

Understanding the dynamics of pathogenic infection in a population

Sangam Banerjee

Saha Institute of Nuclear Physics, 1/AF, Bidhannagar, Kolkata, 700064 India

author's email: sangam.banerjee@saha.ac.in

Abstract

In this article we have presented a new perception of herd immunity threshold (HIT) which considers that only a “band of population” are susceptible to any pathogenic infection. This is termed as the “effective herd immunity threshold” (EHIT) and the progression of the disease (caused by this pathogenic infection) is mainly determined by this EHIT value. We have argued here that this EHIT value (considering the immunity band picture in the population) will be substantially lower than the estimated canonical HIT values obtained from various existing models. We propose that the actual prediction of the disease progression should now be calculated using the EHIT values.

Keywords: Pathogen, Herd Immunity Threshold, R-naught, Infection dynamics.

It is now well known that development of herd immunity is key to control or eliminate communicable diseases that are air, droplet or fomite borne. Herd immunity is developed in a region only after a given threshold value of the population is infected by the pathogen (termed as herd immunity threshold - HIT). If the regions are densely populated, HIT value would be attained very rapidly as compared to regions with sparse populations.

HIT is related to the basic reproduction number R-naught (R_0) which is a measure of the transmission potential of a disease by a single infected individual to the susceptible population [1]. R_0 is proportional to and depends mainly on three factors such as (1) the rate of contacts (R_c) the infectious person makes in the host population, (2) the probability of infection (P_i) being transmitted during contact and (3) the duration of infectiousness (T_i), thus $R_0 = R_c P_i T_i$. R_0 being a number is calculated by various mathematical models. Two most popular models being used are SIR and SEIR, where “S” stands for Susceptible population with zero immunity, “I” stands for Infectiousness of the pathogen and “R” stands for people Removed (Recovered and attained immunity or died) and in SEIR model “E” stands for ‘people exposed’. Infectiousness of pathogen may change according to geographical, environmental and climatic conditions and so R_0 for any pathogen is not constant because of these factors and as well as time as the disease progresses. So, R_0 is not a biological constant for a pathogen nor a measure of severity of the disease. R_0 is rarely measured directly, it is calculated using various models as mentioned above. Thus, the modelled R_0 values are dependent on model structures and assumptions. Some R_0 values reported in the scientific literature are very likely obsolete. R_0 must be applied with great caution because this basic metric is very complex.

This reminds us of George Box's wise words "All models are wrong but some models are useful".

Immunity to any infection is the state of being non-susceptible to it, for the reason of being able to neutralize the pathogen when challenged by it. We also know that in real world we find groups of population having different susceptibility to infection due to varying innate immunity and this also varies geographically as suggested by us earlier [2, 3]. This tolerance to infection can be due to innate immunity acquired genetically [4-7], or as a result of immunity acquired by previous infections or by immunization. Therefore, not all contacts will become infected and the average number of secondary cases per infectious case will be lower than the basic reproduction number (R_0). To incorporate "only" the susceptible and exclude non-susceptible individuals, we can now introduce the "Effective Reproduction Number R_E ". This can be calculated by multiplying R_0 by the fraction of the population who are susceptible 's', leading to the equation:

$$R_E = sR_0 \quad (1)$$

Herd immunity occurs only if a critical fraction of population becomes immune. This results in breaking the chain of infection by the contagious disease. This results in providing protection to the remainder of the 'unprotected' community (herd) having weak immunity.

The herd immunity threshold value (V_{TH}) is the proportion of the population that needs to be immune by natural infection or immunisation in order for an infection to be brought under control. This is reached only when each new infected case leads to only a new single infection, i.e., $R_E = 1$. (We know that when $R_E > 1$ then the infection will grow exponentially.)

To find the value of V_{TH} we use Eq. (1) to get:

$$sR_0 = 1. \quad (2)$$

Considering that some fractions of population are susceptible (s) and the rest are immune (i) where the sum will be equal to 1 i.e.,

$$s + i = 1. \quad (3)$$

Rearranging this to get $s = 1 - i$ and substituting into Equation (2) and considering the number of immune reaching the threshold value i.e., $i = V_{TH}$ and hence:

$$(1 - V_{TH})R_0 = 1 \quad (4)$$

Thus, V_{TH} in terms of R_0 is:

$$V_{TH} = 1 - 1/R_0 \quad (5)$$

If i has surpassed the threshold V_{TH} value for herd immunity (i.e., $i > V_{TH}$) then $sR_0 < 1$, i.e., $R_E < 1$ and the number of cases of infection decreases (Note, s and i are fraction and hence takes value between 0 and 1). This is an important measure used in infectious disease control for immunisation programmes to eradicate the disease. For example, if R_0 is 2.5 in a population where one quarter of the population is immune and hence three quarter will be

susceptible ($s = 0.75$), the effective reproductive number R_E is then $2.5 \times 0.75 = 1.875$. Under these circumstances, a single case would produce an average of 1.875 new secondary cases. To successfully eliminate a disease from a population, R_E needs to be less than 1, which can be achieved by reducing the value of s (i.e., increasing the value of i). For R_E to be < 1 we require the value of s to be < 0.4 i.e., susceptible to be less than 40% of the population. In other words, the amount of immune population should be greater than 60% by either getting the population infected or by immunization.

The above-mentioned concept is well accepted in the epidemiology community. With the above background and understanding we can now move forward and add a new concept in epidemiology i.e., what if a given pathogen can infect only a section (stratum) of the population having weak or compromised immunity. This can be realized by introducing the concept of “level of immunity strength” and the population can be graded and scaled into different strata or bands as shown in Fig. 1(a).

Before discussing Fig.1 we shall discuss briefly on the degree of immunity among individuals. First and foremost, “Can one measure immunity?”. The immune system acts to protect the host from infectious pathogens that exist in the environment. Immune system is constantly active in discriminating “nonself” from “self.” The immune system consists of two functional arms: the innate and the adaptive (acquired). Both involve various blood-borne factors and cells. Recently several methods have been developed to assess and measure various aspects of immune function; these are carried out by ex vivo cell culture. Ex vivo measurements can determine the functional responses of specific immune cell types such as neutrophils, monocytes, T lymphocytes, B lymphocytes, natural killer cells etc... Even among healthy individuals there are large variations of the targeted immunity having varying immune functions. Several factors such as age, gender, nutritional status, gut flora and also mainly genetics determines the variation in immunity in an individual. Individuals with immune responses significantly below “normal” are more susceptible to infections. Thus, among individuals the variation in immune function causes variation in susceptibility to an infection. Moreover, there can be wide variation in immune response in a single individual too. Hence, we can accept that different pathogens may have very different dynamics in the same individual.

We know that to activate the adaptive arm of the immune system, the antigen-presenting cells presents the foreign antigens to antigen-specific T lymphocytes [8] and upon recognition - the cell activation, proliferation and cytotoxic activity takes place. Defects in any of these steps may lead to immune deficiency and the individual becomes susceptible to infections. For example: T-cell deficiencies can lead to suppression of antibody generation. Immunity can be evaluated through a number of routines, simple and complex laboratory test such as complete blood counts (CBC) and differential, serum immunoglobulin test, immunoregulatory T-cell defects test to name a few. With this elementary discussion on immunity in population toward any infection, we can now simply state that any population can be stratified and can be graded in an immunity level scale. Hence, here we can consider that a given pathogen can infect and cause a severe disease to only those individuals who have immunity below a certain threshold level.

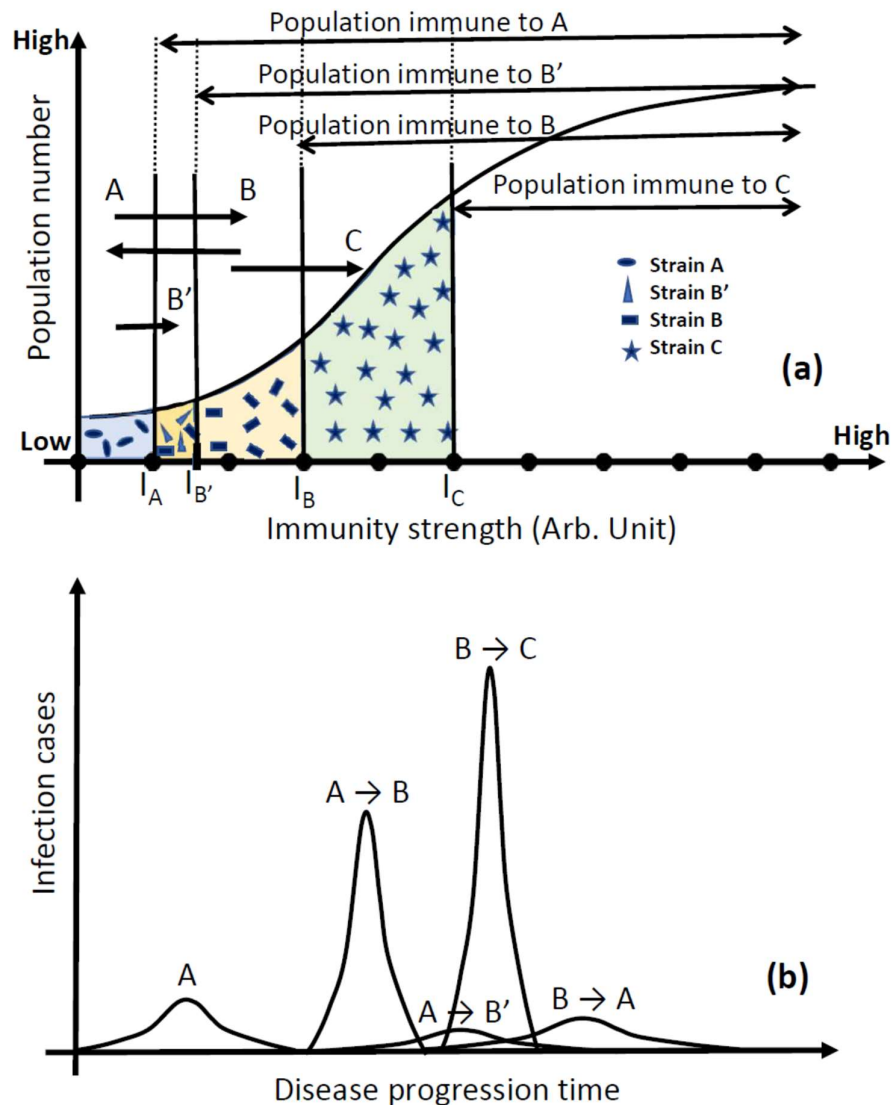


Fig.1 (a) Schematic of expected population number distribution vs immunity strength level. The single ended arrows indicate the jump of the pathogen between the strata and the double ended arrows indicate the portion of the population immune to the particular strain. (b) Schematic of the evolution of infection case as the disease progresses with time. The single ended arrows indicate the jump of the pathogen between the strata.

Now, in Fig. 1(a), we have shown schematically the expected population number distribution vs the immunity strength level. We scale the immunity strength level from low to high (from left to right) in the x-axis. Low level of immunity strength mainly comprises of older population, immunocompromised, comorbidity etc., and high level of immunity mainly comprises of healthy individuals (mainly young children). It should be mentioned here that

this is just a representation to visualise and this may not be the case always. We have shown the x-axis in an arbitrary unit (Arb. Unit) from some low value to high value as a representative. Now let us explain this figure below:

Let us consider a pathogen labelled as A and let us consider that this can infect only that portion of the population having very weak immunity (let us consider a hypothetical threshold value of immunity I_A). In this case this band will consist of very old and immunocompromised share of the population which will be the target of the pathogen for causing infection. It is interesting to point out here that in our earlier paper [2, 3] we had posed a question “Why the progression of the disease stops without reaching the herd immunity threshold value with respect to the total population”? Now we can reason out that the pathogen is not capable of infecting the whole population and consider all non-infected as susceptible, rather it can infect only some portion of the total population i.e., only those individuals who happen to fall in the band and have immunity below a cut-off threshold value. Rest of the population having immunity above this threshold value will remain immune to the pathogen (infection) and hence we had ascribed that population having high natural innate immunity or acquired immunity. This is the main reason for the flattening of the curve compared to if the total population is considered to be susceptible before the commencement of the infection/disease by the pathogen. Thus, in this model we consider that pathogen can only infect certain band of population having their level of immunity below a certain threshold value (I_A) and consists of portion P_A of the total population (P) as indicated in Fig. 1(a). Generally, one observes the progression of the disease to follow a normal distribution as shown in Fig. 1(b). The portion of population P_A is the area under the curve A in Fig. 1(b). The slope of the disease progression will depend directly on the effective reproduction number “ R_E ”. For higher value of R_E the slope of the progression of the disease will be higher. It is to be pointed out here that the thickness of the band will determine the density of susceptible population (Note: the density difference has been depicted in fig. 1(a) for different bands). Narrower the band, the density of susceptible will be low (susceptible individuals will be physically more distant apart because of the narrow range of susceptible population is considered) and for broader band the density of susceptible will be large among the total population (susceptible individuals will be physically closer because of the broad range of susceptible population is considered). Now R_C , P_I and T_I will depend on the band width and hence R_0 is determined by the thickness of the band and in turn the effective reproduction number R_E . It is now obvious that if the band is narrow then the peak height of the progression of the disease will be low because susceptible population involved are less in number and the curve (full width at high maximum) will be broader because the progression of the disease will be slower. Whereas, if the band width is large (broad) then the progression of the disease will rise rapidly and fall equally rapidly resulting in the disease progression curve to be narrower with a higher peak height (as shown in fig. 1(b)). This is so because the number of susceptible populations is larger in the broader band and the transmission will also be quicker. Thus, the width of the band is determined by the infectiousness of the pathogen which can infect only individuals having immunity below a certain threshold level of immunity. Now it may happen that due to immunity pressure from the population in band A, the pathogen evolves by mutation to a new favourable strain/variant B (antigenic drift) having

higher fitness and hence can now move/jump and encompass the next higher strata having a slightly higher immunity strength up to I_B in the total population. Thus, depending on the degree of improvement in the fitness (infectiousness) of the pathogen, some new portion of the population will fall prey to this new strain/variant. The width of the new band of population getting infected will depend on the infectiousness of the new strain/variant. If the band width is broad, then a greater number of susceptible (higher density of susceptible in the total population) leading to increase in the transmission of the new strain causing increase in the infection rate and the disease will spread rapidly. But if the infectiousness by the new variant is a very small incremental increase, then the band will be narrow (low density of susceptible in the total population) and transmission will be lower unlike what happens for the case of broader band. In Fig.1(a) we depict this scenario with variant B broader band and its respective disease progression in Fig.1(b) (see curve A \rightarrow B), similarly, we show for variant B' where the infectiousness increment is very small which will exhibit low peak height and slow progression of the disease resulting to wider slow progressive curve. If simultaneously two variants are present in the same population, then the disease progression curve will have the convoluted effect. As the disease progress with B and B' then now one can have two possibilities (1) either jump back to population P_A (curve B \rightarrow A) or (2) move forward by mutating to variant C (Curve B \rightarrow C) affecting individuals having immunity strength below I_C , its respective disease progression curve is shown in Fig.1(b). If it is the first possibility then the disease will be over soon in the population and if it is the second possibility then the disease will persist for a longer time.

Here we notice that the disease progression has been considered by considering only the population in the band is susceptible and not considering the whole non-infected of total population as susceptible. Thus, we can introduce a new concept to estimate the effective HIT (EHIT) value subject to a particular band only. We can clearly see here that EHIT value will be substantially lower than the presently used HIT value obtained from various models because now only a band of population is playing a role in the progression of the disease rather than the whole population. Thus, by introducing a band picture of immunity among the population as described in this article, the real value of EHIT can be obtained which generally will be much lower than the canonical HIT value.

Having discussed the dynamics of the pathogenic infection and the progression of the disease, now let us ask the question: "Can one predict the time of occurrence of peak, peak height and the duration of the disease?". Here we would like to point out that it is possible to roughly guess these only after the occurrence of the first peak of the disease in the population. As mentioned above, we know that in reality we find groups of population having different susceptibility due to different innate and acquired immunity to infection and this varies geographically. The progression of the first peak gives us a rough idea on the nature of population responding to a particular disease. The slope of the infection rate and the death rate of the subsequent wave can be compared with the first wave and we can find a rough estimate of the peak height, the time range of the occurrence of peak and the duration of the subsequent wave can be estimated along with its EHIT value. This we shall discuss in a separate article.

To summarise, we have introduced a new concept to estimate the effective HIT (EHIT) value like we have considered the effective reproduction rate R_E instead of reproduction rate R_0 . We have argued here that EHIT value will be substantially lower than that presently being used and estimated HIT value obtained from various modelling. By introducing the parameters mentioned in this article the true real value of EHIT can be obtained which generally will be much lower than canonical HIT value. This is an important concept and will help in understanding the disease progression more realistically.

Ethical approval: This study did not require ethical approval

Conflict of interest: None

Declarations of interest: None

References:

- [1] Paul L. Delamater et. al., "Complexity of the Basic Reproduction Number (R_0)", *Emerging Infectious Diseases* Vol. 25, No. 1, pg 1 (2019). Rothman KJ, Lash T, Greenland S. *Modern Epidemiology* (3rd ed.), Lippincott Williams & Wilkins, 2013.
- [2] S. Banerjee, "Can innate immunity flatten the curve?", *Preprints (www.preprints.org)* Posted: 19 April 2020, doi:10.20944/preprints202004.0353.v1 (2020).
- [3] S. Banerjee, "Pandemic or Endemic", *Preprints*, 31 May 2020, doi:10.20944/preprints202005.0515.v1 (2020).
- [4] Scott Alper, Rebecca Laws, Brad Lackford, Windy A. Boyd, Paul Dunlap, Jonathan H. Freedman, and David A. Schwartz, "Identification of innate immunity genes and pathways using a comparative genomics approach", *PNAS*, 7016–7021 105(19) (2008).
- [5] Massimo Mangino, Mario Roederer, Margaret H. Beddall, Frank O. Nestle, and Tim D. Spector, "Innate and adaptive immune traits are differentially affected by genetic and environmental factors", *Nature Communications* |8:13850| (2017).
- [6] Z Du, SK Sharma, S Spellman, EF Reed and R Rajalingam, "KIR2DL5 alleles mark certain combination of activating KIR genes", *Genes and Immunity*, 470–480 9 (2008).
- [7] JL. McCarville and JS Ayres "Disease tolerance: concept and mechanisms", *Current Opinion in Immunology* 88–93, 50 (2018).
- [8] S. Banerjee, "Reviewing Innate vs Acquired", *OSF preprint: 2021-04-08*, DOI 10.17605/OSF.IO/AMUJK (2021).