## Introduction

"The Freshman Plague" is a semester long period in which college freshmen spread the common cold in rapid succession, leaving what seems like half of the freshman class coughing at lecture and troubled by headaches and fatigue. This poster explains how this phenomenon can be modeled on an individual level. The results mimic an SIR (Susceptible, Infected, Recovered) model commonly used by epidemiologists. In order to better understand the spread of the common cold within the freshman class, and hopefully limit it in the future, we have decided to investigate the various factors (cleanliness, degree of interaction with others, etc) that have been proven to affect how one contracts a disease.

We hypothesize that our model will accurately describe the dynamics of the "Freshman Plague" at Johns Hopkins University, as predicted by a professional SIR model, and determined by the Health survey that was distributed to the freshman class.

## Background

Diseases are either acute (characterized by fast immune responses which eradicate the infection readily), or chronic (last far longer). Models of acute diseases are best represented by a system of differential equations collectively known as an **SIR model**, which stands for **S**usceptible (previously unexposed to the pathogen), **I**nfected (currently colonized by the pathogen), and **R**ecovered (successfully cleared the infection). As these categories represent proportions of the population, S + I + R = 1.

The transition from I to R can be predicted accurately by clinical data of how long a disease typically lasts, and/or its fatality rate. The recovery rate, γ, is the inverse of the mean period of infection.

The transition from S to I is dependent on 1) the prevalence of the infecteds, 2) the underlying population contact structure, and 3) the probability of transmission given contact. For the second constraint, **frequency-dependent** transmission reflects the situation where the number of contacts is independent of the population size, but rather a function of social patterns (whom does one normally see day to day?). In contrast, **density-dependent** transmission assumes that the contact rate increases with population density, which relies on random contacts and a "well-mixed" population. The difference between the two mechanisms becomes important if the population size changes or if a range of population sizes is being considered (for example, rural and urban).

The equations of the SIR model are as follows:

$$\frac{dS}{dt} = -\beta SI,$$

$$\frac{dI}{dt} = \beta SI - \gamma I,$$

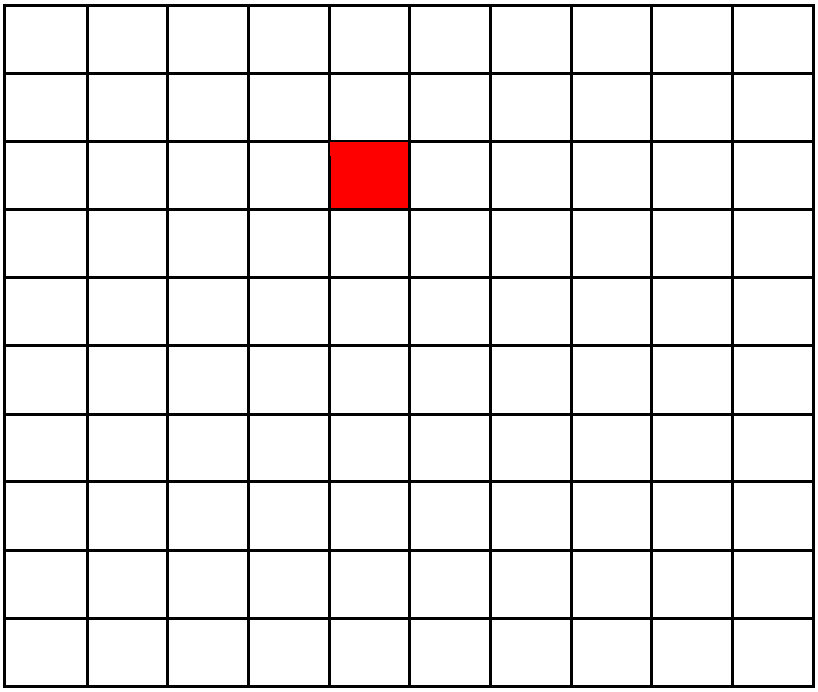
$$\frac{dR}{dt} = \gamma I.$$

| | |
|---|---|
| $\beta$ | is the transmission rate and incorporates the encounter rate between susceptible and infectious individuals together with the probability of transmission. |
| $\gamma$ | is called the removal or recovery rate, though often we are more interested in its reciprocal ($1/\gamma$) which determines the average infectious period. |
| $S(0)$ | is the initial proportion of the population that are susceptible. |
| $I(0)$ | is the initial proportion of the population that are infectious. |

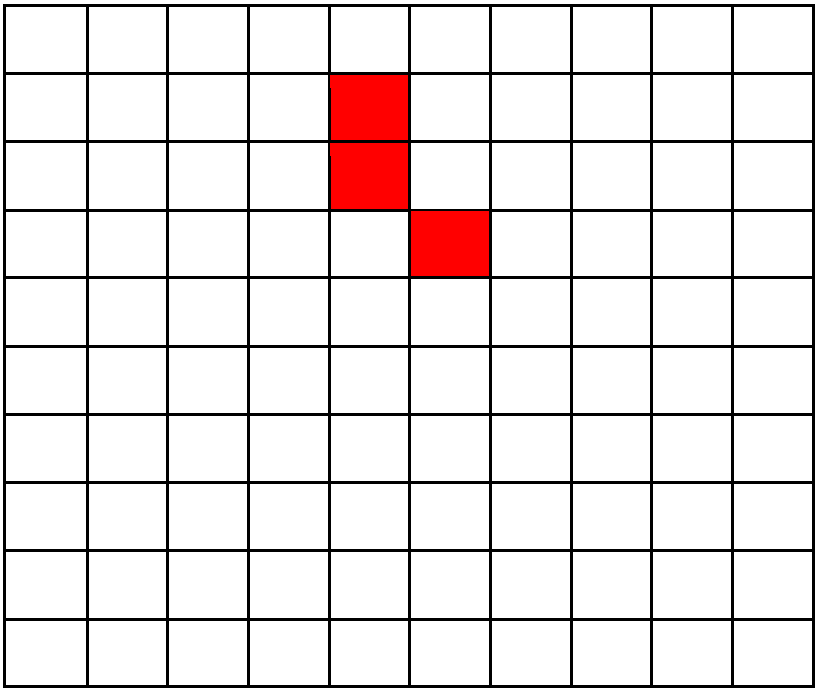
The **Threshold Phenomenon** states that if If the initial fraction of susceptibles, S(0), is less than $\gamma / \beta$ , then $dI/dt < 0$ and the infection "dies out." The inverse of $\gamma / \beta$ is known as the **basic reproductive ratio, $R_0$**. It is defined as: *the average number of secondary cases arising from an average primary case in an entirely susceptible population;* and essentially measures the maximum reproductive potential for an infectious disease. a pathogen can invade only if $R_0 > 1$.

**To include in methods:** The program is coded to do the following: infect a person (paint their cell red) into the matrix that represents a dorm. There is a certain probability that each sick person will make any of 8 neighboring cells sick. Then, a random even number of cells will be swapped to simulate movement of people within the dorm in a single day--some days may be busier than others by this added randomness. When a person becomes infected, a counter starts, and after a certain number of days the sick person becomes immune for the duration of the simulation. The process repeats itself to exhaustion: each sick person may make more people sick, then several people's locations will be swapped.

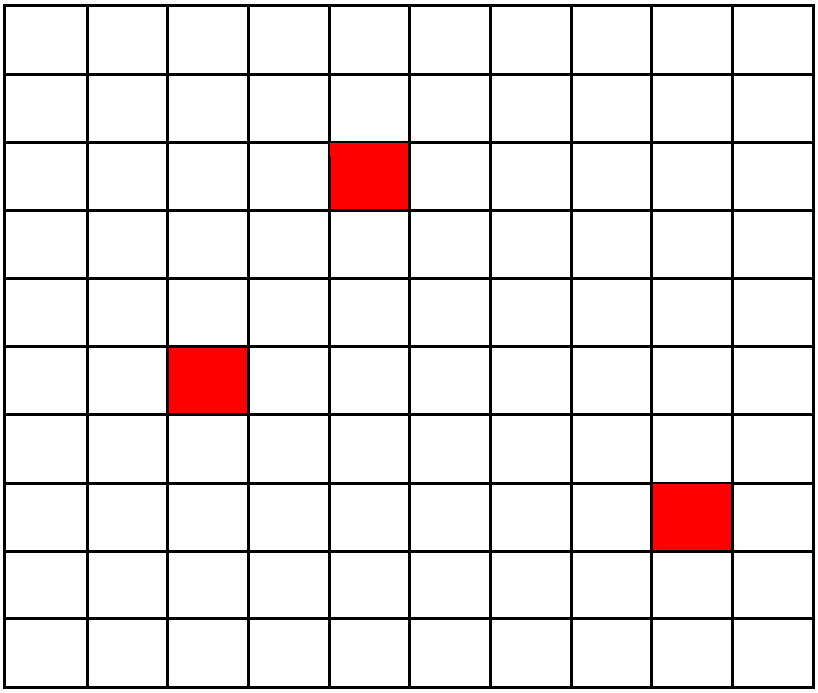
**Pictures for methods section:**



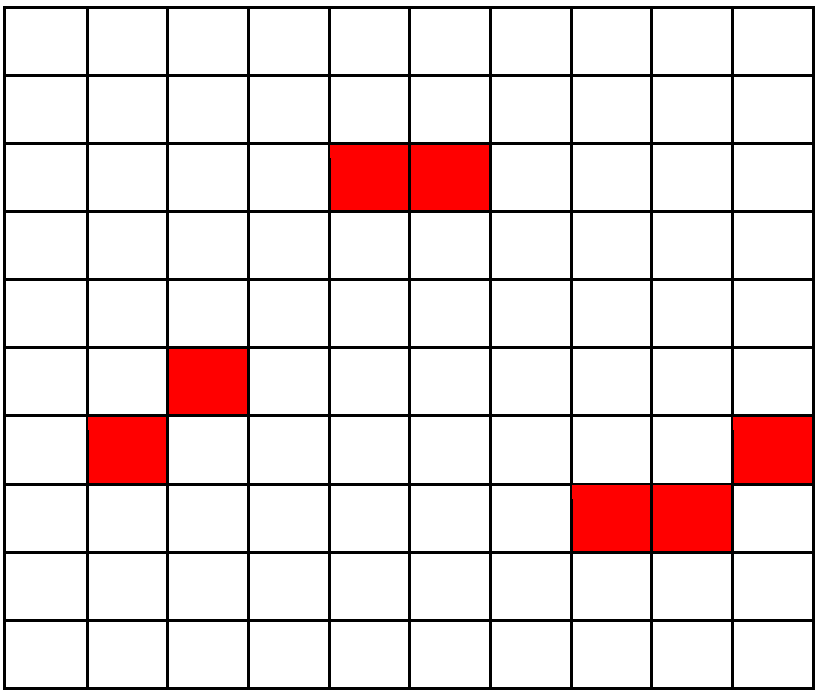
*A single person is infected*



*The sick person infects his/her neighbors.*



*The sick are dispersed.*



*The dispersed sick infect others, if they aren't already immune.*

**Results:**

We first identified three main forms of diseases, namely (I) Standard, (II) Chronic, and (III) Acute, which vary according to their transmission and recovery rates on an individual level. These rates are known in our model as "probability of infection" and "time until recovered" respectively. We used our model to analyze these three forms of disease, and then compared our results to those produced by the conventional SIR model.

(I) Standard Diseases (average transmission rate and recovery rate)

· Time taken for number of infected to return to minimal levels (indicative of disease dying out), is 60 days for our model and 40 days for the SIR model.

· Proportion of population infected peaks between ~40-60% for our model and about 50-70% for the SIR model.

(II) Chronic Diseases (low transmission rate, low recovery rate)

· Our model shows that number of infected takes a longer period of time (100 days) to return to minimal levels as compared to standard diseases, as analysed by the same model.
· This is confirmed by the SIR model which shows that it requires 60 days for the number of infected to return to basal levels.

· Proportion of population infected peaks at 35% for our model and 60% for the SIR model.

(III) Acute diseases (high transmission rate, high recovery rate)

· For our model, time taken for number of infected to return back to minimal levels is 100 days, which is much longer than the other two forms of diseases (standard and Chronic). The SIR model, however, reflects a much shorter period of infectivity, which number of infected returning to minimal levels just after 10 days.

Proportion of population infected peaks at 90% for our model and 75% for the SIR model.

Although our data depends on the probability of infection, P, and the time it takes until a sick person is immune, the following generalities can be made based on our results:

1. The percentage of people sick rises exponentially as the abundance of a susceptible population facilitates a positive feedback cycle. As immunity builds in the population, however, the percentage of the population infected decreases exponentially because those infected come into contact with those susceptible less frequently. That implies that S decreases by $S \simeq S(0)e^{-R(t)R_0}$

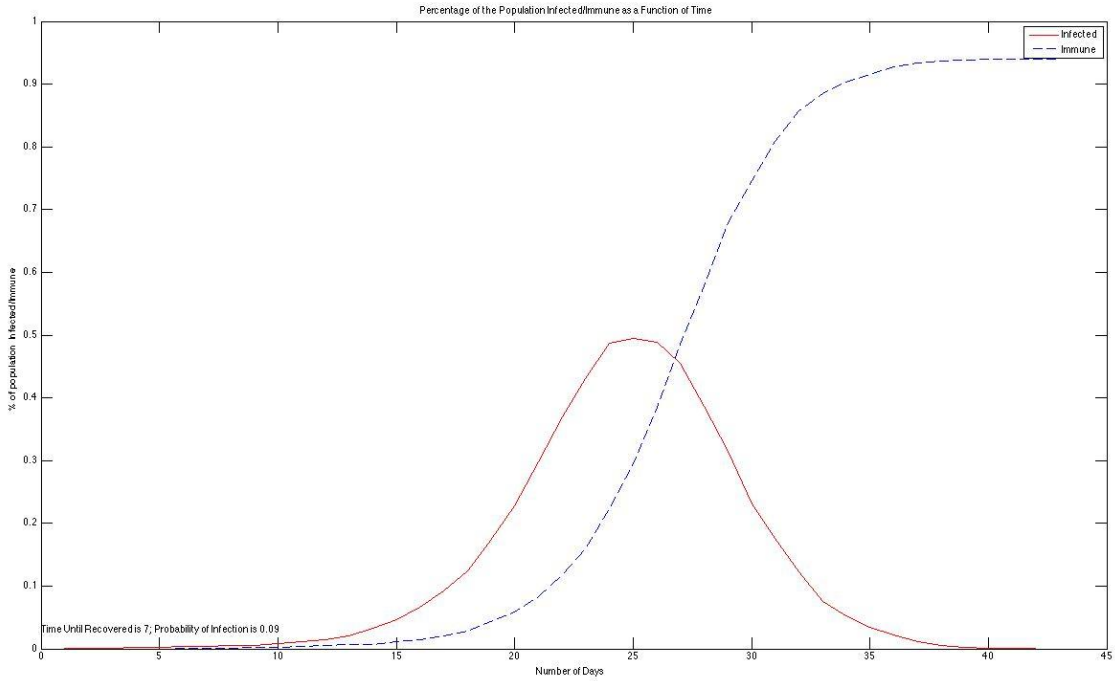
2. The percentage of the population immune increases by logarithmically (except for the acute case), and immunity always lags infection. (The equation of those immune as a function of time can only be calculated numerically)

3. **Chronic diseases**: A disease with low probabilities of infection but high recovery times yields two outcomes: either it dies out quickly, or it infects small quantities of the population for very long periods of time. The more time it takes to reach immunity, the longer the disease persists. This describes diseases like HIV, chlamydia, and herpes.

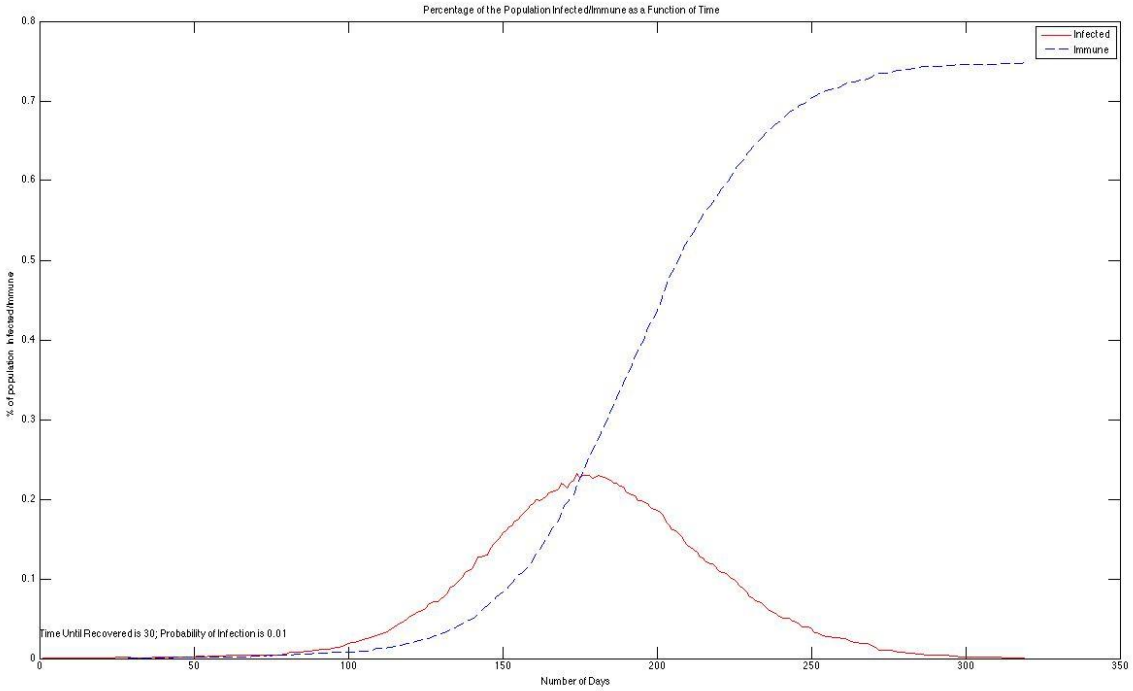
4. **Acute diseases**: with high probabilities of infection and short recovery times, nearly the whole population is infected immediately, but then the whole population recovers promptly. The individual may recover within days or weeks, which characterizes influenza, chickenpox, and rubella.

**Graphs to include in results:**

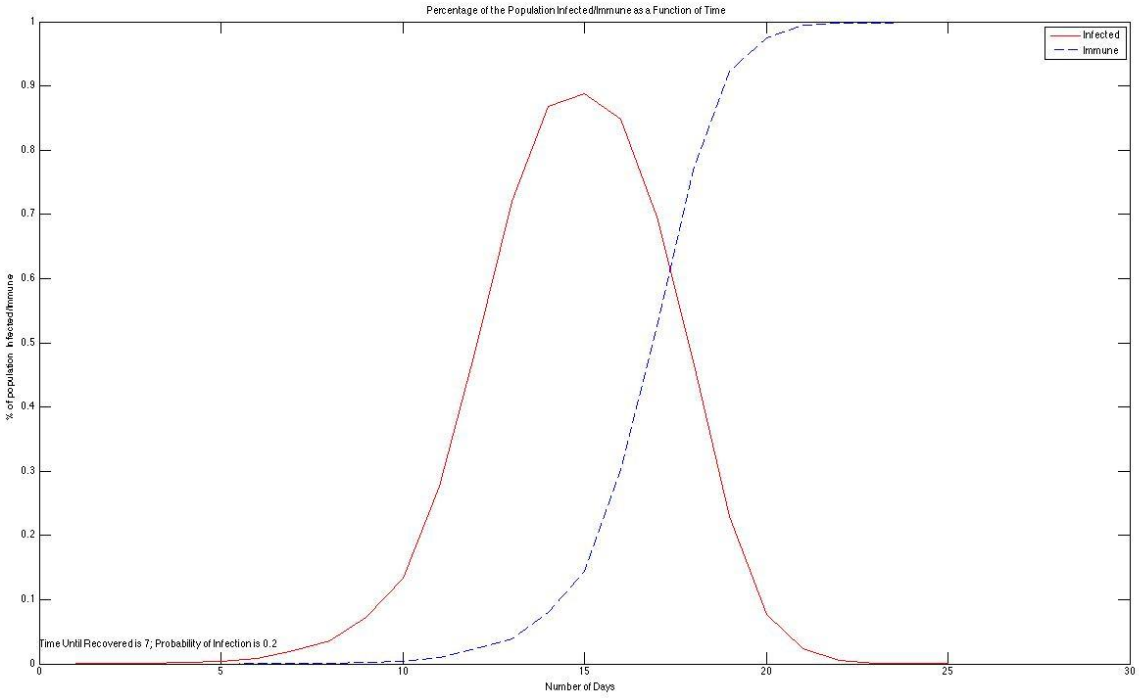
**standard (Time scale to 40)**




**Chronic: (Time scale to 300)**

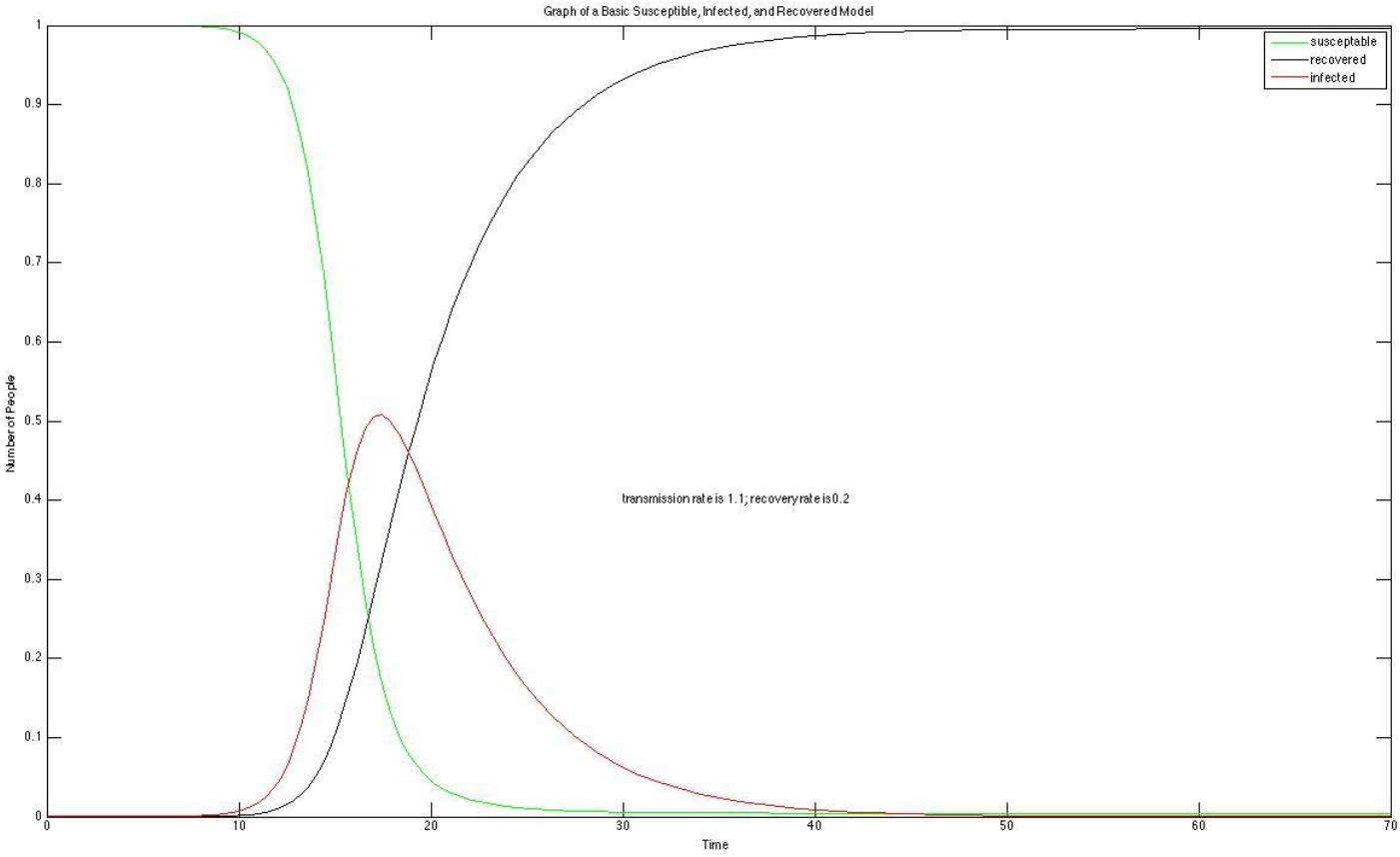



**Acute: (the time scale only goes to 25)**

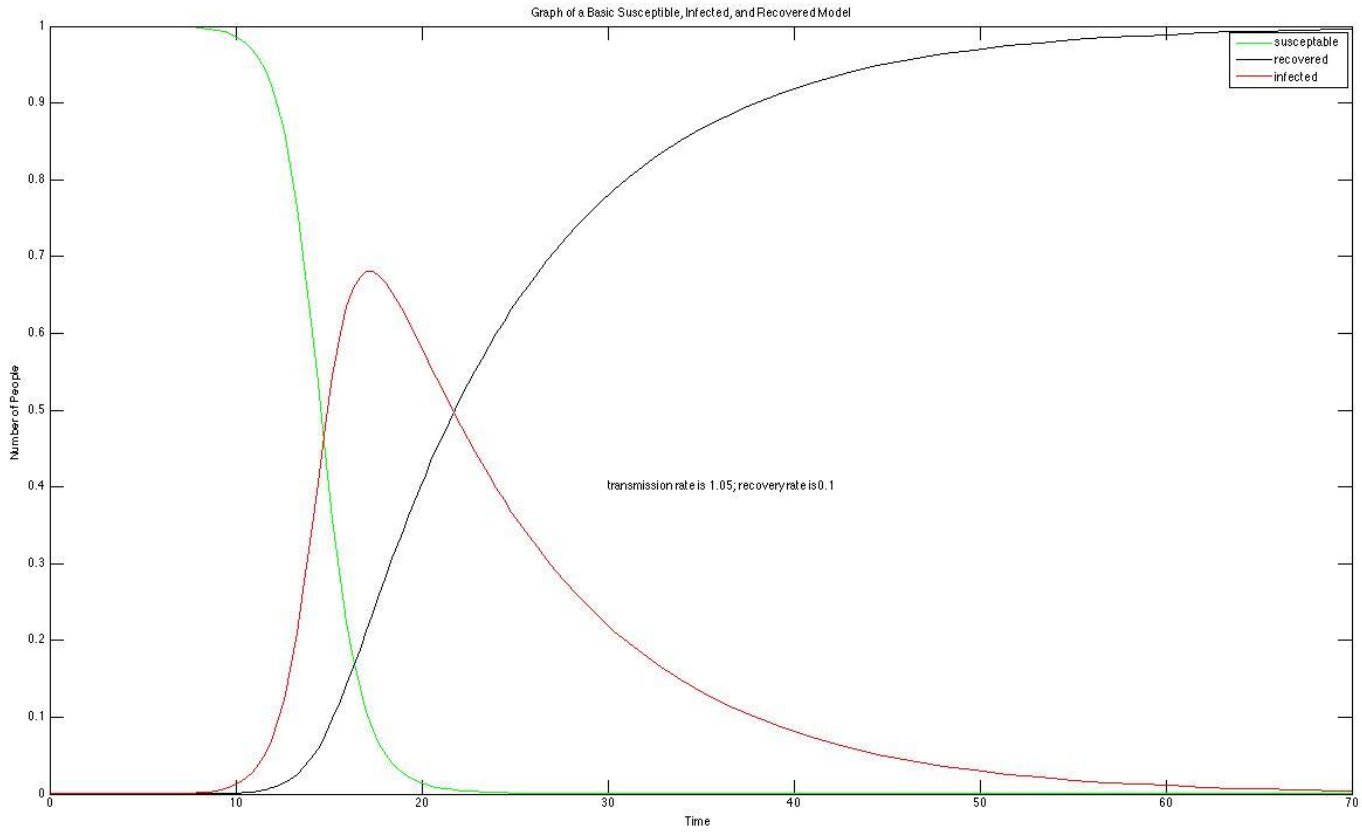



**Compare to SIR model:_____**

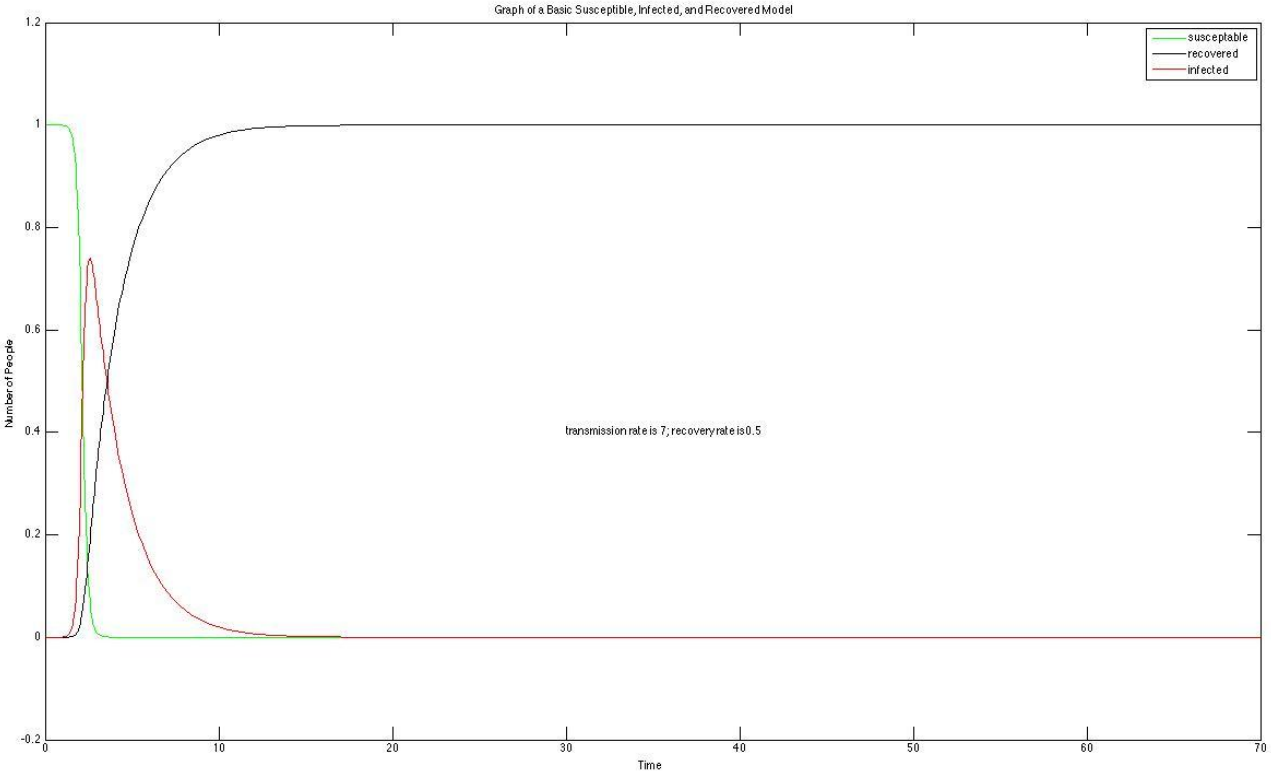
**standard**


Graph of a Basic Susceptible, Infected, and Recovered Model
susceptable
recovered
infected
transmission rate is 1.1; recovery rate is 0.2
Number of People
Time
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
10
20
30
40
50
60
70

## Chronic


Graph of a Basic Susceptible, Infected, and Recovered Model
susceptable
recovered
infected
transmission rate is 1.05; recovery rate is 0.1
Number of People
Time
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
10
20
30
40
50
60
70

## Acute




## Discussion

We derive two important observations from our results:

1. Acute diseases infect a larger percentage of a given population
2. Chronic diseases remain in a given population for a longer period of time than standard diseases.

These two observations can be derived based on either of the two models.

| Type of Disease | Time for Disease to "Die Out" (days) | | Maximum % of Population Infected | |
|---|---|---|---|---|
| | Our Model | SIR Model | Our Model | SIR Model |
| Standard | 40 | ~50 | 48% | 52% |
| Chronic | 300 | ~70-100 | 24% | 70% |
| Acute | 25 | ~20 | 87% | 78% |

Discrepancies between Our Model and the SIR model:

- A chronic disease will die out much sooner in the SIR model than in our model. SPECULATE WHY...Granted, SIR models are geared towards acute diseases (cite the book)
- Chronic diseases infect a much higher percentage of the population at any given time in the SIR model than in our model.
-

Limitations of Our Model:

- It does not take into account genetic, hygienic, or other predispositions that introduce variations in the susceptibility of each individual. We abbreviated risk factors for contracting a disease (such as poor hygiene) into one parameter (probability of infection), and set that constant for each person. Each individual is basically a clone.
- It attempts to distribute members of the population randomly, but movement within a dorm is not necessarily random. Roommates, floormates, and dorm-mates all exhibit varying frequencies of interaction.
- Only one disease is considered. Secondary infections are not considered, even though they may be more likely once a person has becom
- Not all infecteds will take the same amount of time to recover in the real world. Even the SIR model only considers the mean rate of recovery, which is what our model does as well.
- Our model only puts those "touching" a sick person at risk of infection, that means the 8 people bordering a sick cell have a chance of becoming sick. Individuals farther than one cell away are not considered. However, susceptibles may be in double or triple jeopardy, at times, of becoming sick if they border more than one sick person.

## Discussion part 2: Survey

We compare our survey results to both the SIR model and our model, which allows us to derive the predominant disease type that infect students on campus.

## Supplementary Information

| *Infectious Disease* | *Host* | *Estimated $R_0$* | *Reference* |
|---|---|---|---|
| FIV | Domestic Cats | 1.1–1.5 | Smith (2001) |
| Rabies | Dogs (Kenya) | 2.44 | Kitala et al. (2002) |
| Phocine Distemper | Seals | 2–3 | Swinton et al. (1998) |
| Tuberculosis | Cattle | 2.6 | Goodchild and Clifton-Hadley (2001) |
| Influenza | Humans | 3–4 | Murray (1989) |
| Foot-and-Mouth Disease | Livestock farms (UK) | 3.5–4.5 | Ferguson et al. (2001b) |
| Smallpox | Humans | 3.5–6 | Gani and Leach (2001) |
| Rubella | Humans (UK) | 6–7 | Anderson and May (1991) |
| Chickenpox | Humans (UK) | 10–12 | Anderson and May (1991) |
| Measles | Humans (UK) | 16–18 | Anderson and May (1982) |
| Whooping Cough | Humans (UK) | 16–18 | Anderson and May (1982) |

**Box 2.1 The Transmission Term**

Here, we derive from first principles the frequency-dependent (mass action) transmission term (also called proportionate mixing; Anderson and May 1992), which is commonly used in epidemic models. It assumes homogenous mixing in the population, which means everyone interacts with equal probability with everyone else; it discards possible heterogeneities arising from age, space, or behavioral aspects (see Chapters 3 and 7).

Consider a susceptible individual with an average $\kappa$ contacts per unit of time. Of these, a fraction $I = Y/N$ are contacts with infected individuals (where $Y$ is the *number* of infectives and $N$ is the total population size). Thus, during a small time interval (from $t$ to $t + \delta t$), the number of contacts with infecteds is $(\kappa Y/N) \times (\delta t)$. If we define $c$ as the probability of successful disease transmission following a contact, then $1 - c$ is the probability that transmission does not take place. Then, by independence of contacts, the probability (denoted by $1 - \delta q$) that a susceptible individual escapes infection following $(\kappa Y/N \times \delta t)$ contacts is

$$1 - \delta q = (1 - c)^{(\kappa Y/N)\delta t}.$$

Hence, the probability that the individual is infected following any of these contacts is simply $\delta q$.

We now define $\beta = -\kappa \log(1 - c)$ and substitute into the expression for $1 - \delta q$, which allows us to rewrite the probability of transmission in a small time interval $\delta t$ as

$$\delta q = 1 - e^{-\beta Y \delta t/N}.$$

To translate this probability into the *rate* at which transmission occurs, first we expand the exponential term (recalling that $e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} \ldots$), divide both sides by $\delta t$, and take the limit of $\delta q/\delta t$ as $\delta t \to 0$. This gives:

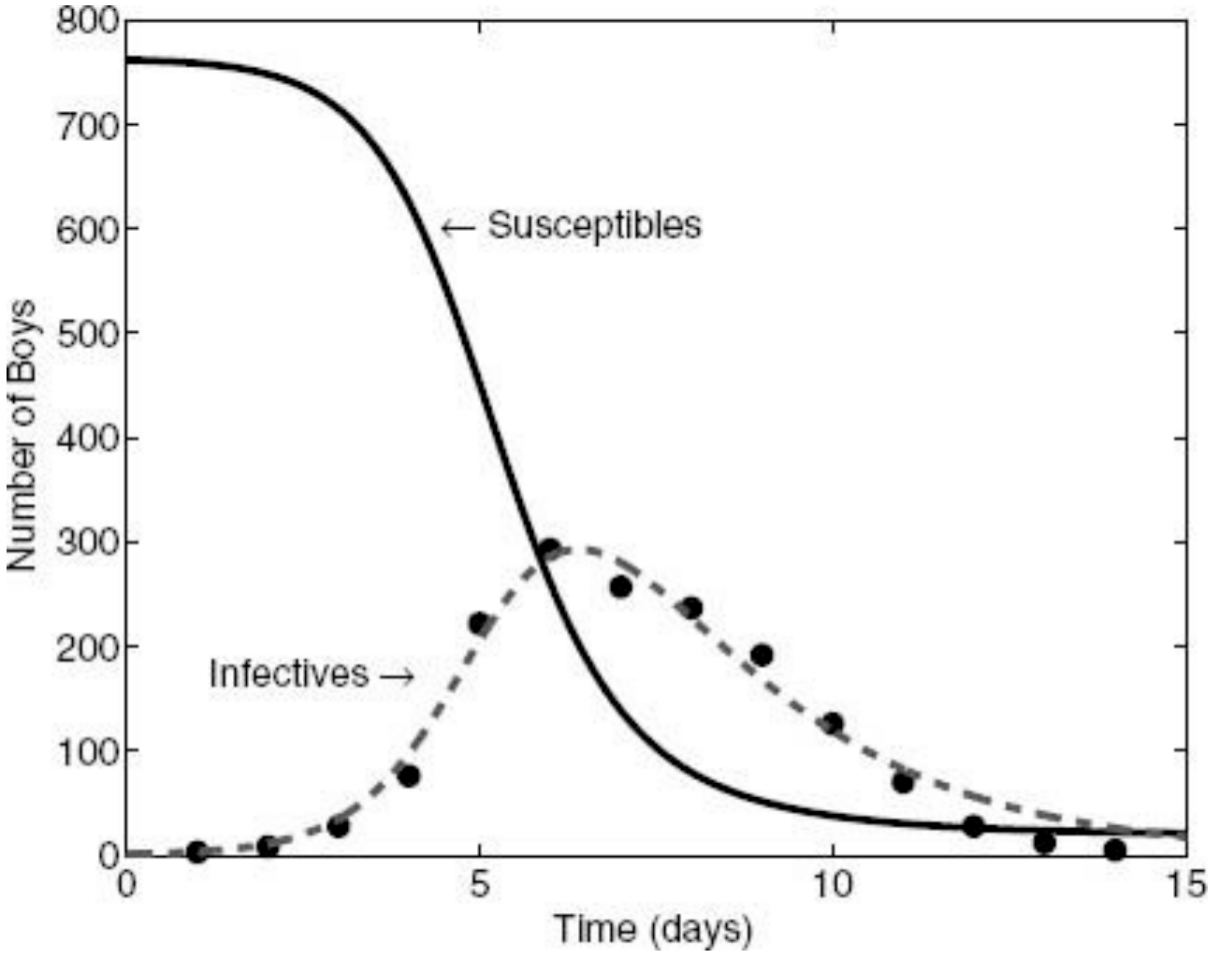
$$\frac{dq}{dt} = \beta Y/N,$$

which is the transmission rate per susceptible individual. In fact, this quantity is often represented by $\lambda$ and referred to as the "force of infection"—it measures the per capita probability of acquiring the infection (Anderson and May 1991). Then, by extention, the total rate of transmission to the entire susceptible population is given by

$$\frac{dX}{dt} = -\lambda X = -\beta XY/N,$$

where $X$ is defined as the *number* of susceptibles in the population. If we rescale the variables (by substituting $S = X/N$ and $I = Y/N$) so that we are dealing with *fractions* (or densities), the above equation becomes

$$\frac{dS}{dt} = -\beta IS.$$


800
700
600
500
400
300
200
100
0
Number of Boys
← Susceptibles
Infectives →
0
5
10
15
Time (days)

### 2.1.1.3. *Worked Example: Influenza in a Boarding School*

An interesting example of an epidemic with no host demography comes from an outbreak of influenza in a British boarding school in early 1978 (Anon 1978; Murray 1989). Soon after the start of the Easter term, three boys were reported to the school infirmary with the typical symptoms of influenza. Over the next few days, a very large fraction of the 763 boys in the school had contracted the infection (represented by circles in Figure 2.4). Within two weeks, the infection had become extinguished, as predicted by the simple $SIR$ model without host demography.

We can get an understanding into the epidemiology of this particular strain of influenza A virus (identified by laboratory tests to be A/USSR/90/77 (H1N1)) by estimating the parameters for the $SIR$ model from these data. Using a simple least squares procedure (minimizing the difference between predicted and observed cases), we find the best fit parameters yield an estimated active infectious period $(1/\gamma)$ of 2.2 days and a mean transmission rate $(\beta)$ of 1.66 per day. Therefore, the estimated $R_0$ of this virus during this epidemic is $\beta/\gamma = 1.66 \times 2.2$, which is 3.65. As shown in Figure 2.4, model dynamics with these parameters is in good agreement with the data. Note, however, that as pointed out by Wearing et al. (2005), the precise value of $R_0$ estimated from these data is substantially affected by the assumed model structure (in Section 3.3 we deal with this issue in more detail).

Sources

someone cite the book with ISBM: 978-0-691-11617-4

paper:

**Biological Sciences - Medical Sciences:**Christophe Fraser, Steven Riley, Roy M. Anderson, and Neil M. Ferguson. **Factors that make an infectious disease outbreak controllable**PNAS 2004 101 (16) 6146-6151; published ahead of print April 7, 2004, doi:10.1073/pnas.0307506101