Vaccine Allocation Simulation Under Covid-19

Project Report for Simulation IEOR 4404

Victor Ye Dong, Ruturaj Bhagat, Tejas Gujar, Yingfei Fan

March 2, 2021

1 Introduction

We aim to simulate a decision-making scenario on who should be vaccinated first. Given the vaccination drive has already begun in many countries. We are simulating vaccine allocation policies for New York City. We take into account practical complexities using demographics of a population in NYC like age, occupation, any existing disease and many such factors which would affect the infection spread and fatality if infected with virus. Our goal is to simulate and compare the performance of different vaccine-distribution policies.

2 Value Proposition

Vaccination is one of the most important activity taking place today in the world across different countries. We through our analysis and simulation try to understand and find ways how this activity should be carried which involves whole of humanity under various circumstances. We specifically target New York City to compare different policies and provide a road-map for policy makers on different scenarios to enhance the effectiveness of policy decisions.

3 Model Methodology

3.1 Population Simulation

To create a representation of the population of New York based on features such as age, gender, occupation, housing category, height, weight and possibility of any chronic diseases. We consider the population distribution according to age and gender[?] and independently assign ethnicity[?] to them. The estimated heights[?] and weights[?] of the generated random variables are a function of their age and normalised over the average height and weight in New York. The housing category of a person can be divided into three for the city, namely, Nursing or home-care, homeless, and neither. Subsequently, we assigned occupation to the generated population independently based on their age[?]. Occupation can be allotted from

21 different categories, and a strict threshold of ages 19-65 can be employed for our model. For policy comparison process, we labelled certain occupation as essential worker or not based on norms and policies defined by New York State. The simulated population also had a metric to denote if the individual posses any chronic deceases such as cancer[?] or diabetes[?]. .The allocation of each individual features was done through inverse transform method.

3.2 Infection Spread

Infection spread is the rate at which a person could infect another person before the person realises that they have been infected and quarantines to begin the treatment. This spread rate in our simulation models is based on a very important assumption that occupation is the most critical factor which results in the contact and then infections spread. We collect occupation wise populations data(Po) for NYC[?] along with rate of COVID-19 spread in that particular occupation(Io)[?] in general. Using both these factors we derived a metric (Po*Io) which gives infection spread in NYC for every occupation.

3.3 Group-wide SIRD Model

To verify the vaccine policy performance, we introduce cross-group risk factor into the original SIR model[?], along with this we also consider deaths of the covid-infected patients as one of the results[?]. Similar as in the original model, every people are in one of the following five states :

$$S_t + I_t + D_t + R_t = N_t \tag{1}$$

where we give our denotation as follows ,

$$\begin{aligned} &S : Susceptible \\ &I : Infected \\ &R : Recovered \\ &D : Dead \end{aligned} \tag{2}$$

we divide our population into d groups based on demographic information. To incorporate the complexity of vaccine distribution , we follow a discrete-form model[?] Mainly following the model by [?], we give our formal specification of the group-wise SIRD model as follows,

$$\begin{aligned} \Delta S_{t+1} &= -\sum_i \sum_j \rho_{i,j} S_{it} I_{jt}/N \\ \Delta I_{t+1} &= \sum_i \sum_j \rho_{i,j} S_{it} I_{jt}/N - \gamma I_t \\ \Delta R_{t+1} &= (1-\delta)\gamma I_t \\ \Delta D_{t+1} &= \delta \gamma I_t \end{aligned} \tag{3}$$

specifically, the new infections in group j is $S_j \sum_k \rho_{j,k} I_k$[1], and $\sum_i^d I_i = I, \sum_i^d S_i = S$, and $d = 6$, while δ is the death rate. To capture the randomness of the infection spread, we further adopt the method in [?] such that the resolving from infection follows poisson distribution with rate of γ, so that it takes on average γ days to resolve for an infected patient, and each infected patients' recovering randomness is independent.

ρ measures the cross-group contact rate, the higher ρ_{ij} is, the more contact between the people in group i and group j. To better measure the cross-group contact rate risk, we categorize the contact rate based on the occupation, in which we assume that all children and seniors are unemployed. The reason for us to categorize the group by occupation is that we notice that contact rate is largely influenced by people's occupation property and it matches our data analysis for infection rate of different occupations. The occupation which requires large mobility and contact with people, health workers, service industry, government officials have higher infection risk than other occupation. Specifically, we categorize our risk group with cross-group risk matrix, with reference to similar idea of [?] and based on the group infection risk, we define our risk matrix as a 6×6 matrix. The assignment rule is as follows : for any row i, we give a core parameter such that $\rho_{i,i} = m_i$, the top-risk group has the highest $\rho_{i,i}$, then for increasing $j \neq i$, we slightly decrease the $\rho_{i,j}$ by step. The intuition is that, a risk group should have relatively lower contact rate with lower-risk group, and relatively high contact rate with people in the same risk group.

3.4 Death Risk Prediction Model

In this part, we modified the death risk prediction algorithm from QCOVID[?]. In the QCOVID model, a linear regression with LASSO is applied to the patients' data in England. The input predictor variables include demographic, lifestyle, and clinical information and the output measures the death risk due to confirmed or suspected COVID-19 given such information of an individual. Due to limitations of data accessibility, we simplified the model by combining variables if clinically similar in nature and by dropping variables that yield low estimated coefficients. To reduce bias, variable coefficients are revised based on the demographic trends of COVID-19 cases and deaths in the US [?]. Thus, in our final model, the input variables are: age, gender, ethnicity, body mass index, housing category(nursing home, homeless, or others), diabetes types, and whether one have cancer. The output is the predicted death rate(%) of an individual.

3.5 Vaccine Allocation

To simplify, we assume that the vaccine allocation only occurs after the next day's infection data has been updated. After the today's covid-related situation has been known, the vaccine starts to be allocated by the government to all the request people, which we assume to be all susceptible people. In the procedure of allocation, normally we assume the effective

[1]in the model of [?], they assume that the new infection is $S_j \frac{\sum_k \rho_{j,k} I_k}{(\sum_k \rho_{j,k}(S_k+I_k+R_k)}) ^{2-\alpha}$, we take $\alpha = 2$ to better match the canonical SIR model

rate is around $80 \sim 90\%$ [2], but we also consider the case when effective rate might have uncertainties. In the procedure, the government follows a consistency policy in each simulation. We give our simulation for policy comparison in Section 4.3.

4 Simulation Experiment

4.1 Basic Model Characterization

In this section, we give a brief characterization of the model input, procedure, and output. For a fixed simulation period, we assume that there is only one vaccine allocation center to implement the vaccine allocation policy, the vaccine number is given at the first day and will be distributed with a capacity among each day. On each day, vaccine distribution occurs after daily infection-resolving-recovery data has been collected. At the end of each day's simulation, we collect the following statistics : 1) number of infected people; 2) number of people with anti-body; 3) death number; 4) number of the susceptible; 5) number of the uninfected (=all citizen number − (1) − (3)).

4.2 No-vaccine model simulation

In this section, we show outputs of the simulation without vaccinating people. See from Figure 1 that the curve of infected number looks similar to the one in a standard SIR model (bell curve). Since there are no vaccines[3], the anti-body number increases mainly due to infected patients recovering from infection.

Figure 1: basic

To end this section, we show how to use control variate method to reduce the variance of a system output[4]. We are interested in how many people gets anti-body after the last day, therefore we use the resolving period as a control variate, which decides whether an infected people will get anti body before the last day, therefore related to the system output.

4.3 Policy effectiveness

In this part, we describe the vaccine allocation policy algorithm and compare the policy performance based on the interested indicators. To measure the performance of each policy, we particularly focus on tracking daily death number and infected number. For each policy, we run at around 50 simulations on 10000 citizens and obtain the estimates by taking the average of the statistics.

[2]https://investors.modernatx.com/news-releases/news-release-details/modernas-covid-19-vaccine-candidate-meets-its-primary-efficacy

[3]the request number in this case is all the susceptible people

[4]see our colab for system output, we skip the result to save space

To fit in the distribution policy guided by CDC [5], we typically categorize the priority group to be a) health workers; b) essential workers; c) seniors ; d) other citizens. Overall, we are interested in the comparison of these four algorithms : 1) Uniform policy (Unif): uniformly sample people to distribute the vaccine; 2) Uniform Priority (UP) : uniformly samples the people from each priority group based on the each group's susceptible number; 3) Strict Priority (SP): strictly gives the time window for each priority group's order to take the vaccine; 4) Mixture Priority (MP) [6]: a modification to the SP policy by giving mixture time window for each priority group's order to be vaccinated; 5) Age-prioritize policy (pA): same procedure as (SP) but with prioritizing seniors first in grouping part.

Currently the SP policy is considered the most by the NYC government. It strictly gives the vaccination order for health workers, essential workers, and then seniors. However, constraint to the target to prioritize several groups, a natural question is that if this framework is really optimal method, if not in certain condition, then what should we consider as an alternative policy.

4.3.1 Effectiveness under different vaccine numbers

The first important uncertainty in the real world is the supply of vaccines. Therefore we compare policy performance given different vaccine numbers. We show our simulation results at Figure 2, Figure 3

ill_num.jpeg

Figure 2: Death number by days

infected_num.png

Figure 3: Infected number by days

From Figure 2, Figure 3, we see that policies perform similarly given a limited vaccine number. The intuition is that, when the vaccine supply is limited, the overall effect for the optimal policies is small. When the daily vaccine capacity is relatively large, there is more space for the policy optimization to "flatten the curve" before the infected arrives at the peak since after the peak a large portion of people either has anti-body or going to the severe illness or death. With the increasing number of provided vaccines, we can see from the plots that policies SP, MP, and pA yield better results in controlling infected cases than uniformly allocation without prioritizing each group. Particularly, MP performs well in mixture the priority by giving mixed time window for each priority group.

[5] https://www.cdc.gov/vaccines/acip/meetings/downloads/slides-2020-12/COVID-02-Dooling.pdf

[6] each priority group has a priority sliding window, if each other's time window does not intersect, then the vaccination policy would be strictly in order, i.e., SP. If not, then would be mixture policy, i.e., largely in order but allow mixture, e.g [0.1, 0.3] and [0.27, 0.5]

4.3.2 Robustness under different vaccine effective rate

For this section, we assume that the real effectiveness of the vaccine might have uncertainties. We would like to see if the policy is robust again lower or medium vaccine effectiveness rate. We compare three policies that prioritize certain groups : SP, MP and pA.

effective_rate.png

Figure 4: policy performance under different effective rate

From the graph, we see that the MP policy is still robust compare to the two strict priority policy.

4.3.3 Policy optimization for mixture policy

Finally, we would like to know if we could optimize the mixture policy, the main parameters we would like to optimize on is the priority period [7], $\{t_1, ..., t_d\}$ for d priority groups, which decides the time window period that allows the mixture of different priority group. Since we have 4 priority groups, we fix the first two and try to loop over the left two. For each simulation, we compare the output indicator (infected number) compared to the other policies. Our final optimal parameters would be a set that beats the rest of the policies and also the best compared to itself. Our finding is that actually we could find out an optimal sliding window set that allows that vaccination for multiple priority groups. Therefore we recommend that the government should adopt more flexible mixture while maintaining the priority order of the priority groups.

5 Conclusion and limitations

In this project, we are focused on the analysis of vaccine allocation policies in the covid-19 period. Our contribution are mainly threefold : 1) we simulate a group-wise SIRD infection model based on demographic information ; 2) we compare the policy performance from multiple dimensions, including effectiveness and robustness; 3) we found out that even under the priority framework the current policy implemented by the NYC government might not be the optimal, we claim that in certain conditions the optimal policy would be using mixture policy to mix the priority groups instead of only giving to one group strictly.

Meanwhile, we also notice that there exists certain limitations and constraints in our work. The first is that it would be interesting to see how our model truly match the real data; Secondly, with a complex model in our case, it takes us around several hours to implement a complete [8]. Given more time, it would be interesting to see how our simulations result will have change because of lack of simulations rounds. Thirdly, a natural conjecture of an

[7] see details in colab

[8] $20 \sim 50$ simulations out of $10000 \sim 20000$ people for $30 \sim 60$ days with $16 \sim 20$ combinations of parameters and policies

optimal policy form is more adaptive (allocate likes a bandit system) based on the updated infection and the death rate and tries to control either one if certain group of people perform really bad, it would be interesting to see if this policy could be applied to the real world policy design.

6 Links

Colab:https://colab.research.google.com/drive/1v8-i4CzN13xfRQ70GmKO3yK2Uiu3LmHw#scrollTo=Cnhjae1VgyfR
Drive:https://drive.google.com/drive/folders/1o7QiqzadiCyE4zTiuXZYP393-h0Y_pzI?usp=sharing
Video:https://columbia.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=65fe9374-3bd1-49e

7 Contribution

Equal contribution to general system structure and policy implementation. Also everyone has focused one work component independently as follows : **Victor:** System **Yingfei:** Death risk model **Tejas:** Infection Spread **Ruturaj:** Population Simulation