MATHEMATICALLY SIMULATING AN EPIDEMIC SPREAD TO IDENTIFY THE BETTER VACCINATION STRATEGY BETWEEN TWO ALTERNATIVES

Kapse, Gauri R; Jain, Shilpi

ABSTRACT

In this paper, we simulate an epidemic, first without any preventative measures, and then with preventative measures. We aim to contrast two vaccine distribution strategies of a two-dose vaccine against each other to identify the more effective strategy, when coupled with other preventative measures like masking, distancing, etc. The two strategies contrasted against each other are labelled vaccination strategy A (the first dose of the vaccine is distributed indiscriminately, and once 70% of the population has received the first dose, then distribution of the second dose begins) and vaccination strategy B (the first dose is distributed indiscriminately and the second dose is distributed immediately after a short priming period).

We find that vaccination strategy B results in fewer total infections than vaccination strategy A. We also find neither vaccination strategy reduces the total length of the epidemic compared to the other.

1 INTRODUCTION

1.1 Overview

The project followed requirements outlined in Problem #15 in the Spring 2022 Projects list for ISYE 6644. Model building was conducted in Python, and hypothesis testing was conducted in R. The team coordinated all communication on MS Teams.

The report begins with a literature review, a short description of the math behind our models, our model development procedure, our hypothesis testing procedures, and a section on our results and conclusions.

1.2 Background and Description

1.2.1 Types of Epidemiological models

Individuals in a population are mobile. When infected individuals carrying the pathogen move around and come into contact with susceptible individuals who are also mobile, transmission occurs and the disease spreads (Arino, 2021).

At the population level, this type of spread can be mathematically modelled by a compartmental model characterised by an initial value problem (Tulu et al, 2017). This type of modelling approach can be useful when simulating population level effects of an epidemic – for example, when assessing the impact of public health measures on the length of an epidemic, or the infection attack rate (defined later) of an epidemic (Karimi et al, 2015).

Compartmental models contrast with agent-based models, where individuals are assigned certain characteristics, and the simulation as a whole contains paths of movement, and a set of rules (Arino, 2021). For example, only a particular type of individual may enter a particular location (e.g., only employees may enter an office), or only a particular “path” may be taken to move between two

locations (e.g., a boat can't be sailed across land). Network-based models, where nodes in a graph represent individuals and edges represent contact between individuals can also be used to model epidemics (Arino, 2021). These types of modelling approaches are useful when the goal is to identify the patterns of disease spread, transmission hotspots, etc (Aleman et al, 2011).

Since this project aims to identify the better vaccination strategy between two alternatives, we are not concerned with the patterns of spread, but rather with the extent of spread, i.e., we are concerned with the overall population dynamics. Hence, we approach the problem using a compartmental model. All further discussion pertains to this type of modelling technique. Additionally, we only consider diseases that are transmitted through contact between individuals of the same population – vector borne or zoonotically contracted infectious diseases are out of the scope of this project.

1.2.2 The SIR Model

In 1927, Kermack and McKendrick put forth as a series of three differential equations describing a special case of the "process by which epidemics in limited populations run their peculiar courses" (Kermack and McKendrick, 1927, p.712). This series of ODEs became a classic compartmental model for the spread of infectious disease, commonly known today as the SIR model, where individuals progress unidirectionally between three states (compartments) – susceptible (S), infectious (I) and removed (R) – depending on their interactions with other individuals. In this scenario, individuals move freely (there are no pre-established movement patterns), the population is assumed to be wholly mixed, and each individual's 'actions' are randomly assigned (Kermack and McKendrick, 1927).

When a susceptible individual comes into contact with an infectious individual, and disease transmission occurs between them, the susceptible is re-assigned into the infectious compartment. Now, this person may infect further susceptible individuals they come into contact with. When an infectious person recovers or dies, that person is re-assigned to the removed compartment, and conferred immunity against the disease. This person may not get re-infected. Over time, as the epidemic progresses, the number of susceptible individuals declines, the number of removed individuals increases, and transmission rates decrease. This shifting distribution of the population is responsible for the end of the epidemic (Kermack and McKendrick, 1927).

Figure 1: In the SIR model described by Kermack and McKendrick in 1927, Individuals move unidirectionally from the S, to the I and to the R compartments.

1.2.3 The SEIR Model

There are many variations of the SIR model, but the one of interest to us is the SEIR model (Linka et al, 2020). In the SEIR model, once a susceptible individual contracts the disease from an infected individual, they are not immediately infectious themselves. Instead, they undergo a short incubation period, where the pathogen propagates within their bodies until its numbers reach a critical threshold. Once this threshold is reached, the infected individual becomes infectious. Hence, upon contracting the disease, we move the susceptible individual to the exposed (E) compartment, before moving them to the infectious compartment.

Figure 2: In the SEIR model, the SIR model is modified to include an 'Exposed' compartment. Infected individuals move into the Exposed compartment before they become infectious themselves. Individuals move unidirectionally from the S, to the E, to the I and to the R compartments.

A mathematical description of the scenario follows. This 'scenario' is a list of rules or assumptions that the disease spread follows, and consequently, constrains all models that simulate the disease:

- Since we are simulating a short-term epidemic (which we expect to come to an end, rather than become endemic), we ignore vital dynamics of the population (e.g. natural birth and death rates). The total population size is constant.
- Some individuals in the population become infected, and begin infecting the susceptible individuals they come into contact with.
- The number of susceptible individuals who get infected depends on
    - the number of infectious individuals in the system,
    - the contact rate between infected and susceptible individuals, and
    - the probability of disease transmission from an infected to a susceptible individual.
- Newly infected individuals take a few days to become infectious themselves.
- Once an individual becomes infectious, they remain infectious for a few days.
- When a person recovers or dies of the disease, that person is moved to the removed compartment, i.e., they cannot contract the disease again or spread it further.

Notation Guide:

- N is the total population
- S(t), E(t), I(t) and R(t) are the number of susceptible, exposed, infectious, and removed individuals respectively in the population at time t
- $\beta = \alpha * \tau$ where
    - α is the probability of disease transmission upon contact between an infectious individual and a susceptible individual, and
    - τ is the average number of contacts per person per unit time
- σ is the rate at which individuals transition from the exposed compartment to the infectious compartment. σ is calculated as the inverse of the average incubation period in days.
- γ is the rate at which individuals transition from the infectious compartment to the removed compartment. γ is calculated as the average infectious period in days.

Differential Equations:

Assuming that the population is constant,

$$N = S(t) + E(t) + I(t) + R(t)$$

By the Cauchy-Lipschiz Theorem, this initial value problem ensures that the system of ODEs has a unique solution.

The number of susceptible individuals each infectious individual can infect is the number of susceptible individuals among all the all the individuals with whom contact is made. Since the population is assumed to be wholly mixed and homogenous, this can be calculated as

$$\frac{\tau S}{N}$$

Therefore, the total number of susceptible individuals infected per unit time can be given as the product of the transmission probability, the number of susceptible individuals each infectious individual can infect, and the number of infectious individuals

$$\frac{\alpha \tau S I}{N}$$

From this, it follows that the rate at which individuals transition out of the susceptible compartment is

$$\frac{dS}{dt} = \frac{-\beta S I}{N}$$

Individuals transition into the exposed compartment at the same rate as they transition out of the susceptible compartment. However, individuals also leave the exposed compartment when they become infectious. So, the rate of change of the number of individuals in the exposed compartment is given by

$$\frac{d}{dt}E = \frac{\beta S I}{N} - \sigma E$$

By the same logic, the rate of change of the number of individuals in the infectious compartment is given by

$$\frac{d}{dt}I = \sigma E - \gamma I$$

Finally, individuals move into the removed compartment at a rate of

$$\frac{d}{dt}R = \gamma I$$

Additionally, since the total population size is constant, it follows that

$$0 = \frac{dS}{dt} + \frac{dE}{dt} + \frac{dI}{dt} + \frac{dR}{dt}$$

An important implication of these ODEs, is that if the exposed and infectious compartments were considered together, we see that

$$\frac{d}{dt}(E+I) = (\frac{\beta S}{\gamma N} - 1)\gamma I = (\frac{RoS}{N} - 1)\gamma I$$

Where R_o is the 'basic reproduction number.' Clearly, the spread of the infectious disease will only develop into an epidemic if $R_oS > N$ (Kretzschmar and Wallinga, 2009), i.e.,

$$Ro > 1$$

1.2.4 Disease-free equilibrium

The disease-free equilibrium is the time at which the number of infected (infectious + exposed) individuals is zero, and the population only contains susceptible and immune individuals (Tulu et al, 2017). Mathematically,

$$I(t)+E(t)=0$$

Generally, this is considered the time at which the epidemic comes to an end, as there are no more individuals who may transmit the disease to the population (Tulu et al, 2017).

1.2.5 Impact of Preventative Behaviours

When preventative measures are added into the mix, the model parameters change. As a consequence of masking and partial vaccination, transmission probability is reduced, and as a consequence of social distancing and quarantining, average contact rate is reduced. Thus, a reduction in α and τ causes an overall reduction in β.

1.2.6 Impact of Vaccination – the SEIRV Model

Vaccination reduces both, the rate at which individuals transition to the exposed (and subsequently infectious) compartment, and the number of susceptible individuals present in the population.

Because even partially vaccinated individuals are no longer equivalent to fully susceptible individuals, the SEIR model is no longer adequate when vaccinations are introduced during an epidemic. Hence, we modified the SEIR model to include vaccinations, and created a new SEIRV model.

When describing the SEIRV model mathematically, we make a few more assumptions:

- For our purposes, we consider a 2-dose vaccine
- Individuals who receive the first dose of the vaccine have a lower probability of contracting the disease upon contact with an infectious individual, than individuals who are merely susceptible
- Upon receiving the second dose of the vaccine, individuals become completely immune to contracting the disease. This means that there is no separate compartment for fully vaccinated individuals; rather, they move directly into the removed compartment
- The vaccination effort begins after the epidemic has already begun, once the proportion of infectious individuals has passed a certain threshold
- Receiving the first dose of the vaccine confers partial immunity to the receiver and reduces the risk of transmission upon contact with an infectious individual by 60%. There is no effect on the average contact rate in the population.

We also introduce new notation, in addition to what has been previously defined for the SEIR model:

- $V_1(t)$ represents the number of individuals who have received the first dose of the vaccine and have not contracted the disease.
- η denotes the rate at which individuals receive the first dose of the vaccine.
- μ denotes the product of the reduced transmission probability due to vaccination, and the average contact rate.
- ψ denotes the rate at which individuals receive the second dose of the vaccine.

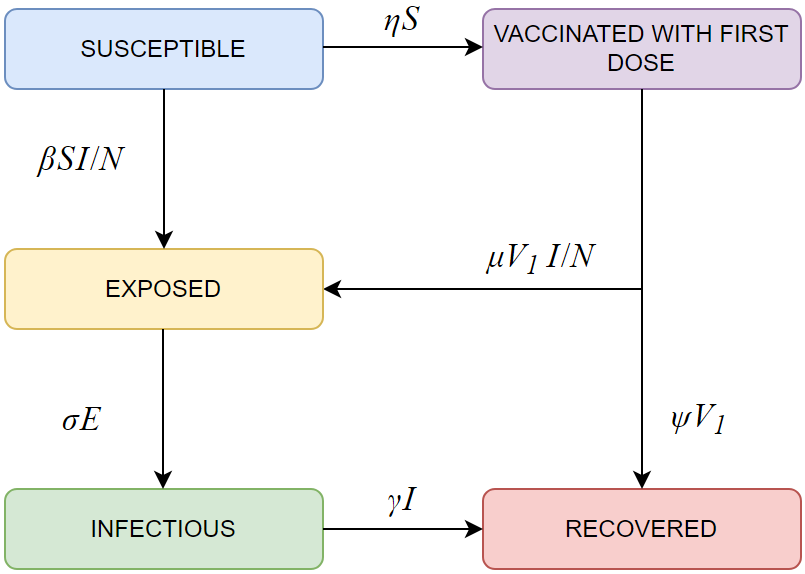

Figure 3: We modified the SEIR model to include a two-step vaccination course. The flowchart displays the direction and rates of transitions between compartments.

Given this new scenario, the differential equations are adjusted as:

$$\frac{dS}{dt} = \frac{-\beta IS}{N} - \eta S$$

$$\frac{dV_1}{dt} = \eta S - \frac{\mu V_1 I}{N} - \psi V_1$$

$$\frac{dE}{dt} = \frac{\beta IS}{N} - \sigma E + \frac{\mu V_1 I}{N}$$

$$\frac{dI}{dt} = \sigma E - \gamma I$$

$$\frac{dR}{dt} = \gamma I + \psi V_1$$

In line with the above assumptions, there are two vaccination strategies considered in this project. In both cases, susceptible individuals indiscriminately receive the first dose. The differences arise in the plan for distributing the second dose to partially vaccinated individuals. The two plans are:

1. Vaccination strategy A: Once all (or a certain threshold proportion) of susceptible individuals are partially vaccinated, provide the second dose to them.
   In this case, the rate at which the first dose is distributed is not necessarily equal to the rate at which the second dose is distributed. We assume that the rate at which the second dose is rolled out is lower than the rate at which the first dose is rolled out – in the real world, this may be due to supply chain issues, non-compliance, or partially vaccinated individuals contracting the disease and becoming ineligible for the second dose.

2. Vaccination strategy B: After a standard delay, all those who have received the first dose must receive the second dose as well.
   This necessitates that the rates of first and second inoculation are equal, i.e.,
   $$\eta = \psi \in (0,1)$$
   but the distribution of the second dose begins a few days after the distribution of the first dose.

1.2.7 Assessing the Impact of Vaccination

When a vaccine becomes available, it is of interest to public health authorities to implement the most effective distribution strategy (Kretzschmar and Wallinga, 2009). In this context, the most effective vaccine distribution strategy may be one that can curb the spread of the epidemic the fastest, or one that can minimize the total number of infections. Depending on the disease itself and the socio-economic conditions in which the epidemic arises, one interpretation may be favoured over the other. In this project, we evaluate the two vaccination strategies explained above, and compare their effectiveness based on both interpretations.

The metrics we use are the infection attack rates and the epidemic durations. The infection attack rate of an epidemic is defined as "the total proportion of the population that's eventually infected during the epidemic" (Kretzschmar and Wallinga, 2009). The duration of an epidemic is simply the number of days until the epidemic comes to an end, i.e., the number of days until the disease-free equilibrium is achieved.

There are other metrics for evaluating vaccine distribution efficiencies, and translating mathematical theory into actionable insights is a complex problem (Kretzschmar and Wallinga, 2009). Public health priorities may be impacted based on the spatial distribution of transmission hotspots, considerations for healthcare workers, complications due to supply chain issues, and the evolution of the pathogen into vaccine-resistant strains. In these cases, our simplistic evaluation would not be sufficient. However, these issues are out of the scope of this project.

2 METHODS

2.1 Model Development

Our code is based on a tutorial by Jay Gopalakrishnan, which he created for Portland State University's course titled Mathematical Computing with Data (Gopalakrishnan, 2020).

In order to compare the two vaccination strategies, we simulated four situations. First, we simulated a natural epidemic spread, i.e., a situation with no preventative measures in place, as a simple SEIR model.

We then adjusted it to create a second model, also SEIR, adjusted for preventative measures such as masking, social distancing and quarantines – the only difference is a lowered β.

The third and fourth models are built upon the second one. Model 3 corresponds to vaccination strategy A, and model 4 corresponds to vaccination strategy B – the vaccinations occur in addition to the self-initiated preventative measures of model 2.

Each model was initialised with a population of 1 million individuals,

$$N = 1000000$$

Among these, the number of infectious individuals is chosen as a random number between 1 and 50, i.e.,

$$I(0) \in [1,50]$$

There are no exposed, recovered or vaccinated individuals at the beginning of the pandemic. The total population is made up of susceptible and infectious individuals,

$$I(0) + S(0) = N$$

Depending on the model, we use the appropriate system of differential equations and solve the initial value problem to get the number of individuals in each compartment at time t = 1. Then, we reset the IVP, and solve for the number of individuals at time t = 2. We repeat this to calculate the number of individuals in each compartment for all subsequent time periods until the end of the epidemic.

We use the solve_ivp() function provided by the SciPy module in Python to facilitate these calculations.

We mark the end of the epidemic as the time when the disease-free equilibrium is achieved, i.e., the earliest time period when there are no more infected individuals in the population,

$$I(t) + E(t) = 0$$

For each model, we ran 50 replications, and the infection attack rate and the epidemic duration were recorded for each run of each model.

The disease spread parameters chosen for our model reflect our attempt to model a hypothetical disease spread similar to the spread of COVID-19 in the USA in April 2021. For the natural epidemic, we chose the basic reproduction number as $R_0 = 4$ based on the lower limit of proposed reproduction numbers for COVID-19, as stated in Linka et al. (Linka et al, 2020). This was reduced to 3 for models with preventative measures.

Similarly, during the course of the epidemic, isolation guidelines have consistently warned that the total of incubation and infectious period lasts for around 14 days. Based on this and the ranges described in Linka et al, we chose:

- Incubation period = 5 days, which implies that $\sigma = 0.2$
- Infective period = 9 days, which implies that $\gamma = 0.1$

β is calculated from its relationship to R_0 and γ.

Vaccination rates are calculated based on the daily count of first and second doses administered as of April 7th 2021 in the USA (2563505 and 2305364 respectively), as reported by the CDC (Centers for Disease Control and Prevention, 2022).

These daily counts of vaccine deliveries divided by the total population of the USA, taken as 331,893,745 (US Census Bureau, 2021) gives us the vaccination rates, η and ψ.

The exact epidemic spread parameters used in each model are tabled below (please refer to notation guidelines provided previously).

	R_0	β	μ	ψ	η
Model 1: Epidemic Spread with No Preventative Measures	4	0.4	0	0	0
Model 2: Epidemic Spread with Masking and Social Distancing	3	0.3	0	0	0
Model 3: Epidemic Spread with Vaccination Strategy A	3	0.3	0.12	0.007	0.008
Model 4: Epidemic Spread with Vaccination Strategy B	3	0.3	0.12	0.008	0.008

Table 1: The epidemic parameters used while building our simulations.

2.2 Hypothesis testing

We would prefer the vaccination strategy that delivers the greatest reduction in one or both of the infection attack rate, and the time taken to reach the disease-free equilibrium, as compared to the scenario where only preventative measures are used.

We used a Kruskal-Wallis test to determine which of the two vaccination strategies gives us the greatest reduction in the infection attack rate, and an ANOVA F-test to determine if there is any significant difference between the epidemic durations all four models.

3 RESULTS AND CONCLUSIONS

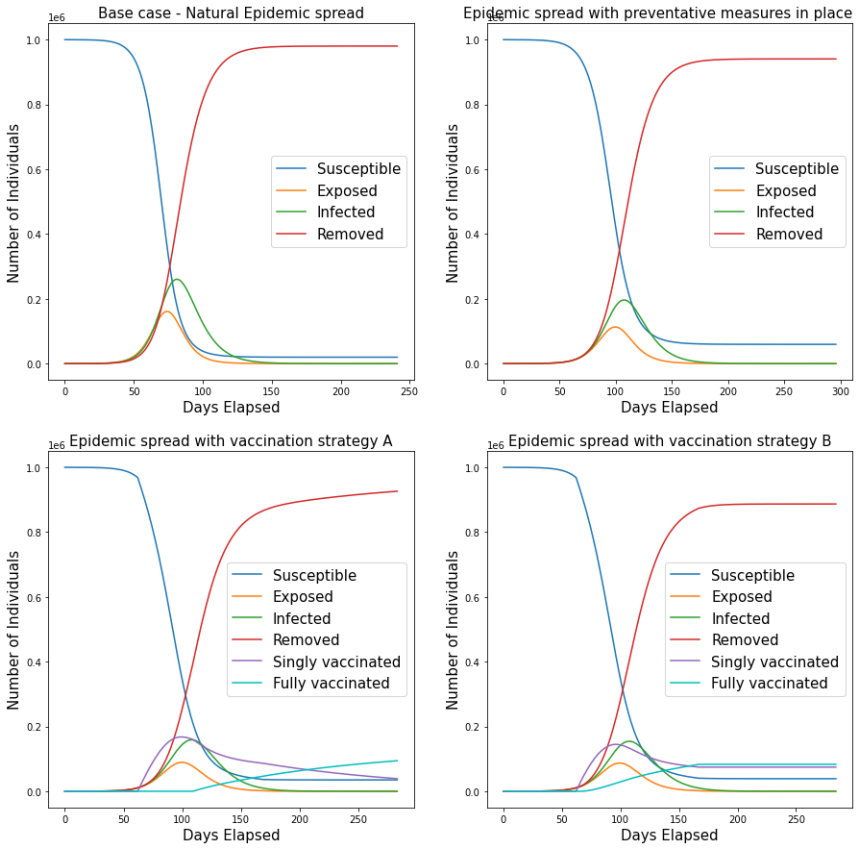

Figure 4: Example epidemic spread across the four scenarios, starting from the same initial distribution.

Figure 4 shows an example of the evolution of an epidemic spread over time, across the four different scenarios. In all four panes, the epidemic starts with the same initial conditions.

Note that in the base case, at the end of the simulated epidemic, the number of susceptible individuals remaining is almost zero. This is because most susceptible individuals rapidly get infected, then recover. In all other scenarios, the epidemic ends much before the number of susceptible individuals goes to zero.

Visually, we can see that the peak of the infectious compartment for model 2 is lower than the peak for model 1. The peaks for models 3 and 4 cannot be visually distinguished, but they are lower than the peak for model 2, as expected. However, from this plot, we cannot claim any difference between the two vaccination cases.

Instead, we calculate the infection attack rate, and repeat the simulation 50 times. The, we create box plots to examine the differences between the four scenarios, as shown below.

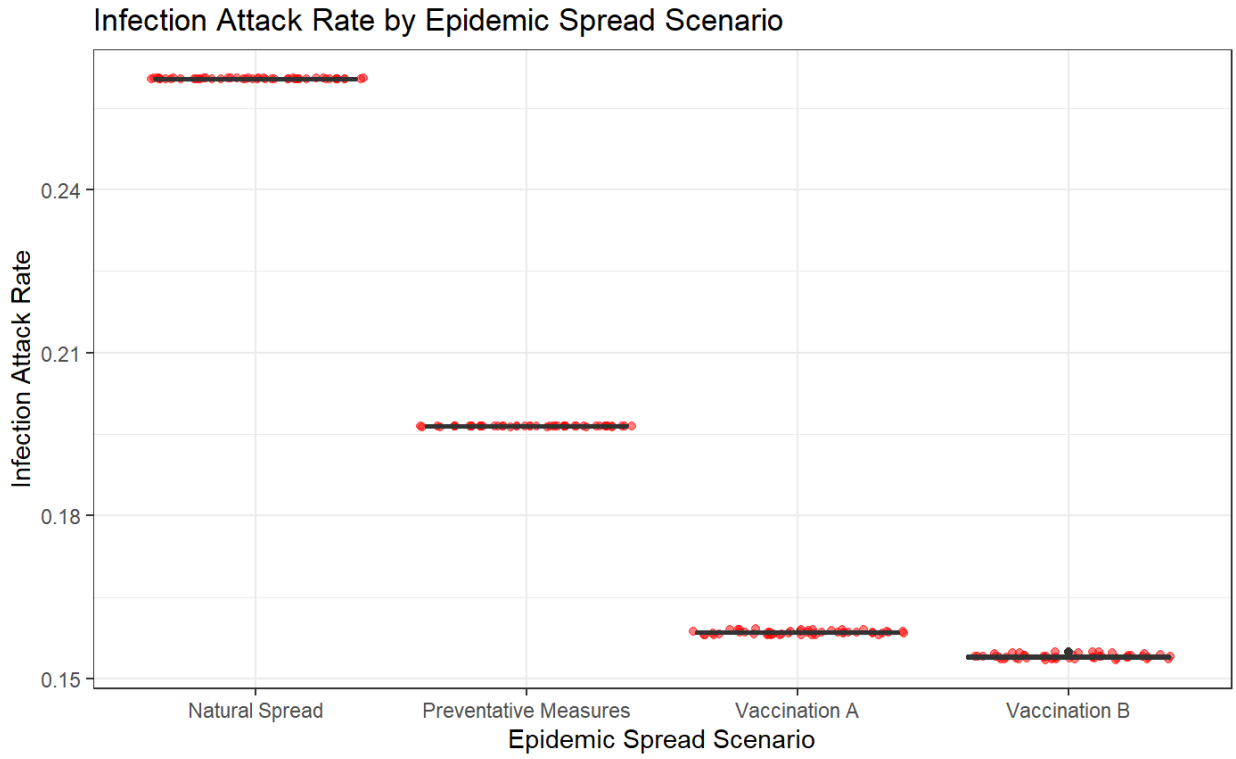

Figure 5: Box plot of infection attack rate for each epidemic scenario.

Figure 5 shows that despite the differences in the initial conditions, across 50 replications, the infection attack rates remain close to the median for all four scenarios. Visually, it seems like the infection attack rate for vaccination strategy B is lower than vaccination strategy A. This is confirmed by the Kruskal-Wallis test, which shows us that the mean infection attack rate for vaccination strategy A is statistically significantly higher than the mean infection attack rate for vaccination strategy B (p-value < 2e-16, code provided in attachment).

Similarly, Figure 6 examines the differences in epidemic durations across 50 replications of each of the four scenarios. From Figure 6, we see that the natural spread leads to a quick end to the epidemic. This may be because in the base case, the disease spreads quickly and infects a majority of the susceptible population. Once they recover, the disease cannot spread further as the number of susceptible individuals remining is low. By applying preventative measures, the epidemic is prolonged, but when preventative measures are coupled with vaccinations, this delay in reaching the disease-free equilibrium is tempered a bit.

As can be seen from Figure 6, vaccination strategy B has a higher median epidemic duration than vaccination strategy A. A Tukey-HSD test for pairwise mean differences shows us that this difference is not statistically significant.

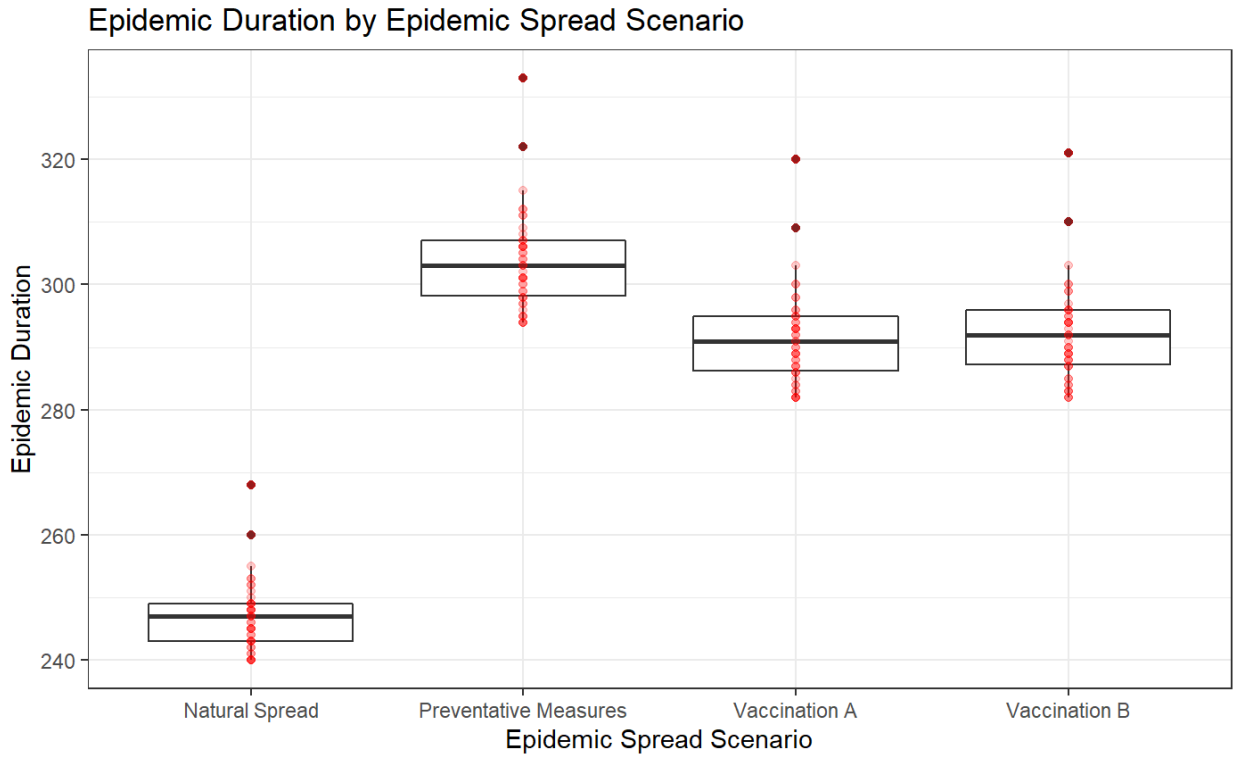

Figure 6: Box plot of epidemic duration for each epidemic scenario.

Based on these results, we can conclude that vaccination strategy B is 'better' as it results in a significantly lower number of infections than vaccination strategy A. Since both strategies result in equally long epidemics, we don't make any decisions based on epidemic durations.

These results are only valid for our hypothetical disease. If the epidemic parameters were changed, perhaps we would get a different result.

4 REFERENCES

1. Karimi, E., Schmitt, K., & Akgunduz, A. (2015). Effect of individual protective behaviors on influenza transmission: An agent-based model. *Health Care Management Science*, *18*, 318–333. https://doi.org/10.1007/s10729-014-9310-2
2. Tulu, T. W., Tian, B., & Wu, Z. (2017). Mathematical Modeling, analysis and Markov chain Monte Carlo simulation of ebola epidemics. *Results in Physics*, *7*, 962–968. https://doi.org/10.1016/j.rinp.2017.02.005
3. Aleman, D. M., Wibisono, T. G., & Schwartz, B. (2011). A Nonhomogeneous Agent-Based Simulation Approach to Modeling the Spread of Disease in an epidemic Outbreak. *Interfaces*, *41*(3), 301–315. http://www.jstor.org/stable/23016307
4. Kretzschmar, M., & Wallinga, J. (2009). Mathematical models in infectious disease epidemiology. *Modern Infectious Disease Epidemiology*, 209–221. https://doi.org/10.1007/978-0-387-93835-6_12
5. Jay, G. (2020, April 22). The SEIR model of infectious diseases. *Portland State University*. http://web.pdx.edu/~gjay/teaching/mth271_2020/html/_CONTENTS.html [Accessed 19 March 2022].
6. Wikimedia Foundation. (2022, March 18). *Compartmental models in epidemiology*. Wikipedia. https://en.wikipedia.org/wiki/Compartmental_models_in_epidemiology#The_SIR_model [Accessed 19 March 2022].
7. Arino, J. (2021). Describing, modelling and forecasting the spatial and temporal spread of COVID-19 – A short review. https://arxiv.org/pdf/2102.02457v1.pdf [Accessed 12 April 2022].
8. Kermack, W. and McKendrick, A., 1927. Contributions to the mathematical theory of epidemics—I. *Proceedings Royal Society London, A 115 700-721*. https://royalsocietypublishing.org/doi/10.1098/rspa.1927.0118 [Accessed 12 April 2022].
9. Linka, K., Peirlinck, M. and Kuhl, E., (2020). The reproduction number of COVID-19 and its correlation with public health interventions. *medRxiv preprint*. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7359536/pdf/nihpp-2020.05.01.20088047.pdf [Accessed 12 April 2022].
10. Centers for Disease Control and Prevention. (2022, April 12). CDC Covid Data tracker. Centers for Disease Control and Prevention. https://covid.cdc.gov/covid-data-tracker/#vaccination-trends_vacctrends-fully-daily [Accessed 12 April 2022].
11. Census Bureau, U.S. (2021, December 21). New Vintage 2021 population estimates available for the nation, states and Puerto Rico. Census.gov https://www.census.gov/newsroom/press-releases/2021/2021-population-estimates.html [Accessed 12 April 2022].