Far East Journal of Applied Mathematics

http://www.pphmj.com
http://dx.doi.org/10.17654/AM098010001
Volume 98, Number 1, 2018, Pages 1-30 **ISSN: 0972-0960**

A GENERAL EPIDEMIC MODEL SUITABLE FOR PLANNING

Alan Washburn

Naval Postgraduate School
U. S. A.

Abstract

This paper introduces EPIMOD, a generalization of the susceptible-infected-recovered (SIR) epidemic model. The generalization is essentially to expand the infected compartment of SIR into a transient class of compartments in a continuous-time Markov chain. EPIMOD is general enough to encompass a variety of communicable diseases, while still being simple enough to enable planning for their consequences and containment. Stochastic and deterministic analyses are made and compared. Examples include one based on an Ebola epidemic in the USA.

1. Introduction

Kermack and McKendrick [10] introduced a deterministic epidemic model that is now referred to as the susceptible-infective-recovered (SIR) model. Subsequent work has generalized SIR to include more "compartments" of the disease, including the SEIR model that adds the *E*-compartment of victims who are exposed but not yet infective and the SEQIJR model of Gumel et al. [6] that has two additional compartments for

Received: October 17, 2017; Accepted: November 25, 2017
2010 Mathematics Subject Classification: 92D30, 60J28, 60H10.
Keywords and phrases: epidemic, Ebola, SEIR.

intermediate stages of the SARS disease. In addition to these generalizations, it has been observed that the data for any of these models can be interpreted either deterministically, as a system of ordinary differential equations, or stochastically as a continuous-time, finite-state Markov chain. This paper introduces a general epidemic model EPIMOD that encompasses SIR, SEIR, SEQIJR and other similar models. Section 3 is a deterministic analysis and Section 4 is a stochastic analysis. Section 5 includes a rough Ebola epidemic planning model.

By "epidemic" is meant the introduction of a new or mutated communicable disease into a population. History has many examples of epidemics, including the introduction of smallpox to the New World, the worldwide 1918 influenza epidemic, the 2014 Ebola epidemic in West Africa and the 2014 Ebola epidemic in the USA. The first two were disastrous, whereas the fourth died out soon after it was introduced ("fizzled", as we shall say). Ebola is apparently endemic in Africa, but in the other three examples the causative agent has by now died out, a feature that will always be the case in EPIMOD. As we use the term epidemics always die out in the sense that the initial causative agent ultimately leaves no offspring, even though the number of cases caused in the meantime can be explosively large. In fact, as we use the term epidemics die out quickly enough that birth rates and immigration rates within the affected susceptible population can be (and are) assumed to be negligible. Explosiveness is a possible but not necessary feature; in fact, one of our objects in Section 4 is to determine the probability of "fizzling", a term we use to mean "does not explode".

Difficult questions arise in planning for epidemics. By now we are familiar with the properties of smallpox and know how to make effective vaccines, but, considering the cost and shelf life of vaccines and the unlikelihood (but not impossibility) of a smallpox epidemic, how much vaccine should be stockpiled, and under what circumstances should it be administered to whom? Should vaccines for other potential epidemic agents be developed, and if so should they be produced or is it sufficient to simply

have a plan for producing them? Is it worth the cost to construct facilities for isolation that will be useless unless there is an epidemic? In the particular case of Ebola, one of the current problems is the lack of a cheap and fast diagnostic tool – without one, isolation facilities could easily be overwhelmed. Should such a tool be developed or improved? Although questions like these can be difficult, epidemics must still be planned for. Our goal for EPIMOD is that it should be sufficiently flexible to deal with a variety of epidemics and responses to them, while at the same time being sufficiently simple to enable these crucial analyses.

2. EPIMOD, a Completely Mixed Epidemic Model

In this paper, vectors and matrices will be represented in bold, non-italic type, while scalars will be italic but not bold. The components of a vector or matrix will be given the same alphabetic symbol as the vector or matrix. Thus, the vector $(x_1, ..., x_n)$ will be named **x**, and if **A** is a matrix, then the element in row *i* and column *j* is A_{ij}. The symbol "≡" means "by definition".

EPIMOD is a continuous-time Markov chain whose state is a vector $(s, \mathbf{x}, r)$ of counts of individuals who are either "susceptible"(*s*), "removed"(*r*), or in one of a finite number of infected "compartments" of the disease (row vector **x**). An infected susceptible enters one of *n* compartments, circulates among the compartments for a while, possibly infecting other susceptibles in the process, and ultimately enters the absorbing state *R*, which can stand for either "recovered" or "removed". Once in *R* the victim can never infect anybody again, nor ever again be infected. The susceptible population decreases by unity with each new infection, and never increases.

A probability distribution represented by row vector **p** is given. A newly infected susceptible enters compartment *i* with probability p_i. Also given are *n*-vectors **v** (a row with all positive components) and **c** (a column with no negative component) whose meaning will be explained shortly, and an $(n+1)\times(n+1)$ Markov transition matrix **P** whose states correspond to the

compartments augmented by the absorbing state *R*, which has index $n+1$. Except for $P_{n+1,\,n+1}$, which is 1, all diagonal elements of **P** are 0, and it is assumed that all states of **P** except for *R* are transient-all infected individuals ultimately enter state *R*.

Every occupant of compartment *i* leaves compartment *i* at rate v_i; that is, he remains in compartment *i* for an exponentially distributed time with mean $1/v_i$. The SIR model described by Kermack and McKendrick [10] includes the possibility that the compartmental dwell times are not exponentially distributed, but dwell times are always exponential in EPIMOD. A partial remedy for this restriction is that EPIMOD can have many compartments. Any dwell time that is a sum of independent exponentials can be modeled, so all Erlang-type dwell time distributions are available.

While in compartment *i*, each victim infects each susceptible at rate c_i. When the victim leaves compartment *i*, he instantly transitions to state $j \neq i$ with probability P_{ij}, moving among the compartments until state *R* is finally encountered. It is assumed that the fate of every occupant of a compartment is independent of the fate of every other occupant, and that infections of susceptibles are also independent of each other.

The total rate at which susceptibles become infected is assumed to be

$$s\sum_{i=1}^{n} c_i x_i = s\mathbf{xc}; \tag{1}$$

that is, each susceptible is infected by each occupant of compartment *i* at the rate c_i. This is the "completely mixed" model. The reason for the special notation for susceptibles (*s* is not one of the compartments) is the presence of a product of variables in (1), which distinguishes susceptible transitions from compartmental transitions.

The effective rate of transferring from compartment *i* to state *j*, for each occupant of compartment *i*, is $A_{ij} \equiv v_i P_{ij}$; $i = 1, ..., n$; $j = 1, ..., n+1$; $i \neq j$.

The vector **v** and the matrix **P** can be dispensed with in favor of the $n \times n$ transition rate matrix $\mathbf{A} = (A_{ij})$ whose diagonal elements are $A_{ii} = -v_i$. Thus, for example, instead of dealing with $\mathbf{v} = (1, 2, 4)$ and

$$\mathbf{P} = \begin{bmatrix} 0 & 0.2 & 0.5 & 0.3 \\ 0.1 & 0 & 0.9 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

we can deal with the more compact $\mathbf{A} = \begin{bmatrix} -1 & 0.2 & 0.5 \\ 0.2 & -2 & 1.8 \\ 0 & 0 & -4 \end{bmatrix}$. We will make no further reference to either **v** or **P**, referring only to the analytically more convenient matrix **A**, which is a diagonally dominant matrix and also irreducible because of our assumption that all states of **P** except for *R* are transient. Matrix **A** will therefore always have an inverse (Horne and Johnson [8]). In fact, if $\mathbf{T} \equiv -\mathbf{A}^{-1}$, then T_{ij} is the average amount of time spent in compartment *j* before *R* is encountered, starting in compartment *i*. Diekmann et al. [5] pointed out that for some purposes the dimensionality of **A** can be reduced to the number of positive components of **p**, since these are the possible compartments at *infection*, but we will not exploit that feature here.

The matrix **A** might be upper triangular, which corresponds to the idea that the compartment index of an infected individual can only increase with time. All of our examples will be of this form, but EPIMOD does not require it. The generality could be useful. It sometimes happens, for example, that a victim temporarily goes from being symptomatic to being asymptomatic, as well as vice versa.

Initial susceptible and compartment populations are given, with at least one of the compartment counts being nonzero. In the long run the compartment populations will all become 0, at which point the disease will disappear. The primary goal of our analyses will be to determine the number of susceptibles that are infected in the meantime (the “case count”).

We will also consider a different epidemic model where the number of susceptibles remains constant at s_0, regardless of how many infections have occurred. Call this model LINMOD to distinguish it from EPIMOD. Brauer [3] referred to the linearized model as "standard". The only difference between LINMOD and EPIMOD is that s in (1) is replaced by its initial value s_0. Although we think of LINMOD as an approximation to EPIMOD, LINMOD might actually be a better model of an epidemic than EPIMOD, depending on how mixing actually occurs. Imagine an experiment conducted before the epidemic that is designed to estimate one of the components of **c**. We take a random citizen and track him through a day of his life, noting all of the susceptibles that he infects or would infect if he were an occupant of component i. Then we repeat the experiment often enough to establish that the average number of victims per day is m_i. All susceptibles are equally infectable, so we might argue that the probability that any particular susceptible is infected has to be m_i/s_0. This reasoning leads to EPIMOD with $c_i = m_i/s_0$. But one might also argue that the average number of infections caused by one person-day of component i is m_i, regardless of how many susceptibles remain. This reasoning leads to LINMOD. The essential question is whether the infective activity of a compartment occupant can be expected to decrease in proportion to the remaining population of susceptibles (EPIMOD) or not (LINMOD). Depending on social habits and the infection mechanism, either model (or something in between as in Brauer [3]) might be more realistic. The two models are indistinguishable at the beginning of an epidemic, but differ once the population of susceptibles in EPIMOD has been significantly reduced, with LINMOD being the more pessimistic of the two.

LINMOD is entirely linear and easier to analyze than EPIMOD, but it does have the unfortunate feature that there is no upper limit on the number of cases, which in large epidemics can exceed s_0 or even be infinite. Even so, LINMOD can serve as good approximation to EPIMOD for certain purposes.

3. Deterministic Analysis

Let $\mathbf{x} = (x_1, ..., x_n)$ be the compartment populations and let s be the population of susceptibles, with given nonnegative initial values s_0 and $\mathbf{x}_0$. As functions of time, $\mathbf{x}$ and s are assumed to satisfy a set of ordinary differential equations:

$$\dot{s} = -s\mathbf{x}\mathbf{c}, \text{ and} \tag{2}$$

$$\dot{\mathbf{x}} + \dot{s}\mathbf{p} = \mathbf{x}\mathbf{A}. \tag{3}$$

In (2) and (3), the superimposed dot means "derivative with respect to time". We omit the equation for $\dot{r}$ because r is not involved in either (2) or (3). These equations are a fluid approximation to EPIMOD, so the variables are not integer-valued. Equation (2) states that every occupant of compartment i infects every susceptible at the rate c_i, and equation (3) states that population i increases at a rate that is p_i of the infections, otherwise being determined by the dynamics represented by $\mathbf{A}$.

There is a problem interpreting what $\mathbf{x}$ means, as is often the case in fluid approximations. The components of $\mathbf{x}$ are not integers, which is at odds with the notion that its components are populations of individuals. One might suppose that $\mathbf{x}$ is the mean population vector, but that turns out not to be true in general. Even so, the equations are easily solved numerically, and ought to be a good approximation to epidemics that are in some sense "large".

Equations (2) and (3) do not constitute a linear system because of the product of variables in (2), so we cannot easily construct an analytic solution. However, the solution must still have certain predictable properties. Since s_0 is nonnegative and $\dot{s} = 0$ when $s = 0$, s will always be nonnegative. Our assumptions also imply that $\dot{x}_i \geq 0$ if ($s \geq 0$ and $\mathbf{x} \geq 0$ and $x_i = 0$), so $\mathbf{x}$ will also remain nonnegative for all $t \geq 0$. Therefore, s must decrease to some limiting nonnegative value s_∞, and $\dot{s}$ must approach 0 in the limit as

$t \to \infty$. All compartment populations must also approach 0 in the limit, since every infected susceptible will ultimately end up in state *R*.

It would be useful if we could compute s_∞, the limiting population of susceptibles, without having to solve the differential equations numerically. Toward this end, let

$$\mathbf{M} = \int_0^\infty \mathbf{x} dt. \tag{4}$$

M is necessarily nonnegative, and represents the total number of victim-hours spent in each of the various compartments. Dividing (2) by *s* and integrating, we have

$$\int_0^\infty (\dot{s}/s) dt = \ln(s_\infty/s_0) = -\mathbf{Mc}. \tag{5}$$

Integrating (3), we have, since $\mathbf{x}_\infty = \mathbf{0}$,

$$\int_0^\infty (\dot{\mathbf{x}} + \dot{s}\mathbf{p}) dt = -\mathbf{x}_0 + (s_\infty - s_0)\mathbf{p} = \mathbf{MA}. \tag{6}$$

Since $\mathbf{A}^{-1}$ exists, (6) can be solved for **M**, after which (5) determines s_∞.

Now let $f \equiv (s_0 - s_\infty)/s_0$ be the fraction of original susceptibles who are ultimately infected by the disease, the "case fraction". Given f, s_∞ is simply $s_0(1 - f)$. To find *f*, solve (6) for **M**, recall the definition of **T**, and substitute the result into (5) to obtain

$$\ln(1 - f) + (\mathbf{x}_0 + s_0 f\mathbf{p})\mathbf{Tc} = 0. \tag{7}$$

Equation (7) can be rewritten as

$$a \equiv \mathbf{x}_0\mathbf{Tc},$$

$$b \equiv s_0\mathbf{pTc},$$

$$g(f) \equiv \ln(1 - f) + a + bf = 0. \tag{8}$$

Parameter a is a scalar measure of the initial size of the epidemic, and parameter b is the "basic reproductive number" that is often called R_0. As long as the number of susceptibles remains at s_0, the average number of additional infections caused by a newly infected individual before he enters state R is b. When $b > 1$, at least until the population of susceptibles is substantially reduced, there is clearly the potential for the epidemic to explode.

Consider the concave function $g(f)$ defined in (8). Since $g(0) = a$ and the function approaches $-\infty$ as f approaches 1, (8) will have a unique solution in the interval $[0, 1]$ as long as $a \geq 0$. The equation is transcendental, but can nonetheless be solved numerically by (say) the Newton-Raphson technique. The case fraction is therefore well-defined and computable. When $b < 1$ and $a \ll 1$, the logarithm in (8) can be approximated by $-f$, and the solution is therefore approximately $f \cong a/(1-b)$, provided that the ratio is small. A somewhat more accurate solution can be obtained by approximating the logarithm by $-f - f^2/2$ and solving a quadratic equation: $f \cong \sqrt{(1-b)^2 + 2a} - (1-b)$. When $a = 0.1$ and $b = 0.5$, the two approximations are 0.200 and 0.171, while the exact f is 0.168.

This completes the deterministic analysis of EPIMOD.

The deterministic analysis of LINMOD is comparatively simple. With $\mathbf{m} = s_0\mathbf{c}$, (3) can be replaced by

$$\dot{\mathbf{x}} - \mathbf{xmp} = \mathbf{xA}. \tag{9}$$

Letting $\mathbf{B} \equiv \mathbf{A} + \mathbf{mp}$, (9) can be expressed as

$$\dot{\mathbf{x}} = \mathbf{xB}. \tag{10}$$

Since $\mathbf{mp}$ is a nonnegative matrix, $\mathbf{x}(t)$ must be nonnegative for all t, as in EPIMOD. The solution of (10) is $\mathbf{x}(t) = \mathbf{x}_0 \exp(\mathbf{B}t)$, where exp() is the

matrix exponential function. This analytic solution for the compartment levels is of course an advantage of LINMOD, since the corresponding solution for EPIMOD must be obtained by numerical integration.

Depending on **B**, the compartment populations might approach infinity because there is no limit on the number of infections in LINMOD. We assume in the following that s_0 is small enough that **B** shares with **A** the property that an entirely negative inverse exists, in which case the compartment populations must remain finite. In fact, integrating (10), we find that

$$-\mathbf{x}_0 = \left(\int_0^\infty \mathbf{x}(t)\,dt \right) \mathbf{B} = \mathbf{M}\mathbf{B}, \tag{11}$$

where **M** is defined as before as a vector of victim-hours, and therefore

$$\mathbf{M} = -\mathbf{x}_0 \mathbf{B}^{-1}. \tag{12}$$

Equation (12) determines **M**, and the case fraction f is simply **Mc**. If f exceeds 1, it means that there will be more cases than the initial number of susceptibles.

The case fraction will generally be very small in planning countermeasures to an epidemic, and in that case EPIMOD and LINMOD will produce nearly identical solutions.

Whether the epidemic is modeled with EPIMOD or LINMOD, there still remains the problem of interpreting the meaning of **x**, which is not a vector of integers in the deterministic analysis. The remedy is a stochastic interpretation of the same data, which we consider next.

4. Stochastic Analysis

Consider a continuous-time Markov chain where the state is $(\mathbf{x}, s)$, a vector where as usual **x** is an n-vector of compartment populations and s is the remaining number of susceptibles. All variables are now required to be

nonnegative integers. States where $\mathbf{x} = \mathbf{0}$ are absorbing, while all other states are transient. Our main interest is in the long run occupancy probabilities of the absorbing states, since these probabilities determine the probability distribution of the number of cases.

Every transient state $(\mathbf{x}, s)$ has a set $SUC(\mathbf{x}, s)$ of successor states that can be reached after one transition. There are three transition categories that need to be accounted for: (A) a new infection can occur, (B) an occupant of some compartment can disappear into R, and (C) an occupant of some compartment can move into a different compartment. Let $r((\mathbf{x}, s), \mathbf{y})$ be the rate of transition from $(\mathbf{x}, s)$ to successor state $\mathbf{y}$. Also let $\mathbf{e}_i \equiv (0, ..., 0, 1, 0, ..., 0)$, where the single 1 is in the ith of n places, and let $A_{iR} \equiv -\sum_{j=1}^{n} A_{ij}; i = 1, ..., n$. A_{iR} is the rate of transference from compartment i to the absorbing state R, nonnegative by assumption. Then for the three categories we have

$$r((\mathbf{x}, s), (\mathbf{x} + \mathbf{e}_i, s - 1)) = s\mathbf{x}\mathbf{c}p_i;\ i = 1, ..., n,$$

$$r((\mathbf{x}, s), (\mathbf{x} - \mathbf{e}_i, s)) = x_i A_{iR};\ i = 1, ..., n,$$

$$r((\mathbf{x}, s), (\mathbf{x} - \mathbf{e}_i + \mathbf{e}_j, s)) = x_i A_{ij};\ i, j = 1, ..., n;\ i \neq j, \qquad (13)$$

except that the rate is 0 if any component of the successor state is negative. The total rate out of state $(\mathbf{x}, s)$ is $R(\mathbf{x}, s) \equiv \sum_{\mathbf{y} \in SUC(\mathbf{x}, s)} r((\mathbf{x}, s), \mathbf{y})$. It is easily shown that $R(\mathbf{x}, s) = \sum_{i=1}^{n} (sc_i - A_{ii}) x_i$, which is always positive for transient states.

Every state also has a set $PRE(\mathbf{x}, s)$ of transient predecessor states from which $(\mathbf{x}, s)$ can be reached after one transition. We have $PRE(\mathbf{x}, s) = A \cup B \cup C$, where

$$A = \{(\mathbf{x} - \mathbf{e}_i,\ s+1);\ i = 1, ..., n\},$$

$$B = \{(\mathbf{x} + \mathbf{e}_i,\ s);\ i = 1, ..., n\},$$

$$C = \{(\mathbf{x} + \mathbf{e}_i - \mathbf{e}_j,\ s);\ i,\ j = 1, ..., n;\ i \neq j\}, \tag{14}$$

except that set elements are missing if any component of the predecessor state is negative.

Now let $P(\mathbf{z}, t)$ be the probability that the state is $\mathbf{z}$ at time t, and let the derivative with respect to time be $\dot{P}(\mathbf{z}, t)$. Then Kolmogorov's forward differential equations for the stochastic process are

$$\dot{P}(\mathbf{z}, t) = \sum_{\mathbf{y} \in PRE(\mathbf{z})} P(\mathbf{y}, t) r(\mathbf{y}, \mathbf{z}) - P(\mathbf{z}, t) R(\mathbf{z}). \tag{15}$$

Equation (15) states that the time derivative is the average rate coming into state $\mathbf{z}$ minus the average rate going out. It applies to all states, transient or not.

We know that all the derivatives in (15) are 0 in the limit as $t \to \infty$. That observation is often used to determine stationary probabilities, but it is useless here because only absorbing states have nonzero limiting probabilities, and the corresponding probabilities are all multiplied by 0 in (15). However, (15) can still be used to determine the case distribution by using a method pioneered by Bailey [1].

Except for the initial state, we know that the initial probability of being in any given state is 0. For transient states we must therefore have $\int_0^\infty \dot{P}(\mathbf{z}, t) dt = -\delta(\mathbf{z})$, where $\delta(\mathbf{z})$ is 0 in all states except the initial state $(\mathbf{x}_0, s_0)$, where it is 1. For all transient states $\mathbf{z}$ define $M(\mathbf{z}) \equiv \int_0^\infty P(\mathbf{z}, t) dt$; this is the mean amount of time spent in state $\mathbf{z}$ before some absorbing state is reached. We then have, after integrating (15) over the time interval $[0, \infty)$ and rearranging terms,

$$R(\mathbf{z})M(\mathbf{z}) = \delta(\mathbf{z}) + \sum_{\mathbf{y}\in PRE(\mathbf{z})} M(\mathbf{y})r(\mathbf{y}, \mathbf{z}). \tag{16}$$

The function $M(\mathbf{x}, s)$ is 0 unless $(\mathbf{x}, s)$ is a feasible transient state. Since the number of susceptibles cannot increase, feasibility requires $0 \le s \le s_0$. Since compartment shifts never create new occupants, feasibility also requires $s + \sum_{i=1}^{n} x_i \le s_0 + \sum_{i=1}^{n} x_{i0}$, so the total number of transient states is finite. Thus, (16) is a finite system of linear equations to be solved for $M(\mathbf{x}, s)$.

Now let $Q(s)$ be the probability that the terminal number of susceptibles is *s*. Since state $(\mathbf{0}, s)$ can only be reached from some state $(\mathbf{e}_i, s)$ for some compartment index *i*, and since the rate at which state $(\mathbf{e}_i, s)$ moves to absorbing state $(\mathbf{0}, s)$ is A_{iR}, we have

$$Q(s) = \sum_{i=1}^{n} M(\mathbf{e}_i, s)A_{iR}. \tag{17}$$

The probability that the epidemic results in *m* cases is $Q(s_0 - m)$; $0 \le m \le s_0$. This case distribution is the desired end of analysis.

If **A** is upper triangular, the states can be numbered so that every transition increases the state number, starting with 1 for $(\mathbf{x}_0, s_0)$. If (16) is employed in state number sequence, the right-hand side will always be known when $M(\mathbf{z})$ is to be computed. The initial computation leads to $M(\mathbf{x}_0, s_0) = 1/R(\mathbf{x}_0, s_0)$ because the initial state has no predecessors, and the rest of the computations can be made in a single pass (Black and Ross [2]).

An alternate method of computing the case distribution can be based on the associated embedded discrete Markov chain that models state transitions without regard to the time of occurrence. The transition probability from

state **z** to state **y** in this chain is $Pr(\mathbf{z}\,|\,\mathbf{y}) = r(\mathbf{y}, \mathbf{z})/R(\mathbf{y})$. Define $N(\mathbf{z})$ to be the mean number of entrances to state **z**. Then the desired probability that the terminal number of susceptibles is s is $Q(s) = N(\mathbf{0}, s)$, a number that cannot exceed 1 because the maximum number of entrances to an absorbing state is 1. Furthermore, according to the conditional expectation theorem we have

$$N(\mathbf{z}) = \delta(\mathbf{z}) + \sum_{\mathbf{y} \in PRE(\mathbf{z})} N(\mathbf{y}) Pr(\mathbf{z}\,|\,\mathbf{y}). \tag{18}$$

The term $\delta(z)$ in (18) accounts for the initial entrance to state $(\mathbf{x}_0, s_0)$. This equation applies to all states, transient or not, so (18) is an alternate, more direct method of computing the case distribution. In fact $N(\mathbf{z}) = R(\mathbf{z})M(\mathbf{z})$ for transient states, so (18) is just a compact restatement of (16) and (17). In the upper triangular case $N(\mathbf{x}_0, s_0) = 1$.

We now turn to the stochastic case count analysis for LINMOD.

It is not necessary to include the number of susceptibles in the state vector in LINMOD because that number remains constant at s_0. Define the transition rates $r(\mathbf{x}, \mathbf{y})$ and $R(\mathbf{x})$ as in EPIMOD, except that the state is now **x**, rather than $(\mathbf{x}, s)$, and s is replaced by s_0 in the first transition rate of (13). Also let $\mathbf{m} \equiv s_0\mathbf{c}$, a vector of infection rates. As before, the transition probability from state **x** to state **y** in the embedded Markov chain is $r(\mathbf{x}, \mathbf{y})/R(\mathbf{x})$. For the three types of transition possible, define $q_{iR} \equiv A_{iR}/R(\mathbf{e}_i)$, $q_{ij} \equiv A_{ij}/R(\mathbf{e}_i)$, $q_{im} \equiv m_i/R(\mathbf{e}_i)$, and note that

$$q_{iR} + \sum_{j \neq i} q_{ij} + q_{im} = 1;\ i = 1, \ldots, n.$$

The embedded Markov chain in LINMOD is a *multitype Galton-Watson process* (Harris [7]).

Unlike EPIMOD, LINMOD permits a clear definition of the “fizzle” event. There are only two possibilities: either the state will eventually be **0**

(which is absorbing) or not (the compartment populations might grow indefinitely). In the former case we say that the epidemic fizzles. Even epidemics where the reproductive number *b* exceeds 1 can fizzle.

The distinguishing feature of LINMOD is that every victim creates a mini-epidemic that is independent of all the others, so that the total number of cases is a sum of independent random variables. Sums of independent random variables are well handled by generating functions, so let C_i be the number of cases spawned by a single new victim in compartment *i*, and define the generating functions $\phi_i(z) \equiv E(z^{C_i})$; $i = 1, ..., n$; $0 \le z \le 1$. We take z^{C_i} to be 0 if C_i is not finite, even if $z = 1$, and we take z^0 to be 1, even if $z = 0$. This convention makes z^{C_i} a continuous function of *z* throughout the unit interval, regardless of the number of cases. The probability that $C_i = 0$ is $\phi_i(0)$, and the generating function $\phi_i(z)$ in principle determines the rest of the distribution of C_i through its derivatives at $z = 0$. That distribution will sum to $\phi_i(1)$, the fizzle probability. Since the generating function for a sum of independent random variables is the product of the generating functions, the generating function of the total number of cases, starting from $\mathbf{x}_0$, is $\phi(z) \equiv \prod_{i=1}^{n} \phi_i(z)^{x_{0i}}$.

We will determine the generating functions by using the conditional expectation theorem. There are only three possibilities for a mini-epidemic starting in compartment *i*: the victim may disappear, the victim may transfer from compartment *i* to compartment *j*, or the victim may cause an additional infection. In the latter case there will then be two victims, the original victim in compartment *i* and a new companion in compartment *j* who will start his own mini-epidemic. Thus

$$z^{C_i} = \begin{cases} 1 & \text{if the victim disappears, or} \\ z^{C'_j} & \text{if the victim changes compartments, or} \\ z^{1+C'_i+C'_j} & \text{if the victim finds a companion.} \end{cases} \qquad (19)$$

The primed notation in (19) is to emphasize that all random variables are independent of each other, with the subscript alone determining the distribution. Note that (19) is consistent with respect to the meaning of 0^{C_i} and 1^{C_i} defined above; e.g., 1^{C_i} is 0 if C_i' is finite, but C_j' is not. Taking expectations of both sides,

$$\phi_i(z) = q_{iR} + \sum_{j \neq i} q_{ij} E(z^{C_j'}) + z q_{im} \sum_{j=1}^{n} p_j E(z^{C_i' + C_j'}); \; i = 1, ..., n. \qquad (20)$$

If we define the function $\mathbf{F}(\boldsymbol{\phi})$ by

$$F_i(\phi_1, ..., \phi_n; z) \equiv q_{iR} + \sum_{j \neq i} q_{ij} \phi_i + z q_{im} \phi_i \sum_{j=1}^{n} p_j \phi_j; \; i = 1, ..., n, \qquad (21)$$

then, since the two expected values in (20) are first $\phi_j(z)$ and second $\phi_i(z)\phi_j(z)$, (20) can be compactly stated as $\boldsymbol{\phi}(z) = \mathbf{F}(\boldsymbol{\phi}(z); z)$, a system of second order equations to be solved for $\boldsymbol{\phi}(z)$.

Multiplying through by $R(\mathbf{e}_i)$, (20) is equivalent to

$$(m_i - A_{ii})\phi_i(z) = A_{iR} + \sum_{j \neq i} A_{ij} \phi_j(z)$$

$$+ z m_i \phi_i(z) \sum_{j=1}^{n} p_j \phi_j(z); \; i = 1, ..., n. \qquad (22)$$

Equation (22) may be more appealing than (20) because it avoids appeal to the embedded Markov chain. The most important special case is when $z = 1$, for which case define $r_i \equiv \phi_i(1)$. As mentioned earlier, r_i is the fizzle probability for a mini-epidemic starting in compartment *i*. Equation (22) reduces to

$$(m_i - A_{ii})r_i = A_{iR} + \sum_{j \neq i} A_{ij} r_j + m_i r_i \sum_{j=1}^{n} p_j r_j;\ i = 1, ..., n. \qquad (23)$$

This equation can be rearranged so that

$$m_i r_i \sum_{j=1}^{n} p_j (1 - r_j) + A_{iR}(r_i - 1) + \sum_{j \neq i} A_{ij}(r_i - r_j) = 0;\ i = 1, ..., n, \qquad (24)$$

in which form it is clear that $\mathbf{r} = \mathbf{1}$ will always be a solution. Harris ([7], Corollary 1 to Theorem 7.2 in Chapter II) proved that there can be at most two solutions to (23) whose components are all in the unit interval, with the smaller of the two having the correct meaning. If there is only one such solution, fizzling is inevitable regardless of the starting state.

An epidemic can also fizzle "directly", by which is meant that no further susceptibles are ever infected. If we let u_i be the direct fizzle probability, then $u_i = \phi_i(0)$ and (22) reduces to

$$(m_i - A_{ii})u_i = A_{iR} + \sum_{j \neq i} A_{ij} u_j;\ i = 1, ..., n. \qquad (25)$$

This system of linear equations in **u** can be solved by simple substitution when **A** is upper triangular. Direct fizzling is comparatively unlikely, but the probability is easier to compute than the general fizzle probability and has the additional advantage that direct fizzling is the same in EPIMOD as it is in LINMOD. The event "fizzle" is hard to define in EPIMOD because the compartment populations always vanish in any case, but whatever is meant by "fizzling" in EPIMOD, it should surely include direct fizzling.

As long as the only solution of (23) is $\mathbf{r} = \mathbf{1}$, fizzling is guaranteed and we can investigate the mean number of cases. Let w_i be the mean number of cases spawned by a victim in compartment *i*. Then $w_i = \phi_i'(1)$, where now the prime notation denotes the first derivative. After differentiating all terms in (22) and setting $z = 1$, we find that (22) reduces to

$$(m_i - A_{ii})w_i = \sum_{j \neq i} A_{ij} w_j + m_i(1 + w_i) + m_i \sum_{j=1}^{n} p_j w_j;\ i = 1, ..., n. \qquad (26)$$

This system of linear equations has the solution $\mathbf{w} = -\mathbf{B}^{-1}\mathbf{m}$, where **B** is the matrix defined in Section 3. This is in agreement with the deterministic analysis of LINMOD. Thus, the deterministic solution of LINMOD is also the mean of the stochastic solution of LINMOD, at least as far as case count is concerned.

An approximation to the average case count in the stochastic version of a large epidemic, starting from a single occupant of compartment *i*, would be to multiply the deterministic case count by $1 - r_i$, the probability that the LINMOD epidemic does not fizzle. The justification is that the deterministic analysis is accurate enough if the epidemic does not fizzle, and the case count is negligible if it does fizzle. This approximation will be tested in the special cases described below.

We close this section with two comments:

(1) Known algorithms cannot compute the exact case count distribution for large susceptible populations in reasonable amounts of time. In the next section, we will give an example where $s_0 = 400$. The associated computations using (18) are not taxing, but would be impractical if (say) $s_0 = 400{,}000$.

(2) The observation above is not important to the applications we have in mind, which will always plan for epidemics where $b < 1$ and $a \ll 1$. For such epidemics the fizzle probability is 1 and $a/(1 - b)$ is an adequate summary of the case fraction in either EPIMOD or LINMOD.

5. Special Cases and Examples

In this section, we will use uppercase symbols for the compartment populations, as is customary in SIR and its derivatives. All of the

computations referenced in this section are available from an Excel™ spreadsheet EPIMOD.xlsm that can be found at the downloads link at http://faculty.nps.edu/awashburn/.

5.1. Susceptible-infective-removed (SIR)

The deterministic equations for the SIR special case of EPIMOD are

$$\dot{S} = -\alpha IS,$$

$$\dot{I} = \alpha IS - \gamma I,$$

$$\dot{R} = \gamma I.$$

As usual we ignore the R category, so the only variables in EPIMOD are the number of susceptibles (S) and the number of infecteds (I). Omitting the bold notation we take $A = -\gamma$ and $c = \alpha$, in which case (see (8)) $T = 1/\gamma$, $a = \alpha I_0/\gamma$, and $b = \alpha S_0/\gamma$.

Since A is trivially upper triangular, (18) can be coded for sequential computation of the case count distribution. Figure 1 shows the EPIMOD case count distribution for $(S_0, I_0, \alpha, \gamma) = (400, 1, 0.005, 1)$, an epidemic where $b = 2$. In spite of appearances in that figure, the case count probability is always positive–the least likely number of cases is 117, where the probability is 0.00004. Generating Figure 1 takes less than a second on a modern computer. S_0 could be increased to a few thousand without much difficulty, but it could not be increased to the population of any city or country with a population in the millions. In spite of considerable previous work, there is still no practical method of computing the exact, complete SIR case count distribution for large value of S_0. Whittle's (1953) method requires less arithmetic than (18), but is numerically unstable because of the need for subtraction (the only operations involved in (18) are multiplication and addition). Britton [4] also commented on this instability, and gave an asymptotically accurate normal approximation (his formula 12) to the bell-shaped part of Figure 1.

The generating function for the case count distribution for one initial infective in LINMOD can be found by solving (22):

$$\phi(z) = \frac{1 + b - \sqrt{(1+b)^2 - 4bz}}{2bz}; \; 0 < z \leq 1. \tag{27}$$

The direct fizzle probability u can be found either by solving (25) or by using L'Hopital's rule on (27): it is $u = \phi(0) = 1/(1+b)$. This is the probability that the total number of cases (as usual not counting the initial infective) is zero. A Taylor series expansion of $\phi(z)^{I_0}$ reveals (Diekmann et al. [5]) the distribution of C, the total number of cases if there are I_0 initial infectives:

$$P(C = x) = u^{x+I_0}(1-u)^x \frac{I_0(I_0 + 2x - 1)!}{x!\,(I_0 + x)!}; \; x \geq 0. \tag{28}$$

Equation (28) sums to $\phi(1)$, which is 1 if $u \geq 0.5$, or otherwise $1/b$, the fizzle probability. Equation (28) is not shown in Figure 1 because it is nearly 0 throughout the bell-shaped region and otherwise in close agreement with the EPIMOD distribution.

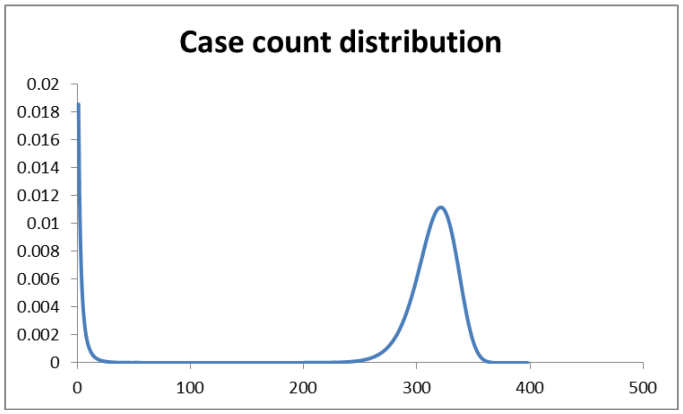

Figure 1. Case count distribution for an epidemic among 400 susceptibles.

The mean of the distribution shown in Figure 1 is 158.02, whereas the deterministic case count is 319.4, right in the middle of the bell-shaped

region. The two differ so strongly because the true mean takes account of the possibility of fizzling. The probabilities of 0, 1 and 2 cases are not shown in Figure 1 because they are large enough to take over the vertical scale, but the sum of just those three is already 0.44. The LINMOD fizzle probability is 0.5, and the resulting product of the nonfizzle probability with the deterministic case count is 159.7, remarkably close to the true mean. Kendall [9] suggested also including the product of the LINMOD fizzle probability with the expected number of cases, given that the epidemic fizzles. This results in the slightly larger estimate of 160.2. It should be obvious that summarizing Figure 1 with any single number is bound to be misleading. It is much more informative to report that the fizzle probability is 0.5, and that about 319 cases can be expected if the epidemic does not fizzle.

The various models come much closer to agreeing with each other when $b < 1$. If α is reduced from 0.005 to 0.001 in the above example (so $b = 0.4$), the case count distribution becomes a decreasing function that closely agrees with (28)–the maximum difference between the EPIMOD case probabilities and (28) is 0.0002. The true stochastic mean shrinks to 0.662, while the deterministic case count in EPIMOD is 0.666. The LINMOD mean is $b/(1-b)$ in general, which is 0.667 in this instance.

While it is difficult to define fizzling in EPIMOD, it is natural to seek some level of cases that will separate the left-hand slide in Figure 1 from the right-hand hump. Let $F(x)$ be the probability that there are x or fewer cases. Table 1 shows $F(x)$ for three values of x and five values of S_0, in all cases adjusting α to keep $\alpha S_0 = 2$ as in the base case. The rows of the table might be regarded as different case cutoff levels, the first row being the direct fizzle probability and the other two as increasingly generous allowances for cases in defining "fizzle". It should be evident that there is little difference between an allowance of 20 and an allowance of 40 when S_0 is large, with both resulting in approximately the LINMOD fizzle probability (0.5). It should also be evident that all large values of S_0 result in nearly the same fizzle probability.

The total computational effort required to produce Table 1 is small, in spite of our earlier observation that producing the complete case distribution is problematic for large values of S_0. The total number of arithmetic operations required in SIR to compute $F(x)$; $x = 1, ..., C$ via (18) is $O(C^2)$, regardless of S_0. It is only when $C = S_0$ that the computational problem emerges. For large epidemics like the one in Figure 1, this means that the fizzling end of the case distribution is easy to compute exactly, even though the nonfizzling end is not.

Table 1. Fizzle probability for three definitions of "fizzle" and five values of S_0

$x\downarrow$, $S_0 \rightarrow$	**40**	**400**	**4,000**	**40,000**	**400,000**
0 (direct)	0.333333	0.333333	0.333333	0.333333	0.333333
20	0.566071	0.502401	0.499344	0.499052	0.499023
40	1.000000	0.503975	0.500332	0.499997	0.499963

5.2. SEIR+ and Ebola

The SEIR model differs from SIR only in having an additional compartment (E) that has contracted the disease, but is not yet infective. The SEIR+ model is a modification of SEIR that might be used as a model for an Ebola epidemic in the USA. We will first briefly review the SEIR model and then explain the modifications.

The SEIR model has four variables: S is the susceptibles, E is the exposed, I is the infected, and R is the removed. The differential equations are

$$\dot{S} = -\alpha IS,$$

$$\dot{E} = \alpha IS - \beta E,$$

$$\dot{I} = \beta E - \gamma I,$$

$$\dot{R} = \gamma I.$$

As usual omit R and let $\mathbf{X} = (E, I)$. This is then a special case of EPIMOD where

$$\mathbf{A} = \begin{bmatrix} -\beta & \beta \\ 0 & -\gamma \end{bmatrix}, \mathbf{T} \equiv -\mathbf{A}^{-1} = \begin{bmatrix} 1/\beta & 1/\gamma \\ 0 & 1/\gamma \end{bmatrix}, \mathbf{c} = \begin{bmatrix} 0 \\ \alpha \end{bmatrix}, \mathbf{p} = [1, 0]. \tag{29}$$

Employing (8), we find that

$$a = (E_0 + I_0)\alpha/\gamma \quad \text{and} \quad b = \alpha S_0/\gamma. \tag{30}$$

As long as b is smaller than 1 and $a/(1-b)$ is much smaller than 1, any epidemic with a small number of initially infected victims (both compartments have the same ultimate impact) will have a small value for the case fraction. Note that parameter β is not involved in (30). It is also not involved in (22), so the LINMOD generating function and case distribution are the same as for SIR.

The SEIR assumption is that the E compartment is not contagious (there is no E term in the first equation). An E has to wait for a time that averages $1/\beta$ before he becomes infective and can expose other susceptibles, and even then he will have only a time that averages $1/\gamma$ before he disappears from the infected population and can cause no further harm.

SEIR+ is a modification of SEIR that is intended to be a model for the possible spread of Ebola in the USA. It is not an arbitrary example, but rather one that is as realistic as this medically amateur writer can make it, inspired by the USA epidemic that actually occurred in September of 2014. Some of the parameters are characteristic of the disease, and are consistent with West African epidemic(s) (Quammen [13]). Other parameters, notably the transmission rate, are little more than guesses. To emphasize that the model is entirely my own, I will switch from the usual “we” to “I” in describing it.

I assume that the Ebola-vaccinated fraction of the population is negligible, so essentially all of the USA population starts out in the susceptible category. I also assume that the main tactical countermeasure is

isolation of the victims from the susceptibles. This can only be done once the victims are recognized as such, which is initially difficult with Ebola because initial symptoms (mild temperature, nausea, …) are similar to the flu. Advanced symptoms (high temperature, pain, vomit, bleeding…) are more suggestive of Ebola, as well as more dangerous because the disease is spread by contact with bodily fluids. I assume that any victim known to have Ebola, regardless of the stage of the disease, will be isolated and therefore play no further role in its spread. This was not true at first in West Africa, where isolation facilities were sometimes overwhelmed, but I assume that it will be true in the USA, at least initially. I therefore define E and I as follows:

- E is the population that has unidentified Ebola and is at most mildly symptomatic.

- I is the population that has unidentified Ebola and is highly symptomatic.

I assume that all of the I-compartment is infective. There are conflicting reports about whether the E-compartment is infective, possibly because of vagueness in the word “mildly” in its definition. I assume that a fraction q of the E-compartment is infective, as well as all of the I-compartment.

The SEIR+ equations are, as usual omitting the one for R:

$$\begin{aligned} \dot{S} &= -\alpha S(I + qE), \\ \dot{E} &= \alpha S(I + qE) - \beta E - \mu E, \\ \dot{I} &= \beta E - \gamma I - \delta I. \end{aligned} \tag{31}$$

Countermeasures to Ebola are represented by the parameters μ and δ, each being the rate at which victims in one of the compartments are recognized and isolated. Recognizing victims in the E compartment would require a test for Ebola that is quick and inexpensive, a test that does not exist as of this writing. My baseline estimate of μ is therefore 0. Recognizing victims in the I compartment is easier, either by medical staff or by the

victims themselves, so my baseline estimate of δ is positive. Only the sum $g \equiv \gamma + \delta$ is significant, but it is convenient to retain both parameters because δ is controllable, while γ is not. SEIR+ is an upper triangular model. The SEIR model is a specialization where $q = \mu = \delta = 0$.

My baseline values for the parameters are given in () in the following list:

- β is the rate at which an *E* becomes an *I*, a disease property (0.1/day).
- μ is the rate at which an *E* is discovered to have Ebola (0/day).
- γ is the rate at which an *I* dies or recovers, a disease property (0.2/day).
- δ is the rate at which an *I* is discovered to have Ebola (0.1/day).
- *q* is the fraction of *Es* that are infective (0.2).
- S_0 is 320,000,000, the current population of the USA.
- E_0 is the initial number of exposed individuals (1).
- I_0 is the initial number of infected individuals (0).
- *m* is the initial mean number of susceptibles infected by each *I*, per day (0.32/day).
- $\alpha = m/S_0 = 10^{-9}/\text{day},$ the transmission rate.

I estimated parameter *m* by asking my friends how many people they thought they would give Ebola to per day, given that they were infective but did not know it, and given that Ebola is spread by bodily contact. Their answers varied a lot, but 0.32/day seemed to be in the collective ball park. Then I simply divided by S_0 to estimate the transmission rate α. I am not proud of this method, and suspect that my estimate of α is the worst of all those in the list. The transmission rate depends as much on social habits (there is less touching and very little funereal laying on of hands in the USA) as it does on the nature of Ebola, so extrapolating from the West African

experience would be risky. In fact, I have difficulty imagining how to get a good estimate of α without first having an extensive Ebola epidemic in the USA. This is especially unfortunate because results are very sensitive to this parameter, as will be seen.

The EPIMOD parameters are

$$\mathbf{A} = \begin{bmatrix} -\beta - \mu & \beta \\ 0 & -g \end{bmatrix}; \quad \mathbf{x}_0 = [1,\ 0]; \quad \mathbf{c} = \begin{bmatrix} q\alpha \\ \alpha \end{bmatrix}; \quad \mathbf{p} = [1\ 0]. \tag{32}$$

Define four dimensionless parameters $Q \equiv qm/(\beta + \mu)$, $G \equiv g/m$, $W \equiv \beta/(\beta + \mu)$ and $\Delta \equiv G(1 - Q) - W$. In the baseline model these are 0.64, 0.9375, 1, and –0.6625, respectively. Using (8) we find that

$$b = (W + GQ)/G; \quad a = (E_0 b + I_0/G)/S_0. \tag{33}$$

Parameter Δ is negative if and only if the basic reproductive number b exceeds 1. That number b is 1.71 in the baseline case, characteristic of the early stages of the West African epidemic (WHO [12]). Given a and b, the deterministic case fraction f can be computed from (8); it is 0.694 in the baseline model.

The deterministic solution of the EPIMOD differential equations can be obtained by numerical integration. Figure 2 graphs E and I as a function of time measured in days after the arrival of the single exposed individual. Both compartments rise abruptly to a peak after about 1 year, and then collapse. At the peak there are almost 10 million victims in compartment I, and almost 30 million in compartment E. At the end of the epidemic most of the population of the USA would have contracted Ebola. The epidemic would be even worse in reality because isolation facilities would be completely overwhelmed, thus falsifying my optimistic assumption that known Ebola cases can all be isolated.

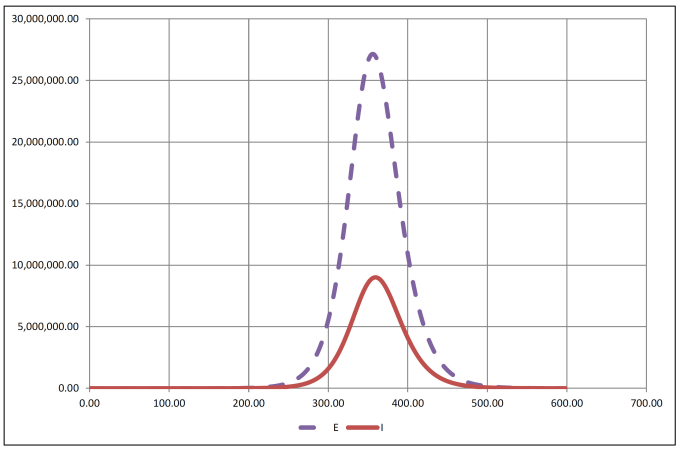

Figure 2. Exposed (dashed) and Infected (solid) in the nominal case. Time is measured in days on the horizontal axis.

The USA recently had an epidemic of exactly this type. I refer to the epidemic that started on September 20, 2014 when T. E. Duncan arrived mildly symptomatic in Texas after having been infected with Ebola in Liberia. Two nurses were subsequently infected while he was hospitalized, but neither those nurses nor T. E. Duncan himself ever gave Ebola to anyone else. In other words, the epidemic fizzled, albeit not directly. This might be taken as evidence that my model is wrong, and that Ebola epidemics on the scale of the one described above are not a threat in the USA. However, it would be premature to conclude that without knowing the fizzle probability.

Equations (24) for the LINMOD fizzle probabilities are

$$[1]: mqr_1(1-r_1)+\mu(r_1-1)+\beta(r_1-r_2)=0,$$

$$[2]: mr_2(1-r_1)+\gamma(r_2-1)=0. \tag{34}$$

Recall the dimensionless quantities Q, G, W, and Δ that are defined above. Solving [2] for r_2 we have $r_2 = G/(1+G-r_1)$ and thus $r_1 - r_2 = (G-r_1)(r_1-1)/(1+G-r_1)$. When this is substituted into [1] the factor (r_1-1) can be cancelled, after which [1] can be rearranged to the quadratic equation

$$Qr_1^2 - r_1(1 + Q + QG) + 1 + QG + \Delta = 0. \tag{35}$$

The only solution of (35) that is smaller than 1 when $\Delta \le 0$ is

$$r_1 = \frac{2(1 + QG + \Delta)}{1 + Q + QG + \sqrt{(1 - Q + QG)^2 - 4Q\Delta}}. \tag{36}$$

Once r_1 is known, r_2 can be determined by substitution. The fizzle probabilities in the baseline case are $\mathbf{r} = (0.49, 0.65)$, which makes my fizzle probability for the 2014 epidemic (T. E. Duncan being the one initial victim in the E compartment) 0.49. This is not a small probability, so it is plausible that the USA just got lucky when the 2014 epidemic fizzled, and that, barring interference that would change model parameters, an Ebola epidemic on the scale of Figure 2 could very well happen next time. The direct fizzle probabilities from (25) are of course smaller: $\mathbf{u} = (0.23, 0.38)$. The 2014 USA epidemic did not fizzle directly.

The good news is not that my model is necessarily wrong, but rather that the epidemic shown in Figure 2 allows time for interference. After 60 days there are only 52 cases, a number that would not strain facilities for isolation and medical care. Suppose that on day 60 some way were found to increase the rate at which E victims are removed from the population and isolated, perhaps by massive application of a cheap, fast Ebola test. Change μ from 0 in the baseline case to $0.1/\text{day}$, reflecting the idea that about 10% of the E compartment will be identified and isolated on each day. Also make $(E_0, I_0) = (19, 6)$, the numbers of exposed and infective victims on baseline day 60, and reset the clock to be 0 on day 60. The effect of this is to reduce the number of additional cases from over 223,000,000 in the baseline case to a mere 154 in the modified model, thus demonstrating the importance of isolating victims in the E-compartment. Reductions in the transmission rate have an even more dramatic effect. There are echoes here of the Ebola epidemic in West Africa, which in its early stages was forecast to have millions of cases in total based on initial observations of transmission rates.

The number of cases tuned out to be thousands, rather than millions, primarily because of actions taken after the epidemic was well under way (Onishi [11]).

6. Summary

EPIMOD is a generalization of SIR with an arbitrary number of compartments and arbitrary transition rates. We have developed both deterministic and stochastic analyses of EPIMOD, in the process introducing the LINMOD approximation where the number of susceptibles remains fixed throughout the epidemic. Emphasis is placed throughout on estimating the average case count, since it is anticipated that minimizing that count will be the goal of any countermeasures taken. EPIMOD is used make a rough model of the Ebola epidemic started by T. E. Duncan's arrival in the USA in 2014, making the point that the epidemic could have been very dangerous even though it fizzled.

References

[1] N. Bailey, The total size of a general stochastic epidemic, Biometrika 40 (1953), 177.

[2] A. Black and J. Ross, Computation of epidemic final size distributions, Journal of Theoretical Biology 367 (2015), 159-165.

[3] C. Brauer, The Kermack-McKendrick epidemic model revisited, Mathematical Biosciences 198 (2005), 119-131.

[4] T. Britton, Stochastic epidemic models: a survey, Mathematical Biosciences 225 (2010), 24-35.

[5] O. Diekmann, J. Heesterbeek and T. Britton, Mathematical Tools for Understanding Infectious Disease Dynamics, Princeton University Press, 2013.

[6] A. Gumel, S. Ruan, T. Day, J. Watmough, P. van den Driessche, F. Brauer, D. Gabrielson, C. Bowman, M. E. Alexander, S. Ardal, J. Wu and B. Sahai, Modeling strategies for controlling SARS outbreaks based on Toronto, Hong Kong, Singapore and Beijing experience, Proc. R. Soc. Lond. B Biol. Sci. 271 (2004), 2223-2232.

[7] T. Harris, The Theory of Branching Processes, Springer Verlag, Berlin, 1963, pp. 36-43.

[8] R. Horne and C. Johnson, Matrix Analysis, Cambridge University Press, 1985, 363.

[9] D. Kendall, Deterministic and stochastic epidemics in closed populations, Proc. 3rd Berkeley Symp. Math. Statist. Prob. 4 (1956), 149-165.

[10] W. Kermack and A. G. McKendrick, A contribution to the mathematical theory of epidemics, Proc. R. Soc. Lond. B Biol. Sci. 115 (1927), 700-721.

[11] N. Onishi, Schools Open-region Closer to Normal, February 1 New York, Times, 2015, 1.

[12] WHO, Ebola Virus Disease in West Africa - The First 9 Months of the Epidemic and Forward Projections, New England J. of Med., 2014, pp. 1-15.

[13] D. Quammen, Ebola: The Natural and Human History of a Deadly Virus. WW Norton & Company, 2014.

Alan Washburn: awashburn@nps.edu