Predicting Dynamics of an Infectious Disease

Tesfaye Demissie

Data Science Practitioner
Bristol, UK
Email: tesfad {at} hotmail.com

ABSTRACT: - ***In a broader sense, disease is either infectious or not infectious. Such a classification purely rely weather a disease can be passed or transmitted between individuals or not. Infectious diseases are those diseases that can spread or transmitted easily and influenza is an example. Non- infectious diseases seem to develop over an individual's lifespan and are not transmittable as such. Disease such as arthritis is an example of non-infectious disease. Classification of diseases is not an easy touch and in principle we have to admit to some overlapping situations. Some infectious diseases although not necessary for developing can be a cause for non-infectious disease such as cancer. There is however a clear distinction in carrying out epidemiological study for each disease type. The epidemiologic interest in studying non-infectious disease lies primarily on identifying the risk factors that are prominent in the chance of developing a particular disease. This could for example be learning the risk of contracting lung cancer attributed to smoking. In studying infectious disease on contrary, the epidemiologic interest would be identifying the infectious cause, origin and pattern of the disease in a population of interest. A study of disease outbreak in a local population where the primary risk factor for catching an infectious disease is the presence of an outbreak itself is example in this case [1]. This paper focuses on learning the great predictive power of models for infectious disease at population scale over the short period of time. The paper is a build-up on the same idea to examine infectious disease transmission mechanisms from host to another caused by micro parasitic pathogens. The paper does not cover cases where a disease is transmitted by macro parasite intermediate vectors enter the dynamic of the model. The examination of the model is based on constant and susceptible healthy, infected and immune population without no vital (birth and death) statistics.***

Keywords— Infectious Disease, Epidemiology, SIR Model

1. INTRODUCTION

Public health study in its generality covers two broader areas: Epidemiology and Clinical Trials. Epidemiology is a systematic study of using observational data collected from a study population that are not under the influence of experimental settings to learn about disease cause and origin (ethology). It is a science of multidisciplinary in nature. It encompasses disciplines such as clinical epidemiology, behavioral epidemiology, occupational epidemiology, chronic disease epidemiology, infectious disease epidemiology, and environmental epidemiology. In this regard, my intention in writing this paper is to shine a light in the dynamics of infectious epidemiology through practical example. Nelder and Wedderburn argued that such a study could for instance be carried on to learn the casual relationship between smoking and lung cancer, air pollution and respiratory illness, heart disease and diet, childhood leukemia and water contamination, and investigating the prevalence and incidence of HIV infection and AIDS, etc. [14] [18]. Its function is mainly aimed at improving overall health of the population.

Clinical Trials on the other hand are specifically designed in a controlled experimental settings to evaluate specific type of medical treatment or intervention. Examples of Clinical trial study could include comparing the effect of applying HIV drug versus placebo on patients survival length who contracted AIDS, learning the effectiveness of new drug on athletes foot fungus development, evaluating hormonal remedy on the lessening of breast cancer, etc.

Modelling is an act of scientific investigation that allows precise, rigorous analysis and quantitative prediction without claiming complete certainty. It is about expressing ideas mathematically to clarify thinking. Cliff & Murray and Spicer have discussed that modelling the dynamics of infectious disease can have a direct bearing on the choice of curious measures, optimal allocation of resources and deployment of best medical intervention techniques [10] [20]. The scientific field epidemiology has entered its exiting time. This is happening not by chance. It is happening due to the high demanding nature of public health study for expertise in epidemiology and its advanced methods. As argued in [3] by Black, epidemiologic methods are capable of handling sophisticated ways of evaluating public health risk indicators that result from many exposure and environmental pollutants of our modern society. Factors and enablers for

epidemiologic methods are being emerged as powerful as never seen before. The 21st century information technology advances including super powered microcomputers, the Internet, software developments, and the exhilarating prospects paved the path to execute wider array of studies. The way health care is delivered today, in particular the emergence and grow of organized health care system in a digital world has created chances and opportunities for epidemiology and epidemiologists to shine and involve in evidence-based public health and the valuation of health care operation and excellence. Decisions made and policies formed in public health exploration without sound epidemiologic data analysis and reasoning are becoming the things of the past. The public and medical practitioners will benefit from the awareness brought as a result of evidence based epidemiologic study in regard to important diseases [9] [17]. Knowledge about a disease incidence and prevalence rates, its morbidity and mortality rates, its importance (e.g. population-attributable risk fraction or the global burden of disease), time (trends – whether the incidence is rising or falling), place (whether there are areas where the disease is particularly common or rare), person (the type of person who is most at risk, with regard to demographics, lifestyle, health status and workplace), and prevention (primary, secondary and tertiary) are all contribute to enhance the wellbeing of the society [8] [12].

Epidemiological Modelling has its root at early 18th century when Daniel Bernoulli designed a model to investigate the effectiveness of inoculating healthy people against the smallpox virus in 1760 [26]. Hamer also studied recurrence of measles epidemics in 1906 and carried analysis on discrete time model that he formulated [7]. The mathematical epidemiologic model that almost won universal acceptance was developed by Ronald Ross in 1911 [22] where he developed differential equation models for malaria (Ronald Ross, 1857-1932). Since then mathematical models were developed in 1927 by Kermack and Mckendrick as an extension of Ross's model and epidemic threshold results were derived [26]. The study made by Ross was to indicate that a disease can go extinct not only by eliminating all the pathogen carrier insects but also by satisfying certain conditions. Hethcote, 1976 and Fred, 2008 have given detailed discussion of such a model, that a model without vital dynamics is termed as SIR model [25] [27].

Studies in the past have demonstrated the requirement to establish stronger link between traditional epidemiology that focus only on methods of determining disease etiologic such as study design, source of bias, and casual reasoning; and applied epidemiology that synthesizes and applies the results of etiologic studies to set priorities for intervention, evaluates public health interventions and policies, measures the quality and outcome of medical care, and effectively communicates epidemiologic findings to health professionals and the public is paramount [2] [15] [21].

Standard World Health Organization guideline [4] highlights that the core principle in conducting an epidemiological study rests around three basic values. These are case finding, increase public health disease knowledge and identification of important diseases.

Case finding is a strategy for targeting resources at individuals or groups who are suspected to be at particular risk of a disease. It involves actively searching systematically for high-risk people, rather than waiting for them to present themselves to medical attention after symptoms or signs of active disease have occurred. Note the similarities between case finding and screening: both seek to risk stratify the population using a simple and cheap procedure, and assume that better outcomes can be achieved through identifying the early stages of disease and offering prompt treatment. As an example, case finding may be used as part of the investigations into an outbreak of a communicable disease (e.g. syphilis) to identify potential sources of the disease. It may also be employed during food-borne outbreaks to identify as many at-risk individuals as possible. The advantage of case finding include that it is cheap and incur low personnel demand, case finding improves the positive predictive value of a diagnostic test by targeting high-risk patients with higher underlying prevalence. By targeting preventive care, case-finding tools can help improve care of individuals and reduce costs for the state. The prime disadvantage might be the presence of a potential to widen health inequalities because some high-risk groups are hard to reach (homeless, refugees, etc.)

Knowledge in epidemiologic study refers defining clinical features, distribution, causes, behavioral features and determinants of diseases that currently make a significant impact on the health of local populations, with particular reference to those that are potentially preventable, or require the planned provision of health services at individual, community and structural levels, or are otherwise of particular public concern, e.g. mental health. The World Health Organization's global burden of disease project provides an estimate of the relative importance of all communicable and non-communicable diseases, together with intentional harms (e.g. suicide and war). The global burden of disease does not account for the degree to which illnesses are preventable or can be treated, but it does provide a useful guide to which illnesses have the greatest impact globally – and are thus of public health importance [24].

1.1 Infectious Disease Characterization

The speed and development of an infectious disease can be qualitatively defined in terms of the causes of the disease. The causes of an infectious disease are either microscopic or macroscopic pathogens that are potentially capable of replicating themselves and invade human body tissues; further producing toxins to poison the cells. The interaction of these pathogens and their growth rate within human body and the human body's immune response are vital to determine the progress of an infectious disease. The conclusion made from the study by [6] and [16] is that

understanding the whole process is the basic principle in infectious disease epidemiology and gaining an insight as to how particular interventions at different stages could prevent or control the disease spread.

A disease occurs when infectious pathogen finds its way and enters human's body through what is known as route of entry. Potential routes of entry for successful disease transmission are respiratory tract, gastrointestinal tract and skin. Infectious pathogen such as mycobacterium tuberculosis enters into human body through air breathed into the lungs. Pathogens that cause diarrhea for instance enters into human body through contaminated food and water taken by mouth, or unhygienic hands. Naturally human skin is capable of serving as a barrier against many infectious pathogens but in some cases such as malaria parasites, infectious pathogens can enter into human body when infected mosquito bites through the skin to suck blood.

At an initial stage the host becomes susceptible to infection. This is the stage where there is no pathogen in anyone's system at all, but a low-level unidentifiable and suspicious host immunity exist. Example of a person entering this stage could include for example a person shaking hands with someone suffering from a common cold, a child living in the same room as an adult with tuberculosis.

Then the host gets an exposure for infection. A parasite duplicates and grows over time and enter the host, but host might not exhibit any clear sign of infection and the number of pathogens might be small to cause further transmission. This stage puts an individuals at the exposed stage. The exposure is the stage immediate after infectious pathogen enters and assume multiplying. Example is when a person has consumed food that has been contaminated with bacteria causing typhoid fever (Salmonella typhii), it is said to be exposed. But when the bacteria gets to the lining of the intestine and started multiplying, the person is said to have entered infected stage. However, there might not necessarily be clinical manifestation of the disease at this stage. The clinical manifestation happened when there is a match between the disease symptoms (complaints of a person such as headache, vomiting, dizziness, etc.) and disease signs (features like high temperature, high pulse rate, swelling of organs inside the body) that can only be detected by trained health professional. Once at this stage the pathogens will become abundant enough to spread themselves and gain the potential to transmit to another susceptible individual and the disease enters its infectious stage. Infected people can be carriers but not infectious themselves. If they are infectious, they are termed as active cases. After the clearing of the pathogens from diseased individuals and host gets cleared of its infectious stage, the individuals enters recovered stage. Recovered stage is a general term to infer complete recovery from disease, being disabled or dead.

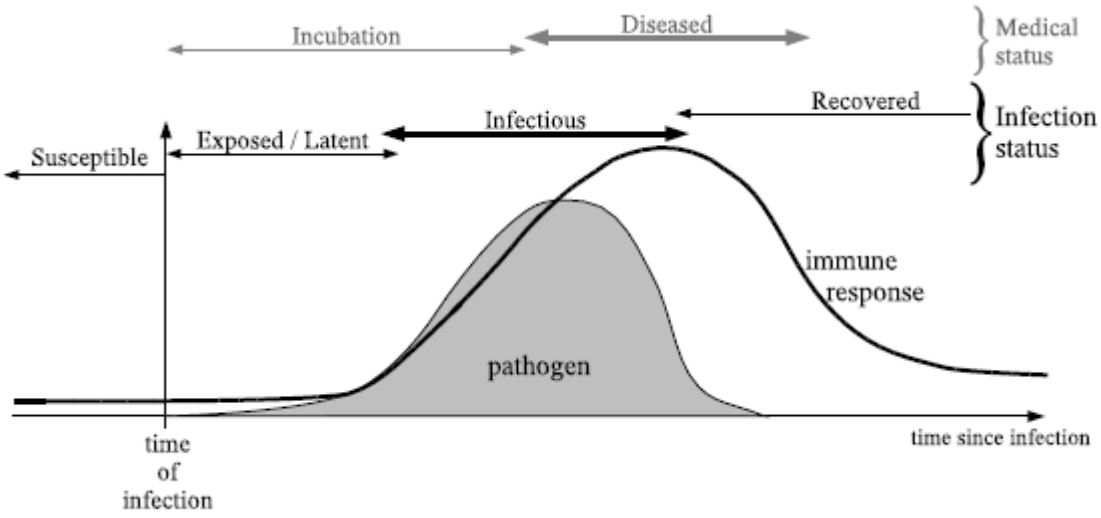

Figure 1: Infectious disease classification

This ultimate infectious disease classification (as susceptible, exposed, infectious, or recovered) exclusively depends on the disease's ability (host in this case) to pass or transmit the pathogen. The takeaways here are that the host's status regarding the disease is irrelevant, that is an individual who actually has a perfect healthy feeling with no symptoms can be releasing huge amount of pathogen; and boundaries between exposed and infectious (and

infectious and recovered) is somehow grey and the tendency to transmit is not as simple as turning buttons on and off. This is an addition to the complicated nature of infectious disease in understanding the variability in response to disease of individuals and level of pathogens over the infection period. Important to note that diseased period, when symptoms are experienced, is not necessarily correlated with any particular infection stage.

2. METHOD

The study from [11] [13] [19] [23] illustrates that achieving an iconic objective of an epidemiologic study requires properly crafted public health study method that:

a. Discovers the cause, origin and environmental factors which has an impact on health so as to provide the scientific basis for the prevention of disease and injury and the promotion of health.

b. Determines the relative importance of causes of illness, disability, and death, in order to establish priorities for research and action.

c. Identifies those sections of the population which have the greatest risk from specific causes of ill health, in order that the indicated action may be directed appropriately.

d. Evaluates the effectiveness of health programs and services in improving the health of the population.

The people in epidemiological study are mainly interested in finding the features that are prominent in determining the pattern of the disease and its way of transmission or spread.

The assumption is that we have a constant population, *N*, and that the population is divided into the three states: susceptible *S*, infected *I*, and recovered or immune *R*. Most specifically, the model covers the simplest form of epidemic *SIR* model.

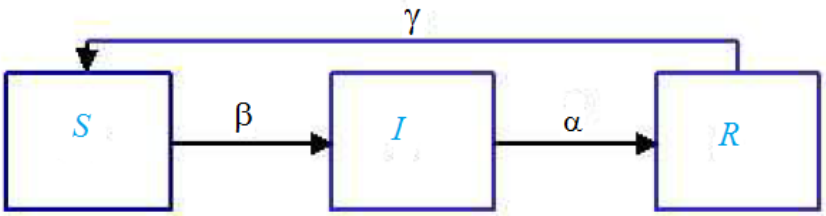

Figure 2: Epidemic *SIR* Model

The first group are the individuals who are capable of becoming infected with a particular disease. The second group consists of individuals who are infected and can infect others. Sometimes these models include a class of exposed individuals, *E*, who are infected but cannot yet pass along the disease. Finally, the class *R* represents those who have recovered from the disease and are immune to infection. Most viral diseases, such as measles or chickenpox, cause the body to mount an immune response [5]. Once the body sees a particular disease, then a future infection is highly unlikely. After a host becomes infected, then they develop a permanent immunity to the disease, *R*.

2.1 Modelling considerations

Modelling an epidemic should take into consideration factors such as population structure and demography (stratification by age, sex, location, etc.), natural history of the infection (latency, infectious period, immunity, etc.), and intervention (at what stage of disease transmission).

2.1.1 Transmission rate

Consider an individual susceptible to disease:

- Rate of contracting other individual ‘c’ is the contact rate that applies to all individuals irrespective of infection status.

- Transmission requires contact with infected individuals and rate of contacting infectious individuals is 'cI/N', where I/N is proportion of infectious population, I is no of infected, and N is total population.
- Rate of transmission from infectious individuals is given by 'pcI/N' usually termed as force of infection, where p is the probability of transmission when an infectious individual contacts a susceptible.
- If we consider all susceptible individuals, the total transmission rate in population is pcSI/N, where S is the number of susceptible individuals. Most often, 'pc' is written as 'β'.

2.2 Simulation of Epidemic model (SIR)

A derivative approach to calculate time derivatives of *S, I* and *R* is implemented. Given a value of S, I and R at time t, the derivative calculates the time derivatives of *S, I* and *R*; and parameters of the model like the recovery period and the transmission rate.

The population size, N is always *S+I+R* because there are no births or deaths in the model.

$$dS/dt = -\beta SI/N + \gamma R,$$
$$dI/dt = \beta SI/N - \alpha I,$$
$$dR/dt = \alpha I - \gamma R$$

Like many processes associated with living organisms, the spread of a disease caused by a microorganism through a population can be modelled mathematically using differential equations. Although numerous models of varying complexity have been developed to describe the dynamics of disease spread in a population, the SIR model presented here combines relative simplicity with good modelling of diseases that are spread from person-to-person and are familiar to public, such as measles, smallpox, and influenza.

In the SIR model, members of a population are categorized into one of three groups: those who are susceptible to being infected, those who have been infected and are able to spread the disease to susceptible individuals, and those who have recovered from the disease and are immune to subsequent re-infection. Movement of individuals is one-way only, $S \rightarrow I \rightarrow R$ and the two fundamental parameters of the model, α (the daily infection rate) and β (the recovery rate), act as rate constants that control how fast members progress into the *I* and R groups, respectively. A composite parameter, $\gamma = \alpha/\beta$ is often used and is referred to as the contact number. The SIR model is described by the differential equations.

Solving such an equation is difficult algebraically and therefore integration technique is used. Doing so is used to see the change in the different rates at each stage of the model over time. In differentiating an equation, the derivatives indicate how the slopes (changes in rate) relate to the model at any point in time.

Initially, S(0) = 1.

$$dI/dt = \beta si - \alpha i = (\frac{\beta s}{\alpha} - 1)\alpha i, \quad i = I/N,\ s = S/N$$

Now, an epidemic occurs if the number of infected increases.

$$dI/dt > 0.$$

This is true when $\beta/\alpha > 1$.

On the contrary, the disease dies out if the number of infected decreases.

$$dI/dt < 0$$

This is true when $\beta/\alpha < 1$.

$\beta/\alpha = R_0$ is the base reproduction number. It is the mean number of secondary infections generated by single infected case in a completely susceptible population.

When initial conditions for these groups are specified, the change in size of these groups may be plotted over time.

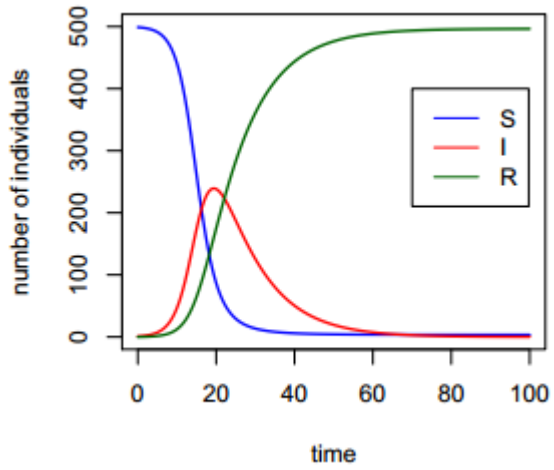

Figure 3: Anatomy of Epidemic disease

3. SIMULATION RESULTS

Whether an epidemic will ensue under certain initial conditions can now be discussed in terms of the contact number, and we may reasonably be expected to empirically determine that the transition between epidemic and non-epidemic states occurs when the initial fraction of the population in the susceptible group (lower line) is equal to the reciprocal of the infected number (upper line) (Fig 4). The recovery rate,β can also be indirectly introduced as the more accessible duration of the disease $1/\beta$.

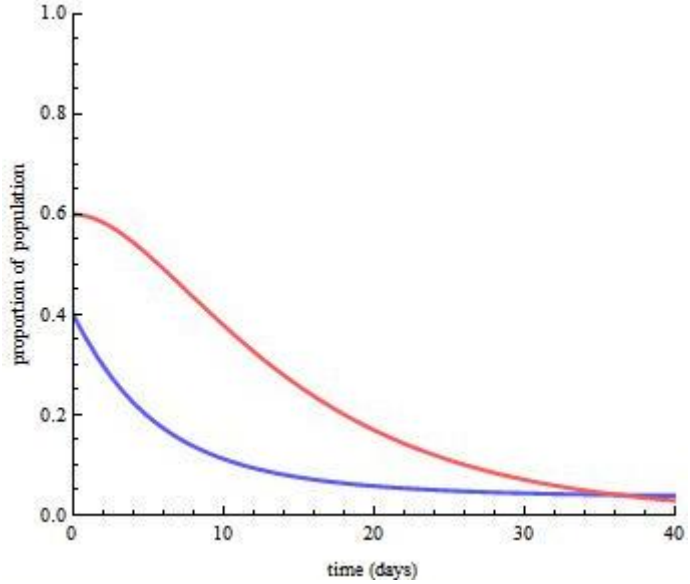

Figure 4

By discussing epidemic dynamics in terms of these more easily understandable parameters and allowing *R* to convert to the actual model parameters behind the scenes, it is possible for discussions of an important topic to be tailored to the public. The dynamic nature of the output also facilitates discussions of the effect of different parameters on the nature of disease spread in a population without necessarily resorting to the equations governing the

model. In particular, the importance of infected number (Figures 5a and 5b) and the effect of artificially moving members of the population directly from the susceptible group to the recovered (and therefore immune) group through immunizations (Figure 6) can be easily investigated by manipulating the appropriate rates of the model.

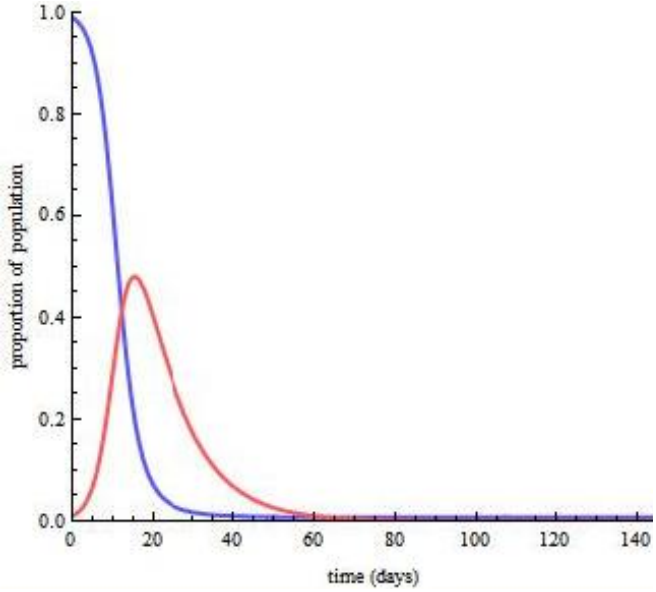

Figure 5a

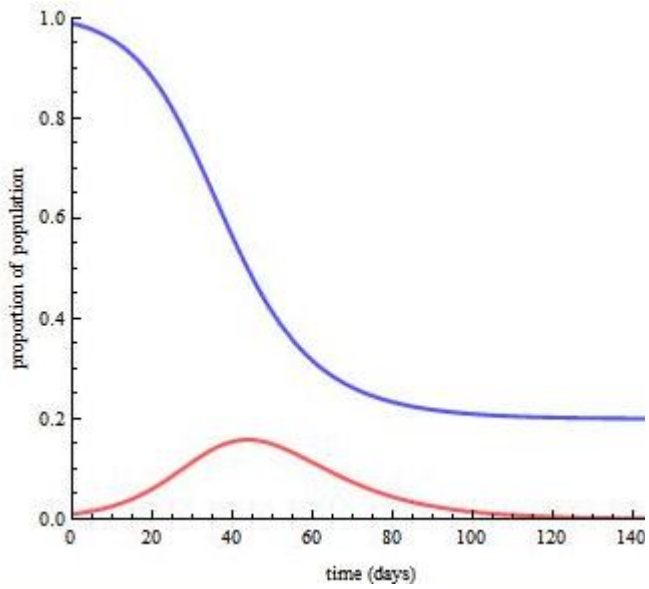

Figure 5b

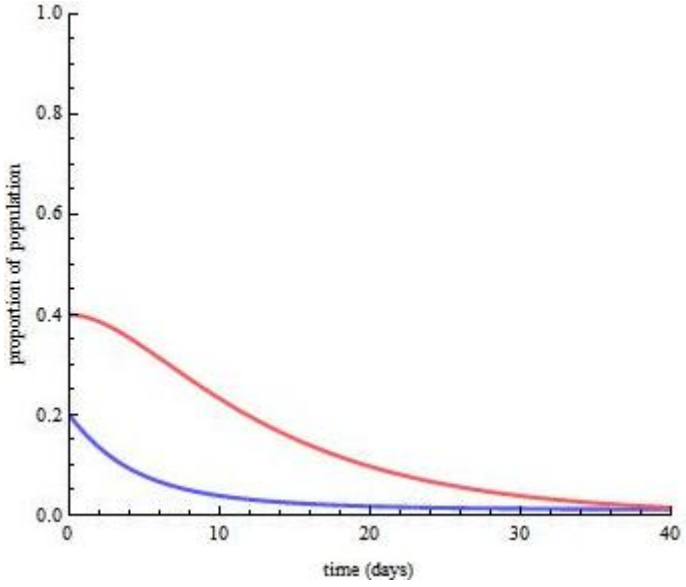

Figure 6

4. DISCUSSION

Careful inspection of SIR model will reveal insights into the dynamics of the disease in a population. For example, if the fraction of the population in the infected group is initially increasing (i.e., $dI/dt > 0$ at $t = 0$), it means an epidemic has begun. The transition between an epidemic and a non-epidemic spread of a disease then occurs when $dI/dt = 0$, and inspection of the differential equations will quickly reveal that this transition point results when $s_o = \beta/\alpha$. Likewise, the peak of an epidemic occurs when $s = \beta/\alpha$ and the rate of change of the infected group stops increasing and

starts decreasing. The contact number also has an easily understood "real-world" interpretation: the average number of susceptible members of the population. An infected individual spreads the disease while that individual is in the infected group. The anatomy of an epidemic is such that initially the number of infections will not be extreme and are small that follow stochastic nature. The infection then start to experience increased prevalence and at increased speed. As infection depletes the number of susceptible, the spreading rate declines through time.

4.1 Limitations

The classic SIR models presented here assumes that the total population size remains constant and the population is uniform and homogeneously mixing. Mixing depends on many factors including age, sex, geographical location, etc. Different geographic and social-economic groups have different contact rates. Also the models ignores random effects, which can be very important when *s* or *i* are small.

5. CONCLUSIONS

In efforts to control the spread of the disease, we must select the optimal solution for the maximum public health benefits. Mathematical models can help us to better understand the spread of an infectious disease and to test the control strategies. In this paper, the epidemic problem can be solved by using SIR model and through R statistical package program and simulating the epidemic problem. Different deterministic models can be constructed by choosing different number and types of epidemic models. The approach of the analysis is based on theory of dynamical systems. It is reasonably enough to justify the modelling approach clarifies what the underlying assumptions are. For optimum results model analysis and simulation predictions suggest crucial data that should be gathered and control strategies that could be implemented. Estimates of R_0 for various diseases are useful for comparing diseases. If $R_0 > 1$, an epidemic is prevented when $R_0 S(0) < 1$. Thus, if the initial susceptible fraction has been reduced to less than $1/R_0$, for example by immunization, then an epidemic can be prevented.

6. REFERENCES

[1]. Bailey Ntj. The mathematical theory of infectious diseases. London: Griffin, 1975.

[2]. Anderson RM, May RM. Infectious diseases of humans. Oxford: Oxford University Press, 1991.

[3]. Black FL Measles. In: Evans AS ed. viral infections of humans: epidemiology and control. New York: Plenum Medical, 1984: 397.

[4]. World Health Organization. Epidemiological and vital statistics report. 1952; 5: 332.

[5]. Chapin CV. Measles in Providence, R.I. American Journal of Hygiene 1925; 5: 635-55.

[6]. Evans AS. Epidemiological concepts and methods. In: Evans AS ed. viral infections of humans: epidemiology and control. New York: Plenum Medical, 1984: 20.

[7]. Hamer. WH Epidemic diseases in England. Lancet 1906; i: 733-39.

[8]. Soper. HE Interpretation of periodicity in disease prevalence. Journal of the Royal Statistical Society A 1929; 92: 34-73.

[9]. Greenwood M. On the statistical measure of infectiousness. Journal of Hygiene 1931-31: 336-51.

[10]. Cliff AD, Murray GD. A stochastic model for measles epidemics in a multi-region setting. Institute of British Geographers, Transactions New Series 1977; 2: 158-74.

[11]. Fox JP. Herd immunity and measles. Reviews of Infectious Diseases 1983; 5: 463-66.

[12]. Bailey Njt. The elements of stochastic processes. New York: Wiley, 1964: 183.

[13]. Bartlett MS. Deterministic and stochastic models for recurrent epidemics. Proceedings of the Third Berkeley Symposium on Mathematical Statistics and probability, 1956; 4:81-108.

[14]. Nelder JA, Wedderburn Rwm. Generalised linear models. Journal of the Royal Statistical Society, 1972; 135: 370-84.

[15]. Semple AB. Epidemiology of the influenza epidemic in Liverpool in 1950-51.Proceedings of the Royal Society of Medicine 1951; 44: 794-96.

[16]. Kilbourne ED. The molecular epidemiology of influenza. Journal of Infectious Diseases1973; 127: 478-87.

[17]. Fox JP, et al. Herd immunity: basic concept and relevance to public health immunization practices. American Journal of Epidemiology 1971; 94: 179-89.

[18]. Elveback LR, et al. An influenza simulation model for immunization studies. American Journal of Epidemiology 1976; 103: 152-65.

[19]. Longini I., et al. An optimization model for influenza A epidemics. Mathematical Biosciences 1978; 38: 141-57.

[20]. Spicer CC. Mathematical modelling of influenza epidemics. British Medical Bulletin1979; 35: 23-28.

[21]. Rvachev LA, Longini IM. A mathematical model for the global spread of influenza. Mathematical Biosciences 1985; 75: 3-22.

[22]. Ross. R (1911). The Prevention of Malaria. London: John Murray. 651-686p.

[23]. Lyle D. Broemeling (2014). Bayesian Methods in Epidemiology. Taylor & Francis Group, LLC.

[24]. Alan D. Lopez, et al. (2006). Global Burden of Disease and Risk Factors. The International Bank for Reconstruction and Development / The World Bank

[25]. Fred Brauer, et al. (2008). Mathematical Epidemiology, Springer.

[26]. Nicolas Bacaër (2011). A Short History of Population Dynamics, Springer-Verlag London Limited.

[27]. Herbert W. Hethcote (1978). An immunization model for a Heterogeneous Population, Academic press, Vol. 14, No 3, p338 – 349.