


# Mathematical Models and Infectious Disease Dynamics

Mark Lewis[†]

Mathematical models can be used to understand what factors govern infectious disease outbreaks including HIV/AIDS, West Nile virus, and even the bubonic plague! The purpose of the model is to take facts about the disease as inputs and to make predictions about the numbers of infected and uninfected people over time as outputs.

Factors that can go into the models include the length of time one is ill, the length of time one can infect others (often different than the total length of time one is ill), the level of contagiousness of the disease (i.e., the likelihood of infecting another individual if one comes into close contact), the number of uninfected (susceptible) individuals, and so forth. Mathematical modellers then feed these facts into a set of equations.

In the three accompanying articles (see pages 5–17), the equations used are referred to as differential equations. Differential equations involve derivatives of functions. For example, if the number of infected individuals at a given time $t$ is represented by function $I(t)$, then the time derivative of $I$, denoted by $I'(t)$, says how quickly $I(t)$ is changing. If $I'(t) > 0$, the disease is increasing with time and if $I'(t) < 0$, the disease is decreasing with time.

Being able to make predictions about disease dynamics is really helpful for public health. If we know there will be an outbreak we can prepare for it. Alternatively, if we have a reliable model, we can study how to prevent an outbreak and save lives by changing the factors we can control using public health means. These factors include education, immunization, quarantine regulations, and health treatment strategies.

The application of mathematical models to infectious disease dynamics has been a real success story in $20^{\text{th}}$ century science. Even though the dynamics of disease appear to be very complex, surprisingly simple mathematical models can be used to understand features governing the outbreak and persistence of infectious disease.

Early models, such as the one by Kermack–McKendrick (see Fred Brauer's article on pages 10–13), were applied to understand the dynamics of historical diseases, such bubonic plague. Amongst other things, the models could be used to predict the fraction of the population that would survive a major disease outbreak.

Here the form of the mathematical model is a system of differential equations (say, one equation to track the levels of each of the susceptible $S(t)$, infective $I(t)$, and previously infected but now recovered or dead $R(t)$ portion of the population). The equations are 'coupled' to each other, because the growth of infectives relies on having new susceptibles to infect, the growth of removed (recovered or dead) relies on having had infected individuals, and so forth.

Such models can be easily put on the computer, using software such as Maple, to make predictions about changing levels of disease outbreak over time. Alternatively, as outlined in the article by Fred Brauer, mathematics can play a more central role. Analysis of the models yields useful statistics about the disease, such as the 'basic reproduction number'—a measure of infectivity, expressed as the number of secondary infections arising directly from a single infective individual surrounded by susceptibles. Methods to control the disease are then summarized by the degree to which they able to reduce the basic reproduction number to a number less than one.

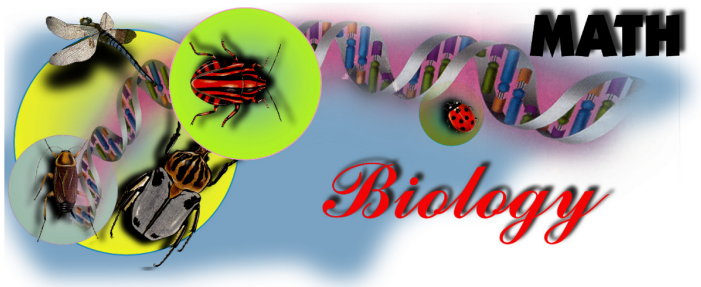
Recurrent $20^{\text{th}}$ century diseases, such as childhood measles, can be analyzed using a similar framework (see David Earn's article on pages 14–17). Here models are extended to include birth and death of individuals and seasonal and other changes in the levels of contact, for example, between school-age children. Trends in birth rates and effects of immunization can also be incorporated into the models. With the analysis of these augmented models, new questions can be posed and answered, varying from: "why do diseases exhibit complex temporal patterns, ranging from cycles to erratic, seemingly chaotic fluctuations?" to "what is the optimal immunization strategy for a given disease, given finite resources and vaccine availability?"

Much of the recent mathematical work in disease modelling has focused on emerging diseases, ranging from HIV/AIDS, to SARS and West Nile virus (see Marjorie Wonham's article on pages 5–10 and the last part of Fred Brauer's article). Here recommendations regarding control methods are needed urgently. For most emerging diseases, modellers do not have the luxury of comprehensive data sets showing outbreak levels over time. Therefore the models must be developed based on detailed understanding of the components of the disease dynamics and from our experiences with modelling previous diseases. Here predictions and recommendations for control stem from the mathematical and numerical analysis of the models (see Marjorie Wonham's article). The modelling of emerging disease is the current challenge for mathematical epidemiologists, and it is one that will be with us for some time to come.




©Copyright 2004
Wieslaw Krawcewicz

[†] **Mark Lewis** is the Canada Research Chair in Mathematical Biology and a Professor in the Department of Mathematical and Statistical Sciences at the University of Alberta. His e-mail address is mlewis@math.ualberta.ca.




# The Mathematics of Mosquitoes and West Nile Virus

Marjorie Wonham[†]

Lying on my sleeping pad, I warily eye the mosquito perched above my head. I could reach up and squash it, but that would require extracting my arm from the warmth of my sleeping bag. So for now, it clings to the yellow nylon of my tent, unaware of its reprieve. Although I may triumph over this particular mosquito, I am all too aware of being vastly outnumbered outside my tent.

Until recently, my interest in mosquitoes was largely pragmatic: avoid, repel, or swat. Lately, though, I have developed a grudging curiosity about how they make a living. Fact: a female mosquito overwinters with fertilized eggs so the first thing she does in spring, before even feeding, is lay eggs. Fact: Even if she doesn't find a blood meal, she can survive by sucking plant juices. Fact: the combined meals of a mosquito horde can (and do) bleed a newborn calf to death. Fact: at best a mosquito bite simply itches; at worst, it means disease transmission—malaria, yellow fever, dengue fever, and West Nile virus are all mosquito-borne.

It is West Nile virus that has piqued my interest of late. First identified in Uganda in 1937, the virus is well established in its native Africa where it lives primarily in birds and is transmitted among them by mosquitoes. Only occasionally does a mosquito transmit the infection to a mammal. From time to time a West Nile virus outbreak occurs in Europe and Africa—in Israel in the 1950s and South Africa in 1974, and more recently in Romania, Morocco, Tunisia, Italy, France, and Russia. Just recently, West Nile virus made its first known, and headline grabbing, North American appearance.

In the summer of 1999, the birds of New York City began mysteriously to die, their bodies appearing conspicuously in the city zoo, parks, and backyards. At first the cause was unknown, but by December of the same year it had been identified, in two reports published in the same issue of Science magazine, as West Nile virus, a disease never before seen on this continent. In subsequent summers, West Nile virus spread west across the continent reaching Ontario in 2001, California and Washington in 2002, and Alberta in 2003.

Corvids—crows and jays—were the hardest hit among the birds; other passerines such as sparrows also carried the virus but were dying in smaller numbers. Among mammals, horses appeared especially vulnerable, with a mortality rate of approximately 40%. Human cases were less common and less likely to be fatal, but were a growing health concern nonetheless. By the end of 2003, the virus had been identified in 7 Canadian provinces and 46 U.S. states, in at least 10 mosquito, 150 bird, and 17 other vertebrate species, and in a total—for that year alone—of over 11,000 human cases in Canada and the U.S.

In Alberta, West Nile virus was first reported in 2003. The year before, I had moved to Edmonton to join a group of mathematical biologists at the University of Alberta. There, I was immersed in a world of mathematical modelling used to tackle biological questions. Two years later, here in my tent on a canoe trip in the Northwest Territories, I have a chance to reflect on the unpredicted collaboration that developed with my mathematical colleagues on the dynamics and control of West Nile virus.

At first glance, we made an unlikely trio for this project. Tomás was a graduate student and programming whiz who studied chamomile invasions on farmland. Mark was a mathematical biologist who had modelled the movement of birds and wolves. And I was a marine ecologist with a mathematical background largely limited to reading tide tables. But together we were galvanized by a question from a colleague in Ontario: *"How come no one is modelling West Nile virus?"* Hugh asked. And with that casual question began our Year of the Mosquito.

How could mathematical modelling help us understand West Nile virus dynamics? The virus was spreading, and control proposals were beginning to include spraying adult and larval mosquitoes, removing larval mosquito habitat, and even removing birds. Since all of these strategies would cause additional impacts on the environment, perhaps modelling could help maximize control effectiveness while minimizing unwanted effects? I had no idea where to begin, but luckily, I was in good mathematical hands.

Mark and Tomás introduced me to a class of disease models known by their acronym as SIR models (see accompanying articles by Fred Brauer and David Earn). These models were first extended to vector-borne diseases by R. Ross in the early 1900s and G. Macdonald in the 1950s, to combat malaria. Since then, a large associated body of mathematical theory, and an impressive history of contributing to disease control, have both evolved.

In an SIR model, the host population is divided into three groups: Susceptible (healthy uninfected individuals), Infectious (infected and capable of transmitting the disease), and Removed (immune, dead, or otherwise removed individuals). The rates at which an average individual moves from Susceptible to Infectious and from Infectious to Removed are determined, and the relevant birth and death rates are incorporated.

Once constructed, a key piece of information can be extracted from an SIR model, called the disease *reproduction number*, or $\mathcal{R}_0$ ("R-zero" or "R-naught"). $\mathcal{R}_0$ tells us the number of new infections that would result from the introduction of a single infectious individual into an entirely susceptible population. For example, if a student with chicken pox walked into a classroom of individuals with no previous exposure to the disease, how many new cases would be caused by direct contact with the initial infectious student? The answer is given by $\mathcal{R}_0$. Or for West Nile virus, if an infectious bird arrived in a new city, how many other birds would be infected (via mosquito bites) by that original bird? Again, $\mathcal{R}_0$ tells us the answer.

The expression for $\mathcal{R}_0$ is constructed, according to a particular formula, from variables and parameters in the model. Reasonably enough, it takes into account factors such as how long the first individual remains infectious, the likelihood of contact between the infectious and susceptible individuals (either directly or via another species), and how often contact leads to disease transmission. Details of the mathematics underlying the calculation of $\mathcal{R}_0$ are given in the article by Fred Brauer.

[†] **Marjorie Wonham** is a Postdoctoral Fellow in the Centre for Mathematical Biology and the Department of Mathematical & Statistical Sciences at the University of Alberta. Her e-mail address is mwonham@ualberta.ca.



BY SWITCHING TERMS IN HER EQUATIONS, MARJORIE FOUND SHE COULD REVERSE THE WEST NILE VIRUS EPIDEMIC.

©Copyright 2004
Gabriela Novakova

For disease control, the value of $\mathcal{R}_0$ is key. $\mathcal{R}_0 < 1$ means that an infectious individual will, on average, generate fewer than one new infection, so the disease will die out even without control efforts. On the other hand, $\mathcal{R}_0 > 1$ means that an infectious individual will generate more than one new infection, so a disease outbreak will occur. In this case, the disease can be controlled by methods that alter one or more of the model components—such as mortality, contact, or transmission rates—to reduce $\mathcal{R}_0$ below one.

Armed with this mathematical background, we were ready to develop an SIR model for West Nile virus in North America, calculate $\mathcal{R}_0$, and identify how and how much to control the disease to prevent an outbreak. First, we had to define the biological and geographic scope of our efforts. We had already learned that the virus persisted in transmission cycles between mosquitoes (vectors) and birds (reservoir hosts). Although it occasionally spread to other vertebrates (including humans), it seemed not to return to mosquitoes. In other words, although an infection might be deeply significant to the human in question, it would not influence the overall disease dynamics.

This biological fact helped us simplify the mathematical description: by viewing the virus outbreak level in mosquitoes and birds as a proxy for the human infection risk, we could limit our model to only the vector and the reservoir host. For mosquitoes, we focused on the species group that was emerging as the dominant North American vector (*Culex pipiens* spp.). For birds, we focused on the species with the best available infection and mortality data, the American crow (*Corvus brachyrhyncos*). And since the best virus prevalence data were available where West Nile had first appeared, we confined ourselves to modelling the New York City outbreak. Finally, since the disease outbreak at this latitude showed a marked seasonality, appearing in summer and disappearing in winter, we confined the model to a single season from spring through fall.

Soon we had the skeleton of a model. We defined three groups ($S$, $I$, and $R$) for both mosquitoes and birds. From the literature, we obtained estimates of the mosquito biting rate and the transmission probabilities that allowed a mosquito to infect a bird (which was around 88%) and a bird to infect a mosquito (which was only around 16%). We had recovery and mortality rates for birds, but not for mosquitoes since they didn't seem to be affected by the disease. Since we had birth and death rates for mosquitoes but not for birds, we assumed the birds reproduced once in spring before the model began, and their background (natural) mortality rate would be negligible in the one summer.

To refine the model and assign numerical values to the parameters, we divided up the work according to our expertise. Tomás investigated how to solve and simulate these models on a computer, and Mark explored the mathematical theory underlying this approach. I searched the biological literature for parameter values for mosquitoes, birds, and virus transmission. I didn't envy Tomás and Mark, as my task seemed much the easiest; I was surprised to learn later that they felt the same way about their roles.

Nonetheless, tracking down the biological data took some sleuthing. Today, from the safety of my tent, I can make a guess as to the local mosquito abundance outside: it's very, very high. If I were to stick out a bare arm, it would be covered almost instantly; if I left it out, I could watch an individual probe repeatedly, biting several times before finding her blood meal. (I don't do this often.) As I canoe down the river, I sometimes see a cloud of mosquito larvae rising to the surface and hatching.

For our model, though, we needed observations like this to be quantified. Just how many mosquitoes were out there? How many eggs did one female lay, and how many larvae hatched? How many crows were there? How many mosquito bites per crow in a day? And how contagious was the virus? My quest was something of a scavenger hunt travelling back through biological history, with each paper leading me to an older one. At the same time, new reports about the virus were appearing almost daily in print and online. I divided my time between dusty library shelves and internet listserves.

New biological information, new equations, new parameter estimates, new model analyses, and new simulations surfaced. For several months, we worked to tailor the model as best we could to the biological information. In the interests of tractability, the model had to remain as simple as possible.



WE DO NOT RECOMMEND SWITCHING VARIABLES IN THE EQUATIONS!

©Copyright 2004
Wieslaw Krawcewicz

But the biological complexity seemed almost infinite. At times, I wondered if we were simplifying too much for the model to be informative, while Tomás and Mark wondered if the biology was making the model too complicated to be useful.

In the end, the biology dictated two substantial additions to the model. We learned that mosquitoes could spend quite a long time, up to 14 days, as aquatic larvae that don't bite birds and therefore don't transmit the disease. It can also take quite a long time, perhaps 10–12 days, for an infected mosquito to develop a viral load high enough to transmit the disease back to a bird. These two time periods could add up to almost half a mosquito's lifespan, so they could substantially alter the disease dynamics. We therefore added two new groups to the population: one for larval $L$ and one for exposed $E$ mosquitoes.

With these additions, it seemed we finally had a model we all felt was realistic and tractable. This is the model illustrated in Figure 1. Now that we had the model, it was time for a test: would it behave realistically? We had to know this before we could use $\mathcal{R}_0$ to make any predictions.

As a test, we chose the records of West Nile virus incidence in both mosquitoes and birds from New York City in 2000. We plugged our literature-based parameter estimates into the model, crossed our fingers, and ran the simulations. Sure

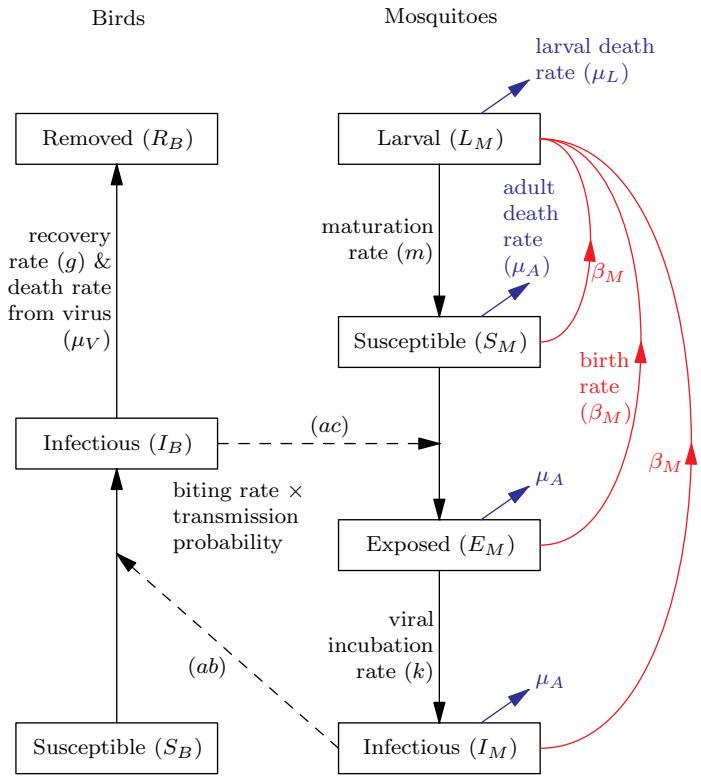



**Figure 1:** Cartoon illustration of SIR model for West Nile virus cross-infection between birds and mosquitoes. The boxes represent *variables*: the number (or proportion) of birds that are susceptible, infectious, and removed, and mosquitoes that are larval, susceptible, exposed, and infectious. The arrows represent *parameters*: the daily rates at which the value of each variable changes. Solid black lines show the rates (and reasons) for individual birds and mosquitoes moving from one category to another. The two dashed lines represent the rates at which mosquitoes infect birds and birds infect mosquitoes. The red and blue arrows respectively show the birth and death rates for mosquitoes (which can reproduce multiple times during a single season).

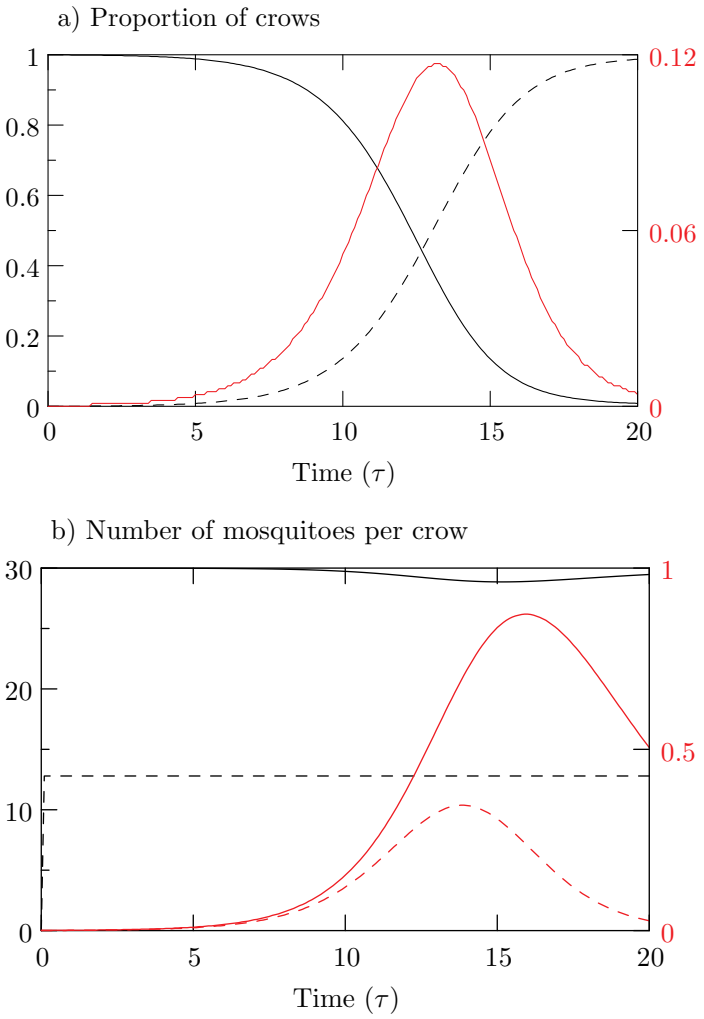



**Figure 2:** Numerical simulations of crow proportions and mosquito numbers predicted by the West Nile virus model for a steady-state mosquito population ($S_{M_0}/N_{B_0} = 30$) over a season lasting $14\tau$ days, where $\tau$ is a time unit representing 9.4 days (this time unit may seem funny, but it is a result of manipulating the equations to a nondimensional form—a detail that is not shown here). a) the proportion of susceptible (solid line) and removed birds (dashed line) on the left axis and infectious birds (red line) on the right. b) the number of susceptible (solid line) and larval mosquitoes per bird (dashed line) on the left axis and exposed (dashed red line) and infectious mosquitoes per bird (solid red line) on the right.

enough, for a reasonable range of starting values the model predicted a range of mosquito and bird infection levels that matched the observed data (Figure 2). We were now in a position to use the model to investigate control strategies.

Following methods developed by other mathematicians, we calculated $\mathcal{R}_0$ for our model, which turned out to be made up of three elements:

$$\mathcal{R}_0 = \frac{ab}{\mu_A} \cdot \frac{ac}{\mu_V + g} \cdot \frac{S_{M_0}}{N_{B_0}} \left( \frac{k}{k + \mu_A} \right).$$

The parts work together like this. The first factor represents the number of new crow infections caused by an infectious mosquito. This is the number of bites per day per mosquito $a$ multiplied by the probability of the virus being transmitted from an infectious mosquito to a crow $b$ multiplied by the average number of days $1/\mu_A$ a mosquito would live.

The second factor is the mirror image: the number of new mosquito infections caused by an infectious crow. The mosquito biting rate $a$ is multiplied by the transmission probability from crows to mosquitoes $c$ and by the average number of days until the infectious crow either recovered $(1/g)$ or died $(1/\mu_V)$.

The third factor represents, generally speaking, the number of infectious mosquitoes per crow. Specifically, it is the number of initially susceptible mosquitoes $S_{M_0}$ that survive the virus exposure period $k/(k+\mu_A)$ for every bird $N_{B_0}$. Together, these three elements give us the expression for the total disease $\mathcal{R}_0$ from birds to birds (via mosquitoes) or from mosquitoes to mosquitoes (via birds). Taking the square root of the right hand side of the expression for $\mathcal{R}_0$ is a common convention that gives the geometric mean $\mathcal{R}_0$ from bird to mosquito and *vice versa*.

(If you're looking at Figure 1 and wondering why the mosquito birth rate and larval death rate don't seem to show up in $\mathcal{R}_0$, rest assured. They are accounted for by the equality $L_{M_0} = \beta_M S_{M_0}/(m+\mu_L)$, a simplifying assumption in the model that ensures a constant mosquito population.)

Our parameter values gave $\mathcal{R}_0$ greater than one, predicting a disease outbreak. We were now able to return to our original questions, namely, how could West Nile virus be controlled, and how much control was needed?

One possible answer was obvious: if every mosquito and bird were removed, the virus could not persist. But we were hoping to find a more palatable answer. Examining the expression for $\mathcal{R}_0$, we found the ratio $S_{M_0}/N_{B_0}$ in the numerator, telling us that reducing the number of mosquitoes $S_{M_0}$ would reduce $\mathcal{R}_0$, but reducing the number of birds $N_{B_0}$ would not. In fact, reducing bird abundance would only make things worse by increasing the value of $\mathcal{R}_0$. This was our first lesson: *reducing mosquitoes could help control the virus, but removing birds would only increase the chances of an outbreak.*

Would every mosquito have to be removed to prevent an outbreak? This was our second lesson. Plotting a graph of $\mathcal{R}_0$ *vs.* $S_{M_0}/N_{B_0}$ showed us that *the virus could be controlled simply by reducing mosquito abundance, without requiring that every last individual be eliminated* (Figure 3).

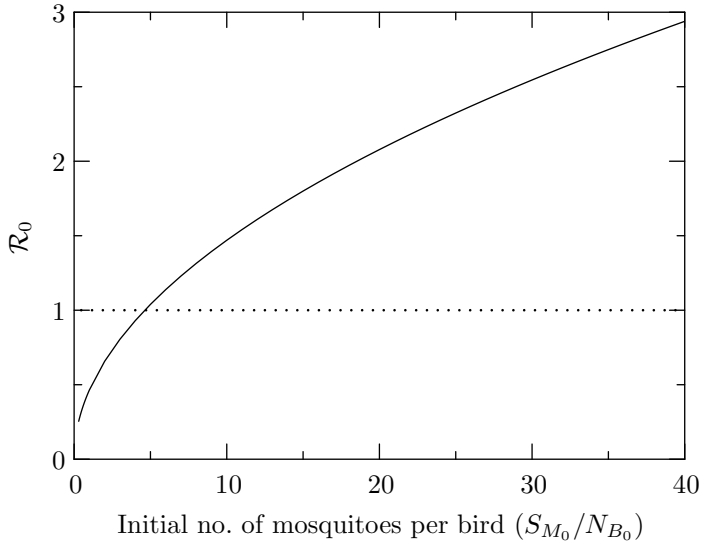



**Figure 3:** Plot of $\mathcal{R}_0$ *vs.* the initial number of mosquitoes per bird $S_{M_0}/N_{B_0}$, showing that the mosquito population can simply be reduced, and not completely eliminated, to bring $\mathcal{R}_0$ below one and therefore prevent a West Nile outbreak.

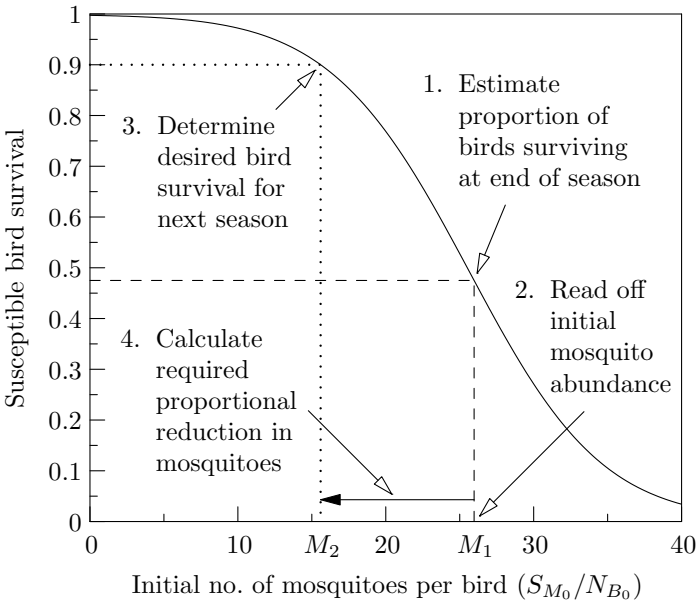



**Figure 4:** Plot of susceptible bird survival at the end of the season $S_B$ *vs.* the initial number of mosquitoes per bird at the beginning of the season $S_{M_0}/N_{B_0}$, showing how future mosquito control efforts can be estimated from previous bird mortality data.

This begged the final question of exactly how much mosquito reduction was enough. It took us a while to work this out, but eventually we came up with an indirect route to the answer. It requires the following four steps, which are illustrated in Figure 4.

1. Obtain from public health authorities an estimate of the proportion of crows that survived the virus in the last season. (These data are becoming increasingly available.)

2. Determine how many mosquitoes per crow $M_1$ there must have been at the start of the last season to cause the observed bird mortality.

3. Determine the maximum allowable bird mortality for the following season, and the corresponding maximum allowable initial number of mosquitoes per crow $M_2$.

4. The ratio of these two mosquito-per-crow estimates, $M_2/M_1$, gives the proportional reduction in mosquitoes needed to control the virus in the next year.

This was our third lesson from the model, that *we could calculate the required amount of mosquito reduction.* In fact, we found that a 40–70% reduction in the New York City mosquito population would have prevented the 2000 outbreak. There is, of course, a catch with this approach: *the virus has to have already caused the initial mortalities that are used for subsequent control estimates.* However, as the virus continues to spread, bird mortality data are increasingly available. Perhaps the experience of cities and regions with the virus (and therefore with the data) can inform the control programs in areas where a virus outbreak has not yet occurred.

Like many research projects, this one raised more questions than it answered. How fast could West Nile virus spread across the continent and where would it go? How important was climate in determining the survival of birds, mosquitoes, and the virus? How would controlling the mosquitoes in an

urban area influence the surrounding rural area? How was our model similar or different compared to other models of similar diseases? Was there anything more general that could be learned about the epidemiology of vector-borne viruses? By working with some additional collaborators, we have begun to address these questions too.

In the meantime, I have an active interest in some very local mosquito control. Earlier this evening, I noticed that the spiders in the rocks behind my tent were making a killing, literally. Mosquitoes were landing in their webs so fast the spiders could hardly keep up. So now I'm starting to wonder how many spiders I would need to keep in my tent to control the mosquitoes. Or if not spiders, perhaps a bat would be more efficient? Clearly, I need another model. I'll have to consult with my collaborators.

## Appendix

| **Bird equations:** | |
|---|---|
| Susceptible | $\frac{dS_B}{dt} = -abI_M\frac{S_B}{N_B}$ |
| Infectious | $\frac{dI_B}{dt} = abI_M\frac{S_B}{N_B} - \mu_V I_B - gI_B$ |
| Removed | $\frac{dR_B}{dt} = (g+\mu_V)I_B$ |
| **Mosquito equations:** | |
| Larval | $\frac{dL_M}{dt} = \beta_M(S_M+E_M+I_M) - mL_M - \mu_L L_M$ |
| Susceptible | $\frac{dS_M}{dt} = -acS_M\frac{I_B}{N_B} + mL_M - \mu_A S_M$ |
| Exposed | $\frac{dE_M}{dt} = acS_M\frac{I_B}{N_B} - kE_M - \mu_A E_M$ |
| Infectious | $\frac{dI_M}{dt} = kE_M - \mu_A I_M$ |

**Table 1:** Ordinary differential equations (ODEs) representing the transmission of West Nile virus between birds and mosquitoes. Each equation keeps track of the change in a variable: the number of susceptible $S_B$, infected $I_B$, and removed $R_B$ birds, and the number of larval $L_M$, susceptible $S_M$, exposed $E_M$, and infected $I_M$ female mosquitoes per bird. Change in these variables is represented by parameters, which are probabilities and rates. (Note that all variables are indicated by capital letters, and all parameters by lower-case letters.) These equations match up with the boxes and arrows in Figure 1. To create your own numerical simulation of this model, you will need to choose the initial conditions, or values, for each variable. For a West Nile virus invasion, you might start with all variables equal to zero, except for $I_B$ or $I_M$, which could be a very small positive number. Then you can set up the ODEs to change the values of the variables in each time step. What initial value of $I_B$ or $I_M$ do you need in order for the infection to increase? How do your results change if you allow birds to recover (i.e., set $g > 0$)?

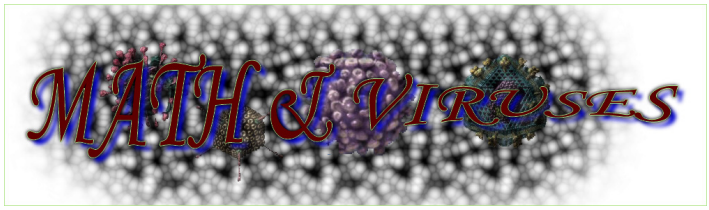
| Parameter | Mean (range) | Biological meaning |
|---|---|---|
| $a$ | 0.09 (0.03–0.16) | Mosquito daily per capita biting rate on birds |
| $b$ | 0.88 (0.80–1.00) | Probability per bite of West Nile virus being transmitted by an infectious mosquito biting a bird |
| $c$ | 0.16 (0.02–0.24) | Probability per bite of West West Nile virus being transmitted by a mosquito biting an infectious bird |
| $\beta_M$ | 0.537 (0.036–42.5) | Mosquito daily per capita birth rate |
| $m$ | 0.068 (0.051–0.093) | Mosquito larva daily per capita maturation rate to susceptible adults |
| $\mu_A$ | 0.029 (0.016–0.070) | Adult mosquito daily per capita mortality rate |
| $\mu_L$ | 1.191 (0.213–16.9) | Larval mosquito daily per capita mortality rate |
| $k$ | 0.106 (0.087–0.125) | Mosquito daily per capita transition rate from exposed to infectious |
| $\mu_V$ | 0.143 (0.125–0.200) | Bird daily per capita mortality rate from West Nile virus |
| $g$ | 0 | Bird daily per capita recovery rate from West Nile virus |

**Table 2:** Sample parameter values (and ranges) estimated from the literature for the West Nile virus model shown in Figure 1 and Table 1. The phrase *per capita* is a common biological term meaning per individual.

## Further Information


[1] This article is based on the paper: Wonham, M., T. de-Camino-Beck and M. Lewis, 2004: An epidemiological model for West Nile virus: Invasion analysis and control applications. *Proceedings of the Royal Society of London B*, **271**, 501–507.

[2] Read more about this and other research at the University of Alberta's Centre for Mathematical Biology: http://www.math.ualberta.ca/~mathbio/.




# What does Mathematics Have to do with SARS?

Fred Brauer[†]

At least since the beginning of recorded history there have been epidemics. One of the plagues that Moses brought down upon Egypt described in the Book of Exodus was murrain, an infectious cattle disease, and there are many other biblical descriptions of epidemic outbreaks. The Black Death (thought to be bubonic plague) spread from Asia through Europe in several waves beginning in 1346, causing the death of one-third of the population of Europe between 1346 and 1350 and recurring regularly in Europe for more than 300 years, notably as the Great Plague in London of 1665–1666. Recurring invasions of cholera killed millions in India in the $19^{th}$ century. The influenza epidemic of 1918–19 killed 20 million people overall and more than half a million in the United States. More recently, the SARS epidemic of 2002–3 caused worldwide concern and even more recently several strains of avian flu have forced the killing of millions of birds and worries about spread to humans.

Diseases that are *endemic* (always present), especially in less developed countries, have effects that are probably less widely known but may be of even more importance. Every year millions of people die of measles, respiratory infections, diarrhoea and other diseases that are easily treated and not considered dangerous in the Western world. Diseases such as malaria, typhus, cholera, schistosomiasis, and sleeping sickness are endemic in many parts of the world. The effects of high disease mortality on mean life span and of disease debilitation and mortality on the economy in afflicted countries are considerable. There are many useful practical conclusions that have been drawn from models for endemic diseases. One example was the possibility of eliminating smallpox worldwide by vaccination; this was successfully achieved in 1977. However, to keep the mathematics relatively simple in this article, we shall discuss only models for epidemic diseases. We hope that the reader who finds this introduction to the modelling of epidemics interesting may be motivated to study enough additional mathematics to learn about models for endemic diseases.

Perhaps the first epidemic to be examined from a modelling point of view was the Great Plague in London. The Great Plague killed about one-sixth of the population of London. The village of Eyam near Sheffield, England suffered an outbreak of bubonic plague in 1665–1666 whose source is believed to be the Great Plague. The Eyam plague was survived by only 83 of an initial population of 350 persons. There were actually two epidemics in Eyam and the first phase was survived by 261 persons. As detailed records were preserved and as the community was persuaded to quarantine itself to try to prevent the spread of disease to other communities, the second phase of the epidemic in Eyam has been used as a case study for modelling. The actual data for the Eyam epidemic are remarkably close to the predictions from the simple model as described below.

[†] Fred Brauer is a professor in the Department of Mathematics, University of British Columbia. His e-mail address is brauer@math.ubc.ca.

If there is no vaccine or treatment available for a disease, the only control strategies available are isolation of individuals diagnosed with the disease and quarantine of suspected infectives. The rates of isolation and quarantine may be varied, depending on decisions about the amount of effort to invest in these strategies. It is rarely possible to compare possible control strategies such as the division of efforts into isolation and quarantine during an actual epidemic. For this reason, mathematical modelling of epidemics is a promising tool for comparison of possible strategies. If and when a vaccine is developed, models might indicate whether it is more urgent to concentrate on vaccination or isolation.

One of the early triumphs of mathematical epidemiology was the formulation in 1927 of a simple model by a public health physician, W. O. Kermack, and a biochemist, A. G. McKendrick, predicting behaviour similar to that observed in countless epidemics, namely that diseases develop suddenly and then disappear just as suddenly without infecting the entire community. Kermack and McKendrick considered the class $S$ of individuals susceptible to the disease, that is, not yet infected, and the class $I$ of infected individuals, assumed infective and able to spread the disease by contact with susceptibles. In their model, individuals who have been infected and then removed from the possibility of being re-infected or of spreading infection are ignored. Removal is carried out through isolation from the rest of the population, immunization against infection, recovery from the disease with immunity against reinfection, or through death caused by the disease. These characterizations of removed members are quite different from an epidemiological perspective and of course also from a human point of view, but are equivalent from a modelling point of view that takes into account only the state of an individual with respect to the disease.



Angry SARS virus

©Copyright 2004
Wieslaw Krawcewicz

The Kermack–McKendrick epidemic model makes very simple assumptions about the rates of disease transmission and removal. One of the assumptions is that the disease is transmitted from one individual of a population to another by direct contact. Thus it is not applicable to diseases that are transmitted by a *vector*, that is, diseases transmitted back and forth between two populations such as mosquitoes and birds, as in West Nile virus. However, the ideas that go into the formulation of the Kermack–McKendrick model are also useful for the formulation of more complicated epidemic models.

The model contains only two parameters (the values of which are to be determined from observed data) and could be applied to many diseases transmitted by direct contact. While a more detailed model might be a better description of a specific disease, it would require more parameters. Since data are often incomplete and inaccurate because of under-reporting and mis-diagnosis at the beginning of an epidemic, a simple model may give better predictions.

The Kermack–McKendrick model, which is a deterministic compartmental model, is formulated in terms of the rates of flow of members of the population between compartments. There is a flow from $S$ to $I$ representing the rate of new infections and a flow out of $I$ representing the rate of recovery or disease death. Mathematically, these rates of change are described as *derivatives* with respect to time $t$. We will use $dS/dt$ to denote the derivative of $S$ and $dI/dt$ the derivative of $I$, thinking of $S$ and $I$ as functions of time $t$. There are techniques for calculating the derivative of a given function, but our situation is that we want to set up equations for the derivatives of the functions $S$ and $I$ and draw some conclusions about the behaviour of the functions.

The specific assumptions about the flow rates are as follows:

(i) An average infective member of the population makes contact sufficient to transmit infection with $\beta N$ others per unit time, where $N$ represents total population size.

(ii) A fraction $\gamma$ of infectives leave the infective class per unit time.

(iii) There is no entry into or departure from the population, except possibly through death from the disease.

According to (i), since the fraction of contacts by an infective with a susceptible, who can then transmit infection, is $S/N$, the number of new infections in unit time per infective is $(\beta N)(S/N)$, giving a rate of new infections $(\beta N)(S/N)I = \beta SI$. Fortuitously, we need not give an algebraic expression for $N$ since it cancels out of the final model. The hypothesis (ii) says that the infective periods are exponentially distributed with mean infective period $1/\gamma$. The hypothesis (iii) really says that the time scale of the disease is much faster than the time scale of births and deaths, so that demographic effects on the population may be ignored.

When these assumptions are translated into mathematical statements of the transition rates between classes, the result is a pair of equations, called differential equations, for the derivatives $dS/dt$ and $dI/dt$. These equations are

$$\frac{dS}{dt} = -\beta SI,$$
$$\frac{dI}{dt} = (\beta S - \gamma)I.$$

In words, there is a rate of flow $\beta SI$ of new infections out of $S$ and into $I$, and a rate of flow $\gamma I$ out of $I$.

Let us think of a population of initial size $N$ into which a small number of infectives is introduced, so that $S(0) \approx N$, $I(0) \approx 0$. If $S(0) \approx N < \gamma/\beta$, then $I$ decreases to zero (no epidemic), while if $N > \gamma/\beta$, then $I$ first increases to a maximum attained when $S = \gamma/\beta$ and then decreases to zero (epidemic). The quantity $\beta N/\gamma$ is a threshold quantity, called the *basic reproduction number* and denoted by $\mathcal{R}_0$, which determines whether there is an epidemic or not.

The definition of the basic reproduction number $\mathcal{R}_0$ is that the basic reproduction number is the number of secondary infections caused by a single infective introduced into a wholly susceptible population of size $N$ over the course of the infection of this single infective. In this situation, an infective makes $\beta N$ contacts in unit time, all of which are with susceptibles and thus produce new infections, and the mean infective period is $1/\gamma$; thus the basic reproduction number is $\beta N/\gamma$.

Initially, the number of infectives grows if and only if $\mathcal{R}_0 > 1$ because the equation for $I$ may be approximated by

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

and thus $dI/dt$ is positive when $t = 0$ if $\mathcal{R}_0 > 1$ and negative if $\mathcal{R}_0 < 1$. If we could solve the above pair of differential equations for the functions $S$ and $I$, we would have a prediction of the numbers of susceptibles and infectives as functions of time. Unfortunately, we cannot solve them analytically, but there is some information that we can deduce directly from the differential equations. Since $dS/dt < 0$, the function $S(t)$ is always decreasing. Since $dI/dt < 0$ if $\beta S < \gamma$, the function $I(t)$ decreases after $S$ reaches the value $S = \gamma/\beta$. Thus, eventually $I$ decreases to zero. If $\mathcal{R}_0 < 1$, then $I(t)$ is always decreasing, which means that the initial infection does not develop into an epidemic. On the other hand, if $\mathcal{R}_0 > 1$, then $I(t)$ increases initially until $S(t)$ decreases to the value $S = \gamma/\beta$ and then decreases to zero, an epidemic.

There is a useful mathematical trick that will give us some additional information about the behaviour of solutions. The trick is that instead of thinking of $S$ as a function of the time $t$, we think of $t$ as a function of $S$. Then $I$ being a function of $t$, which in turn is a function of $S$, makes $I$ indirectly a function of $S$. Calculus tells us that the derivative of $I$ with respect to $S$, which we denote by $dI/dS$, is given by the rule

$$\frac{dI}{dS} = \frac{\dfrac{dI}{dt}}{\dfrac{dS}{dt}},$$

which allows us to calculate

$$\frac{dI}{dS} = \frac{(\beta S - \gamma)I}{-\beta SI}$$
$$= -1 + \frac{\gamma}{\beta S}.$$

More calculus rules (techniques of integration) give $I$ as a function of $S$, namely

$$I = -S + \frac{\gamma}{\beta}\ln S + c,$$

where $c$ is a constant that we still need to determine. We use the initial state of the population to do this. The constant $c$ is determined by the initial values $S(0)$, $I(0)$ of $S$ and $I$, respectively. With $S(0) + I(0) = N$ we substitute the value $t = 0$ into the solution to give

$$c = N - \frac{\gamma}{\beta}\ln S(0).$$

Thus

$$I = -S + \frac{\gamma}{\beta}\ln S + N - \frac{\gamma}{\beta}\ln S(0).$$

If we use the fact that $I(t) \to 0$ as $t \to \infty$ and let $S_\infty$ be the limiting value of $S(t)$ as $t \to \infty$, we obtain

$$N - \frac{\gamma}{\beta} \ln S(0) = S_\infty - \frac{\gamma}{\beta} \ln S_\infty,$$

the *final size equation*. In particular, this equation tells us that $S_\infty > 0$, so that some members of the population escape the epidemic. This has frequently been observed in real life. After an epidemic passes there are always some members of the population that do not have disease antibodies, which means that they are not immune and have not been infected.

It is convenient to depict the epidemic model by showing the *phase portrait*, the orbit shown as a curve in the $(S, I)$ plane in Figure 1. Since $S$ is a decreasing function of time, the orbit is traversed from right to left. The vertical line in Figure 1 has the equation $S = \gamma/\beta$, so that $\mathcal{R}_0 > 1$ if $S(0)$ is to the right of this line and $\mathcal{R}_0 < 1$ if $S(0)$ is to the left of this line. This shows graphically that if $\mathcal{R}_0 > 1$, there is an epidemic with the number of infectives first increasing to a maximum $I_{\max}$ and then decreasing to zero, while if $\mathcal{R}_0 < 1$, the infection dies out without spreading. The graph in Figure 1 was actually drawn using a computer algebra system (Maple) to solve the pair of differential equations numerically with the parameter values

$$\beta = 0.0178, \quad \gamma = 2.73, \quad S(0) = 254, \quad I(0) = 7, \quad N = 261,$$

which correspond to the Eyam plague epidemic data. It is worth noting that we may use these data to calculate $I_{\max} = 30.26$ from the relation between $I$ and $S$ with $S = \gamma/\beta$. We may also use the final size equation to estimate numerically that $S_\infty = 76.05$, again using a computer algebra system (because the final size equation gives $S_\infty$ implicitly). We should remember that since the actual observed results have integer values, we should interpret the model as predicting $S_\infty = 76$, $I_{\max} = 30$, compared with the observed values $S_\infty = 83$, $I_{\max} = 29$.

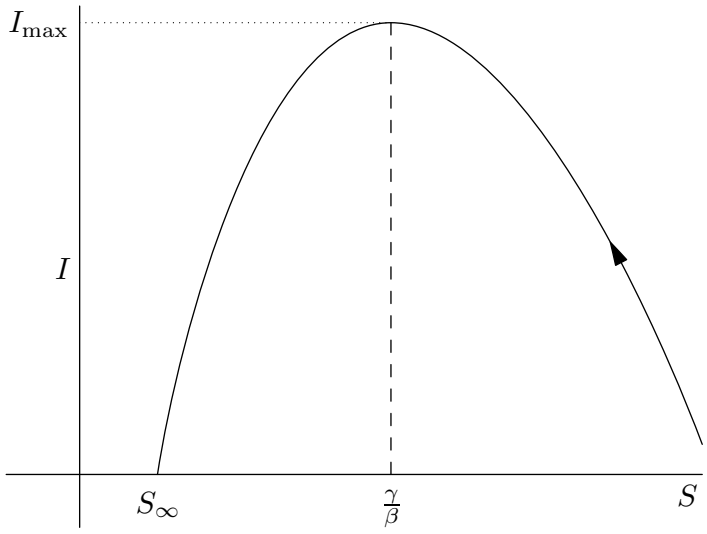



**Figure 1:** The phase portrait.

The actual data for the Eyam epidemic are remarkably close to the predictions of this very simple model. However, the model is really too good to be true. Our model assumes that infection is transmitted directly between people. While this is possible, bubonic plague is transmitted mainly by rat fleas. When an infected rat is bitten by a flea, the flea becomes extremely hungry and bites the host rat repeatedly, spreading the infection in the rat. When the host rat dies its fleas move on to other rats, spreading the disease further. As the number of available rats decreases, the fleas move to human hosts; this is how plague starts in a human population (although the second phase of the epidemic may have been the pneumonic form of bubonic plague, which can be spread from person to person.) One of the main reasons for the spread of plague from Asia into Europe was the passage of many trading ships; in medieval times ships were invariably infested with rats. An accurate model of plague transmission would have to include flea and rat populations, as well as movement in space. Such a model would be extremely complicated and its predictions might well not be any closer to observations than our simple unrealistic model.

In the village of Eyam the rector persuaded the entire community to quarantine itself to prevent the spread of disease to other communities. One effect of this policy was to increase the infection rate in the village by keeping fleas, rats, and people in close contact with one another, and the mortality rate from bubonic plague was much higher in Eyam than in London. Further, the quarantine could do nothing to prevent the travel of rats and thus did little to prevent the spread of disease to other communities. One message this suggests to mathematical modellers is that control strategies based on false models may be harmful, and it is essential to distinguish between assumptions that simplify but do not alter the predicted effects substantially, and wrong assumptions that make an important difference, even if they seem to fit observed data. To model an epidemic we must know the transmission route, even if we can not yet identify a virus or a method of transmission.

A real epidemic differs considerably from an idealized model. One difference is that when it is realized that an epidemic has begun, some individuals will modify their behaviour by avoiding crowds to reduce contacts and by being more careful about hygiene to reduce the risk that a contact will produce infection. Diagnosed infectives may be hospitalized, both for treatment and to isolate them from the rest of the population. Contact tracing of diagnosed infectives may identify people at risk of becoming infective, and these may be quarantined (instructed to remain at home and avoid contacts) and monitored so that they may be isolated immediately if and when they become infective. Isolation may be imperfect and there may be infections transmitted in hospitals.

Soon after the beginning of the SARS epidemic of 2002–3 scientists began looking at the formulation of models to describe this epidemic. Usually, this work was carried out by groups of scientists, including mathematicians, statisticians, epidemiologists, and microbiologists. For example, a network was set up by MITACS (Mathematics of Information Technology and Complex Systems) including scientists from Health Canada and universities across Canada. Some early observations were that the SARS epidemic was an example of a general class of models, not a completely separate situation, and that it was possible to formulate more general and more inclusive models of Kermack–McKendrick type. For this reason, the MITACS group, which was originally set up to study SARS, became a group studying the modelling of infectious diseases in general. Models have been developed that incorporate the additional features described above and have been used to try to answer such questions as whether control of an epidemic should concentrate on isolation of diagnosed infectives or quarantine of contacts of infectives, or a combination of both. For example, it appears that in an epidemic that can be controlled by isolation, the additional gains from a quarantine are slight. Of course, it is important to remember that a model is only a model and will not be an accurate description of all aspects of an epidemic. Nevertheless, although mathematics has not yet cured any diseases (except

possibly math anxiety), it may help in controlling future epidemics. In the event of future outbreaks, you can expect that epidemiologists and mathematical modellers will collaborate to suggest the best control strategy. This kind of mathematical modelling requires some knowledge of calculus, differential equations, and linear algebra. Further discussion of the above epidemic model and more realistic models is contained in [1, Chapter 7] and in [2, Section 6.6].

## Acknowledgements

*I would like to acknowledge David Earn and Pauline van den Driessche who made some very helpful suggestions in the preparation of this article.*

## References

[1] Brauer, F. and C. Castillo-Chavez, 2001: Mathematical Models in Population Biology and Epidemiology. *Texts in Applied Mathematics*, **40**, Springer-Verlag, New York.

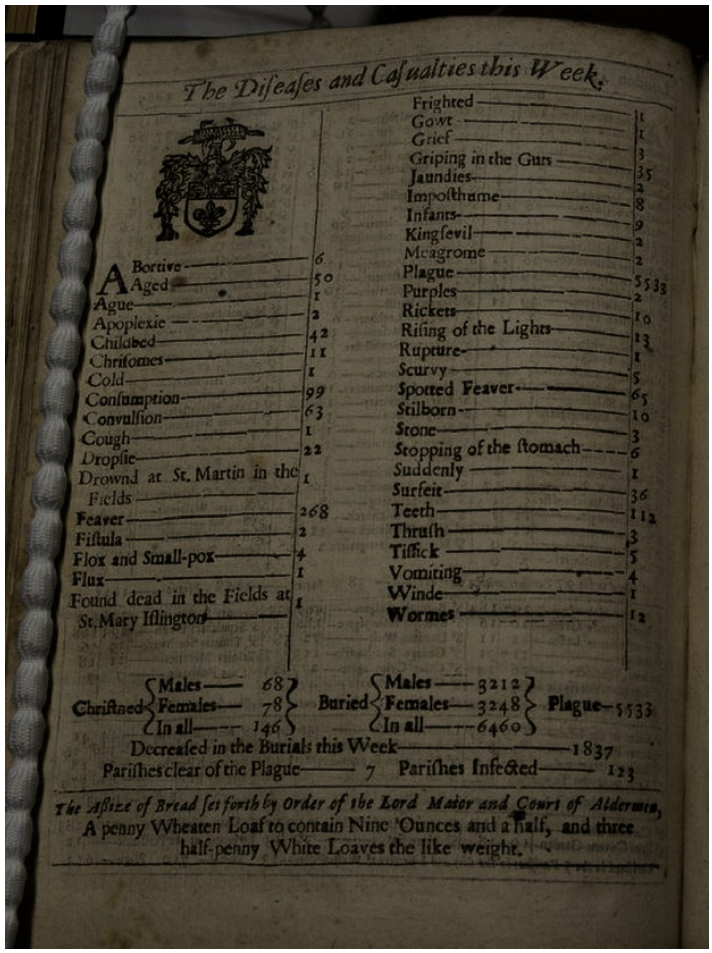
[2] Keshet, L., 2004: Mathematical Models in Population Biology. *SIAM Classics in Applied Mathematics*, **46**, SIAM.




# Mathematical Modelling of Recurrent Epidemics

David J. D. Earn†

One of the most famous examples of an epidemic of an infectious disease in a human population is the Great Plague of London, which took place in 1665–1666. We know quite a lot about the progression of the Great Plague because weekly bills of mortality from that time have been retained. A photograph of such a bill is shown in Figure 1. Note that the report indicates that the number of deaths from plague (5533) was more than 37 times the number of births (146) in the week in question, and that wasn't the worst week! (As Fred Brauer notes in his article in this issue, an even worse plague occurred in the 14th century, but no detailed records of that epidemic are available.)


The Diſeaſes and Caſualties this Week.

| | | | |
|---|---|---|---|
| Abortive | 6 | Frighted | 1 |
| Aged | 50 | Gowt | 1 |
| Ague | 1 | Grief | 3 |
| Apoplexie | 2 | Griping in the Guts | 35 |
| Childbed | 42 | Jaundies | 2 |
| Chrifomes | 11 | Impoſthume | 8 |
| Cold | 1 | Infants | 9 |
| Conſumption | 99 | Kingſevil | 2 |
| Convulſion | 63 | Meagrome | 2 |
| Cough | 1 | Plague | 5533 |
| Dropſie | 22 | Purples | 2 |
| Drowned at St. Martin in the Fields | 1 | Rickets | 10 |
| Feaver | 268 | Riſing of the Lights | 13 |
| Fiſtula | 2 | Rupture | 1 |
| Flox and Small-pox | 4 | Scurvy | 5 |
| Flux | 1 | Spotted Feaver | 65 |
| Found dead in the Fields at St. Mary Iſlington | 1 | Stilborn | 10 |
| | | Stone | 3 |
| | | Stopping of the ſtomach | 6 |
| | | Suddenly | 1 |
| | | Surfeit | 36 |
| | | Teeth | 112 |
| | | Thruſh | 3 |
| | | Tiſſick | 5 |
| | | Vomiting | 4 |
| | | Winde | 1 |
| | | Wormes | 12 |

Chriſtned: Males 68, Females 78, In all 146. Buried: Males 3212, Females 3248, In all 6460. Plague 5533.

Decreaſed in the Burials this Week 1837

Pariſhes clear of the Plague 7 Pariſhes Infected 123

*The Aſſize of Bread ſet forth by Order of the Lord Maior and Court of Aldermen,*
A penny Wheaten Loaf to contain Nine Ounces and a half, and three half-penny White Loaves the like weight.

**Figure 1:** A photograph of a bill of mortality for the city of London, England, for the week of 26 September to 3 October 1665.

† **David J. D. Earn** is a professor in the Department of Mathematics & Statistics at McMaster University. His web site is http://www.math.mcmaster.ca/earn.

Putting together the weekly counts of plague deaths from all the relevant mortality bills, we can obtain the *epidemic curve* for the Great Plague, which I've plotted in the top left panel of Figure 2. The characteristic exponential rise, turnover and decline is precisely the pattern predicted by the classic susceptible-infective-recovered (SIR) model of Kermack and McKendrick [1] that I describe below (and Fred Brauer also discusses in his article). While this encourages us to think that mathematical modelling can help us understand epidemics, some detailed features of the epidemic curve are not predicted by the simple SIR model. For example, the model does not explain the jagged features in the plotted curve (and there would be many more small ups and downs if we had a record of daily rather than weekly deaths). However, with some considerable mathematical effort, these "fine details" can be accounted for by replacing the differential equations of Kermack and McKendrick with equations that include stochastic (i.e., random) processes [2]. We can then congratulate ourselves for our modelling success... until we look at more data.

The bottom left panel of Figure 2 shows weekly mortality from plague in London over a period of 70 years. The Great Plague is the rightmost (and highest) peak in the plot. You can see that on a longer timescale, there was a complex pattern of plague epidemics, including extinctions and re-emergences. This cannot be explained by the basic SIR model (even if we reformulate it using stochastic processes). The trouble is likely that we have left out a key biological fact: there is a reservoir of plague in rodents, so it can persist for years, unnoticed by humans, and then re-emerge suddenly and explosively. By including the rodents and aspects of spatial spread in a mathematical model, it has recently been possible to make sense of the pattern of 17th century plague epidemics in London [3]. Nevertheless, some debate continues as to whether all those plagues were really caused by the same pathogenic organism.

A less contentious example is given by epidemics of measles, which are definitely caused by a well-known virus that infects the respiratory tract in humans and is transmitted by airborne particles. Measles gives rise to characteristic red spots that are easily identifiable by physicians who have seen many cases, and parents are very likely to take their children to a doctor when such spots are noticed. Consequently, the majority of measles cases in developed countries end up in the office of a doctor (who, in many countries, is required to report observed measles cases to a central body). The result is that the quality of reported measles case data is unusually good, and it has therefore stimulated a lot of work in mathematical modelling of epidemics.

An epidemic curve for measles in New York City in 1962 is shown in the top right panel of Figure 2. The period shown is 17 months, exactly the same length of time shown for the Great Plague of London in the top left panel. The 1962 measles epidemic in New York took off more slowly and lasted longer then the Great Plague of 1665. Can mathematical models help us understand what might have caused these differences?

Using the same notation as Fred Brauer uses in his article in this issue, the basic SIR model is

$$\frac{dS}{dt} = -\beta SI, \tag{1}$$

$$\frac{dI}{dt} = \beta SI - \gamma I. \tag{2}$$

Here, $S$ and $I$ denote the numbers of individuals that are susceptible and infectious, respectively. The derivatives $dS/dt$ and $dI/dt$ denote the rates of change of $S$ and $I$ with

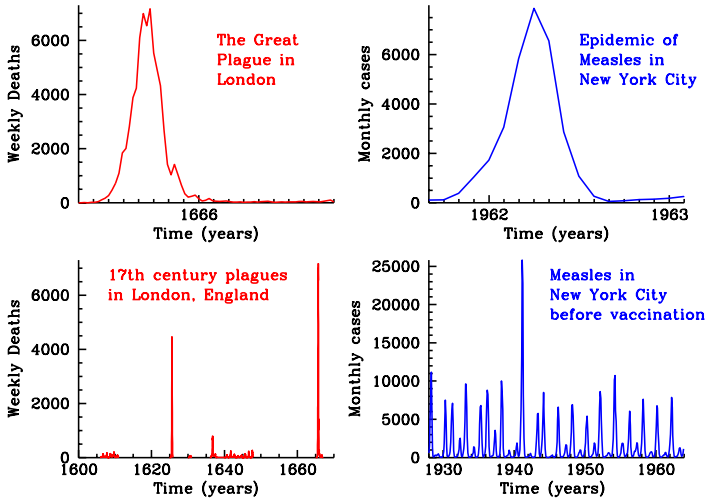



**Figure 2:** Epidemic curves for plague in London (left panels) and measles in New York City (right panels). For plague, the (red) curves show the number of deaths reported each week. For measles, the (blue) curves show the number of cases reported each month. In the top panels, the small ticks on the time axis occur at monthly intervals.

respect to time. The *mean transmission rate* is $\beta$ and the *mean recovery rate* is $\gamma$ (so the *mean infectious period* is $1/\gamma$). As Fred Brauer discusses, if the total population size is $N$, and everyone is initially susceptible ($S(0) = N$), then a newly introduced infected individual can be expected to infect $\mathcal{R}_0 = \beta N/\gamma$ individuals (the *basic reproduction number* $\mathcal{R}_0$ is also discussed at length by Marjorie Wonham in her article in this issue). You can find a discussion of the SIR model together with the mathematical ideas it is based on in some introductory calculus textbooks (see, for example, [4]).

As Fred Brauer notes, we cannot solve the SIR equations and obtain formulae for the functions $S(t)$ and $I(t)$. Yet the epidemic curves that we are trying to explain are essentially given by $I(t)$, so it is hard to proceed without it! Fortunately, computers come to our rescue. Rather than seeking an explicit formula for $I(t)$, we can instead obtain a numerical approximation of the solution. One simple approach is Euler's method, which we can implement as follows (using a spreadsheet or any standard programming language).

The derivative $dS/dt$ is defined as the ratio of the change in $S$ in a given short time interval $dt$, divided by that time interval, in the limit that $dt$ approaches zero. Dealing with that limit is tricky, but at any time $t$ we can approximate the derivative by writing $dS = S(t + dt) - S(t)$ and solving for the number of susceptibles at a time $t + dt$ in the future,

$$S(t + dt) = S(t) - \beta S(t)I(t)dt\,. \qquad (3)$$

Similarly, we can approximate the number of infectives at time $t + dt$ as

$$I(t + dt) = I(t) + \beta S(t)I(t)dt - \gamma I(t)dt\,. \qquad (4)$$

Equations (3) and (4) together provide a scheme for approximating solutions of the basic SIR model. To implement this scheme on a computer, you need to decide on a suitable small time interval $dt$. If you want to try this, I'd suggest taking $dt$ to be one tenth of a day. I should point out that I am being extremely cavalier in suggesting the above method. Do try this at home, but be forewarned that you can easily generate garbage using this simple approach if you're not careful. (To avoid potential confusion, include a line in your program that checks that $S(t) \geq 0$ and $I(t) \geq 0$ at all times. Another important check is to repeat your calculations using a much smaller $dt$ and make sure your results don't change.)

In order for your computer to carry out the calculations specified by equations (3) and (4), you need to tell it the parameter values ($\beta$ and $\gamma$, or $\mathcal{R}_0$, $N$ and $\gamma$) and initial conditions ($S(0)$ and $I(0)$). For measles, estimates that are independent of the case report data that we're trying to explain indicate that the mean infectious period is $1/\gamma \sim 5$ days and the basic reproduction number is $\mathcal{R}_0 \sim 18$ [5]. The population of New York City in 1960 was $N = 7\,781\,984$. If we now assume one infectious individual came to New York before the epidemic of 1962 ($I(0) = 1$), and that everyone in the city was susceptible ($S(0) = N$), then we have enough information to let the computer calculate $I(t)$. Doing so yields the epidemic curve shown in the top panel of Figure 3, which does *not* look much like the real data for the 1962 epidemic in New York. So is there something wrong with our model?

No, but there is something very wrong with our initial conditions. The bottom right panel of Figure 2 shows reported measles cases in New York City for a 36 year period, the end of which includes the 1962 epidemic. Evidently, measles epidemics had been occurring in New York for decades with no sign of extinction of the virus. In late 1961, most of New York's population had already had measles and was already immune, and the epidemic certainly didn't start because one infectious individual came to the city. The assumptions that $I(0) = 1$ and $S(0) = N$ are ridiculous. If, instead, we take $I(0) = 123 \cdot (5/30)$ (the number of reported cases in September 1961 times the infectious period as a proportion of the length of the month) and $S(0) = 0.065N$, then we obtain the epidemic curve plotted in the middle panel of Figure 3, which is much more like the observed epidemic curve of Figure 2 (top right panel). This is progress—we have a model that can explain a single measles epidemic in New York City—but the model cannot explain the recurrent epidemics observed in the bottom right panel of Figure 2. This is not because we still don't have exactly the right parameter values and initial conditions: no parameter values or initial conditions lead to recurrent epidemics in this simple model. So, it would seem, there must be some essential biological mechanism that we have not included in our model. What might that be?

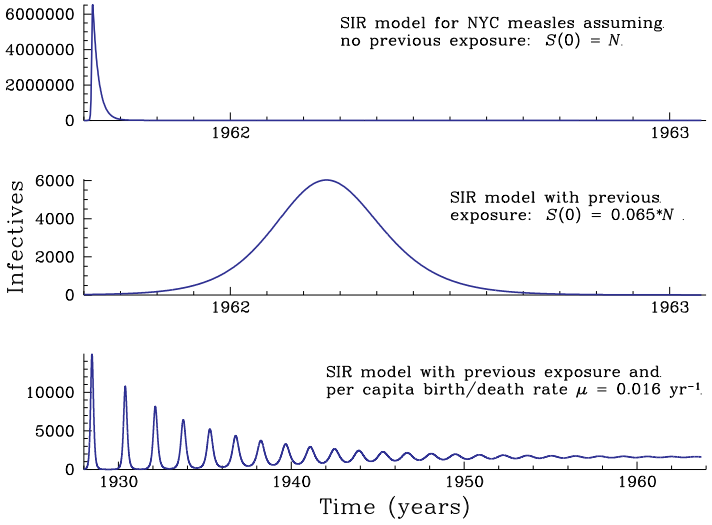



**Figure 3:** Epidemic curves for measles in New York City, generated by the basic SIR model. The curves show the number of infectives $I(t)$ at time $t$. In the top two panels, the small ticks on the time axis occur at monthly intervals.

Let's think about why a second epidemic cannot occur in the model we've discussed so far. The characteristic turnover and decline of an epidemic curve occurs because the pathogen is running out of susceptible individuals to infect. To stimulate a second epidemic, there must be a source of susceptible individuals. For measles, that source cannot be previously infected people, because recovered individuals retain lifelong immunity to the virus. Newborns typically acquire immunity from their mothers, but this wanes after a few months. So births can provide the source we're looking for.

If we expand the SIR model to include $B$ births per unit time and a natural mortality rate $\mu$ (per capita), then our equations become

$$\frac{dS}{dt} = B - \beta SI - \mu S\,, \tag{5}$$

$$\frac{dI}{dt} = \beta SI - \gamma I - \mu I\,. \tag{6}$$

The timescale for substantial changes in birth rates (decades) is generally much longer than a measles epidemic (a few months), so we'll assume that the population size is constant (thus $B = \mu N$, so there is really only one new parameter in the above equations; we can take it to be $B$). As before, we can use Euler's trick to convert the equations above into a scheme that enables a computer to generate approximate solutions. An example is shown in the bottom panel of Figure 3, where I have taken the birth rate to be $B = 126\ 372$ per year (the number of births in New York City in 1928, the first year for which we have data). The rest of the parameters and initial conditions are as in the middle panel of the figure.

Again we seem to be making progress. We are now getting recurrent epidemics, but the oscillations in the numbers of cases over time damp out, eventually reaching an equilibrium. While the graph is just an approximate solution for a single set of initial conditions, it can actually be proved that all initial conditions with $I(0) > 0$ yield solutions that converge onto this equilibrium. So we still don't have a model that can explain the real oscillations in measles incidence from 1928 to 1964, which showed no evidence of damping out. Back to the drawing board?

Don't give up. We've nearly cracked it. So far, we have been assuming implicitly that the transmission rate $\beta$ (or, equivalently, the basic reproduction number $\mathcal{R}_0$) is simply a constant and, in particular, that it does not change in time. Let's think about that assumption. The transmission rate is really the product of the rate of contact among individuals and the probability that a susceptible individual who is contacted by an infectious individual will become infected. But the contact rate is *not* constant throughout the year. To see that, consider the fact that in the absence of vaccination, the average age at which a person is infected with measles is about five years [5]; hence most susceptibles are children. Children are in closer contact when school is in session, so the transmission rate varies seasonally. A crude approximation of this seasonality is to assume that $\beta$ varies sinusoidally,

$$\beta(t) = \beta_0(1 + \alpha\cos 2\pi t)\,. \tag{7}$$

Here, $\beta_0$ is the mean transmission rate, $\alpha$ is the *amplitude* of seasonal variation and the time $t$ is assumed to be measured in years. If, as above, $\beta$ is assumed to be a periodic function (with a period of one year) then the SIR model is said to be *seasonally forced*. We can still use Euler's trick to solve the equations approximately, and I encourage you to do that using a computer for various values of the seasonal amplitude $\alpha$ (you must have $0 \le \alpha \le 1$: why?).

You might think that seasonal forcing is just a minor tweak of the model, but in fact this forcing has an enormous impact on the epidemic dynamics that the model predicts. If you've taken Physics and studied the forced pendulum, then you might already have some intuition for this. A pendulum with some friction will exhibit damped oscillations and settle down to an equilibrium. But if you tap the pendulum with a hammer periodically then it will never settle down and it can exhibit quite an exotic range of behaviours including chaotic dynamics [6] (oscillations that look random). Similarly complex dynamics can occur in the seasonally forced SIR model.

Most importantly, with seasonal forcing, the SIR model displays *un*damped oscillations similar to the patterns seen in the real measles case reports. But we are left with another puzzle. If you look carefully at the New York City measles reports in the bottom right panel of Figure 2 you'll see that before about 1945 the epidemics were fairly irregular, whereas after 1945 they followed an almost perfect two-year cycle. While the SIR model can generate both irregular dynamics and two-year cycles, this happens for different parameter values, not for a single solution of the equations. How can we explain *changes over time* in the *pattern* of measles epidemics?

Once again, the missing ingredient in the model is a changing parameter value. This time it is the birth rate $B$, which is not really constant. Birth rates fluctuate seasonally, but to such a small extent that this effect is negligible. What turns out to be more important is the much slower changes that occur in the average birth rate over decades. For example, in New York City the birth rate was much lower during the 1930s (the "Great Depression") than after 1945 (the "baby boom") and this difference accounts for the very different patterns of measles epidemics in New York City during these two time periods [7].

A little more analysis of the SIR model is very useful. It is possible to prove that changes in the birth rate have exactly the same effect on disease dynamics as changes of the same relative magnitude in the transmission rate or the proportion of the population that is vaccinated [7]. This equivalence makes it possible to explain historical case report data for a variety of infectious diseases in many different cities [8].

One thing that you may have picked up from this article is that successful mathematical modelling of biological systems tends to proceed in steps. We begin with the simplest sensible model and try to discover everything we can about it. If the simplest model cannot explain the phenomenon we're trying to understand, then we add more biological detail to the model, and it's best to do this in steps because we are then more likely to be able to determine which biological features have the greatest impact on the behaviour of the model.

In the particular case of mathematical epidemiology, we are lucky that medical and public health personnel have painstakingly conducted surveillance of infectious diseases for centuries. This has created an enormous wealth of valuable data with which to test hypotheses about disease spread using mathematical models, making this a very exciting subject for research in applied mathematics.

## Acknowledgements

*It is a pleasure to thank Sigal Balshine, Will Guest, and the anonymous referee for helpful comments. The photograph in Figure 1 was taken by Claire Lees at the Guildhall in London, England. The weekly plague data plotted in Figure 2 were digitized by Seth Earn.*

## References


[1] Kermack, W. O. and A. G. McKendrick, 1927: A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London, Series A*, **115**, 700–721.

[2] Bailey, N. T. J., 1975: *The mathematical theory of infectious diseases and its application*. Griffin, London, 2nd edition.

[3] Keeling, M. J. and C. A. Gilligan, 2000: Metapopulation dynamics of bubonic plague. *Nature*, **407**, 903–906.

[4] Hughes-Hallett D. et al, 2002: *Applied Calculus*. John Wiley & Sons, Toronto, 2nd edition.

[5] Anderson, R. A. and R. M. May, 1991: *Infectious Diseases of Humans: Dynamics and Control*. Oxford University Press, Oxford.

[6] Gleick, J., 1987: *Chaos*. Abacus, London.

[7] Earn, D. J. D., P. Rohani, B. M. Bolker and B. T. Grenfell, 2000: A simple model for complex dynamical transitions in epidemics. *Science*, **287**, 667–670.

[8] Bauch, C. T. and D. J. D. Earn, 2003: Transients and attractors in epidemics. *Proceedings of the Royal Society of London, Series B*, **270**, 1573–1578.