

Research Article

Simulation of Herd Immunity in Covid-19 Using Monte Carlo Method

Farshi E*

Department of Respiratory Diseases, Anahita Inc, San Diego, USA

***Corresponding author:** Farshi E, Department of Respiratory Diseases, Anahita Inc, San Diego, Californian, USA**Received:** October 01, 2020; **Accepted:** November 12, 2020; **Published:** November 19, 2020**Abstract**

Background: Coronavirus (COVID-19) is number one hot topics in the world. It seems Herd Immunity is going to play very important role in COVID-19 issue because of lack of treatment drugs and vaccination. Herd immunity happens when so many people in a community become immune to an infectious disease that it stops the disease from spreading. It is one of two ways to achieve high levels of immunity. The other is a vaccine. Please note that herd immunity thresholds constitute convenient targets for vaccination coverage too. Therefore, present simulation is very important and applicable for both herd immunity and vaccination scenario. Calculation of important factors, especially basic reproductive ratio R_0 , may be used in both vaccine scenario and herd immunity. It may be also used for prediction of COVID-19 behaviour in any city or society.

Method: There is no any valid simulation in such topics in the world. Our simulation is based on Monte Carlo Method. We used a Monte Carlo software developed for random walk plus virus transfer Monte Carlo code inside a city or society.

Results: We applied those software's to our purpose for calculation of herd immunity, vaccination factors, basic reproductive ratio, and prediction of COVID-19 behaviour.

Conclusion: We applied our simulation for New York city as an example but it could be applied to any city or society. Please note that ability of our software is very much but we present here only some parts of its abilities.

Introduction

At the end of 2019, a novel coronavirus (COVID-19), was identified as the cause of a cluster of pneumonia cases in Wuhan, China. It rapidly spread, resulting in an epidemic throughout China, with sporadic cases reported globally [1].

Monte Carlo methods, or Monte Carlo experiments, are a broad class of computational algorithms that rely on repeated random sampling to obtain numerical results. The underlying concept is to use randomness to solve problems that might be deterministic in principle. They are often used in physical and mathematical problems and are most useful when it is difficult or impossible to use other approaches. Monte Carlo methods are mainly used in three problem classes: optimization, numerical integration, and generating draws from a probability distribution [2].

Simple example for showing a Monte Carlo method is calculation of probability of getting 6 in playing dice by mathematics that is $1/6$. In Monte Carlo method we play dice many times (for example 100,000 times) then get statistics of every number including 6 then we reach to same $1/6$ without using any mathematical formula.

There are two ways to achieve high levels of immunity against COVID-19: 1) Vaccination, 2) Herd Immunity. Herd immunity is a concept in epidemiology that describes how people can collectively stave off infections if some percentage of the population has immunity to a disease.

We developed a software program to simulate herd immunity and a vaccination in a city. We apply our simulation to New York city as a typical city because we have many statistics from that city and its high-resolution map is available [3-7].

Our Monte Carlo simulation is based on random walk method. A random walk is a mathematical object, known as a stochastic or random process, that describes a path that consists of a succession of random steps on some mathematical space such as the integers.

An elementary example of a random walk is the random walk on the integer number line, Z which starts at 0 and at each step moves $+1$ or -1 with equal probability [8,9]. The term random walk was first introduced by Karl Pearson in 1905 [10]. If the state space is limited to finite dimensions, the random walk model is called simple bordered symmetric random walk, and the transition probabilities depend on the location of the state because on margin and corner states the movement is limited [11]. Dimensional random walk, but we will apply 3 dimensions simulation inside buildings with different floors. To visualize the two-dimensional case, one can imagine a person walking randomly around a city (New York city in our simulation).

At every intersection, the person randomly chooses one of the four possible routes (including the one originally travelled from). Formally, this is a random walk on the set of all points in the plane with integer coordinates.

While roads and streets of our assumed city are almost square

lattice but inside every lattice we have applied finite element meshes. Person walks or drives in square grids but inside lattice the person can walk in any side and in non-square grids. Our problem will be different than random walk because in our case we have many people who are walking and we have possibility of colliding people to each other (actually not really colliding but becoming more close to each other's safe social distance) to infect virus to each other.

When an infected pedestrian walks in streets (border of lattices), there is very low chance that person infects other person who is walking in same street. But in cross sections when people waiting for red light, there is possibility of infection if social distance is low there. However, main infections happen inside meshes of lattices. These meshes are inside downtown, shopping centers, metro stations, stadiums, malls and such crowded places.

As mentioned before, we have used New York city as our model. Our computer simulation has been developed to decide in random walk directions mostly toward high population places of the city (like metro stations, cinemas, concerts, shopping centers, malls and other crowded places). However, our program has ability to use Google trace information for more exact simulations and results. It is evidence we may apply any city to our program also.

In addition to random walk we have used simulation of neutron transportation inside nuclear reactor using Monte Carlo method. We revised that software to our purpose for calculation of herd immunity and vaccination scenario. Neutrons collide to nucleons and results are releasing 1-3 more neutrons that we may simulate it for infection of COVID-19 by a person to several other people. Inside reactor, geometry, and area of Uranium and moderator has been revised to our assumed city for our purpose. One may see the flowchart of computer programming, that is combination of random walk and reproductive of virus in the city.

We define "R naught" or R_0 to estimate how many other people one sick person is likely to infect. In very simple mathematics, the Herd Immunity Threshold (HIT) is defined by $1-1/R_0$ (e.g., if $R_0=4$, the corresponding herd immunity threshold is 0.75) [3]. Once the HIT is passed, the rate of new infections starts to decline, which should ensure that health systems will not thereafter be overwhelmed and makes it more practicable to take steps to eliminate the disease.

This R_0 for COVID-19 is roughly about 2-6 or even more. Scientist have neither an experimental number nor calculation number for R_0 and it is one of biggest problems in Simulation and understanding of COVID-19 behaviour. In very simple mathematics, the herd immunity threshold is defined by $1-1/R_0$ (e.g., if $R_0=4$, the corresponding herd immunity threshold is 0.75) [3]. Therefore, the more communicable a pathogen, the greater its associated R_0 and the greater the proportion of the population that must be immune to block sustained transmission. A similar parameter important for understanding population-level immunity is the effective reproduction number (R_e or R_t). R_e is defined as the average number of secondary cases generated by a single index case over an infectious period in a partially immune population [4].

R_0 is not a biological constant for a pathogen, a rate over time, or a measure of disease severity, and R_0 cannot be modified through vaccination campaigns. R_0 is rarely measured directly, and modeled

R_0 values are dependent on model structures and assumptions.

Ultimately, the goal of vaccination programs is to bring the value of R_e below

This occurs when the proportion of the population with immunity exceeds the herd immunity threshold. At this point, pathogen spread cannot be maintained, so there is a decline in the number of infected individuals within the population [5]. As R_0 increases, the proportion of the population that must be immune to generate herd immunity increases ($1-1/R_0$). The above interpretation of R_0 and its relation to the herd immunity threshold is the simplest understanding of these terms. It relies on several key assumptions, including homogeneous mixing of individuals within a population and that all individuals develop sterilizing immunity—immunity that confers lifelong protection against reinfection—upon vaccination or natural infection. In real-world situations, these epidemiological and immunological assumptions are often not met, and the magnitude of indirect protection attributed to herd immunity will depend on variations in these assumptions, then in our Monte Carlo calculation we will not assume such epidemiological and immunological assumptions. R_0 is defined by both the pathogen and the particular population in which it circulates. Thus, a single pathogen will have multiple R_0 values depending on the characteristics and transmission dynamics of the population experiencing the outbreak [4]. This inherently implies that the herd immunity threshold will vary between populations, which is a well-documented occurrence [4]. For any infectious disease, communicability depends on many factors that impact transmission dynamics, including population density, population structure, and differences in contact rates across demographic groups, among others [3]. All of these factors will directly or indirectly impact R_0 and, consequently, the herd immunity threshold. Since the onset of COVID-19 spread, various studies have estimated the basic reproductive number (R_0) of the virus to be in the range of 2 to 6. From an initial cohort of 425 confirmed cases in Wuhan, China, an R_0 of approximately 2.2 was estimated, meaning that, on average, each infected individual gives rise to 2.2 other infections [6]. More recent estimates place the R_0 higher at 5.7, although many estimates fall within this range [7]. This variation reflects the difficulty of obtaining accurate R_0 estimates in an ongoing pandemic, and the current estimated COVID-19 R_0 values likely do not indicate a complete picture of the transmission dynamics across all countries. In our Monte Carlo simulations we will not assume a certain R_0 or R_e , but our software will calculate it according to the characteristics and transmission dynamics of the population experiencing the outbreak. Then, our software is also a good method for estimation of R_0 or R_e in any society experiencing outbreak.

We define n as the potential number of susceptible neighbours. The precise form of this parameter can only be calculated by knowing the exact contact network, but in general, networks with predominantly long-range connections (such as the random graph) initially, whereas networks with many local connections (such as lattice-based models plus small mesh networks) have far lower n . These local networks have reduced n for two reasons: Firstly infectious individuals are aggregated so each infectious case has fewer susceptibles in its neighborhood, and secondly, some susceptibles are shared between many sources of infection again reducing the potential n . Although



Figure 1: a) Preliminary estimation of R_0 by some scientists, but we will not use this data in our calculation and simulation. Our software will calculate R_0 dependent on many factors that impact transmission dynamics, including population density, population structure, and differences in contact rates across demographic groups and other factors. b) Map of Manhattan New York used for our test simulation.

we could calculate n from the contact network, as is selectively done in ref. [12], there is only limited data from which to estimate this network; but our Monte Carlo simulation will calculate it.

Please note in Monte Carlo simulation we don't need to assume a certain R_0 , R_e , or n because nature of Monte Carlo method is based on random sampling to obtain numerical results to be deterministic in principle. In ordinary methods a person is assumed and probability of his moving direction and virus infection to other people are calculated. In Monte Carlo simulation we consider population of sample in NYC and randomly estimate their walk directions and locations.

Therefore, we will find collide and contact of people in our simulation. If sample was 100 people the result was very far of real but since we do it for huge number of people then final result will be deterministic. As mentioned before, one example is calculation of probability of getting 6 in playing dice by mathematics that is $1/6$. In Monte Carlo method we play dice many times (for example 100,000 times) then get statistics of every number including 6 then we reach to same $1/6$.

This example show how Monte Carlo Method works and why we don't need exact numbers of R_0 , R_e , or n in Monte Carlo Simulation but program will give us ultimately number of such factors.

Therefore, we will have reliable results using Monte Carlo method to get important numbers and factors of herd immunity to give such reliable data to governments for their plans against COVID-19.

Results

Our main purpose is find how much percent of people in New York City become immune to COVID-19 disease that it stops the disease from spreading. It means herd immunity or Vaccination. Our other purpose is to find how long time takes to reach herd immunity in New York city. Our last purpose is to find which factors are affecting herd immunity, then we may use them to reach herd immunity in low percent of immune people. It gives idea about vaccination scenario in the city. Please note it is impossible to vaccinate 7 billion people of the world after making vaccine for COVID-19. We need simulation to find how much percent of people in a city should be vaccinated.

Like reference [13-17], we use a simple 'SEIR' epidemiological symbols, in which the population is divided into four compartments: Susceptible (uninfected), Exposed (latent: infected but not yet infectious), Infectious (typically when diseased), and Recovered (and thus immune and harmless).

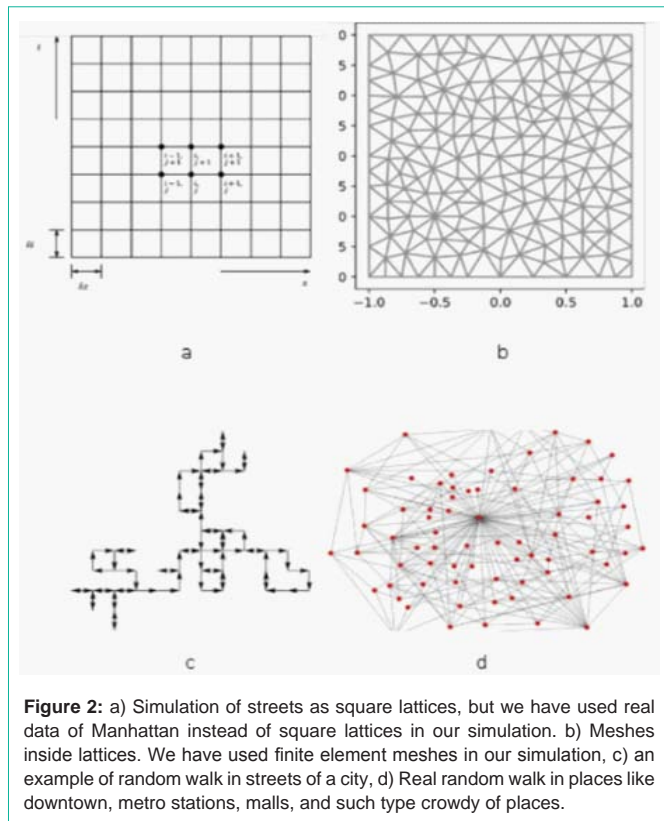
People from above four compartments walk or driving (generally moving) is through lattices but they can enter inside lattices too. We have used finite elements meshes inside lattices to cover all area of the city. The exact way for finding their moving is to use Google moving trace but here our lattice and finite elements meshes give enough estimation for our purpose. We don't assume any average social distance because Monte Carlo simulation will find it in simulation experimentally. We again emphasize that Monte Carlo method is not a formula, mathematical methods or model. It is like hundreds of thousands staffs are going inside city and get experiments and measure all parameters for several million people of that city. Then they give such results to computer. Actually instead of hundreds thousands staffs, our computer makes those works as simulation.

Infection rate in downtown of NYV will be high because of high density of people and small social distances.

It is clear such simulation is very complicated then we will use super computer. One may see brief flowchart of Our computer program in. This flowchart is mix of random walk and neutron transport inside a nuclear reactor. Neutrons behavior inside reactors are similar to people who infect virus to other people. When a neutron collides nucleon, it may cause fission plus releasing 1-3 additional neutrons. We simulate such behavior of neutron as transferring infection from one person to several people if they become in same mesh (type of collide of neutron to nucleon).

1) Zero percent Immune people in city

It means almost near 100% of people are susceptible and rarely immune. In this situation few infected people enter in the city, and there is no quarantine and social distance rules in the city, all people will be infected after a while. It theoretically will take about 3 months that all New York city people to become infected but in reality because most of infected people will recover and will become immune, we will never experience of 100% of people get infected. We will discuss



details later.

- 2) 100% of immune people in city

It is clear in this situation infected people can't infect disease to any people. The result shown in Figure 4b. This part and previous section were just for first testing of our Monte Carlo simulation. This is typical first test of some routine models to be sure our software works well.

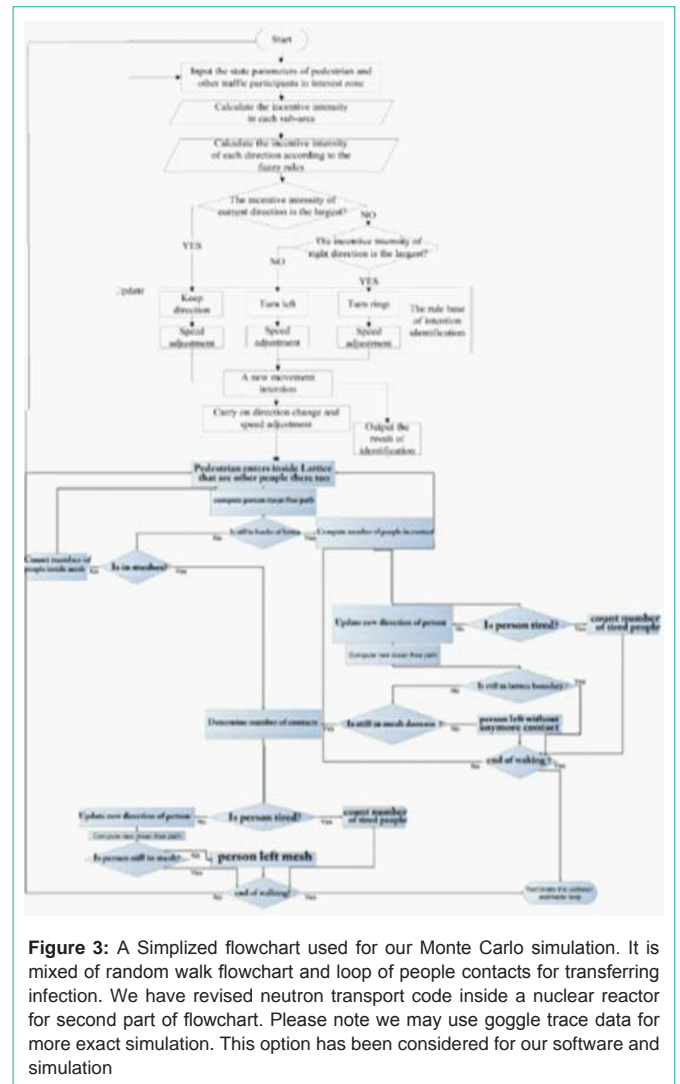
We may also test reliability of our Monte Carlo simulation through result of. This Figure shows infected people in NYC in period of March 1-21, 2020. We have chosen this period because of lack of lockdown and quarantine in the city.

According to official reports of NY government, official number of infected people should be multiplied 10. Positive cases track only people with a positive COVID-19 test result- because testing has been limited, the number of people who have actually been infected with COVID-19 is much higher- about 10 times the positive cases, according to New York State's antibody screenings.

Therefore, our simulation predicts infectious behaviour of the city very well in that period of time. This is another reliable result to get confidence to our Monte Carlo simulation.

- 3) We have four groups in the city

We had four compartments in the city, 1) Susceptible (uninfected), 2) Exposed (latent: infected but not yet infectious), 3) Infectious (typically when diseased), and 4) Recovered (and thus immune and harmless). In reality, the Recovered compartment includes people who instead die, which has the same effect on the model dynamics.



The entire population starts in the Susceptible compartment, save for a tiny proportion that are transferred to the Infectious compartment to seed the epidemic. The seed infectious individuals infect Susceptible individuals, who move to the Exposed compartment. Exposed individuals gradually transfer to the Infectious compartment, on average remaining as Exposed for the chosen latent period. Infectious individuals in turn gradually transfer to the Recovered compartment, on average remaining as Infectious for the selected infectious period [13].

Our Monte Carlo simulation shows when 41% of population of NYC become immune against COVID-19, the city will reach to the Herd Immunity Threshold (HIT) and 65% of the population ultimately becomes infected. This result seems strange and very far from prediction of the simple model of $(1-1/R_0)$ that shows herd immunity threshold will be near 60% and finally 82% of population will be infected.

We will explain the reason of difference and why our result is correct. There are two reasons: 1) inhomogeneous structure of population in NYC, 2) n as the potential number of susceptible neighbors. Figure 8a shows only population inhomogeneity of land

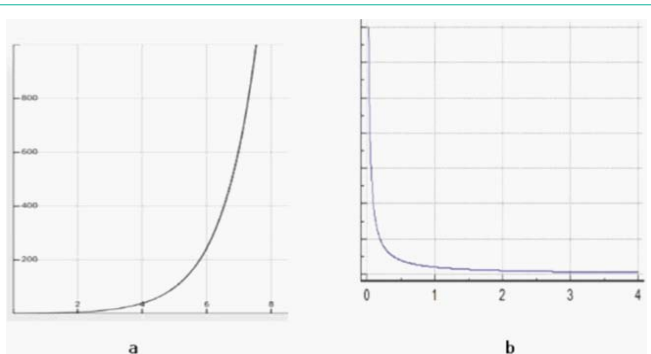


Figure 4: a) Increasing infected people as progression when all people are vulnerable without immune people. b) Rapid decrease of infected people when near 100% of people are immune. It is routine to test any simulation in the first step with simple situations that we know results. It shows in the first step our simulation and calculations are reasonable.

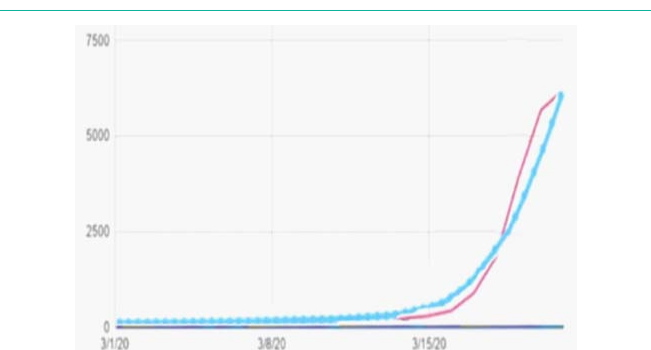


Figure 5: Number of infected people in NYC between March 1 and 21, 2020. Red is official report and blue is result of simulation with our Monte Carlo code. Please note we considered this period of time because no any social distance or quarantine was applied in NYC in that period. The second point, we have multiplied our Monte Carlo result at 10 because official reports were about 10 times less than real infected people.

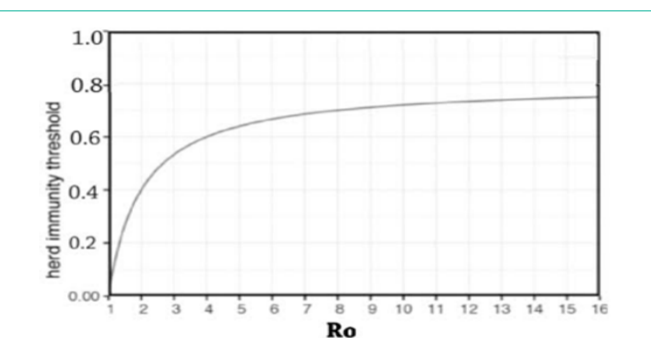


Figure 6: This shows R_0 number versus herd immunity threshold (HIT) for NYC obtained from Monte Carlo simulation. It is different from simple formula $(1-1/R_0)$ calculated by others because in our simulation are several key factors, including non-homogeneous mixing of individuals within a population and that all individuals develop sterilizing immunity—immunity that confers lifelong protection against reinfection—upon vaccination or natural infection.

size in NYC. The Coefficient Variant (CV) related to it is about 0.65-0.70 according to mathematical calculation. If we consider other inhomogeneity of population, it is reasonable to have CV around Herd Immunity Threshold (HIT) versus CV plotted from a gamma function. If $CV=1$ then HIT will be about 40% that confirms

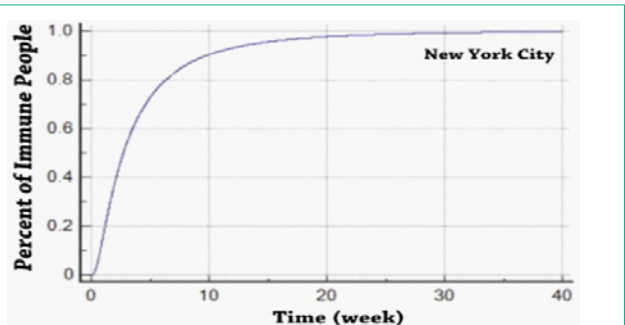


Figure 7: Percent of Immune people versus time (week) in NYC 90% Immune people after about 10 weeks.

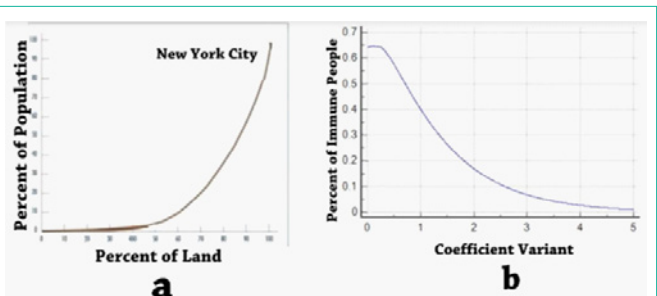


Figure 8: a) Heterogeneity of NYC population versus land size. This heterogeneity alone shows the coefficient variant (CV) about %70. Then it is reasonable to have about $ICV=1$ for total non-homogeneous of NYC. b) Immune population percentage versus CV. When $CV=1$ then the herd immunity threshold (HIM) will be around %40 that confirms our Monte Carlo simulation result.

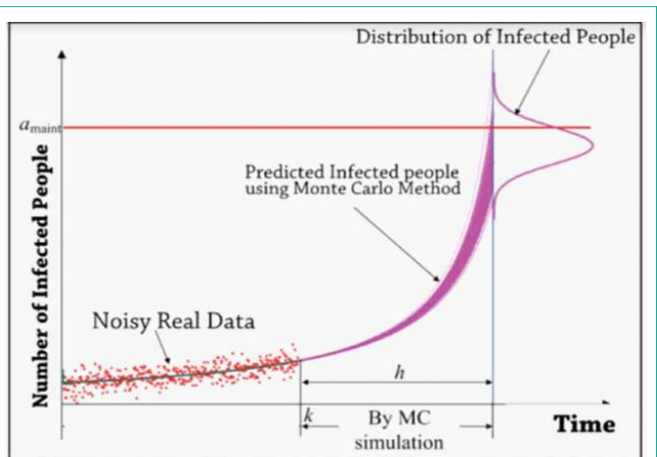


Figure 9: Prediction by Monte Carlo Simulation. This is one of hundreds options of Monte Carlo. Even with preliminary noisy real data we may predict next behavior with good accuracy using Monte Carlo simulation. I

our Monte Carlo simulation result. The effect of inhomogeneous population on HIT is very important, because governments may have plan and programs to reach herd immunity in shorter percent of infected people.

In one may see R_0 versus Herd Immunity Threshold (HIT) of NYC obtained from our Monte Carlo simulation. This is different from R_0 obtained from $(1-1/R_0)$ in some references. For example

when $R_0=2.5$, then HIT will be 0.75 according to $(1-1/R_0)$, but it will be about 0.43 according to our Monte Carlo simulation. As we mentioned before, several key factors are implicitly in our Monte Carlo simulation, for example inhomogeneous mixing of individuals within a population and that all individuals not develop sterilizing immunity—immunity that confers lifelong protection against reinfection—upon vaccination or natural infection.

One may see infected people's percent in Manhattan versus time (month). It takes about two months to reach Herd Immunity Threshold (HIT) in Manhattan. But interesting factor is that HIT in Manhattan area is 50% while in whole NYC is about 40%. Difference may be related firstly the population in Manhattan is more homogeneous comparing other areas of NYC, structure of population, and far low n number in Manhattan. New York city officials confirm surprisingly that number of infected people in Manhattan is less than other places of NYC.

Time needed for herd immunity depends on some parameters like population of city, area of city, time needed recovery after infection, distribution of infected people in city (inhomogeneous of infected people). Our simulation shows if quarantine was not applied to NYC, it could reach to herd immunity threshold in 70 days.

4) Prediction by Monte Carlo Simulation

Prediction by Monte Carlo simulation is one of hundreds useful features of this type simulation. It is applicable in epidemics especially COVID-19 because preliminary data are not complete and there are noisy in such input data.

Please note that time running a super computer for present Monte Carlo simulation may take one or two weeks but the results are thousands factors that need several months analysis. Here we presented only few.

Conclusion

Monte Carlo simulation is so far only method for exact investigation of herd immunity and vaccination scenario. Other methods are not exact and reliable.

There are two possible approaches to build widespread COVID-19 immunity: (1) a mass vaccination campaign, which requires the development of an effective and safe vaccine, or (2) natural immunization of global populations with the virus over time. However, the consequences of the latter are serious and a large fraction of the human population would need to become infected with the virus, but it is still an option in the absence of a vaccination program, then we need exact simulation for it. Such simulation may be used for vaccination scenario too. Then our Monte Carlo simulation may be applied to both herd immunity and vaccination scenario. Please note for example it is impossible to have vaccination campaign for 7 billion people in the world. Our Monte Carlo simulation may help to decide how much percent of people should be vaccinated to reach immunity of society against COVID-19.

While there is ethical path to reach herd immunity, but some cities (like Bergamo in Italy) have already reached to herd immunity as unwanted. In case of delay for reaching vaccination, many cities in the world will reach automatically to herd immunity then we need a

real and exact simulation for herd immunity scenario of such cities.

In reference [16] estimations of COVID-19 have been used mortality and time until death from a recent study published in *The Lancet Infectious Diseases* to test the quality of official case records. Their data shows that countries have only discovered on average about 6% of coronavirus infections and the true number of infected people worldwide may already have reached several tens of millions. It means the world moves fast toward the herd immunity.

Actually the number of true cases in the world is likely to be between one and two orders or magnitude higher than it is officially reported. This means many cities in the world are already in herd immunity threshold automatically and unwanted but we don't have any plan and simulation for them.

While in places and cities with high density people an contagion or infection spreads fast but reaches to herd immunity fast with low percentage of immune people. The reason is related to n (the potential number of susceptible neighbors). Networks with predominantly long-range connections initially, whereas networks with many local connections (such as lattice-based models plus small mesh networks) have far lower n . These local networks have reduced n for two reasons: firstly infectious individuals are aggregated so each infectious case has fewer susceptibles in its neighbourhood, and secondly, some susceptibles are shared between many sources of infection again reducing the potential n [12].

Incorporating, in a reasonable manner, inhomogeneity in susceptibility and infectivity in a standard epidemiological model, rather than assuming a homogeneous population, causes a very major reduction in the herd immunity threshold, and also in the ultimate infection level if the epidemic thereafter follows an unconstrained path.

Therefore, the number of fatalities involved in achieving herd immunity is much lower than it would otherwise be [13].

The great bulk of COVID-19 transmission is thought to occur directly from symptomatic and pre-symptomatic infected individuals, with little transmission from asymptomatic cases or from the environment [14]. There is strong evidence that a small proportion of individuals account for most infections – the 'superspreaders'.

According to *Clinical Infectious Diseases*, there is a smaller probability that non-immune individuals will come into contact with the virus if there is a greater proportion of immune individuals in a community.

Finally please note that Monte Carlo program used here is a huge program with many capabilities. Here we showed only small parts of its abilities as fundamental parts. Like Monte Carlo code used in nuclear reactors many scientists use that code for decades to simulate new factors and conditions, our program may be applied to many different cities, societies, countries, different circumstances, and calculation of many key factors related to COVID-19 or other epidemics.

References

1. Zhou P, Yang X, Wang X. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020; 579: 270–273.

2. Kroese DP, Brereton T, Taimre T, Botev ZI. Why the Monte Carlo method is so important today. *Wiley Int Rev Comput Stat*. 2014; 6: 386–392.
3. Anderson RM, May RM. Vaccination and herd immunity to infectious diseases. *Nature*. 1985; 318: 323-329.
4. Delamater PL, Street EJ, Leslie TF, Yang YT, Jacobsen KH. Complexity of the basic reproduction number (R0). *Emerg Infect Dis*. 2019; 25: 1-4.
5. Haley E, Randolph 1, Luis B. Barreiro. Herd Immunity: Understanding COVID-19, Immunity Cell Press. 2020; 52: 737-741.
6. Q Li, X Guan, P Wu, X Wang, L Zhou, Y Tong, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus- infected pneumonia. *N Engl J Med*. 2020; 382: 1199-1207.
7. Sanche S, Lin YT, Xu C, Romero- Severson N, Hengartner K. High contagiousness and rapid spread of severe acute respiratory syndrome coronavirus 2. *Emerg Infect Dis*. 2020; 26.
8. Wirth E, Szabó G, Czinkóczy A. "Measure Landscape Diversity with Logical Scout Agents". *ISPRS – International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. XLI-B2: 491–495. Bibcode. 2016.
9. Wirth E. Pi from agent border crossings by NetLogo package. Wolfram. 2015.
10. Pearson K. "The Problem of the Random Walk". *Nature*. 1905; 72: 294.
11. Kohls, Moritz, Hernandez, Tanja. "Expected Coverage of Random Walk Mobility Algorithm". 2016.
12. Matthew J Keeling. "The effects of local spatial structure on epidemiological invasions" *Proceedings of the Royal Society of London. Series B: Biological Sciences*. 1999; 266: 859-867.
13. Nic Lewis. Why herd immunity to COVID- 19 is reached much earlier than thought. 2020.
14. Ruiyun L, Sen P, Bin C. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV- 2), *Science*. 2020; 368.
15. M Ferguson N. Impact of Non- Pharmaceutical Interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. Imperial College COVID-19 Response Team Report. 2020; 9: 16.
16. Bommer C, Vollmer S. Average detection rate of SARS-CoV-2 infections is estimated around six percent, April. 2020.
17. Gomes MGM. Individual variation in susceptibility or exposure to SARS-CoV-2 lowers the herd immunity threshold. *Med Rxiv* 2 May. 2020.