1. **The basic “SIR” model**



**Define the variables:**



**The most basic model is:**



**Two key terms move people between categories:**

1. Number of newly infected (subtract from S, add to I):



* “*t*” is the TRANSMISSION coefficient
* the number infected is proportional to BOTH the number infected and the number susceptible
  + more people infected the more likely the disease spreads
  + more susceptible means more people the disease can infect

1. Number of recovered (subtract from I, add to R):



* “*w*” is the RECOVERY coefficient

1. **Example (see spreadsheet)**
   1. Initial values: school of 500 students with 2 infected on day 1:



* 1. Parameter values:



* 1. Example calculations for Day 2:



*So, we have 1.992 (round to 2) new infections but 1 recovers meaning I = 3 and R = 1 after day 2.*

* 1. Notes:
* In spreadsheet, susceptible formula modified to ensure cannot have a negative number (“max” function) which can happen if the transmission rate is very high
* Since we want a discrete (i.e. whole person) number we round the S,I in the spreadsheet (so day two we have 496 and 3 respectively in the S, I for the example)
* Rather than use the “R” formula, easiest to just use 500-S-I (otherwise the rounding can cause the number of students in the school to differ from 500)
* The results in this example are typical for how epidemics work:
  + an initial rise in the number infected which eventually reaches a peak, in our example around day 12/13
  + at this point the number infected begins to decrease – this is due to the fact that there are not enough remaining in the susceptible category
  + eventually an “equilibrium” is reached when the epidemic is essentially over
  + in our example, nearly all are infected (only 13 are not) since we just let the epidemic run its course without intervention
* The rounding to ensure integer numbers sometimes prevents the disease to reach 0 infected (as in our example); in a game if the number infected levels at some value for several days the epidemic could be ended and the infected set to 0 (move the remaining to the R category)

1. **Game considerations**

It is important to consider what the player knows and how the underlying model is used to inform and update at each time point. A few examples (there are more):

* Among the “S” group are some who might have symptoms, particularly in cases where the early disease symptoms are similar to other illnesses (common cold, flu). Some might even think they have symptoms due to fear. The attached press release is an example of what the epidemic disease might look like
* The “I” group has subjects in various stages of the disease. If early, they could be infected/contagious but either without any symptoms or just the symptoms that are more like a cold. At some point, the disease reaches a state where it is obvious that they have it and the player would then know it.
* The choice of which subjects move from S to I and from I to R could/should be impacted by other factors. In the example, there are 2 new infections. These are most likely students with most contact with the first 2 infected (i.e. in the same class/grade etc). Likewise, the 1 recovered would be the one infected the longest (although not always as people recover at different paces).
* Note that we assume once in the “R” status the student does not get the disease again – this is true for diseases like measles/mumps for the most part. In a deadly outbreak some in the “R” die but we want to have a less fatal game
* Some in the “S” group might be IMMUNE to the disease and therefore not truly “susceptible”
* Within the basic model are places where randomness could be introduced
* In real outbreaks a number of possible methods to address. A few:
  + Quarantine – can cause more panic and is disruptive as it would, in the case of schools, mean missed classes and even a shutdown. 100% effective IF people actually follow the rules AND if quarantine people before they are contagious and contact others (may be difficult to identify when to quarantine if symptoms not present; quarantine of healthy people will cause them to be upset!)
  + Vaccines – not always available, particularly immediately. May not be 100% effective and must be administered prior to becoming infected. Some people don’t like and concerns about side effects etc. Cost trade off – how many fewer infected for the price?
  + Treatments – may be drugs already on the market that move people more rapidly from “I” to “R” (ie improve the recovery rate). If a new disease, efficacy of these may be low and it will take time/money to produce a better drug.

1. **Simple statistics ideas**

* Very simple statistics: use of descriptive statistics/graphs to help identify where the disease is centered to make decisions about how to address the outbreak
* Hypothesis testing: this was the simple game we tried to produce in previous grant. The idea was there were two drugs that might be effective in treating the disease but with a better efficacy – the player is collecting data and testing the hypothesis that the drugs are essentially the same. The efficacy/costs for each could be variables that change each time the game is played. From a teaching standpoint, this a great way to introduce the idea of statistical vs. clinical significance. For example, the player might determine one drug is statistically better…but if it costs a lot more and is only going to slightly improve the recovery rate is it worth it? There is a bit of a moral question here too…we could reduce the number infected by a handful but at a great price.

1. **More advanced statistics ideas (a game with these possibilities would be potentially very interesting to many I think!)**

* Decision analysis type of issues, using conditional probability and introducing “Bayes Rule”
  + In this version we have one or more tests that can be used to determine if students have the disease (“screening tests”)
  + This is important given that we have a number of infected without symptoms that may be contagious – we want to identify and treat them quickly
  + The issue is that tests, particularly the less expensive “screening” type have error rates (false positives, negatives)
  + The underlying concept of interest is PREVALENCE (ie denominator) is important! An example. Suppose a screening test will return the correct result 95% of the time (pretty good!). In our school with 500 students on day 1 we have 2 infected. If we test ALL 500 with the screening test, the test will likely be positive for both the infected students. However, it will incorrectly identify about 5% of the 498 healthy students as “I” (that is, 25). So, we will think 27 are “I” but only 2 of them actually are! If the disease if more prevalent the ratio of false to true “I” will not be as bad.
  + In the game, students understanding Bayes Rule/conditional probability might be more cautious in treating when the likelihood of false positives is high.
  + In real disease detection a good screening test (95%) typically is used to help identify those for a more accurate test such as a blood test. The reason not to use the more accurate test initially and for all people is cost/resources (must draw blood and send to the lab)