

Submitted to *University of Amsterdam*

# VACCINATION STRATEGIES IN EPIDEMICS WITH MUTATING INFECTIVITY ON SMALL-WORLD NETWORKS

EMIL JAVŮREK

13331124, University of Amsterdam

Simulations of epidemic spread are an active area of research with the practical aim of advancing our understanding of real world epidemic spread and ability to mitigate it. Here, another step is taken towards improving our understanding of the performance of various vaccination strategies. Specifically, the focus is on investigating vaccination strategy performance in a basic SIR model and its extension with mutating infectivity of the pathogen. The simulation is constrained to discrete time-step SIR model of epidemic spread on small-world networks. Beyond confirming the findings of past research, no difference between baseline and extended model is found, suggesting a robustness of the relative vaccination strategy performance with regards to the model extension.

#### 1. STATEMENT OF ORIGINALITY

This document is written by Emil Javůrek who declares to take full responsibility for the contents of this document.

I declare that the text and the work presented in this document are original and that no sources other than those mentioned in the text and its references have been used in creating it.

The Faculty of Economics and Business is responsible solely for the supervision of completion of the work, not for the contents.

## 2. INTRODUCTION – RESEARCH QUESTION

With the recent emergence of the global COVID-19 pandemic, the importance of understanding the spread of infectious diseases has been reinstated within the scientific community and beyond. A part of this sought after knowledge is the analysis of performance of various vaccination strategies. Strategies can do better or worse than random vaccinations and in the context of real world disease spread these differences can be measured in human lives. With other avenues of research of epidemic spread being difficult or outright unethical, simulations are a relatively simple way for researchers to explore this important topic. The difficulty of simulation lies in the need to simplify in a way that still allows for accurate conclusions about real disease spread. The focus of this paper will thus be to improve this understanding using one method, namely the analysis of simulating discrete time SIR model epidemics on small-world networks. Specifically, the performance of various vaccination strategies will be studied under the extension of the simulation model with a mutating infectivity of the pathogen.

## 3. THEORETICAL FRAMEWORK

The study of networks has its theoretical foundation laid out by the mathematical field of graph theory. On its own, network science is a fairly new research field, and its beginnings are attributed to Erdős et al. (1960) and Granovetter (1983). An extensive coverage of the science can be found in Barabási (2013). While the mathematical objects "network" and "graph" are the same, the focus of network science is on a certain process occurring on a network. The understanding of the network is then used to infer characteristics of the studied process.

For the purposes of epidemic simulation, the underlying network must accurately represent human interactions that lead to disease spread. Rather than mapping the connections in any real human society, a general class of networks that possesses characteristics typical of human networks is used. Apart from avoiding the near-impossible task of mapping all connections in a society, a general set of networks has the advantage of allowing the researcher to easily sample a large collection of such networks and generalize found properties to the whole set.

For these reasons, simulation of epidemics is conducted on (but not limited to) small-world networks (Zanette and Kuperman (2002), Hartvigsen et al. (2007), Xu and Sui (2009), Rüdiger et al. (2020), Azizi et al. (2020), Kiss et al. (2017)). The family of small-world networks has first been described by Watts and Strogatz (1998). Defined with a parameter of randomness, these networks lie somewhere between regular lattices and completely random graphs. While retaining a high clustering coefficient (CC) characteristic of the regular lattices, they also share

the random graphs property of low average path length (APL; see [Figure 1](#)). Their efficient generation, necessary for large-scale simulations such as those of epidemics, has been developed by [Batagelj and Brandes \(2005\)](#).

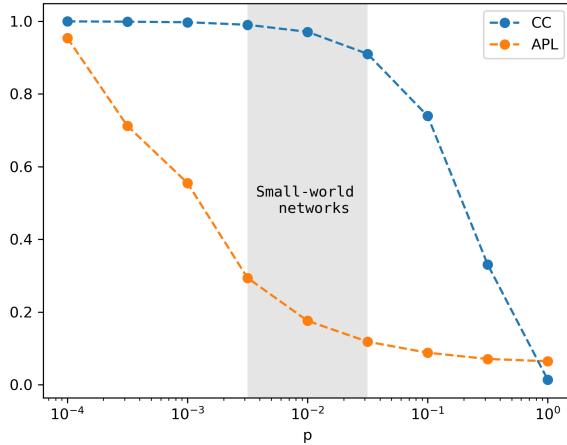


FIGURE 1.—Depiction of the properties defining the small-world networks. The figure shows the averages of APL and CC values over 20 Watts-Strogatz ( $n = 1000, k = 10$ ) networks rewired with probability  $p$  ranging from 0.0001 to 1. Both APL and CC averages are normalized against the APL and CC values of a regular lattice ( $p = 0$ ). We can see that for certain values of  $p$ , the networks retain high CC but have low APL, characteristics defining the small-world networks, as first described by [Watts and Strogatz \(1998\)](#).

The spread of infectious diseases is modeled with standard model structures, most notably using the SIR model ([Zanette and Kuperman \(2002\)](#), [Hartvigsen et al. \(2007\)](#), [Xu and Sui \(2009\)](#), [Rüdiger et al. \(2020\)](#), [Azizi et al. \(2020\)](#)). The name of this model is an acronym for the different stages of disease – "S" for susceptible, "I" for infectious, and "R" for recovered – that an individual can be found in. The model allows transitions between the states according to the graphic:

$$S \xrightarrow{\beta} I \xrightarrow{\gamma} R \quad (1)$$

where  $\beta$  is the infection probability/rate and  $\gamma$  is the recovery probability/rate (for discrete/continuous time). On this foundation, more complex models can be built that include more stages and/or more dynamics<sup>1</sup>.

<sup>1</sup>for an overview, see [Kiss et al. \(2017\)](#)

Within the general framework of epidemic simulation described above, the impact of various vaccination strategies is investigated. [Zanette and Kuperman \(2002\)](#) have shown a negative relationship between the prior immunization of a fraction  $\rho$  of the population and  $r$ , the final fraction of the susceptible population that was infected, and a positive relationship between small-world network parameter  $p$  and  $r$ . Furthermore, comparing random and targeted immunization (targeted prioritizing the most connected nodes), the targeted strategy is found to outperform the random one. However, it is recognized that the targeted strategy requires full knowledge of the network and thus its perfect implementation is not realistic.

Extending this, [Hartvigsen et al. \(2007\)](#) have also explored several other strategies: HCC, LCC (prioritizing immunization of nodes with high / low CC) and CCE (prioritizing nodes with maximum distance cross-cut edges). Beyond confirming the relationships between  $\rho$ ,  $p$  and  $r$  and the comparison of the targeted and random strategies, they concluded that HCC performs worse than random, and LCC and CCE comparable to random. [Xu and Sui \(2009\)](#) have also investigated the strategies of tracing (vaccination of susceptible nodes neighboring an infected one) and acquaintance immunization (vaccination of random nodes and their immediate neighbours until fraction  $\rho$  is reached) and have found the latter to have similar effects to the targeted strategy. The advantage of the acquaintance strategy is that it does not require any knowledge of the network for implementation, and thus represents a more realistic substitute for the targeted strategy.

The vaccination strategies mentioned above have been explored in the basic SIR model. Meanwhile, however, further research simulated more realistic extensions of the model. [Azizi et al. \(2020\)](#) have introduced peer communication into the model and compared various network structures, while [Rüdiger et al. \(2020\)](#) considered the mutating infectivity of the pathogen.

Although the effect of vaccination strategies in simple epidemic models is understood and so are the dynamics of the extensions of these models, the effects of vaccine strategies on the extended models are not. The aim of this paper is to narrow this gap in research by investigating the performance of vaccination strategies in a model with mutating infectivity of the pathogen.

#### 4. METHODOLOGY

In this section, the structure of the simulation will be given, including the descriptions of the used networks, epidemic model, and vaccination strategies.

In accordance with previous research, the epidemic spread is simulated on small-world networks. Starting with a ring of  $N = 10,000$  nodes, each node is connected to its  $K = 4$  nearest neighbors (2 on each side). Subsequently, each edge is randomly rewired anywhere in the network with a probability  $p$  (starting node stays the same, end node uniformly random). Furthermore, each node is connected to any other node by at most one edge (i.e. no multiple edges are allowed), no edges go from and to the same node (i.e. no loops), and resulting network must be connected (i.e. no standalone islands, we can reach any node from any starting node).

Networks are generated for rewiring probabilities  $p \in \{10^k, k = -2, -1.9, \dots, 0\}$ , that is, equidistant points on the logarithmic scale between  $p = 0.01$  and  $p = 0$  (matching [Zanette and Kuperman \(2002\)](#)). For each value of  $p$ , a collection of  $2^{10}(1024)$  networks is generated and used as the sample for simulations. The number of networks is selected as a reasonable compromise between achieving a convergence of results and computational feasibility (the total "network library" is approximately 8GB in size, reaching the practical limits of personal computers).

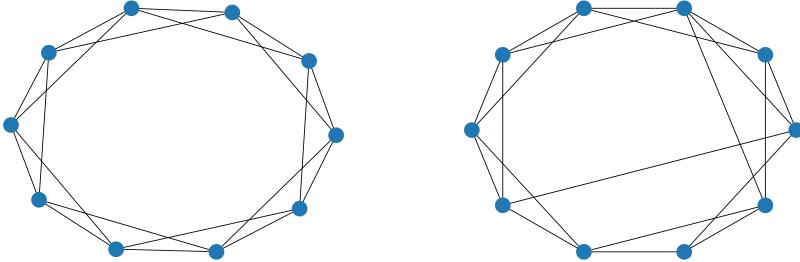


FIGURE 2.—Creation of a small-world network: starting with a regular lattice on the left ( $N = 10, K = 4$ ), each edge is rewired with a probability  $p$ . Here, with  $p = 0.05$ , 2 edges are rewired (this is random).

The model of epidemic spread used falls under the SIR framework with a pathogen gene mutation extension. The simulation is initialized with  $\rho N$  of the nodes vaccinated – those are thus immediately put into the recovered set. The value of  $\rho$  is varied throughout the simulations. The remaining nodes are all susceptible apart from one that is the initial infected. Note that this is quite a general arrangement since the network can be perceived as a subset of a larger population. This initial infected node is then just the first infection to reach from the general population into our subset. The only assumption made is that this is the only time an infection

from the outside reaches our subset; any further infection in the network can be traced back to the first infected node. For simplicity, the option of modelling with discrete time steps is chosen.

The epidemic progresses along the existing edges between nodes. First, at each time step, infected nodes transmit the disease to neighboring susceptible nodes with infection probability  $\beta$ . To be precise, each edge with an infected node at one end and a susceptible node at the other end has a probability  $\beta$  of transmitting the disease. The infection probability is set at  $\beta = 0.45$ .

Second, after the transmission stage, each infected node has a recovery probability  $\gamma$ . This excludes all nodes that have become infected at the most recent step (i.e. an infection lasts at least one time step). Recovered nodes are no longer susceptible (or infectious) and cannot be reinfected. Distinction between types of recovery is not made – the recovered set includes all those no longer susceptible or infected, be it those vaccinated, those achieving natural immunity through survival of disease, or those who have died. The simulation ends when there are no more infected nodes left. For the sake of simplicity the recovery probability is set at  $\gamma = 1$ <sup>2</sup>.

Extending this version of the basic SIR model on networks, gene mutation of the pathogen is enabled. The structure of the discrete case analyzed by [Rüdiger et al. \(2020\)](#) is followed. The pathogen has a set of gene variations  $\alpha \in \{-1, 0, 1\}$  that correspond with infection rates  $\beta \in \{0.45, 0.35, 0.95\}$ , respectively. The baseline mutation for the first infected node is set to be  $\alpha = -1$  (such that baseline  $\beta = 0.45$ , just like in the original model), with every newly infected inheriting the gene from those who infected them. At the end of each time step, all infected are subject to a chance that their pathogen will mutate. The gene mutations are a discrete Markov chain described in the diagram below:

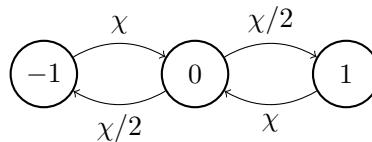


FIGURE 3.—The pathogen mutates across the gene states  $\alpha \in \{-1, 0, 1\}$  according to the probabilities specified next to the directed arrows that represent allowed state transitions.

---

<sup>2</sup>Additionally, it is believed that fixing  $\gamma = 1$  avoids a potential parameter identification issue. Intuitively it can be seen that only the ratio of the infection and recovery probabilities is relevant/identified, especially as the length of one time step is arbitrary. This paper does not formally prove this since the presence of these probabilities within a network creates additional nuance.

Here  $\chi$  is the probability that represents the mutation parameter, set to 0.004 (middle value analyzed by [Rüdiger et al. \(2020\)](#)). The infection probabilities for the gene states are specifically chosen to represent two peaks and a valley in between. The infectivity at the initial, lower peak state is set at a subcritical level (in the sense of percolation theory), and the other, higher peak, is set at a supercritical level of infectivity. This setup is, for example, said to correspond to different biological routes of transmission between hosts ([Rüdiger et al. \(2020\)](#)).

Finally, the simulation includes a selection of vaccination strategies that are to be compared. The most basic strategy is the random one:  $\rho N$  nodes are randomly selected (according to uniform probability distribution) and immunized. Due to its simplicity in both idea and real-world implementation, its performance serves as a good benchmark for the other strategies. In line with the strategies analyzed in previous research by [Zanette and Kuperman \(2002\)](#), [Xu and Sui \(2009\)](#), [Hartvigsen et al. \(2007\)](#), the "highest degree" strategy (sometimes called "targeted" or "Hubs") is included. This strategy prioritizes vaccination of nodes based on their degree (number of edges). Lastly, we also include the "highest CC" and "lowest CC" strategies that sort nodes based on their clustering coefficient (called "HCC" and "LCC" by [Hartvigsen et al. \(2007\)](#)).

The outcome of the simulation is the observed severity of the disease spread which is measured with the spread level  $r$  – the fraction of the susceptible population that has been infected. Denoting  $N_R$  as the number of recovered in the final time step, the infection level  $r$  is calculated as follows:

$$r = \frac{N_R - \rho N}{(1 - \rho)N} \quad (2)$$

Other measures of epidemic spread are available, such as maximum epidemic size (MES) and the time to it (tMES), however, the spread level  $r$  is chosen since it is a reasonable single number statistic that captures the overall size of the epidemic.

The above specified structure of epidemic simulation has been implemented and executed in Python. For pseudocode of the algorithms, please refer to [Appendix A](#).

## 5. RESULTS

The presentation of results will begin with investigating the relationship between the vaccination rate  $\rho$  and the spread level  $r$  as shown in Fig. 4.

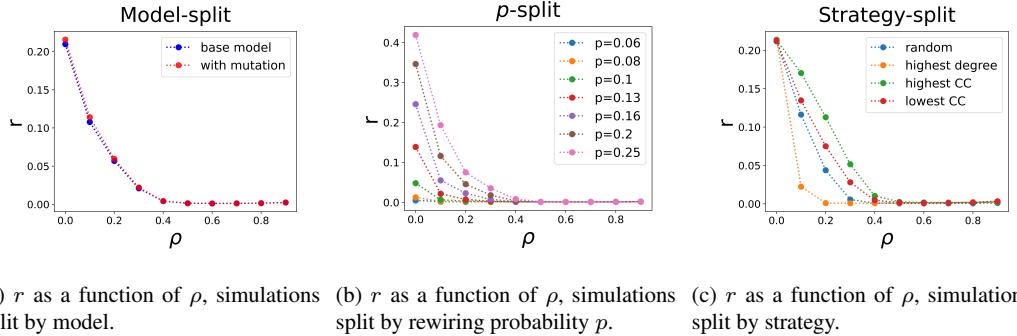


FIGURE 4.—Fraction  $r$  of the susceptible/unvaccinated population that has been infected during the disease spread, as a function of the vaccination rate  $\rho$ . For each level of  $\rho$ , the plotted spread level is the average over all simulations, with the data being partitioned in subplots (a),(b) and (c) by base model/with extension, by rewiring probability  $p$  and by strategy, respectively. Subplot (b) only shows rewiring probabilities  $p \in [0.06, 0.25]$ , since any value less is indistinguishable from  $p = 0.06$  and any value more is indistinguishable from  $p = 0.25$ .

As most clearly visible in Fig. 4a, an overall negative relationship between vaccination rate and spread level is observed. This relationship is further shown to be true for both the model with and without the mutation extension, for every level of rewiring probability  $p$  and for every vaccination strategy simulated. Albeit rather obvious, the simulations thus show undeniable (theoretical) evidence of the effectiveness of vaccination as an instrument to fight disease spread. Fig. 4c additionally serves as a first look into the effectiveness of various strategies. The more effective a strategy is, the lower its curve at any level of vaccination. The highest degree strategy is shown to clearly outperform all other strategies, including the benchmark set by the random strategy. In comparison both highest and lowest CC strategies appear to be worse than random, highest CC being the worst. It should be noted that these plots all show averages of spread level across thousands of simulations. Awareness is especially brought to the fact that Fig. 4a and Fig. 4c show averages over simulations on networks with varying levels of  $p$ . The conclusions available from these plots are thus limited to the most general trends.

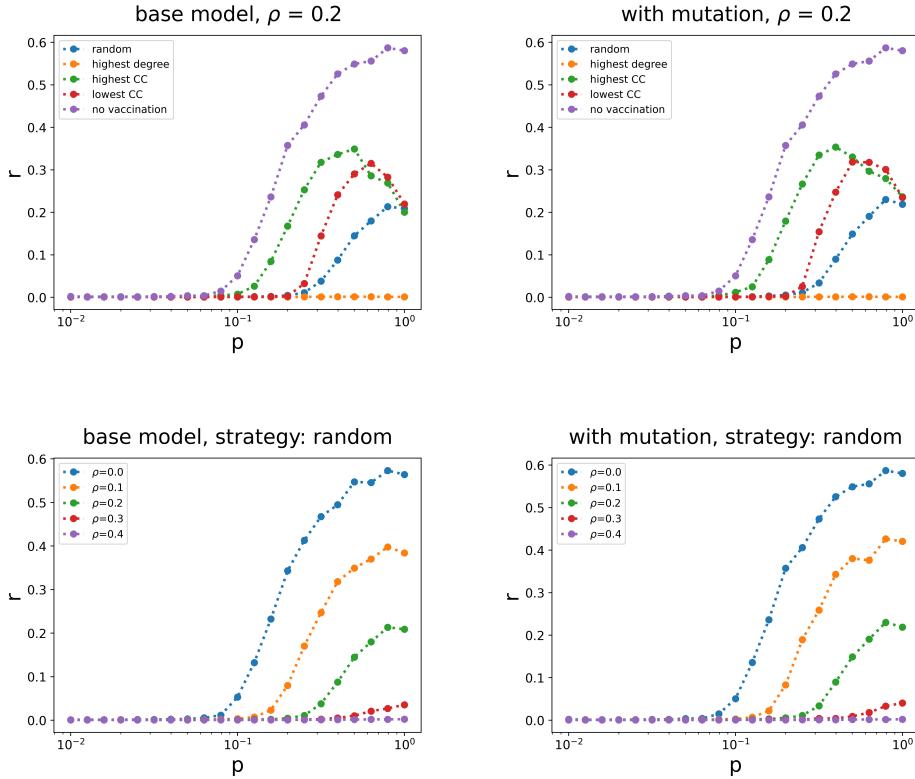


FIGURE 5.—Spread level  $r$  as a function of rewiring probability  $p$ . Top plots show this relationship for all the vaccination strategies at a vaccination rate  $\rho = 0.2$  and for a baseline curve of no vaccination ( $\rho = 0$ ). Top left plot shows this for the base model and top right plot shows this for the extended model with mutation. Bottom plots show this relationship at vaccination rates  $\rho$  ranging from 0 to 0.4, where again, the left plot is from the base model and the right plot is with data from a model with mutation. Vaccination rates  $\rho \geq 0.5$  are omitted as they appear as a flat line at the bottom of both plots. Please note the x axis of all plots is on a logarithmic scale.

Next, the relationship between  $p$  and  $r$  is investigated. Starting with the purple curves of no vaccination simulations in the top plots of Fig. 5, we note a positive relationship between them. That is to say, the more random the network is, the higher the spread level. This could be best explained by the decreasing average path length associated with higher randomness (as shown in Fig. 1). The top row plots also show a comparison between the vaccination strategies for networks of all levels of randomness. For  $p \leq 0.1$ , the disease spread is completely suppressed by all the strategies (at  $\rho = 0.2$ ). For larger values of  $p$ , differences emerge. The ordering by performance is the same as observed in Fig. 4, with highest degree being best, followed by the random strategy and with highest and lowest CC last. For  $0.5 \geq p \geq 0.1$  the lowest CC strategy is better than highest CC, however, at  $p \geq 0.6$ , they perform similarly well. Curiously,

for these high levels of randomness of the networks, these two strategies actually improve. It is important to relate this performance to the random strategy, noting that they perform worse than random, with a peak negative performance at  $p = 0.6$  and with a gradual improvement towards being only as bad as random at  $p \geq 0.6$ . A possible explanation of this improvement could be that prioritising vaccination by the clustering coefficient (both by highest and lowest) is worse than random, where with increasing randomness of the network this ordering becomes less different from random (i.e. with increasing  $p$  the distribution of CCs of nodes will have less variance) and thus closer in performance. The bottom plots of Fig. 5 show the improvements in spread level  $r$  with increasing vaccination rate  $\rho$  for all the levels of randomness in the network. With the specified infection and recovery probabilities, a vaccination rate  $\rho = 0.4$  appears to be sufficient to completely suppress disease spread. By comparison of the left and right column plots of Fig. 5 we see no notable difference between the model with and without the mutation extension.

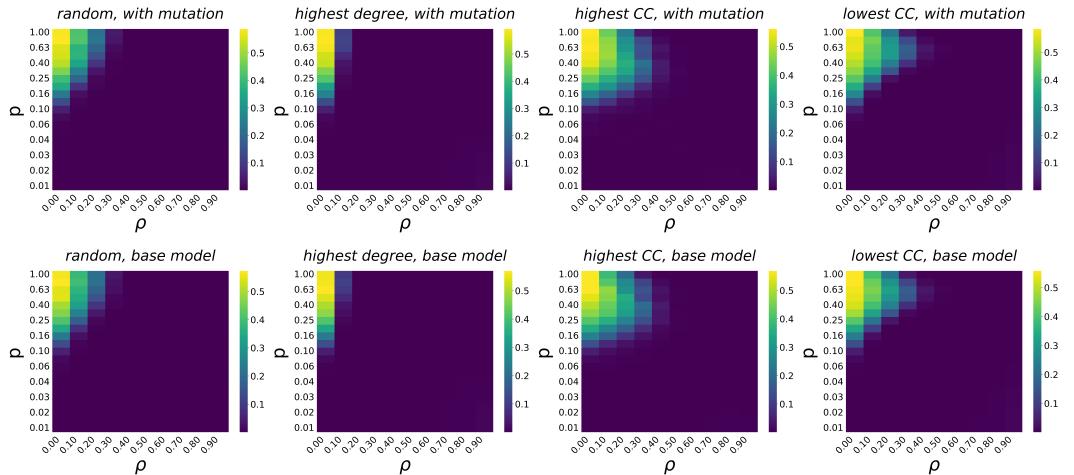


FIGURE 6.—Heat maps of vaccination strategy performance. The spread level  $r$  of the disease is indicated on a color scale for all combinations of vaccination rate  $\rho$  (x axis) and network rewiring probability  $p$  (y axis, log scale). The brighter the color, the higher the spread level. The simulation data is divided into subplots based on vaccination strategy (columns) and whether the model includes a mutation extension (top row), or not (bottom row). Each pixel in the heat maps represents an average of the spread level  $r$  over 1024 networks on which the disease spread was simulated (for that parameter combination).

Lastly, the overall performance of various strategies is assessed in Fig. 6. Comparing across columns, the highest degree strategy is undoubtedly the best, followed by the random strategy and with highest and lowest CC in joint last place. Additionally, as already seen in Fig. 4a and Fig. 5, no clear difference between the model with and without the mutation extension is

observed. Following this result, the heat maps were recreated for  $\chi = 0.01$ , the highest value of the mutation probability parameter studied by Rüdiger et al. (2020). At this level of mutation, Rüdiger et al. (2020) have found most of the simulations to end with a majority of the gene pool at  $\alpha = 1$ , associated with the supercritical level of infectivity ( $\beta = 0.95$ ). The heat maps still show no difference between vaccination strategy performance ordering for the model with and without the mutation extension (see Fig. B1).

## 6. DISCUSSION

Findings about the base model of this paper are in agreement with past research. A negative relationship between vaccination rate  $\rho$  and spread level  $r$  has been found by Zanette and Kuperman (2002) (Fig. 3), Hartvigsen et al. (2007) (Fig. 2) and Xu and Sui (2009) (Fig. 11). This finding is the most intuitive (and maybe obvious), since larger vaccination rates leading to lower infection spread (for most diseases and most vaccinations) is a well proven fact (Andre et al. (2008)). The mathematical explanation for why it is indeed observed in the simulation lies in the reduction in the number of edges in the network. Each vaccinated node renders all its connecting edges unable to transmit the disease and thus apart from protecting itself, each vaccinated node also reduces the number of avenues for infection for all its neighbours. The positive relationship between network rewiring probability  $p$  and spread level  $r$  is also a well researched phenomenon (Zanette and Kuperman (2002), Fig. 1; Hartvigsen et al. (2007), Fig. 1; Xu and Sui (2009), Fig. 13) and the ranking of the vaccination strategies' relative performance of this paper is in agreement with both Zanette and Kuperman (2002) and Hartvigsen et al. (2007).

While the simulation of a model with mutating infectivity of the virus has been researched (Rüdiger et al. (2020)), the comparison of vaccination strategies in that model is novel. Beyond no qualitative difference in outcome of the simulation in the extended model, there appears to be no significant difference between modelling with and without the mutation extension. This finding is of course conditional on the exact setup of the simulation, especially on the mutation model with its associated infection probabilities for each gene state and with the mutation probability  $\chi$ . The simulation has been conducted with  $\chi = 0.004$ , the middle value of mutation probability tested by Rüdiger et al. (2020). Following the finding of no difference between the original and extended model, the simulation is repeated with  $\chi = 0.01$ . This is the highest value of mutation probability investigated in the discrete model by Hartvigsen et al.

(2007), where they have shown it, on average, to lead to most of the nodes to end in the supercritical state (where  $\beta = 0.95$ ). In spite of that this simulation has still found no qualitative difference in outcome between the original and extended model. If confirmed, these findings suggest a robustness of the simulations to the mutation extension. From the point of view of real world consequences this would be the better outcome, since it means that when deciding on a vaccination strategy, the mutations of the disease need not be taken into account. It must be remembered that the mutations in question are only those in level of infectivity, not in resistance to the vaccine or in any other aspects.

The simulation conducted in this paper has encountered some practical limitations. Given the computational power available to the author, each parameter configuration has been simulated only on  $2^{10}$ (1024) networks. The majority of the previous research cited here has used  $10^4$  networks per configuration. For the same reason additional values of any of the parameters (including vaccination strategy) have not been simulated, despite a discovered interest in the mutation probability  $\chi$ <sup>3</sup>. Further research could overcome these issues either by brute computational strength or by improved design. Particularly worthy of mention is the idea of a sample of samples, that is, reusing each of  $n$  networks for  $m$  simulations (per parameter configuration). The potential advantage of this comes from the fact that a majority of the code runtime is actually spent on opening the pre-created networks and not the epidemic simulation itself<sup>4</sup>.

Multiple directions of potential future research are available. First, the findings should be repeated to be confirmed. Such research could in addition contain a larger parameter space by having more networks, more vaccination strategies and/or more values of mutation probability. Second, the limits of the found robustness of relative vaccination strategy performance to mutation extension could be explored. Of particular interest could be the continuous case of the model and different types of mutation, including mutating vaccine resistance. Last, a model extension with various susceptibility of individuals to infection could be explored. This is relevant in the context of vaccination strategies because a viable strategy is the prioritisation of

---

<sup>3</sup>The conducted simulation already consisted of 1680 distinct parameter configurations. With  $\chi = 0.01$  also explored, this number and the code runtime have risen by 50%.

<sup>4</sup>With  $N = 10^4$  nodes in each network, the network creation is about 8 times slower than loading a pre-created network. Also, networks have already been reused for every parameter configuration (of which there are 80 per rewiring probability  $p$ ). Discounting the time spent to pre-create a "network library", this setup was thus 640 times faster and more flexible than without using the network library. Network generation was done optimally (implemented in the `networkx` python package) and so was the loading of the files (using the `cpickle` python package).

susceptible individuals and since exactly this strategy has largely been implemented across the whole world in the recent COVID-19 pandemic.

## 7. CONCLUSION

This paper has investigated the performance of vaccination strategies in a SIR model with the extension of mutating infectivity of the pathogen. The research took the form of a simulation of the disease spread on small-world networks. Consistent with past research on the base model, a negative relationship was found between the vaccination rate and the spread level, and a positive relationship between the network rewiring probability and the spread level. Furthermore, the relative performance of vaccination strategies was evaluated and their ranking was also found to be consistent. Novel was the comparison between the original model and one with a mutation extension. No significant difference was found, suggesting the conclusion that relative performance of vaccination strategies is robust to the model extension of mutating infectivity. Additional research is required to verify this finding beyond the specific setup of the simulation conducted in this paper. Directions of further research have been suggested, such as investigating other forms of pathogen mutation.

---

APPENDIX A: CODE

---

The pseudocode of the algorithms used in the simulations is given below. The overall structure of the function simulating the SIR model of epidemic spread is inspired by the python EoN package ([Miller and Ting \(2019\)](#)) , created as complement to [Kiss et al. \(2017\)](#). For the use of networks in code, the package `networkx` is used ([Hagberg et al. \(2008\)](#)).

---

**Algorithm 1** Epidemic simulation

---

**Input:** Network  $G$ , recovery rate  $\gamma$ , vaccination rate  $\rho$ , mutation probability  $\chi$ , vaccination strategy  $strategy$ , maximum time  $t_{max}$

**Output:** Lists  $times$  and  $S, I, R$ ; giving a number in each state at each time step.

```

1: function SIR( $G, \gamma, \rho, \chi, strategy, tmax = 10^3$ )
2:    $N \leftarrow$  number of nodes in  $G$ 
3:    $t, S, I, R \leftarrow [0], [N(1 - \rho) - 1], [1], [\rho N]$ 
4:    $vaccinated \leftarrow$  using  $strategy$  select  $N\rho$  nodes from  $G$ 
5:    $infected \leftarrow$  with uniform probabilities select randomly 1 node from  $G$  not in  $vaccinated$ 
6:    $alpha[infected] \leftarrow -1$ 
7:   for all  $u \in G$  do
8:      $susceptible[u] \leftarrow$  True if  $u$  not in  $vaccinated$  or  $infected$ 
9:   end for
10:  while  $infected$  is not empty and  $t[-1] < tmax$  do
11:     $new\_infected \leftarrow$  empty set
12:    for all  $u \in infected$  do
13:      for all neighbors  $v$  of  $u$  do
14:        if  $v$  is  $susceptible$  and  $test\_infection(\beta, mutation, alpha[u])$  then
15:          add  $v$  to  $new\_infected$ 
16:           $v$  no longer susceptible
17:           $alpha[v] \leftarrow alpha[u]$ 
18:        end if
19:      end for
20:      if  $RandomFloat(0,1) < \gamma$  then
21:         $u$  becomes recovered
22:      else
23:        add  $u$  to  $new\_infected$ 
24:      end if
25:    end for
26:    for all  $u \in new\_infected$  do
27:      update gene  $alpha[u] \leftarrow mutation\_test(alpha[u], \chi)$ 
28:    end for
29:    update lists  $t, S, I, R$ 
30:  end while
31:  return  $t, S, I, R$ 
32: end function

```

---

FIGURE A.1.—The function to simulate an epidemic using the SIR model with the extension of gene mutation.

To have the original SIR model without mutation, it is enough to simply remove all lines concerning genes/mutation and to replace `test_infection` function by a simple sampling of the  $\beta$  infection probability.

Complementary to the `SIR` function are the `test_mutation` and `test_infection` functions specified below. The `SIR` function also requires a vaccination strategy. The options used in this paper are specified in the [Methodology](#) section, however, their pseudocode is not provided.

---

### Algorithm 2 Gene mutation

---

**Input:** mutation probability  $\chi$ , gene  $\alpha$   
**Output:** new gene value  $\alpha$

```

1: function MUTATION_TEST( $\alpha, \chi$ )
2:    $P \leftarrow$  transition matrix of the mutation Markov Chain in Fig.3
3:    $states \leftarrow [-1, 0, 1]$ 
4:    $i \leftarrow$  index in vector  $states$  of current gene  $\alpha$ 
5:    $\alpha \leftarrow$  select value from  $states$  at random according to distribution  $p = row_i(P)$ 
6:   return  $\alpha$ 
7: end function

```

---

FIGURE A.2.—The mutation function. For any given state  $\alpha$ , one step along the specified gene mutation Markov Chain is taken to get the gene state  $\alpha$  in the next time step.

---

### Algorithm 3 Transmission test

---

**Input:** gene  $\alpha$   
**Output:** boolean giving whether transmission occurs or not

```

1: function TEST_INFECTON( $\alpha$ )
2:    $p \leftarrow [0.45, 0.35, 0.95]$                                  $\triangleright$  for  $\alpha = [-1, 0, 1]$ , respectively
3:    $i \leftarrow$  index in vector  $[-1, 0, 1]$  of current gene  $\alpha$ 
4:   return RandomFloat(0,1)  $< p[i]$ 
5: end function

```

---

FIGURE A.3.—The function to test possible transmission from one node to another. Probability of transmission selected according to gene of pathogen.

## APPENDIX B: ADDITIONAL FIGURES

The heat maps below show the result of repeating the whole simulation with mutation parameter  $\chi = 0.01$ . Even at this high level of mutation, no significant difference is found between the base model and the extended one. This is cause to suggest a certain robustness of vaccination strategy performance to the mutating infectivity extension.

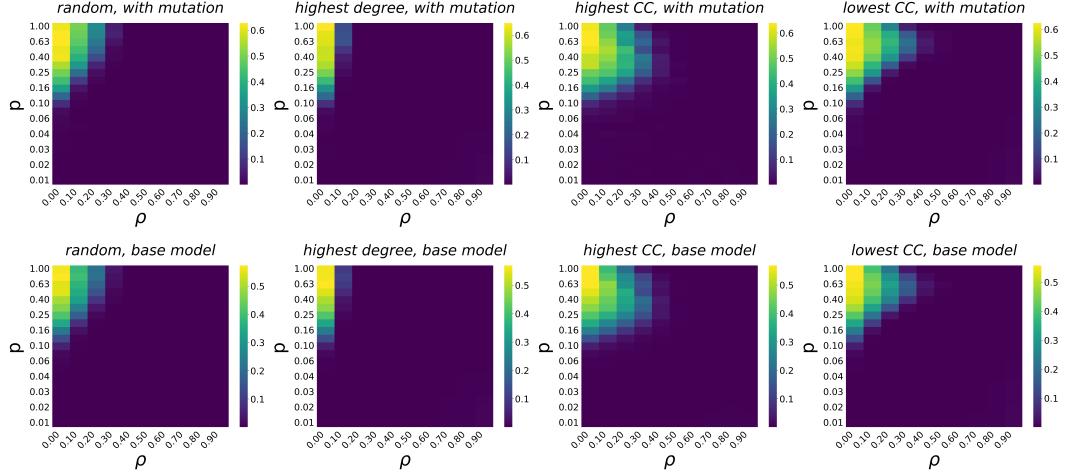


FIGURE B.1.—Heat maps of vaccination strategy performance at  $\chi = 0.01$ . The spread level  $r$  of the disease is indicated on a color scale for all combinations of vaccination rate  $\rho$  (x axis) and network rewiring probability  $p$  (y axis, log scale). The brighter the color, the higher the spread level. The simulation data is divided into subplots based on vaccination strategy (columns) and whether the model includes a mutation extension (top row), or not (bottom row). Each pixel in the heat maps represents an average of the spread level  $r$  over 1024 networks on which the disease spread was simulated (for that parameter combination).

## REFERENCES

- ANDRE, FRANCIS E, ROBERT BOOY, HANS L BOCK, JOHN CLEMENS, SIBNARAYAN K DATTA, THEKKEKARA J JOHN, BEE W LEE, S LOLEKHA, HEIKKI PELTOLA, TA RUFF, ET AL. (2008): “Vaccination greatly reduces disease, disability, death and inequity worldwide,” *Bulletin of the World health organization*, 86, 140–146. [12]
- AZIZI, ASMA, CESAR MONTALVO, BALTAZAR ESPINOZA, YUN KANG, AND CARLOS CASTILLO-CHAVEZ (2020): “Epidemics on networks: Reducing disease transmission using health emergency declarations and peer communication,” *Infectious Disease Modelling*, 5, 12–22. [3, 4, 5]
- BARABÁSI, ALBERT-LÁSZLÓ (2013): “Network science,” *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371 (1987), 20120375. [3]
- BATAGELJ, VLADIMIR AND ULRIK BRANDES (2005): “Efficient generation of large random networks,” *Physical Review E*, 71 (3), 036113. [4]
- ERDŐS, PAUL, ALFRÉD RÉNYI, ET AL. (1960): “On the evolution of random graphs,” *Publ. Math. Inst. Hung. Acad. Sci.*, 5 (1), 17–60. [3]
- GRANOVETTER, MARK (1983): “The strength of weak ties: A network theory revisited,” *Sociological theory*, 201–233. [3]
- HAGBERG, ARIC A., DANIEL A. SCHULT, AND PIETER J. SWART (2008): “Exploring Network Structure, Dynamics, and Function using NetworkX,” in *Proceedings of the 7th Python in Science Conference*, ed. by Gaël Varoquaux, Travis Vaught, and Jarrod Millman, Pasadena, CA USA, 11 – 15. [15]
- HARTVIGSEN, G, JM DRESCH, AL ZIELINSKI, AJ MACULA, AND CC LEARY (2007): “Network structure, and vaccination strategy and effort interact to affect the dynamics of influenza epidemics,” *Journal of theoretical biology*, 246 (2), 205–213. [3, 4, 5, 8, 12]
- KISS, ISTVÁN Z, JOEL C MILLER, PÉTER L SIMON, ET AL. (2017): “Mathematics of epidemics on networks,” *Cham: Springer*, 598, 31. [3, 4, 15]
- MILLER, JOEL AND TONY TING (2019): “EoN (Epidemics on Networks): a fast, flexible Python package for simulation, analytic approximation, and analysis of epidemics on networks,” *J. Open Source Softw.*, 4 (44), 1731. [15]
- RÜDIGER, STEN, ANTON PLIETZSCH, FRANCESC SAGUÉS, IGOR M SOKOLOV, AND JÜRGEN KURTHS (2020): “Epidemics with mutating infectivity on small-world networks,” *Scientific reports*, 10 (1), 5919. [3, 4, 5, 7, 8, 12]
- WATTS, DUNCAN J AND STEVEN H STROGATZ (1998): “Collective dynamics of ‘small-world’networks,” *nature*, 393 (6684), 440–442. [3, 4]
- XU, ZENGWANG AND DANIEL Z SUI (2009): “Effect of small-world networks on epidemic propagation and intervention,” *Geographical Analysis*, 41 (3), 263–282. [3, 4, 5, 8, 12]
- ZANETTE, DAMIÁN H AND MARCELO KUPERMAN (2002): “Effects of immunization in small-world epidemics,” *Physica A: Statistical Mechanics and its Applications*, 309 (3-4), 445–452. [3, 4, 5, 6, 8, 12]