Journal of Combinatorial Mathematics
and Combinatorial Computing
www.combinatorialpress.com/jcmcc

Infectious disease prevention strategies for social health: a reinforcement learning-based assessment of the effectiveness of social capital investments

Li Jin[1,✉]

[1] *Shanghai Xinnasongshan Biopharmaceutical Technology Co., Ltd., Qingdao, Shandong, 266100, China*

ABSTRACT

Disease prevention has always had an important impact on the development of human life and health. The integration of complex network theory and disease has become one of the major trends in epidemiologic research. However, aspects such as individual vaccination behavior and vaccination costs are affected by social capital investment. Based on this, the article investigates a reinforcement learning model of social capital investment on disease prevention. Based on the mechanism of infectious disease dynamics on complex networks, the article investigates the Markov decision process and composition of the reinforcement learning model, and utilizes the theory related to infectious disease dynamics and reinforcement learning to study the mechanism of voluntary vaccination based on epidemic perception. It was found that when the ratio of two kinds of investment (partial investment and full investment) reaches the set maximum value, the full investment policy of targeting selection is more effective in reducing the scale of disease infection in the whole social network and reducing the total social cost, followed by partial investment, and the full investment policy of random selection brings the smallest effect. However, the results may differ for different population investment ratios and partial investment ratios, and both the full investment policy and partial investment policy can effectively control disease prevention, which is conducive to the healthy and prosperous development of the whole society.

Keywords: SIR model, social capital investment, disease prevention, reinforcement learning model

1. Introduction

The community belongs to the population density point, with the characteristics of large population flow and complex types of households, there are many health risks within the community environment, especially with the advancement of the aging society and the increase of middle-aged and old-aged

✉ Corresponding author.
E-mail address: MA2024shannon@163.com (L. Jin).
Received 20 February 2024; Revised 15 April 2024; Accepted 20 November 2024; Published Online 15 April 2025.
DOI: 10.61091/jcmcc127a-215

population, the problem of health risks is gradually obvious [1-3]. In order to minimize health risks, it is necessary to utilize information technology for community health management design work in the modern perspective to systematically provide health management services and create a good community environment.

With the aging of the population, economic globalization, and the rising prevalence level of chronic disease risk factors, the number of disease patients is increasing year by year, and for the country, the burden caused by disease is serious [4-5]. China has a huge number of disease patients, but the disease community health management content is numerous, the workload of the primary health care force is very weak, and there is a great need for a new and effective way of disease management [6-9]. Social capital reflects the relationship between the individual and the community, which is not only the embodiment of the social relationship network, but also includes the degree of individual recognition and integration of collective and social relationships [10-12]. The resources brought to the subject by the social relationship network linked behind social capital are multidimensional, which can help community residents enjoy a more convenient life, save the cost of living, provide richer employment opportunities, access to more sources of income channels, and exposure to a wider range of information [13-15]. In this way, the introduction of social capital investment in community health management can help community residents significantly improve disease prevention.

Literature [16] examined social capital in public health services to influence population health interventions and assessed the extent and role of social capital interventions in health at the individual and community levels. Literature [17] investigated the relationship between social capital and health in different domains and found that various types of social capital were significantly correlated to health outcomes such as death and illness, i.e., social capital can significantly contribute to the level of health. Literature [18] analyzed social capital and its components such as social support and social participation in older adult communities and assessed that social capital interventions showed mixed effects on on quality of life, well-being and self-perceived health. Literature [19] introduced the concept of anti-viral social capital, where community residents with strong and extensive social capital outperformed others in terms of behavioral responses and quality of life under physical isolation conditions during the COVID-19 pandemic. Literature [20] applied the behavior change and care coordination models to construct a social capital intervention model that included the creation of a network linking community health centers and various community assets, as well as a group support program to promote social support and participation among older adults, aimed at alleviating depressive symptoms and disease incidence levels among older adults. Literature [21] suggests a strong correlation between social cohesion and better health and health behaviors, further exploring the social cohesion perspective as an important factor influencing the use of preventive health care services among community residents in order to improve the quality of life of people in the current society.

The article first introduces the basics of infectious disease and SIR dynamics, and investigates the effect of two parameters, the infection rate and and inoculation rate, on the transmission process. The kinetic equations of propagation can be established by mean-field method in fully connected networks, and the propagation process is mainly stochastically simulated by Gillespie's algorithm in other complex networks. Then the basic concepts of reinforcement learning such as definition, general process and objective function are elaborated. On this basis, the article proposes a model of disease prevention mechanism based on the perception of social capital investment, where individuals perceive the risk of infection based on different local or global investment information. In the experimental testing part, the reinforcement learning model proposed in this article is quizzed from two perspectives: loss-based and rule-based. After that, the effects of partial and full investment strategies on disease prevention are explored through computer simulation.

2. Method

2.1. Infectious disease dynamics on complex networks

2.1.1. Overview of Infectious Diseases and Introduction to the SIR Model. Infectious diseases in the most widely studied when the SIR model and SIS model, which SIR model can be used to describe the infected person in the recovery of the infectious disease transmission process to obtain a certain degree of immunity, such as a computer infected with a virus after the virus may be installed antivirus software for the virus to be free of killing and real-time monitoring to effectively prevent re-infection. We focus on the SIR model [22]. In this model, individuals in a group are in one of three states: susceptible (S), infected (I), and recovered (R). They have the following behavioral characteristics:

An infected individual is an individual who has already been infected with the infectious disease and will infect susceptible individuals, causing them to become sick and recovering themselves over time to become recovered individuals.

Susceptible state individuals are individuals who have not yet been infected during the current round of transmission, and it may be possible for them to be infected by the infectious state during the current round of transmission and become infected state individuals.

Recovered individuals are individuals who have recovered from the infectious state, and because they have acquired antiviral antibodies in their bodies after recovery, they are no longer infected during the current round of transmission, and, of course, they will not infect other individuals.

2.1.2. SIR Modeling in Fully Connected Networks. In a fully connected network, the propagation dynamics behavior can be described by differential equations approximated by mean field theory. Let the infection rate be r, the recovery rate be g, and the proportion of susceptible, infected and recovered state individuals in the population at moment t be S, I, R, respectively, then the kinetic evolution equation can be written as:

$$\frac{dS}{dt} = -rNSI \tag{1}$$

$$\frac{dI}{dt} = rNSI - gI \tag{2}$$

$$\frac{dR}{dt} = gI \tag{3}$$

Assuming that there is only one infectious seed at $t = 0$, the initial condition becomes $S(0) = 1 - 1/n \square = 1, I(0) = 1/N, R(0) = 0$, which can be obtained:

$$\frac{dS}{dR} = -R_0 S, R_0 = rN/g \tag{4}$$

Solving the differential equation yields the transcendental equation:

$$S(\infty) = S(0)e^{-R_0[R(\infty)-R(0)]} \tag{5}$$

This can be obtained by utilizing the asymptotic condition $I(\infty) = 0, S(\infty) = 1 - R(\infty)$ at $t \to \infty$ again:

$$R(\infty) = 1 - e^{-R_0 R(\infty)} \tag{6}$$

Since every infected person eventually recovers, $R(\infty)$ represents the scale of transmission of the infection.

Taking the derivative of $R(\infty)$ on both sides, there:

$$1 = R_0 e^{-R_0 R(\infty)} \tag{7}$$

Because of $e^{-R_0 R(\infty)} < 1$, the infection can only spread when $R_0 > 1$. Intuitively, when $R_0 < 1$, the rate at which susceptible people become infected rN is less than the rate at which infected people recover g, so that soon all infected people in the system recover, and because of the lack of infectious seeds, the process of spreading the contagion ends quickly. In general, infectious diseases cannot spread.

Prior vaccination is an important means of effectively preventing the spread of infectious diseases. In a fully connected network, a prior vaccination rate of x is equivalent to the initial condition of the transmission dynamics equation becoming $S(0) = (1-x)(1-1/N) \square = 1-x, I(0) = (1-x)/N, R(0) = 0$ and the final condition of transmission becoming $I(\infty) = 0, S(\infty) = (1-x)(1-R(\infty))$. Similarly, the final scale of transmission can be obtained:

$$R(\infty) = (1-x)(1-e^{-R_0 R(\infty)}) \tag{8}$$

obtained for the variant of Eq:

$$\frac{R(\infty)}{(1-e^{-R_0 R(\infty)})} = (1-x) \tag{9}$$

The derivative of $R(\infty)$ is obtained:

$$\frac{(1-e^{-R_0 R(\infty)}) + R(\infty) R_0 e^{-R_0 R(\infty)}}{(1-e^{-R_0 R(\infty)})^2} > 0 \tag{10}$$

And on the right-hand side $1-x$ is a decreasing function of x, so $R(\infty)$ is a decreasing function of x, i.e., it is intuitively clear that the scale of transmission of infectious diseases decreases as the inoculation rate increases. This being the case, we not only ask at what level of advance vaccination rate does the scale of transmission of infectious diseases fall to 0. It is easy to know that $R(\infty) = 0$ is a solution to the formula, and below we consider the asymptotic properties in the neighborhood of the solution. A Taylor expansion for each side of Eq. near solution 0 yields:

$$\grave{o} = (1-x)[1-(1-R_0\grave{o})] \tag{11}$$

The significant relationships between the parameters in the neighborhood of the 0 solution are thus obtained:

$$x_c = 1 - \frac{1}{R_0} \tag{12}$$

Since the scale of transmission is a decreasing function of the inoculation rate, we can conclude that when $x \geq x_c = 1 - \frac{1}{R_0}$, the scale of transmission is $R(\infty) = 0$, i.e., at this point in time, over-inoculation is occurring and the infectious disease is not able to spread. When $x \leq x_c = 1 - \frac{1}{R_0}$, the scale of transmission is $R(\infty) > 0$, i.e., the inoculation rate is insufficient at this point and the infectious disease has some ability to spread.

2.1.3. SIR modeling in complex networks. In SIR infectious disease transmission dynamics, the central role of the network is to characterize the role of infection by infected individuals with respect to susceptible individuals. We stipulate that if there exists: a connecting edge between i, j nodes, an

individual in state I infects an individual in state S at an infection rate τ when the individual represented by one node in i, j is in state I and the other node is in state S [23].

We present here the Gillespie algorithm stochastic simulation algorithm for the SIR model.

1) At moment $I=0$, randomly select I_0 nodes in the network at random to serve as transmission seeds for the infectious disease, marking their state as I : then randomly select Nr nodes as immunized individuals, marking their state as V : the remaining individuals' state is marked as S .

2) Calculate the state transition rate of individuals that can undergo state transitions during a single infection. For S -state individuals $p_i(t)=r\times C_i$, where C_i denotes the number of I -state individuals in i -node neighbors. For I -state individuals $p_i(t)=g$. Such that the total transition rate of the system is $\lambda(t)=\sum_i p_i(t)$. The system undergoes an asynchronous update at the moment of $t+\Delta t$, where Δt sampling is performed with an exponential distribution with parameter $1/\lambda(t)$.

3) Among the individuals that can undergo a state transition, randomly select an individual to undergo a state transition with a probability that is proportional to the state transition rate of that individual. Specifically, a uniformly distributed random number μ is generated between $[0,\lambda(t))$. Then the state transition rate of each state transitionable individual is accumulated starting from the state transitionable individual numbered 1 until the sum of the accumulated state transition rates exceeds μ for the first time, at which time the corresponding individual k is selected to undergo a state transition. If an individual was previously in state S , then now that it is infected, its own state changes to state I , its state transition rate changes to g , and the state transition rates of all S -state individuals in its neighborhood increase by r. If the individual was previously in state I , then now that the individual recovers from the dye, its own state changes to state R , its state transition rate becomes 0, and the state transition rates of all S state individuals in its surrounding neighborhood decrease by r.

4) The process of 2-3 is repeated until all I -state individuals in the system disappear, i.e., the system reaches equilibrium. The final scale of contagion of the infectious disease undergoes a phase change as r increases. For lower r, the final scale of transmission of the infectious disease grows very slowly because of local clustering. On all subsequent two-dimensional periodic lattice networks, we choose the phase transition threshold of $r=0.46$ as the infection rate of the infectious disease, which corresponds roughly to a final infection size of 0.893.

2.2. Reinforcement Learning Models and Processes

2.2.1. Basic concepts. Reinforcement learning corresponds to supervised learning and unsupervised learning, which is a kind of learning method of machine learning, but it does not rely on a large amount of labeled data, and does not need to give the “correct” policy as supervisory information, but in the interaction with the environment through the feedback of the environment to continuously adjust the policy to find the optimal decision. The general process of reinforcement learning is shown in Figure 1. Reinforcement learning mainly includes the following key factors:

1) Intelligent body (Agent). Intelligent body is the main part of the interaction with the environment, with the ability to learn and make decisions. It can sense the state of the environment (State, s), make decisions through the state to give feasible actions (Action, a), and then receive rewards from the environment (Reward, r), and constantly adjust the strategy of generating actions according to the rewards from the environment.

2) Environment. The environment is a general term for external things that interact with the intelligent body, which can change the current state it is in according to the action generated by the intelligent body's decision-making, and feed back the corresponding rewards obtained from the state transformation to the intelligent body.

3) State. The state is a description of the characteristics of the interaction environment of the intelligent body, and the values in the state space s can be discrete values or continuous values according to the specific task characteristics.

4) Action. Intelligent body strategic decision-making behavior is usually characterized by actions. For a specific task, the action space A can be discrete or continuous.
5) Strategy $\pi(a|s)$. The strategy is a function of the intelligent body's decision on the next action a based on the state s in which the environment is located. Specifically, it can be expressed as the probability distribution of actions computed by the intelligent body in a certain state of the environment.

$$\pi(a|s) \square p(a|s) \tag{13}$$

$$\sum_{a\in A}\pi(a|s)=1 \tag{14}$$

6) State transfer probability $p(s'|s,a)$. The state transfer probability can be described as the probability that the environment transitions from the original state s to a new state s' after the intelligence performs action a.
7) Immediate Reward $r(s,a,s')$. The immediate reward can be regarded as a scalar function that evaluates the change of the environment state. When the intelligent body executes action a based on the environment state s, the environment will give the intelligent body corresponding feedback based on the state change, and the feedback value is the instant reward, the value of which is generally related to the next state s' of the environment transition.

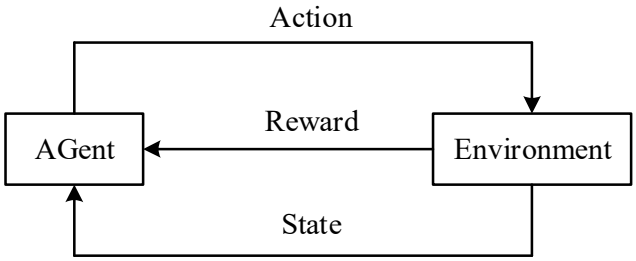


Fig. 1. General reinforcement learning process

2.2.2. Markov decision process. The multiple rounds of interaction between an intelligent body and its environment can be formalized as a Markov Decision Process (MDP). A Markov process consists of a set of sequences of random variables with Markovian properties: $s_0,s_1,...,s_t \in S$, and for the states in sequence S, the state s_{t+1} at the next moment is only related to the current state s_t, as shown in Eq.

$$p(s_{t+1}|s_t,...,s_0)=p(s_{t+1}|s_t) \tag{15}$$

where $p(s_{t+1}|s_t)$ is called the state transfer probability, $\sum_{s_{t+1}} p(s_{t+1}|s_t)=1$.

The Markov decision process introduces an additional variable to the Markov process: action a, i.e., the state s_{t+1} at the next moment in the sequence generated by the decision process is related not only to the current state s_t but also to the action a taken in the current state, as shown in Eq:

$$p(s_{t+1}|s_t,a_t,...,s_0,a_0)=p(s_{t+1}|s_t,a_t) \tag{16}$$

where $p(s_{t+1}|s_t,a_t)$ is the state transfer probability, and the MDP abstract conceptual model is shown in Figure 2.

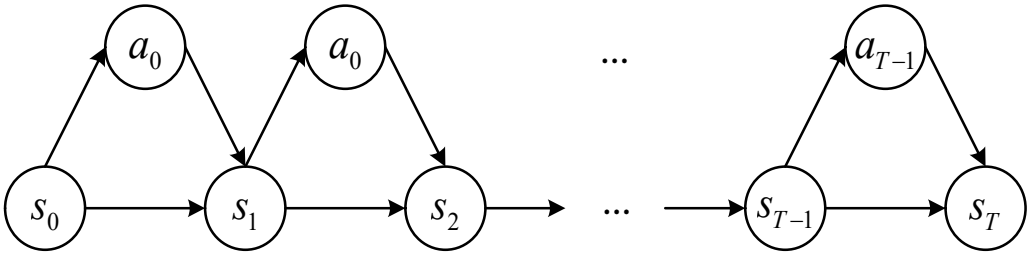


Fig. 2. Markov decision process

From the above theory, if strategy $\pi(a|s)$ is given, the probability of an intelligent body interacting with the environment to produce trajectory $\tau = s_0, a_0, s_1, r_1, a_1, ..., s_{T-1}, r_{T-1}, a_{T-1}, s_T, r_T$ is shown in Eq:

$$p(\tau) = p(s_0, a_0, s_1, a_1, ..., s_t, a_t) = p(s_0)\Pi_{t=0}^{T}\pi(a_t | s_t)p(s_{t+1} | s_t, a_t) \tag{17}$$

2.2.3. Objective function. The goal of reinforcement learning is to maximize the expected return, taking a trajectory τ generated by an intelligent body interacting with its environment as an example, given policy $\pi(a|s)$, the total return of the trajectory is shown in Eq:

$$G(\tau) = \sum_{t=0}^{T-1} r_{t+1} = \sum_{t=0}^{T-1} r(s_t, a_t, s_{t+1}) \tag{18}$$

Generally reinforcement learning tasks are turn-based tasks, i.e., there is a termination state in the environment, and when the interaction between the intelligent body and the environment reaches the termination state, the round of this interaction ends. If there is no termination state set in the environment, i.e., $T = \infty$, then it is a persistent task for persistent tasks, the total return increases over time and may be infinite, in which case reinforcement learning cannot make good decisions and optimization. To solve this problem, the focus on long-term returns can be effectively reduced by introducing a discount factor γ. Eq. represents the discounted return of trajectory τ after adding the discount factor:

$$G(\tau) = \sum_{t=0}^{T-1} \gamma^t r_{t+1} \tag{19}$$

where the discount rate $\gamma \in [0,1]$, the influence of short-term returns on the total returns of the trajectories obtained by Eq. becomes more prominent when the discount rate tends to 0, while the importance of long-term returns on the total returns increases significantly when the discount rate tends to 1.

Since the state and action spaces are not purely one-to-one relationships, the trajectories captured in each experiment are randomly generated, which results in trajectories with varying total returns. For reinforcement learning, the goal is to learn a policy $\pi(a|s)$ such that the average return obtained from all interaction trajectories is maximized under this policy decision, i.e., maximizing the expected return as mentioned above, with the objective function shown in Eq.

$$\mathrm{J}\ (\theta) = E_{\tau \sim P_\theta(\tau)}[G(\tau)] = E_{\tau \sim P_\theta(\tau)}\left[\sum_{t=0}^{T-1} \gamma^t r_{t+1}\right] \tag{20}$$

where θ is an argument to the strategy function.

2.2.4. Value functions. In order to effectively evaluate the expected benefit of strategy π, reinforcement learning introduces two value functions, the state value function and the state-action value function. The state-value function is used to evaluate the value of a state in the state space, which

can be obtained by decomposing the trajectory expectation lookback, and the specific splitting process and calculation results are shown in Eq:

$$\begin{aligned} & E_{\tau \sim p(\tau)}[G(\tau)] \\ & = E_{s \sim p(s_0)}\left[E_{\tau \sim p(\tau)}\left[\sum_{t=0}^{T-1}\gamma^t r_{t+1} \mid \tau_{s_0}=s_0\right]\right] \\ & = E_{s \sim P(s_0)}[V^\pi(s)]V^\pi(s) \\ & = E_{\tau \sim p(\tau)}\left[\sum_{t=0}^{T-1}\gamma^t r_{t+1} \mid \tau_{s_0}=s_0\right] \end{aligned} \tag{21}$$

where $V^\pi(s)$ is the state value function, which is used to calculate the expected value of the total return obtained by executing the strategy function π starting from state s, and τ_{s_0} denotes the initial state of trajectory τ. If $\tau_0:T$ is used to denote trajectory $s_0,a_0,s_1,...,s_T$ and $\tau_1:T$ is used to denote trajectory $s_1,a_1,...,s_T$, there is $\tau_0:T=s_0,a_0,\tau_1:T$. According to the Markov property, $V^\pi(s)$ can be further decomposed as shown in Eq:

$$\begin{aligned} V^\pi(s) &= E_{\tau_0:T \sim p(\tau)}[r_1+\gamma\sum_{t=1}^{T-1}\gamma^{t-1}r_{t+1} \mid \tau_{s_0}=s_0] \\ &= E_{a\sim\pi(a|s)}E_{s'\sim p(s'|s,a)}E_{\tau_1:T\sim p(\tau)}[r(s,a,s')+\gamma\sum_{t=1}^{T-1}\gamma'^{-1}r_{t+1} \mid \tau_{s_1}=s'] \\ &= E_{a\sim\pi(a|s)}E_{s'\sim p(s'|s,a)}[r(s,a,s')+\gamma E_{\tau_1:p(\tau)}[\sum_{t=1}^{T-1}\gamma^{t-1}r_{t+1} \mid \tau_{s_1}=s']] \\ &= E_{a\sim\pi(a|s)}E_{s'\sim p(s'|s,a)}[r(s,a,s')+\gamma V^\pi(s')] \end{aligned} \tag{22}$$

When $\pi(a \mid s), p(s' \mid s,a)$ and $r(s,a,s')$ are given, $V^\pi(s)$ can be computed iteratively and $V^\pi(s)$ will reach convergence after a certain number of iterations due to the presence of γ. The state-action value function or Q function is a measure of the feasibility of taking an action in a particular environmental state, which helps the model to choose the most appropriate behavior according to the situation it is in so as to maximize the expected gain.

$$Q^\pi(s,a)=E_{s'\sim p(s'|s,a)}[r(s,a,s')+\gamma V^\pi(s')] \tag{23}$$

$$V^\pi(s)=E_{a\sim\pi(a|s)}[Q^\pi(s,a)] \tag{24}$$

$$Q^\pi(s,a)=E_{s'\sim p(s'|s,a)}[r(s,a,s')+\gamma E_{a'\sim\pi(a'|s')}[Q^\pi(s',a')] \tag{25}$$

The value function can be regarded as a method of evaluating strategy π, which is the main basis for decision optimization. If in state s, there exists a^* and there is $Q^\pi(s,a^*)>V^\pi(s)$, which means that the return obtained by executing a^* in state s is higher than the return corresponding to the current strategy $\pi(a \mid s)$, then the parameter can be adjusted to make $p(a^* \mid s)$ increase in order to achieve the purpose of optimizing the strategy.

2.3. Disease prevention mechanisms based on capital investment

To clearly represent the network structure used in this section, the social network is denoted by $G=(V,E)$, where $V=\{1,...,N\}$ denotes the set of individuals and $E=\{e_{ij} \mid i,j\in V\}$ denotes the set of social connections between individuals. $e_{ij}=1$ if i and j are socially connected, and $e_{ij}=0$ otherwise.In addition, the first-order neighbors of individual i are denoted as $N_i=\{j \mid e_{ij}=1, j\in V\}$. To simulate the interactions between clusters, each individual i treats his first-order neighbors $\{j \mid j\in N_i\}$ as a locally well-mixed group G_i, where the individual interacts to the same extent as the other individuals in the group. Moreover, each individual i can obtain information related to disease

transmission and vaccination coverage from his second-order neighbor group G_i^2, as well as global information related to disease transmission and vaccination coverage from the public health bureau, where the second-order neighbor group is composed of the neighbors of the individual's neighbors, i.e., $G_i^2 = \{G_j^1 \mid j \in N_i\}$ [24]. Therefore, the model in this chapter focuses on the question of how individuals' risk perception affects their vaccination behavior, where individuals i can perceive the severity of a disease from local information based on (i) G_i^1 (ii) localized information in G_i^2 to perceive the severity of the disease. (iii) global information in the entire population to perceive disease severity.

2.3.1. Outbreak-awareness models based on local or global environments. The basic regeneration number R_0 represents the average number of infections caused by a pathogen during its period of infection in a fully susceptible and well-mixed population. According to the SIR model, $R_0 = \beta N/\gamma$, β denotes the rate of disease transmission, and γ denotes the rate of disease recovery. R_0 is used to denote disease severity, and although β and γ are known prior to modeling disease transmission with inoculation, obtaining precise values for disease severity during a disease epidemic is not possible. In general, individuals judge the severity of a disease based on local or global prevalence. Assuming that the proportion of individuals vaccinated in a mixed population is f_v, existing studies have shown that the risk of infection r satisfies the following relationship with the basic regeneration number R_0 and the immunization coverage f_v:

$$r = 1 - \frac{1}{R_0(1-f_v)} \tag{26}$$

When $f_v \geq 1 - 1/R_0$, the risk of infection is O. Thus, for seasonal epidemics, individuals can perceive the severity of the disease in season t based on the proportion of infections (i.e., disease prevalence $f_{I,t-1}^i$) and the size of infections $f_{V,t-1}^i$ in the previous season:

$$\tilde{R}_0^i(t) = \begin{cases} \dfrac{1}{(1-f_{I,t-1}^i)(1-f_{V,t-1}^i)} & if\ 0 \leq f_{I,t-1}^i < 1 \\ +\infty & if\ f_{I,t-1}^i = 1 \end{cases} \tag{27}$$

where the values of $f_{I,t-1}^i$ and $f_{V,t-1}^i$ depend on the source from which individual i obtains information related to disease prevalence and vaccine coverage.

Based on the different sources from which individuals obtain information related to disease vaccination and infection, this chapter proposes three types of static decision-making mechanisms, assuming that all individuals perceive disease severity based on the same source of information. They are described in detail as follows:

1) Local- G^1 mechanism: individual i can only obtain the prevalence and vaccination coverage of the disease from his first-order neighbors G_i^1. Thus, $f_{I,t-1}^i = f_I(t-1; G_i^1), f_{V,t-1}^i = f_V(t-1; G_i^1), f_I(t-1; G_i^1)$ and $f_V(t-1; G_i^1)$ denote the infection size and vaccination coverage of cluster G_i^1, and the perceived disease severity under the Local- G^1 mechanism can be calculated by bringing $f_{I,t-1}^i$ and $f_{V,t-1}^i$ into Eq.

2) Local- G^2 mechanism: individual i can interact with his social neighbors and can obtain disease-related information from his second-order neighbor cluster G_i^2. Based on this, in season t, the perceived disease severity of individual i is considered to be the maximum value in all his second-order neighbor clusters and in his first-order neighbor cluster, i.e., $\tilde{R}_0^I(t; G_i^2) = max_k \{\tilde{R}_0(t; G_i^1), \tilde{R}_0(t; G_k^1)\} \mid k \in N_i\}$.

3) Global-G mechanism: Individual i obtains the global disease prevalence $f_I(t-1; G)$ and vaccine

coverage $f_V(t-1;G)$ from the Department of Public Health. Therefore, $f_I, t-1 = f_I(t-1;G), f^i_{V,t-1} = f_V(t-1;G)$, the perceived disease severity under the Global-G mechanism can be calculated by bringing $f^i_{I,t-1}$ and f^i_{Vt-1} into Eq.

Based on the calculated perceived disease severity $\tilde{R}^l_0(t)$, an individual can estimate the risk of being infected in his first-order neighborhood as:

$$\tilde{r}_i(t) = 1 - \frac{1}{\bar{R}^i_0(t)(1 - f^{G^1_i}_{V,t})} \tag{28}$$

Before the start of each season, individuals decide whether to vaccinate by weighing the cost of vaccination c_V against the cost of contracting the disease. $p_i(t)$ denotes the probability of vaccination in season t, then the cost of vaccination is denoted as $p_i(t)c_V$ and the cost of contracting the disease is denoted as $(1-p_i(t))\tilde{r}_i(t)c_i$. A rational individual should be vaccinated with probability $p_i(t)$ because of $p_i(t)c_V = (1-p_i(t))\tilde{r}_i(t)c_I$. Therefore the individual's probability of being vaccinated in season t should be:

$$p_i(t) = \frac{\tilde{r}_i(t)}{\tilde{r}_i(t) + c} \tag{29}$$

$c = c_V / c_I$ is the relative cost of vaccination, and $\tilde{r}_t(t)$ is the risk of infection, which is calculated by the formula.

2.3.2. Reinforcement Learning Based Adaptive Decision Modeling. In fact different individuals rely on different sources of information to make inoculation decisions, some are more inclined to make inoculation decisions based on local information from their social neighbors, while others are more inclined to make inoculation decisions based on global information. To explore human self-organizing behavior, this section introduces a reinforcement learning mechanism that allows individuals to adaptively choose whether to make vaccination decisions based on local or global information based on their vaccination experience. Similarly, this section treads with a ò greedy algorithm: an individual randomly chooses whether to base its decision on local or global information with probability ò. The individual chooses with probability $1-ò$ the source of information that yields a high average return [25]. This section focuses on how individuals adjust their strategies between the two mechanisms Local- G^2 and Global-G. According to the mechanism of reinforcement learning individual's perceived severity of the epidemic in season t is calculated as follows:

$$\bar{R}^i_0(t) = \begin{cases} R^i_0(t; G^e_l) \textit{ with probability } 1-ò \textit{ and } U_i(t-1; G^e_l) > U_i(t-1; G) \\ \bar{R}^i_0(t; G) \textit{ with probability } 1-ò \textit{ and } U_i(t-1; G^2_l) < U_i(t-1; G) \\ \bar{R}^i_0(t; G^2_i) \textit{ or } \tilde{R}^i_0(t; G) \textit{ with probability } ò \end{cases} \tag{30}$$

$U_i(t-1; G^2_i)$ and $U_i(t-1; G)$ represent the average returns from the first to the $n-1$th season based on the Local- G^2 and Global-G mechanisms, respectively. The average returns for season t are updated as follows:

$$U_i(t; G^2_i) = \begin{cases} \dfrac{U_i(t-1; G^2_i)(n_i(t; G^2_i) - 1) + u_i(t)}{n_i(t; G^2_i)} & \text{If the } Local - G^2 \text{ mechanism is used} \\ U_i(t-1; G^2_i) & \text{If the } Global - g \text{ mechanism is used} \end{cases} \tag{31}$$

$$U_i(t; \mathrm{G}) = \begin{cases} \dfrac{U_i(t-1; \mathrm{G})(n_i(t; \mathrm{G}) - 1) + u_i(t)}{n_i(t; \mathrm{G})} & \text{If the } Global - g \text{ mechanism is used} \\ U_i(t-1; \mathrm{G}) & \text{If the } Local - G^2 \text{ mechanism is used} \end{cases} \tag{32}$$

where $n_i(t;G_i^2)$ and $U_i(t-1;G)$ denote the number of times the Local-G^2 mechanism and Global-G mechanism were adopted by individual i from the first to the t season, respectively.

3. Results and discussion

3.1. Analysis of Disease Prevention Decisions under Different Update Rules

The experiments in this section experimentally validate the disease prevention reinforcement learning model proposed in this paper from the perspectives of loss-based updating rules and environment-based updating rules.

3.1.1. Loss-based update rule results. To demonstrate the effect of loss A (loss tolerance) on vaccination behavior, we show the joint effect of K (sensitivity) and c (relative vaccination cost) on the vaccination level and epidemic size as a function of relative vaccination cost under different sensitivities as shown in Fig. 3 (Figs. function, Figures a to d show the effect of vaccination level on relative vaccination cost, and Figures e to h show the effect of epidemic size on relative vaccination cost). The figure clearly illustrates the variation of vaccination level (top) and final disease size (bottom) with relative vaccination cost c. The figure also shows the variation of vaccination level (top) and final disease size (bottom) with relative vaccination cost c. Both transition thresholds for vaccination level and final disease size exist for each tolerance. Specifically, in Figures 3(a) and (e), higher K delays the point of phase transition in relative cost c. When the second transition occurs, even extensive vaccination coverage fails to curb disease transmission in the larger K and A scenarios. In addition, it was found that in all cases, higher sensitivities at higher relative costs also produced higher levels of inoculation. As can be seen in Figure 3(d), when K is large and A=0, the morphological changes in inoculation levels gradually converge to a U-shape as c increases.

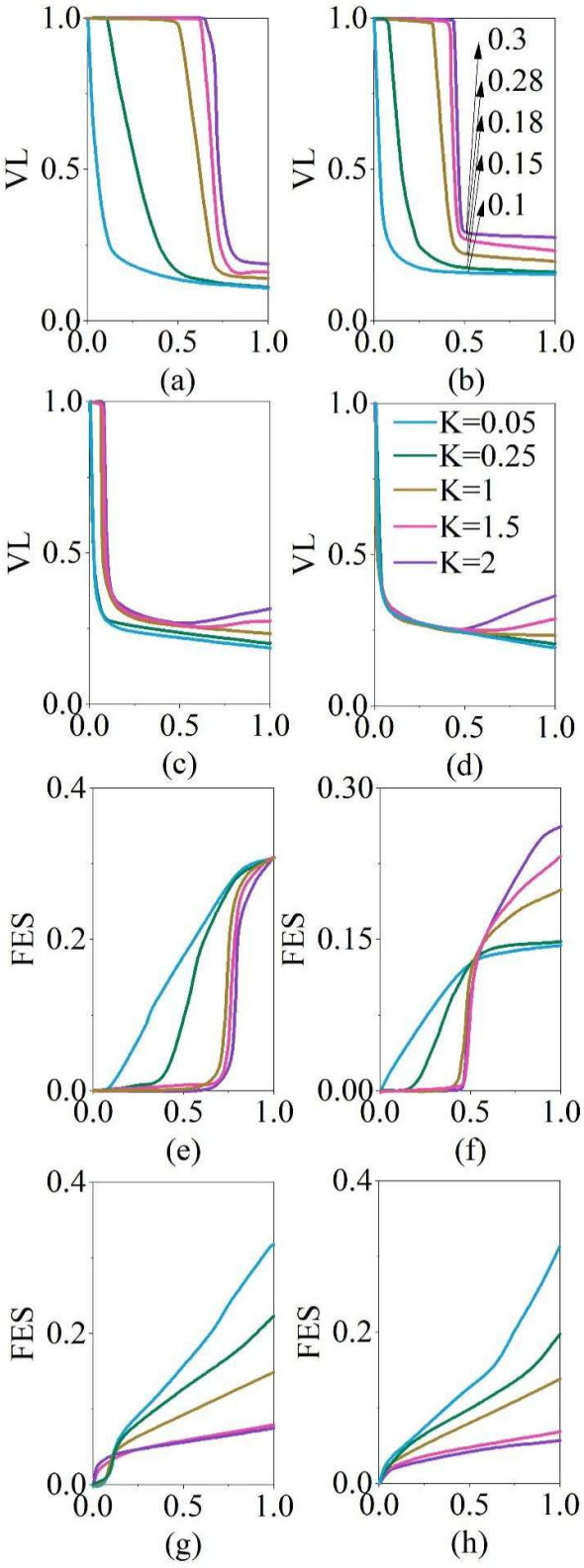


Fig. 3. Relative inoculation costs

To further explore the microstructure that emerges from reinforcement, the temporal evolution of the probability and mean probability of vaccination for the three types of vaccinated, infected, and susceptible individuals is shown in Fig. 4 (denoted by pv, pi, and ps, respectively), and the distribution of the expected probability of an individual to be vaccinated is shown in Fig. 5, and the distribution of

probabilities on the lattice in the steady state is shown in Fig. 6.

Considering that risk averse individuals are extremely averse to economic loss and dominate in the real world, it is crucial to explore the scenario when A = 0. Specifically, given the extreme value of K (0.01 of the approximate minimum), the probability of all individuals' propensity to vaccinate increases over time. In turn, the probability distribution of their propensity to vaccinate is right-skewed, with a median of 0.65 and a mean of approximately 0.7236. With a high propensity for group vaccination, the vaccination level and prevalence size were 0.75 and 0.001, respectively. The median and mean become 0.55 as the relative cost of vaccination increases c. The probability converges to 0.55 for all individuals. In Figure 4, we give the expected probability distribution for all individuals, which ranges from [0.43,0.57]. The figure illustrates the case where K and c both increase. An increase in these parameters shifts the distribution of inoculation propensity to the left. In the figure, we observe that this situation occurs due to the low inoculation probability propensity. Observation of the pattern of expected probability of vaccination for each type of individual shows that higher levels of vaccination occur when the expected probability of vaccination is greater than 0.5 dominates. Different distributions of vaccination propensity can be obtained by varying K and c: left (right) skewed and normal phases.

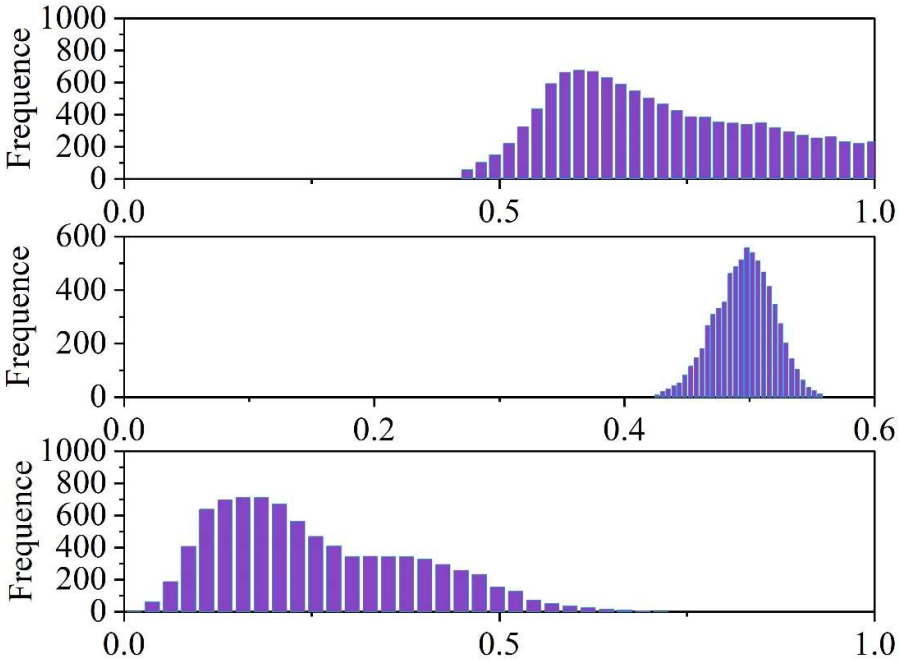


Fig. 4. The time of the probability of a single type is a time of probability

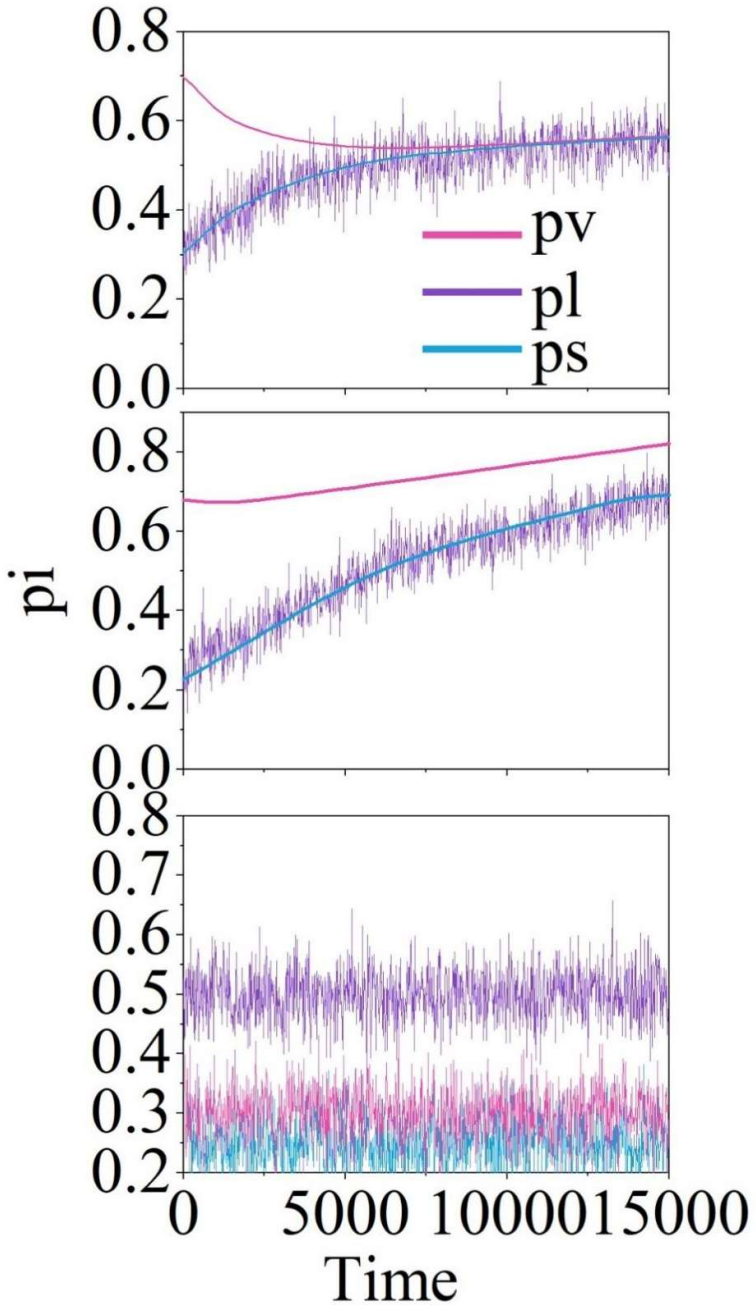
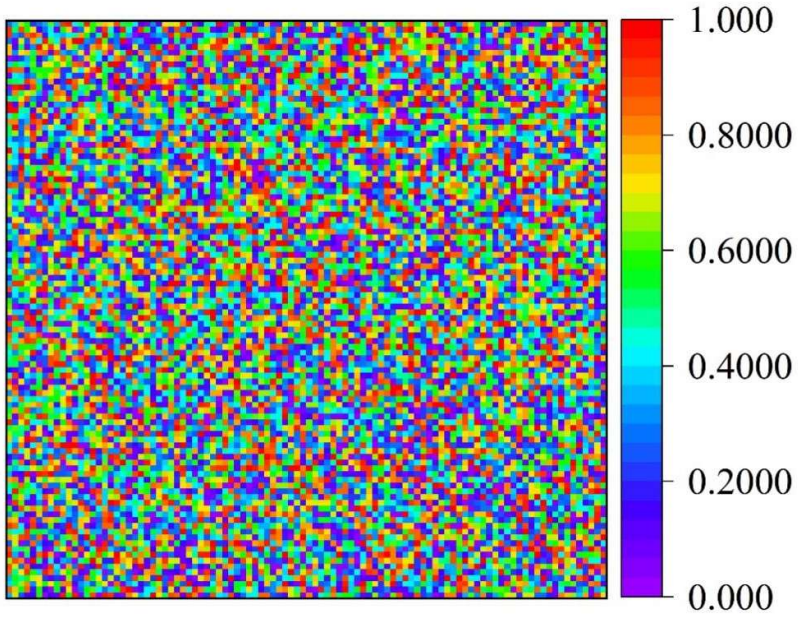
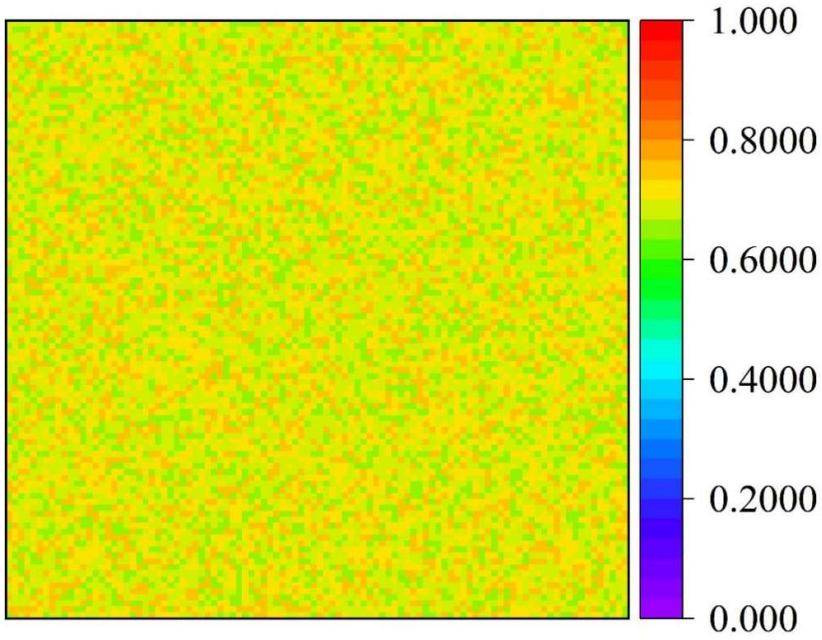


Fig. 5. The distribution of the probability of vaccination for individuals


1.000
0.8000
0.6000
0.4000
0.2000
0.000


1.000
0.8000
0.6000
0.4000
0.2000
0.000

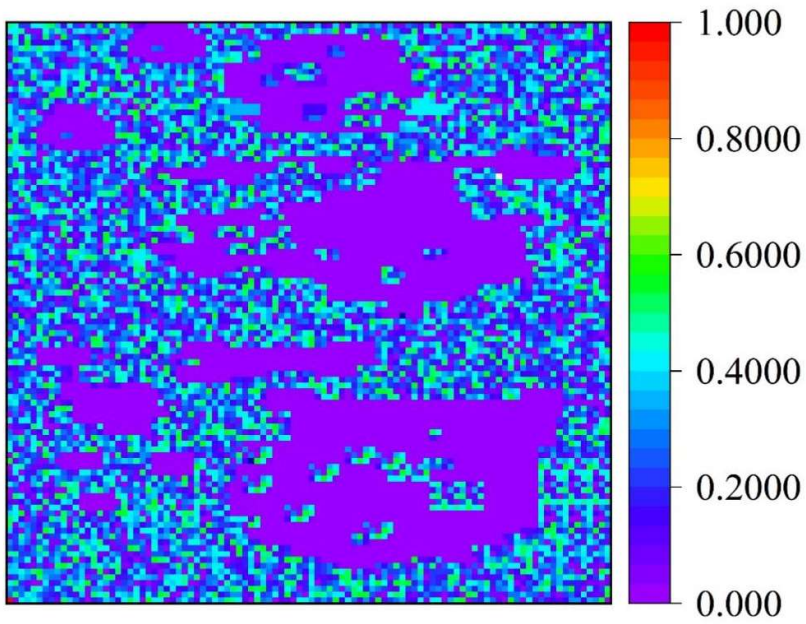


Fig. 6. The probability distribution on the grid in a stable state

The vaccination level and epidemic size as a function of the relative vaccination cost c for different sensitivities K are shown in Fig. 7 (Figs. a to d show the effect of vaccination level on the relative vaccination cost, and Figs. e to h show the effect of epidemic size on the relative vaccination cost). Similarly, we observe that the inoculation level (decreasing from 1 to 0) exhibits three distinct phase transition points as the cost c increases. The first critical point for cost occurs at 0.7 when A = -0.8 and K = 2, when the inoculation level begins to decline. After the second critical point, the inoculation level increases significantly. Finally, as the cost approaches 0.85, the inoculation level begins to decrease again. In the BA network, for the case of loss intolerance, the inoculation level at low K is higher than the inoculation level at high K. The inoculation level at high K is higher than the inoculation level at high K. In this case, a high level of vaccination ensures that the outbreak remains mild. Thus, lower K values in the lattice and BA networks help to alleviate the dilemma faced by vaccination.

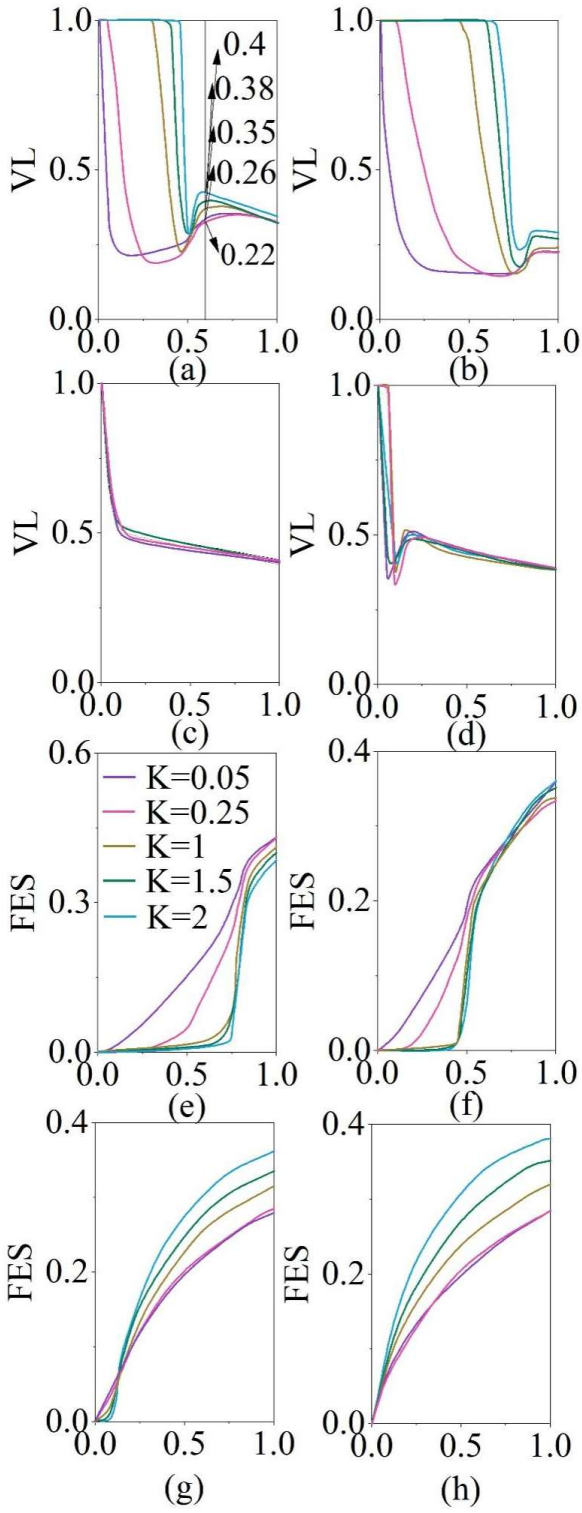


Fig. 7. The level of vaccination and the scale of the epidemic as a function

3.1.2. Results of environment-based updating rules. In this subsection, we show the results of the environment-based updating rule. The synergistic effect of K and c on vaccination levels and disease size is shown in Figure 8. In particular, when K = 0, individuals are not sensitive to differences in returns, and thus vaccination levels do not depend on the relative cost of vaccination c. For the system as a whole, a higher K promotes higher vaccination levels. That is, endowing individuals with higher

sensitivity promotes higher vaccination levels. When K > 0, the vaccination level is always negatively correlated with the relative cost c: as the cost of vaccination increases, the vaccination level decreases.

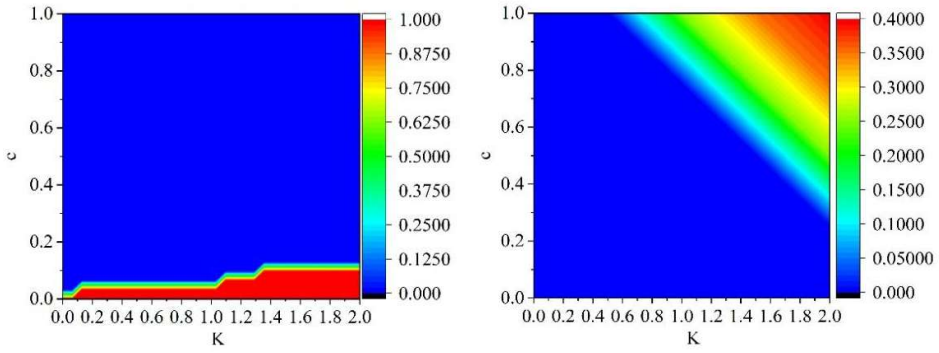


Fig. 8. Synergism

In addition, we investigated the relationship between the level of disease prevention and the size of the disease with c on the lattice network and the BA network under fixed K conditions as shown in Figure 9. As shown in the left panel, on the lattice network, the relatively expensive and with higher K condition, the system maintained a higher inoculation level than in the other conditions. Interestingly, it was found that even though inoculation levels were higher in such conditions, for K=2, disease was still prevalent. In addition, it was found that the effect of sensitivity K on inoculation levels and disease magnitude was not significant in the BA network.

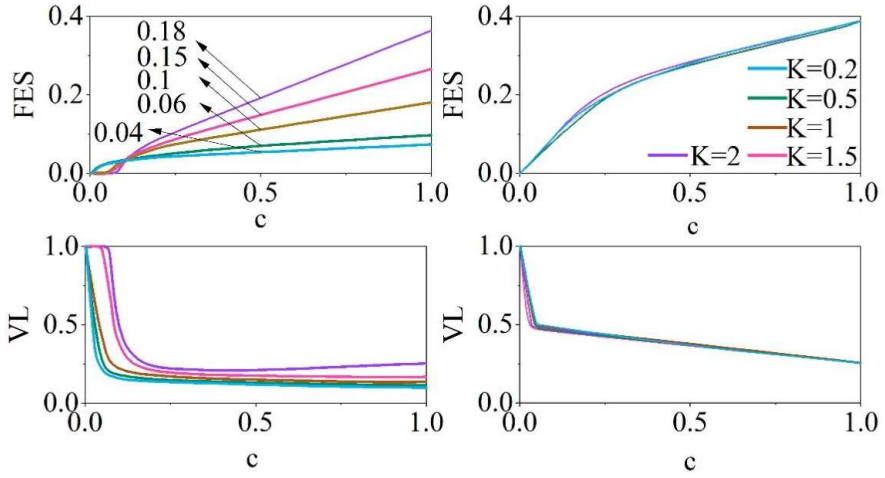


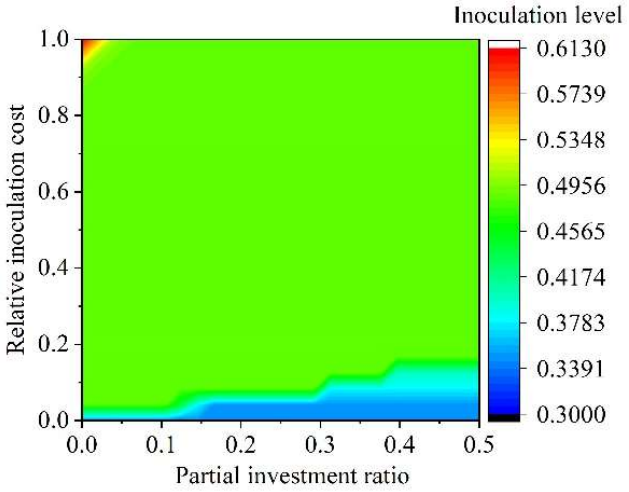
Fig. 9. The relationship between the level of vaccination and the scale of the disease

3.2. Simulation results and analysis

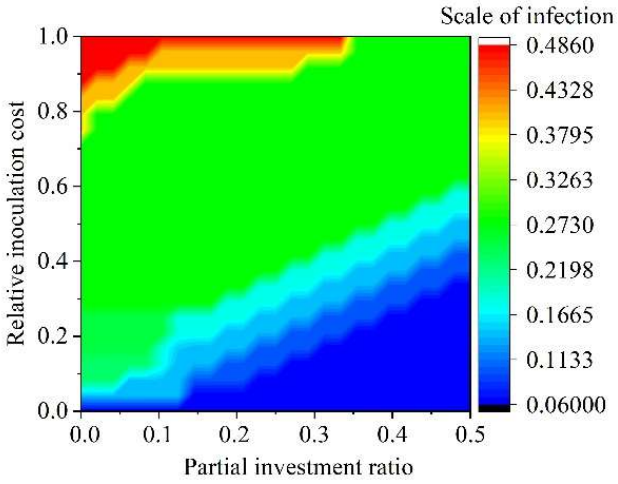
3.2.1. Impact of relative investment costs on results. Firstly, through the simulation results, the effect of relative investment cost c and partial investment ratio J on the results under the partial subsidy policy was explored, and the hotspot diagrams of relative investment cost and partial investment ratio about the vaccination level and the infection scale, respectively, are shown in Fig. 10 (Fig. a is the effect of relative investment cost and partial investment ratio on the vaccination level, and Fig. b is the effect of relative investment cost and partial investment ratio on the infection scale).

From Figure 10(a), it can be seen that as the relative investment cost increases, the vaccination level first decreases and then increases, this is because as the investment cost increases, the willingness of community residents to be vaccinated decreases, but when the cost of vaccination is very high, the investment makes the benefits of vaccinating the individual much smaller than the benefits of not

being vaccinated, and therefore the vaccination level increases. Partial investment ratios have a particular effect on vaccination levels. When the relative cost of vaccination is human or small, the level of vaccination is getting larger as the proportion of partial investment increases. Then, looking at Figure 10(b), it is found that the size of the infection is getting larger as the relative cost of inoculation increases. Looking horizontally, it is found that the size of the infection is becoming progressively smaller as the proportion of partial investment increases. Theoretically, the particular phenomenon of vaccination levels in the intermediate region is due to the fact that partial investment, while incentivizing more people to increase their willingness to be vaccinated, also leads to more people successfully piggybacking on it, so that the scale of epidemics in society as a whole decreases in the face of reduced or fluctuating levels of disease prevention.



(a) The impact of relative investment cost and partial investment ratio



(b) The impact of relative investment costs on the scale of infection

Fig. 10. The relative investment cost is related to the level of vaccination and the scale

3.2.2. Partial versus full investment policies. This section will further analyze the impact of different government subsidy policies on the results. Therefore, in this section, the partial government investment policy will be compared with the full investment policy, so as to analyze the effect of different investment policies on controlling the spread of epidemics in the whole social network and reducing the total social cost. The effects of the three investment policies on the scale of infection and total social cost when the relative cost of vaccination c= 0.2, 0.5, and 0.9 are shown in Figure 11. Among

them, Figure 11(a) depicts the variation of infection size and total social cost under the randomly selected full investment policy. Figure 11(b) depicts the change in infection size versus total social cost under the full investment policy of targeted selection. Figure 11(c) depicts the variation of infection size versus total social cost under partial investment policy. As can be seen in Figure 11, all three investment policies lead to a reduction in the scale of epidemic infections across the social network, as well as a reduction in total social costs. First, look at the three images in the left column. They depict the effects of a randomly selected full investment policy, a target-selected full investment policy, and a partial investment policy on the scale of epidemic infections. It can be seen that when the corresponding investment proportion increases, the infection size decreases subsequently, with the curve in Figure 11(b1) showing a more significant decreasing trend, faster and larger in magnitude. The curve in Fig. 11(c1) shows the second largest change, and the curve in Fig. 11(a1) shows a slower and smaller change compared to the previous two. That is, when the subsidy effect is maximized, a target-selected full investment policy is likely to be the most effective in controlling the spread of epidemics in the entire social network, with partial investment the next most effective, and a randomly-selected full investment policy bringing about the least effect. However, if one wants the partial investment policy to achieve the same effect as the targeting-selected full investment policy, one simply needs to increase the proportion of partial investments appropriately while keeping the proportion of targeting-selected investments constant. Then, look at the three images in the right-hand column. They depict the effects of a randomly selected full investment policy, a target-selected full investment policy, and a partial investment policy on total social costs. The decreasing trend of the curve in Figure 11(b2) is more pronounced, faster, and larger. The change in the curve in Figure 11(c2) is the next most significant, and the change in the curve in Figure 11(a2) is slower and smaller in magnitude compared to the previous two. Therefore, when set in accordance with the range of the two investment ratios in this paper, when the investment situation is maximized, the full investment policy of targeting choice is likely to be the best in terms of reduction of total social costs, partial investment is the second best, and the investment subsidy policy of random choice brings the smallest effect.

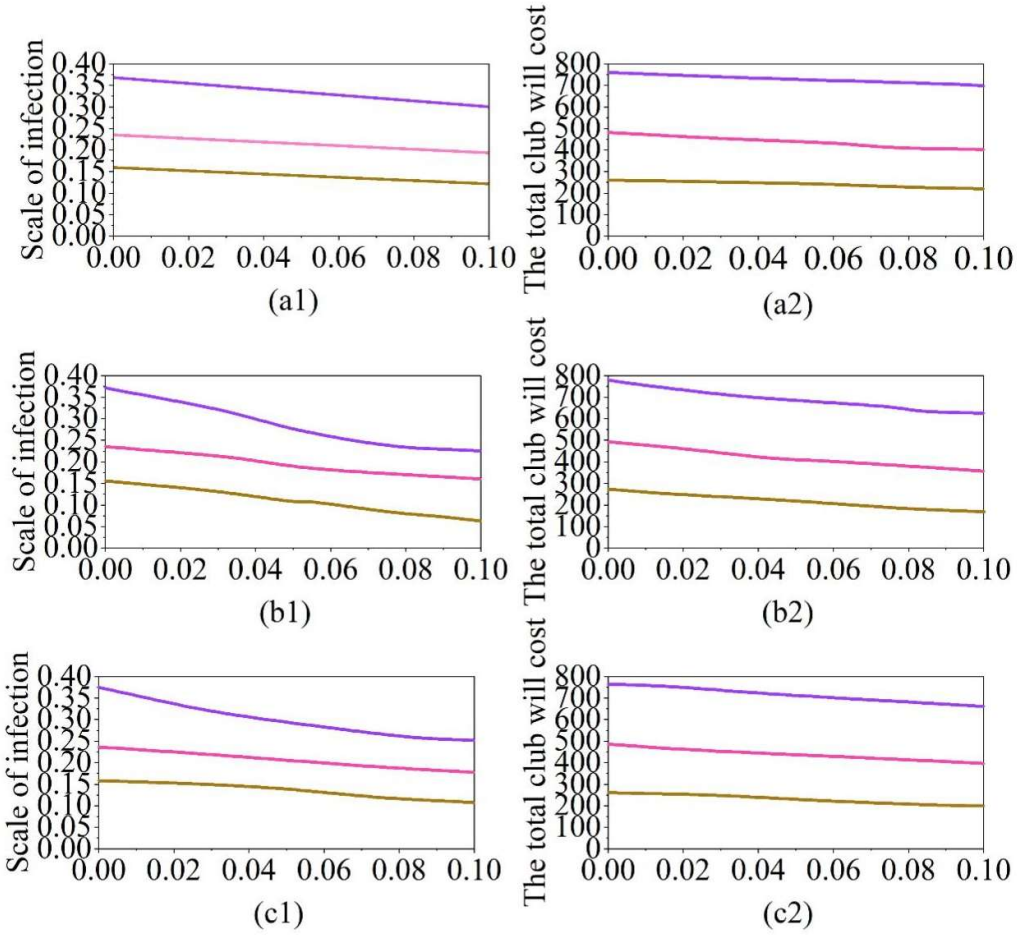


Fig. 11. The impact of three investments on scale

4. Conclusion

The article investigates the interrelationship between social capital investment and disease prevention using theories related to communication dynamics and reinforcement learning.

Both loss-based and environment-based updating rules are beneficial in stopping the spread of, diseases in community health. When considering the loss-based renewal rule, sensitivity K plays a positive role in stopping the spread of disease in most cases. When the environment-based renewal rule is considered, sensitivity plays a crucial role in ending epidemic diseases.

After analyzing the changes in the size of infection and total social cost for three different investment policies for three scenarios, c=0.2, c=0.5, and c=0.9, it is found that the policy of partial social investment is effective in controlling the spread of epidemics and reducing the total cost. In terms of reducing the infection scale as well as reducing the total cost, the full investment of the targeting choice is optimal.

In summary, the reinforcement learning model of social capital investment for disease prevention proposed in this paper has good operation.

References

[1] Shiell, A., Hawe, P., & Kavanagh, S. (2020). Evidence suggests a need to rethink social capital and social capital interventions. Social science & medicine, 257, 111930.

[2] Zhang, X., Liu, S., Chen, X., & Gong, Y. (2017). Social capital, motivations, and knowledge sharing intention in health Q&A communities. Management Decision, 55(7), 1536-1557.

[3] Flores, E. C., Fuhr, D. C., Bayer, A. M., Lescano, A. G., Thorogood, N., & Simms, V. (2018). Mental health impact of social capital interventions: a systematic review. Social psychiatry and psychiatric epidemiology, 53, 107-119.

[4] Gupta, R., Kaur, M., Islam, S., Mohan, V., Mony, P., Kumar, R., ... & Yusuf, S. (2017). Association of household wealth index, educational status, and social capital with hypertension awareness, treatment, and control in South Asia. American journal of hypertension, 30(4), 373-381.

[5] Zoorob, M. J., & Salemi, J. L. (2017). Bowling alone, dying together: The role of social capital in mitigating the drug overdose epidemic in the United States. Drug and alcohol dependence, 173, 1-9.

[6] Cockerham, W. C., Hamby, B. W., & Oates, G. R. (2017). The social determinants of chronic disease. American journal of preventive medicine, 52(1), S5-S12.

[7] Raymond-Flesch, M., Auerswald, C., McGlone, L., Comfort, M., & Minnis, A. (2017). Building social capital to promote adolescent wellbeing: a qualitative study with teens in a Latino agricultural community. BMC public health, 17, 1-9.

[8] Ford, J. A., Sacra, S. A., & Yohros, A. (2017). Neighborhood characteristics and prescription drug misuse among adolescents: The importance of social disorganization and social capital. International Journal of Drug Policy, 46, 47-53.

[9] Muntaner, C., Lynch, J., & Smith, G. D. (2020). Social capital, disorganized communities, and the third way: understanding the retreat from structural inequalities in epidemiology and public health. Political And Economic Determinants of Population Health and Well-Being:, 427-450.

[10] Buttice, V., Colombo, M. G., & Wright, M. (2017). Serial crowdfunding, social capital, and project success. Entrepreneurship theory and practice, 41(2), 183-207.

[11] Török, E., Clark, A. J., Jensen, J. H., Lange, T., Bonde, J. P., Bjorner, J. B., ... & Rod, N. H. (2018). Work-unit social capital and long-term sickness absence: a prospective cohort study of 32 053 hospital employees. Occupational and Environmental Medicine, 75(9), 623-629.

[12] Buehler, J. W., Castro, J. C., Cohen, S., Zhao, Y., Melly, S., & Moore, K. (2019). Personal and neighborhood attributes associated with cervical and colorectal cancer screening in an urban African American population. Preventing chronic disease, 16, E118.

[13] Holt-Lunstad, J. (2022). Social connection as a public health issue: the evidence and a systemic framework for prioritizing the "social" in social determinants of health. Annual Review of Public Health, 43(1), 193-213.

[14] Gómez, C. A., Kleinman, D. V., Pronk, N., Gordon, G. L. W., Ochiai, E., Blakey, C., ... & Brewer, K. H. (2021). Addressing health equity and social determinants of health through healthy people 2030. Journal of public health management and practice, 27(Supplement 6), S249-S257.

[15] Amoah, P. A. (2018). Social participation, health literacy, and health and well-being: A cross-sectional study in Ghana. SSM-population Health, 4, 263-270.

[16] Villalonga-Olives, E., Wind, T. R., & Kawachi, I. (2018). Social capital interventions in public health: A systematic review. Social Science & Medicine, 212, 203-218.

[17] Xue, X., Reed, W. R., & Menclova, A. (2020). Social capital and health: a meta-analysis. Journal of Health Economics, 72, 102317.

[18] Coll-Planas, L., Nyqvist, F., Puig, T., Urrútia, G., Solà, I., & Monteserín, R. (2017). Social capital interventions targeting older people and their impact on health: a systematic review. J Epidemiol Community Health, 71(7), 663-672.

[19] Bian, Y., Miao, X., Lu, X., Ma, X., & Guo, X. (2020). The emergence of a COVID-19 related social capital: the

case of China. International Journal of Sociology, 50(5), 419-433.

[20] Coll‐Planas, L., del Valle Gomez, G., Bonilla, P., Masat, T., Puig, T., & Monteserin, R. (2017). Promoting social capital to alleviate loneliness and improve health among older people in S pain. Health & social care in the community, 25(1), 145-157.

[21] Kim, E. S., & Kawachi, I. (2017). Perceived neighborhood social cohesion and preventive healthcare use. American journal of preventive medicine, 53(2), e35-e40.

[22] Javier López de la Cruz & Alexandre N. Oliveira Sousa. (2025). SIR models with vital dynamics, reinfection, and randomness to investigate the spread of infectious diseases. Communications in Nonlinear Science and Numerical Simulation(P1),108359-108359.

[23] Li Jia,Hongwu Tan & Hui Cao. (2024). Analysis of Dynamics of a Recurrent Infectious Disease SIRS Model with Age Structure and Two Delays. International Journal of Bifurcation and Chaos(10).

[24] Yang Wu,Haixiang Guo,Yong Shi,Wenkai Zhang & Lei Wang. (2024). Vaccination subsidy allocation under budget constraints considering the human interpersonal contact pattern and vaccine protection effect in epidemics. Computers & Industrial Engineering110679-110679.

[25] Shio GaiQuek,GaneshsreeSelvachandran,VimalaJayakumar,PhetDuong & Le HoangSon. (2024). A New Decision Making Model Based on Complex Intuitionistic Fuzzy Soft Lattice for Traffic Monitoring in the Pandemic Scenarios. Advanced Intelligent Systems(11),2400145-2400145.