bivariate effects of network properties on epidemics.** Plots show the effects of homophily (row 1), network clustering (row 2), and average path length (row 3) on final size (column 1), duration (column 2), and peak size (column 3) of epidemics, as well as number of network changes (column 4) in adaptive networks. There are 100 points per group and plot, with each point showing the average of $\approx 803$ $(\frac{80{,}305}{100})$ observations. Measures of variance are omitted, as they merely exceed the size of the dots. Data for duration and peak size have been grouped by final size: 0–10% (blue), 10–90% (orange), and 90–100% (green). Additional effects (average degree, average risk perception, disease severity, infectivity) and effects of all parameters on epidemics in static networks can be found in Figures B2 and *B.3* of *Online appendix B – Additional analyses*. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

final size is mostly stable at high levels in networks with low levels of homophily ($< 0.45$) and clustering ($< 0.45$). In contrast, final size drops off quickly for low levels of average path length and remains stable for higher levels of average path length ($> 3.0$). Interestingly, multivariate regression analyses show that the effect of average path length on final size vanishes when controlled for all other independent variables (Table 2).

The number of network changes helps to understand the complex dynamics in adaptive networks. That is, all independent variables have similar effects on final size and the number of network changes (Fig. 2, column 4). It follows that the further the disease spreads in the network, the more agents reposition themselves to avoid or delay infections. If we, however, control for all independent variables (Table 2), the number of network changes increases, although final size decreases as network clustering increases. That is because infections inside densely connected clusters require agents to dissolve comparatively many ties to distance

**Table 2**
Regression analysis summary.

| | Adaptive networks | | | | Static Networks | | |
|---|---|---|---|---|---|---|---|
| | Final size | Duration | Peak size | Network changes | Final size | Duration | Peak size |
| **I. Main effects** | | | | | | | |
| Number of network changes ($t_G^{+/-}$) | ⊕ | ⊕ | + ++ | | | | |
| Homophily ($\mathcal{H}_G$) | − −− | − | − −− | − | | | − |
| Clustering ($\mathcal{C}_G$) | − − | − | − −− | + + | − | + + + + | − |
| Average path length ($\mathcal{L}_G$) | | + + | | − −−− | − | + + | − − |
| Average risk perception ($r_{\sigma,\pi}$) | − | − | − | − | | | |
| Disease severity ($\sigma$) | − | − | − | | | | |
| Infectivity ($\gamma$) | + + + + | − − | ⊕ | ⊕ | ⊕ | − | ⊕ |
| **II. Interaction effects** | | | | | | | |
| $t_G^{+/-}$ x $\mathcal{H}_G$ | − − | | − | | | | |
| $t_G^{+/-}$ x $\mathcal{C}_G$ | − − | − − | − − | | | | |
| $t_G^{+/-}$ x $\mathcal{L}_G$ | | + ++ | | | | | |
| $t_G^{+/-}$ x $\gamma$ | + + | − − | + + | | | | |
| $\mathcal{H}_G$ x $\mathcal{C}_G$ | + | − | + | | | | |
| $\mathcal{H}_G$ x $\mathcal{L}_G$ | | | | − | | | |
| $\mathcal{H}_G$ x $\gamma$ | − − | | − − | + | | | |
| $\mathcal{L}_G$ x $\gamma$ | | | − − | − | − | ⊕ | − |
| Adjusted $R^2$ | 0.92 | 0.80 | 0.85 | 0.39 | 0.76 | 0.07 | 0.87 |
| Number of observations | 80,305 | 80,305 | 80,305 | 80,305 | 80,305 | 80,305 | 80,305 |

Note: all variables are standardized; all effects shown are significant at $p < 0.001$; effect direction is shown as sign (+ for positive effects, − for negative effects), while effect sizes are shown in relation to the largest effect (⊕/⊖): +/− = 0–20%, + +/− − = 20–40%, …, + + + ++/− −−−− = 80–100%). Individual regression models can be found in *Online appendix B – Additional analyses, Regression models*.

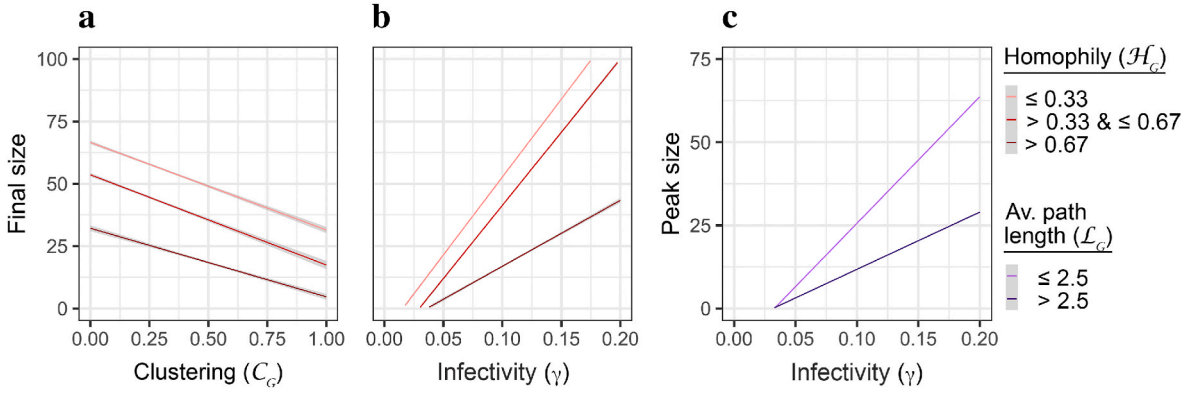


**Fig. 3. Interaction effects in adaptive networks.** Final size of epidemics decreases the more clustered the networks (**a**) and the lower the infectivity (**b**). At the same time, larger degrees of health behavior homophily (darker red lines) create lower numbers of infections. Epidemic peak size increases the higher the infectivity (**c**) and the shorter the average path length (light purple line). (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

themselves from the disease. Due to the low number of ties bridging any two clusters, however, it is less likely that the disease spreads through the entire network.

Furthermore, the interaction effect between homophily and clustering (Fig. 3**a**) reveals that the combination of high clustering and high health behavior homophily produces the lowest final sizes. As the negative effect of risk perception (in combination with disease severity and infectivity) on final size (Table 2) suggests, the less risk taking the agents are, the quicker they dissolve their ties to infectious neighbors. Consequently, a cluster of low risk taking agents is especially hard to infiltrate due to the quick dissolution of ties to comparatively few infectious neighbors outside the cluster. Consequently, the average epidemic in clustered networks is also shorter and has a lower peak size (Table 2). Considering final sizes of 90% and above (Fig. 2**a**, column 2, green points), however, the effect is inverted: If all agents get infected at some point, the disease requires more time to bridge between loosely connected clusters. Fig. 3 reveals further that even for highly infectious diseases, higher levels of homophily produce smaller final sizes (**b**) and larger average path lengths produce lower epidemic peaks (**c**).

Finally, we observe that epidemics with final sizes up to 90% were insensitive to variation in every independent variable (Fig. 2, column 3). Peak sizes for final sizes of more than 90%, however, follow the shape of the final size curves. That is, only if the disease spreads through almost the entire network, the number of simultaneously infected agents correlates with the final size of the epidemic.

### 5.2. *Effects of health behavior homophily on epidemics in static small-world networks*

Dynamics in static networks show several significant differences to their adaptive counterparts. First, there is no observable effect of homophily and risk perception on final size in static networks (Table 2). This is, of course, due to the operationalization of homophily in our study, which renders agents incapable to act upon their perceptions in static networks. Second, we observe a positive effect of clustering on epidemic duration in static networks (Table 2). That is, while the capability to cut off bridging ties quickly in adaptive networks causes epidemics to die out quickly (unless the disease spreads to the entire networks), the incapability to distance oneself from the disease leads to slow spread and on average larger final sizes in static networks. Third,

we observe negative effects of average path length on final size and peak size in static networks only (Table 2).

## 6. Conclusion & implications

Although the literature contains numerous publications on how diffusion processes are shaped by small-world properties and homophily individually (Badham and Stocker, 2010; d'Andrea et al., 2022; Keeling, 1999; Kitchovitch and Lì, 2010; Koku and Felsher, 2020; Miller, 2009; Watts and Strogatz, 1998), it is still unknown how the diffusion of infectious diseases is shaped by their combination. We therefore asked: *What are the effects of health behavior homophily on epidemics in adaptive small-world networks?* and *How do epidemics differ between adaptive and static networks?*

To answer these questions, we created the so-called *Small-Worlds Infectious Disease Model* (*SWIDM*), a specific model case of the *Networking during Infectious Diseases Model* (*NIDM*; Nunner et al., 2021). The SWIDM is a network formation model that considers 1. Subjective risk perception regarding infectious diseases, 2. The propensity of forming closed triads, and 3. The likelihood of forming ties between individuals similar in risk perception. That is, autonomous agents form, maintain, or break social ties based on the trade-off between the benefits and costs of social relations, and potential harm of infections. Agent-based simulations of epidemics enabled us to consider a wide range of infectious diseases (mild vs. severe, low vs. high infectivity), health behaviors (low risk taking vs. high risk taking), homophily (low vs. high chance of ties between agents similar in risk perception), and small-world networks (fixed average degree, low vs. high clustering, short vs. long average path length).

In line with previous studies (Leung et al., 2018), our results suggest that epidemics in static small-world networks are on average larger, thus longer, and have higher peaks compared to epidemics in adaptive networks. Furthermore, we see that the dynamics of disease transmission are heavily shaped by network properties. Take the effect of clustering on an epidemic, for example. While we confirm a negative effect of clustering on final size in both adaptive and static networks (Badham and Stocker, 2010; Keeling, 1999; Miller, 2009), we observe opposing effects of clustering on duration in adaptive (−) and static (+) networks. In both cases, the disease needs to cross one of the few available bridges to reach another cluster. Consequently, only a few ties need to be cut to isolate a cluster, causing the disease to die out quickly in the adaptive networks. In the static networks, however, intrusion of a cluster is more likely, while spreading across one of the few bridges slows down disease spread.

Furthermore, our results suggest that health behavior homophily in adaptive networks may cause the epidemic to end earlier than compared to randomly mixed networks. That is, not only the properties of individual actors but also their composition in heterogeneous networks affect disease spread. Consider, for example, a risk averse ego connected to an infectious alter. The preference to avoid the infectious alter lowers not only the probability to get infected for the ego but also the probability for all of its susceptible connections. Consequently, even the connections that are less risk averse benefit from the ego's decision to isolate from infectious alters. Additionally, our results suggest that the combination of network and actor properties shape disease spread in combination. That is, on the one hand, a cluster composed of mostly risk averse agents increases the chance for actors bridging two clusters to isolate the entire cluster from infectious alters. On the other hand, a cluster containing only a few risk averse agents increases the chance for actors bridging two clusters to maintain ties to infected alters from outside the cluster. This in turn increases the probability for the disease to spread into the cluster. Consequently, the overall large impact of homophilous clusters on epidemic size results from a comparably large chance of entire clusters being isolated from the disease. In contrast to multiple clusters, isolation from infections in networks consisting of a single large cluster requires more ties to be cut and therefore such networks show a higher likelihood for the disease to spread to any part of the network (see also Watts and Strogatz, 1998).

These results are good news for two reasons. First, they support the notion that behavioral adaptation to health risks is a natural mechanism to not only reduce personal health risks, but also to mitigate disease spread. Dönges et al. (2021), for example, used a similar mechanism in a simulation study, concluding that non-pharmaceutical measures ought to leave sufficient room to exercise self-imposed measures resulting from risk perception, as strict measures of social distancing may result in rebound or compensatory effects when measures are lifted. Furthermore, social distancing may reinforce the mitigating effect of homophily. That is, limiting contacts needs to undergo an active selection process after which contacts with similar traits are more likely to remain (McPherson et al., 2001). Yamaguchi (1990), for example, showed that people with fewer friends show a stronger preference for similarity, such as regarding educational level, a trait that coincides with health behavior and risk perceptions (Bish and Michie, 2010). Second, we can use this mechanism to design targeted information campaigns addressing groups with lower risk perception to increase awareness of health risks. A remaining challenge is how to determine such groups. A promising way could be to find suitable proxies for scenarios of interest. From the literature, we know that risk perception often coincide with other personal characteristics, such as age, gender, ethnicity, educational level, or marital status (Bish and Michie, 2010). More specifically, a recent study among heterosexuals at high risk for HIV infection used individual and network data to determine predictors of self-perceived HIV risk (Koku and Felsher, 2020). The results show that men, individuals with lower education, and individuals in ethnically heterogeneous groups are on average more likely to perceive high risks.

Our study compares to earlier work in a few ways. Take Watts and Strogatz (1998), for example, who investigated the effects of clustering and path length on disease spread in static networks. Our findings also indicate that epidemic size decreases and epidemic duration increases with increasing values for clustering and average path length in static networks. In adaptive networks, however, we cannot confirm the effect of average path length on final size (when controlled for all other independent variables). As described earlier, this is a result of only a few bridges needing to be cut, thus causing the disease to die out quickly. Just as suggested by (Kitchovitch and Lì, 2010), we observe that populations with on average less risk taking agents produce epidemics with smaller final size, shorter duration, and lower peak size. However, we considered homogeneous degree distributions, while Kitchovitch and Lì (2010) used scale-free networks.

It is important to note that our study comes with a few limitations. First, it is not in and of itself reasonable to assume that agents know the disease state of others before creating social ties. Even at the time of evaluating a tie, a disease may be transmitted, affecting final size. Second, although research on COVID-19 has shown that infections are primarily transmitted through social network ties (e.g., Li et al., 2020; Shen et al., 2020), it is important to consider transmissions that are not based on cost-benefit considerations (e.g., airborne diseases on public transport). This could be done in the form of a random component for spontaneous infections. Third, we only consider personal properties as a factor of risk perceptions. That is, agents are initialized to be either more or less risk averse throughout the entire course of a simulation run. Risk perceptions, however, may change according to changes in the prevalence of a disease or exchange of opinions with friends, which in turn may affect individual decision-making and thus disease spread (Lau et al., 2005; Teslya et al., 2022). To account for such effects, our model could be extended so that risk perception reflects a personal tendency (base value) that is influenced by current events (e.g., variation according to prevalence). Fourth, in our model social relations are not "paused" but tried to be replaced. Although in some contexts this is unlikely to happen (e.g., families), in others it is to be expected (e.g., friendships break because one friend is not supportive in times of need). Fifth, the operationalization of homophily (in terms of tendency for

social distancing) does not allow effects on epidemics in static networks. Other operationalizations, however, might be interesting extensions for future work.

To improve the model, other health behaviors lowering infectivity (e.g., handwashing, mask wearing) may be added. This would, most likely, also create an observable effect of health behavior homophily in static networks. Another approach to consider the effect of individual differences is to integrate a specific utility component expressing the preference to connect to others alike (*inbreeding* homophily). This has been neglected in the current work to ensure a better control of fixed average degree, and thus the isolated study of network components of interest (clustering, average path length). Furthermore, although our results show the general effect of health behavior homophily on a generic disease, a model could be informative for a specific scenario by fitting parameters to empirical data (e.g., disease severity, infectivity, degree distribution, clustering, recovery times) and refining compartments. Additionally, such an approach would allow considering additional factors that may affect individual susceptibility, treatment, and recovery, such as age, socio-economic status, or capacities of regional healthcare services. Finally, while the theories our model is based on have empirical grounding, our theoretical findings need to be corroborated by empirical experiments. This could be done by creating an experimental game-like study, in which human subjects take the role of the agents in our simulations. Another option is to fit the model to a specific scenario and study how well estimations fit to ongoing observations.

In conclusion, our model suggests that a combination of high network clustering and high health behavior homophily can mitigate the spread of infectious diseases. That is, adaptive agents can cause an epidemic to die out quickly by severing only a few bridging ties. The more risk averse a cluster is on average, the more likely bridging ties are severed, and the more likely it is that the entire cluster is isolated from the disease. Although clusters slow down disease spread in static networks, infections can still reach any part of the network, as agents are rendered incapable of reacting to their risk perceptions. Neglect of these fundamentally different dynamics may, therefore, lead to inaccurate estimations when applied to a specific case. Considering the adaptive dynamics, however, can support the design of non-pharmaceutical interventions such as targeted information campaigns using proxies for risk perception.

## Author contributions

HN and VB developed the mathematical model; MK and AT contributed to its conceptualization. HN programmed the simulation, performed analyses, created visualizations, took care of data curation, wrote the original draft of the manuscript, and integrated the feedback of all co-authors. All authors reviewed and discussed the manuscript repeatedly. MK and VB acquired funding and supervised the project.

## Submission declaration

The submitted manuscript has not been published and is not under consideration for publication elsewhere. All authors approved the manuscript for submission. If accepted, the manuscript will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder.

## Declaration of competing interest

None.

## Acknowledgments

This research was supported by grants from the Dutch organization for health research and care innovation (ZonMw, grant number 91216062). We thank our colleagues from the Sociology department, the Infectious Disease Modeling Group, and the Centre for Complex Systems Studies (CCSS) at Utrecht University for comments that greatly improved the manuscript.

## Appendices A and B. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.socscimed.2022.115350.

## References

Badham, J., Stocker, R., 2010. The impact of network clustering and assortativity on epidemic behaviour. Theor. Popul. Biol. 77 (1), 71–75.
Bansal, S., Read, J., Pourbohloul, B., Meyers, L.A., 2010. The dynamic nature of contact networks in infectious disease epidemiology. J. Biol. Dynam. 4, 478–489.
Bedson, J., Skrip, L.A., Pedi, D., et al., 2021. A review and agenda for integrated disease models including social and behavioural factors. Nat. Human Behav. 5 (7), 834–846.
Bish, A., Michie, S., 2010. Demographic and attitudinal determinants of protective behaviours during a pandemic: a review. Br. J. Health Psychol. 15 (4), 797–824.
Breban, R., 2011. Health newscasts for increasing influenza vaccination coverage: an inductive reasoning game approach. PLoS One 6 (12) e28,300.
Bults, M., Beaujean, D.J., Richardus, J.H., Voeten, H.A., 2015. Perceptions and behavioral responses of the general public during the 2009 influenza A (H1N1) pandemic: a systematic review. Disaster Med. Public Health Prep. 9 (2), 207–219.
Burt, R.S., 1992. Structural Holes: the Social Structure of Competition.
Buskens, V., Snijders, C., 2016. Effects of network characteristics on reaching the payoff-dominant equilibrium in coordination games: a simulation study. Dynam. Games Appl. 6 (4), 477–494.
Buskens, V., Yamaguchi, K., 1999. A new model for information diffusion in heterogeneous social networks. Socio. Methodol. 29 (1), 281–325.
Coleman, J.S., 1994. Foundations of Social Theory. Harvard university press.
d'Andrea, V., Gallotti, R., Castaldo, N., Domenico, M.D., 2022. Individual risk perception and empirical social structures shape the dynamics of infectious disease outbreaks. PLoS Comput. Biol. 18 e1009,760.
Danon, L., Ford, A.P., House, T., et al., 2011. Networks and the epidemiology of infectious disease. Interdiscipl. Perspect. Infect. Dis. 2011.
Danon, L., Read, J.M., House, T.A., et al., 2013. Social encounter networks: characterizing Great Britain. Proc. Biol. Sci. 280 (1765), 20131, 037.
Dönges, P., Wagner, J., Contreras, S., et al., 2021. Interplay between risk perception, behaviour, and covid-19 spread. Front. Phys. 68, 0.
Feld, S.L., 1981. The focused organization of social ties. Am. J. Sociol. 86, 1015–1035.
Funk, S., Gilad, E., Watkins, C., Jansen, V.A., 2009. The spread of awareness and its impact on epidemic outbreaks. Proc. Natl. Acad. Sci. U.S.A. 106 (16), 6872–6877.
Funk, S., Salathé, M., Jansen, V.A., 2010. Modelling the influence of human behaviour on the spread of infectious diseases: a review. J. R. Soc. Interface 7 (50), 1247–1256.
Granovetter, M.S., 1977. The strength of weak ties. In: Social Networks. Elsevier, pp. 347–367.
Grimm, V., Berger, U., Deangelis, D.L., et al., 2010. The ODD protocol: a review and first update. Ecol. Model. 221, 2760–2768.
Gross, T., D'Lima, C.J., Blasius, B., 2006. Epidemic dynamics on an adaptive network. Phys. Rev. Lett. 96, 208,701.
Jackson, M.O., 2008. Social and Economic Networks, vol. 3. Princeton University Press, Princeton and Oxford.
Jackson, M.O., Wolinsky, A., 1996. A strategic model of social and economic networks. J. Econ. Theor. 71 (1), 44–74.
Jones, J.H., Salathé, M., 2009. Early assessment of anxiety and behavioral response to novel swine-origin influenza A(H1N1). PLoS One 4 (12), e8032.
Kasl, S.V., Cobb, S., 1966. Health behavior, illness behavior and sick role behavior: I. health and illness behavior. Arch. Environ. Health 12 (2), 246–266.
Keeling, M.J., 1999. The effects of local spatial structure on epidemiological invasions. Proc. R. Soc. Lond. Ser. B Biol. Sci. 266, 859–867.
Kitchovitch, S., Lì, P., 2010. Risk perception and disease spread on social networks. Procedia Comput. Sci. 1–10, 00.
Klovdahl, A.S., 1985. Social networks and the spread of infectious diseases: the aids example. Soc. Sci. Med. 21 (11), 1203–1216.
Koku, E., Felsher, M., 2020. The effect of social networks and social constructions on hiv risk perceptions. AIDS Behav. 24, 206–221.
Lau, J.T., Yang, X., Tsui, H.Y., Kim, J.H., 2005. Impacts of sars on health-seeking behaviors in general population in Hong Kong. Prev. Med. 41, 454–462.
Leppin, A., Aro, A.R., 2009. Risk perceptions related to SARS and avian influenza: theoretical foundations of current empirical research. Int. J. Behav. Med. 16 (1), 7–29.
Leung, K., Jit, M., Lau, E.H., Wu, J.T., 2017. Social contact patterns relevant to the spread of respiratory infectious diseases in Hong Kong. Sci. Rep. 7 (1), 4–8.
Leung, K.Y., Ball, F., Sirl, D., Britton, T., 2018. Individual preventive social distancing during an epidemic may have negative population-level outcomes. J. R. Soc. Interface 15 (145), 20180,296.
Li, Q., Guan, X., Wu, P., et al., 2020. Early transmission dynamics in wuhan, China, of novel coronavirus–infected pneumonia. N. Engl. J. Med. 382 (13), 1199–1207 pMID: 31995857.

Long, S.J., 1997. Regression Models for Categorical and Limited Dependent Variables. SAGE Publications, Inc., Thousand Oaks, California.
Longini Jr., I.M., Nizam, A., Xu, S., et al., 2005. Containing pandemic influenza at the source. Science 309 (5737), 1083–1087.
Mao, L., Yang, Y., 2012. Coupling infectious diseases, human preventive behavior, and networks – a conceptual framework for epidemic modeling. Soc. Sci. Med. 74, 167–175.
May, R.M., Anderson, R.M., 1987. Transmission dynamics of hiv infection. Nature 326 (6109), 137–142.
McPherson, M., Smith-Lovin, L., Cook, J.M., 2001. Birds of a feather: homophily in social networks. Annu. Rev. Sociol. 27 (1), 415–444.
Milgram, S., 1967. The small-world problem. Psychol. Today 1 (1), 61–67.
Miller, J.C., 2009. Percolation and epidemics in random clustered networks. Phys. Rev. E - Stat. Nonlinear Soft Matter Phys. 80, 020,901.
Mossong, J., Hens, N., Jit, M., et al., 2008. Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Med. 5, 0381–0391.
Newman, M.E., 2002. Assortative mixing in networks. Phys. Rev. Lett. 89 (20), 208,701.
Nunner, H., Buskens, V., Kretzschmar, M., 2021. A model for the co-evolution of dynamic social networks and infectious disease dynamics. Comput. Soc. Network 8, 19.
Nunner, H., van de Rijt, A., Buskens, V., 2022. Prioritizing high-contact occupations raises effectiveness of vaccination campaigns. Sci. Rep. 12 (1), 1–13.
Nyabadza, F., Chiyaka, C., Mukandavire, Z., Hove-Musekwa, S.D., 2010. Analysis of an hiv/aids model with public-health information campaigns and individual withdrawal. J. Biol. Syst. 18, 357–375.
Pastor-Satorras, R., Castellano, C., Van Mieghem, P., Vespignani, A., 2015. Epidemic processes in complex networks. Rev. Mod. Phys. 87 (3), 925.
Risau-Gusman, S., Zanette, D.H., 2009. Contact switching as a control strategy for epidemic outbreaks. J. Theor. Biol. 257, 52–60.
Shen, Y., Xu, W., Li, C., et al., 2020. A cluster of novel coronavirus disease 2019 infections indicating person-to-person transmission among casual contacts from social gatherings: an outbreak case-contact investigation. Open Forum Infect. Dis. 7 (6), ofaa231.
Simmel, G., 1950. The Sociology of Georg Simmel, vol. 92892. Simon and Schuster.
Teslya, A., Nunner, H., Buskens, V., Kretzschmar, M.E., 2022. The effect of competition between health opinions on epidemic dynamics. medRxiv.
Tracy, C.S., Rea, E., Upshur, R.E., 2009. Public perceptions of quarantine: community-based telephone survey following an infectious disease outbreak. BMC Publ. Health 9 (1), 470.
Verelst, F., Willem, L., Beutels, P., 2016. Behavioural change models for infectious disease transmission: a systematic review (2010-2015). J. R. Soc. Interface 13, 20160,820.
Watts, D.J., Strogatz, S.H., 1998. Collective dynamics of 'small-world' networks. Nature 393 (June), 440–442.
WHO, 2021. Coronavirus disease (covid-19): how is it transmitted? URL. https://www.who.int/news-room/questions-and-answers/item/coronavirus-disease-covid-19-how-is-it-transmitted.
Willem, L., Verelst, F., Bilcke, J., et al., 2017. Lessons from a decade of individual-based models for infectious disease transmission: a systematic review (2006-2015). BMC Infect. Dis. 17, 612.
Yamaguchi, K., 1990. Homophily and social distance in the choice of multiple friends: an analysis based on conditionally symmetric log-bilinear association model. J. Am. Stat. Assoc. 85 (410), 356–366.