Final Report for Sample Project 1

Brian R. Hunt

January 2011

Abstract

We develop a simple model for the spread of an infectious disease that is limited in duration and is not fatal. We first assume that when an infected person recovers, (s)he becomes susceptible to being infected again, but we also consider the case where some of the recovering people become immune to the disease. We find in the first case that the prevalence of the disease – the proportion of the population that is infected at a given time – approaches a limiting value between 0 and 1 in the long run, and that the predicted limiting value is very sensitive to errors in data taken when the prevalence is small and growing. In the latter case, we find that the prevalence of the disease reaches a peak and then decays toward zero; in the long run, a certain proportion of the population has become immune and a certain proportion remains susceptible.

1 Introduction

An "epidemic" occurs when an infectious disease that is new to a population is introduced to that population and spreads to a significant proportion of the population. The population generally consists of people (or other animals, such as livestock) in a certain geographical area, and the disease generally comes from contact with a population in a different region. The purpose of modeling an epidemic once an outbreak is detected is both to predict how serious the epidemic will be, and to evaluate the effectiveness of various possible responses to the epidemic.

An infectious disease is spread by contact between infected and uninfected people, as opposed to a hereditary disease. We will use a continuous-time model based on approximating the flow of people between different subgroups of the population – those who are infected, those who are susceptible to infection, and those who are immune from infection. The continuous model regards the populations as continuous variables, though in real life they are discrete (integer-valued). This is reasonable as long as the population is large, so that changing the status of a single individual is not too significant.

Many epidemiological models divide the infected population into smaller subgroups based on how long they have been infected or how their symptoms have progressed. To keep the model simple, we will not do this. In particular, we do not attempt to model the facts that in people in different stages of the disease may have markedly different degrees of infectiousness (ability to spread the disease), that different people may have different amounts of contact with others and different degrees of resistiveness to the disease, and that recovery may take a certain period of time. We also do not take into account that the population may modify

its behavior in response to the epidemic; once diagnosed, people who are infected are likely to have their contact with others limited. Finally, we assume that there are no "carriers" – people who are immune but still carry the disease and can infect others.

We now spell out the main assumptions we make and introduce the notation we will use.

(i) The total population is large and constant in time.

(ii) At a given time t, the population can be divided into three groups:

- $N(t)$, the proportion of people who are infected (the "prevalence" of the disease);
- $S(t)$, the proportion of people who are susceptible; and
- $M(t)$, the proportion of people who are immune.

(iii) Each infected person is equally likely to recover on a given day, independent of how long (s)he has been infected. Let r be the probability per unit time that an infected individual recovers.

(iv) Each recovering person is equally likely to become immune to future infections, independent of how many times (s)he has been infected before. Let q be the probability that a recovering individual becomes immune.

(v) For a given prevalence $N(t)$, each susceptible person is equally likely to be infected on a given day, and the likelihood is proportional to $N(t)$. Let the probability per unit time that a susceptible individual becomes infected be $kN(t)$.

Let us discuss the rationale behind the last assumption. Imagine that each susceptible person comes into contact with the same number of people each day on average, and the proportion of these people who are infected is the same as the prevalence $N(t)$ in the entire population. If each contact between and infected person and a susceptible person has an equal chance of spreading the disease, then k represents the probability of catching the disease form a single contact times the average number of people contacted per unit time. An underlying assumption here that the population is relatively homogeneous; not only does each person interact with others at about the same level, but that each person is more or less equally likely to interact with each other person.

In Section 2, we consider the case with no immunity. In this case, our model is a single first-order differential equation, and we are able to find an analytic solution depending on 3 parameters. We show how to determine these parameters from 3 data points, namely the prevalence $N(t)$ at 3 different times. We also express the long-term prevalence of the disease (the limit of $N(t)$ as $t \to \infty$) in terms of the parameters. Finally, we show by means of a specific example how small changes in the data can have a substantial impact on the predicted limit.

In Section 3, we allow for immunity, leading to a system of 2 first-order differential equations. We study a few scenarios by solving this system numerically. In all scenarios, we find that the prevalence of the disease eventually decays to 0, while the proportion of people who are immune grows but does not approach the entire population. Thus, according to our model at least, even though the disease dies out, part of the population remains susceptible to a future outbreak. In Section 4 we discuss the results and conclusions we draw from them.

2 Model with no Immunity

If nobody can become immune to the disease, then in the notation introduced above, we have $M(t) = q = 0$ for all t. It follows that $N(t) + S(t) = 1$ for all t. In particular, we need only determine one of the two functions $N(t)$ and $S(t)$ since we can find the other function by subtracting from 1. It turns out that we get a simpler differential equation for $N(t)$, which we derive now.

Based on the large population assumption, we assume that the number of people who recover in a given short time span is proportional to the current number of infected people, with the proportionality constant being r times the time span. Thus the rate of decrease of $N(t)$ due to recovery is $rN(t)$. Similarly we model the rate of increase of $N(t)$ due to new infections as $kN(t)S(t) = kN(t)(1 - N(t))$. Thus our model for this section is

$$\frac{dN}{dt} = -rN + kN(1 - N) = (k - r - kN)N. \tag{1}$$

(We will often write N instead of $N(t)$ for brevity.)

Before giving the general solution to this equation, let us consider its equilibrium solutions. We have $dN/dt = 0$ if either $N = 0$ or $N = (k - r)/k$. For the situation we are modeling, the only realistic values of N are from 0 to 1, so the latter equilibrium is only relevant if $k > r$. In that case, $dN/dt > 0$ if $0 < N < (k - r)/k$ and $dN/dt < 0$ if $N > (k - r)/k$, so that $N = 0$ is an unstable equilibrium and $N = (k - r)/k$ is a stable equilibrium. Thus if N is positive, it will approach a limiting value of $(k - r)/k$ as $t \to \infty$. If on the other hand $k < r$, then $dN/dt < 0$ for all $N > 0$, so that N decays to 0 as $t \to \infty$. In this case the epidemic would never break out in the first place, since the rate at which people recover exceeds the rate at which they can infect new people.

The general solution of equation (1)

$$N(t) = \frac{k - r}{k + Ce^{(r-k)t}} \tag{2}$$

where C is a constant. (There is also the equilibrium solution $N(t) = 0$ for all t; this corresponds in a manner of speaking to setting $C = \infty$.) If $r > k$, then for $N(t)$ to be positive C must be negative. Then the denominator approaches $-\infty$ as $t \to \infty$, and hence $N(t)$ approaches 0 as we determined before. If $k > r$, then the denominator of equation (2) approaches k as $t \to \infty$, and $N(t)$ approaches $(k - r)/k$. Again this agrees with our earlier analysis.

From now on, we will assume that $k > r$, so that an outbreak of the disease is possible ($N = 0$ is unstable). Several solution curves are shown in Figure 1, with hypothetical values chosen for k and r and different curves corresponding to different values of C – in essence, different initial conditions.

To determine the values of the parameters k, r, and C appropriate to a given scenario, we need some data. While we could attempt to measure k and r directly if we had good data on new infections and recoveries, we can also infer them if we just know the value of N at 3 different times. By plugging these times into equation (2), we get 3 equations to solve for the 3 unknown parameters. If we had less data, then a variety of parameter values would fit the data perfectly, and we would have no good basis to choose which values to make predictions from. If we had more data, we would have more equations than unknowns, and it is unlikely that we could find values for k, r, and C that exactly fit all of the data. Instead, we would find the

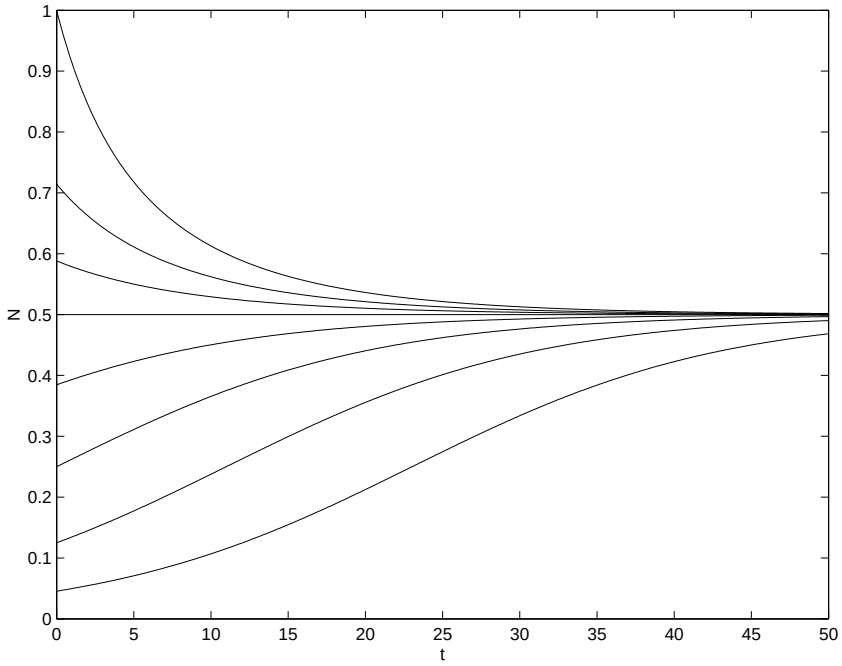

Figure 1: Solutions to equation (1) with $k = 0.2$ and $r = 0.1$, indicating the disease prevalence (according to the model) as a function of time for various initial conditions.

values that "best" fit the data by some statistical criterion (such as least-square error). How well we could fit the data would give us an idea of how realistic our model is.

To fit the data given in the problem statement, we let t be measured in years, with $t = 0$ representing the present time. Then the data are

$$N(-10) = 0.002, \qquad N(-5) = 0.008, \qquad N(0) = 0.03.$$

From equation (2) we get

$$\begin{aligned} \frac{k-r}{k+Ce^{10(k-r)}} &= 0.002 \\ \frac{k-r}{k+Ce^{5(k-r)}} &= 0.008 \\ \frac{k-r}{k+C} &= 0.03. \end{aligned}$$

Though these equations can be solved algebraically, the solution is long and tedious. They can also be solved numerically, by MATLAB for instance. The solution is

$$k \approx 1.036, \qquad r \approx 0.754, \qquad C \approx 8.356. \tag{3}$$

From this we obtain the long-term prevalence of the disease,

$$N(\infty) = \frac{k-r}{k} \approx 0.272.$$

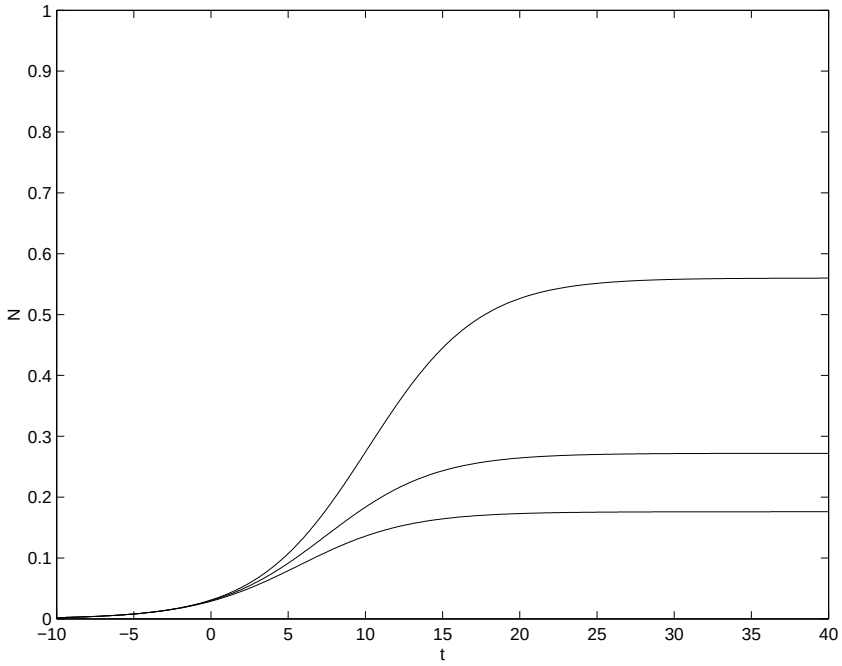

Figure 2: Prevalence as a function of time for $N(-10) = 0.002$, $N(-5) = 0.008$, and: [upper curve] $N(0) = 0.031$; [middle curve] $N(0) = 0.030$; [lower curve] $N(0) = 0.029$.

That is, eventually about 27% of the population will be infected at any given time. We emphasize that this does not mean that the remaining 73% of the population never gets infected. Instead, with 27% of the population infected, the rate of new infections and the rate of recovery balance out, but the specific people who are infected continues to change over time.

To assess the uncertainty in our prediction of the long-term prevalence, suppose the uncertainty in $N(0)$ is 0.001. Even without any uncertainty in the other data points, the effect on our prediction is dramatic. If we change $N(0)$ to 0.029 and follow the same procedure as above, we get $N(\infty) \approx 0.176$. And if we change $N(0)$ to 0.031, we get $N(\infty) \approx 0.560$. The solutions of our model in all 3 cases are shown in Figure 2.

3 Model with Immunity

Now assume $q > 0$, so that M will be positive too. In this case $N + S + M = 1$, so that if we know any 2 of the populations at a given time, we know all 3. Our derivation of the differential equation for $N(t)$ from the previous section still holds, except that we can no longer write $S(t) = 1 - N(t)$. That is, $dN/dt = -rN + kNS$. We can augment this equation with a differential equation for either S or M; we choose M because nobody can leave the immune population, leading to a simpler equation. Of the rN people recovering per unit time, a proportion q of them become immune, so the rate of change of M is qrN. Writing $S = 1 - N - M$, we get our model for this section:

$$\frac{dN}{dt} \;=\; -rN + kN(1 - N - M) = (k - r - kN - kM)N$$

$$\frac{dM}{dt} = qrN.$$

If $q = 1$, this model is equivalent (though with a difference choice of notation) to the SIR model formulated by Kermack and McKendrick in 1927 [1] (see also [2]).

For this model to be in equilibrium, we must have $N = 0$ for dM/dt to be 0, in which case $dN/dt = 0$ too. Thus as long as $q > 0$, the population can only be in equilibrium when the prevalence of the disease goes to 0. This would not necessarily be the case in a more realistic model that takes into account changes in the population; we discuss this point further in the next section.

Having no data that will allow us to determine q, let us consider various values of q along with the values of k and r from equation (3) in the previous section. We use $N(0)$ = 0.03, and to determine a somewhat realistic value for $M(0)$ we consider that in the distant past, nobody was immune, so we should have $M(-\infty) = 0$. Then

$$M(0) = \int_{-\infty}^{0} qrN(t)dt \approx 0.1128qr \approx 0.0851q.$$

Here we assumed that for N is not much different than in the previous section for $t < 0$; in other words, immunity has not has much effect on N so far. (This should be reasonable if q is small, at least.) Then we used the solution from the previous section, namely equation (2) with the values from equation (3), to compute the integral of N.

Given the setup in the previous paragraph, if $q = 0$ then $M(t) = 0$ for all t and $N(t)$ is the same as in the previous section (the middle curve in Figure 2). We computed solutions numerically for $q = 0.01$, 0.1, and 1 using MATLAB. The results are shown in Figures 3, 4, and 5. Notice that the graphs cover very different ranges of t.

When $q = 0.01$, the prevalence N grows to about 25%, nearly as high as when there was no immunity, but then slowly decays toward 0 over a period of several hundred years, while the proportion M of people who become immune grows to about 28%. This slow time scale makes sense mathematically because on average one has to get the disease 100 times before becoming immune. (Notice also that with a recovery rate of about 75% per year, the lifetime of the disease is usually but not always less than a year.) Of course we are somewhat beyond the bounds of reality here, because people don't actually live hundreds of years. It would be vital to take into account changes in the population with these parameters.

With $q = 0.1$, the prevalence N grows to about 16% in 10 to 15 years and then decays more rapidly to 0. The proportion M of people who become immune approaches roughly 34%, more than in the previous case but not radically more. With $q = 1$, the prevalence N grows to about 4.5% in about 5 years and then decays, with M approaching roughly 50%. The fact that M and S approach nearly equal values seems to be a coincidence; it is not a consequence of setting $q = 1$ because it does not happen if we set different values for the other parameters.

4 Discussion

Since we have used relatively simple models for the spread of an epidemic, we should consider to what extent our results are a just property of the particular model and to what extent they may be expected to

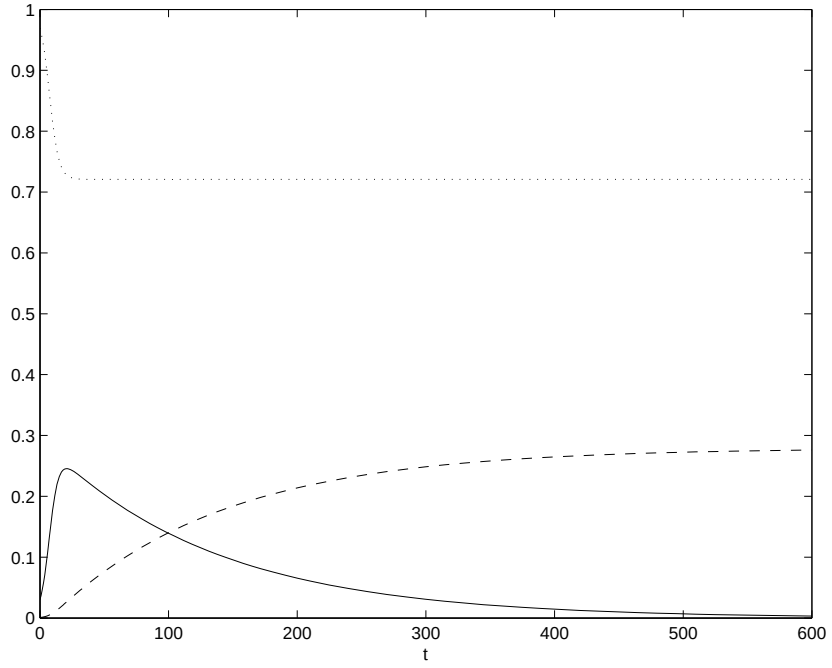

Figure 3: Graphs of $N(t)$ [solid line], $M(t)$ [dashed line], and $S(t)$ [dotted line] for $q = 0.01$.

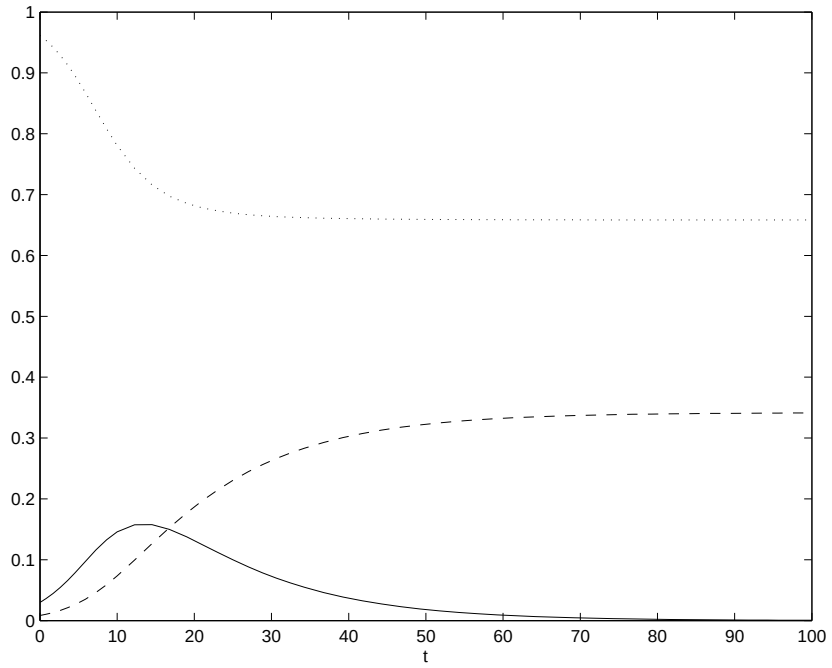

Figure 4: Graphs of $N(t)$ [solid line], $M(t)$ [dashed line], and $S(t)$ [dotted line] for $q = 0.1$.

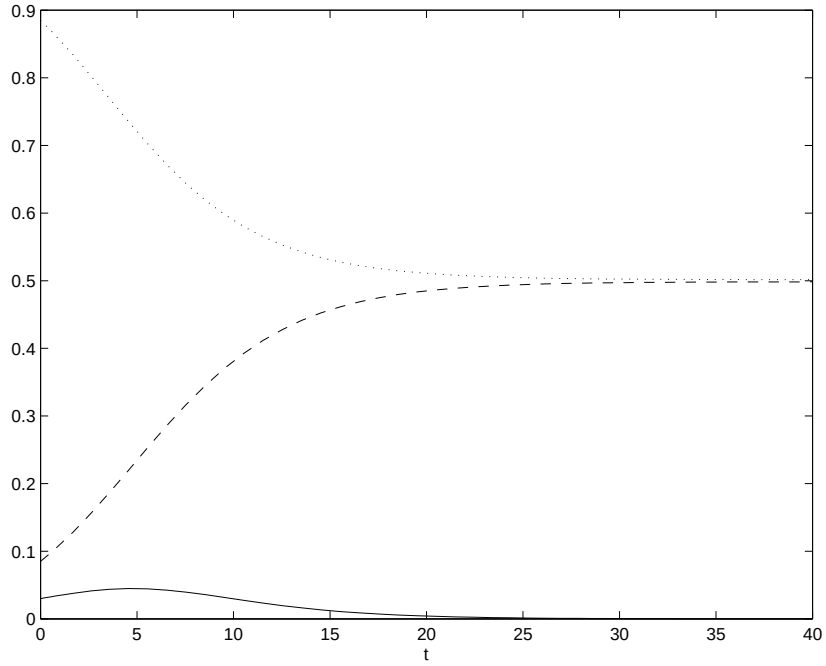

Figure 5: Graphs of $N(t)$ [solid line], $M(t)$ [dashed line], and $S(t)$ [dotted line] for $q = 1$.

apply more generally.

In both cases (with and without immunity) we found that the prevalence of the disease tends to an equilibrium value in the long run. In real life many factors may prevent such an equilibrium, most importantly changes in behavior and in preventive care once the disease has been identified, and ultimately perhaps a vaccination or cure for the disease. However, relatively mild diseases like chicken pox may indeed persist at a more or less constant prevalence for a long time. On the other hand, having chicken pox generally results in immunity ($q \approx 1$) and the prevalence is not going to zero like the model in Section 3 would predict. We will discuss this discrepancy below.

Our main result in Section 2 was that the predicted limiting prevalence can depend sensitively on the data used to determine the parameters of the model. Though we did not do a general study of this effect, we found that in at least one case, a small uncertainty in data at a time when the prevalence of the disease is small but growing rapidly can lead to a very large uncertainty in the predicted long-term prevalence. This is probably true of more sophisticated models as well.

In Section 3 we found that by allowing for immunity, we could simulate an epidemic where the prevalence of the disease rises rapidly to affect a significant percentage of the population, then falls off in the long run. As we suggested there, the fact that the prevalence falls off to 0 may not be realistic for too many diseases, but the fact that it can fall off to 0 in this model without having the entire population become immune suggests that short-lived epidemics may be followed by further outbreaks in the future. And as we will show now, a modified version of the model from Section 3 can lead to a nonzero long-term prevalence even with immunity by taking into account changes in the population.

To incorporate births and deaths into our model, we would abandon the assumption that the total popu-

lation is fixed and let N, S, and M represent total populations rather than proportions (because it is harder to derive a differential equation for a quantity whose numerator and denominator are both changing). We would end up with a system like

$$\begin{aligned}
\frac{dN}{dt} &= -rN + kNS - d_N N \\
\frac{dS}{dt} &= (1-q)rN - kNS - d_S S + b_N N + b_S S + b_M M \\
\frac{dM}{dt} &= qrN - d_M M,
\end{aligned}$$

where we have different birth rates b_N, b_S, b_M and death rates d_N, d_S, d_M for the different populations and we assume nobody is born infected or immune. Probably it would be best to set some of these rates equal, at least $b_S = b_M$ and $d_S = d_M$, to reduce the number of parameters. To be more realistic, we should divide the population into age groups, since birth and death rates and susceptibility to the disease will depend on age.

We could handle immigration and emigration similarly, though probably we would want to assume that some of the immigrants are infected and/or immune. In any case, with births and deaths and/or immigration and emigration it should be possible even with immunity to have an equilibrium without having the prevalence of the disease decay to 0, because now dM/dt can be 0 without having $N = 0$. Actually, with a growing population equilibrium would not require M to be constant anyhow, rather the ratio of M to the total population should remain constant, so that M would grow along with the population. From this point of view, it would be impossible in fact for a growing population to have equilibrium with $M > 0$ and $N = 0$; either both M and N would have to go to 0 or both should remain positive.

Probably the most serious limitation in our model is that we treated both the entire population and our 3 subgroups – infected, susceptible, and immune – as being homogeneous. To be more realistic, we could divide the infected population into subgroups based on what stage of the disease they are in, as mentioned in the introduction, and we could divide the population into age groups, as mentioned above. We could also divide the population by sex, race, geography, and any other grouping we could think of that would affect parameters like susceptibility to the disease, interaction rates with other groups, etc. Of course the model could get very complicated then, and we might have more parameters than we could hope to estimate from the available data. A reasonable approach would be to start with a simple model like those in this report, see how well it fits the data, and then build additional factors into the model until (hopefully) a good fit is achieved without having a ridiculous number of parameters.

Finally, another approach that might be of interest would be to model a heterogeneous population by a computer simulation that keeps track of each individual, assigning each one certain characteristics for sociability, susceptibility, etc., and using a random number generator to simulate chance interactions, infections, etc. This type of approach is sometimes called "agent-based" modeling and is becoming increasingly popular and increasingly feasible with advances in computer speed and memory; see for instance [3].

References

[1] W. O. Kermack & A. G. McKendrick, *A Contribution to the Mathematical Theory of Epidemics*, Proc. Roy. Soc. Lond. A **115** (1927), 700–721.

[2] `http://en.wikipedia.org/wiki/Epidemic_model#The_SIR_Model`, version of 25 January 2011.

[3] J. M. Epstein & R. Axtell, *Growing Artificial Societies: Social Science from the Bottom Up*, MIT Press (1996).