Mathematical Modeling of Diseases: Susceptible-Infected-Recovered (SIR) Model

Teri Johnson

Math 4901 Senior Seminar
Advisor: Peh Ng
Second Reader: Barry McQuarrie
University of Minnesota, Morris
Spring 2009

Abstract:
In this paper we will look at the SIR model for the mathematical modeling of diseases. We will discuss the mathematics behind the model and various tools for judging effectiveness of policies and control methods. We will complete the paper with an example using the infectious disease Varicella, commonly known as the Chicken Pox.

1. Introduction

One of the most basic procedures in the modeling of diseases is to use a compartmental model, in which the population is divided into different groups. The SIR Model is used in epidemiology to compute the amount of *susceptible, infected*, and *recovered* people in a population. It is also used to explain the change in the number of people needing medical attention during an epidemic. It is important to note that this model does not work with all diseases. For the SIR model to be appropriate, once a person has recovered from the disease, they would receive lifelong immunity. The SIR model is also not appropriate if a person was infected but is not infectious [1,2].

2. S-I-R Model

2.1. Assumptions

The SIR Model is used in epidemiology to compute the amount of susceptible, infected, recovered people in a population. This model is an appropriate one to use under the following assumptions [3]:

1) The population is fixed.
2) The only way a person can leave the susceptible group is to become infected. The only way a person can leave the infected group is to recover from the disease. Once a person has recovered, the person received immunity.
3) Age, sex, social status, and race do not affect the probability of being infected.
4) There is no inherited immunity.
5) The member of the population mix homogeneously (have the same interactions with one another to the same degree).

2.2. SIR Formulas

The model starts with some basic notation:

$S(t)$ is the number of susceptible individuals at time t

$I(t)$ is the number of infected individuals at time t

$R(t)$ is the number of recovered individuals at time t

N is the total population size

The assumptions lead us to a set of differential equations.

$$\frac{dS}{dt} = -\beta S(t)I(t) \tag{1}$$

$$\frac{dI}{dt} = (\beta S(t) - k)I(t) \tag{2}$$

$$\frac{dR}{dt} = kI(t) \tag{3}$$

where k is the recovery rate (with k greater or equal to zero), α is the probability of becoming infected, γ is the number of people infected person comes in contact with in each period of time on average, β is the average number of transmissions from an infected person in a time period (with β greater or equal to zero), and

$$S(t) + I(t) + R(t) = N \tag{4}$$

From these equations [3,4], we can discover how the different groups will act as $t \to \infty$. We can see from equation (1), that the susceptible group will decrease over time and approach zero. From equation (3), we know that the recovered group increase and will approach N over time. How the infected group behaves is more complicated. We start by taking the integral of equations (3) from 0 to t, which gives us

$$k\int_0^0 I(s)ds = R(t)\,. \tag{5}$$

We then manipulate equation (4) to get

$$R(t) = N - S(t) - I(t)\,. \tag{6}$$

By combining equations (5) and (6) we get

$$k\int_0^t I(s)ds = R(t) = N - S(t) - I(t) \tag{7}$$

When we take the integral from zero to infinity of right hand side i.e. $k\int_0^\infty I(s)ds$, that this integral is less than infinity, since the amount of people in a group must be finite. By combining this integral with equation (7), we get that as t goes to infinity

$$I(t) \to N - S(\infty) - k\int_o^\infty I(t)ds\,.$$

Since $S(\infty)$ goes to zero, and $k\int_0^\infty I(s)ds$, which is equal to $R(\infty)$, goes to zero. Thus as t goes to infinity $I(t) \to 0$ as given by Mimmo Iannelli [5].

The rate of change of the infected group is not always negative or zero as it is in the susceptible group, nor is strictly positive or zero like the recovered group. Whether the rate of change is positive or negative depends on k, β, and $S(t)$. We can see from equation (2) when $\beta S(t)$ is less than k then the rate of change for the infected group is

negative. If $\beta S(t)$ is greater than k then the rate of change for the infected group is positive. Finally, if $\beta S(t)$ is equal to k, then the rate of change for the infected group is zero.

By applying Euler's method of systems, we can solve the differential equations. The solutions to the differential equations are:

$$S_{n+1} = S_n - \beta S_n I_n \Delta t \quad (8)$$

$$I_{n+1} = I_n(1 + \beta S_n - k)\Delta t \quad (9)$$

$$R_{n+1} = R_n + kI_n \Delta t \quad (10)$$

where S_{n+1}, I_{n+1}, and R_{n+1} are the number of susceptible, infected and recovered people at time (n+1). Δt is a small change in time, and will be equal to one from now on[6]. It is important to note, that researchers and health officials first collect data on who is in what group at a given period of time. The amount of people in a group does not come from equations (8),(9),and (10). These equations are primarily used to calculate β and k.

The recovered group includes people who receive life-time immunity, however it does not specify if the person is alive with life-time immunity or dead. We can therefore replace equation (3), with one equation for people who that live and one equation for those who die. To do this we actually start by splitting the recovery rate, k, into two recovery rates. These rates are k_V (the recovery rate for those who live) and k_D (the recovery rate for those that die). We now can replace the rate of change for the recovered group. In its place, we have two equations, one for death (D_t) and the other for immunity (V_t) [6]. Specifically

$$\frac{dV}{dt} = k_V I(t)$$

$$\frac{dD}{dt} = k_D I(t)$$

Using Euler's method for systems, the solutions to the above equations become

$$V_{n+1} = V_n + kI_n \Delta t$$

$$D_{n+1} = D_n + (1-k)I_t \Delta t$$

where V_{n+1} and D_{n+1} are the number of immune and dead people at time (n+1) Δt. Again, we will have Δt be equal to one.

2.3. Basic Reproductive Ratio

An important part of modeling diseases is the *Basic Reproductive Ratio*, denoted as B_R. The Basic Reproductive Ratio is important since it tells us if a population is at risk from a disease. B_R is affected by the infection and removal rates, i.e. β, k, and is obtained by $B_R = \frac{\beta}{k} S_0$. When $B_R > 1$, the occurrence of the disease will increase. When $B_R < 1$, the occurrence of the disease will decrease and the disease will eventually be eliminated. When $B_R = 1$, the disease occurrence will be constant [7].

The Basic Reproductive Ratio also helps us predict who will not become infected at all. This is done by looking how the SIR model behaves as $t \to \infty$. Mathematicians Kermack and McKendrick came up with the equation $S_\infty = \exp((1 - S_\infty)B_R)$, where S_∞ is the amount of people who will always remain in the susceptible group [8].

2.4. Herd Immunity Threshold

Keely [8] defined the term *Herd Immunity* as the process where "for each person that is vaccinated the risk of infection for the rest of the community decreases." In other words, Herd immunity is when almost everyone has the disease, and there are not enough of those who have not had the disease to cause an epidemic. One purpose of vaccination is to create herd immunity while having the amount of infected people to be very small [8]. The *Herd Immunity Threshold* (H_I) is percentage of the population that needs to be immune to control transmission of a disease, i.e. B_R equal to one. The equation given by Diekmann and Heesterbeek [4] for figuring out the Herd Immunity Threshold is

$$H_t = \frac{B_R - 1}{B_R} = 1 - \frac{1}{B_R}$$

As the amount of vaccinations increase, the herd immunity threshold also increases. By decreasing the amount of susceptible people, the herd immunity threshold decreases.

2.5. Effective Reproductive Number

Effective Reproductive Number, denoted E_R, is the average number of secondary cases generated by an infectious case during the epidemic. To calculate this number, we multiply the basic reproductive ratio by how many people are susceptible at time *t*, that is

$$E_R = B_R \frac{S_t}{N}.$$

The Effective Reproductive Number is important since it helps researchers and health officials determine how effective their policies on controlling the disease are. When $E_R < 1$, the policies concerning the containing of the disease are effective as given by the UC Berkeley School of Public Health [9].

2.6. Control Vaccination Number

The *Control Vaccination Number*, denoted C_V, is the average number of secondary cases generated by an infectious case during epidemic with control measures, i.e. vaccinations. To calculate this number using following formula

$$C_V = B_R[1 - hf] \tag{11}$$

Where *h* is the vaccine efficacy (the effectiveness of the vaccine) and *f* is vaccination coverage (the fraction of the population that has been vaccinated). The goal of researchers and health officials to have $C_V < 1$. With having $C_V < 1$, we are able to calculate the fraction of the people in a population that need to be vaccinated as given by the UC Berkeley School of Public Health [9]. To this we have $C_V = 1$ and use basic algebra to manipulate (11) to get

$$f > \frac{1-(1/B_R)}{h}.$$

3. Varicella Example

3.1. Varicella Background

Varicella, also known as the Chicken Pox, is a common infectious disease featuring an itching rash and red poxes. It is spread person to person, by way of sneezing, coughing, sharing food or drinks, the touching of the fluids from an open sore, and from exposure of five or more minutes. The incubation period of Varicella is about fourteen to sixteen days. A person is infectious from one or two days prior to the onset of poxes until the last pox has crusted over (total about eight days). Varicella is an extremely contagious disease, with the probability of becoming infected 65-85%, and 90% when in close contact [10]. Before the vaccine, there was a high mortality rate. Now that the vaccine is available, the probability of dying from Varicella is .000093 as seen in the Morbidity and Mortality Weekly Report by the Center for Disease Control [10]. However, the vaccination is only 99% effective the first year, and decreases after that.

3.2. SIR Model and Varicella

Varicella is a disease that we can model using the SIR model. The assumptions in section 2.1 are all satisfied. For our example, population is made up of 100 people that mix homogeneously. Since the disease is highly infectious, everyone will eventually become infected. In Tables 1 and 2 are the cases we calculated, where we can see how many people will be in each state at a given period of time. We start with everyone being susceptible to the disease, and then one person suddenly becomes infected. Also a period lasts 8 days. In these cases, we had everyone recover in one period, meaning that the recovery rate, *k*, is equal to one. It is important to note that the equations (8), (9), and (10), were not used to make these cases, as these are just fabricated scenarios.

	State		
Period	S	I	R
0	100	0	0
1	99	1	0
2	35	64	1
3	12	23	65
4	4	8	88
5	1	3	96
6	0	1	99
7	0	0	100

Table 1: The number of cases in each state per period for α =0.65

	State		
Period	S	I	R
0	100	0	0
1	99	1	0
2	15	84	1
3	2	13	85
4	0	2	98
5	0	0	100

Table 2: The number of cases in each state per period for α =0.85

From Tables 1 and 2, we can calculate β. To do this we manipulate equation (8) using $\Delta t = 1$ to get the following

$$\beta = \frac{S_n - S_{n+1}}{S_n I_n}$$

Using this equation we can get the β for each period. These β are shown in the following tables:

Period	Beta
0	0
1	0
2	0.646465
3	0.010268
4	0.028986
5	0.09375
6	0.333333
7	0
Average	0.1391

Table 3: The different $\beta's$ for each period and the average β and α =0.65

Period	Beta
0	0
1	0
2	0.848485
3	0.010317
4	0.076923
5	0
Average	0.155954

Table 4: The different $\beta's$ for each period and the average β and α =0.85

3.3. The Effects of Infectious Rate and the amount of Initial Infectious persons

One of the most important part of disease modeling is the infectious rate. Varicella's infection rate is somewhere between 65-85% as calculated by Debby Golonka [11]. This number affects the amount of people in the susceptible, infected, and recovered groups, and how long it takes until everyone that will get the disease, recovers from it. The next set of graphs shows how the infectious rates affects the amount of people in the susceptible, infected, and recovered groups, controlling for initial amount of people that are infected for our two cases ($\alpha = 0.65$ and $\alpha = 0.85$).

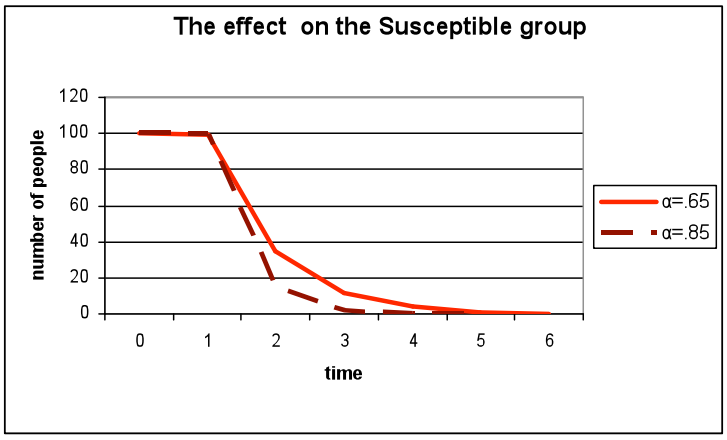

Figure 1: The effect of the infectious rate on the susceptible group

We can see from Figure 1, the population with higher alphas, the amount of susceptible people decreases faster than that of a smaller alpha.

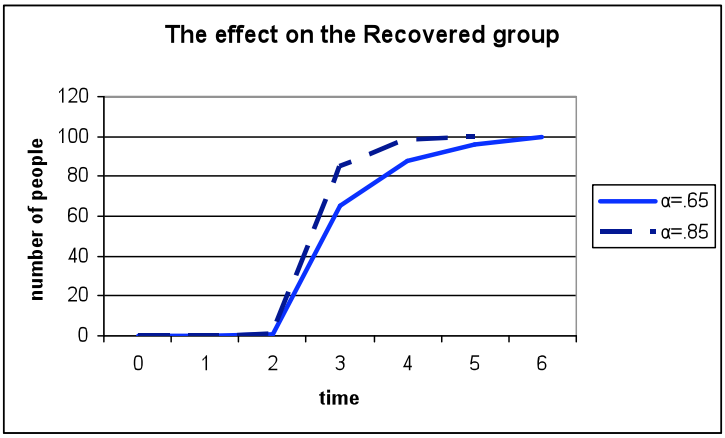

Figure 2: The effect of the infectious rate on the recovered group

We can see from Figure 2, the population with the higher the alpha, the recovered group increases sooner than that of a small alpha.

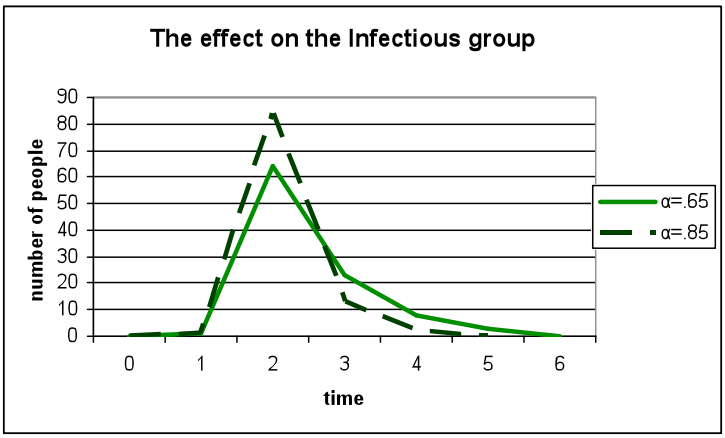

Figure 3: The effect of the infectious rate on the infectious group

We can see from Figure 3, the population with the higher alpha has a higher peak, i.e. when the infectious group's population reaches its peak, there is more in the infectious group with α at .85, than that of a smaller α. We can also see that it takes longer for the infectious population to reach zero with the smaller α.

Another important factor in disease modeling is the amount of people infected initially. The next series of graphs shows how this number affects the amount of people in the susceptible, infected, and recovered groups, while keeping the infectious rate at .65 for our two populations.

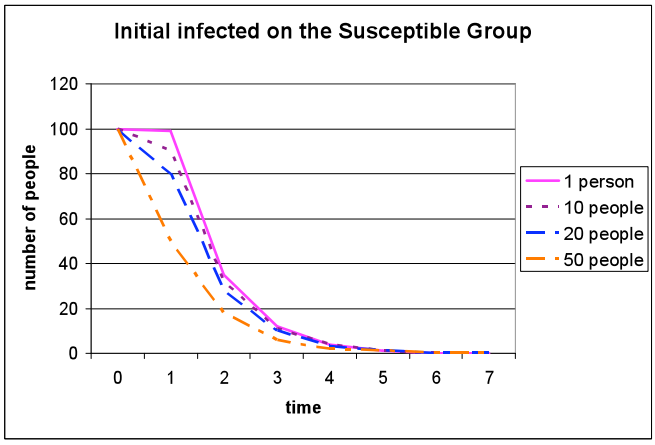

Figure 4: The effect of the amount of initial infectious persons on the susceptible group

From Figure 4, we can see that with the increase in the people who are initially infected, the time it takes for the susceptible to converge is less. As the initial amount of people infected increases, the line showing the population becomes more curved and less jagged.

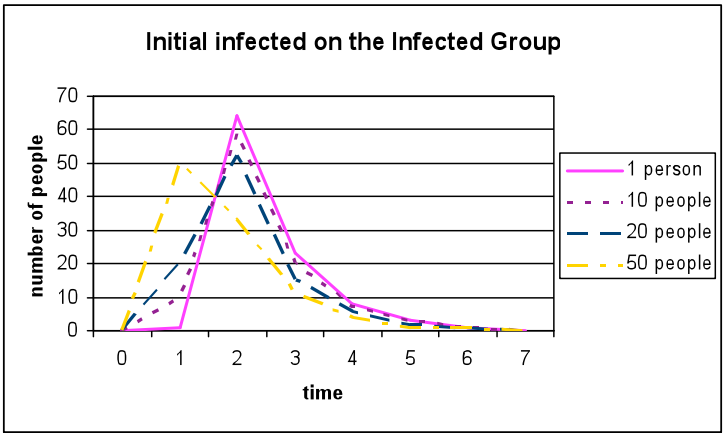

Figure 5: The effect of the amount of initial persons on the infected group

From Figure 5, we can as we increase the amount of initial infected people, the faster the infected group goes to zero. It is interesting that as we decrease the initial amount of infected people, the peak is increased. In addition, when the initial amount of infected people, is half the population, the infected group's peak comes before the peaks of the groups where the initial amount of infected people is less than half the population.

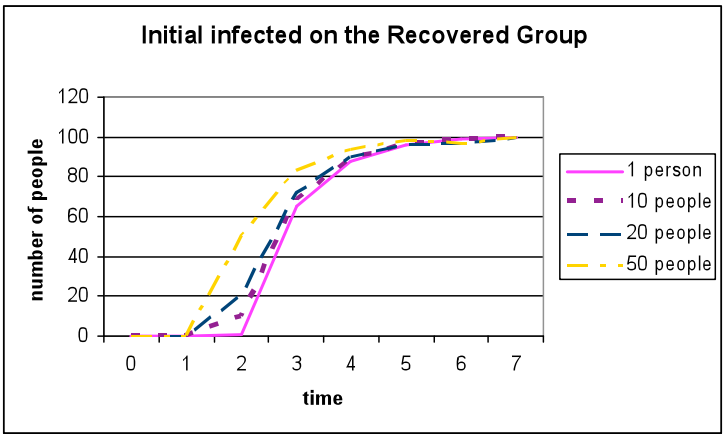

Figure 6: The effect of the amount of initial persons on the recovered group

As with the susceptible group, we can see from Figure 6 when we increase the amount of people who are initially infected that the time it takes for the recovered group to converge, is less. As the initial amount of people infected increases, the line showing the population becomes more curved and less jagged.

3.4. Varicella's Basic Reproductive Ratio

Since we now know β for our outbreak of Varicella, we can calculate the Basic Reproductive Ratio. When the probability of becoming infected is 65%, $B_R = \frac{\beta}{k} S_0 = \frac{.1391}{1} * 100 = 13.91$. When the probability of becoming infected is 85%, $B_R = \frac{\beta}{k} S_0 = \frac{.15595}{1} * 100 = 15.595$. In general for Varicella, B_R is usually between ten and twelve [13]. The reason that our numbers are different from the actual B_R, is because ours are from just 2 cases, while the actual B_R's calculated from many different cases. As we can see from our B_R's (and from the actual B_R's), that $B_R > 1$. Thus the disease can not be eliminated (though in our case no more people can become infected).

3.5. Varicella's Herd Immunity Threshold

Since we now know the Basic Reproductive Ratio, we can calculate the Herd Immunity Threshold, H_I. From our cases, when the probability of becoming infected is 65%, $H_I = \frac{B_R - 1}{B_R} = \frac{13.91 - 1}{13.91} = .928$. When the probability of becoming infected is 85%, $H_I = \frac{B_R - 1}{B_R} = \frac{15.595 - 1}{15.595} = .936$. If we use the established basic reproductive ratios for Varicella, we can calculate the herd immunity threshold to be 0.90-0.9167.

3.6. Varicella's Effective Reproductive Number

We can calculate the Effective Reproductive Number, E_R, for our two cases. We get the following tables (Tables 5 and 6) that show the Effective Reproductive Number at each period

t	S	E_R
0	100	13.91
1	99	13.7709
2	35	4.8685
3	12	1.6692
4	4	0.5564
5	1	0.1391
6	0	0
7	0	0

Table 5: E_R for each period for α =.65

t	S	E_R
0	100	15.595
1	99	15.43905
2	15	2.33925
3	2	0.3119
4	0	0
5	0	0

Table 6: E_R for each period for α =.85

We notice E_R<1, at period 4 when the probability of becoming infected is 0.65, and period 3 when the probability of becoming infected is 0.85. This means that any policies that were implemented were effective. We can calculate the maximum fraction of susceptible people for E_R<1 in general for Varicella. When B_R=10, then $\frac{S_t}{N} = \frac{E_R}{B_R} = \frac{1}{10} = 10\%$. That is, less than ten percent of a population would be susceptible for E_R<1 and implemented polices for control are effective. When B_R=12, then $\frac{S_t}{N} = \frac{E_R}{B_R} = \frac{1}{12} = .08\overline{3}$. That is, less than 8.3 percent of a population would be susceptible for E_R<1 and implemented polices for control are effective.

3.7. Varicella's Control Vaccination Number

We can calculate the Control Vaccination Number, C_V, for Varicella. Research has shown that that the vaccine has 99% effectiveness in the first year, and after eight years the effectiveness drops to 87% [12]. The vaccination coverage for Varicella among teenagers in 2007 was 75.7% for one dose and 18.8% for the second dose [14]. We have calculated C_V for various B_R and doses for different vaccine efficiency as shown in Tables 7 and 8.

B_R	C_V, 1 dose	C_V, 2 doses
10	2.5057	8.1388
11	2.75627	8.95268
12	3.00684	9.76656

Table 7: C_V for different B_R and doses when *h*=0.99

B_R	C_V 1 dose	C_V 2doses
10	3.4141	8.3644
11	3.75551	9.20084
12	4.09692	10.03728

Table 8: C_V for different B_R and doses when *h*=0.87

In Tables 7 and 8, we can see that among teenagers in 2007, researchers and health officials have not reached their goal of having $C_V < 1$.

We can calculate what the coverage would need to be in order to have $C_V < 1$. This coverage would be the same for one and/or two doses. This is shown in the following tables (Tables 9 and 10).

B_R	f
10	0.90909
11	0.91827
12	0.925926

Table 9: Vaccine Coverage needed for various B_R and h=0.99

B_R	f
10	1.03448
11	1.0449
12	1.05393

Table 10: Vaccine Coverage needed for various B_R and h=0.87

We can see from Tables 9 and 10 that when the effectiveness is 99%, then between 90.9-92.6% of the population need to be vaccinated in order for $C_V < 1$. When the effectiveness is 87%, then between 103.4%-105.4% of the population need to be vaccinated in order for $C_V < 1$. However, it is not possible to vaccinate over 100% of the population. Therefore when the vaccination is 87% effective the researcher's and health officials are unable to meet their goal of having $C_V < 1$.

We can calculate what the effectiveness needs to be for various vaccination coverage percentages and still have $C_V < 1$. This is shown on the following table.

	Vaccination Coverage									
B_R	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%
10	9	4.5	3	2.25	1.8	1.5	1.286	1.125	1	0.9
11	9.090	4.545	3.030	2.273	1.182	1.515	1.299	1.136	1.010	0.909
12	9.166	4.583	3.056	2.292	1.833	1.528	1.310	1.146	1.019	0.917

Table 11: The effectiveness needed to be for various vaccination coverage and still have $C_V < 1$

We can see from Table 11, no matter the Basic Reproductive Ratio is, the coverage needs to be 90% or more to have the vaccination efficiency be possible. The following table (Table 12) gives what the coverage needs to be in order to have $C_V < 1$ and 100% efficiency for various Basic Reproductive Ratios.

B_R	vaccination coverage for 100% vaccination efficiency
10	90%
11	90.90909%
12	91.66667%

Table 12: The coverage needed to be in order to have $C_V < 1$ and 100% efficiency for various Basic Reproductive Ratios

Thus in order to have a $C_V < 1$ with 100% vaccination effectiveness, we would only need to vaccinate 90-91.666% of the population.

4. Conclusion

The SIR Model is used in the modeling of infectious diseases by computing the amount of people in a closed population that are *susceptible, infected*, or *recovered* at a given period of time. The model is also used by researchers and health officials to explain the increase and decrease in people needing medical care for a certain disease during an epidemic. From numbers generated by the SIR model researcher's health officials can calculate different numbers that allow them to see if policies are effective and if occurrence of the disease is increasing, decreasing, or stable. We have used Varicella as an example of how the SIR model works. However, the SIR model has some serious disadvantageous. The population has to be fixed and the population needs to mix homogeneously. The model does not take into account any variation in the disease among people of different sexes, races, or ages.

The SIR model is the basis for other similar models. The SI model, also known as the SIS model, is the model where once a person is no longer infectious, this person becomes susceptible once again. The common cold can be modeled with the SI model. There is also the SEIR model, where people are categorized as *susceptible, exposed, infected*, or *recovered.* The SIR model can be adjusted to include variation due to seasonal changes as seen by Bauch and Earn [2].

5. References

1. Bauch, Chris, and David Earn. "Interepidemic Intervals in Forced and Unforced SEIR models." Dynamical Systems and Their Applications in Biology. Ed. Shigui Ruan, Gail S. Wolkowicz and Jianhong Wu. New York: American Mathematical Society, 2003. 33-43.

2. Rhodes, John A., and Elizabeth S. Allman. Mathematical Models in Biology : An Introduction. New York: Cambridge UP, 2003. 280-301.

3. Smith, David, and Lang Moore. "The SIR Model for Spread of Disease." MathDL. Dec. 2001. MMA. Fall 2008 < http://www.math.duke.edu/education/ccp/materials/diffcalc/sir/contents.html>.

4. Diekmann, O., and J. A. Heesterbeek. Mathematical Epidemiology of Infectious Diseases : Model Building, Analysis and Interpretation. New York: John Wiley & Sons, Incorporated, 2000.

5. Iannelli, Mimmo. "The Mathematical Modeling of Epidemics." Mathematical Models in Life Science: Theory and Simulation. 1 July 2005. Florida Gulf Coast University. <http://itech.fgcu.edu/faculty/pfeng/teaching/epidemics.pdf>.

6. Getz, Wayne M., and James O. Lloyd-Smith. "Basic Methods for Modeling the Invasion and Spread of Contagious Diseases." DIMACS Series in Discrete Mathematics and Theoretical Computer Science (2005): 1-23.

7. Hackborn, Bill. "Susceptible, Infected, Recovered: the SIR model of an Epidemic." University of Alberta: Augustana. Fall 2008 <http://www.augustana.ab.ca/~hackw/mat332/exhibit/sir.ppt>.

8. Keeling, Matt. "The mathematics of diseases." Plus Magaazine: Living Mathematics. Mar. 2001. Fall 2008 <http://plus.maths.org.uk/issue14/features/diseases/index.html>.

9. School of Public Health. "Concepts for the Prevention and Control of Microbial Threats." Center for Infectious Diseases and Emergency Readiness. June 2006. University of California Berkeley. <www.idready.org/slides/01epiconceptsII-slides.pdf>.

10. Center for Disease Control. "Summary of Notifiable Diseases --- United States, 2006." Morbidity and Mortaility Weekly Report. 31 Mar. 2008. Center for Disease Control and Prevention. <http://www.cdc.gov/mmwr/preview/mmwrhtml/mm5553a1.htm>.

11. Golonka, Debby."Chickenpox (Varicella)." Web MD. 21 May 2008. Healthwise. <http://www.webmd.com/a-to-z-guides/chickenpox-varicella-topic-overview?page=2>.

12. Peary, Karen. "Study shows drop in effectiveness of chicken pox vaccine." Yale Bulletin and Calender. 27 Feb. 2004. Yale. <http://www.yale.edu/opa/arc-ybc/v32.n20/story3.html>.

13. Anderson, Roy M., Robert M. May, and B. Anderson. Infectious Diseases of Humans Dynamics and Control (Oxford Science Publications). New York: Oxford UP, USA, 1992.

14. "USA Teen Vaccination Coverage Increasing, But Below 2010 Goals." Medical News Today. 13 Oct. 2008. <http://www.medicalnewstoday.com/articles/125241.php>.