A Mathcad SIR Model of the Spread of the Corona Virus (SARS-CoV-2)
By Jeff Henning

Mathematical models can be useful for predicting the spread of a disease. This model can help us to understand the behavior of the epidemic and what can be done to control it. A basic epidemic model, called the SIR, model was proposed by Kermack and McKendrick in 1927 [1][2][3] and has been successfully used to predict the behavior of previous historical epidemics. The SIR model examines a total population, N, and divides it into three sub-populations:

- **S** (Susceptible) - Those who are susceptible to being infected by SARS-CoV-2.
- **I** (Infected) - Those who have contracted COVID-19; may or may not be showing symptoms; are infectious; and have not yet recovered.
- **R** (Removed) - Those who were previously infected and have now either recovered or expired.

The concept here is that initially, some portion of the total population is susceptible to infection. As susceptible people get infected, they are removed from the S population and added to the I population. Without an external cure, the infected population, I, will after some period of time, either develop an immunity to the disease and recover (being no longer infectious or susceptible) or succumb to complications from the disease and expire. Either way, they are no longer infectious and are removed from the I population and placed in the R population. Hence, the S-I-R model.

To simplify the model, we have to make a few underlying assumptions:

1. Assume that for the time period under consideration is short enough that the total population under consideration, N, is constant; ignoring any underlying births and deaths that would normally occur for other reasons. This means then that the sum of the three model populations is equal to a constant total population.

$$S+I+R=N$$

2. The rate that the disease is transmitted, rate of infection, is proportional to the number of interactions between the I and S populations and the proportionality constant, the rate of contact (β), is a constant.

3. The removal rate, γ, is constant.

4. Lastly, and this is a big one because, for the novel corona virus, this is largely unknown, we assume that once an infected person recovers, the are immune to the SARS-CoV-2 virus and cannot be re-introduced into the the susceptible population, S.

If we write rate equations for each population:

$$\frac{\mathrm{d}}{\mathrm{d}t}S=\frac{-\beta\cdot I\cdot S}{N} \tag{1}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}I=\frac{\beta\cdot I\cdot S}{N}-\gamma\cdot I \tag{2}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}R=\gamma\cdot I \tag{3}$$

This is a set of differential equations. It is a non-linear set of differential equations because it contains a multiplication of two of the dependent variables, $I \cdot S$. While these types of equations can be difficult to solve, they are solvable. However, we first need some initial conditions.

Initially, at time $t = 0$, the populations are as follows:

$$S(0) = S_0 = N - I_0 = \sim N \qquad \text{Very nearly } N \text{ when } I_0 \text{ is small.}$$

$$I(0) = I_0 \qquad \text{Usually a very small number, even just 1.}$$

$$R(0) = 0 \qquad \text{Assumes no one has immunity at beginning of outbreak.}$$

Additionally, since the total population is constant at all times,

$$S + I + R = S_0 + I_0 \qquad \text{and} \qquad \frac{\mathrm{d}}{\mathrm{d}t}S + \frac{\mathrm{d}}{\mathrm{d}t}I + \frac{\mathrm{d}}{\mathrm{d}t}R = 0$$

Units

It is very important to understand the units on these values and equations as we are dealing with people and time. PTC Mathcad does not have a base unit to represent people, but we can hijack Mathcad's monetary unit and call it $persons := ¤$. This is the unit that will be applied to each of the model populations. We now have units for each of the constants in the model. Each of our rate equations above has to have units of $persons$ per time (say day or $week$). Mathcad already has these time units.

Since β, the interaction parameter is multiplied by $I \cdot S$ and must result in our rate, it must have units of

$$\beta := 0.3 \cdot day^{-1}$$

We'll worry about determining the **actual** value of β from known data later. Let's say, that at the beginning of our model, at time $t_0 := 0 \cdot day$, we have an initial population of $N := 10^6 \ persons$ (1 million people) and one infected person, $I_0 := 1 \ persons$, then $S_0 := N - I_0$. We see that our rate of susceptible population decrease at time zero, using equation (1) is,

$$dS/dt := -\beta \cdot \frac{I_0 \cdot S_0}{N} = -0.3 \ \frac{persons}{day}$$

The removal parameter, γ, has to have units of inverse time. For now, lets assume that this removal parameter is $\gamma := (14 \ day)^{-1}$; we'll discuss setting this parameter later. For an infected population, of say $I := 100 \ persons$ early in the epidemic ($S := N - I$), the removal rate according to equation (3) would be

$$dR/dt := \gamma \cdot I = 7.143 \ \frac{persons}{day} \quad ,$$

the infection rate is

$$dI/dt := \frac{\beta \cdot I \cdot S}{N} - \gamma \cdot I = 22.854 \ \frac{persons}{day} \quad ,$$

and the susceptible decrease rate has increased from time zero to

$$dS/dt := \frac{-\beta \cdot I \cdot S}{N} = -29.997 \ \frac{persons}{day} \quad .$$

So our Mathcad units work out!

Epidemic or Not?

One very important question to be asked is, will the disease spread? A disease is **endemic** if it maintains a presence in a fixed, closed population (no additions to the population that might be infected). For a disease to be endemic, each infected person has to infect exactly one other person to perpetuate the disease; any more than that and the disease spreads exponentially (an **epidemic**), any less and the disease will die out. So what are the conditions in our SIR model that would tell us if we have an epidemic or not?

Firstly, since our populations are always positive and the rate of infection, β, is a positive constant (infected members cannot move back to the susceptible population), the right hand side of equation (1) is always negative. This means that S is always decreasing and less than or equal to S_0.

$$S \leq S_0 \tag{4}$$

If we plug this relation into equation (2), we get an inequality

$$\frac{\mathrm{d}}{\mathrm{d}t} I \leq I \cdot \left(\beta \cdot \frac{S_0}{N} - \gamma \right) \tag{5}$$

This tells us that the rate of change in the infected population, I, is always less than or equal to the right-hand-side (RHS) of equation (5). Since I is always positive or zero, the rate of infections is dependent on the sign of the parenthetical $(\beta \cdot S_0 - \gamma)$. If zero, then the disease is endemic and the infected population won't change. If negative, the disease will die out. If positive,

$$\beta \cdot \frac{S_0}{N} > \gamma \tag{6}$$

then the infection rate is greater than the removal (recovery/death) rate and the disease will grow exponentially; we have an epidemic! Dividing each side by the removal rate, γ, gives a dimensionless measure of the mobility of the disease called the **Basic Reproduction Number**, R_0.

> Note that at time zero, the susceptible population, S, is very nearly N, or $S_0 \approx N$.

$$R_0 = \frac{\beta}{\gamma} \cdot \left(\frac{S_0}{N} \right) = \frac{\beta}{\gamma} > 1 \tag{7}$$

This Reproduction Number (sometimes called the reproduction ratio or, incorrectly, the reproduction *rate*) has been in the news concerning the news lately concerning the novel corona virus outbreak. Early estimates from the outbreak in Wuhan China were between 2.2 and 2.7. As more data has been collected and analyzed, the R_0 estimate has been updated to between 4.7 and 6.6 [4][5][6][7]. This number is critical to modeling the spread of the disease, but can be very hard to measure early on in a rapidly spreading epidemic because the actual number of infected population is not well known.

In an excerpt from a recent CDC Article [8]:

The basic reproduction number (R0), also called the basic reproduction ratio or rate or the basic reproductive rate, is an epidemiologic metric used to describe the contagiousness or transmissibility of infectious agents. R0 is affected by numerous biological, socio-behavioral, and environmental factors that govern pathogen transmission and, therefore, is usually estimated with various types of complex mathematical models, which make R0 easily misrepresented, misinterpreted, and misapplied. R0 is not a biological constant for a pathogen, a rate over time, or a measure of disease severity, and R0 cannot be modified through vaccination campaigns. R0 is rarely measured directly, and modeled R0 values are dependent on model structures and assumptions. Some R0 values reported in the scientific literature are likely obsolete. R0 must be estimated, reported, and applied with great caution because this basic metric is far from simple.

Setting the SIR Model Parameters

For the purposes of this exercise, we'll set the SIR model parameters based on the currently published values. The removal rate can be estimated based on data currently in the news. Given an arbitrary infected population, each individual will have been infected on a different day and will be at a different point in the disease progression. Given this continuum of infected subjects and an estimate of the mean recovery period of $d := 14 \ day$, we can assume that one fourteenth of a closed, infected population would recover each day. Similarly, the if a person's condition worsens and they do not recover in that 14 days, they would end up in the hospital and still be effectively removed from the population. This is a simplification, but we can now set the recovery rate to,

$$\gamma := \frac{1}{d} \rightarrow \frac{1}{14 \cdot day} = 0.071 \ day^{-1}$$

From current estimates, and for the sake of at least some marginal optimism, we'll set the reproductive number based on the data presented above.

$$R_0 := 4.7$$

$$\beta := R_0 \cdot \gamma = 0.336 \ day^{-1}$$

Now we can make some calculations with our model.

Solving the Equations

We can use PTC Mathcad's numeric solver to get a time varying solution for this system of differential equations.
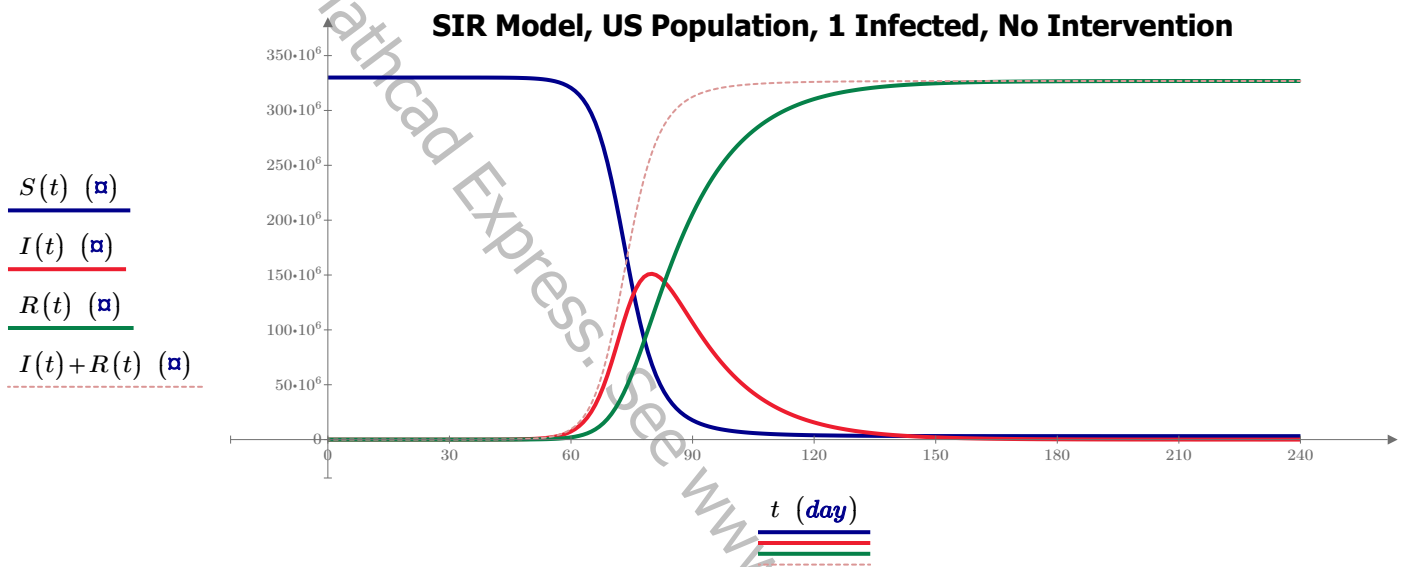
Guess Values

$$N := s_0 + i_0 + r_0 \qquad\qquad \beta := R_0 \cdot \gamma$$

Constraints

$$\frac{\mathrm{d}}{\mathrm{d}t} S(t) = \frac{-\beta \cdot I(t) \cdot S(t)}{N}$$

$$\frac{\mathrm{d}}{\mathrm{d}t} I(t) = \frac{\beta \cdot I(t) \cdot S(t)}{N} - \gamma \cdot I(t)$$

$$\frac{\mathrm{d}}{\mathrm{d}t} R(t) = \gamma \cdot I(t)$$

Initial Conditions: $\qquad S(0) = s_0 \qquad I(0) = i_0 \qquad R(0) = r_0$

Solver

$$SIR\left(s_0, i_0, r_0, R_0, \gamma, T\right) := \mathbf{odesolve}\left(\begin{bmatrix} S(t) \\ I(t) \\ R(t) \end{bmatrix}, T\right)$$

US Population Model

For a total population of the US ($N_{US} := 330 \; M \cdot persons$) and ($I_0 := 1 \; persons$) infected, assuming no initial immunity (no vaccine) so $S_0 := N_{US} - I_0$, over a time period of $T := 240 \; day$ and the model parameters described above, :

$$\begin{bmatrix} S \\ I \\ R \end{bmatrix} := SIR\left(N_{US}, I_0, 0 \; persons, R_0, \gamma, 960 \; day\right)$$

$$t := 0 \; day, 1 \; day .. 240 \; day$$

SIR Model, US Population, 1 Infected, No Intervention

Legend:
$S(t) \; (\text{¤})$
$I(t) \; (\text{¤})$
$R(t) \; (\text{¤})$
$I(t) + R(t) \; (\text{¤})$

$t \; (day)$

Note that I represents the instantaneous number of infected at any one time. The total infected is:

$$I_{total}(t) := I(t) + R(t)$$

At some point, the number of susceptible people has dropped below a threshold such that the epidemic dies out, with most of the population removed. The total infected at the end of the epidemic would be:

$$I_{total}(240 \; day) = 327 \; Million \cdot persons$$

$$\frac{I_{total}(240 \; day)}{N_{US}} = 99.0 \; \% \quad \text{of all US citizens.}$$

Flattening the Curve

So basically, everyone is going to get the virus if we do nothing. With a mortality rate estimated at even 1% (*it's probably higher*), you do the math! This explains why it is so important to "flatten the curve" with social distancing. This means that we limit the number of infectious interactions and reduce the R_0 value, ideally to something less than 1.0, causing the epidemic to cease. An NIH article [9] showed that in the 2003 SARS outbreak, the R_0 was reduced from between 2.4 and 3.6 to between 0.3 and 1.0 in the four regions where the infection was introduced in February 2003. Reducing the R_0 is the only factor that we can affect at this point, however, we're starting at a higher R_0 than the 2003 SARS outbreak to begin with; maybe double!

So, what happens to our curve if we can reduce the reproductive number to something like $R'_0 := 1.5$ through social distancing? On March 22, there are 35,000 confirmed cases in the US [10] (it's probably a lot higher than that because we're not testing everyone). This corresponds to about day 38 on the previous model.

$$I_{total}(39 \; day) = 38034 \; persons$$

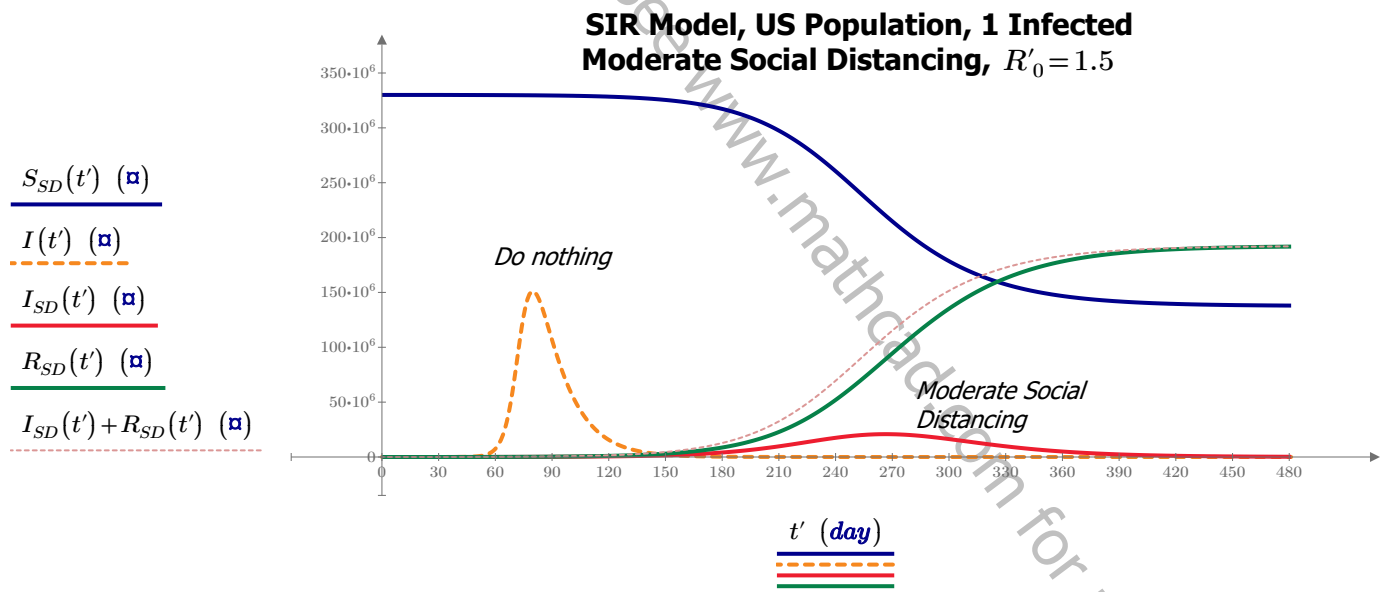$$\begin{bmatrix} S' \\ I' \\ R' \end{bmatrix} := SIR\left(S(39 \; day), I(39 \; day), R(39 \; day), R'_0, \gamma, 960 \; day - 39 \; day\right)$$

$$t' := 0 \; day, 1 \; day .. 960 \; day$$

Adjusting the curve past day 39 for social distancing:

$$S_{SD}(t) := S(min(t, 39 \; day)) + \Phi(t - 39 \; day) \cdot \left(S'(\max(0, t - 39 \; day)) - S(39 \; day)\right)$$
$$I_{SD}(t) := I(min(t, 39 \; day)) + \Phi(t - 39 \; day) \cdot \left(I'(\max(0, t - 39 \; day)) - I(39 \; day)\right)$$
$$R_{SD}(t) := R(min(t, 39 \; day)) + \Phi(t - 39 \; day) \cdot \left(R'(\max(0, t - 39 \; day)) - R(39 \; day)\right)$$

This effort definitely flattens the curve; however, millions of lives are still lost with over half the population still infected and we are living with this thing for years. It is not enough as the R_0 value is still greater than 1 and there is still an epidemic, if even a smaller one. Efforts like were implemented in China are needed to enforce strict distancing and isolation, dropping the R_0 value to something less than 1.0. Let's see what happens with aggressive social distancing at $R''_0 := 0.5$.
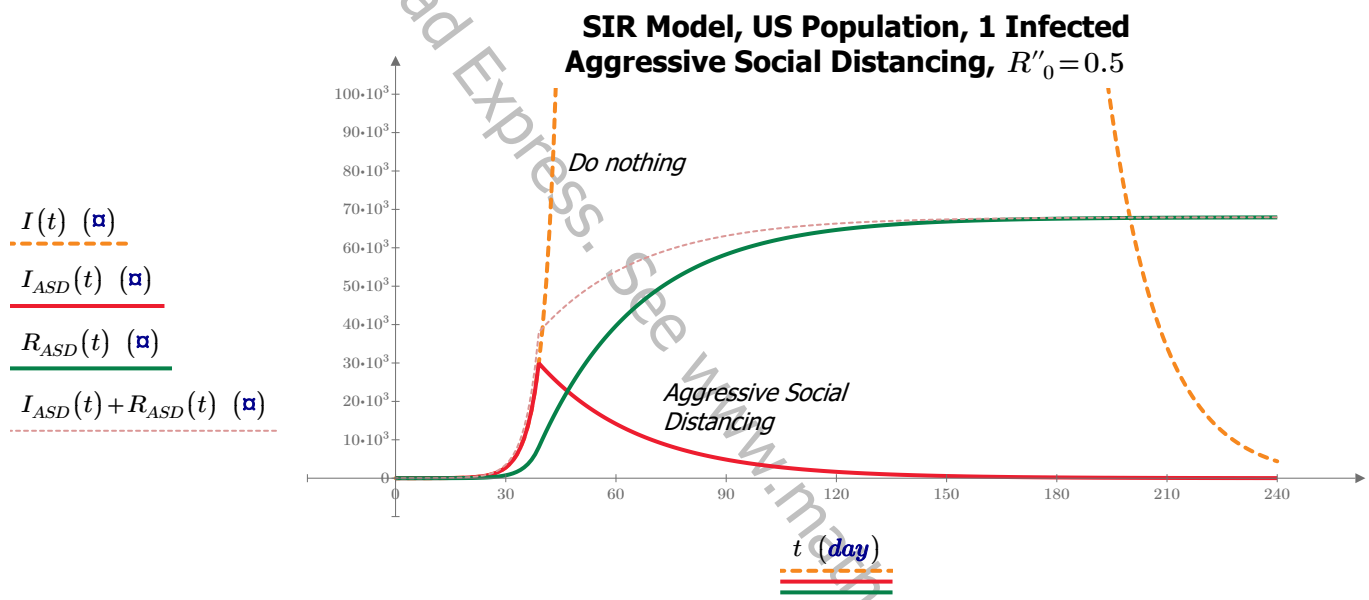
$$\begin{bmatrix} S'' \\ I'' \\ R'' \end{bmatrix} := SIR\left(S\left(39\ \textbf{\textit{day}}\right), I\left(39\ \textbf{\textit{day}}\right), R\left(39\ \textbf{\textit{day}}\right), R''_0, \gamma, 365\ \textbf{\textit{day}} - 39\ \textbf{\textit{day}}\right)$$

Adjusting the curve past day 39 for aggressive social distancing:

$$S_{ASD}(t) := S\left(min\left(t, 39\ \textbf{\textit{day}}\right)\right) + \Phi\left(t - 39\ \textbf{\textit{day}}\right) \cdot \left(S''\left(\max\left(0, t - 39\ \textbf{\textit{day}}\right)\right) - S\left(39\ \textbf{\textit{day}}\right)\right)$$

$$I_{ASD}(t) := I\left(min\left(t, 39\ \textbf{\textit{day}}\right)\right) + \Phi\left(t - 39\ \textbf{\textit{day}}\right) \cdot \left(I''\left(\max\left(0, t - 39\ \textbf{\textit{day}}\right)\right) - I\left(39\ \textbf{\textit{day}}\right)\right)$$

$$R_{ASD}(t) := R\left(min\left(t, 39\ \textbf{\textit{day}}\right)\right) + \Phi\left(t - 39\ \textbf{\textit{day}}\right) \cdot \left(R''\left(\max\left(0, t - 39\ \textbf{\textit{day}}\right)\right) - R\left(39\ \textbf{\textit{day}}\right)\right)$$

SIR Model, US Population, 1 Infected
Aggressive Social Distancing, $R''_0 = 0.5$

Legend:
- $I(t)$ (¤) — orange dashed
- $I_{ASD}(t)$ (¤) — red solid
- $R_{ASD}(t)$ (¤) — green solid
- $I_{ASD}(t) + R_{ASD}(t)$ (¤) — dashed

Do nothing

Aggressive Social Distancing

t ($\textbf{\textit{day}}$)

Now, we have made a HUGE difference! The epidemic is squelched and total infected is limited to:

$$I_{ASD\_total}(t) := I_{ASD}(t) + R_{ASD}(t)$$

$$I_{ASD\_total}\left(365\ \textbf{\textit{day}}\right) = 67966\ \textbf{\textit{persons}}$$

$$I_{ASD}\left(365\ \textbf{\textit{day}}\right) = 0.262\ \textbf{\textit{persons}} \qquad \text{Has to get to less than one!}$$

$$S_{ASD}\left(365\ \textbf{\textit{day}}\right) = 329.9\ M \cdot \textbf{\textit{persons}} \qquad \text{Remaining susceptible population.}$$

This curve resembles results that China is already seeing under strict isolation measures. Drastic measures to stem the spread of the virus are needed to make this kind of difference. As of March 20, only *"23% of the US Population is ordered to "stay at home" after governors in California (40 million residents), New York State (20 million), Illinois (13 million), and Connecticut (3.5 million) ordered nonessential workers to remain at home to slow the spread of coronavirus."* [10]

Immunization

It is too late for immunization to stem the tide as most of the population will be infected before a vaccine can be developed and administered. Vaccination does not, in fact, change the value of R_0, but reduces the size of the susceptible population, S_0. The model above demonstrates what happens when just one single infected person is introduced in to the susceptible population. Even if the world is successful through distancing measures, wiping this out completely will still surely require a vaccine against SARS-CoV-2 to ensure that the outbreak does not occur again with so many people remaining susceptible to another outbreak.

Conclusion

This is a simple model of the current Novel Corona Virus outbreak. More complex models (SEIR, etc.) can and are being used around the globe to figure out how to respond to the novel coronavirus outbreak. However, it clearly demonstrates how dire the situation is and how long it will take to get through it, even with aggressive isolation measures, which are not even in full swing yet around the globe. Unfortunately, because of the incubation/infectious period, all numbers in this analysis are at least two weeks behind the curve. Social distancing measures may not even be noticed in the reported numbers for two weeks after the measures have gone in to place.

China has made great strides to clamp down on the outbreak, many countries are following suit. However, it is fairly clear that new outbreaks could occur if distancing/isolation measures are not kept up. Without a vaccine/treatment, social distancing will need to be maintained; it will be the new "normal".

While US data was used in this analysis, the model can be applied to any population and any outbreak. The current crisis is a global one and its eradication will require global cooperation on a level never seen before. Our planet has suddenly gotten very small.

References

1. Kermack, W. O.; McKendrick, A. G. (1927). "*A Contribution to the Mathematical Theory of Epidemics*". Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences. 115 (772): 700. Bibcode:1927RSPSA.115..700K. doi:10.1098/rspa.1927.0118. JSTOR 94815.

2. Moore, Lawrence C.; Smith, David A., "The SIR Model for Spread of Disease" Duke development group of the Connected Curriculum Project (Duke CCP), Copyright 1997-2003, https://services.math.duke.edu/education/ccp/materials/diffcalc/sir/sir2.html

3. Jones, James Holland, "Notes on R0", Department of Anthropological Sciences, Stanford University, May 1, 2007, https://web.stanford.edu/~jhj1/teachingdocs/Jones-on-R0.pdf.

4. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. (January 2020). *"Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia"*. The New England Journal of Medicine. doi:10.1056/NEJMoa2001316. PMID 31995857.

5. Riou, Julien and Althaus, Christian L. (2020). "Pattern of early human-to-human transmission of Wuhan 2019 novel coronavirus (2019-nCoV), December 2019 to January 2020". *Eurosurveillance.* **25** (4). doi:10.2807/1560-7917.ES.2020.25.4.2000058. PMC 7001239. PMID 32019669.

6. Liu T, Hu J, Kang M, Lin L (January 2020). *"Time-varying transmission dynamics of Novel Coronavirus Pneumonia in China"*. bioRxiv. doi:10.1101/2020.01.25.919787.

7. Read JM, Bridgen JRE, Cummings DAT, et al. (January 28, 2020). *"Novel coronavirus 2019-nCoV: early estimation of epidemiological parameters and epidemic predictions"*. MedRxiv. doi:10.1101/2020.01.23.20018549

8. Delamater, Paul L., Erica J. Street, Timothy F. Leslie, Y. Tony Yang, Kathryn H. Jacobsen, "*Complexity of the BasicReproduction Number (R0)*", Emerging Infectious Diseases • www.cdc.gov/eid • Vol. 25, No. 1, January 2019 https://wwwnc.cdc.gov/eid/article/25/1/17-1901_article

9. Wallinga J., P. Teunis, "Different Epidemic Curves for Severe Acute Respiratory Syndrome Reveal Similar Impacts of Control Measures", American Journal of Epidemiology, vol 160: no. 6, September 15, 2004.

10. WorldOMeters web site: https://www.worldometers.info/coronavirus/

Some additional units:

$$Million \equiv 10^6 \qquad M \equiv Million$$

$$\% \equiv 1\%$$