An Introduction to Deterministic Infectious Disease Models

Table of Contents

1. Introduction

Deterministic infectious disease models are a powerful tool in epidemiology. They are flexible enough to generate approximations of many types of infectious disease outbreaks yet simple enough to require minimal computational resources. They are easy to understand and therefore helpful to policy makers who are faced with difficult decisions during potentially deadly outbreaks. This aspect of deterministic models makes them sometimes preferred to more complex models with stochasticity or "black box" structures that are difficult to dissect.

Despite their simplicity, deterministic infectious disease models are versatile in their application and have been used to understand the dynamics of many different types of diseases. For one, these models can be very useful at predicting the future spread of an infectious disease early on during an outbreak. For example, Fisman, et al. used a simple disease model to understand the epidemic dynamics of Ebola in 2014 [1]. Understanding disease dynamics can help public health officials respond appropriately to try to ensure community wellbeing. Deterministic infectious disease models can also be used to understand why there are surges in cases of endemic diseases which may otherwise have consistent levels of prevalence. Feng, et al. implemented a deterministic infectious disease model to understand the potential causes of resurging cases of tuberculosis [2]. Their model helped shape the argument for the importance of exogeneous reinfection as a cause of climbing tuberculosis cases. Often, these models influence important decisions in health policy. Granich et al. developed a deterministic model to evaluate the effectiveness of universal voluntary HIV testing on mitigating HIV/AIDS epidemics [3].

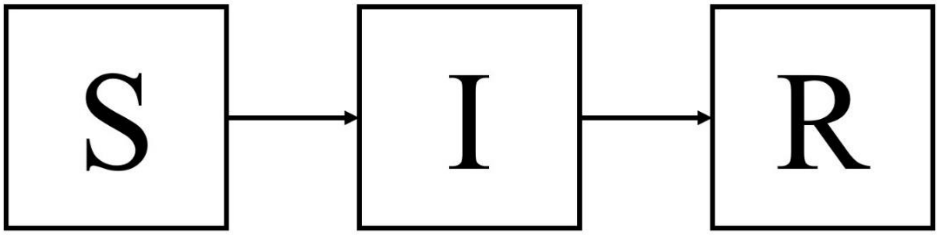
The purpose of this paper is to enable the reader to understand and implement a deterministic infectious disease model and is organized as follows. Section 2 uses the SIR model to explain the concepts and mathematics underlying deterministic infectious disease modeling, including the parameters of the model, the reproduction ratio, and vaccination thresholds to prevent outbreaks. Section 3 explains how to implement an SIR model in Excel, including creating graphs. Section 4 presents other common model structures, such as SI and SEIR models. Section 5 describes the differences between deterministic and stochastic infectious disease models. Section 6 discusses a stochastic simulation model of future Filovirus outbreaks. Section 7 concludes.

2. The SIR Model

The SIR model is the quintessential deterministic infectious disease model first described by Kermack and McKendrick [4] and more recently by Keeling and Rohani [5]. The SIR model is based on the idea that a population during an outbreak can be divided into three groups or compartments: S – susceptible, I – infected, and R – recovered. The susceptible compartment represents the portion of the population that has not yet been exposed nor infected with the disease but could be infected in the future. The infected compartment represents the portion of the population that is both infected with the disease and infectious. The recovered compartment represents the portion of the population that has recovered from infection and developed immunity to reinfection.

The way that individuals transition between compartments in the model is dependent on several factors. Transition from the susceptible to the infected compartment requires disease transmission and is based on the transmission rate and the prevalence of the disease. Transition from the infected to the recovered compartment is based on the recovery rate of infection. These transitions are outlined in Figure 1.

Figure 1. SIR model structure. Disease state transitions are possible from susceptible to infected and from infected to recovered compartments.



Differential equations describe the SIR model mathematically. The equations represent the rate of change of each compartment over time rather than the absolute number of individuals in a compartment at any given time. The SIR model equations are as follows [5]:

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

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

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

The notation $\frac{dX}{dt}$ means the rate of change of function X over time t. In this case, the rate is measured at infinitely small increments of time dt. For example, $\frac{dS}{dt}$ is the instantaneous rate of change of the size of the compartment of susceptible individuals at a given time t. In these equations, N is the size of the total population, β is the transmission rate, and γ is the recovery rate. The equations make the relationships between compartments more obvious. For example, the rate of overall infection is dependent on the number of susceptible individuals and infected individuals at a given time, as well as the fixed transmission rate parameter, β. When you plug in all the numbers, the result of each equation gives units in number of people per time step.

The transmission rate, β, is the product of the average number of contacts per individual per time step and the probability of transmission between a susceptible and infected individual. While the average number of contacts per individual per time step and the probability of transmission between a susceptible and infected individual are sometimes known, the β parameter is more often calibrated using observed incident cases of an infectious disease over time.

The recovery rate, γ, is equivalent to the inverse of the infectious period. For example, if an infectious disease has an infectious period of 5 days, then $\gamma = \frac{1}{5\ days} = 0.2\ days^{-1}$.

The equations reveal several ways to mitigate the spread of an infectious disease. The first of which includes interventions to reduce the transmission rate, β. This could include reducing the contact rate through interventions such as social distancing [6] or reducing the transmission probability of each contact through interventions such as wearing masks [7]. Additionally, given that the rate of decline in the number of infected individuals is directly related to the recovery rate, γ, then treatments aimed at reducing the infectious period of a disease can also help to mitigate the spread of an outbreak [8].

2.1 The Reproduction Ratio

The reproduction ratio is an important measure of the strength of an infectious disease outbreak. The basic reproduction ratio, R_0, is the average number of secondary infections that arise from an average primary infection in a fully susceptible population [5]. Notably, if $R_0 > 1$, then the conditions allow for an outbreak to occur, as the average primary infection leads to more than one additional secondary infection. Conversely, if $R_0 < 1$, then conditions do not allow for an outbreak to occur, as the average primary infection does not lead to an additional secondary infection.

This outbreak threshold, and thus R_0, can be derived mathematically from the SIR equations we described earlier. Another way to think about the threshold for an outbreak to occur is to consider the conditions that must be present for the rate of change of the infected compartment to be positive when the outbreak begins. Mathematically, $\frac{dI}{dt} > 0$ when $t = 0$. For the SIR model:

$$\frac{\beta S(0)I(0)}{N} - \gamma I(0) > 0$$

Simplifying:

$$\frac{\beta S(0)}{N} > \gamma$$

In this case, $S(0) \approx N$ as the population is fully susceptible (save for the primary infection). Therefore, the threshold for an outbreak to occur is re-written:

$$\frac{\beta}{\gamma} > 1$$

Thus, combining the two ways to express the outbreak threshold, for the SIR model, $R_0 = \frac{\beta}{\gamma}$.

The reproduction ratio of an outbreak at a given time, t, is known as the effective reproduction ratio, R_t. The difference between R_0 and R_t is that R_0 describes the outbreak in a fully susceptible population, whereas R_t describes the outbreak in the current conditions at time t. Therefore, for the SIR model $R_t = \frac{\beta S(t)}{\gamma N}$. In other words, as the size of the susceptible compartment decreases, so too does the effective reproduction ratio. This makes intuitive sense, as there are fewer possible susceptible individuals to serve as secondary infections for the average primary infection.

The concept of the reproduction ratio is not limited to the SIR model. It is possible to derive the R_0 and R_t for more complicated models with additional compartments. Diekman et al. developed a method called the next generation method which utilizes linear algebra to derive the R_0 and R_t as functions of the model parameters [9]. Additionally, the R_0 and R_t can be estimated from an observed epidemic curve. Wallinga and Teunis developed a likelihood-based estimation method to approximate R_0 and R_t of an outbreak [10]. This method does not require a formalized model structure and is therefore useful in evaluating R_0 and R_t for emerging disease outbreaks where historical data are sparse.

The R_0 of an outbreak is often one of the first metrics that researchers seek to quantify. For example, estimates of R_0 for COVID-19 range from 2.2-5.7 [11]. Estimates of R_0 for Ebola range from 1.5-1.9 [12]. Estimates of R_0 for smallpox range from 3.5-6 [13]. Finally, estimates of R_0 for measles range from 12-18 [14]. This high basic reproduction ratio is one of the reasons that measles is frequently at the center of debates around vaccination and vaccine hesitancy.

2.2 Vaccination in SIR Models

The SIR model elucidates powerful information about vaccination in a relatively simple way. The basic idea is that the vaccination moves individuals from the susceptible compartment to the recovered compartment by conferring immunity without infection. Using the concept of the effective reproduction ratio, we can determine the level of vaccination required to prevent an outbreak of an infectious disease. The key concept again is to reduce the effective reproduction ratio to below one. The notation for the effective reproduction ratio under vaccination will be $R_{0,vax}$. For the SIR model, $R_{0,vax} = \frac{\beta S(0)}{\gamma N}$. To derive the level of vaccination required to prevent an outbreak, we want $R_{0,vax} < 1$, thus the fraction of the population that can remain susceptible is:

$$\frac{S(0)}{N} < \frac{\gamma}{\beta}$$

When $t = 0$, we assume the population is composed of susceptible $S(0)$ or vaccinated $R(0)$ individuals and that $N \approx S(0) + R(0)$. With substitution:

$$\frac{N - R(0)}{N} < \frac{\gamma}{\beta}$$

After reducing with algebra:

$$\frac{R(0)}{N} > \frac{\beta - \gamma}{\beta}$$

The proportion of the population that needs to be vaccinated to prevent an outbreak from starting must be greater than $\frac{\beta-\gamma}{\beta}$. With further algebra, $\frac{R(0)}{N} > 1 - \frac{1}{R_0}$. Therefore, the portion of the population which should be vaccinated to prevent disease outbreak is at least $1 - \frac{1}{R_0}$.

2.3 The SIR Model with Birth and Death

Demography is a simple addition to the SIR model and is described in detail by Keeling and Rohani [5]. In this model, we assume a natural mortality rate μ where $\frac{1}{\mu}$ is the average lifespan of an individual in the model. For simplicity, in this implementation there are equal mortality rates in each compartment of the model (i.e. infection does not lead to greater mortality rates). Historically, the birth rate, $\delta = \mu N$ so that the overall population remains the same; mathematically, $\left(\frac{dS}{dt} + \frac{dI}{dt} + \frac{dR}{dt} = 0\right)$ [5]. The updated SIR model is as follows:

$$\frac{dS}{dt} = \delta - \frac{\beta SI}{N} - \mu S$$

$$\frac{dI}{dt} = \frac{\beta SI}{N} - (\gamma + \mu)I$$

$$\frac{dR}{dt} = \gamma I - \mu R$$

Using the same method as before, we can again derive R_0 and R_t for an SIR model with birth and death. First, we find the threshold at which an outbreak occurs in a completely susceptible population using $\frac{dI}{dt} > 0$:

$$\frac{\beta S(0)}{N} - \gamma - \mu > 0$$

Simplifying algebraically the conditions required for an outbreak to occur are:

$$\frac{\beta}{\gamma + \mu} > 1$$

Therefore, $R_0 = \frac{\beta}{\gamma+\mu}$ in an SIR model with birth and death. Similarly, $R_t = \frac{\beta S(t)}{(\gamma+\mu)N}$. Thus, when accounting for mortality in an SIR model, R_0 and R_t are lower compared to an SIR model without birth and death. This is an intuitive result because by adding mortality into the model, the overall length of the infectious period is reduced despite holding transmission and recovery constant. Likewise, the proportion of individuals needed to be vaccinated to prevent an outbreak $(1 - \frac{1}{R_0})$ is lower because the value of R_0 is lower.

3. Implementing an SIR Model in Excel

An Excel spreadsheet is a great platform to implement and experiment with an SIR model. The first thing to note is that implementation of the SIR model in Excel will require difference equations rather than differential equations. The SIR difference equations define the rate of change of the susceptible, infected, and recovered populations over very small but finitely small increments of time (Δt), whereas the SIR differential equations define the rate of change of the susceptible, infected, and recovered compartments over an infinitely small increment of time (dt). The difference equations for the SIR model are as follows:

$$\frac{\Delta S}{\Delta t} = \frac{-\beta SI}{N}$$

$$\frac{\Delta I}{\Delta t} = \frac{\beta SI}{N} - \gamma I$$

$$\frac{\Delta R}{\Delta t} = \gamma I$$

Likewise, the difference equations for the SIR model with birth and death are as follows:

$$\frac{\Delta S}{\Delta t} = \delta - \frac{\beta SI}{N} - \mu S$$

$$\frac{\Delta I}{\Delta t} = \frac{\beta SI}{N} - (\gamma + \mu)I$$

$$\frac{\Delta R}{\Delta t} = \gamma I - \mu R$$

If Δt is very small, then the difference equations will be close approximations of the differential equations. This simple modification will greatly reduce the computational complexity of the SIR model and allow for seamless integration into Excel without sacrificing much accuracy.

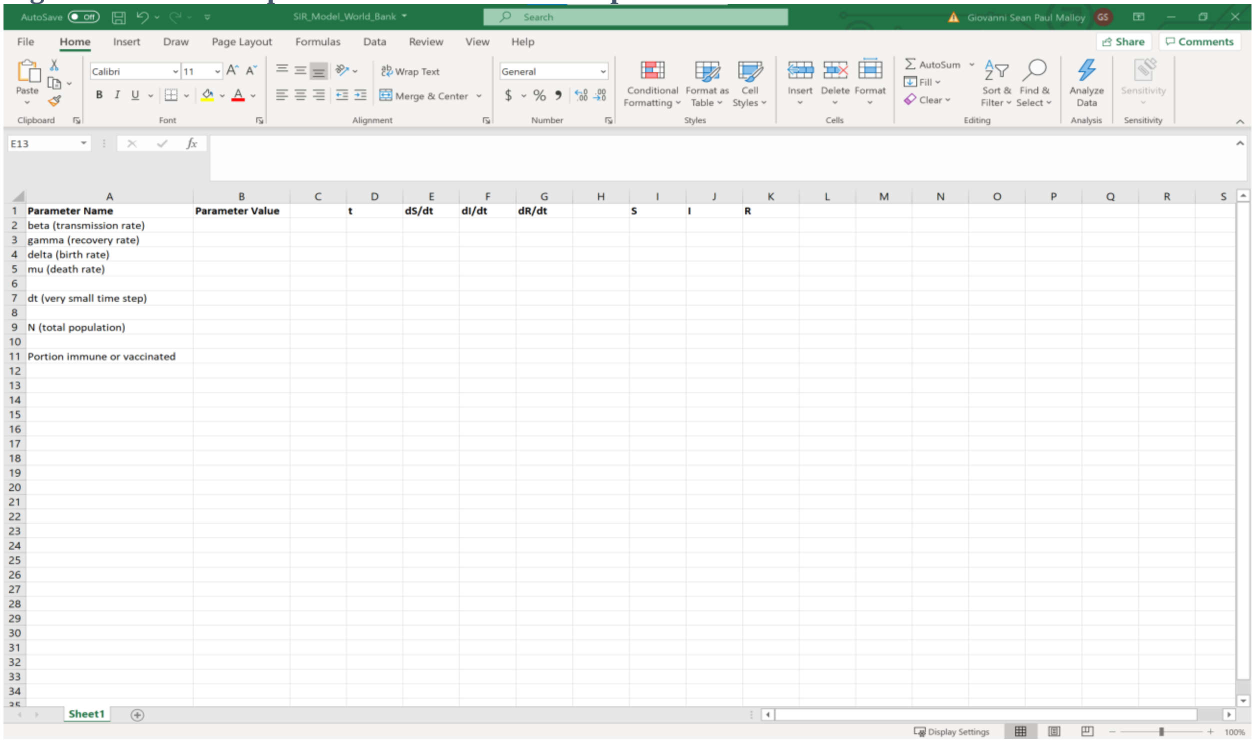
To learn how to implement an SIR model in Excel, we are going to model a new infectious disease called disease X. First, open Excel and set up a spreadsheet to include columns titled "Parameter Name", "Parameter Value", "t", "dS/dt", "dI/dt", "dR/dt", "S", "I", and "R". The first two columns will serve as the user inputs for the model parameters. The other columns will be the model output. Then, list the model parameters under the "Parameter Name" column. These parameters include "beta (transmission rate)", "gamma (recovery rate)", "delta (birth rate)", "mu (death rate)", the time step "dt (very small time step)", "N (total population)", and "Portion immune or vaccinated". The spreadsheet should now look something like Figure 2.

Next, populate the parameter values in column B. Choose a value for β between 0 and 1. Mathematically, β is restricted to this interval because the infection term is normalized to the population when we divide by N in $\frac{-\beta SI}{N}$. In this case, disease X has transmission rate $\beta = 0.74$. Then, populate a value for γ. Remember that the recovery rate is the inverse of infection period. Disease X has an infectious period of 10 days, therefore, $\gamma = 0.1$. For now, set $\delta = 0$ and $\mu = 0$. The very small time step Δt for disease X is 1 day. Finally, model a population of 100,000 individuals with no immunity or vaccination.

After the spreadsheet is setup, it is time to input the initial conditions for disease X. On day 0, there is 1 infected individual.

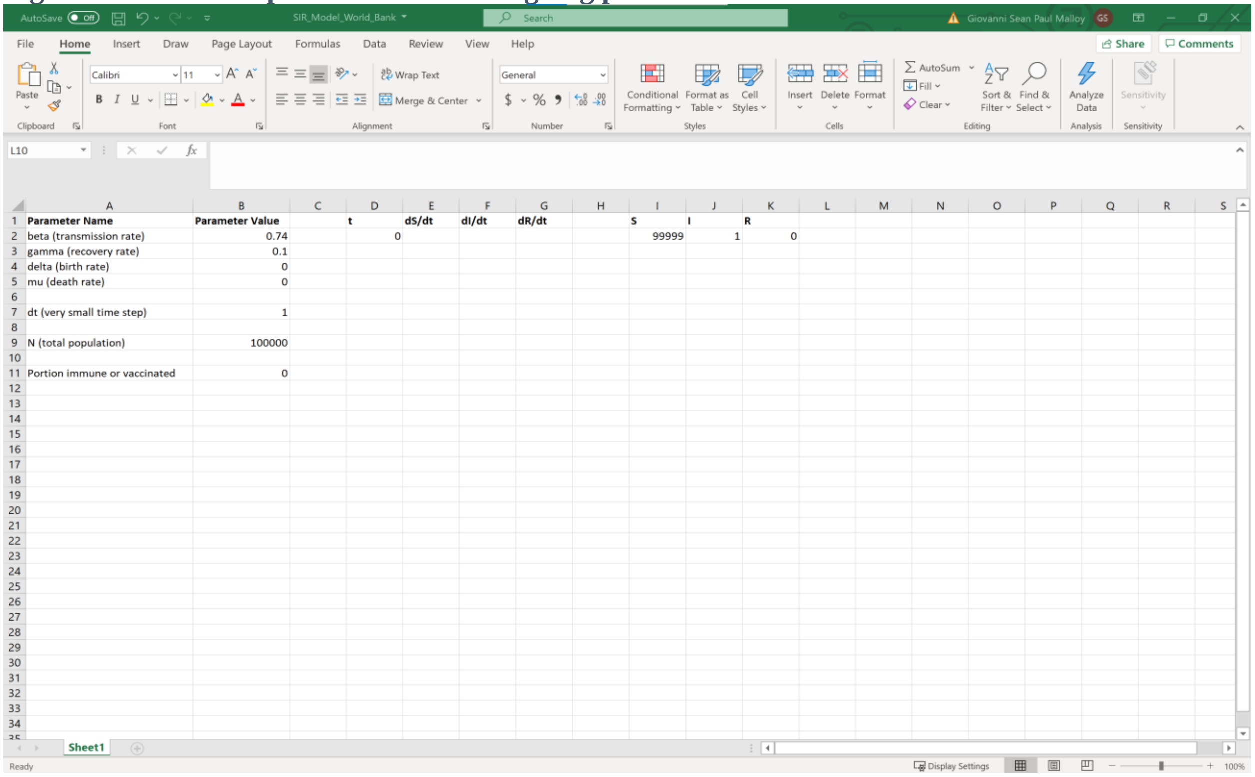
Then, to finish implementing the initial conditions of disease X, use formulas to improve the flexibility of the model. Rather than treating cells I2 and K2 as user inputs, make them formulas relative to the total population. When $t = 0$, the only individuals in the recovered population are either immune or vaccinated which is equal to the product of the total population and the portion of the population which is immune or vaccinated. Populate cell K2 with "=B9*B11" this is equivalent to N times the portion of the population that is immune or vaccinated. Likewise, since $S = N - I - R$, populate cell I2 with "=B9-J2-K2". The spreadsheet should now look like that of Figure 3.

Figure 2. The Excel spreadsheet after initial setup



	A	B	C	D	E	F	G	H	I	J	K
1	**Parameter Name**	**Parameter Value**		**t**	**dS/dt**	**dI/dt**	**dR/dt**		**S**	**I**	**R**
2	beta (transmission rate)										
3	gamma (recovery rate)										
4	delta (birth rate)										
5	mu (death rate)										
6											
7	dt (very small time step)										
8											
9	N (total population)										
10											
11	Portion immune or vaccinated										

Figure 3. The Excel spreadsheet after assigning parameter values



	A	B	C	D	E	F	G	H	I	J	K
1	**Parameter Name**	**Parameter Value**		**t**	**dS/dt**	**dI/dt**	**dR/dt**		**S**	**I**	**R**
2	beta (transmission rate)	0.74		0					99999	1	0
3	gamma (recovery rate)	0.1									
4	delta (birth rate)	0									
5	mu (death rate)	0									
6											
7	dt (very small time step)	1									
8											
9	N (total population)	100000									
10											
11	Portion immune or vaccinated	0									

Now, it is time to implement the SIR model itself. First, implement the difference in time Δt in cell D3 using the formula “=D2+$B$7”. This is the initial time $t = 0$ plus our very small time increment Δt. The cells which are referenced with dollar signs “$” indicate that the reference is “frozen” to that cell. This means that when we drag down the formulas to copy them to other cells, the reference to the Δt parameter will remain the same.

Then, implement $\frac{\Delta S}{\Delta t} = \delta - \frac{\beta SI}{N} - \mu S$ in cell E3 using the formula “=$B$4 - ($B$2*I2*J2)/$B$9 - $B$5*I2”. While the current parameter values for birth rate and death rate are 0, it is nice to implement this feature now rather than later. Next, implement $\frac{\Delta I}{\Delta t} = \frac{\beta SI}{N} - \gamma I - \mu I$ in cell F3 using the formula “=($B$2*I2*J2)/$B$9 - $B$3*J2 - $B$5*J2”. Finally, implement $\frac{\Delta R}{\Delta t} = \gamma I - \mu R$ in cell G3 using the formula “=$B$3*J2 - $B$5*K2”.

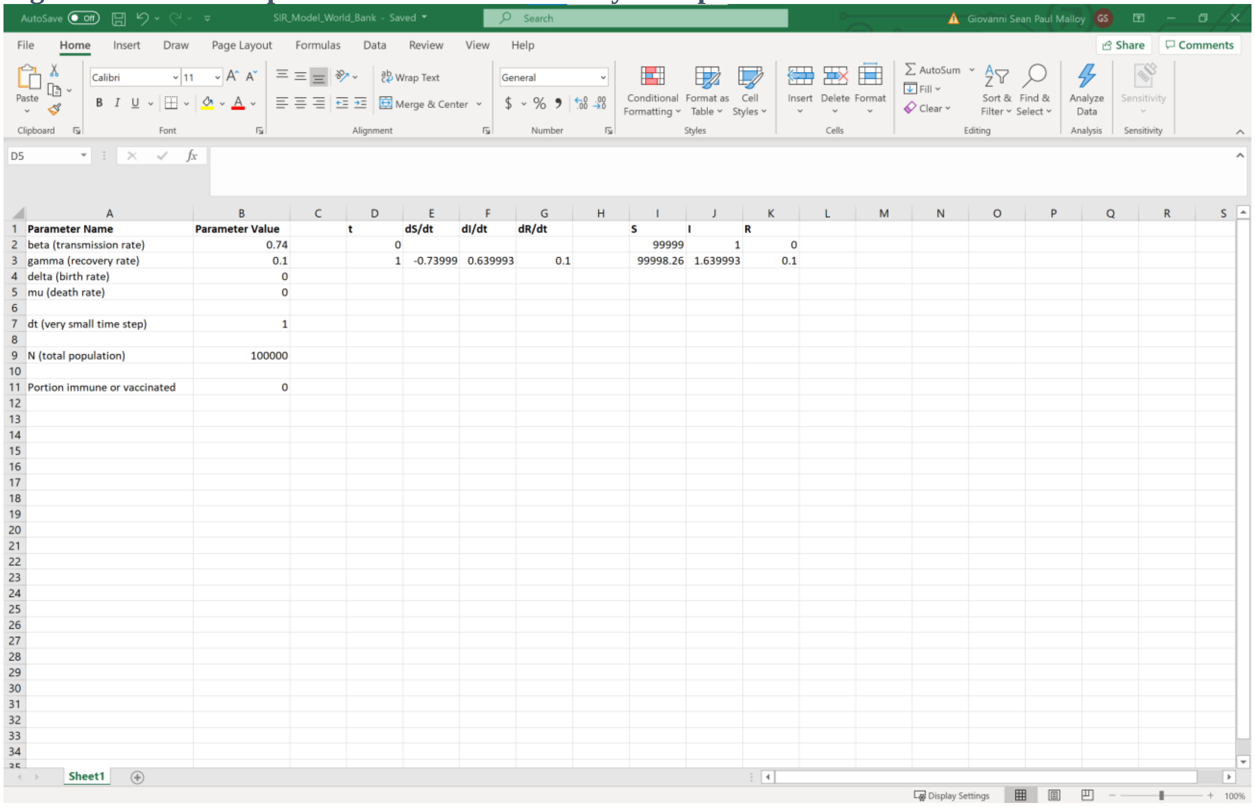
Use the SIR formulas to update the number of individuals in each compartment. To update the susceptible compartment, use the formula “=I2 + E3” in cell I3. This will add the total number of susceptible individuals when $t = 0$ to the change in the number of susceptible individuals that we calculated as $\frac{\Delta S}{\Delta t}$. The result is the total number of susceptible individuals when $t = 1$. In a similar way, to update the infected compartment use the formula “=J2 + F3” in cell I3, and to update the recovered compartment use the formula “=K2 + G3” in cell I3. The spreadsheet should now look like that of Figure 4.

To complete the implementation of the SIR model, select cells D3 through K3 and drag the formulas down until you reach $t = 100$. By day 100, the change in population between groups should be minimal and almost all the population should be in the recovered compartment. The spreadsheet should resemble that of Figure 5.

To visualize the model you just created, highlight columns I, J, and K. Go to the “Insert” tab at the top, click on the icon called “Insert Line or Area Graph”, and choose the first option under “2-D Line” called “Line”. This will automatically generate a line graph of the SIR model you created. To improve the x axis, click on and select the graph you created and then navigate to the “Chart Design” tab at the top. Under the “Chart Design” tab, click on the “Select Data” option. A pop-up window should appear titled “Select Data Source” as seen in Figure 6.

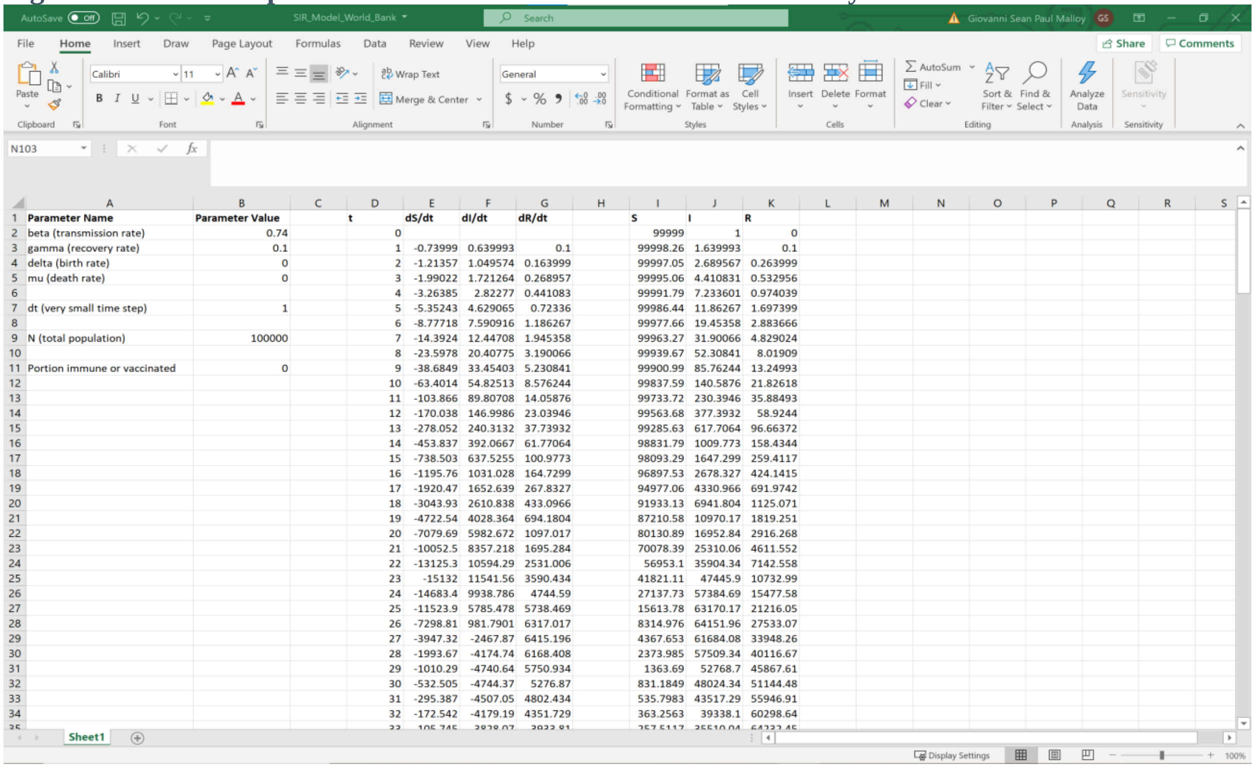
In the pop-up window, under the column titled “Horizontal (Category) Axis Labels” click the box labeled “Edit”. Then select (or manually enter) cells D2 through D102 as pictured in Figure 7. Then, click “OK” and “OK” again. At this point you have now created your first epidemic curves in Excel. Feel free to change the “Chart Title” and to add “Axis Titles” as in Figure 8. The x axis represents the time in terms of days and the y axis represents the number of individuals in each compartment.

Figure 4. The Excel spreadsheet with the first day of implementation of the SIR model



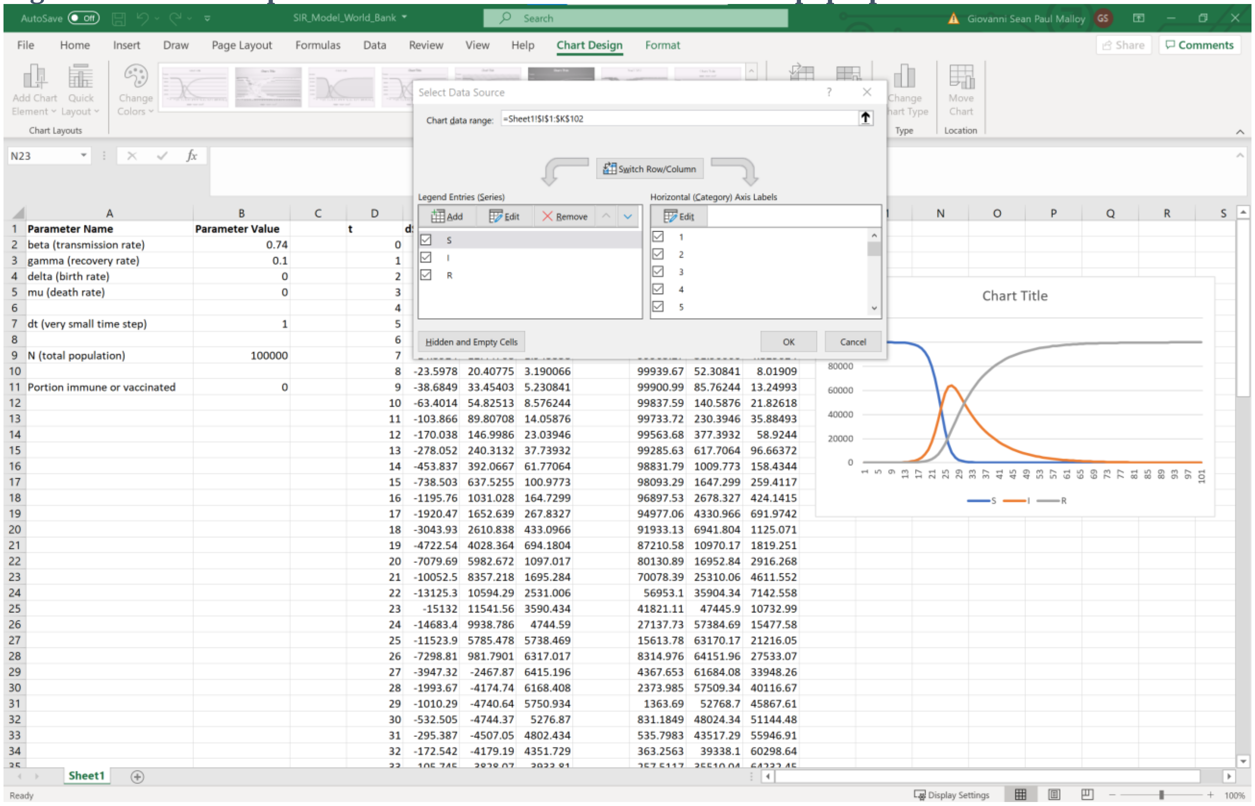
	A	B	C	D	E	F	G	H	I	J	K
1	**Parameter Name**	**Parameter Value**		**t**	**dS/dt**	**dI/dt**	**dR/dt**		**S**	**I**	**R**
2	beta (transmission rate)	0.74		0					99999	1	0
3	gamma (recovery rate)	0.1		1	-0.73999	0.639993	0.1		99998.26	1.639993	0.1
4	delta (birth rate)	0									
5	mu (death rate)	0									
6											
7	dt (very small time step)	1									
8											
9	N (total population)	100000									
10											
11	Portion immune or vaccinated	0									

Figure 5. The Excel spreadsheet with SIR model extended to 100 days



	A	B	C	D	E	F	G	H	I	J	K
1	**Parameter Name**	**Parameter Value**		**t**	**dS/dt**	**dI/dt**	**dR/dt**		**S**	**I**	**R**
2	beta (transmission rate)	0.74		0					99999	1	0
3	gamma (recovery rate)	0.1		1	-0.73999	0.639993	0.1		99998.26	1.639993	0.1
4	delta (birth rate)	0		2	-1.21357	1.049574	0.163999		99997.05	2.689567	0.263999
5	mu (death rate)	0		3	-1.99022	1.721264	0.268957		99995.06	4.410831	0.532956
6				4	-3.26385	2.82277	0.441083		99991.79	7.233601	0.974039
7	dt (very small time step)	1		5	-5.35243	4.629065	0.72336		99986.44	11.86267	1.697399
8				6	-8.77718	7.590916	1.186267		99977.66	19.45358	2.883666
9	N (total population)	100000		7	-14.3924	12.44708	1.945358		99963.27	31.90066	4.829024
10				8	-23.5978	20.40775	3.190066		99939.67	52.30841	8.01909
11	Portion immune or vaccinated	0		9	-38.6849	33.45403	5.230841		99900.99	85.76244	13.24993
12				10	-63.4014	54.82513	8.576244		99837.59	140.5876	21.82618
13				11	-103.866	89.80708	14.05876		99733.72	230.3946	35.88493
14				12	-170.038	146.9986	23.03946		99563.68	377.3932	58.9244
15				13	-278.052	240.3132	37.73932		99285.63	617.7064	96.66372
16				14	-453.837	392.0667	61.77064		98831.79	1009.773	158.4344
17				15	-738.503	637.5255	100.9773		98093.29	1647.299	259.4117
18				16	-1195.76	1031.028	164.7299		96897.53	2678.327	424.1415
19				17	-1920.47	1652.639	267.8327		94977.06	4330.966	691.9742
20				18	-3043.93	2610.838	433.0966		91933.13	6941.804	1125.071
21				19	-4722.54	4028.364	694.1804		87210.58	10970.17	1819.251
22				20	-7079.69	5982.672	1097.017		80130.89	16952.84	2916.268
23				21	-10052.5	8357.218	1695.284		70078.39	25310.06	4611.552
24				22	-13125.3	10594.29	2531.006		56953.1	35904.34	7142.558
25				23	-15132	11541.56	3590.434		41821.11	47445.9	10732.99
26				24	-14683.4	9938.786	4744.59		27137.73	57384.69	15477.58
27				25	-11523.9	5785.478	5738.469		15613.78	63170.17	21216.05
28				26	-7298.81	981.7901	6317.017		8314.976	64151.96	27533.07
29				27	-3947.32	-2467.87	6415.196		4367.653	61684.08	33948.26
30				28	-1993.67	-4174.74	6168.408		2373.985	57509.34	40116.67
31				29	-1010.29	-4740.64	5750.934		1363.69	52768.7	45867.61
32				30	-532.505	-4744.37	5276.87		831.1849	48024.34	51144.48
33				31	-295.387	-4507.05	4802.434		535.7983	43517.29	55946.91
34				32	-172.542	-4179.19	4351.729		363.2563	39338.1	60298.64

Figure 6. The Excel spreadsheet with the "Select Data Source" pop-up window



Select Data Source

Chart data range: =Sheet1!$I$1:$K$102

Switch Row/Column

Legend Entries (Series)

Add | Edit | Remove

☑ S

☑ I

☑ R

Horizontal (Category) Axis Labels

Edit

☑ 1

☑ 2

☑ 3

☑ 4

☑ 5

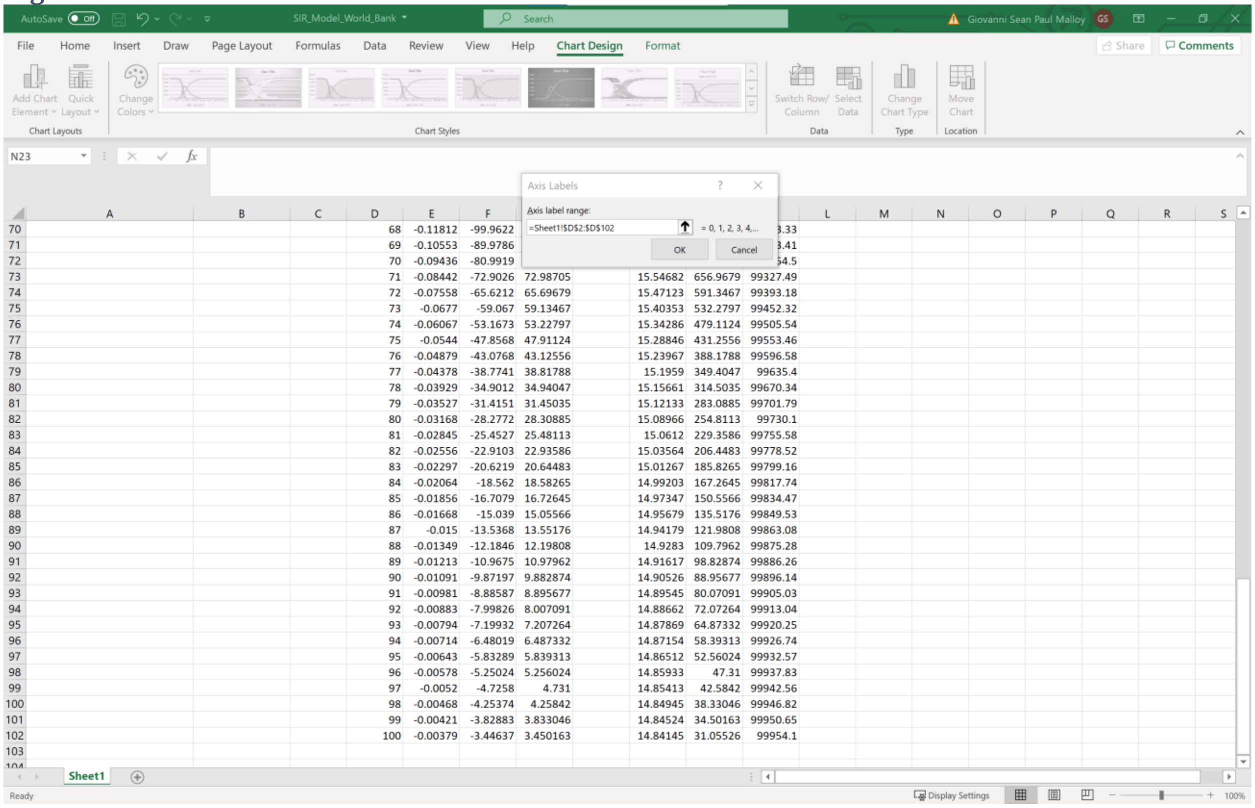
Hidden and Empty Cells

OK

Cancel

	A	B	C	D							
1	**Parameter Name**	**Parameter Value**		t	d						
2	beta (transmission rate)	0.74		0							
3	gamma (recovery rate)	0.1		1							
4	delta (birth rate)	0		2							
5	mu (death rate)	0		3							
6				4							
7	dt (very small time step)	1		5							
8				6							
9	N (total population)	100000		7							
10				8	-23.5978	20.40775	3.190066		99939.67	52.30841	8.01909
11	Portion immune or vaccinated	0		9	-38.6849	33.45403	5.230841		99900.99	85.76244	13.24993
12				10	-63.4014	54.82513	8.576244		99837.59	140.5876	21.82618
13				11	-103.866	89.80708	14.05876		99733.72	230.3946	35.88493
14				12	-170.038	146.9986	23.03946		99563.68	377.3932	58.9244
15				13	-278.052	240.3132	37.73932		99285.63	617.7064	96.66372
16				14	-453.837	392.0667	61.77064		98831.79	1009.773	158.4344
17				15	-738.503	637.5255	100.9773		98093.29	1647.299	259.4117
18				16	-1195.76	1031.028	164.7299		96897.53	2678.327	424.1415
19				17	-1920.47	1652.639	267.8327		94977.06	4330.966	691.9742
20				18	-3043.93	2610.838	433.0966		91933.13	6941.804	1125.071
21				19	-4722.54	4028.364	694.1804		87210.58	10970.17	1819.251
22				20	-7079.69	5982.672	1097.017		80130.89	16952.84	2916.268
23				21	-10052.5	8357.218	1695.284		70078.39	25310.06	4611.552
24				22	-13125.3	10594.29	2531.006		56953.1	35904.34	7142.558
25				23	-15132	11541.56	3590.434		41821.11	47445.9	10732.99
26				24	-14683.4	9938.786	4744.59		27137.73	57384.69	15477.58
27				25	-11523.9	5785.478	5738.469		15613.78	63170.17	21216.05
28				26	-7298.81	981.7901	6317.017		8314.976	64151.96	27533.07
29				27	-3947.32	-2467.87	6415.196		4367.653	61684.08	33948.26
30				28	-1993.67	-4174.74	6168.408		2373.985	57509.34	40116.67
31				29	-1010.29	-4740.64	5750.934		1363.69	52768.7	45867.61
32				30	-532.505	-4744.37	5276.87		831.1849	48024.34	51144.48
33				31	-295.387	-4507.05	4802.434		535.7983	43517.29	55946.91
34				32	-172.542	-4179.19	4351.729		363.2563	39338.1	60298.64

Figure 7. Select the cells in column D to serve as x axis labels



Axis Labels

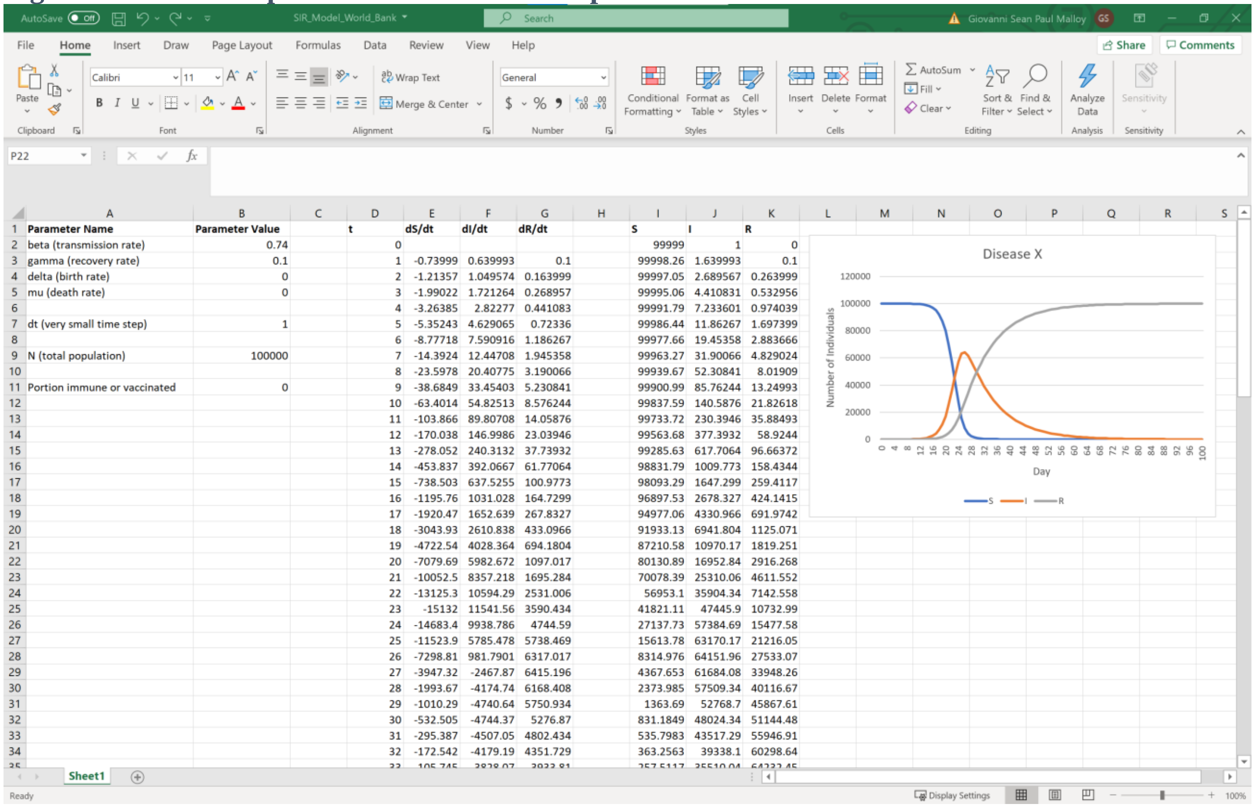
Axis label range: =Sheet1!$D$2:$D$102 = 0, 1, 2, 3, 4,...

OK

Cancel

	A	B	C	D	E	F						L
70				68	-0.11812	-99.9622					3.33	
71				69	-0.10553	-89.9786					3.41	
72				70	-0.09436	-80.9919					54.5	
73				71	-0.08442	-72.9026	72.98705		15.54682	656.9679	99327.49	
74				72	-0.07558	-65.6212	65.69679		15.47123	591.3467	99393.18	
75				73	-0.0677	-59.067	59.13467		15.40353	532.2797	99452.32	
76				74	-0.06067	-53.1673	53.22797		15.34286	479.1124	99505.54	
77				75	-0.0544	-47.8568	47.91124		15.28846	431.2556	99553.46	
78				76	-0.04879	-43.0768	43.12556		15.23967	388.1788	99596.58	
79				77	-0.04378	-38.7741	38.81788		15.1959	349.4047	99635.4	
80				78	-0.03929	-34.9012	34.94047		15.15661	314.5035	99670.34	
81				79	-0.03527	-31.4151	31.45035		15.12133	283.0885	99701.79	
82				80	-0.03168	-28.2772	28.30885		15.08966	254.8113	99730.1	
83				81	-0.02845	-25.4527	25.48113		15.0612	229.3586	99755.58	
84				82	-0.02556	-22.9103	22.93586		15.03564	206.4483	99778.52	
85				83	-0.02297	-20.6219	20.64483		15.01267	185.8265	99799.16	
86				84	-0.02064	-18.562	18.58265		14.99203	167.2645	99817.74	
87				85	-0.01856	-16.7079	16.72645		14.97347	150.5566	99834.47	
88				86	-0.01668	-15.039	15.05566		14.95679	135.5176	99849.53	
89				87	-0.015	-13.5368	13.55176		14.94179	121.9808	99863.08	
90				88	-0.01349	-12.1846	12.19808		14.9283	109.7962	99875.28	
91				89	-0.01213	-10.9675	10.97962		14.91617	98.82874	99886.26	
92				90	-0.01091	-9.87197	9.882874		14.90526	88.95677	99896.14	
93				91	-0.00981	-8.88587	8.895677		14.89545	80.07091	99905.03	
94				92	-0.00883	-7.99826	8.007091		14.88662	72.07264	99913.04	
95				93	-0.00794	-7.19932	7.207264		14.87869	64.87332	99920.25	
96				94	-0.00714	-6.48019	6.487332		14.87154	58.39313	99926.74	
97				95	-0.00643	-5.83289	5.839313		14.86512	52.56024	99932.57	
98				96	-0.00578	-5.25024	5.256024		14.85933	47.31	99937.83	
99				97	-0.0052	-4.7258	4.731		14.85413	42.5842	99942.56	
100				98	-0.00468	-4.25374	4.25842		14.84945	38.33046	99946.82	
101				99	-0.00421	-3.82883	3.833046		14.84524	34.50163	99950.65	
102				100	-0.00379	-3.44637	3.450163		14.84145	31.05526	99954.1	
103												

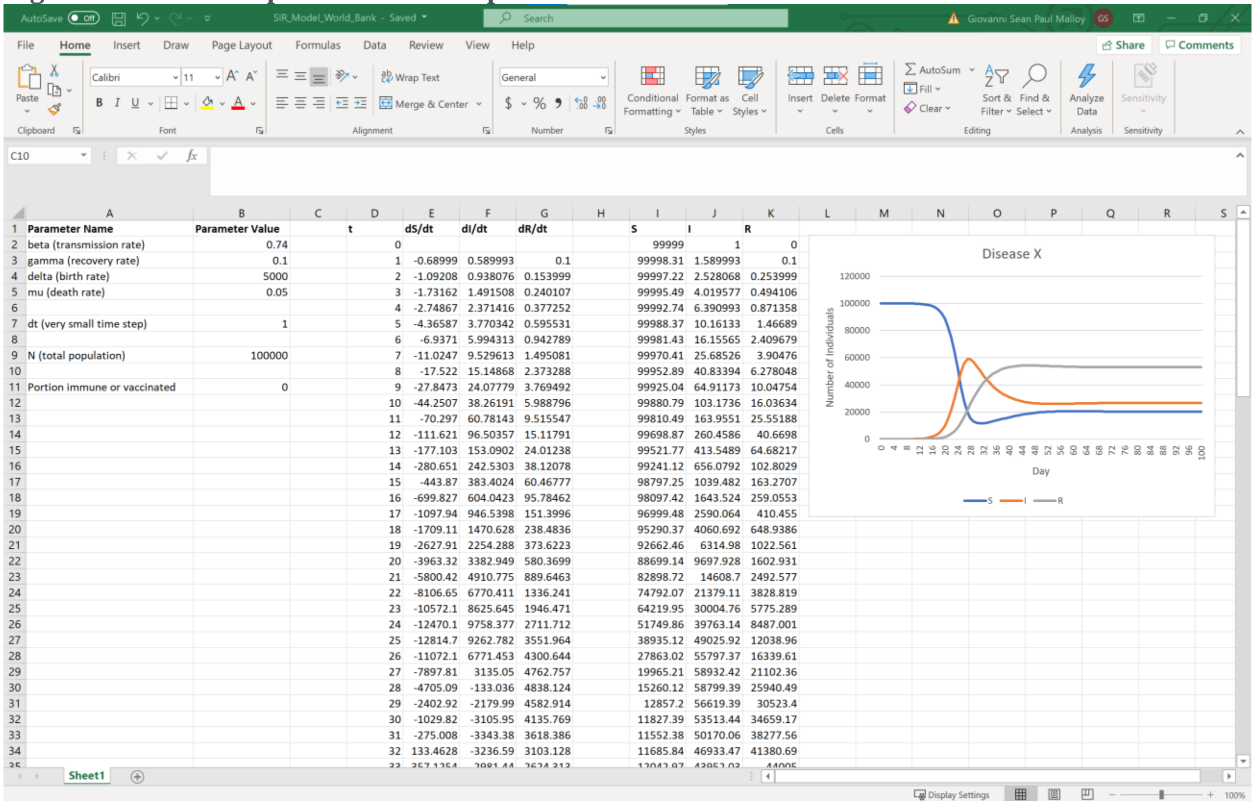
Figure 8. The Excel spreadsheet with the SIR epidemic curves for Disease X



	A	B	C	D	E	F	G	H	I	J	K
1	**Parameter Name**	**Parameter Value**		**t**	**dS/dt**	**dI/dt**	**dR/dt**		**S**	**I**	**R**
2	beta (transmission rate)	0.74		0					99999	1	0
3	gamma (recovery rate)	0.1		1	-0.73999	0.639993	0.1		99998.26	1.639993	0.1
4	delta (birth rate)	0		2	-1.21357	1.049574	0.163999		99997.05	2.689567	0.263999
5	mu (death rate)	0		3	-1.99022	1.721264	0.268957		99995.06	4.410831	0.532956
6				4	-3.26385	2.82277	0.441083		99991.79	7.233601	0.974039
7	dt (very small time step)	1		5	-5.35243	4.629065	0.72336		99986.44	11.86267	1.697399
8				6	-8.77718	7.590916	1.186267		99977.66	19.45358	2.883666
9	N (total population)	100000		7	-14.3924	12.44708	1.945358		99963.27	31.90066	4.829024
10				8	-23.5978	20.40775	3.190066		99939.67	52.30841	8.01909
11	Portion immune or vaccinated	0		9	-38.6849	33.45403	5.230841		99900.99	85.76244	13.24993
12				10	-63.4014	54.82513	8.576244		99837.59	140.5876	21.82618
13				11	-103.866	89.80708	14.05876		99733.72	230.3946	35.88493
14				12	-170.038	146.9986	23.03946		99563.68	377.3932	58.9244
15				13	-278.052	240.3132	37.73932		99285.63	617.7064	96.66372
16				14	-453.837	392.0667	61.77064		98831.79	1009.773	158.4344
17				15	-738.503	637.5255	100.9773		98093.29	1647.299	259.4117
18				16	-1195.76	1031.028	164.7299		96897.53	2678.327	424.1415
19				17	-1920.47	1652.639	267.8327		94977.06	4330.966	691.9742
20				18	-3043.93	2610.838	433.0966		91933.13	6941.804	1125.071
21				19	-4722.54	4028.364	694.1804		87210.58	10970.17	1819.251
22				20	-7079.69	5982.672	1097.017		80130.89	16952.84	2916.268
23				21	-10052.5	8357.218	1695.284		70078.39	25310.06	4611.552
24				22	-13125.3	10594.29	2531.006		56953.1	35904.34	7142.558
25				23	-15132	11541.56	3590.434		41821.11	47445.9	10732.99
26				24	-14683.4	9938.786	4744.59		27137.73	57384.69	15477.58
27				25	-11523.9	5785.478	5738.469		15613.78	63170.17	21216.05
28				26	-7298.81	981.7901	6317.017		8314.976	64151.96	27533.07
29				27	-3947.32	-2467.87	6415.196		4367.653	61684.08	33948.26
30				28	-1993.67	-4174.74	6168.408		2373.985	57509.34	40116.67
31				29	-1010.29	-4740.64	5750.934		1363.69	52768.7	45867.61
32				30	-532.505	-4744.37	5276.87		831.1849	48024.34	51144.48
33				31	-295.387	-4507.05	4802.434		535.7983	43517.29	55946.91
34				32	-172.542	-4179.19	4351.729		363.2563	39338.1	60298.64
35				33	-105.745	-3828.07	3933.81		257.5117	35510.04	64232.45

To model an outbreak with birth and death in Excel, first decide on an appropriate death rate (or 1/lifespan). For this example, set $\mu = 0.05$. Then, set the birth rate δ such that the overall population remains stable. To do so, use the formula "=B5*B9" in cell B4 this is equivalent to μN as we discussed earlier. The spreadsheet will now look like that of Figure 9. Notice that the epidemic curves find a new long-term stable level of individuals in each compartment. This new equilibrium has a larger number of susceptible and infected individuals at any given time. This makes sense because the births into the model provide the population with 5,000 new susceptible individuals every day.

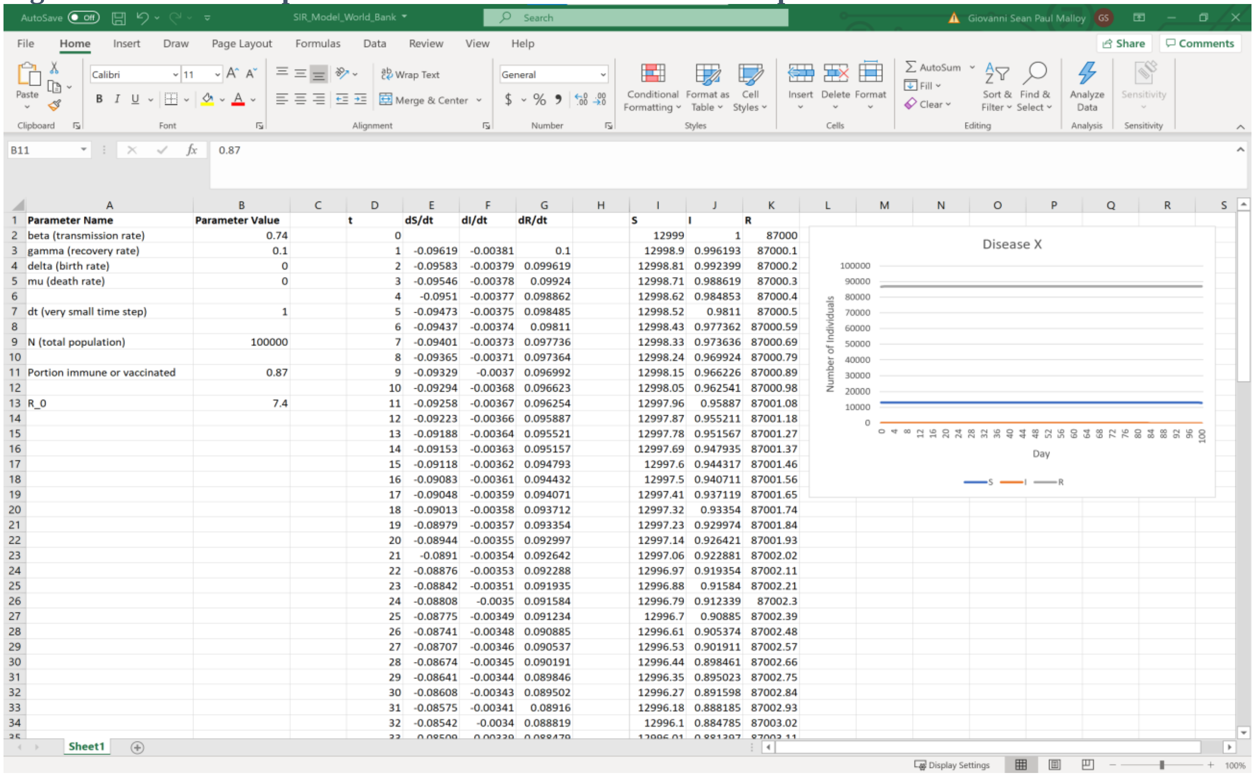
Figure 9. The Excel spreadsheet with epidemic curves for SIR model with birth and death



C10

	A	B	C	D	E	F	G	H	I	J	K
1	**Parameter Name**	**Parameter Value**		**t**	**dS/dt**	**dI/dt**	**dR/dt**		**S**	**I**	**R**
2	beta (transmission rate)	0.74		0					99999	1	0
3	gamma (recovery rate)	0.1		1	-0.68999	0.589993	0.1		99998.31	1.589993	0.1
4	delta (birth rate)	5000		2	-1.09208	0.938076	0.153999		99997.22	2.528068	0.253999
5	mu (death rate)	0.05		3	-1.73162	1.491508	0.240107		99995.49	4.019577	0.494106
6				4	-2.74867	2.371416	0.377252		99992.74	6.390993	0.871358
7	dt (very small time step)	1		5	-4.36587	3.770342	0.595531		99988.37	10.16133	1.46689
8				6	-6.9371	5.994313	0.942789		99981.43	16.15565	2.409679
9	N (total population)	100000		7	-11.0247	9.529613	1.495081		99970.41	25.68526	3.90476
10				8	-17.522	15.14868	2.373288		99952.89	40.83394	6.278048
11	Portion immune or vaccinated	0		9	-27.8473	24.07779	3.769492		99925.04	64.91173	10.04754
12				10	-44.2507	38.26191	5.988796		99880.79	103.1736	16.03634
13				11	-70.297	60.78143	9.515547		99810.49	163.9551	25.55188
14				12	-111.621	96.50357	15.11791		99698.87	260.4586	40.6698
15				13	-177.103	153.0902	24.01238		99521.77	413.5489	64.68217
16				14	-280.651	242.5303	38.12078		99241.12	656.0792	102.8029
17				15	-443.87	383.4024	60.46777		98797.25	1039.482	163.2707
18				16	-699.827	604.0423	95.78462		98097.42	1643.524	259.0553
19				17	-1097.94	946.5398	151.3996		96999.48	2590.064	410.455
20				18	-1709.11	1470.628	238.4836		95290.37	4060.692	648.9386
21				19	-2627.91	2254.288	373.6223		92662.46	6314.98	1022.561
22				20	-3963.32	3382.949	580.3699		88699.14	9697.928	1602.931
23				21	-5800.42	4910.775	889.6463		82898.72	14608.7	2492.577
24				22	-8106.65	6770.411	1336.241		74792.07	21379.11	3828.819
25				23	-10572.1	8625.645	1946.471		64219.95	30004.76	5775.289
26				24	-12470.1	9758.377	2711.712		51749.86	39763.14	8487.001
27				25	-12814.7	9262.782	3551.964		38935.12	49025.92	12038.96
28				26	-11072.1	6771.453	4300.644		27863.02	55797.37	16339.61
29				27	-7897.81	3135.05	4762.757		19965.21	58932.42	21102.36
30				28	-4705.09	-133.036	4838.124		15260.12	58799.39	25940.49
31				29	-2402.92	-2179.99	4582.914		12857.2	56619.39	30523.4
32				30	-1029.82	-3105.95	4135.769		11827.39	53513.44	34659.17
33				31	-275.008	-3343.38	3618.386		11552.38	50170.06	38277.56
34				32	133.4628	-3236.59	3103.128		11685.84	46933.47	41380.69

Figure 10. The Excel spreadsheet when vaccination threshold prevents an outbreak of disease X



B11 fx 0.87

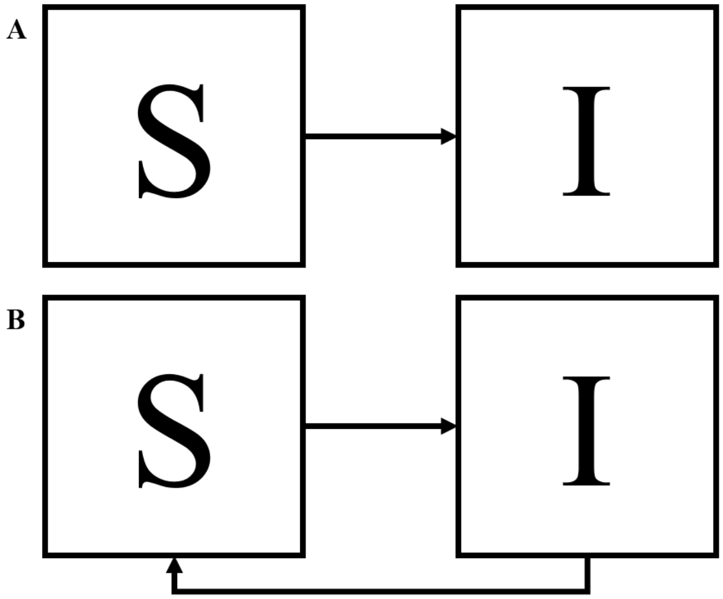
	A	B	C	D	E	F	G	H	I	J	K
1	**Parameter Name**	**Parameter Value**		**t**	**dS/dt**	**dI/dt**	**dR/dt**		**S**	**I**	**R**
2	beta (transmission rate)	0.74		0					12999	1	87000
3	gamma (recovery rate)	0.1		1	-0.09619	-0.00381	0.1		12998.9	0.996193	87000.1
4	delta (birth rate)	0		2	-0.09583	-0.00379	0.099619		12998.81	0.992399	87000.2
5	mu (death rate)	0		3	-0.09546	-0.00378	0.09924		12998.71	0.988619	87000.3
6				4	-0.0951	-0.00377	0.098862		12998.62	0.984853	87000.4
7	dt (very small time step)	1		5	-0.09473	-0.00375	0.098485		12998.52	0.9811	87000.5
8				6	-0.09437	-0.00374	0.09811		12998.43	0.977362	87000.59
9	N (total population)	100000		7	-0.09401	-0.00373	0.097736		12998.33	0.973636	87000.69
10				8	-0.09365	-0.00371	0.097364		12998.24	0.969924	87000.79
11	Portion immune or vaccinated	0.87		9	-0.09329	-0.0037	0.096992		12998.15	0.966226	87000.89
12				10	-0.09294	-0.00368	0.096623		12998.05	0.962541	87000.98
13	R_0	7.4		11	-0.09258	-0.00367	0.096254		12997.96	0.95887	87001.08
14				12	-0.09223	-0.00366	0.095887		12997.87	0.955211	87001.18
15				13	-0.09188	-0.00364	0.095521		12997.78	0.951567	87001.27
16				14	-0.09153	-0.00363	0.095157		12997.69	0.947935	87001.37
17				15	-0.09118	-0.00362	0.094793		12997.6	0.944317	87001.46
18				16	-0.09083	-0.00361	0.094432		12997.5	0.940711	87001.56
19				17	-0.09048	-0.00359	0.094071		12997.41	0.937119	87001.65
20				18	-0.09013	-0.00358	0.093712		12997.32	0.93354	87001.74
21				19	-0.08979	-0.00357	0.093354		12997.23	0.929974	87001.84
22				20	-0.08944	-0.00355	0.092997		12997.14	0.926421	87001.93
23				21	-0.0891	-0.00354	0.092642		12997.06	0.922881	87002.02
24				22	-0.08876	-0.00353	0.092288		12996.97	0.919354	87002.11
25				23	-0.08842	-0.00351	0.091935		12996.88	0.91584	87002.21
26				24	-0.08808	-0.0035	0.091584		12996.79	0.912339	87002.3
27				25	-0.08775	-0.00349	0.091234		12996.7	0.90885	87002.39
28				26	-0.08741	-0.00348	0.090885		12996.61	0.905374	87002.48
29				27	-0.08707	-0.00346	0.090537		12996.53	0.901911	87002.57
30				28	-0.08674	-0.00345	0.090191		12996.44	0.898461	87002.66
31				29	-0.08641	-0.00344	0.089846		12996.35	0.895023	87002.75
32				30	-0.08608	-0.00343	0.089502		12996.27	0.891598	87002.84
33				31	-0.08575	-0.00341	0.08916		12996.18	0.888185	87002.93
34				32	-0.08542	-0.0034	0.088819		12996.1	0.884785	87003.02

4. Other Common Model Structures

The SIR model is not the only common type of deterministic model. Two other common models are the SI/SIS model and the SEIR model. They have various uses and may be more appropriate for a given infectious disease than the SIR model.

The SI, or SIS, model is a deterministic model with two compartments: S – susceptible and I – infected. The difference between the SI model and the SIS model is that the SI model is intended for lifelong infections from which there is no recovery, and the SIS model is intended for infections from which there is recovery but not immunity to reinfection. In the SI model (Figure 11a), susceptible individuals become infected and remain infected until death. Human infectious diseases that exhibit this behavior include HIV [15]. In addition, there are a variety of animal and plant infectious diseases that are lifelong and fatal including Feline Infectious Peritonitis, Spongiform Encephalopathy, Leishmaniasis, Rabbit Hemorrhagic Disease, and Highly Pathogenic Avian Influenza [5]. In the SIS model (Figure 11b), susceptible individuals can become infected and infected individuals recover without immunity to reinfection. Some examples of infectious diseases that do not confer immunity include rotaviruses, many sexually transmitted infections, and many bacterial infections [5].

Figure 11. SI (A) and SIS (B) model structure. Disease state transitions are possible from susceptible to infected compartments in both models and from infected to susceptible compartments in the SIS models



The mathematical formulations of these models are similar [5]. For the SI model with birth and death, the equations are as follows:

$$\frac{dS}{dt} = \delta - \frac{\beta SI}{N} - \mu S$$
$$\frac{dI}{dt} = \frac{\beta SI}{N} - \mu I$$

For the SIS model with birth and death, the equations are as follows:

$$\frac{dS}{dt} = \delta + \gamma I - \frac{\beta SI}{N} - \mu S$$
$$\frac{dI}{dt} = \frac{\beta SI}{N} - (\gamma + \mu)I$$

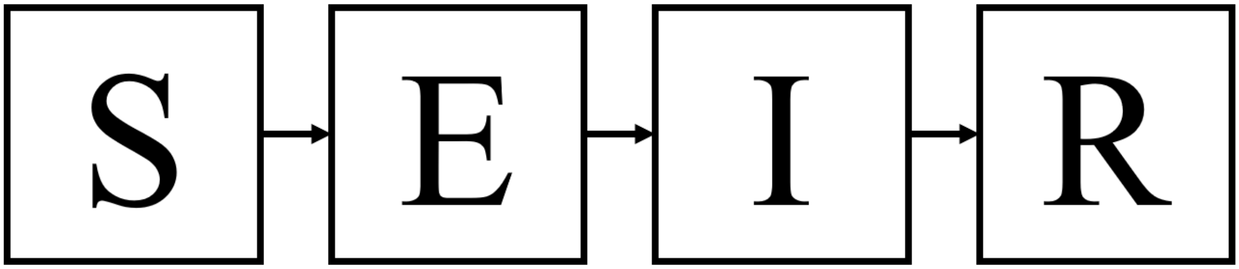
For the SI model, $R_0 = \frac{\beta}{\mu}$, and for the SIS model $R_0 = \frac{\beta}{\gamma+\mu}$ using the same methodology as for the SIR model previously.

Another popular model is called the SEIR model. The SEIR model is a deterministic infectious disease model with four compartments: S – susceptible, E – exposed, I – infected, R – recovered. In this case, the exposed population is comprised of individuals who are infected with the disease but not yet infectious, whereas the infected population is comprised of individuals who are both infected and infectious. The period of time between transitioning to the exposed compartment and transitioning to the infected compartment is called the latent period. Individuals in the model progress from the susceptible compartment to the exposed compartment to the infected compartment and finally to the recovered compartment. The diagram of the model can be seen in Figure 12. The SEIR model has been used to describe diseases such as COVID-19 [16] and Ebola [17]. The equations that define the SEIR model are similar to that of the SIR model [5]. The latent period is defined as $\frac{1}{\sigma}$ in this model:

$$\frac{dS}{dt} = \delta - \frac{\beta SI}{N} - \mu S$$
$$\frac{dE}{dt} = \frac{\beta SI}{N} - (\sigma + \mu)E$$
$$\frac{dI}{dt} = \sigma E - (\gamma + \mu)I$$
$$\frac{dR}{dt} = \gamma I - \mu R$$

For the SEIR model, the next generation method is needed to derive R_0 [9]. The next generation method is beyond the scope of this document, but the result is that $R_0 = \frac{\sigma\beta\delta}{\mu(\sigma+\mu)(\gamma+\mu)}$ for the SEIR model.

Figure 12. SEIR model structure. Disease state transitions are possible from susceptible to exposed from exposed to the infected and from infected to recovered compartments



5. Deterministic versus Stochastic Models

While deterministic models have a long and successful history in epidemiology, they are not without their limitations. The first and most notable limitation of deterministic models is that they assume homogeneous mixing. This assumption means that all individuals within the model have equal probability of coming into contact with one another [18]. This assumption can be reasonable when you are solely interested in the population level outcomes of a disease but does not reflect the real world contact structures between human beings. Another important limitation of deterministic models is that they do not incorporate uncertainty [18]. Deterministic models produce the same output every time given the same inputs. Roberts et al. point out that it can therefore be difficult to incorporate the possibility of events like multi-strain infections, infections with time-varying infectivity, and infections where superinfection is possible [19].

Unlike deterministic models, stochastic models incorporate an element of randomness [5]. This element of randomness can be implemented through various methods. Two common ways include adding randomization to parameters and adding randomization to events. In deterministic models, each parameter is a constant point estimate, however, in stochastic models, any number of parameters can be described by a distribution [5]. For example, Malloy et al. describe several parameters of COVID-19 with distributions instead of point estimates and sample from these distributions over many Monte Carlo simulations [16]. Another similar approach introduces noise to the model events, such as the terms describing infection ($\frac{-\beta SI}{N}$) or recovery (γI). Allen outlines this approach in an example model of the spread of malaria [20]. The result of both processes produces uncertainty bounds on important epidemiological outcomes, such as the number of cases or R_0.

Despite the obvious perks of stochastic models, they can be more difficult to implement, understand, and describe. The requirement to run many simulations of a stochastic model means that they are computationally more expensive than deterministic models. Stochastic models should be thought of as a robust way to capture uncertainty in disease dynamics or heterogeneities of a population. However, Muller et al. shows that it is possible to develop deterministic models that approximate stochastic processes [21], and Allen and van den Driessche confirm that deterministic and stochastic epidemiological thresholds are similar [22]. Depending on the application, deterministic models will often suffice.

6. Stochastic Simulation to Predict Outbreak Size

An additional goal of public health preparedness and infectious disease modeling is to predict when new outbreaks might occur and how big they will be. It is possible to simulate this phenomenon using principles of stochastic modeling and the computing power of Excel. Accompanying this document is another Excel spreadsheet containing data on Filovirus (Ebola and Marburg) outbreaks between 1996 and 2014. The data describe the start date of the outbreak ("Start Year or Date of First WHO Report"), the disease of the outbreak ("Pathogen"), the location of the outbreak ("Territory(ies)"), and the total number of cases and deaths during the outbreak. Ultimately, the question that we will answer is *how many outbreaks of at least a* *certain size, z, will we expect over a given time period, y?* To answer our question, we will first need to develop a model for the time between outbreaks and the size of each outbreak. Therefore, the two most important data needed to answer this question are the start date of the outbreak and the total number of cases of the outbreak.

The time between outbreaks is called “Delay” in the attached spreadsheet. Since these outbreaks occur continuously and independently, we assume that the delay between outbreaks occurs at a constant average rate, λ, and therefore, the occurrence of an outbreak can be modeled as a Poisson process [23]. The time between events in a Poisson process is modeled using an exponential distribution, which is a continuous and memoryless distribution. A continuous distribution is one in which there is an associated probability for an infinite number of outcomes over a given range, and a memoryless distribution is one in which the probability of an event occurring is independent of the time that has elapsed [23]. The exponential distribution is defined by only one parameter, λ, the rate at which an event occurs[24]. In this case, the rate of outbreaks is estimated as the inverse the delay between outbreaks. The average observed delay between outbreaks is 296 days which makes $\lambda = 1/296$. In column H of the spreadsheet, we assess the goodness of fit of the chosen distribution using the likelihood function. The likelihood function is a way determining how well a proposed distribution fits observed data [24]. It is calculated by multiplying together the likelihoods of choosing each data point from the exponential distribution [24]:

$$\mathcal{L}_n(\lambda) = \prod_{i=1}^{n} f(X_i, \lambda)$$

For the exponential distribution, the likelihood of choosing a random value from the distribution is $f(X_i, \lambda) = \lambda e^{-\lambda x}$. Due to the properties of both the exponential distributions and logarithms, it is easier to instead use the log likelihood function to measure goodness of fit. Simplifying, we get:

$$\ln\big(\mathcal{L}_n(\lambda)\big) = \sum_{i=1}^{n} \ln\big(f(X_i, \lambda)\big) = \sum_{i=1}^{n} [\ln(\lambda) - \lambda x_i]$$

In the spreadsheet formulas, the rate of new outbreaks occurring, λ, is replaced by the equivalent inverse of the average delay $\frac{1}{E[X]}$ which simplifies to:

$$\ln\big(\mathcal{L}_n(\lambda)\big) = \sum_{i=1}^{n} \left[- \ln(E[X]) - \frac{x_i}{E[X]}\right]$$

The likelihood of each x_i is shown as an example in cells H3-H25.

The size or severity of the outbreak is measured in terms of the total number of cases in cells D2-D25. The Weibull distribution is a great choice to model the number of cases because it is a very flexible distribution [25]. By calibrating the shape parameter, α, and the scale parameter, β, of the distribution, we can leverage the Weibull to model many types of data. In the spreadsheet, α is parameter 1 and β is parameter 2. The log-likelihood of each data point is located in cells I2-I25 of the spreadsheet calculated using the pre-existing “Weibull.Dist” function in Excel. In cells

A31 – O10142, 10,110 sample parameter sets are generated using a Markov Chain Monte Carlo method (the Metropolis-Hastings algorithm) and using the likelihood function to calculate an acceptability threshold. The details of the Metropolis-Hastings algorithm are laid out in a digestible manner by Wakefield [26]. The important takeaway is that the algorithm gives us a structured format to sample the distribution state space.

Of the many parameter sets generated by the Metropolis-Hastings algorithm, 200 are randomly selected in cells Q31 – AZ231. The sampled $1/\lambda$ parameters (average delay between outbreaks) for the exponential distribution are in cells R32 – R231. The sampled parameter sets for the Weibull distribution are in cells T32 – U231. Then, we use the distributions to generate outbreaks. In cells V32 – AJ231 of the spreadsheet, we simulate the delays between outbreaks for up to 15 outbreaks each column "di" denotes the start day of the ith outbreak. In cells AK32 – AY231 of the spreadsheet, we simulate the number of cases per outbreak for all outbreaks in the "Policy term" in years defined in cell W13. Each column "oi" denotes the number of cases of the ith outbreak. Now that we have simulated Filovirus outbreaks for the duration of our policy term, we want to identify how many of the simulated outbreaks meet our minimum case threshold defined as the "Payout case threshold" in cell W14. In cells AZ32 – AZ231, we count the number of outbreaks in our policy term which have at least the number of cases defined in our payout case threshold. This column is called "# payouts."

Finally, we use each simulation of outbreak events to generate our own probability distribution. This probability distribution will help us answer our original question. It defines the likelihood of having a given number of filovirus outbreaks of at least our payout case threshold size over the time horizon of our policy term. In cells Y2 – Y17, are the number of Filovirus outbreaks that meet our threshold in our time horizon from 0 – 15 outbreaks. Since the range of possible outcomes is finite, we call this a discrete probability distribution. In cells Z2 – Z17, we count the frequency of occurrences of a given payout in our 200 simulations. For example, cell Z2 is the number of times there were 0 Filovirus outbreaks that met our case threshold size within our policy term. Finally, in cells AA2 – AA17, we include the cumulative distribution function. For example, cell AA2 is the probability that there will be 0 Filovirus outbreaks of at least our threshold size in our policy term, cell AA3 is the probability that there will be 0 or 1 Filovirus outbreaks of at least our threshold size in our policy term, cell AA4 is the probability that there will be at most 2 Filovirus outbreaks of at least our threshold size in our policy term, and so on.

We can easily construct our probability distribution function. Feel free to try this yourself. First, label cell AB1 "pdf". Next, enter the formula "=Z2/200" into cell AB2. Then, drag this formula down to cell AB17. You can now see more easily which outcome is most likely to occur.

7. Conclusions

After reading this document, you should have a good understanding of some of the fundamental deterministic infectious disease models including the SIR model, SI model, SIS model, and SEIR model. You should feel comfortable with the mathematics behind these models and be able to evaluate key outbreak features, such as R_0 and the portion of the population needed to be vaccinated to prevent an outbreak. Additionally, you should be able to continue to explore how these parameters interact using an Excel spreadsheet model and to communicate the limitations of such models and how they differ from stochastic infectious disease models.

Deterministic infectious disease models are simple to understand, implement, and communicate which makes them very powerful tools in epidemiology. They can be expanded and manipulated to approximate the spread of almost any infectious disease on earth with minimal computational requirements. Their simplicity and flexibility mean that they will continue to be ubiquitous in the world of infectious disease modeling, and those who are able to understand these models will be able to effectively engage with the infectious disease epidemiology community.

References

[1] Fisman D, Khoo E, Tuite A. Early epidemic dynamics of the west african 2014 ebola outbreak: estimates derived with a simple two-parameter model. PLoS Curr. 2014; 6.

[2] Feng Z, Castillo-Chavez C, Capurro AF. A model for tuberculosis with exogenous reinfection. Theor Popul Biol. 2000; 57(3):235-47.

[3] Granich RM, Gilks CF, Dye C, De Cock KM, Williams BG. Universal voluntary HIV testing with immediate antiretroviral therapy as a strategy for elimination of HIV transmission: a mathematical model. Lancet. 2009; 373(9657):48-57.

[4] Kermack WO, Woolhouse MEJ. A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society A. 1927; 115:700-21.

[5] Keeling MJ, Rohani P. Modeling Infectious Diseases in Humans and Animals. Princeton: Princeton University Press 2008.

[6] Reluga TC. Game theory of social distancing in response to an epidemic. PLoS Comput Biol. 2010; 6(5):e1000793.

[7] Eikenberry SE, Mancuso M, Iboi E, Phan T, Eikenberry K, Kuang Y, et al. To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic. Infect Dis Model. 2020; 5:293-308.

[8] Towers S, Vogt Geisse K, Zheng Y, Feng Z. Antiviral treatment for pandemic influenza: assessing potential repercussions using a seasonally forced SIR model. J Theor Biol. 2011; 289:259-68.

[9] Diekmann O, Heesterbeek JAP, Metz JAJ. On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations. J Math Biol. 1990; 28(4):365-82.

[10] Wallinga J, Teunis P. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. Am J Epidemiol. 2004; 160(6):509-16.

[11] Sanche S, Lin YT, Xu C, Romero-Severson E, Hengartner N, Ke R. High Contagiousness and Rapid Spread of Severe Acute Respiratory Syndrome Coronavirus 2. Emerg Infect Dis. 2020; 26(7):1470-7.

[12] Khan A, Naveed M, Dur EAM, Imran M. Estimating the basic reproductive ratio for the Ebola outbreak in Liberia and Sierra Leone. Infect Dis Poverty. 2015; 4:13.

[13] Gani R, Leach S. Transmission potential of smallpox in contemporary populations. Nature. 2001; 414(6865):748-51.

[14] Guerra FM, Bolotin S, Lim G, Heffernan J, Deeks SL, Li Y, et al. The basic reproduction number (R0) of measles: a systematic review. Lancet Infect Dis. 2017; 17(12):e420-e8.

[15] Long EF, Vaidya NK, Brandeau ML. Controlling Co-Epidemics: Analysis of HIV and Tuberculosis Infection Dynamics. Oper Res. 2008; 56(6):1366-81.

[16] Malloy GSP, Puglisi L, Brandeau ML, Harvey TD, Wang EA. Effectiveness of interventions to reduce COVID-19 transmission in a large urban jail: a model-based analysis. BMJ Open. 2021; 11(2):e042898.

[17] Lekone PE, Finkenstadt BF. Statistical inference in a stochastic epidemic SEIR model with control intervention: Ebola as a case study. Biometrics. 2006; 62(4):1170-7.

[18] Tolles J, Luong T. Modeling Epidemics With Compartmental Models. JAMA. 2020; 323(24):2515-6.

[19] Roberts M, Andreasen V, Lloyd A, Pellis L. Nine challenges for deterministic epidemic models. Epidemics. 2015; 10:49-53.

[20] Allen LJS. A primer on stochastic epidemic models: Formulation, numerical simulation, and analysis. Infect Dis Model. 2017; 2(2):128-42.

[21] Muller J, Kretzschmar M, Dietz K. Contact tracing in stochastic and deterministic epidemic models. Math Biosci. 2000; 164(1):39-64.

[22] Allen LJ, van den Driessche P. Relations between deterministic and stochastic thresholds for disease extinction in continuous- and discrete-time infectious disease models. Math Biosci. 2013; 243(1):99-108.

[23] Ross SM. Introduction to probability and statistics for engineers and scientists. 3rd ed. Amsterdam ; Boston: Elsevier Academic Press 2004.

[24] Wasserman L. All of statistics : a concise course in statistical inference. New York: Springer 2004.

[25] Thomopoulos NT. Probability Distributions With Truncated, Log and Bivariate Extensions. 1st ed:1 online resource (171 pages).

[26] Wakefield J. Bayesian and Frequentist Regression Methods. *Springer Series in Statistics,*. 1st ed:1 online resource (699 p.).