Spring 2015

An Examination of Mathematical Models for Infectious Disease

David M. Jenkins
dmj40@zips.uakron.edu

An Examination of Mathematical Models for Infectious Disease

David Jenkins

Abstract:

Starting with the original 1926 formulation of the SIR (Susceptible-Infected-Removed) model for infectious diseases, mathematical epidemiology continued to grow. Many extensions such as the SEIR, MSIR, and MSEIR models were developed using SIR as a basis to model diseases in a variety of circumstances. By taking the original SIR model, and reducing the system of three first-order equations to a single first-order equation, analysis shows that the model predicts two possible situations. This analysis is followed by discussion of an alternative use of the SIR model which allows for one to track the amount of sustainable genetic variation in a population of pathogens as a quantity π. Finally, using the 2014 West African Ebola epidemic as an example, an alternative model to SIR is presented and discussed.

Table of Contents

§1 Introduction

Infectious disease is a major source of strife for developing countries (less so of late for post-modern nations), so it is not surprising that much effort has been placed in attempts to soften the suffering they bring. Be they in the form of vaccinations, antibiotics, or improved health initiatives, all attempts at alleviating the effects of these diseases require accurate data and projections about their epidemiology. In the 1960's much success was achieved in managing these diseases to the point where many were largely thought to be eradicated [1]. However, with the emergence of the human immunodeficiency virus (HIV) in the 1980's, as well as various strains of hepatitis in the 1990's, a new interest has emerged in epidemiology, which is the study of patterns, causes, and effects of diseases in defined populations [7]. Mathematical epidemiology seeks to predict an infection's behavior via the use of mathematical models in order to provide insight into how to prevent undesirable outcomes. Mathematical models of how diseases spread, while having been used since the eighteenth century [1], became a primary tool for analyzing how diseases like HIV spread, and suggesting how to control them. While certainly inspired by infections in human populations, many of these models (including the classic Susceptible-Infected-Removed model and its derivatives) are also employed by ecologists to study non-human population dynamics, assuming the disease does not behave anomalously in the population of interest [7].

Mathematical models for infectious diseases are not a recent development in epidemiology. As early as the eighteenth century Daniel Bernoulli developed a model to predict the effects of inoculating smallpox sufferers with infected material from other infected individuals; however, disease modeling did not really start to take off until the twentieth century [1]. In the early twentieth century, various models for diseases such as measles and malaria were developed, and in 1926 the Susceptible-Infected-Removed (sometimes called Susceptible-Infected-Resistant or SIR) model was developed by Kermack

and McKendrik. The SIR model introduced the idea that there exists a threshold density of susceptibles which must be surpassed in order for an infection to take hold and a true epidemic to occur [1].

Many variations of SIR have been developed since its introduction in 1926. One variation is the Susceptible-Exposed-Infected-Removed (SEIR) model, which includes a latency period where the host is infected, but not yet infectious. The MSIR model assumes that resistant mothers confer some of their immunity to their children, so the model contains a M category for maternally resistant susceptibles. The MSEIR model includes both maternally derived immunity as well as a latency period [1]. The choice of which "compartments" to include in a model widely depends on the situation and pathogen being modeled. In fact, models other than SIR derivatives also vary and different ones work better for different problems. Some models depend on very accurate estimates of the parameters in the model and require extensive analysis of preexisting data before they can be useful. Other models may not provide the most sophisticated or complex predictions, but require little data to parametrize effectively. Such models may be useful in new or not well understood situations.

New mathematical models of infectious disease are still being developed. New strains of influenza have evolved, alongside antibiotic-resistant strains of pneumonia, tuberculosis, and gonorrhea [1]. Mathematical disease models have been modified to account for and analyze the recent evolution of pathogens as understanding of evolutionary theory has matured [3, 4]. With changes in human and animal habitation, climate, environmental degradation, travel patterns, and even socioeconomic conditions, pathogens evolve to exploit emerging niches. Models that describe the infectious diseases associated with such pathogens must also be further developed to keep up lest these models lose their efficacy [1]. In the following sections an analysis of the basic SIR model will be provided, followed by a presentation of an alternative use of SIR, as well as an alternative model to SIR.

§2 The Basic SIR Model and Epidemics

The SIR model (Susceptible-Infected-Removed) can be presented in a variety of ways; a simple version is presented by Strogatz [2] as

(1)
$$\dot{x} = \frac{dx}{dt} = -kxy,$$

(2)
$$\dot{y} = \frac{dy}{dt} = kxy - ly,$$

(3)
$$\dot{z} = \frac{dz}{dt} = ly.$$

Here k and l are positive constants, $x(t)$ denotes the number of susceptible individuals in the population at time t, $y(t)$ denotes the number of infected individuals at time t, and $z(t)$ denotes the number of dead individuals at time t. Note that $z(t)$ could also denote the number of recovered individuals, in which case the model assumes that recovery confers total immunity to the individual with no chance of reinfection. It is assumed that $z(0) = 0.$ The model as it is presented here assumes that healthy individuals become infected at a rate proportional to the product of x and y, thereby assuming that healthy and sick individuals come in contact at a rate proportional to their numbers with a constant probability k that an infected individual will infect a healthy individual. It is also assumed that infected individuals die at a constant rate $l.$ The model also assumes that the total population N is fixed with no entries or departures due to births or migration. There is no latency period in this model, meaning that a host is immediately able to infect others upon being infected. This system of three first-order equations can easily be reduced to a single first-order equation in the following manner.

First note that

$$\dot{x}+\dot{y}+\dot{z}=\frac{d}{dt}(x+y+z)=0.$$

Thus

(4)
$$x+y+z=N,$$

where N is a constant representing the total population size. Note that, $\dot{z}=ly>0.$ This implies z is increasing and 1-1 which further implies $z(t)$ is invertible and has an inverse denoted $t(z).$ Since $t(z)$ exists, then the composition

$$\hat{x}(z)=x(t(z))$$

is valid and by the chain rule one obtains

$$\frac{d\hat{x}}{dz}=\frac{dx}{dt}\frac{dt}{dz}=\frac{dx}{dt}\frac{\frac{1}{dz}}{dt}=\frac{-kxy}{ly}=\frac{-kx(t(z))}{l}=\frac{-k\hat{x}}{l}.$$

Therefore $\hat{x}$ satisfies

$$\frac{dx}{dz}=\frac{-k}{l}x.$$

Separation of variables then yields

$$\frac{1}{x}dx=\frac{-k}{l}dz.$$

So,

(5)
$$\hat{x}=x_0e^{\frac{-k}{l}z}$$

where x_0 is the initial susceptible population. From (4) one can rewrite the equation for $\dot{z}$ as

$$\dot{z}=ly=l(N-x-z).$$

Substituting (5) for x provides

(6)
$$\dot{z}=l\left(N-z-x_0 e^{\frac{-k}{l}z}\right).$$

This is a single nonlinear first-order equation for z which can now be analyzed.

This equation can be nondimensionalized using the rescalings

$$u=\frac{z}{k^{-1}l} \quad \text{and} \quad \tau=\frac{t}{\hat{T}},$$

where $\hat{T}$ is some combination of parameters with dimension of time that is specified below. If populations are considered to already be dimensionless then one can see that the dimension of u denoted by $[u]$ is

$$[u]=\frac{[z]}{[k^{-1}][l]}=\frac{1}{(1/T)^{-1}T^{-1}}=1.$$

Thus

$$\frac{du}{d\tau}=\frac{d\left(\frac{1}{k^{-1}l}z\right)}{d\left(\frac{1}{\hat{T}}t\right)}=k\frac{\hat{T}}{l}\frac{dz}{dt}.$$

Substituting these rescaled variables into (6) produces

$$\frac{l}{k\hat{T}}\frac{du}{d\tau}=l\left(N-\frac{l}{k}u-x_0 e^{\frac{-kl}{kl}u}\right),$$

$$\frac{du}{d\tau}=k\hat{T}\left(N-\frac{l}{k}u-x_0 e^{-u}\right),$$

$$\frac{du}{d\tau}=k\hat{T}N-l\hat{T}u-k\hat{T}x_0 e^{-u}.$$

Note that

$$[kx_0]=1[k]=\frac{1}{T}.$$

Thus $\hat{T}=(kx_0)^{-1}$ i.e. the time it takes for the initial susceptible population to come in contact with the infected individuals, is a valid choice of $\hat{T}$ and produces

$$\frac{du}{d\tau} = \frac{N}{x_0} - \frac{l}{kx_0}u - e^{-u},$$

or

(7)
$$\frac{du}{d\tau} = a - bu - e^{-u},$$

where

$$a = \frac{N}{x_0}, \qquad b = \frac{l}{kx_0}$$

and both a and b are dimensionless. Clearly $a \geq 1$ since $N \geq x_0 > 0$ and $b > 0$ since

$l, k, x_0 > 0.$

Using the rescaled model (7), one can then determine the number and stability of fixed points

(values of u where constant solutions to (7) exist) by studying the graphs of $a - bu$ and e^{-u}.

These two curves are shown in Figure 1 for typical values of a and b.

Graphs of e^{-u} and a-bu for typical values of a and b
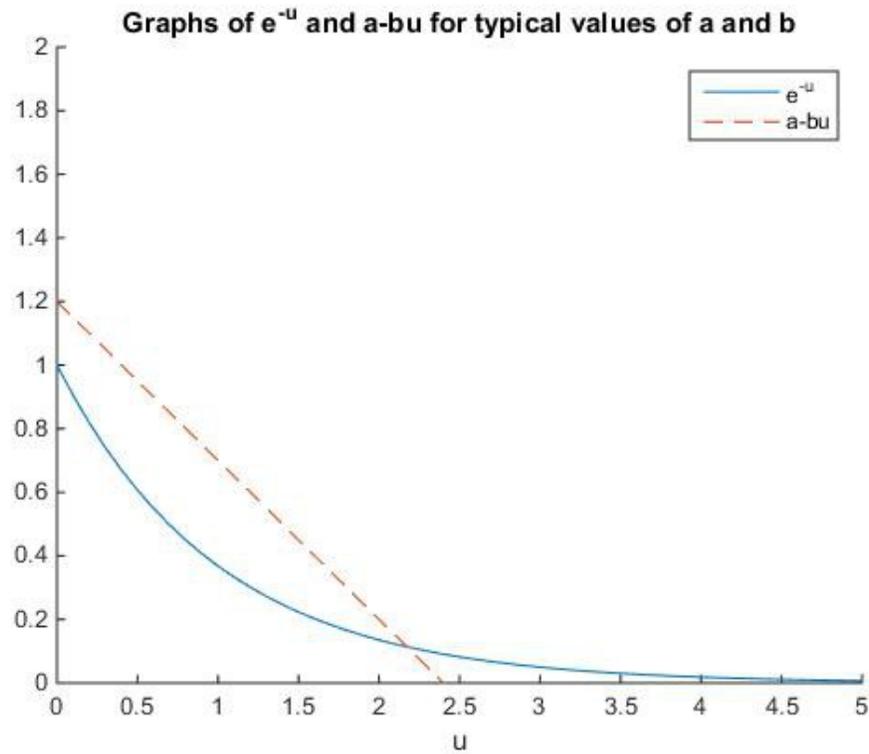
Fig. 1. The value of u corresponding to the intersection of the two curves represents the unique stable fixed point ũ. One can

see that for values of a>1 and b>0 there is always one unique intersection and thus one unique fixed point.

From Figure 1, one can see that there exists a unique fixed point for some u value $u=\tilde{u}$ where the

two curves intersect. If $u<\tilde{u}$, then $a-bu>e^{-u}$ and thus $\dfrac{du}{d\tau}=a-bu-e^{-u}>0,$ which means that

solutions to (7) will increase to the fixed point for values of u less than the fixed point. If $u>\tilde{u}$ then

$a-bu<e^{-u}$ and solutions will decrease to $\tilde{u}$. Hence $\tilde{u}$ is a stable fixed point.

Next the concept of an epidemic is introduced and analysis of (7) provides what conditions are

necessary for one to occur. Recall that z is the number of dead individuals and u is z rescaled. So

$\dfrac{du}{d\tau}$ is the death rate. From (7) one can see that the death rate can be described as the difference

between the two functions $f_1=a-bu$ and $f_2=e^{-u}.$ In Figures 2 and 3 one can see the difference of

these two functions as the length of the vertical lines spanning from $f_1(u)$ to $f_2(u)$ for arbitrary

values of u. Both figures are for the same fixed $a \geq 1$. In Figure 2 $0 < b < 1$, while in Figure 3

$b > 1$.

Qualitative estimate of death rate
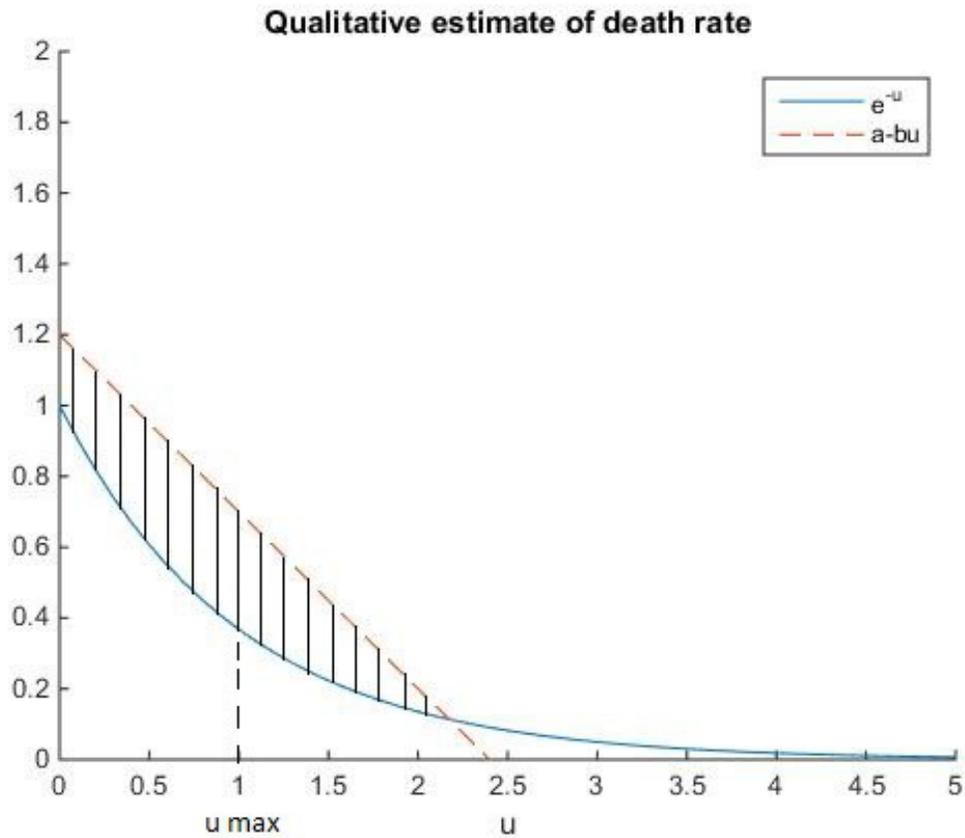
Fig. 2. The lengths of the vertical lines indicate the death rate. One can see that as u increases from 0 to the fixed point ũ the

lengths increase to a maximum length at a point u_{max} and then decrease to zero. Note that u_{max} is not exactly 1.
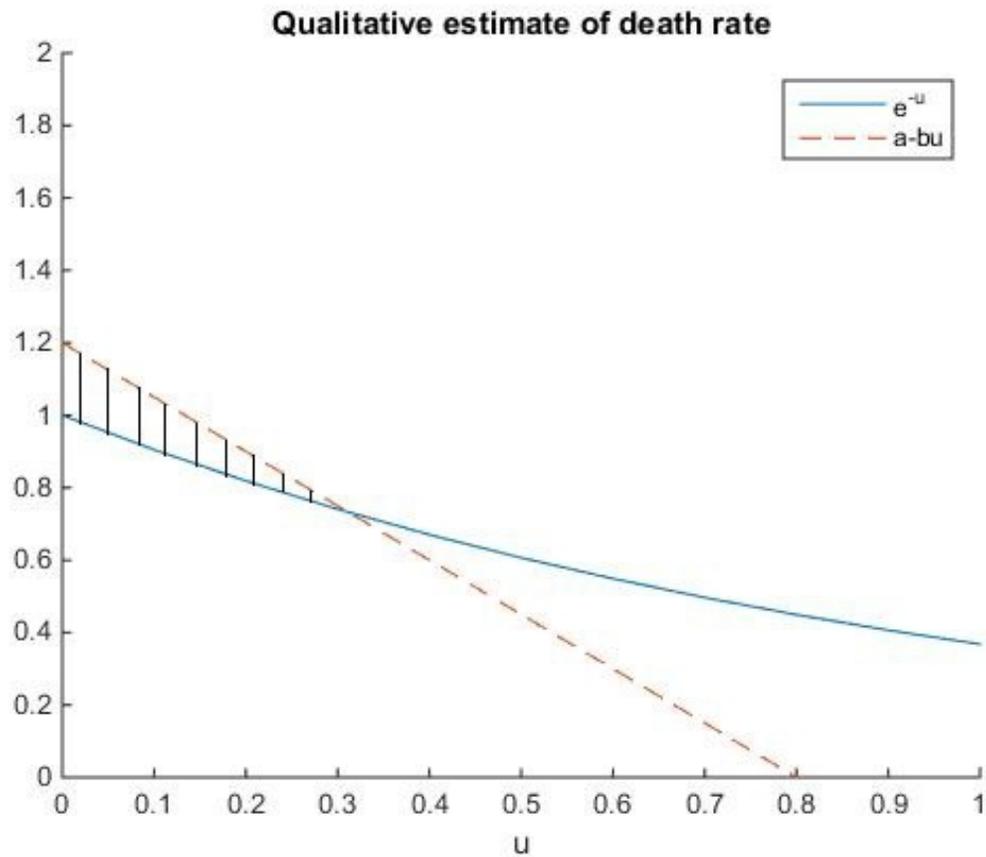
Fig. 3. The length of the vertical lines again indicate the magnitude of the death rate. Here one can see that lengths decrease

to 0 as u increases to the fixed point.

To understand what these two figures say about the death rate, note that $u(0)=0$ and that u

increases monotonically to the unique fixed point $\tilde{u}$ (the intersection of the curves in Figures 2 and

3) as τ increases. So the change in the death rate is seen in Figures 2 and 3 as the change in the

lengths of the vertical lines as one moves from left to right. It is clear from Figure 2 that the death rate

$\dfrac{du}{d\tau}$ achieves a unique maximum at some u value u_{max} corresponding to a unique time denoted

τ_{max}. In the case where $b>1$, $\tau_{max}=0$ since the longest vertical line occurs at $u=0$.

Since u is simply a rescaling of z, then $\dfrac{dz}{dt}$ must also achieve a unique maximum when

$\dfrac{du}{d\tau}$ does at time t_{max} where $t_{max} = \dfrac{\tau_{max}}{kx_0}$. The case described by Figure 2 with $0 < b < 1$ refers to the situation in which an epidemic occurs. The rate at which individuals die increases before it decreases; or stated simply, things get worse before they get better. On the other hand, if $b > 1$ the peak of the epidemic occurs at time 0, thus it is inappropriate to say an epidemic occurs at all since things only get better.

It follows that $b = 1$ is the threshold value for the occurrence of an epidemic. Recall that

$b = \dfrac{l}{kx_0}$, where l is the rate at which infected individuals die, and kx is the rate at which healthy individuals are infected. When b is smaller than 1, it follows that at times t close to 0, so that $x(t)$ is close to x_0, the rate at which infected individuals die is smaller than that at which healthy individuals become infected and thus the number of infected individuals y is increasing. However the death rate, ly, depends on the number of infected individuals, and it increases as y increases. At the same time, the infection rate, $kxy - ly$, decreases as the death rate increases and becomes negative once ly exceeds kxy. Once the infection rate becomes negative, the number of infected individuals decreases, and as such the death rate tends toward zero and the epidemic comes to a halt.

To illustrate the notion of an epidemic consider the situation where a population consisting of 1000 individuals initially has 50 that are infected and 950 are susceptible to the infection. Also assume that infected and susceptible individuals encounter each other at a rate of .15 encounters per unit time, and those infected recover, with complete immunity, at a rate of 12 individuals per unit time. In the context of the original system $x_0 = 950, y_0 = 50, z_0 = 0, k = .15, l = 12.$ Calculating b for this situation

one obtains

$$b=\frac{l}{kx_0}=\frac{12}{.15(950)}\approx.084.$$

In this case b is certainly less than 1 so one would expect an epidemic to occur. Figure 4 shows the

solutions to x, y, and z with this set of parameters and initial conditions.
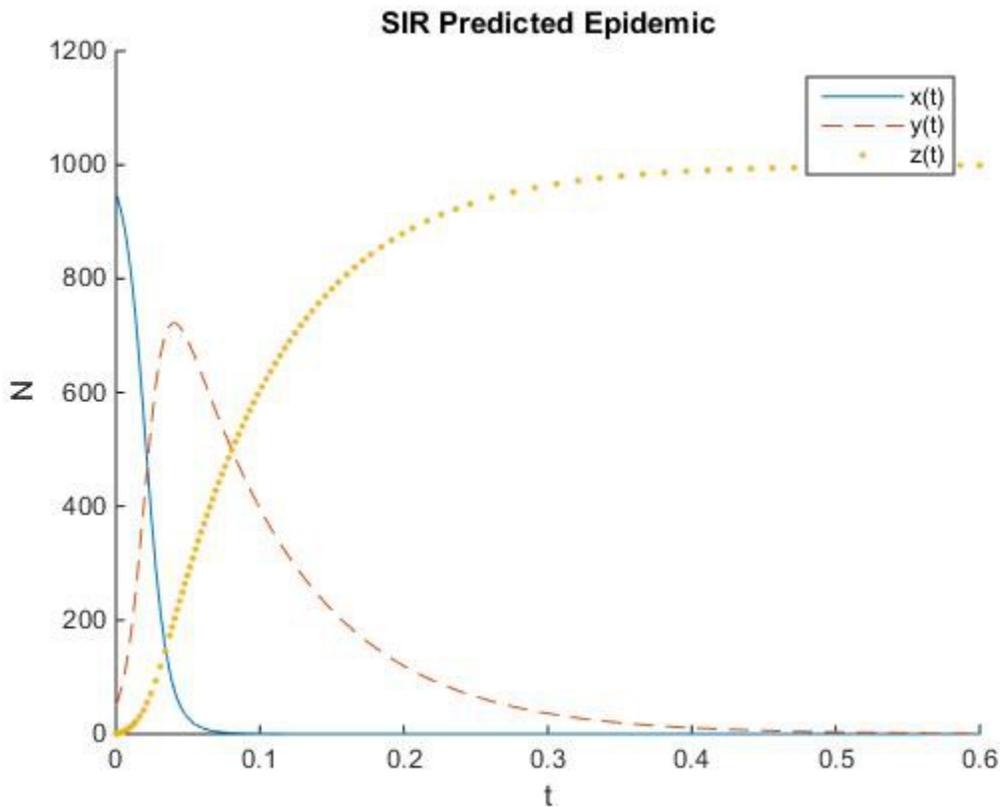
Fig. 4. Here one sees the solutions x(t), y(t), z(t) for the original SIR system described at the beginning of this section with

k=.15, l=12, x_0=950, y_0=50, and z_0=0.

The solutions observed in Figure 4 comply with the expected solution behavior in an epidemic. Here

one can see the number of removed, or recovered, individuals stabilizing as the removal rate goes to its

equilibrium point. The number of infected individuals reaches a unique maximum when the removal

rate is at its highest, and then goes to 0 as the removal rate goes to 0.

§3 The SIR Model and Genetic Diversity of Pathogens

Mathematical epidemiology has been a powerful component of the population biologist's repertoire since the SIR model was first proposed by Kermack and McKendrik in 1927 [2]. Some years later, quantitative biology was further augmented when biologists such as Sewall Wright, Ronald Fischer, and Ernst Mayr attempted to integrate methods of population biology and quantitative genetics with evolutionary theory during a period known as the Modern Synthesis. The Modern Synthesis created a new paradigm by which evolutionary theory could be explored and studied by biologists. Perhaps as an extension of the Modern Synthesis of the twentieth century, a new synthesis has emerged combining the concepts proposed during the Modern Synthesis with mathematical epidemiology. This new framework, dubbed "Phylodynamics" by Grenfell et al. combines methods of mathematical epidemiology and evolutionary biology in order to better assess the dynamics of an epidemic by accounting for selection on pathogen populations [3].

In particular, a 2009 paper by Gordo et al. uses the SIR model to assess the levels and patterns of genetic diversity in pathogen populations [4]. They introduce a SIR model that has been modified to allow for births and deaths of hosts for reasons not related to the infection. The SIR model is then used to track the infection as it spreads throughout a population. However, the ultimate goal of the paper is to track the pathogens themselves and not the individuals or hosts. Via numerical simulation using the SIR model the researchers are able to develop a mathematical framework that monitors the genetic diversity of pathogens whose epidemiology is compliant with the SIR structure as the status of the infection changes [4].

In their paper the pathogen population is treated as a metapopulation where each subpopulation (deme) consists of those pathogens within a given infected host. A metapopulation is simply a

collection of demes of a given organism, where interbreeding and interactions between demes is so small that it is considered negligible. The demes then are nothing more than the largest interbreeding groups within the metapopulation. In the case of asexual organisms, such as the pathogens tracked in this paper, a deme is the largest group of collectively mutating and reproducing organisms in the metapopulation. In their model, births of hosts (and thus the creation of empty demes or niches to be occupied by pathogens) are accounted for, and the contact structure of hosts is such that any host may come in contact with another host with equal chances of infection. It should be noted however that the model does not allow for reinfection, therefore only empty demes in the susceptible state may be occupied and those in the recovered state are always empty. It is also assumed that all reproduction of pathogens following an infection event is done with some rate of mutation per genome per generation [4, 5]. Since many pathogens reproduce at extremely high rates [4], mutations are also occurring quickly thus positive selection must be occurring in some portion of the genome of a given deme. As a genome is altered more rapidly, new and beneficial collections of mutations become more likely to occur and will thus be selected for.

Ultimately the researchers examine the parameter R_0 in the model, known as the basic reproductive number [4], which is essentially an average of the number of infection events a particular infected host can cause; alternatively, R_0 measures the number of new demes a given deme can create. Extrapolating from R_0 one can obtain some estimate of a given deme's fitness since any demes created by it would share the same basal genome and set of mutations. Those demes with a higher R_0 would implicitly have a higher fitness and thus would have greater positive selection pressure. In order to understand the effects R_0 and selection had in the model, the researchers allowed for a slightly distinct variant with higher fitness to the population [4]. A randomly chosen host is infected with the more fit strain, and then that strain's fate is tracked as the simulation runs its course

until one of two scenarios occur: either the strain is lost or the strain is fixed. Fixation refers to the

event where one deme excludes the others as it continues to be selected for until it is the only strain

present in the population [4]. From this, the probability that the particular mutant will fixate is given by

the number of simulations in which fixation of that mutant occurred over the total number of

simulations. The authors also study the time to fixation which can be measured using the number of

succeeding generations it takes for a mutant to attain fixation [4]. High fixation probabilities and short

times to fixate correspond to strong selective forces on a particular pathogen strain and thus indicate

ultimately lower levels of genetic diversity. Alternatively, in situations where fixation probabilities

were low, genetic diversity was greater and selective forces weaker [4].

One problem with the researcher's simulations that ran to fixation was that they assumed an

infected state conferred total immunity to the host, thus no reinfection could occur. Since most

pathogens do not behave in this manner [4], the model had to be modified using H3N2 (Influenza A) as

a model organism. H3N2 does not confer total immunity, and its evolution is characterized by periods

of increasing diversity due to mutations punctuated by periods of decreasing diversity when fixation of

an advantageous strain occurs [4]. Advantageous strains often emerge when a strain is able to overcome

the antigenic resistance conferred upon previously infected hosts. The SIR variant used was further

modified to allow for such reinfection by allowing recovered hosts to be reinfected with the some

probability distinct from that which susceptible hosts are infected [4]. In an analysis similar to before, a

new, more advantageous strain was introduced into a random host. The advantage of this strain came

from the fact that it created a longer infection in hosts where a reinfection event had occurred although

the same duration in first-time infection events as the other strains [4]. The selective advantage was

accounted for by establishing a threshold number of mutations d_c by which any strain that had

acquired a number of mutations greater than d_c is considered to have the selective advantage over

those below this threshold [4]. After running simulations using the modifications to permit reinfection, the researchers observed an evolutionary pattern similar to that of H3N2, thus adding support to their framework.

Using phylodynamic techniques the researchers were able to develop a framework by which genetic diversity among pathogen metapopulations can be studied via analysis of both within and between host evolutionary processes [4]. In particular, their analyses allowed them to derive and use the following expression,

$$\pi = \frac{IU}{e+b},$$

where π represents the level of genetic diversity, I is an estimation of the infected proportion of the host population or the number of demes, U is the mutation rate, and $e+b$ is a measure of the probability a deme will be removed [4]. According to this expression, genetic variability in pathogen metapopulations is proportional not only to the level of an infection but also its duration, that is, pathogens which have the capacity to cause longer infections also have the capacity to sustain higher amounts of genetic variation. By allowing reinfection into the model, the researchers developed a mathematical framework derived from a simple modified SIR epidemiology model that can be used to study the evolution of pathogens which behave in such a manner [4].

§4 An Alternative to SIR Based Modeling

Not all epidemiology models are appropriately applied to any disease event. To illustrate this point, consider the recent Ebola epidemic in West Africa in 2014. The SIR model works very well for outbreaks and epidemics where there is much data about the epidemiology of the particular pathogen. The SIR model requires information about incidence, immunity, and contact patterns of the host population. When there is little or unreliable data on an epidemic, the SIR model tends to be ineffective as its parameters cannot be determined with any appreciable confidence. For the Ebola outbreak that began in 2014, there was very little past data available on the epidemiology of Ebola. In fact, fifty percent of all the documented cases of Ebola occurred during the outbreak of 2014, which is still ongoing [6]. Reports coming out of West Africa were also highly erratic or admittedly conjectural. Because of this lack of confidence in the data, or a complete lack of data thereof, modelers were forced to use different methods to make predictions.

In 2013 a model known as the Incidence Decay with Exponential Adjustment or "IDEA" model was proposed and got to see its first major use during the Ebola epidemic of 2014 [6]. The model is the simple equation

$$I_t = \left(\frac{R_0}{(1+d)^t} \right)^t,$$

where t is time scaled in generation time, R_0 is the basic reproductive number, and d is a control parameter used to cause a decay in incidences [6]. Values of d change with control initiatives such as increased public sanitation or other public health gains, behavioral change, increased immunity, and other changes that slow transmission [6]. I_t gives the number of incidences within a given generation. One notable feature of this model is that in the absence of control, when $d=0,$ the model becomes

$$I_t = R_0^t,$$

which is always exponential growth in time. However when $d \neq 0$ the number of incidences decreases

by a power of t^2, which then causes transmission to cease [6]. This model works very well for

analyzing short term outbreaks or epidemics that have little data available, since R_0 can be fairly

easily estimated via empirical analysis of what data is available.

Researchers Fisman et al. decided to use the IDEA model and what data were available to

generate quick projections for the outbreak as well as determine what course of action would be most

beneficial to stopping it. The only reliable data available to them were cumulative incidence data

compiled by the World Health Organization (WHO), which was sufficient for the IDEA model to work.

They began by parametrizing the model by fitting it to the data obtained from the WHO; choosing the

parameter values that minimized the distance between the model projections and the acquired incidence

data [6]. They then established a range of parameters that most accurately described the current status

of the epidemic, and found $R_0 \approx 1.8$ and d to be only around .010 [6]. As it stood the model

projected that the number of incidences would continue to increase, peaking in April 2015, and

carrying forward into 2016, with a final size of more than 140,000 total cases [6]. However the

researchers did find that a small increase in d (.014 instead of .010) caused a constant decrease in

incidences with a peak of less than 10,000 total cases.

The researchers were able to determine that immediate action was required in order to

potentially decrease the number of victims in this outbreak tenfold. Had the researchers relied solely on

the SIR model, or other models like it, they would have undoubtedly been unable to acquire such

alarming results, and any results obtained would have been highly unreliable. While the IDEA model

does not give any indication as to what specific interventions are needed to increase the value of d,

and thus drastically slow the spread of the virus, the researchers were able to use the model to

determine that intervention early on was critical in preventing the outbreak from spreading to extremely catastrophic dimensions.

Despite being the classic epidemiology model, from which many models are then derived, SIR is not without its weaknesses—as made apparent in 2014. SIR is useful for diseases which have much reliable data readily available, which is often not the case in modeling epidemics. Most diseases for which reliable data are available are also already well understood, having already been modeled in typical conditions after years of data collection. For quick and meaningful results (as are often required in these situations) a simple model such as IDEA is often more desirable. That is not to say SIR is without its uses. SIR still remains as the basis on which many infectious disease models are built, and can be further used to study things such as pathogen population dynamics and evolution, thereby extending its potential.

References

[1] Hethcote, Herbert W. *The Mathematics of Infectious Diseases*. SIAM Review 42.4 (2000): 599. Web.

[2] Strogatz, Stephen H. *Nonlinear Dynamics and Chaos.* Cambridge: Perseus Books, 1994. Print.

[3] Grenfell, Bryan T, et al. *Unifying the Epidemiological and Evolutionary Dynamics of Pathogens.* (2004) Science 303, 327 DOI: 10.1126/science.1090727

[4] Gordo I, Gomes MGM, Reis DG, Campos PRA *Genetic Diversity in the SIR Model of Pathogen Evolution.* (2009) PLoS ONE 4(3): e4876. DOI:10.1371/journal.pone.0004876

[5] Combadão Jaime, Campos Paulo R. A., Dionisio Francisco, Gordo Isabel. *Small-world networks decrease the speed of Muller's ratchet.* (2007) Genetics Research 89(7)DOI**:** 10.1017/S0016672307008658

[6] Fisman, David, Edwin Khoo, and Ashleigh Tuite. *Early Epidemic Dynamics of the West African 2014 Ebola Outbreak: Estimates Derived with a Simple Two-Parameter Model.* PLoS Currents. Public Library of Science, 8 Sept. 2014. Web.

[7] Ricklefs, Robert E., and Rick Relyea. *Ecology: The Economy of Nature*. 7th ed. Gordonsville, VA: Freeman, W.H., 2014. Print.