A Review of COVID-19 Mathematical Models and an Implementation of Vaccination Policies

Samir Gouin[1], Amy Li[1], Gloria Hanxi Ma[2], Dimitri Yang[1]

Abstract

Cellular automata (CA) models have been used to simulate the behaviour of infectious diseases and can offer valuable information concerning the spread of infection and population susceptibility to sickness. Furthermore, CA models can be used to evaluate various response strategies, such as herd immunization and vaccination. By assessing CA approaches for COVID-19, we identified a lack of an age-based symptom severity score and death probability (1-3). We used a Susceptible Infected Removed (SIR) model and distribution of population age groups from the Canadian government to determine the effect of vaccinating older (60+ years old) versus a younger population (30-59 years old). Our findings show a significant decrease in the total number of deaths and peak number of infections when the older population was vaccinated. This is a result of the higher probabilities of death and severe symptoms in older age groups. While this simulation is based on a small scale, the findings provide evidence to prioritize vaccination of the elderly.

Keywords: COVID-19, pandemic, mathematical modeling, epidemiology, cellular automata

COVID-19 alongside other modern epidemics has demonstrated a need for mathematical models to better reflect the realistic characteristics of disease spread. These models offer methods to test various preventative and control strategies for disease spread, as well as project outcomes under specific circumstances.

Background

The foundation of modern mathematical epidemiology was established by Ross' 1902 paper regarding the discrete-time dynamics of malaria transmission from mosquitoes to humans (4). Following Ross' research, Kermack and McKendrick published three papers in 1927, 1932 and 1933, which developed the SIR model. In this model, the population was divided into susceptible (S), infected (I), and recovered (R) groups whereas each group is represented by a differential equation. Since 1933, many models have continued to use the Kermack-McKendrick SIR model to investigate the dynamics of epidemic transmission, including the SARS outbreak in 2003 (5), the Ebola outbreak of 2014-2016 (6), and most recently, the COVID-19 pandemic.

Building on these seminal studies, cellular automata (CA) models were developed by Stanislaw Ulam, who was investigating the growth of crystals, and John Von Neumann, who was studying self-replicating artificial systems (7). Since then, cellular automata have shown a remarkable ability to model physical and biological phenomena in a spatial context. CAs consist of a grid of individual cells. Each cell is surrounded by other cells, referred to as neighbors, and can represent an individual or group of individuals of the population at a discrete spatial point. As used in Conway's Game of Life CA model, the Moore neighborhood is commonly used to define a set of cells (Figure 1). It consists of eight neighboring cells (blue squares). Cells change their states according to a set of transition rules and the states of their neighbors. Transition rules can be deterministic, if neighbor x is infected, then cell y will be infected, or probabilistic, if neighbor x is infected, then there is a probability that cell y will be infected (similar to the equation for the rate of infection in the SIR model). All cells within a CA are governed by the same set of transition rules, but by implementing an initial distribution of cells with different states and transition rules based on probability, CAs can be used to model realistic systems (8). Ultimately, probabilistic CAs can simulate how local interactions between neighboring cells contribute to the spread of a disease, while taking into account spatial, temporal and probabilistic characteristics.

COVID-19

COVID-19 can be described by its timeline of infection and age-dependent severity. The disease progresses in three main steps: a latent period, an infectious period and a recovery phase. Post infection, the individual is in a latent period. This period varies from 4-6 days (9)[1]. The individual

---

1 Please note that these estimates are continually changing as new research is released
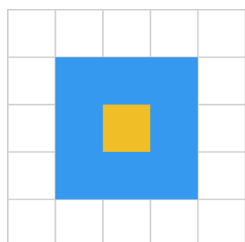
Figure 1. Moore Neighbourhood

Figure 2. Age distribution of the hospitalized and deceased from COVID-19 in Canada

may or may not start to show symptoms after the latent period as they enter the infectious period. Notably, recent research has shown that people infected with COVID-19 can be contagious 2 days prior to symptom onset (10). Regardless of symptom severity, individuals can transmit the disease to others. While symptoms may last after the infectious period, the individual is no longer considered capable of transmitting COVID-19.

Findings regarding reinfection of COVID-19 remain contentious. Viral infections are typically followed by pro-inflammatory and anti-inflammatory mechanisms that offer long-term immunity. However, a recent study among various reports have found cases of reinfection. While results reported by Tillett et al. provide preliminary evidence of a possibility of reinfection, the sample sizes are too low to generalize the results and further investigation is required (11).

During the infection period, individuals of different ages have different survival rates as well as severity of symptoms. Data2 from March to December 2020 show a greater number of hospitalized and deceased cases among older age groups (Fig. 2). Mortality rates increase drastically in older populations, with the highest mortality rates observed in patients aged 80 years and older (12). It is unclear whether and by how much infection rates change across age; however, some studies indicate that children have a lower susceptibility to infection (13).
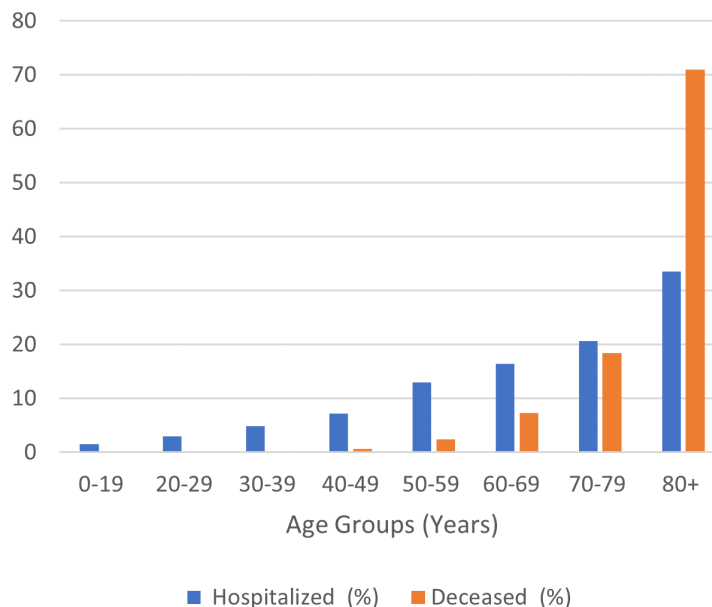
COVID-19 & CAs

Various CA approaches have been used to model the spread of COVID-19 in a population including Schimit (1), Mondal et al., (2) and Ghosh and Bhattacharya (3)3.

Mondal et al. created a model of COVID-19 to examine two strategies to control COVID-19; the use of lockdown, social distancing and quarantine or transmission for a large proportion of the population to develop immunity. The latter is the concept of herd immunity (HI) and when achieved through infection, it is associated with a large number of deaths in vulnerable populations. Notably, the more readily transmitted a disease is, the higher the proportion of the population that needs to be immunized to achieve HI.

Mondal et al. employed a probabilistic SIR CA model with the Moore definition of a neighborhood to determine the costs of HI without a vaccine and the effect of the rate of HI attainment on death rates. The eight differential equations (representing infectious populations of various severity, susceptible individuals and the recovered) have been fitted using data from the government of India to model the spread of infection and death rates in the entire country. They divided the initial population into two main groups, the vulnerable (60+ years old or those with preexisting medical conditions) and the resilient (the remainder of the population) and used recovery rates respective to these groups.
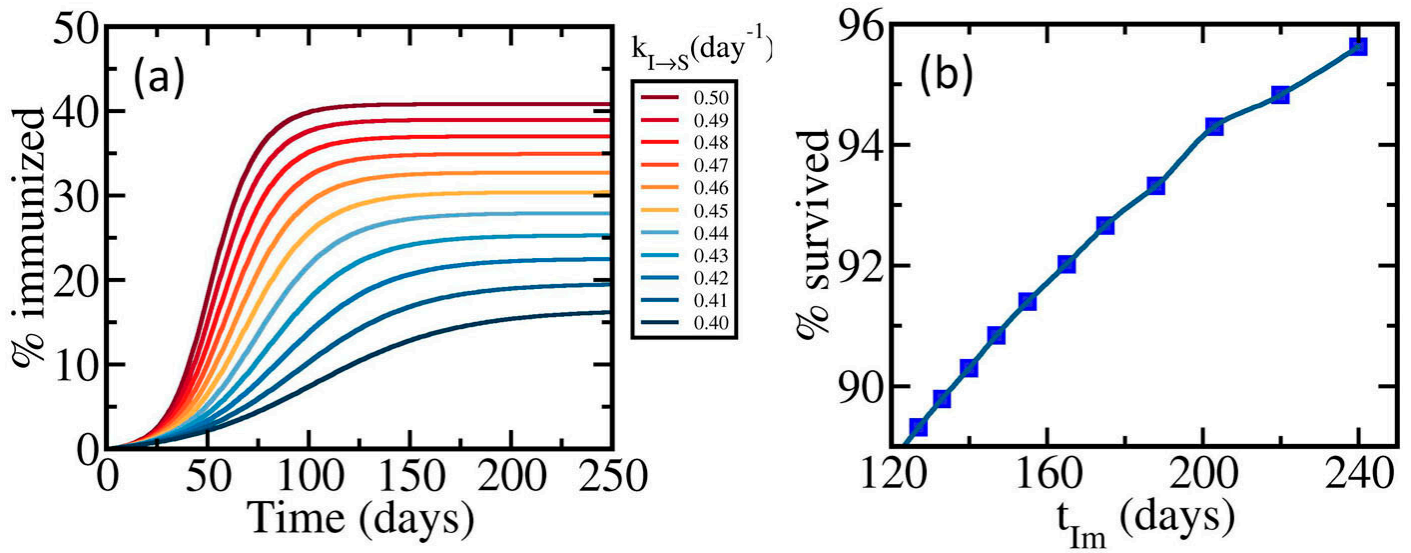
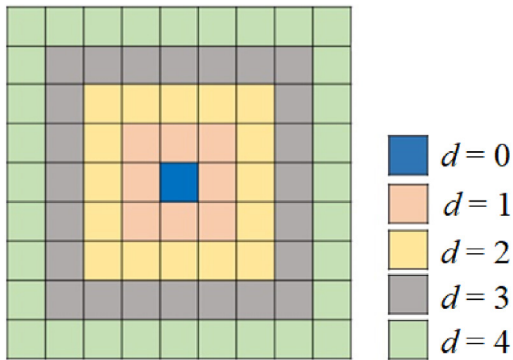Figure 3. The effect of different rates of attaining HI on survival % (2)

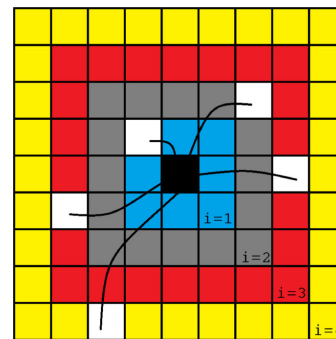Figure 4. Spheres of Movement: d = 0 staying home, d = 1 movement extends to the nearest neighbors ...(17)

Figure 5. An expanded Moore neighborhood with five connections (white squares) and four layers (I=4)

As hypothesized, the simulation revealed that HI should be attained gradually, otherwise, the mortality rate, particularly for the vulnerable population, increases dramatically. For example, a rate of infection of 0.5 (the proportion of susceptible cases becoming infected) led to a mortality rate of 4.7% for the resilient population and 34.3% for the vulnerable population over 250 days. In contrast, when the infection rate was raised to 0.78, deaths increased to 7.9% for resilient population and 57.1% for the vulnerable population. These findings highlight that the longer it takes to reach immunity saturation (tm) as a result of a lower infection rate (transition of a susceptible (S) individual to an infected (I) one: $kS \rightarrow I$), the greater the survival % (Figure 3). In addition, the simulation showed a linear relationship between the percentage of recovered individuals and the HI percentage of the population. This relationship was positive for resilient groups and negative for the vulnerable population. Thus,

these results stress the importance of a low infection rate to protect vulnerable populations when adopting a HI approach. Notably, these results are specific to the simulated population and cannot be generalized.

Mondal et al. have explored the dynamics of COVID-19 progression by employing an HI approach in vulnerable and resilient populations. Their simulation does not take into account realistic distributions of vulnerable populations; rather, they used a fixed ratio (1 vulnerable : 4 resilient) to simulate the entire population. In addition, population groups can be further distinguished by movement and social activities. Although harder to quantify, by using a distribution of mobility and interaction, the simulation could better approach realistic conditions. For example, an asymptomatic individual may be prone to more movement as compared with a severely ill individual. While Mondal et al. have considered

uniform infection rates as part of their HI approach, in reality, adherence to safety policy, a main determinant of infection rate, differs across the population.

Ghosh and Bhattacharya have simulated COVID-19 spread across numerous countries by implementing the SEIQR framework in a synchronous probabilistic CA model. The SEIQR model is based on the SIR model, with the additions of an E group consisting of exposed and asymptomatically infected individuals, and a Q group consisting of quarantined or hospitalized individuals. They initialized the populations with varying population densities. In contrast to Mondal's simulation, Ghosh and Bhattacharya considered individual movement and implemented a sphere of influence that sets boundaries in which an individual can come into contact with others (Fig. 4). Through this approach, they explored the spread of COVID-19 by varying population density, healthcare efficiency and mitigation strategies.

Their results show that movement restriction and the early implementation of a lockdown are critically important. For example, when the sphere of movement was increased to the distance of three cells, almost everyone came into contact with COVID-19. Likewise, the initial number of infected and exposed individuals had a significant impact on the proportion of total population infection. Exposure probability also played an important role in transmission dynamics. A reduction of this probability from 0.8 to 0.2 linearly lowered the spread of infection. Among the effects of other parameters, such as testing time and recovery probability, these findings emphasize the importance of large-scale initiatives to reduce spread, especially in countries with high population density.

In comparison to Mondal's simulation, Ghosh and Bhattacharya better accounted for individual movement but neglected age-dependent symptom severity and case fatality rate.

Moreover, individuals in their model were only in contact with others in their sphere of influence, which limited the fluctuation of movement as well as the impact of super spreaders (highly mobile infected individuals). Despite severe lockdown conditions, the sphere of influence of an individual is often dispersed and not concentrated. Even when governmental regulations reduced movement to isolate social circles, essential activities still exposed individuals to others outside their sphere.

Similar to Ghosh's and Bhattacharya's approach, Schimit used a SEIR probabilistic CA to analyze the effect of healthcare responses and social isolation on the number of deaths due to COVID-19 in Brazil4. Schimit expanded the Moore neighborhood by allowing individuals to make connections outside of the radius of eight neighboring cells (Fig. 5). The probability of a connection decreased as the distance between cells increased. This created a clustering effect

to mimic social bubbles. Schimit decreased the number of connections allowed per individual to model the effect of social isolation.

Schimit split the infected population into five subgroups, each representing a level of symptom severity ranging from asymptomatic (IA) to individuals in hospital beds or in an intensive care unit (ICU) (IS2 and IS3). The number of connections with different types of infective individuals determined the probability of infection of an individual. Asymptomatic individuals (IA) were considered twice as infectious as symptomatic individuals (IS) and those with severe symptoms (IS2 and IS3) were considered half as infectious as symptomatic individuals due to reduced movement. Notably, infected individuals in the hospital still retained a degree of movement to account for contact with health care personnel and their visitors.

By restricting the availabilities of ICUs and the transfer of the infected to hospitals, Schimit showed that about 2 million deaths would result from the ICU limit and 11.5 million deaths would be due to the lack of hospital beds. While this is much higher than the current number of deaths in Brazil5 (~175 000), these statistics reflect the total number of deaths over approximately 8 years without consideration of a vaccine or control policies. In a subsequent analysis, Schimit highlighted the effect of social isolation, both in the reduction of infection and alleviation of stress on the healthcare system. For example, increasing social contact reduction from 40% to 50% would decrease the rate of new ICUs needed from 200 per day to 150 to accommodate all the patients.

In contrast to Mondal et al.'s and Ghosh's and Bhattacharya's approaches, each cell in Schimit's model was filled by an individual. The other models used empty cells to simulate the space between individuals. The latter approach is closer to the heterogenous distribution of a population. Like the other models, Schimit used a probability of infection uninfluenced by age and pre-existing vulnerabilities, both important determinants of COVID-19 infection severity.

Model Description

Out of the gaps identified in Mondal et al.'s, Ghosh's and Bhattacharya's and Schimit's models, we have focused our attention on age; specifically, how age influences COVID-19 spread and the impact of vaccination of specific age groups. To explore these factors, we modified a publicly available probabilistic CA SIR model6 by Max Brenner in which the population is divided into four possible states:

1. Susceptible (S) individuals who do not have the disease but are capable of contracting it

2. Infected (I) individuals who are capable of spreading the disease

3. Recovered (R) individuals who have recovered and are immune to the disease

4. Deceased (D) individuals are removed from the simulation

These states were selected based on the progression of the COVID-19 infection and the previously mentioned COVID-19 CA approaches. (I) individuals are able to infect their neighbors independent of their symptoms. This includes individuals in the infectious period. Symptoms can worsen and lead to death or the individual can recover. (R) individuals can no longer contract the virus and are removed from the spread of infection but are still part of the population total.

In comparison to the previously mentioned models, Brenner's model allows cells to move. Instead of stationary cells with different connections, the cells could be displaced and come in contact with new neighbors. A cell's likelihood of movement was lowered if they were infected and decreased further if symptoms worsened.

Through the following rules, Brenner simulated the dynamics of the spread of COVID-19 and by extending the third rule, we were able to implement an age distribution.

1. Exposure: An (S) individual comes into contact with individuals as a function of movement probability and spatial proximity (interaction with more than the eight initial neighbors of the Moore neighborhood). The level of movement differs across individuals and is decreased for those practicing social distancing or the severely sick (hospitalized).

2. (S) -> (I): COVID-19 is contracted based on exposure at p_i (the probability of infection).

$$p = p_{base} - (m * p_{mask})$$

If a mask (m) is worn, m=1, otherwise, m=0. The efficacy of the mask (p_{mask}) is subtracted from the base probability of infection (p_{base}) to obtain the individual probability of infection (p).

$$p_i = 1 - (1 - p)^R$$

(p) is subtracted from 1 to calculate the probability of remaining uninfected and then raised to the power of the number of infectious neighbors (R). This result is subtracted from 1 and is the pi.

3. (I) -> (R): individuals can progress from mild to severe symptoms at an age-dependent probability. As individuals progress to severe symptoms, their movement decreases and can die based on age-dependent case fatality rates. To implement this, we developed profiles per age group by the probability of severe symptoms and death, based on rates of hospitalization and death rates respectively.

The simulation lattice represented 1 square km populated by 890 people, proportional to the urban density of Montreal. Data7 from the most recent census profile of Montreal was used to weigh age groups, each representing a range of 10 years, except the upper limit of 80+ (Figure 6). The model had 50 time steps, each one representing a day.

Brenner implemented two diagnostic tests to determine the success of safety policies, mask wearing and social distancing. The Secondary Attack Rate (SAR) is a ratio of the infected population to the total number of susceptible individuals. The basic reproductive number (R0) is the number of secondary cases for one infectious case in a population with only susceptible individuals. These measures, in addition to the total number of deaths and peak infection number, were used to compare various vaccination policies and the effect of age-dependent severity and death probability. Specifically, the means of deceased cases, peak number of infections, SAR and R0 were calculated for the original model (D), the model with age distributions (D_{age}), vaccinated 30-59 year olds (V_{mid}) and vaccinated 60+ year olds (V_{old}).

To simulate the effect of vaccination, death and severity probabilities were set to 0 in the middle age group, 30-59, and older group, 60+. We used a medium level of safety policy across all trials, as such, the probability of social distancing and the probability of wearing a mask were set to 25%.

Results

Within each tested group (D, D_{age}, V_{mid} and V_{old}), there was a downward trend with highest values of deceased cases, peak number of infections, SAR, and R0 reported in the D group and the lowest values in the Vold group (Figure 7).

In regards to the total number of deaths, a significant difference between groups was found (H(3) = 69.444, p < 0.001). Post hoc analysis revealed significant differences between V_{old} (M = 0, SD = 0) and D_{age} (M = 3.30, SD = 1.03) (p < 0.001) as well as between V_{old} and D (M = 15.65, SD = 2.87) (p < 0.001). There were also significant differences between V_{mid} (M = 2.35, SD = 1.42) and D (p < 0.001), between V_{old} and V_{mid} (p < 0.001), and between D_{age} and D (p < 0.001).

For peak number of infections, a significant difference between groups was found (H(3) = 25.575, p < 0.001). Post hoc analysis indicated a significant difference between V_{old} (M = 112.15, SD = 12.09) and D_{age} (M = 129.00, SD = 19.45) (p = 0.006), between V_{old} and D (M = 139.30, SD = 14.42) (p < 0.001), and between V_{mid} (M = 123.30, SD = 13.52) and D (p = 0.009).

The mean SAR values range from 0.214-0.216. For SAR, a significant difference between groups was found (p = 0.013). A significant difference was found between V_{old} (M = 0.214, SD = 0.037) and D (M = 0.261, SD = 0.040) (p = 0.003). V_{mid}
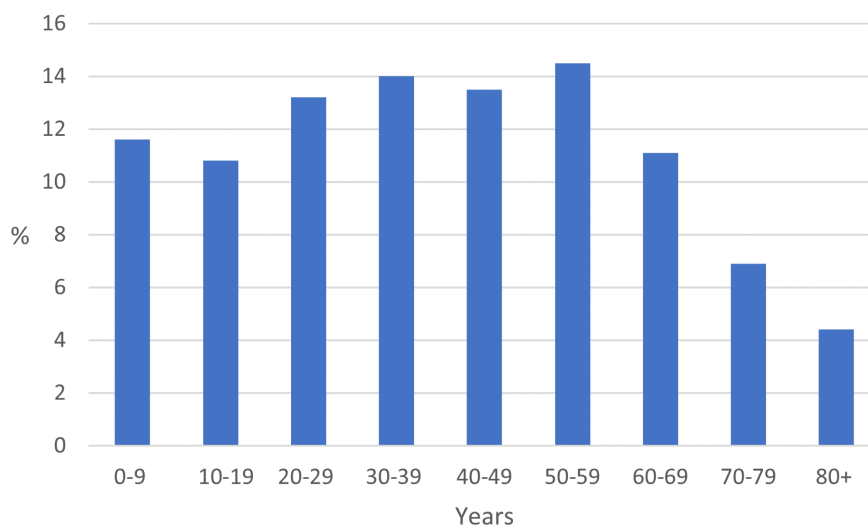
Figure 6. Distribution of population age groups in Montreal, 2016

(M = 0.238, SD = 0.039) and D_{age} (M = 0.240, SD = 0.056) had insignificant group differences.

The group difference for R0 was insignificant (H(3)= 4.871, p = 0.182) - V_{old} (M = 1.815, SD = 0.148), V_{mid} (M = 1.853, SD = 0.092), D_{age} (M = 1.853, SD = 0.095) and D (M = 1.875 , SD = 0.115).

Discussion

Through leveraging Brenner's approach, we were able to implement an age distribution based on Montreal's urban population and test two vaccination policies. Our data showed a consistent trend - the vaccination of the V_{old} group caused the largest decrease in total number of deaths, peak number of infections, SAR, and R_0 relative to the other tested groups. Notably, V_{mid} represented almost double the population size compared to V_{old}, respectively, 42% and 22%. This effect on a smaller population supports prioritization of the vaccination of older populations. A caveat of this finding is that movement depended on severity and policy adherence rather than age.

The Vmid group compared to the Dage group showed little difference across the measured deceased cases, peak number of infections, SAR, and R0. We attributed this result to death rates highly concentrated in the older population. In this instance, it appeared that altering the model had a larger effect than vaccinating the middle age group. Prior to implementing the age distribution, the same death rates and probability of infection were used throughout, inflating the total number of deaths, peak number of infections, SAR, and R0. Subsequently, it is expected that altering the model and reducing death rates and probability of infection overall has a

larger effect than decreasing these values to 0 for solely the Vmid group.

The mean SAR values (0.214-0.216) fall within the commonly reported range of 0.147-0.27 for household contacts; non-household data are currently not available (13). Implementing vaccination was expected to decrease SAR values, which are used as a measure of vaccine efficacy (14). There are small decreases of SAR post-vaccination, with the greatest decrease following vaccination of the 60+ age group; however, the size of the change in SAR post-vaccination did not reflect the magnitude of the changes in peak number of infections and total COVID-19 related deaths. The lack of a significant variation could be a result of a smaller sample size, as SAR and R0 are both ratios and are less prone to individual fluctuations.

Since our simulation is on a small scale, a larger lattice would be needed to model hubs of activity and an uneven distribution of the population in order to generalize our results. This includes the clustering of elderly individuals in long-term facilities where movement probability would be low. In doing so, governments could test disease control methods on a larger scale. Additionally, an observation of 0 deaths following vaccination of the 60+ age group is unrealistic when generalized. However, given the scale of our simulation, this is an expected result.

While our findings are not necessarily surprising, the factors we have identified should be integrated in infectious disease models. The parameters of this approach could be adjusted to simulate other infectious diseases as age-dependent characteristics are common but have been largely unaccounted for in previous models. Although we focused
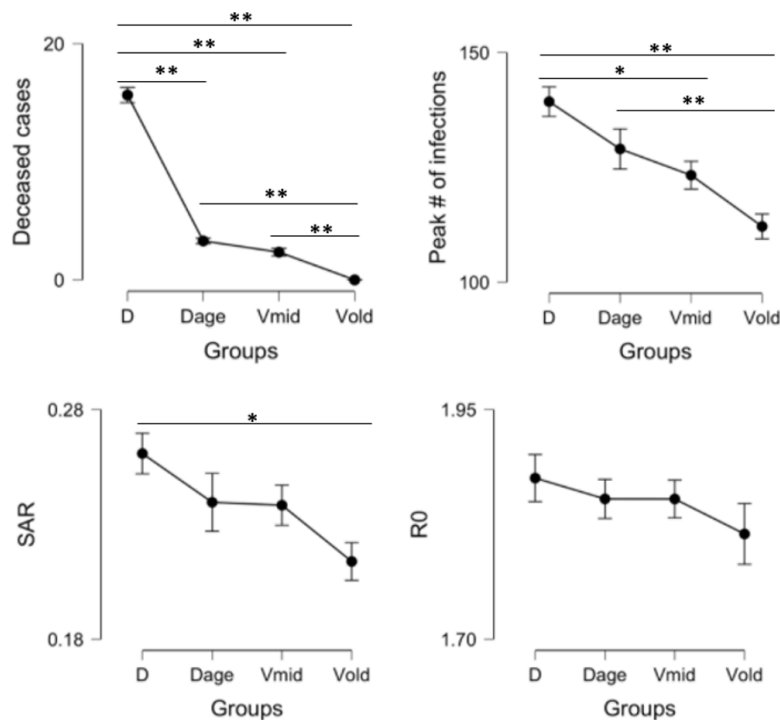
Figure 7. Graphical trends of deceased cases, peak # of infections, SAR, and R0 per tested group (n=20) are shown. The error bars shown represent the standard deviation. * P < 0.05 (Dunn test)

on vaccinations of different age groups, adjustments to the model can be made to test different regulations, such as quarantine or over-exposure. This model can be extended to test regulations on groups differentiated by factors other than age, such as underlying medical conditions or occupation. Through tweaking the model, untested methods for disease control can be tested for safety and efficacy before implementation in real life.

The significant decreases of total deaths, peak infection number, and SAR when vaccinating the older population demonstrates the importance of prioritizing vaccination of the elderly for the COVID-19 vaccine. Through prioritization, we can drastically reduce the spread of such diseases.

References

1. Schimit PHT. A model based on cellular automata to estimate the social isolation impact on COVID-19 spreading in Brazil. Computer Methods and Programs in Biomedicine. 2020:105832.

2. Mondal S, Mukherjee S, Bagchi B. Mathematical modeling and cellular automata simulation of infectious disease dynamics: Applications to the understanding of herd immunity. The Journal of Chemical Physics. 2020;153(11):114119.

3. Ghosh S, Bhattacharya S. Computational model on COVID-19 Pandemic using Probabilistic Cellular Automata. arXiv preprint. 2020. https://arxiv.org/pdf/2006.11270.pdf.

4. Brauer F. Mathematical epidemiology: Past, present, and future. Infectious Disease Modelling. 2017;2(2):113-27.

5. Mkhatshwa T, Mummert A. Modeling Super-spreading Events for Infectious Diseases: Case Study SARS. IAENG International Journal of Applied Mathematics. 2010;41.

6. Okyere E, De-Graft Ankamah J, Hunkpe AK, Mensah D. Deterministic Epidemic Models for Ebola Infection with Time-Dependent Controls. Discrete Dynamics in Nature and Society. 2020;2020:2823816.

7. White SH, Rey AM, Sanchez GR. Modeling epidemics using cellular automata. Applied Mathematics and Computation. 2007;186(1):193-202.

8. Agapie A, Andreica A, Giuclea M. Probabilistic cellular automata. J Comput Biol. 2014;21(9):699-708.

9. Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, et al. The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. Annals of internal medicine. 2020;172(9):577-82.

10. He X, Lau EHY, Wu P, Deng X, Wang J, Hao X, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nature Medicine. 2020;26(5):672-5.

11. Tillett RL, Sevinsky JR, Hartley PD, Kerwin H, Crawford N, Gorzalski A, et al. Genomic evidence for reinfection with SARS-CoV-2: a case study. The Lancet Infectious Diseases. 2020.

12. Goldstein E, Lipsitch M, Cevik M. On the effect of age on the transmission of SARS-CoV-2 in households, schools and the community. medRxiv. 2020:2020.07.19.20157362.

13. Davies NG, Klepac P, Liu Y, Prem K, Jit M, Pearson CAB, et al. Age-dependent effects in the transmission and control of COVID-19 epidemics. Nature Medicine. 2020;26(8):1205-11.

14. Halloran ME, Préziosi M-P, Chu H. Estimating Vaccine Efficacy from Secondary Attack Rates. Journal of the American Statistical Association. 2003;98(461):38-46.