



Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Original Research Article

Modeling of network based digital contact tracing and testing strategies, including the pre-exposure notification system, for the COVID-19 pandemic[☆]



Daniel Xu

MIT PRIMES-USA, Massachusetts Institute of Technology, Cambridge, MA, United States of America

ARTICLE INFO

Keywords:

Testing strategies
Digital contact tracing
COVID-19
Infectious disease epidemiology
Mathematical modeling

ABSTRACT

With more than 1.7 million COVID-19 deaths, identifying effective measures to prevent COVID-19 is a top priority. We developed a mathematical model to simulate the COVID-19 pandemic with digital contact tracing and testing strategies. The model uses a real-world social network generated from a high-resolution contact data set of 180 students. This model incorporates infectivity variations, test sensitivities, incubation period, and asymptomatic cases. We present a method to extend the weighted temporal social network and present simulations on a network of 5000 students. The purpose of this work is to investigate optimal quarantine rules and testing strategies with digital contact tracing. The results show that the traditional strategy of quarantining direct contacts reduces infections by less than 20% without sufficient testing. Periodic testing every 2 weeks without contact tracing reduces infections by less than 3%. A variety of strategies are discussed including testing second and third degree contacts and the pre-exposure notification system, which acts as a social radar warning users how far they are from COVID-19. The most effective strategy discussed in this work was combining the pre-exposure notification system with testing second and third degree contacts. This strategy reduces infections by 18.3% when 30% of the population uses the app, 45.2% when 50% of the population uses the app, 72.1% when 70% of the population uses the app, and 86.8% when 95% of the population uses the app. When simulating the model on an extended network of 5000 students, the results are similar with the contact tracing app reducing infections by up to 79%.

1. Introduction

More than 1.7 million people have died of COVID-19 [1]. With such catastrophic loss of life at risk, it is a top priority to identify effective measures to prevent the spread of COVID-19. Preventing the spread of COVID-19 has several challenges. Firstly, 44% of COVID-19 transmissions occur during the asymptomatic stage with infected individuals being most infectious 1 to 2 days before symptom onset [2]. Secondly, according to the Center for Disease Control and Prevention (CDC) Planning Scenarios report [3], 40% of all COVID-19 cases remain asymptomatic and this number could be as high as 79% for those under 20 [4]. The CDC Planning Scenario also estimates that asymptomatic individuals are 75% as infectious as symptomatic individuals. Finally, the false-negative rate of RT-PCR tests on presymptomatic individuals ranges from 67% to 100% [5]. All of these factors make it very difficult to prevent COVID-19 with symptom-based measures.

Contact tracing has existed for decades, helping to reduce tuberculosis [6], sexually transmitted diseases [7], and Ebola [8]. However, manual contact tracing relies on substantial human labor. According to a recent survey by National Public Radio, 39 states do not have enough contact tracers [9]. Additionally, of the contacts reported, only around

50% were reachable by contact tracers [10]. Health officials estimate an additional \$12 billion dollars are required to fund the 180,000 manual tracers needed [11]. Manual contact tracing is also susceptible to errors with the United Kingdom's Test and Trace service losing 15,000 positive COVID-19 cases between September 25 and October 2, 2020 [12].

Digital contact tracing is a relatively new method of fighting pandemics. Although the Bluetooth technology for digital contact tracing was first validated in 2014 [13], contact tracing apps have only recently been implemented to fight the COVID-19 pandemic. Since digital contact tracing apps require a critical mass of a population in order to be effective, a key goal is to convince enough people to use the app. In [14], it is suggested that targeting small communities like universities first will allow the app to be used by enough people within that community. This is more feasible than expecting a significant proportion of the population at large to use the app. Many universities have experienced outbreaks [15] and are looking for ways to prevent further spread. Currently, Georgia Tech, Carnegie Mellon, Grand Valley State University, and even the city of Santa Fe are beginning to adopt contact tracing apps like NOVID [16]. Currently, there are not enough

[☆] This project was completed under the MIT PRIMES program.

E-mail address: danielxu@mit.edu.

users within these communities to prove the effectiveness of contact tracing apps. We will use simulation to evaluate the effectiveness of digital contact tracing.

Efforts to model diseases date back to the 1920s with the creation of the SIR (Susceptible, Infectious, Recovered) model [17] where people in a fully mixed population are modeled as being Susceptible, Infectious, or Recovered. The SEIR (Susceptible, Exposed, Infectious, Recovered) model [18] is a variant of the SIR framework used for diseases with longer incubation periods. It includes the Exposed stage which are infected individuals who do not show symptoms. While all individuals are asymptomatic during the exposed stage, we assume a certain proportion, p , of individuals remain asymptomatic during the Infectious stage as well. More recently, network based models, where humans are vertices and contacts are edges, have been adopted. Epidemiological models have been instrumental in encouraging preventative measures in the COVID-19 pandemic such as masking [19,20], social distancing [21,22], testing [23,24]. In particular, [19,23] are network based models.

Since digital contact tracing relies on the exact contacts that occur in a population, it is best modeled with a contact network where people are vertices and contacts are edges. Recent efforts to simulate digital contact tracing [25,26] have been done using synthetic networks where households and communities are constructed with random processes such as full mixing, and each pair of people have the same probability of contact. However, very little is known about the exact structure of human contacts and how these structures affect digital contact tracing. Additionally, those models assume perfect COVID-19 tests which changes optimal strategies of digital contact tracing and testing.

This paper presents an enhanced network based SEAIR (Susceptible, Exposed, Asymptomatic, Infectious, Recovered) model of the COVID-19 pandemic with digital contact tracing and testing strategies that incorporates variations in infectivity and test sensitivity. Our SEAIR model further refines the population into symptomatic individuals, who show symptoms during infection, and asymptomatic individuals, who do not. In contrast to previous work on digital contact tracing, the networks in the model are generated from a real world data set of interactions among 180 students of a high school in France [27]. Since the data set was only recorded over 7 days, we use a MUNGE-like heuristic to generate additional days for the model. We present a new method to extend temporal weighted graphs in order to perform simulations on a larger population of 5000 people.

Our model incorporates test sensitivity and shows that it has a significant impact on digital contact tracing strategies. The model simulates a new digital contact tracing strategy developed by NOVID [16] called the pre-exposure notification system. The pre-exposure notification system acts like a “social radar” telling app users how close they are to the nearest COVID-19 infections, allowing them to take extra precautionary measures. In contrast with traditional SEIR models, test sensitivity is modeled to change over time depending on the time of symptom onset [5]. The incubation period is sampled from the COVID-19 incubation period distribution [28]. The infectiousness of an individual is modeled to change over the infectious period and is fitted to the function presented in [2].

The model shows that the traditional strategy of quarantining direct contacts reduces infections by less than 10% when more than half the population is asymptomatic. Testing second and third degree contacts reduces infections by up to 40% when 70% of the population uses the app. The pre-exposure notification system reduces infections by an additional 43% and reduces the number of quarantines required by 51%. Quarantining second degree contacts reduces infections but leads to a high number of quarantines. If large proportions of the population are asymptomatic, periodic testing reduces infections by an additional 41%. However, periodic testing without tracing reduces infections by only 3%. The most effective strategy discussed in this work was combining the pre-exposure notification system with testing second and third degree contacts. This strategy reduces infections by

18.3% when 30% of the population uses the app, 45.2% when 50% of the population uses the app, 72.1% when 70% of the population uses the app, and 86.8% when 95% of the population uses the app. When simulating the model on an extended network of 5000 students, the results are very similar with the contact tracing app reducing infections by up to 79%.

1.1. Paper outline

In Section 2, we present the contact network generation process. Section 3 outlines the SEAIR based model of COVID-19 spread. In Section 4, we introduce the contact tracing app into the model including the pre-exposure notification system. Section 5 introduces the contact tracing and testing strategies which are tested in the results. In Section 6 we present the results of 5 simulated scenarios. In Section 6.1, the traditional strategy of quarantine of first degree contacts is investigated. In Section 6.2, testing of second and third degree contacts is incorporated. In Section 6.3, the pre-exposure notification system is investigated. Section 6.4 investigates periodic testing. Section 6.5 simulates the model on an extended version of the graph. In Section 7, we provide a summary and concluding remarks.

2. Model of contact network

The first component of the model is the contact network which encapsulates the interactions between individuals in the simulation. The model iterates through discrete timestamps each representing a day in the simulation. Each day in the model, individuals come into contact, and these contacts ultimately determine how the virus transmits. The contact network on day t , G_t is defined to be a graph where each vertex v_1, \dots, v_n represents an individual in the model and the edge weight e_{ij} between v_i and v_j represents the total amount of time person i and person j have spent in contact with each other during this time step. We do not distinguish between times of the day of the contact or number of contacts between the same pair of individuals. This is in accordance with a recent CDC policy change that defines a close contact to be measured using cumulative contact time over the course of a day [29].

Traditional SEIR models assume uniform mixing where each individual has the same probability of coming into contact with every other individual. Since digital contact tracing involves the exact contacts that occur in a population, a realistic contact network is needed. Temporal networks in the model are generated using a publicly available high-resolution data set [27] which used RFID (Radio Frequency Identification Devices) to record all contacts between 180 students from Lycée Thiers high school in France over the course of 7 days.

Since the simulation is over longer periods of time, we present a heuristic for generating additional days. Define H_a to be the contact network in day a of the data set where $1 \leq a \leq 7$. The edge weight $H_{a,i,j}$ is the duration of contact between people i, j on day a from the data set. The heuristic is similar to the MUNGE algorithm [30] which generates synthetic training data. The algorithm picks an initial data value, and for each feature, with a certain probability, swaps that feature with the feature of its nearest neighbor. In this heuristic, we use a random day rather than the nearest neighbor. Initially, 2 distinct days a, b are chosen at random from the data set. Starting with the contact graph for day a , for the contact duration between v_i and v_j with probability 0.5, we replace that contact duration with $H_{b,i,j}$. On average, the generated contact network G_t has half of its edges equal to the corresponding edges in H_a , and the other half equal to H_b .

To simulate the app on larger networks, we present a method to generate larger networks from the original data set. We create a modified version of the Albert–Barabási process that is adjusted for constructing weighted temporal graphs and maintains the average degree of the vertices.

The principal idea of the Albert–Barabási process is that nodes of high degree are more likely to interact with new nodes. Given a

temporal graph G with n vertices where the graph on timestamp t is G_t , extra people are added one at a time with the following process:

A new vertex v is added, and for each existing w , a random third vertex u is chosen. The temporal edge weight between v and w is determined to be the same as the edge weight between w and u : $G_{t,v,w} = G_{t,u,w}$ for all t . This step is the same as in Albert–Barabási except adjusted to fit a weighted temporal graph. To preserve the average degree of the vertices, each of the original edges of G are deleted with probability $\frac{1}{n-1}$. Note that the generated graph no longer has the scale-free property.

The extended network is initialized to be the original graph in the data set and extended in accordance to the above procedure to a total of 5000 individuals.

3. Model of COVID-19 spread

The SEIR framework organizes the people in the simulation into one of the following four states: Susceptible (S), Exposed (E), Infectious (I), or Recovered (R). Individuals in the Susceptible stage have not been infected and therefore are susceptible to the virus. Exposed individuals have been infected with the virus but do not yet show symptoms. The Infectious stage begins when infected individuals show symptoms. Finally, the Recovered stage contains individuals who have recovered or otherwise removed from the model and are now immune from further spread. Individuals in both the Exposed and Infectious stage can infect others. An individual moves onto the next stage according to the following rules:

- $S \rightarrow E$: Individuals in the susceptible stage can only be exposed if they come into contact with an infected individual. Infections will be discussed in later sections.
- $E \rightarrow I$: After being exposed to a virus, the period of time before symptom onset is called the Incubation period, typically within 2 to 14 days. According to [28], the distribution of incubation periods is approximately the log-normal distribution with parameters $\mu = 1.621$, $\sigma = 0.418$. The log-normal distribution is defined as follows: Let Z be a random variable with normal distribution with mean μ and deviation σ , then $X = e^Z$ where X is the random variable with log-normal distribution with parameters μ, σ . To determine the incubation period of an individual, we take a random sample from this distribution rounded to the nearest positive integer.
- $I \rightarrow R$: Individuals with moderate symptoms stop being infectious around 10 days from symptom onset [31]. As in traditional SEIR models, we assume that every time stamp after symptom onset, there is a probability $\lambda = 0.11$ which is $\frac{1}{9}$ that an Infectious person recovers or is removed from the model.

As in traditional SEIR models, the functions $S(t), E(t), I(t), R(t)$ are the number of individuals in the compartments susceptible, exposed, infections, and recovered respectively at time t . We defined $Q(t)$ to be the number of individuals quarantined at time t but have not received a positive test result and $T(t)$ to be the number of individuals who have received a positive before or during time t but have not recovered. Individuals counted in $T(t)$ are in quarantine as well. Quarantines will be discussed in Section 4. In the rest of the paper, for each individual v_i , we will refer to E_{v_i}, I_{v_i} , and R_{v_i} as the time of exposure, symptom onset, and recovery respectively of v_i . Additionally, T_{v_i} is the time of the first positive test of v_i .

In realistic scenarios, individuals can infect others even before symptom onset and the level of infectiousness changes throughout the infection. As in [2], the infectiousness of an individual is directly proportional to the amount of viral shedding. In [2], the amount of viral shedding of an individual over the course of the infection is calculated using the distribution of the serial interval and distribution of the incubation period. We will denote the function giving the infectiousness profile by $ID(t)$ where the input t is the number of days since symptom

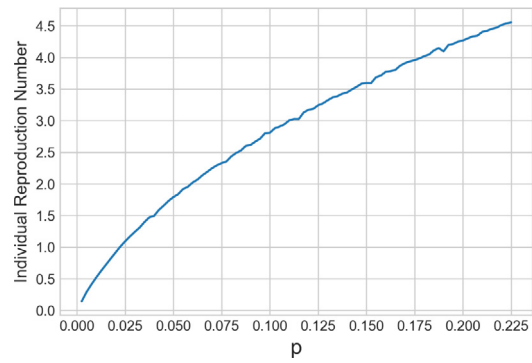


Fig. 1. Probability of infection vs individual reproduction number.

onset, and the output $ID(t)$ is the density of viral shedding on day t . For example, since $ID(1)$ is approximately 0.11, around 11% of the total viral shedding happens 1 day after symptom onset. The values of $ID(t)$ are taken directly from the function in [2] where infectiousness was assumed to start 5 days before symptom onset.

The function $ID(t)$ allows us to introduce variations in infectiousness throughout the infection. The infectiousness of an individual, which is defined to be the probability of infection given a 20-second contact, is some constant multiple, p , of $ID(t)$. The value of p will be calculated based on the R_0 value.

Consider the graph G_t to be the contact network during timestamp t in the model, with vertices v_1, \dots, v_n representing people and edge weight e_{ij} to be the total contact duration between v_i, v_j during day t . Assuming infection to be an event that can happen during the course of a contact, if v_i is in either the exposed or infectious compartments, and v_j is susceptible, we model the probability that v_j becomes exposed to be

$$1 - (1 - pID(t - I_{v_i}))^d$$

where $t - I_{v_i}$ is the number of days since symptom onset of v_i .

The basic reproduction number, or R_0 , is the expected number of secondary infections caused by a single infectious individual. Estimates of the value of R_0 range from 1.5 to 6.7 with the median value being 2.8 [32] and the methods used to estimate R_0 varies widely between studies [33]. Through simulation, we calculate the value of R_0 as a function of p as shown in Fig. 1. The value of p is ranged in increments of 0.0025 from 0 to 0.225 for a total of 100 values. For each value of p , we run the simulation 1,800 times where each individual is the seed infection 10 times and the R_0 value is measured as the average number of secondary infections caused by the seed infection. As shown in Fig. 1, the scenario when $p = 0.10$ yields an R_0 value of 2.8. Based on an additional 18,000 trials for $p = 0.1$ where each individual is the seed infection 100 times, the confidence interval is 2.76 to 2.81 for 95% confidence. In the rest of this work, the value of p will be set to 0.10.

Our model refines the population to include asymptomatic individual. According to [3], 40% of the population remain asymptomatic throughout their infectious periods. Again, estimates vary widely and it is noted that this value could range from 10% to 70%. In [34], it is estimated that over 80% of young individuals are asymptomatic. At the beginning of each simulation, some individuals are randomly selected to be asymptomatic. The amount of asymptomatic individuals is used as an independent variable. Asymptomatic individuals do not show symptoms if they are infected. Note that the status of an individual as asymptomatic is not related to whether they are infected, but whether they will or will not show symptoms if infected. We test the possibilities when the expected proportion of asymptomatic individuals is 20%, 40%, 60%, or 80%. Asymptomatic individuals have similar viral loads to symptomatic individuals [35]. Thus, we will assume that asymptomatic individuals have the same infectiousness as symptomatic individuals.

Testing is a key component of contact tracing. The contact tracing app relies on positive tests to identify cases of COVID-19. We adjust test sensitivities based on days from symptom onset to the distribution calculated in [5]. The false-negative rate of RT-PCR COVID-19 tests are between 67% to 100% before symptom onset, and fall to 20% to 40% after symptom onset [5]. The median number of days from symptom onset to taking a test is 3 days with interquartile range 1 to 6 [3]. We approximate this distribution by assuming that each day there is a probability of 19% that symptomatic individuals are tested. This preserves the median of 3 days and interquartile range 1 to 6. We do not account for false-positive tests or individuals symptomatic with a disease unrelated to COVID-19. Test results are received after a delay of 1 day. This is similar to the delay on university campuses [36]. After receiving a positive COVID-19 test result, we assume the person will remain in quarantine until recovery.

4. Model of contact tracing app

The contact tracing app has incomplete information about the contact network and the states of the individuals. Ultrasound apps such as NOVID can measure distances to the resolution of inches and detect contacts with accuracy over 99.6% [16]. We will assume all contacts between individuals with the app are sensed. Thus, the contact tracing app can detect the subgraph of the contact map induced by the set of app users. Contact tracing apps support the consideration of degree k contacts rather than just direct contacts. This is not possible with manual contact tracing. A contact is defined to be a tuple (p_1, p_2, t, d) that contains the following four pieces of information: p_1, p_2 are the people involved in the contact, t is the day of contact, and d is the duration of the contact measured in units of 20 s. Since the digital contact tracing app cannot determine the time of symptom onset, the estimated transmission probability of the contact is $(1 - (1 - p')^{d_i})$. The value $p' = \widehat{ID}p$ is substituted for $pID(t)$ where $\widehat{ID} = 0.11$ is the average value of $ID(t)$ during the 6 most infectious days. In addition to contacts, users report all positive test results. If a user has tested positive, we assume that this user will report recovery. In the contact tracing app, we define a degree k contact to be a sequence of contacts c_1, \dots, c_k that satisfy the following properties:

1. The individuals p_1, \dots, p_{i+1} involved in the contacts form a chain: $c_i = (p_i, p_{i+1}, t_i, d_i)$.
2. p_1 has reported a positive test: $t_1 \leq T_{p_1} \leq t_1 + 10$. Through simulation of 1800 outbreaks without interventions, 95% of all symptomatic individuals took a test within 8 days of symptom onset. Since infectiousness becomes significant 2 days before symptom onset, the contacts from 10 days before the test must be recalled.
3. p_i, p_{i+1} have not reported recovery and thus can transmit or catch the virus: $R_{p_i}, R_{p_{i+1}} > t_i$ for each i .
4. The serial interval is between 1 and 10 days: $t_i + 1 \leq t_{i+1} \leq t_i + 10$. The serial interval is defined to be the duration from the exposure time of the infector (t_i) to the exposure time of the infected (t_{i+1}). Through simulation of the model 1800 times, 94% of all serial intervals are within 10 days.

The weight of this contact chain is the product of the estimated transmission probabilities of each contact:

$$\prod_{i=1}^k (1 - (1 - p')^{d_i}).$$

First and second degree contacts are computed using contact data from previous days directly after p_1 reports a positive test. Larger degree contacts are computed recursively. For each person v_i in the model, we keep track of a matrix M_i where $M_{i,l,t}$ is the total sum of contact chains of length l that affect person v_i at time t . Using the contacts, the app can calculate M_i recursively.

On day t , if v_i has contacts with v_{a_1}, \dots, v_{a_k} with durations d_1, \dots, d_k respectively, where $v_i, v_{a_1}, \dots, v_{a_k}$ have not reported recovery by day t , then:

$$M_{i,l,t} = \sum_{s=1}^k \left((1 - (1 - p')^{d_s}) \sum_{r=1}^{10} M_{s,l-1,t-r} \right).$$

For all of v_i 's contacts, we take the sum of all contact chains of length $l - 1$ and multiply by the estimated transmission probability $1 - (1 - p')^{d_s}$ to obtain the sum of all contact chains of length l .

A person x is a degree k contact on day t if the sum of all weights of their contacts with degree $\leq k$ is above 10%:

$$\sum_{i=1}^k \sum_{j=1}^{10} M_{x,i,t-j} \geq C.$$

The contact cutoff value, C , only affects the quarantine rules and not simulating the disease transmission. Essentially, direct contacts that are calculated to be infected with probability at least C are quarantined. The value of C is set to 10% by default. Note that contacts are recalculated and updated every day in the simulation. By default, first degree contacts will be quarantined.

4.1. Pre-exposure notification system

The pre-exposure notification system, as implemented in NOVID [16], acts like a "social radar" telling users how close they are to COVID-19 by showing the number of positive cases at each distance in their social network. App users can see the number of neighbors of distance d that are COVID-19 positive. We assume that first, second, and third degree neighbors will take precautionary measures. For second degree contacts, we assume a 75% reduction in contacts. For third degree contacts, we assume a 50% reduction in contacts. Contact reduction is simulated by taking the set of all interactions of a individual, and deleting each contact with a certain probability (75% or 50% depending on whether the individual is a second or third degree contact).

5. Contact tracing strategies and simulation

At the beginning of the simulation, exactly one individual is exposed while the rest are susceptible. The simulation runs for 120 days. There are 180 individuals and for each scenario we run 1800 trials where each individual starts as the seed infection 10 times. At the beginning of each trial, some individuals are selected as asymptomatic, which means they will not show symptoms if infected. Note that the status of an individual as asymptomatic is not related to whether they are infected, but whether they will or will not show symptoms if infected. Thus, asymptomatic is a fixed status of an individual. Every individual has an equal and independent probability of being asymptomatic. These probabilities are shown in the Asymptomatic column in Section 6. The set of app users are determined in a similar way with the probability of being an app user shown in the App Proportion column. As discussed in Section 3, symptomatic individuals that are in the infectious stage have a 19% chance of getting tested everyday. We simulate the cases when an expected 0%, 30%, 50%, 70%, and 95% of the population use the app and when an expected 20%, 40%, 60%, and 80% of the population is asymptomatic.

Each contact tracing strategy is a combination of quarantine and testing rules. We use the definition of degree d contacts in Section 4 to determine which individuals to quarantine or test. For example, first degree contacts are assumed to be at the highest risk of infection while second and third degree contacts are at a lower risk. The first 4 of our scenarios are simulated on the original network of 180 students while the fifth scenario is simulated on the extended network of 5000 students as discussed in Section 2. In all our scenarios, we assume individuals who have tested positive are quarantined for the duration of the infection.

We simulate the following 5 situations:

Table 1
Quarantine direct contacts, no followup testing.

Situation		Infected	Quarantine			Tests used
App proportion	Asymptomatic		False	True	Tested	
0%	40%	106	0	0	292	54
30%	40%	101	79	40	278	51
50%	40%	97	196	94	266	49
70%	40%	87	329	150	243	44
95%	40%	70	477	201	195	35
0%	80%	110	0	0	101	18
30%	80%	108	34	19	99	18
50%	80%	108	88	52	97	18
70%	80%	102	159	92	94	17
95%	80%	100	283	157	91	17

1. Quarantine of first degree contacts:
This is the basic strategy where first degree contacts, people who are at the highest risk, are quarantined. No additional testing strategies are implemented.
2. Quarantine of first degree contacts with followup testing of second and third degree contacts:
As in the first scenario, first degree contacts are quarantined. Additionally, first, second, and third degree contacts are tested every 3 days. The goal is to test individuals who are at moderate risk of infection.
3. Pre-exposure notification system:
As in scenario 2, first degree contacts are quarantined and first, second, and third degree contacts are tested every 3 days. In addition, we implement the pre-exposure notification system as discussed in Section 4. We assume that second and third degree contacts will take precautionary actions and thus reduce contacts by 75% and 50% respectively.
4. Pre-exposure notification system with periodic testing:
In addition to the strategy in scenario 3, we test all individuals every 14 days. This scenario tests whether mass testing is effective in curbing spread.
5. Scenario 4 on the extended graph:
The strategy implemented in scenario 5 is the same as in scenario 4. In this scenario, we simulate using the extended graph of 5000 students as discussed in Section 2. This scenario is used to test whether our results are consistent even for larger populations.

We will measure the contact tracing app and COVID-19 testing configuration by 3 metrics: Total infected, total days spent in quarantine, and total tests used. In the tables in Section 6, the Infected column shows the number of individuals infected after 120 days. Quarantines are split into 3 categories: false, true, tested. False quarantines are quarantined individuals who do not have COVID-19. Individuals in the True quarantine category are infected with COVID-19 but have not received a positive test result. Tested quarantines are individuals placed in quarantine after testing positive. We distinguish the Tested quarantines because these infections have been confirmed by test while True and False quarantines are predicted by the quarantine rules in Section 4. Each of the quarantine columns show the total number of days spent in quarantine across all individuals over 120 days. The Tests Used column shows the total number of tests used in the simulation after 120 days. The values in the tables show the total over 120 days. In the graphs below, we present the time graphs for cumulative infections and active infections. We define active infections as individuals who are infected, not recovered, and have not tested positive.

6. Results

We simulate each scenario 1800 times where each individual is the seed infection 10 times. The data provided in the tables below are averaged over all trials and rounded to the nearest integer. In summary, we simulate the following 5 scenarios:

1. Quarantine of first degree contacts
2. Quarantine of first degree contacts with followup testing of second and third degree contacts
3. Pre-exposure notification system
4. Pre-exposure notification system with periodic testing
5. Scenario 4 on the extended graph

In scenario 1, the most basic strategy of quarantining first degree contacts is not effective at high asymptomatic levels, suggesting that a more comprehensive testing strategy is required. Scenario 2 shows that testing second and third degree contacts greatly increases the effectiveness of the app. The pre-exposure notification system, which warns second and third degree contacts to take extra precautionary measures, can reduce infections by an additional 40% while also reducing the number of quarantines by up to 50%. In scenario 4, it is shown that periodic testing without contact tracing reduces infections by less than 3%. Finally, scenario 5 presents the results for the extended network of 5000 people. We find that the extended network has similar results to the original network. In the graphs in Figs. 5(b) and 5(b), there are large drops in active infections on days when the population is tested. The graphs also display bumps that come from testing the new second and third degree contacts every 3 days. Since the graphs are averaged over thousands of trials, the jaggedness is significant. It is likely caused by many second and third degree contacts being detected at the same time, and thus, the followup testing, which happens every 3 days, becomes synchronized.

6.1. Scenario 1 — Quarantine of first degree contacts (traditional strategy)

This is the most basic strategy where only first degree contacts are quarantined. Every day, symptomatic individuals have a 19% chance of getting a test as discussed in Section 3. The simulation is performed on the original network of 180 students.

From Table 1, we can find out that the effectiveness of the app depends on the percentage of the people use the app. At 40% asymptomatic ratio, the contact tracing app reduces infections by 8.4% at 50% app usage and 17.5% at 70% app usage. The effectiveness of the app increases doubles from adding the extra 20% of app users. This shows that it is very important to have a majority of app users.

A limitation of this strategy is that it is ineffective at high asymptomatic ratios. At 80% asymptomatic ratio the contact tracing app reduces infections by less than 10% in all cases. This is because in this scenario, since COVID-19 testing relies on symptomatic individuals, the effectiveness of the app is greatly reduced at high asymptomatic levels. Less than 20 tests are used while over 100 students are infected which indicates that the majority of infections are undetected. The results in this scenario suggest that more comprehensive testing strategies are needed, especially when large proportions of the population are asymptomatic.

The cost of using the app is that quarantines increase as more people use the app. In all cases, the number of false quarantines is approximately double the number of true quarantines. The time of peak infections is similar for each app usage. In Fig. 2(b), infections rise dramatically, peaking at around 30 days after the initial infection and then falls as herd immunity is reached. Because quarantine rules require a history of contacts to be recalled, the initial 10 days show no difference between the graphs of each app usage level.

6.2. Scenario 2 — Testing of second and third degree contacts

In this scenario, first, second, and third degree contacts are tested every 3 days. All other parameters are the same as in scenario 1. The simulation is performed on the original network of 180 students (see Table 2).

As in scenario 1, the effectiveness of the app increases greatly as the number of app users increases. In Table 2, at 40% asymptomatic

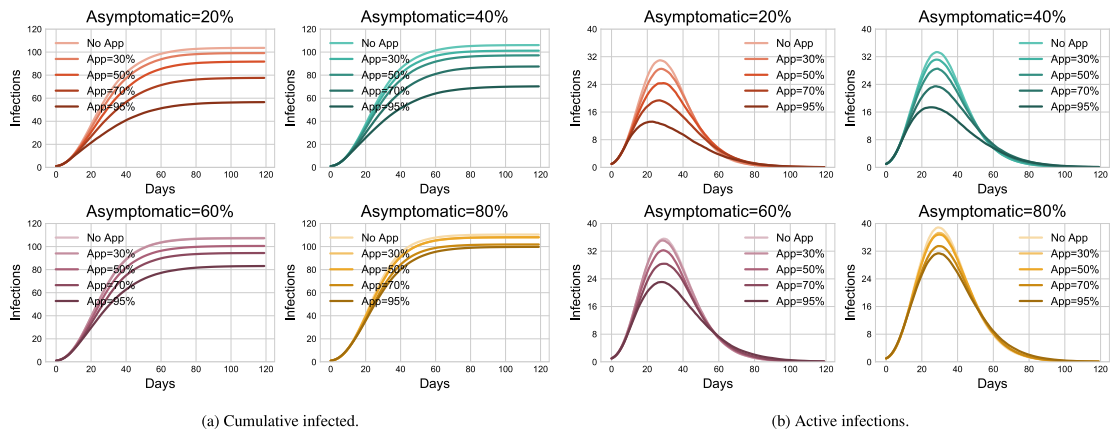


Fig. 2. Quarantine direct contacts, no followup testing.

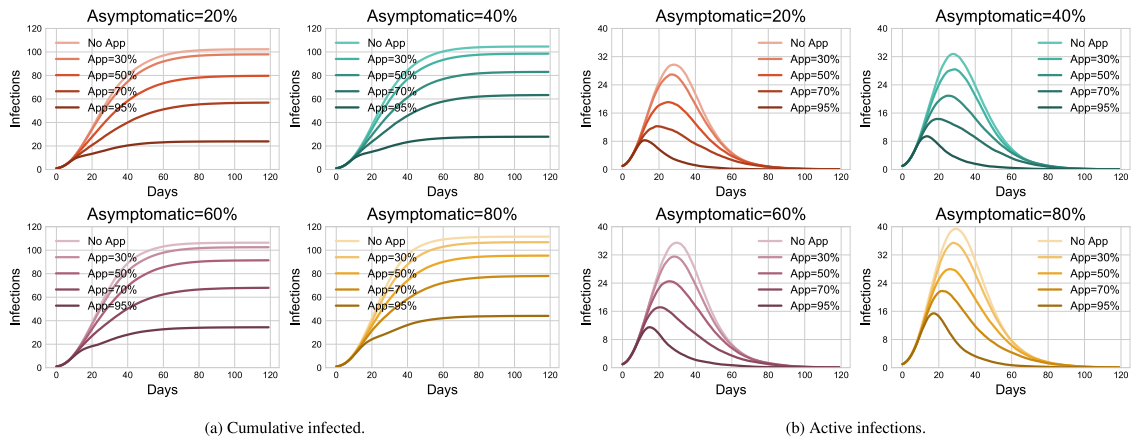


Fig. 3. Test second, third degree contacts.

Table 2
Quarantine direct contacts, test second, third degree contacts.

Situation		Infected	Quarantine		Tests used	
App proportion	Asymptomatic		False	True	Tested	
0%	40%	105	0	0	286	53
30%	40%	99	113	33	326	108
50%	40%	83	259	74	341	234
70%	40%	63	395	101	316	470
95%	40%	28	314	77	172	671
0%	80%	112	0	0	101	19
30%	80%	107	89	24	156	62
50%	80%	95	254	71	250	188
70%	80%	78	419	110	305	401
95%	80%	44	423	106	239	645

ratio, the contact tracing app reduces infections by 20.6% at 50% app usage and 39.5% at 70% app usage. These reductions are over two times greater than in scenario 1. Notably, at 95% app usage, infections are reduced by 73.4%, which is almost twice the reduction at 70% app usage. Compared to scenario 1, the number of true quarantines decreases while the number of tested quarantines increases at 40% asymptomatic ratio. At 80% asymptomatic ratio, the contact tracing app reduces infections by 14.5% at 50% app usage and 30.0% at 70% app usage. Again, this strategy is significantly more effective than scenario 1. However, the number of quarantines increases significantly in all categories compared to scenario 1. In particular, the number of tested quarantines rises by 224% at 70% app usage. This is to be expected since the number of tests used rises significantly, from less than 20 to over 400. Again, the app is significantly more effective at 95%

app usage, with infections being reduced by 60.05%. Although results for 40% and 80% asymptomatic ratio are similar for app proportions less than 70%, at 95% app proportion, the case of 80% asymptomatic yields 58% more infections.

Similar to scenario 1, the graphs in Fig. 3 are similar during the first 10 days and then diverge significantly. However, at high app usage, infections peak before 20 days while no app usage yields peak infections at around 30 days. In conclusion, testing second and third degree contacts greatly increases the effectiveness of the app especially at higher asymptomatic proportions. Thus, the increase in test usage and quarantines is justified.

6.3. Scenario 3 — Pre-exposure notification system

In this scenario, we simulate the pre-exposure notification system. For second degree contacts, we assume a 75% reduction in contacts. For third degree contacts, we assume a 50% reduction in contacts. All other parameters are the same as those in scenario 2. The simulation is performed on the original network of 180 students.

As in the previous scenarios, the effectiveness of the app increases as the number of app users increases. In Table 3, at 40% asymptomatic ratio, the contact tracing app reduces infections by 37.1% at 50% app usage and 66.1% at 70% app usage. At 95% app usage, infections are reduced by 84.5% which is an additional 40.2% reduction compared to scenario 2. Compared to scenario 2, the total number of quarantines is reduced by 51.4% at 70% app usage. There is a 54.7% reduction in the number of tests used which is caused by the reduced interactions for second and third degree contacts. This strategy is more effective even though the number of true and tested quarantines are reduced by 55.2% and 48.4% respectively.

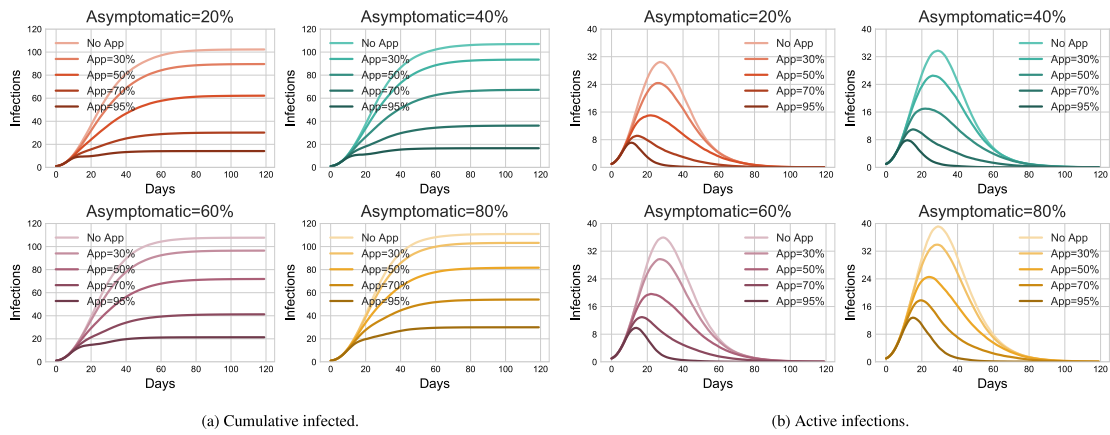


Fig. 4. Quarantine direct contacts, pre-exposure notification system.

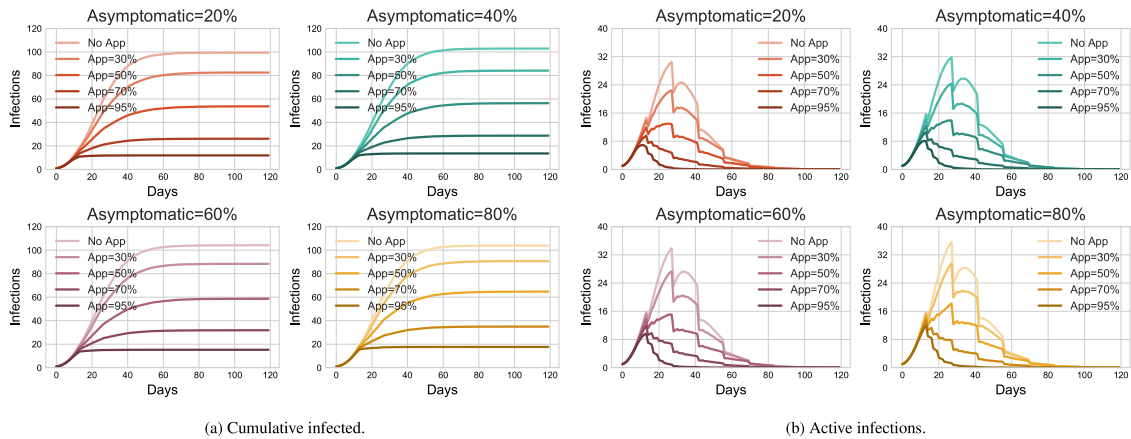


Fig. 5. Pre-exposure notification system with periodic testing.

Table 3
Quarantine direct contacts, pre-exposure notification system.

Situation		Infected	Quarantine		Tests used	
App proportion	Asymptomatic		False	True	Tested	
0%	40%	107	0	0	296	54
30%	40%	93	93	27	303	102
50%	40%	67	162	44	254	161
70%	40%	36	186	45	163	213
95%	40%	17	190	42	98	280
0%	80%	111	0	0	101	19
30%	80%	103	87	22	150	65
50%	80%	82	187	48	190	143
70%	80%	54	251	61	180	221
95%	80%	30	288	66	156	309

At 80% asymptomatic ratio, the contact tracing app reduces infections by 14.5% at 50% app usage and 30.0% at 70% app usage. Again, this strategy is significantly more effective than scenario 1. At 95% app usage, the 80% asymptomatic case yields 80.6% more infections than the 40% case. This is an even more pronounced gap than in scenario 2. The graphs in Fig. 4 are similar in shape to Scenario 2, the main difference being that the number of total infections are lower.

6.4. Scenario 4 — Pre-exposure notification system with periodic testing

In scenario 4, we simulate the Pre-Exposure Notification System with periodic testing every 14 days is investigated. As in scenario 3, second and third degree contacts take extra precautionary measures.

Individuals in the model are tested every 14 days. The simulation is performed on the original network of 180 students.

At 40% asymptomatic ratio, the contact tracing app reduces infections by 45.2% at 50% app usage and 72.1% at 70% app usage (see Table 4). At 95% app usage, infections are reduced by 86.8%. Scenario 3 and 4 show similar results at higher app usage. In this scenario, the app is effective even at 30% app usage. The 30% app usage case reduces infections by 18.1% which is much more effective than in previous scenarios. Additionally, results are similar at all asymptomatic ratios. At 80% asymptomatic ratio, the contact tracing app reduces infections by 37.8% at 50% app usage and 66.4% at 70% app usage. Thus, periodic testing is especially helpful when larger proportions of individuals are asymptomatic.

In Fig. 5(b), the number of infected but not tested individuals drop sharply during days when the population is tested. The graph displays bumps that come from testing the new second and third degree contacts every 3 days. Since the graphs are averaged over thousands of trials, the jaggedness is significant. It is likely caused by many second and third degree contacts being detected at the same time, and thus, the followup testing becomes synchronized. This shows that when used with contact tracing, mass testing can be very helpful in reducing active infections. Surprisingly, periodic testing at 0% app usage reduces infections from 106.05 in scenario 1 to 102.83, only a 3.0% decrease. This is an insignificant change given that the population of 180 students. Thus, periodic testing without contact tracing is not effective.

6.5. Graph extension

In this section, we simulate the strategy in scenario 4 on an extended version of the graph. By the process described in Section 2, the original

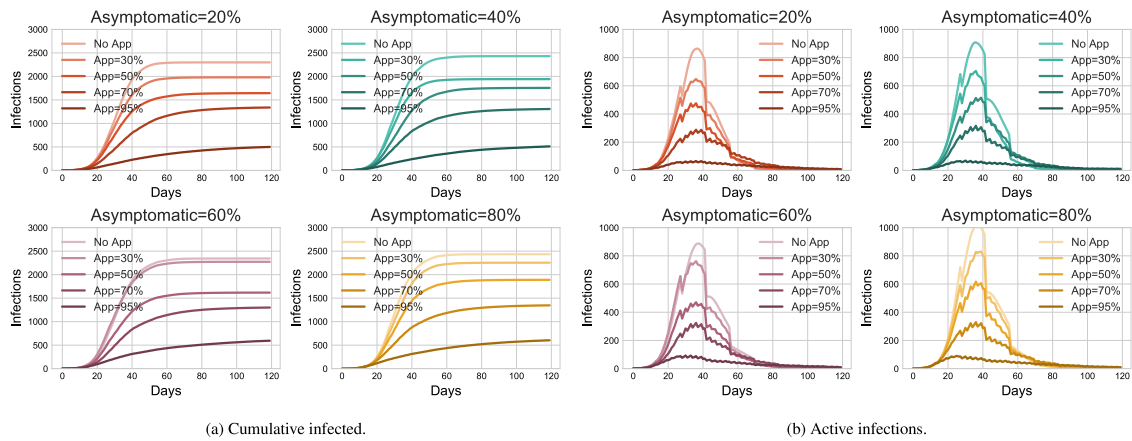


Fig. 6. Graph Extension.

Table 4
Pre-exposure notification system with periodic testing.

Situation		Infected	Quarantine		Tests used	
App proportion	Asymptomatic		False	True	Tested	
0%	40%	103	0	0	466	1021
30%	40%	84	95	27	404	1126
50%	40%	56	153	38	289	1287
70%	40%	29	166	38	159	1441
95%	40%	14	174	35	85	1571
0%	80%	104	0	0	375	1044
30%	80%	91	104	29	358	1118
50%	80%	65	175	44	281	1266
70%	80%	35	195	45	172	1426
95%	80%	18	209	44	106	1562

Table 5
Graph Extension.

Situation		Infected	Quarantine		Tests used	
App proportion	Asymptomatic		False	True	Tested	
0%	40%	2430	0	0	10 992	31 002
30%	40%	1942	5858	869	10 054	36 557
50%	40%	1754	15 411	1686	9847	45 625
70%	40%	1301	26 188	2111	7902	62 003
95%	40%	501	22 810	1439	3319	81 970
0%	80%	2433	0	0	8759	31 576
30%	80%	2253	6641	1017	10 022	36 332
50%	80%	1888	16 594	1852	9563	45 640
70%	80%	1341	26 338	2192	7671	61 131
95%	80%	592	26 708	1723	3861	86 491

weighted temporal graph is extended to 5000 individuals. This population size is closer to larger communities such as schools or universities (see Table 5).

At 40% asymptomatic ratio, the contact tracing app reduces infections by 27.8% at 50% app usage and 46.5% at 70% app usage. This is less effective than scenario 4. At 95% app usage, infections are reduced by 79.4% which is similar to in scenario 4. Since the average degree is preserved on the network of 5000 students, the similar results are not surprising. As in scenario 4, the results are similar for all asymptomatic ratios. At 80% asymptomatic ratio, the contact tracing app reduces infections by 22.4% at 50% app usage and 44.9% at 70% app usage. At 95% app usage, infections are reduced by 75.7%. These numbers are very similar to the corresponding values of 40% asymptomatic ratio.

As in scenario 4, Fig. 6(b) shows noticeable drops in infections on days when the population is tested. Interestingly, at 95% app usage, the number of infected individuals never rises to more than 100 people, or 2% of the population. The graph does not exhibit an obvious peak

but rather a sustained amount of infections which is a reflection of the greater population size.

7. Discussion and conclusion

Given the significant loss of life at risk, finding effective measures to prevent the spread of COVID-19 is a top priority. Challenges in COVID-19 prevention include significant pre-symptomatic transmission, high proportions of asymptomatic cases, and inaccurate tests during the pre-symptomatic stage. This work focuses on solving these challenges by presenting effective strategies in digital contact tracing and testing.

The parameters in this model can be easily changed as more precise values of COVID-19 parameters are measured. As more data on human social patterns are collected, larger social networks can be constructed leading to more accurate predictions in the model. Additionally, as more accurate COVID-19 tests are developed, model results and optimal strategies could change. Finally, this model only considers infections based on close contacts. Although COVID-19 mainly spreads through close contacts [37], the significance of spread through indirect contacts is unclear and could be investigated in future models.

By simulating a variety of tracing and testing strategies, we found that digital contact tracing can be very effective when combined with testing. In scenario 1, the most basic strategy of quarantining first degree contacts is not effective at high asymptomatic levels, suggesting that a more comprehensive testing strategy is required. Testing second and third degree contacts greatly increases the effectiveness of the app. The pre-exposure notification system, which warns second and third degree contacts to take extra precautionary measures, can reduce infections by an additional 40% while also reducing the number of quarantines by up to 50%. While periodic testing with contact tracing is effective, periodic testing without contact tracing reduces infections by less than 3%. We find the results on the extended network of 5000 students to be similar.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

I would like to thank my mentor Dr. Jesse Geneson for his invaluable advice and guidance throughout the project. This would not be possible without him. Thank you to Dr. Tanya Khovanova and Dr. Alexander Vitanov for reviewing the paper draft. I thank the reviewers of Mathematical Biosciences for their contribution, in particular for suggestions to clarify parts of the manuscript. This work was completed under the MIT PRIMES-USA program.

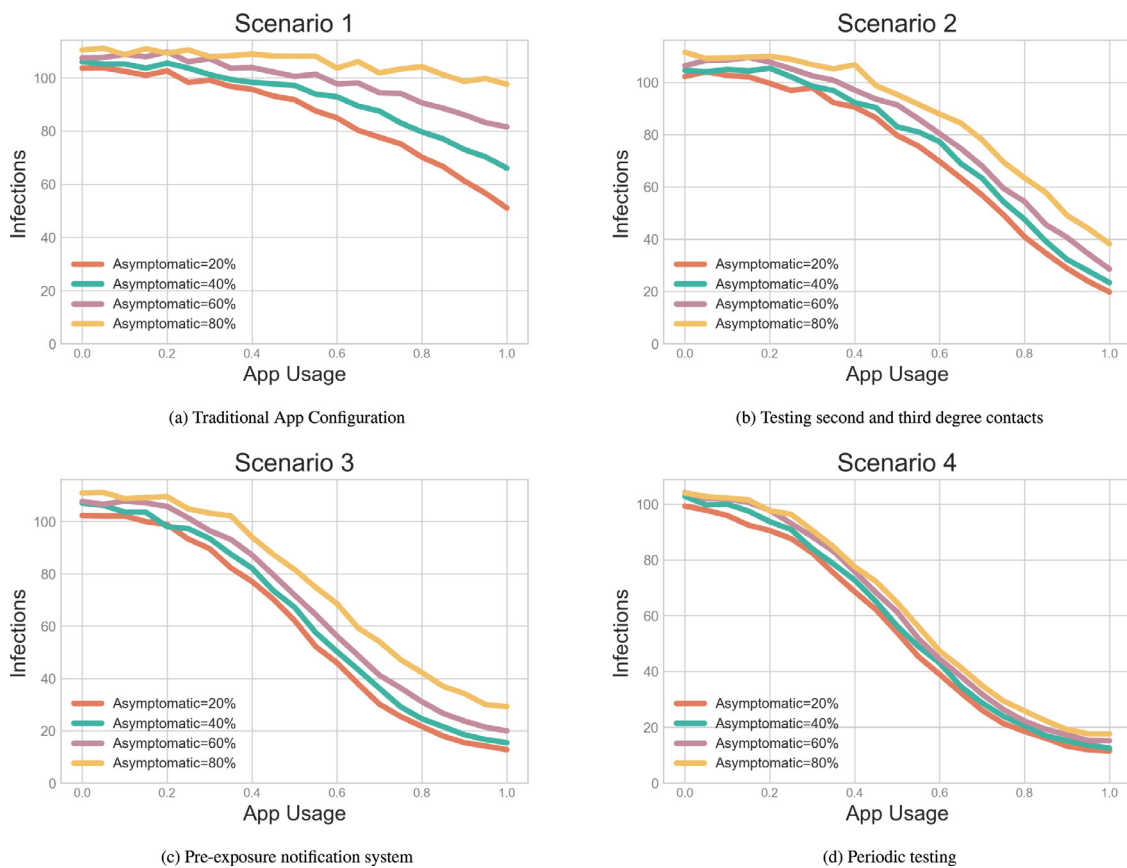


Fig. A.1. App Usage vs Infections.

Appendix. App usage

In the following, we present figures for the number of infections as a function of the app usage. Simulations are run with the proportion of app users ranging from 0% to 100% in increments of 5%. The app becomes much more effective as more users use the app. As shown in Section 6.4, scenario 4 shows very close results for all asymptomatic levels. Scenarios 3 and 4 shows that the effectiveness of the app begins to level off after more than 80% of the population uses the app (see Fig. A.1).

References

[1] E. Dong, H. Du, L. Gardner, An interactive web-based dashboard to track COVID-19 in real time, *Lancet Infect. Dis.* 20 (5) (2020) 533–534.
 [2] X. He, E.H. Lau, P. Wu, X. Deng, J. Wang, X. Hao, Y.C. Lau, J.Y. Wong, Y. Guan, X. Tan, et al., Temporal dynamics in viral shedding and transmissibility of COVID-19, *Nat. Med.* 26 (5) (2020) 672–675.
 [3] CDC, COVID-19 pandemic planning scenarios, 2020, Retrieved from <https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html>. (Accessed 25 December 2020).
 [4] N.G. Davies, P. Klepac, Y. Liu, K. Prem, M. Jit, R.M. Eggo, CMMID COVID-19 working group and others, Age-dependent effects in the transmission and control of COVID-19 epidemics, 2020, *MedRxiv*.
 [5] L.M. Kucirka, S.A. Lauer, O. Laeyendecker, D. Boon, J. Lessler, Variation in false-negative rate of reverse transcriptase polymerase chain reaction–based SARS-CoV-2 tests by time since exposure, *Ann. Internal Med.* (2020).
 [6] M. Andre, K. Ijaz, J.D. Tillinghast, V.E. Krebs, L.A. Diem, B. Metchock, T. Crisp, P.D. McElroy, Transmission network analysis to complement routine tuberculosis contact investigations, *Am J Public Health* 97 (3) (2007) 470–477.
 [7] R.Y. Barrow, F. Ahmed, G.A. Bolan, K.A. Workowski, Recommendations for providing quality sexually transmitted diseases clinical services, 2020, *MMWR Recomm. Rep.* 68 (5) (2020) 1.
 [8] R. Levine, Development of a contact tracing system for Ebola virus disease—Kambia district, Sierra Leone, January–February 2015, *MMWR Morb. Mortal. Wkly. Rep.* 65 (2016).

[9] S. Simmons-Duffin, Coronavirus Cases are Surging. The Contact Tracing Workforce is not, Centers for Disease Control and Prevention, 2020, Retrieved from <https://www.npr.org/sections/health-shots/2020/08/07/899954832/coronavirus-cases-are-surging-the-contact-tracing-workforce-is-not>. (Accessed 25 December 2020).
 [10] R.R. Lash, C.V. Donovan, A.T. Fleischauer, Z.S. Moore, G. Harris, S. Hayes, M. Sullivan, A. Wilburn, J. Ong, D. Wright, et al., COVID-19 contact tracing in two counties—North Carolina, June–July 2020, *Morb. Mortal. Wkly. Rep.* 69 (38) (2020) 1360.
 [11] A. Slavitt, L. Brilliant, B. Frist, B. Kocher, M. Leavitt, M. McClellan, F. Mostashari, V. Murthy, S. Gottlieb, A. Gawande, T. Inglesby, J. Kayyem, M. Osterholm, J. Sharfstein, E. Topol, L. Wen, Bipartisan public health leaders’ letter on COVID-19 tracking and tracing, 2020, Retrieved from <https://www.aei.org/articles/bipartisan-public-health-leaders-letter-on-covid-19-tracking-and-tracing/>. (Accessed 25 December 2020).
 [12] L. Donnelly, Coronavirus cases ‘lost’ in test and trace blunder, 2020, Retrieved from <https://www.telegraph.co.uk/news/2020/10/04/governments-coronavirus-data-chaos-computer-error-blamed-spike/>. (Accessed 25 December 2020).
 [13] K. Farrahi, R. Emonet, M. Cebrian, Epidemic contact tracing via communication traces, *PLoS One* 9 (5) (2014) e95133.
 [14] C. Farronato, M. Iansiti, M. Bartosiak, S. Denicolai, L. Ferretti, R. Fontana, How to get people to actually use contact-tracing apps, 2020, Retrieved from <https://hbr.org/2020/07/how-to-get-people-to-actually-use-contact-tracing-apps>. (Accessed 25 December 2020).
 [15] The New York Times, Tracking the coronavirus at U.S. colleges and universities, 2020, Retrieved from <https://www.nytimes.com/interactive/2020/us/covid-college-cases-tracker.html>. (Accessed 25 December 2020).
 [16] P.-S. Loh, Flipping the perspective in contact tracing, 2020, arXiv preprint [arXiv:2010.03806](https://arxiv.org/abs/2010.03806).
 [17] W.O. Kermack, A.G. McKendrick, A contribution to the mathematical theory of epidemics, *Proc. R. Soc. A* 115 (772) (1927) 700–721.
 [18] A. Abou-Ismaïl, Compartmental models of the COVID-19 pandemic for physicians and physician-scientists, *SN Compr. Clin. Med.* (2020) 1.
 [19] D. Kai, G.-P. Goldstein, A. Morgunov, V. Nangalia, A. Rotkirch, Universal masking is urgent in the COVID-19 pandemic: SEIR and agent based models, empirical validation, policy recommendations, 2020, arXiv preprint [arXiv:2004.13553](https://arxiv.org/abs/2004.13553).
 [20] S.E. Eikenberry, M. Mancuso, E. Iboi, T. Phan, K. Eikenberry, Y. Kuang, E. Kostelich, A.B. Gumel, To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic, *Infect. Dis. Modell.* (2020).

- [21] R. Singh, R. Adhikari, Age-structured impact of social distancing on the COVID-19 epidemic in India, 2020, arXiv preprint [arXiv:2003.12055](https://arxiv.org/abs/2003.12055).
- [22] G. Giordano, F. Blanchini, R. Bruno, P. Colaneri, A. Di Filippo, A. Di Matteo, M. Colaneri, Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy, *Nat. Med.* (2020) 1–6.
- [23] A.D. Paltiel, A. Zheng, R.P. Walensky, Assessment of SARS-CoV-2 screening strategies to permit the safe reopening of college campuses in the United States, *JAMA Netw. Open* 3 (7) (2020) e2016818.
- [24] F. Piguillem, L. Shi, Optimal COVID-19 quarantine and testing policies, CEPR Discussion Paper No. DP14613, 2020.
- [25] R. Hinch, W. Probert, A. Nurtay, M. Kendall, C. Wymant, M. Hall, C. Fraser, Effective configurations of a digital contact tracing app: A report to NHSX, 2020, (Unpublished Results) url: https://github.com/BDI-pathogens/covid-19_instant_tracing/blob/master/Report.
- [26] J. Almagor, S. Picascia, Can the app contain the spread? An agent-based model of COVID-19 and the effectiveness of smartphone-based contact tracing, 2020, arXiv preprint [arXiv:2008.07336](https://arxiv.org/abs/2008.07336).
- [27] J. Fournet, A. Barrat, Contact patterns among high school students, *PLoS One* 9 (9) (2014) e107878.
- [28] S.A. Lauer, K.H. Grantz, Q. Bi, F.K. Jones, Q. Zheng, H.R. Meredith, A.S. Azman, N.G. Reich, J. Lessler, The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application, *Ann. Internal Med.* 172 (9) (2020) 577–582.
- [29] CDC, Case investigation and contact tracing guide, 2020, Retrieved from <https://www.cdc.gov/coronavirus/2019-ncov/php/contact-tracing/contact-tracing-plan/appendix.html>. (Accessed 25 December 2020).
- [30] C. Buciluă, R. Caruana, A. Niculescu-Mizil, Model compression, in: Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2006, pp. 535–541.
- [31] CDC, Duration of isolation and precaution for adults, 2020, Retrieved from <https://www.cdc.gov/coronavirus/2019-ncov/hcp/duration-isolation.html>. (Accessed 25 December 2020).
- [32] Y. Liu, A.A. Gayle, A. Wilder-Smith, J. Rocklöv, The reproductive number of COVID-19 is higher compared to SARS coronavirus, *J. Travel Med.* (2020).
- [33] G. Viceconte, N. Petrosillo, COVID-19 R0: Magic number or conundrum? *Infect. Dis. Rep.* 12 (1) (2020).
- [34] P. Poletti, M. Tirani, D. Cereda, F. Trentini, G. Guzzetta, G. Sabatino, V. Marziano, A. Castrofino, F. Grosso, G. Del Castillo, et al., Probability of symptoms and critical disease after SARS-CoV-2 infection, 2020, arXiv preprint [arXiv:2006.08471](https://arxiv.org/abs/2006.08471).
- [35] S. Lee, T. Kim, E. Lee, C. Lee, H. Kim, H. Rhee, S.Y. Park, H.-J. Son, S. Yu, J.W. Park, et al., Clinical course and molecular viral shedding among asymptomatic and symptomatic patients with SARS-CoV-2 infection in a community treatment center in the Republic of Korea, *JAMA Internal Med.* 180 (11) (2020) 1447–1452.
- [36] University of Michigan Health Service, COVID-19 testing through university health service, 2020, Retrieved from <https://uhs.umich.edu/covid-testing>. (Accessed 25 December 2020).
- [37] CDC, COVID-19 pandemic planning scenarios, 2020, Retrieved from <https://www.cdc.gov/coronavirus/2019-ncov/faq.html>. (Accessed 25 December 2020).