Analyzing Bioterror Response Logistics: The Case of Anthrax

David L. Craft, Lawrence M. Wein, Alexander H. Wilkins

November 12, 2003

Abstract

To aid in understanding how best to respond to a bioterror anthrax attack, we analyze a system of differential equations that includes an atmospheric release model, a spatial array of biosensors, a dose-response model, a disease progression model, and a set of spatially-distributed tandem queues for distributing antibiotics and providing hospital care. We derive approximate closed-form expressions for the number of deaths as a function of key parameters and management levers, including the size of the attack, the time at which the attack is detected via symptomatic patients, the number of days to distribute antibiotics, the efficacy (both for treatment and prevention) of antibiotics, the prophylactic antibiotic distribution strategy, the prioritization of the antibiotics queue, and the detection limit, deployment density and delay time of biosensors.

Operations Research Center, MIT, Cambridge, MA, 02139; dcraft@mit.edu
Graduate School of Business, Stanford University, Stanford, CA, 94306; lwein@stanford.edu
Scientific Computing and Computational Mathematics, Stanford University, Stanford, CA, 94306; awilkins@stanford.edu
1 Introduction

The two most feared biological agents in a terrorist attack are smallpox and anthrax (Henderson 1999). Of the two, anthrax appears to be more readily available, as evidenced by its fatal delivery in the US postal system in the fall of 2001 (Thompson 2003). Absent intervention, anthrax is nearly always fatal (Inglesby et al. 2002). It also is highly durable, having survived in a viable form in soil for 35 years (Manchee et al. 1981). When properly weaponized into spores several microns in size, an aerosol release of anthrax can be dispersed over wide areas, which has led to dire warnings that an airborne attack could kill millions of people (World Health Organization 1970; Office of Technology Assessment, US Congress 1993; Alibek 2000). Anthrax’s combination of deadliness and hardiness makes its threat to the nation’s human and physical assets surpassed only by a nuclear weapon.

Because deterrence is not a reliable strategy against terrorists, and because it is difficult to get biological agents out of the hands of terrorists before they attack, our security against a biological attack rests largely on consequence management, i.e., what can be done after the attack to minimize the number of deaths. This paper is part of a series that analyzes the emergency response to biological attacks with either smallpox or anthrax. Mathematical models were used in Kaplan et al. (2002) to compare the consequences of the government-proposed ring-vaccination strategy and a mass-vaccination strategy in the aftermath of a smallpox attack. A mathematical analysis of these policies in a slightly simplified setting was carried out in Kaplan et al. (2003), which led to some non-obvious scaling relations (Kaplan and Wein 2003). Wein et al. (2003) developed a mathematical model to compare various emergency responses to an airborne anthrax attack. The US Government has no detailed response plan in place for anthrax, and this analysis led to a broad outline of what we believe to be the main elements of an effective response plan (Wein and Kaplan 2003). The present paper provides a mathematical analysis of a slightly simplified version of the
model in Wein et al.

The crux of our model is a spatially-distributed set of two-stage queueing networks for distributing antibiotics and providing care to symptomatic patients. The model also incorporates an atmospheric release model to determine the dose inhaled at each location, an array of biosensors that attempts to detect the attack, a dose-response model to compute the fraction of people infected at each location, and a model for disease progression, which can be truncated by medical intervention if service at each of the two queues is received quickly enough.

The model is presented in §2 and a biosensor analysis, which determines when intervention is initiated, is presented in §3. A single service zone is analyzed in §4 and several generalizations are considered in §5. The deaths are spatially aggregated across service zones in §6 and concluding remarks are offered in §7.

2 The Model

The model is depicted in Figure 1, and consists of an atmospheric release model, a biosensor model, a dose-response model, a disease progression model, and a queueing network model. A comparison of this model and the one in Wein et al. is given at the end of this section.

We use the Gaussian plume model, which is the workhorse of atmospheric models (Hanna et al. 1982), to track the dispersal of \( Q \) anthrax spores that are released at a height of \( z \) meters when the wind speed is \( u \). If we let \( b \) denote the breathing rate, then the number of spores inhaled by a person \( x \) meters directly downstream of the release point, and \( y \) meters crosswind is

\[
s(x, y) = \frac{bQ}{\pi u \sigma_y \sigma_z} e^{-\frac{x^2}{2\sigma_y^2} - \frac{h^2}{2\sigma_z^2}},
\]

where \( \sigma_y = a_1 x^d, \sigma_z = a_2 x^d \), with the Brookhaven(C) parameters (i.e., \( a_1 = 0.32, a_2 = 0.22, d = 0.78 \)). Table 1 contains a list of parameter values used in our computational studies;
unless otherwise noted, these values are taken from Wein et al.

We superimpose on this spatial model a set of biosensors, which are laid out in a square array. The biosensors are characterized by three parameters: the time delay to obtain results ($\tau_b$), the detection limit ($l_b$), which is the minimum number of spores that generates an alarm, and the distance (in meters) between adjacent biosensors ($w_b$), so that the geographical density of biosensors is $w_b^{-2}$. We ignore the possibility of false-positive test results.

Let $\tau$ be the time that the attack is detected; we assume there is no time lag between detection and the initiation of intervention. The attack can either be detected by biosensors or by early symptomatic patients. The detection delay via early symptomatics is given by a constant $\tau_s$, which is assumed to be independent of the attack size. Hence, if we assume that $\tau_b < \tau_s$ then

$$
\tau = \begin{cases} 
\tau_b & \text{if the maximum spore count at a biosensor } > l_b; \\
\tau_s & \text{otherwise.}
\end{cases}
$$

(2)

We use an age-dependent linear dose-response model,

$$
P(s, a) = \min\left(1, \frac{s}{c_1 - c_2 a}\right),
$$

(3)

where $c_1 = 38,000$, $c_2 = 450$, and maximum age $A = 80$. Here, $P(s, a)$ is the probability that someone of age $a$ would get infected from inhaling $s$ spores. This functional form coincides with the data in Table 3 of Webb and Blaser (2002). The age distribution is $U[0, A]$ with pdf $f(a) = A^{-1}$.

Absent intervention, all infected people progress through three exponential stages with means $r_j^{-1}$, for $j = 1, 2, 3$. These three stages respresent incubation, prodromal and fulminant. Symptoms are observed at the end of the incubation period. There are two phases of symptomatic disease: patients are feverish in the prodromal stage, and the toxin released by the anthrax takes over the body in the fulminant (or exploding) stage (Jernigan et al. 2001).
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Q$</td>
<td>Amount released</td>
<td>$10^{15}$ spores</td>
</tr>
<tr>
<td>$u$</td>
<td>Wind speed</td>
<td>5 m/s</td>
</tr>
<tr>
<td>$h$</td>
<td>Release height</td>
<td>100 m</td>
</tr>
<tr>
<td>$b$</td>
<td>Breathing rate</td>
<td>$5 \times 10^{-4}$ m$^3$/s</td>
</tr>
<tr>
<td>$N$</td>
<td>Zone population</td>
<td>90,000</td>
</tr>
<tr>
<td>$I$</td>
<td>Number of infected people in zone (in §4 and §5)</td>
<td>45,000</td>
</tr>
<tr>
<td>$r_1^{-1}$</td>
<td>Mean incubation</td>
<td>12.5 days</td>
</tr>
<tr>
<td>$e_1$</td>
<td>Prophylactic efficacy during incubation</td>
<td>0.9</td>
</tr>
<tr>
<td>$r_2^{-1}$</td>
<td>Mean of prodromal phase</td>
<td>1 day</td>
</tr>
<tr>
<td>$r_3^{-1}$</td>
<td>Mean of fulminant phase</td>
<td>1.5 days</td>
</tr>
<tr>
<td>$e_2$</td>
<td>Probability of infinite post-antibiotic prodromal</td>
<td>0.4</td>
</tr>
<tr>
<td>$\tau$</td>
<td>Detection delay from symptomatics</td>
<td>2 days</td>
</tr>
<tr>
<td>$n_A$</td>
<td>Number of antibiotic servers per zone</td>
<td>109.4</td>
</tr>
<tr>
<td>$\mu_A$</td>
<td>Service rate for antibiotics</td>
<td>$(7 \text{ min})^{-1}$</td>
</tr>
<tr>
<td>$n_H$</td>
<td>Number of hospital servers per zone</td>
<td>12.69</td>
</tr>
<tr>
<td>$\mu_H$</td>
<td>Service rate in hospital</td>
<td>$(6 \text{ hours})^{-1}$</td>
</tr>
<tr>
<td>$p$</td>
<td>Ring parameter</td>
<td>0</td>
</tr>
<tr>
<td>$t_b$</td>
<td>Biosensor time delay</td>
<td>6 hours</td>
</tr>
<tr>
<td>$l_b$</td>
<td>Biosensor detection limit</td>
<td>$10^4$ spores</td>
</tr>
<tr>
<td>$d_b$</td>
<td>Biosensor density</td>
<td>$(9 \text{ km}^2)^{-1}$</td>
</tr>
</tbody>
</table>

Table 1: Parameter values for the model in the base case.
Figure 1: Graphical depiction of the model. Asymptomatics enter the antibiotics queue according to a geographic ring strategy (- - -). Figure taken from Wein et al.
We assume that two-dimensional \((x, y)\) space is broken into \(9\ \text{km}^2\) service zones, each populated by \(N = 90,000\) people. For simplicity, we let the dose in each service zone be fixed at the value of the Gaussian plume model evaluated in the center of each square zone. Let \(I_j(x, y, a, t)\) be the number of people of age \(a\) at location \((x, y)\) and in disease stage \(j\) at time \(t\), where stage 0 corresponds to uninfected people, and stage 4 is death. The above assumptions imply that, just after the attack,

\[
I_1(x, y, a, 0) = N_f(a) P(s(x, y), a) = \min\left(\frac{N}{A}, \frac{k_1 x^{-2d} \exp(-k_2 y^2 x^{-2d} - k_3 x^{-2d})}{c_1 - c_2 a}\right),
\]

where

\[
k_1 = \frac{N b Q}{A \pi u a_1 a_2}, \quad k_2 = \frac{1}{2a_1^2}, \quad k_3 = \frac{h^2}{2a_2^2}.
\]

We now aggregate over age, and after equation (7) suppress the spatial dependence until §6, where we aggregate the deaths over service zones. For a generic service zone, let us denote the number of infected people by

\[
I = \int_0^A I_1(x, y, a, 0) \, da.
\]

For \(t \in [0, \tau]\) and \(j = 1, \ldots, 4\), the system state is given by

\[
I_j(t) = C_j(t)I,
\]

where

\[
C_1(t) = e^{-r_1 t},
\]

\[
C_2(t) = \frac{r_1 (e^{-r_1 t} - e^{-r_2 t})}{r_2 - r_1},
\]

\[
C_3(t) = \frac{e^{-(r_1 + 2r_2 + r_3)t} F_1 F_2 (r_1 - r_2) e^{(r_1 + 2r_2)t} + (r_2 - r_3) e^{(2r_2 + r_3)t} + (r_3 - r_1) e^{(r_1 + r_2 + r_3)t}}{(r_1 - r_2)(r_1 - r_3)(r_2 - r_3)},
\]

\[
C_4(t) = 1 - C_1(t) - C_2(t) - C_3(t).
\]
For notational simplicity, for \( j = 1, \ldots, 4 \), we define \( C_j \) as \( C_j(t) \) evaluated at \( t = \tau \).

Intervention begins at time \( \tau \) and consists of a two-stage multi-server queueing system in each service zone that distributes antibiotics and provides medical care. Infected people continue their disease progression while in the queueing network. Antibiotics have prophylactic and treatment effects: they prevent a fraction \( e_1 = 0.9 \) of people in the incubation period from progressing to symptoms. The 10\% ineffectiveness is due to noncompliance to the drug regimen (as opposed to ineffectiveness of the antibiotics), and these people enter compartment \( U_1 \) in (17) and eventually move to the hospital queue when they progress to the prodromal stage. As soon as these people enter the hospital queue, they are given (and adhere to) antibiotics. Antibiotics also prevent a fraction \( e_2 = 0.4 \) of people in the prodromal phase from ever progressing to the fulminant stage; these people enter compartment \( Q_2^H \) in (21) upon receiving antibiotics. The remaining 60\% of prodromals receiving antibiotics have their prodromal clock restarted (they join compartment \( Q_2^H \) in (20)) and need to complete their hospital care while still in the prodromal stage in order to survive. Patients entering the prodromal stage eventually die, regardless of antibiotics or hospital care.

When intervention is initiated, all symptomatic people immediately enter the antibiotics queue. Hence, the state of the system at time \( \tau \) is \( Q_2^A(\tau) = I_2(\tau), Q_3^A(\tau) = I_3(\tau) \), and \( Q_0^A(\tau) = Q_1^A(\tau) = Q_2^H(\tau) = Q_3^H(\tau) = 0 \). In addition, \( I_2(t) = I_3(t) = 0 \) for \( t > \tau \) since symptomatic people enter the hospital during this time period. The dynamics for \( t \geq \tau \) are given by

\[
\frac{dI_0(t)}{dt} = -A_0(t),
\]

\[
\frac{dQ_0^A(t)}{dt} = A_0(t) - S_0^A(t)Q_0^A(t),
\]

\[
\frac{dI_1(t)}{dt} = -A_1(t) - r_1I_1(t),
\]

\[
\frac{dQ_1^A(t)}{dt} = A_1(t) - [r_1 + S_1^A(t)]Q_1^A(t),
\]
\[
\frac{dU_1(t)}{dt} = (1 - e_1)S_1^A(t)Q_1^A(t) - r_1U_1(t),
\]

(17)

\[
\frac{dQ_2^A(t)}{dt} = r_1I_1(t) + r_1Q_1^A(t) - [r_2 + S_2^A(t)]Q_2^A(t),
\]

(18)

\[
\frac{dQ_3^A(t)}{dt} = r_2Q_2^A(t) - [r_3 + S_3^A(t)]Q_3^A(t),
\]

(19)

\[
\frac{dQ_2^H(t)}{dt} = (1 - e_2)[r_1U_1(t) + S_2^A(t)Q_2^A(t)] - [r_2 + S_2^H(t)]Q_2^H(t),
\]

(20)

\[
\frac{dQ_2^H(t)}{dt} = e_2[r_1U_1(t) + S_2^A(t)Q_2^A(t)] - S_2^H(t)Q_2^H(t),
\]

(21)

\[
\frac{dQ_3^H(t)}{dt} = r_2Q_2^H(t) + S_3^A(t)Q_3^A(t) - [r_3 + S_3^H(t)]Q_3^H(t).
\]

(22)

Those who die progress from states \(Q_2^A\) and \(Q_2^H\), and so the total dead is

\[
D = \int_0^\infty (r_2Q_2^A(t) + r_2Q_2^H(t)) \, dt.
\]

(23)

We complete this section by defining the rates \(A_j(t)\) and \(S_j^i(t)\) appearing in the above equations, which represent the prophylactic antibiotic strategy and the service discipline, respectively. The prophylactic strategy is a ring-based strategy, which tracks the fraction of symptomatic anthrax cases at each location by time \(t\), assuming that this location has not entered the ring by time \(t\). This quantity, which we call the observed anthrax burden, is \(\frac{I_1(0)(1-e^{-r_1t})}{N}\). The ring at time \(t\) consists of all locations that have burdens at least as large as the threshold \(p\); although \(p = 0\) in our base case, we allow \(p\) to vary between 0 and 1 in §5.3. Hence, a given location enters the ring at time

\[
t_p = -\frac{1}{r_1} \ln \left( 1 - \frac{pN}{I_1(0)} \right) \quad \text{if} \quad I_1(0) > pN,
\]

(24)

and \(t_p = \infty\) otherwise. Because intervention does not begin until time \(\tau\), if we let \(I_{\{x\}}\) denote the indicator function of the event \(x\), then

\[
A_0(\tau) = I_0(0) I_{\{\tau \geq t_p\}},
\]

(25)

\[
A_1(\tau) = I_1(0)e^{-r_1\tau}I_{\{\tau \geq t_p\}},
\]

(26)
and, for \( t > \tau \),

\[
A_0(t) = I_0(0) \left. I \right|_{t=t_p}, \\
A_1(t) = I_1(0) e^{-r_1 t_p} \left. I \right|_{t=t_p}.
\]

(27)

(28)

Turning to the service terms \( S^i_j(t) \), define \( Q^H_2(t) = \sum_{j=\{2,\hat{2}\}} Q^H_j(t) \) to be the total number of prodromals at the hospital queue at time \( t \), and \( Q^i(t) = \sum_{j=0}^3 Q^i_j(t) \) to be the total number of people in queue \( i \) at time \( t \). The mass service policy is defined by

\[
S^i_j(t) = \mu_i \min \left( 1, \frac{n_i}{Q^i(t)} \right)
\]

(29)

for \( i = A, j = 0, \ldots, 3 \) and for \( i = H, j = \hat{2}, \hat{2}, 3 \). Hence, the queue departure rates from each disease stage is proportional to the relative numbers in queue. Note that we are not modeling the queue according to a typical first-come first-served discipline, which would not likely be adhered to in the chaotic, post-attack environment.

We also consider the symptomatic priority policy, where asymptomatic people in the antibiotics queue are only served if the number of servers exceeds the number of symptomatics in queue. This policy is defined by (29) for \( i = H, j = \hat{2}, \hat{2}, 3 \), and by

\[
S^A_j(t) = \mu_A \min \left( 1, \frac{n_A - Q^A_j(t) - Q^A_3(t)}{Q^A_0(t) + Q^A_1(t)} \right) \quad \text{for } j = 0, 1,
\]

(30)

\[
S^A_j(t) = \mu_A \min \left( 1, \frac{n_A}{Q^A_2(t) + Q^A_3(t)} \right) \quad \text{for } j = 2, 3.
\]

(31)

The model presented in this section simplifies the model in Wein et al. in nine ways: (i) the Brookhaven version of the functions \( \sigma_y \) and \( \sigma_z \) in the Gaussian plume model are used in place of the Briggs version (Hanna et al.); (ii) the total area of study is unbounded and has a constant population density, whereas the region in Wein et al. is bounded and divided into a rural and urban area; (iii) the dose is fixed for each 9 km\(^2\) service zone, but the dose varies over each square km region in Wein et al.; (iv) a uniform age density is used; (v) the dose-response model is an age-dependent linear model rather than an age-dependent probit.
(vi) disease stages are exponential rather than log-normal random variables; (vii) the progression rate from disease stage 2 is taken to be the same as disease stage 2; (viii) Wein et al. also consider an antibiotic priority policy based on age (people over 55 get priority) and a hospital priority policy that depends on disease stage (prodromals get priority over fulminants); and (ix) Wein et al. also consider a set of mobile hospital servers (representing federal and military resources) that can dynamically move among zones to treat overflow patients. Among the first six simplifications, only simplification (vi) significantly alters the qualitative nature of our model outputs; consequently, this simplification is relaxed in §5.1. The last two simplifications force us to ignore other aspects of the problem that are difficult to address analytically.

3 Biosensor Analysis

Our goal in this section is to derive analytical expressions for the efficacy of biosensors in terms of their detection limit $l_b$ (in terms of inhaled spores) and their spatial density (assuming they are placed on a lattice with a width of $w$ meters).

3.1 The Dense Limit

First, let’s consider the case where $w \to 0$, i.e. there are biosensors everywhere. We can find the maximum spore count by setting $y = 0$ and differentiating (1) with respect to $x$. This yields

$$x^* = \left( \frac{2a_2^2}{h^2} \right)^{-\frac{1}{2a}} \quad (32)$$

and

$$s(x^*, 0) = \frac{2bQa_2}{\pi u h^2 e a_1}. \quad (33)$$
For the parameter values in Table 1, \( x^* = 1637 \) meters, or 1.6 km. By (33), if the detection limit of a biosensor is \( l_b \), then we could detect an attack of size

\[
Q^* = \frac{\pi uh^2 e a_1 l_b}{2ba_2} = 6.2 \times 10^8 l_b.
\]  

Equation (34) provides a lower bound on the size of an attack that could be detected by any spatial deployment.

### 3.2 General Deployment

We assume that the terrorists release the anthrax in a random location that is independent of the locations of the biosensors. A natural performance measure is the likelihood of detecting an attack of a given size. Because

\[
Pr(\text{release of size } Q \text{ detected}| l_b = s) = Pr(\text{max dose at sensors } \geq s|\text{release } = Q),
\]

we could try to derive the probability distribution of the maximum dose received at any sensor. However, to more easily incorporate the biosensor analysis with the rest of the analysis, as we do in §6, we focus instead on estimating the expected maximum dose received at any sensor, which in turn provides an estimate of \( Q^* \), which is the expected minimum attack size that can be detected.

By the unimodal nature of the Gaussian plume model, we only need to consider the four biosensors surrounding the location of the maximum dose. Let us fix these four sensors at \((0,0), (w,0), (0,w), \) and \((w,w)\), and assume that the maximum dose occurs at location \((m_x, m_y)\) that is uniformly distributed in this square; this state of affairs is depicted in Figure 3.2. We need to figure out which of these four biosensors receives the maximum dose, as a function of \((m_x, m_y)\). By the crosswind symmetry of the Gaussian plume model, it is clear that the maximum biosensor is one of the two “southern” sensors (where the wind is blowing east, in the positive \( x \) direction) if \( m_y < w/2 \), and is one of the two northern sensors
if $m_y > w/2$. Hence, we can restrict ourselves to a $w \times w/2$ rectangle, i.e. assume $(m_x, m_y)$ is uniformly distributed in $[0, w] \times [0, w/2]$ and consider only the southern sensors, so that $m_y = 0$ means that the southern biosensors are located precisely downwind of the point of release. Moreover, if $m_x > x^*$ then the upwind (or “western”) sensor (which is located $x^* - m_x$ meters downwind and $m_y$ meters crosswind of the point of release) is upwind of the point of release, and so the downwind sensor (which is located $x^* - m_x + w$ meters downwind and $m_y$ meters crosswind of the point of release) receives the maximum dose. This occurs with probability $(w - x^*)/w$, which goes to 1 as $w \to \infty$. If $m_x < x^*$ then the maximum dose of a biosensor is

$$c_1 \max\{(x^* - m_x)^{-2d}e^{-c_2(m_y)(x^*-m_x)^{-2d}}, (x^* - m_x + w)^{-2d}e^{-c_2(m_y)(x^*-m_x+w)^{-2d}}\},$$  \qquad (36)$$

where

$$c_1 = \frac{bQ}{\pi u a_1 a_2}, \quad c_2(m_y) = \frac{m_y^2}{2a_1^2} + \frac{h^2}{2a_2^2}. \qquad (37)$$

Hence, if we define

$$g(m_x, m_y) = \max\{(x^* - m_x)^{-2d}e^{-c_2(m_y)(x^*-m_x)^{-2d}}, (x^* - m_x + w)^{-2d}e^{-c_2(m_y)(x^*-m_x+w)^{-2d}}\} \quad (38)$$

and

$$f(m_x, m_y) = \begin{cases} (x^* - m_x + w)^{-2d}e^{-c_2(m_y)(x^*-m_x+w)^{-2d}} & \text{if } m_x > x^*; \\ g(m_x, m_y) & \text{if } m_x < x^*. \end{cases} \quad (39)$$

then the expected maximum dose received by any sensor given a release of size $Q$ is

$$\frac{2c_1}{w^2} \int_0^w \int_0^{w/2} f(m_x, m_y) \, dm_y \, dm_x. \quad (40)$$

The value of $w$ could vary from tens of meters in the case of an asset critical to national security, to hundreds of kilometers in a rural region. In this section we attempt three different approximate analyses that are relevant for small $w$, intermediate $w$ and large $w$, respectively.

Beginning with the small-$w$ regime, we try to find the curve that partitions $[0, w] \times [0, w/2]$ into two regions, depending upon whether the maximum dose occurs at the upwind
sensor or the downwind sensor. This curve is found by equating, for each given value of 
$m_y \in [0, w/2]$, the two terms in (36) and solving for $m_x$. The equation that equates the two 
terms in (36) can be expressed as

$$(x^* - m_x + w)^{-2d} = (x^* - m_x)^{-2d} + \frac{2d}{c_2(m_y)} \ln \left( \frac{x^* - m_x}{x^* - m_x + w} \right). \quad (41)$$

We use three approximations to estimate the expected minimum detectable attack size, 
$Q^*$. First, we substitute the Taylor series approximation $(x^* - m_X + w)^{-2d} \sim (x^* - m_x)^{-2d} - 2dw(x^* - m_x)^{-2d-1}$ into the left side of (41) to get

$$(x^* - m_x)^{-2d-1} = -\frac{1}{wc_2(m_y)} \ln \left( \frac{x^* - m_x}{x^* - m_x + w} \right). \quad (42)$$

We then use a crude iterative method: let $m_x = 0$ be an initial estimate, and substitute this 
value on the right side of (42) to get our estimate, $m_x^S$, for the $x$ coordinate of the location 
separating the two regions:

$$(x^* - m_x^S)^{-2d-1} = -\frac{1}{wc_2(m_y)} \ln \left( \frac{x^*}{x^* + w} \right), \quad (43)$$

or

$$m_x^S = x^* - \left( -\frac{1}{wc_2(m_y)} \ln \left( \frac{x^*}{x^* + w} \right) \right)^{-\frac{1}{2d+1}}. \quad (44)$$

Therefore, the maximum dose of a biosensor is approximately

$$Q^* = \begin{cases} 
    c_1(x^* - m_x^S)^{-2d} e^{-c_2(m_y)(x^* - m_x^S)^{-2d}} & \text{if } m_x < m_x^S; \\
    c_1(x^* - m_x^S + w)^{-2d} e^{-c_2(m_y)(x^* - m_x^S + w)^{-2d}} & \text{if } m_x > m_x^S. 
\end{cases} \quad (45)$$

Setting the quantities in (45) equal to $l_b$ and solving for $Q$, we estimate the minimum 
detectable size of a release to be

$$Q^* = \begin{cases} 
    \pi a_1 a_2 b (x^* - m_x)^{-2d} e^{-c_2(m_y)(x^* - m_x)^{-2d}} & \text{if } m_x < m_x^S; \\
    \pi a_1 a_2 b (x^* - m_x + w)^{-2d} e^{-c_2(m_y)(x^* - m_x + w)^{-2d}} & \text{if } m_x > m_x^S. 
\end{cases} \quad (46)$$

As a last approximation, to get a deterministic estimate for $Q^*$ we set $m_x$ and $m_y$ equal to 
the midpoint of their range (i.e., $m_x = w/2$, $m_y = w/4$) and get

$$Q^* = \begin{cases} 
    \frac{\pi a_1 a_2 b}{\pi a_1 a_2 b} \left( x^* - \frac{w}{2} \right)^{-2d} e^{-c_2(w/4)(x^* - w/2)^{-2d}} & \text{if } m_x^S > w/2; \\
    \frac{\pi a_1 a_2 b}{\pi a_1 a_2 b} \left( x^* + \frac{w}{2} \right)^{-2d} e^{-c_2(w/4)(x^* + w/2)^{-2d}} & \text{if } m_x^S < w/2. 
\end{cases} \quad (47)$$

14
which coincides with the dense-limit result (34) as \( w \to 0 \). The exponential term in (47) causes this approximation for \( Q^* \) to increase rapidly for sufficiently large values of \( w \). To find the range of applicability of this approximation, we determine the value of \( w \) that generates a 10\% increase over the dense limit, assuming that the other terms in (47) remain at their dense limit values (i.e., that \( x^* + w/2 \sim x^* - w/2 \sim x^* \)). This occurs when

\[
e^{c_2(w/4)(x^*)^{-2d}} = 1.1e^{c_2(0)(x^*)^{-2d}}.
\]

Using (37) and solving for \( w \), we find that (47) breaks down at

\[
w = 4a_1(x^*)^d\sqrt{2 \ln(1.1)} = 180 \text{ meters}.
\]

Our second line of attack, appropriate for intermediate values of \( w \), also uses three approximations. First, we assume that the downwind biosensor is always the biosensor that receives the maximum dose. While this assumption causes us to overestimate the size of the smallest detectable attack, it is likely to be accurate if \( w \) is significantly larger than \( x^* \). In this case, the expected maximum dose received at a sensor, given a release of size \( Q \), is

\[
\frac{2c_1}{w^2} \int_0^w \int_0^{w/2} (x^* - m_x + w)^{-2d} e^{-c_2(m_y)(x^* - m_x + w)^{-2d}} dm_y dm_x.
\]

Because

\[
\int_0^x e^{-cy^2} dy = \frac{\sqrt{\pi}}{2\sqrt{c}} \Phi(\sqrt{c}x) \to \frac{\sqrt{\pi}}{2\sqrt{c}} \quad \text{as} \quad x \to \infty,
\]

we approximate (50) by

\[
\frac{\sqrt{2\pi c_1 a_1}}{w^2} \int_0^w (x^* - m_x + w)^{-d} e^{-\frac{y^2}{2\sigma^2}(x^* - m_x + w)^{-2d}} dm_x.
\]

We can then substitute the Taylor series approximations

\[
(x^* - m_x + w)^{-2d} \sim (x^* + w)^{-2d} + 2dm_x(x^* + w)^{-2d - 1}
\]

and

\[
(x^* - m_x + w)^{-d} \sim (x^* + w)^{-d} + dm_x(x^* + w)^{-d - 1}
\]
into (52) and integrate to get

\[
\frac{\sqrt{2\pi c_1 a_1^2} (x^* + w)}{\omega^2 \pi c_2 h^2} (x^* + w)^{d+1} e^{-\frac{h^2}{2\pi c_2} (x^* + w)^2} \left[ \left( 1 - e^{-\frac{d}{a_2^2} (x^* + w)^{2d-1}} \right) + \frac{a_2^2}{h^2} (x^* + w)^{2d} \left( 1 - e^{-\frac{d}{a_2^2} (x^* + w)^{2d-1}} \right) \right].
\]

(55)

Solving for \( Q \) gives

\[
Q^* = \frac{\sqrt{2\pi} u w^2 h^2}{2 b a_2} (x^* + w)^{d-1} e^{-\frac{h^2}{2\pi c_2} (x^* + w)^2} \left[ \left( 1 - e^{-\frac{d}{a_2^2} (x^* + w)^{2d-1}} \right) + \frac{a_2^2}{h^2} (x^* + w)^{2d} \left( 1 - e^{-\frac{d}{a_2^2} (x^* + w)^{2d-1}} \right) \right]^{-1}. 
\]

(56)

If we define

\[
c_3 = \frac{d h^2}{a_2^2} (x^* + w)^{2d-1} w
\]

and make the Taylor series approximation \( e^{-c_3} \approx 1 - c_3 + \frac{c_3^2}{2} \), then (56) simplifies to

\[
Q^* = 3\sqrt{2\pi} uu w^2 a_2 (x^* + w)^d e^{-\frac{h^2}{2\pi c_2} (x^* + w)^2} \frac{6 - 3 c_3 + \frac{d w}{(x^* + w)} (3 - 2 c_3)}{b [6 - 3 c_3 + \frac{d w}{(x^* + w)} (3 - 2 c_3)]}. 
\]

(58)

Our last approach starts from (52) and makes two further large-\( w \) approximations. The Taylor series approximation \( e^{-\frac{h^2}{2\pi c_2} (x^* - m_0 + w)^2} \approx 1 \) again allows the integral in (52) to be taken explicitly, giving

\[
\sqrt{2\pi c_1 a_1} \frac{1}{\omega^2} \left( \frac{(x^*)^{1-d} - (x^* + w)^{1-d}}{d-1} \right). 
\]

(59)

The assumption \( x^* \ll w \) allows us to make the further approximations \( (x^* + w)^{1-d} \approx w^{1-d} \) and \( (x^*)^{1-d} \approx -w^{1-d} \), yielding

\[
\sqrt{2\pi c_1 a_1} \frac{1}{w^{1+d}(1-d)}. 
\]

(60)

Setting this result equal to \( l_b \) and solving for \( Q \) gives

\[
Q^* = \frac{\sqrt{2\pi} u a_2 (1-d) l_b}{2 b w^{1+d}} 
\]

(61)

as an estimate for the limiting behavior of \( Q^* \) as \( w \to \infty \).
3.3 Accuracy of Estimates

Figure 3.3 compares our three estimates of $Q^*$ (normalized by the detection limit $l_b$) to the exact value from (40). As predicted, our small-$w$ approximation (47) is valid for $w < 200$ m, but rapidly explodes after that point. Our intermediate-$w$ approximation (56) becomes reasonably accurate around $w = 400$ m and remains so until $w = 20$ km. It remains useful as an upper bound for some time after, but displays extreme swings in value after $w = 10^5$ km. Finally, our large-$w$ approximation (61) starts to be fairly accurate starting at 400 km, and captures the asymptotic behavior of the function. We should note, however, that the Gaussian plume model is typically not applicable in this last regime (Hanna et al.).

4 Analysis of Base Case

In this section, we estimate $D/I$, which is the fraction of infected people who die in a service zone. Our base case uses the ring parameter $p = 0$, so that $A_0(t) = A_1(t)$ for $t > \tau$. The antibiotics queue is analyzed in §4.1, which allows us to characterize the arrivals to the hospital queue. Subsections 4.2 and 4.3 consider the scenarios in which the hospital is overcongested and uncongested, respectively, and §4.4 computationally assesses the accuracy of these approximations.

4.1 The Antibiotics Queue

Summing equations (14), (16), (18) and (19) gives a differential equation for the total number of people in the antibiotics queue,

$$\frac{dQ^A(t)}{dt} = -\mu_A n_A - r_3 Q^A_3(t).$$

(62)
Figure 2: Random location $(m_x, m_y)$ of maximum dose, relative to its four surrounding biosensors, which are located at $(x, y) = (0,0)$, $(w,0)$, $(0,w)$ and $(w,w)$. 

Figure 3: Exact versus approximate values of the expected minimum detectable release size divided by the detection limit, $Q^*/l_b$, versus the biosensor grid width $w$. 

18
Neglecting the second term and assuming no one has exited the system by time $\tau$, we solve (62) and get

$$Q^A(t) = N - n_A \mu_A (t - \tau) \text{ for } t \in [\tau, \tau + \frac{N}{n_A \mu_A}],$$

(63)

$$= N \left( \frac{t_e - t}{t_e - \tau} \right) \text{ for } t \in [\tau, t_e],$$

(64)

where $N$ denotes the zone population and

$$t_e = \tau + \frac{N}{n_A \mu_A}$$

(65)

is the emptying time of the antibiotics queue.

This simple expression for $Q^A(t)$ allows us to solve for $Q^A_1(t)$ and $Q^A_2(t)$, which are needed later in this section. Equation (16) is of the form $\dot{Q}^A_1(t) = -h_1(t)Q^A_1(t)$ where $h_1(t) = \frac{n_A \mu_A}{Q^A(t)} + r_1$. The solution is

$$Q^A_1(t) = Q^A_1(\tau)e^{-\int_{\tau}^{t} h_1(t) \, dt} = Q^A_1(\tau)e^{-r_1(t-\tau)} \left(1 - \frac{n_A \mu_A (t - \tau)}{P}\right) = I e^{-r_1 t} \left( \frac{t_e - t}{t_e - \tau} \right).$$

(66)

The solution for $Q^A_2$ is

$$Q^A_2(t) = \frac{I(t_e - t)}{(t_e - \tau)(r_2 - r_1)} \left(r_1 C_1 (e^{-r_1 (t-\tau)} - e^{-r_2 (t-\tau)}) + (r_2 - r_1)C_2 e^{-r_2 (t-\tau)}\right),$$

(67)

where $C_1$ and $C_2$ are given by (9)-(10), evaluated at $t = \tau$. It turns out that $\left(r_1 C_1 (e^{-r_1 (t-\tau)} - e^{-r_2 (t-\tau)}) + (r_2 - r_1)C_2 e^{-r_2 (t-\tau)}\right) \approx (r_2 - r_1)C_2$ for $t \in [\tau, t_e]$, at least for the base-case parameter values. We therefore employ the following simplification:

$$Q^A_2(t) = IC_2 \left( \frac{t_e - t}{t_e - \tau} \right).$$

(68)

### 4.2 Overcongested Hospital Queue

If we assume that the hospital is so congested that anyone who enters $Q^H_2$ dies, then the number of deaths can be calculated without solving the hospital queue equations. Unfortunately, this is a good approximation in many practical instances (Wein et al.), and allows
us to understand all parameters in the model except the hospital service parameters. Using
this idea, we can categorize the people who die as follows: those who leave $Q_2^A$ via progres-
sion, $(1 - e_2)$ of those who leave $Q_2^A$ via service, and $(1 - e_1)(1 - e_2)$ of those who leave
$Q_1^A$ via service. Add to this the people who are in stages 3 and 4 by time $\tau$ and we have,
by (18), (29), (64), and (65),

$$D = I(C_3 + C_4) + \int_\tau^{\tau_e} \left[ r_2 Q_2^A(t) + (1 - e_2) \frac{1}{t_e - t} Q_2^A(t) + (1 - e_1)(1 - e_2) \frac{1}{t_e - t} Q_1^A(t) \right] \, dt. \quad (69)$$

Substituting equations (18), (66) and (68) into (69) and integrating yields

$$\frac{D}{I} = 1 - C_1 + \frac{r_2}{2} C_2(t_e - \tau) - e_2 C_2 + (1 - e_1)(1 - e_2) \frac{C_1}{r_1(t_e - \tau)}(1 - e^{-r_1(t_e - \tau)}). \quad (70)$$

### 4.3 Uncongested hospital queue

Next is an expression for the number of deaths assuming that the hospital queue is uncon-
gested, i.e., the total number of people in the hospital queue, $Q_H(t)$, satisfies $Q_H(t) < n_H$ for
all $t$. We follow the same logic as the derivation of equation (69), except that now the $(1 - e_2)$
people do not die with certainty, but with probability $r_2/(r_2 + \mu_H)$ due to the competing
exponentials. This gives

$$\frac{D}{I} = 1 - C_1 + \left[ \frac{r_2}{2} (t_e - \tau) - 1 + \frac{r_2(1 - e_2)}{r_2 + \mu_H} \right] C_2 + \frac{r_2(1 - e_1)(1 - e_2)}{r_2 + \mu_H} \frac{C_1}{r_1(t_e - \tau)}(1 - e^{-r_1(t_e - \tau)}). \quad (71)$$

### 4.4 Computational Study

We complete this section with a computational study to determine the accuracy of our
approximations in equations (70)-(71), and to gain an understanding of how these quantities
are influenced by various parameters. Figure 4 compares these approximations to the actual
fraction of deaths, $D/N$, as a function of $I/N$, which is the fraction of people infected in the
service zone. The overcongested approximation (70) is accurate over the entire range, and the difference between the overcongested and uncongested estimates is modest.

![Figure 4: Fraction dead vs. fraction infected.](image)

We now investigate three special cases where the response is rapid: pre-attack antibiotic distribution, instantaneous detection, and a combination of the two. For pre-attack antibiotic distribution, we have $t_e = \tau$. We use l’Hôpital’s rule in (70) and (71) to obtain

$$
\frac{D}{I} = 1 - (e_1 + e_2 - e_1 e_2)C_1 - e_2 C_2 \quad \text{if overcongested,} \tag{72}
$$

$$
\frac{D}{I} = 1 - \left(1 - \frac{r_2(1 - e_1)(1 - e_2)}{r_2 + \mu_H}\right)C_1 - \left(1 - \frac{r_2(1 - e_2)}{r_2 + \mu_H}\right)C_2 \quad \text{if uncongested.} \tag{73}
$$

Setting $\tau = 0$ in (70) and (71) for instantaneous detection, together with the second-order Taylor-series approximation $e^{-x} \approx 1 - x + \frac{x^2}{2}$, gives

$$
\frac{D}{I} = \frac{(1 - e_1)(1 - e_2)(1 - e^{-r_1(t_e - \tau)})}{r_1(t_e - \tau)} \quad \text{if overcongested,} \tag{74}
$$

$$
\approx (1 - e_1)(1 - e_2)\left(1 - \frac{r_1(t_e - \tau)}{2}\right) \quad \text{if overcongested,} \tag{75}
$$
\[
\frac{D}{I} = \frac{r_2}{r_2 + \mu_H} \frac{(1 - e_1)(1 - e_2)(1 - e^{r_1(t_e - \tau)})}{r_1(t_e - \tau)} 
\]

(76)

\[
\approx \frac{r_2}{r_2 + \mu_H}(1 - e_1)(1 - e_2) \left(1 - \frac{r_1(t_e - \tau)}{2}\right) \text{ if uncongested.} 
\]

(77)

The simultaneous use of instantaneous detection and pre-attack antibiotic distribution gives, via l'Hôpital’s rule,

\[
\frac{D}{I} = (1 - e_1)(1 - e_2) \text{ if overcongested,} 
\]

(78)

\[
\frac{D}{I} = \frac{r_2}{r_2 + \mu_H}(1 - e_1)(1 - e_2) \text{ if uncongested.} 
\]

(79)

Equations (72)-(79) offer several useful observations. Equation (78), which optimistically employs instantaneous detection, pre-attack distribution and ample hospital service, provides a lower bound (over all conceivable logistical responses) on the fraction of infected who die, which is 0.012 with our parameter values; that is, almost 99% of the infected people survive. Also, the modest difference between the overcongested and uncongested estimates in Figure 4 depends on having an intervention delay of \(\tau = 2\) days. If \(\tau = 0\), as in (75)-(79), then the uncongested case reduces the fraction dead by a factor of five.

We now turn our attention to how the key parameters influence the fraction dead in the overcongested case. For the fixed value of \(I/N = 0.5\) (Table 1), Figure 5 shows the actual and estimated (according to the overcongested approximation (70)) fraction of infecteds who die as a function of four key parameters: the intervention delay (\(\tau\)), the time it takes to distribute antibiotics \((t_e - \tau)\), the prophylactic efficacy of antibiotics \((e_1)\), and the treatment efficacy of antibiotics \((e_2)\).

Substituting the base-case parameter values aside from \(\tau\) (but keeping \(t_e - \tau\) fixed) into (70) gives

\[
\frac{D}{I} = 1 - 0.81e^{-0.08\tau} - 0.139e^{-\tau}. 
\]

(80)
Figure 5: Fraction of infected people who die vs. four key intervention parameters.
Differentiating (80) gives

\[
\frac{d (D)}{d\tau} = 0.065e^{-0.08\tau} + 0.139e^{-\tau} \approx 0.065e^{-0.08\tau}. \tag{81}
\]

Equation (82) implies that at \(\tau = 2\) days, an additional \(0.065e^{-0.16} \times 100\% = 5.5\%\) of the people die if intervention is delayed an additional day, and the rate at which deaths increase for every day’s delay decreases exponentially, but at the rather slow rate of 8% per day; for example, the increase in deaths from day 6 to day 7 is 4.0%, which is 16% less (8% times two days) than the 4.7% increase from day 4 to day 5. Figure 5a confirms the accuracy of these observations. Our approximation deteriorates for \(\tau < 1\) day because of (68); using (67) in place of (68) increases the accuracy in this parameter range. The deleterious impact of detection delay is heightened as the antibiotics become more effective, and is lessened as the time to distribute antibiotics is shortened.

Substituting the base-case parameter values (aside from \(t_e - \tau\)) into (70) gives, using the second-order Taylor-series approximation \(e^{-x} \approx 1 - x + \frac{x^2}{2}\),

\[
\frac{D}{I} = 0.123 + 0.031(t_e - \tau) + \frac{0.639(1 - e^{-0.08(t_e - \tau)})}{t_e - \tau} \approx 0.174 + 0.029(t_e - \tau). \tag{83}
\]

Approximation (84) holds for antibiotic distribution times up to about one week, and suggests that an additional 2.9% of infected people perish for every day that it takes to deliver antibiotics. Figure 5b confirms the accuracy of this simple formula. As in (82), the effect in (84) is exacerbated as the antibiotics become more effective, and is mitigated as the intervention delay is decreased. Hence, rapid antibiotic distribution and biosensors are partial substitutes.

Figures 5c and 5d confirm equation (70)’s assertion that the number of deaths is linearly decreasing in the prophylactic and treatment efficacies of the antibiotics. Equation (70)
shows that the slope in Figure 5c gets steeper (i.e., prophylactic efficacy has more of an impact) as the number infected increases, as the time to distribute antibiotics decreases and as the treatment efficacy of antibiotics decreases. Similarly, the slope in Figure 5d steepens (i.e., treatment efficacy has more impact) as the number infected increases, the prophylactic efficacy decreases and the time to distribute antibiotics decreases. Hence, the prophylactic and treatment efficacies are partial substitutes for each other. Substituting base-case parameter values into (70) shows that

\[
\frac{D}{T} = 0.685 - 0.437e_1, \quad (85)
\]

\[
\frac{D}{T} = 0.346 - 0.135e_2. \quad (86)
\]

Comparing the coefficients in (85)-(86), we see that the prophylactic efficacy, which is partially controllable via drug compliance, is about three times more influential on a percentage basis than the treatment efficacy, which is an exogenous and largely unknown parameter (www.anthrax.mil estimates its range between 0.2 and 0.55).

5 Generalizations

This section considers four generalizations of our model: non-exponential disease progression, the symptomatic priority policy, the ring policy with \( p > 0 \), and voluntary pre-attack vaccination.

5.1 Non-exponential Disease Progression

In this subsection, our base-case result (69) is generalized to allow for non-exponential disease-stage durations. For \( i = 1, 2, 3 \), let \( X_i \) denote the random time a person spends in disease stage \( i \) (in the absence of medical intervention), and denote by \( f_i \) and \( F_i \) the
pdf and cdf. Also, let $T_s$ denote the random time that a typical infected person receives antibiotics. By equation (64), $T_s \sim U[\tau, t_e]$ under the mass service discipline.

Under mass service, an infected person dies if he (i) reaches stage 3 before receiving antibiotics, (ii) is in stage 2 when receiving antibiotics and then progresses, or (iii) receives antibiotics in stage 1, but needs hospitalization anyways and then progresses in the hospital.

These three disjoint events give

$$
\frac{D}{I} = P(X_1 + X_2 < T_s) + (1 - e_2)P(X_1 < T_s < X_1 + X_2) + (1 - e_1)(1 - e_2)P(X_1 > T_s),
$$

$$
= \frac{1}{t_e - \tau} \left[ \int_\tau^{t_e} \int_0^s f_1(t)F_2(s-t)dt \, ds + (1 - e_2) \int_\tau^{t_e} \int_0^s f_1(t)(1 - F_2(s-t))dt \, ds \right.
$$

$$
+ (1 - e_1)(1 - e_2) \int_\tau^{t_e} (1 - F_1(s))ds. \right] \quad (87)
$$

The non-exponential analog to the uncongested death estimate in equation (71) is to replace $(1 - e_2)$ in (87) with $(1 - e_2) \int_0^\infty F_2(s)\mu_H e^{-\mu_H s} ds$ because, for the fraction $1 - e_2$ for whom antibiotics provide ineffective treatment, these people progress if their prodromal time is less than their service time. Figure 6 shows the overcongested, uncongested and exact number of deaths versus the fraction infected, in the case where the disease durations are log-normal random variables with parameters given in Table 2 of Wein et al. This plot reveals that, compared to the exponential case, the number of hospital servers has a bigger impact on the number of deaths in the log-normal case. The underlying reason for this difference is the fatter left tail of the exponential distribution.

### 5.2 Symptomatic Priority Policy

This subsection considers the symptomatic priority policy in (30)-(31) in the overcongested case, where asymptomatic people in the antibiotics queue are only served if the number of servers exceeds the number of symptomatics in queue. As in §4, we begin our analysis by studying the antibiotics queue. We approximate the queue lengths by dividing the time
interval $[\tau, t_e]$ into two phases. From time $\tau$ to time $t_s$, we assume that the $Q^A_2(\tau) + Q^A_3(\tau)$ people who are symptomatic at time $\tau$ are served. Hence, the first phase ends at time

$$ t_s = \tau + \frac{Q^A_2(\tau) + Q^A_3(\tau)}{n_{A\mu_A}} = \tau + \frac{I(C_2 + C_3)}{n_{A\mu_A}}, \quad (88) $$

at which time we have $Q^A_0(t_s) = Q^A_0(\tau)$ and $Q^A_1(t_s) = Q^A_1(\tau)e^{-r_1(t_s-\tau)}$ because $Q^A_0(t)$ is constant and $Q^A_1(t)$ is exponentially decreasing during phase one. The asymptomatic people are served in the second phase, and we make the simplifying assumption that all servers are always busy distributing antibiotics to the asymptomatic people (in reality, a small fraction of these servers will be processing the new symptomatics), giving for $t \in [t_s, t_e]$,

$$ \frac{dQ^A_0(t)}{dt} = -\frac{Q^A_0(t)}{Q^A_0(t) + Q^A_1(t)} n_{A\mu_A}, \quad (89) $$

$$ \frac{dQ^A_1(t)}{dt} = -\frac{Q^A_1(t)}{Q^A_0(t) + Q^A_1(t)} n_{A\mu_A} - r_1 Q^A_1(t). \quad (90) $$

Letting $g(t) = \frac{n_{A\mu_A}}{Q^A_0(t) + Q^A_1(t)}$, we can solve (89)-(90) in terms of the unknown function $g(t)$,

$$ Q^A_0(t) = Q^A_0(t_s)e^{-\int_{t_s}^{t} g(u) \, du}, \quad (91) $$
Substituting these solutions into the definition of $g(t)$ gives the following integral equation that $g(t)$ must satisfy,
\[ g(t)\left[Q_0^A(t_s)e^{-\int_{t_s}^t g(u) \, du} + Q_1^A(t_s)e^{-\int_{t_s}^t (r_1 + g(u)) \, du}\right] = n_A \mu_A. \tag{93} \]

The substitution $z(t) = \int_{t_s}^t g(u) \, du$ turns (93) into the differential equation
\[ \dot{z}(t)[Q_0^A(t_s)e^{-z(t)} + Q_1^A(t_s)e^{-r_1(t-t_s)-z(t)}] = n_A \mu_A, \tag{94} \]
which can be rewritten in the separable form
\[ e^{-z} \, dz = \frac{n_A \mu_A \, dt}{Q_0^A(t_s) + Q_1^A(t_s)e^{-r_1(t-t_s)}}. \tag{95} \]

The solution, using the initial condition $z(t_s) = 0$, is
\[ z(t) = -\ln \left( \frac{r_1(Q_0^A(t_s) - n_A \mu_A (t - t_s)) + n_A \mu_A \ln \left( \frac{Q_0^A(t_s) + Q_1^A(t_s)e^{-r_1(t-t_s)}}{Q_0^A(t_s)r_1} \right)}{Q_0^A(t_s)r_1} \right). \tag{96} \]

Because $\int g = z$, equations (91)-(92) can be expressed as
\[ Q_0^A(t) = Q_0^A(t_s)e^{-z(t)}, \tag{97} \]
\[ Q_1^A(t) = Q_1^A(t_s)e^{-r_1(t-t_s)-z(t)}, \tag{98} \]
where $z(t)$ is given in (96). Equations (97)-(98) hold until the antibiotics queue empties at time $t_e$, which is computed by solving for when $z(t)$ goes to infinity in (96).

Now we mimic equation (69) to find the number of deaths in the overcongested regime. For simplicity, we assume that during the relatively short duration of the first phase, people in stage 2 are in a competing exponentials situation (i.e., service versus progression), whereas the queue congestion during the first phase implies that an exponential versus uniform
competition would be more realistic. The result is

\[ D = I(C_3 + C_4) + \left( \frac{r_2}{r_2 + \mu_A} + \frac{\mu_A}{r_2 + \mu_A}(1 - e_2) \right) IC_2 + \int_{t_s}^{t_e} r_1 Q_1^A(t) \, dt \]

\[ + (1 - e_1)(1 - e_2) \int_{t_s}^{t_e} g(t) Q_1^A(t) \, dt \]

\[ \approx I(C_3 + C_4) + (1 - e_2)(IC_2 + \int_{t_s}^{t_e} r_1 Q_1^A(t) \, dt) + \]

\[ (1 - e_1)(1 - e_2) \int_{t_s}^{t_e} g(t) Q_1^A(t) \, dt, \]  

(99)

where (99) follows from the fact (see Table 1) that \( \mu_A \gg r_2 \). To carry out the integrations in (99), we observe that \( r_1 Q_1^A(t) + g(t) Q_1^A(t) \) is exactly \( -\dot{Q}_1^A(t) \). However, the constants in front of each integral in (99) are different. We only need to compute the easier of the two integrals because for any constants \( a \) and \( b \) we have

\[ a \int r_1 Q_1^A + b \int g Q_1^A = b \int (r_1 Q_1^A + g Q_1^A) - (b-a) \int r_1 Q_1^A = -b \int \dot{Q}_1^A - (b-a) \int r_1 Q_1^A. \]  

(100)

Turning to the integral of \( Q_1^A(t) \), we have by (96) and (98) that

\[ \int_{t_s}^{t_e} Q_1^A(t) \, dt \]

\[ = \int_{t_s}^{t_e} \frac{Q_1^A(t_s)e^{-r_1(t-t_s)}}{Q_0^A(t_s)r_1} \left( r_1(Q_0^A(t) - n_A\mu_A(t - t_s)) + n_A\mu_A \ln \left( \frac{Q_0^A(t_s) + Q_1^A(t_s)}{Q_0^A(t_s) + Q_1^A(t_s)e^{-r_1(t-t_s)}} \right) \right) dt. \]  

(101)

We could directly integrate this expression, but the result is simpler and quite accurate if we replace the logarithmic term with its first-order Taylor expansion, which gives

\[ \int_{t_s}^{t_e} Q_1^A(t) \, dt \approx \int_{t_s}^{t_e} \frac{Q_1^A(t_s)e^{-r_1(t-t_s)}}{Q_0^A(t_s)r_1} \]

\[ \left( r_1(Q_0^A(t) - n_A\mu_A(t - t_s)) + \frac{n_A\mu_A r_1 Q_1^A(t_s)(t - t_s)}{Q_0^A(t_s) + Q_1^A(t_s)} \right) dt \]

\[ = \frac{Q_1^A(t_s)}{r_1^2} \left( r_1 - \alpha + e^{-r_1(t_e-t_s)} (-r_1 + \alpha (1 + r_1(t_e - t_s))) \right), \]  

(102)

where \( \alpha = n_A\mu_A/(Q_0^A(t_s) + Q_1^A(t_s)) \).
Using (100), we rewrite equation (99) as

\[ D = I(C_3 + C_4) + (1 - e_2)IC_2 + (1 - e_2) \int_{\tau}^{t_s} r_1Q^A_1(t) \, dt + (1 - e_1)(1 - e_2)Q^A_1(t_s) \]
\[ + e_1(1 - e_2)r_1 \int_{t_s}^{t_e} Q^A_1(t) \, dt. \]  

(103)

Finally, substituting in the integrations, we get

\[ D = I(C_3 + C_4) + (1 - e_2)I[C_2 + C_1(1 - e^{-r_1(t_e - \tau)})] + (1 - e_1)(1 - e_2)IC_1e^{-r_1(t_s - \tau)} \]
\[ + e_1(1 - e_2)Q^A_1(t_s) \left( r_1 - \alpha + e^{-r_1(t_e - t_s)}(-r_1 + \alpha(1 + r_1(t_e - t_s))) \right). \]  

(104)

Computational results (not shown here) reveal that (104) is quite accurate (e.g., similar to Figures 4 and 5), except when the time to distribute antibiotics exceeds about 8 days. While the expression in (104) is rather cumbersome, it still allows for some observations. As in the mass service case, the number of deaths is linear in the number infected, and the fraction of infected people who die is linearly decreasing in both the prophylactic and treatment antibiotic efficacies. Substituting base-case parameter values yields

\[ \frac{D}{T} = 0.695 - 0.502e_1, \]  

(105)

\[ \frac{D}{T} = 0.308 - 0.162e_2. \]  

(106)

Substituting the base-case parameter values of \(e_1\) and \(e_2\) into (105)-(106), we find that the fraction of infected people who die is reduced from 0.292 for the mass service policy to 0.243 for the symptomatic priority policy. Numerical computations (not shown here) reveal that the difference in performance between these two policies goes down as \(t_e - \tau\) is reduced. Comparing the coefficients in (105)-(106) with those in (85)-(86) shows that the impact of prophylactic effectiveness is slightly greater under the symptomatic priority policy than the mass service policy. As in the mass service case, prophylactic efficacy is about three times more effective on a percentage basis than treatment efficacy.
5.3 The Ring Policy

This subsection analyzes how the number of deaths depends on the ring parameter \( p \) in the overcongested scenario. Since for a given zone location, \( p \) maps to time \( t_p \) via (24), we instead work with \( t_p \). Symptomatics start receiving medical intervention at time \( \tau \), and at time \( t_p \) mass service is initiated for everyone in the zone. In light of our analysis in §5.2, it is likely that the death-minimizing time to start serving asymptomatics is close to time \( t_s \) as defined in (88). That is, it is probably optimal to clear the queue of symptomatics, and then switch to mass service. Symptomatics may not be served fast enough if \( t_p < t_s \), and capacity is wasted if \( t_p > t_s \).

Hence, we assume that the ring time \( t_p = t_s + \Delta t \), where \( \Delta t \) is the additional wait before switching to mass service. For practical (i.e., not too large) values of \( \Delta t \), the antibiotics queue will be emptied at roughly time \( t_e + \Delta t \), where \( t_e \) is given in (65), because the antibiotic distributors are mostly idle during the interval \([t_s, t_s + \Delta t]\). To compute the number of deaths, we perform a separate analysis over three time intervals: \([\tau, t_s]\), \([t_s, t_s + \Delta t]\), and \([t_s + \Delta t, t_e + \Delta t]\). Letting \( D_1, D_2 \) and \( D_3 \) denote the number of deaths in these three time intervals, we have

\[
D = I(C_3 + C_4) + D_1 + D_2 + D_3. \tag{107}
\]

\( D_1 \) was computed in §5.2, and consists of those who are in stage 2 at time \( \tau \) and all those who enter stage 2 during \([\tau, t_s]\),

\[
D_1 = (1 - e_2)(IC_2 + I(e^{-r_1\tau} - e^{-r_1t_s})). \tag{108}
\]

During the rather dormant time interval, \([t_s, t_s + \Delta t]\), we have \( I(e^{-r_1t_s} - e^{-r_1(t_s + \Delta t)}) \) people who progress from stage 1 to stage 2. These people are apt to experience no wait in the antibiotics queue, and the fraction of these people who die is \( 1 - e_2 \), invoking the
\[ D_2 = (1 - e_2)I(e^{-r_1 t_s} - e^{-r_1 (t_s + \Delta t)}). \] (109)

We cannot apply the results from §4.2 to the third time interval, \([t_s + \Delta t, t_e + \Delta t] \), because \(Q^A_2(t)\) no longer drops linearly over time, but rather rises and falls somewhat parabolically. The bulk of the people in this time interval start out in \(Q^A_0\) and \(Q^A_1\). Let \(L_R\) be the number of people who progress from \(Q^A_1\) all the way to \(Q^A_3\) before getting served, let \(S\) denote the number of people who get served from compartment \(Q^A_2\), and let \(G\) be the number of people who enter compartment \(Q^A_2\). These definitions imply that

\[ D_3 = L_R + (1 - e_2)S + (1 - e_1)(1 - e_2)(Q^A_1(t_s + \Delta t) - G), \]

where the last term in \(D_3\) represents those who are served while in stage 1, but the antibiotics are not effective. Because \(G = L_R + S\), we have

\[ D_3 = e_2L_R + (1 - e_2)G + (1 - e_1)(1 - e_2)(Q^A_1(t_s + \Delta t) - G). \] (110)

We calculate the quantities \(G\) and \(L_R\) from a probabilistic analysis, under the reasonably accurate assumption that \(Q^A_1(t)\) drops linearly to zero throughout this time interval. Given that someone enters stage 2 at time \(t\), we have a competition between an \(\exp(r_2)\) random variable, call it \(X_2\), and a \(U[t, t_e + \Delta t] \) random variable denoted by \(U(t)\). This competition yields

\[ L_R = \int_{t_s + \Delta t}^{t_e + \Delta t} r_1Q^A_1(t_s + \Delta t) \frac{t_e + \Delta t - t}{t_e - t_s} P(X_2 \leq U(t)) \, dt. \] (111)

Substituting

\[ P(U(t) < X_2) = \int_{t}^{t_e + \Delta t} \frac{e^{-r_2(u-t)}}{t_e + \Delta t - t} \, du = \frac{1 - e^{-r_2(t_e + \Delta t - t)}}{r_2(t_e + \Delta t - t)} \] (112)

into (111) and integrating gives

\[ L_R = \frac{r_1Q^A_1(t_s + \Delta t)(t_e - t_s)}{2} + \frac{r_1Q^A_1(t_s + \Delta t)}{r^2_2(t_e - t_s)} (1 - e^{-r_2(t_e - t_s)}) - \frac{r_1Q^A_1(t_s + \Delta t)}{r_2} \]

\[ = Ir_1e^{-r_1(t_s + \Delta t)} \left( \frac{(t_e - t_s)}{2} + \frac{1 - e^{-r_2(t_e - t_s)}}{r^2_2(t_e - t_s)} - \frac{1}{r_2} \right). \] (113)

32
The quantity $G$ is simply all those who enter stage 2 by progressing from $Q_1^A$,

$$G = \int_{t_s+\Delta t}^{t_e+\Delta t} r_1 Q_1^A(t) \, dt = \frac{I r_1(t_e - t_s) e^{-r_1(t_s+\Delta t)}}{2}. \quad (114)$$

Equations (110), (113) and (114) give

$$D_3 = I e^{-r_1(t_s+\Delta t)} \left[ \frac{r_1(t_e - t_s) (e_1 + e_2 - e_1 e_2)}{2} + \frac{e_2 r_1 (1 - e^{-r_2(t_e-t_s)})}{r_2^2(t_e-t_s)} \right] - \frac{e_2 r_1}{r_2} + (1 - e_1)(1 - e_2). \quad (115)$$

By (12), (107), (108), (109) and (115), the total number of deaths is approximately

$$\frac{D}{I} = 1 - C_1 - e_2 C_2 + (1 - e_2) e^{-r_1 \tau} + e^{-r_1(t_s+\Delta t)} \left[ \frac{r_1(t_e - t_s) (e_1 + e_2 - e_1 e_2)}{2} \right. $$

$$+ \left. \frac{e_2 r_1 (1 - e^{-r_2(t_e-t_s)})}{r_2^2(t_e-t_s)} - \frac{e_2 r_1}{r_2} - e_1 (1 - e_2) \right]. \quad (116)$$

Fixing all parameter values except $\Delta t$ in (116) gives

$$\frac{D}{I} = 0.634 - 0.361 e^{-0.08 \Delta t}$$

$$\approx 0.273 + 0.029 \Delta t \quad \text{for small } \Delta t. \quad (117)$$

Note that the coefficient in front of the ring delay parameter $\Delta t$ in (117) is identical to the coefficient for the time it takes to distribute antibiotics in (84). For small $\Delta t$, we directly compute the impact of the ring parameter $p$, via $\frac{dD}{dp} = \frac{dD}{dt} \frac{dt}{dp} = \frac{bN}{r_1(I_1(0)-pN)}$, which follows from (24), (117) and the fact that $\Delta t = t_p - t_s$. Plugging in base values, with $I_1(0) = 0.5N$ (Table 1), we have $\frac{dD}{dp} = \frac{0.725}{1-2p} \approx 0.725 + 1.45p$.

Finally, we note that in the limiting case $p = 1$, where only symptomatic people are given antibiotics, the overcongested estimate is simply

$$\frac{D}{I} = 1 - e_2. \quad (118)$$
5.4 Voluntary Pre-attack Vaccination

Because of the nonnegligible probability of another anthrax attack, and because of the great cost and difficulty of mounting an effective post-attack response, serious consideration should be given to a voluntary pre-attack anthrax vaccination program (Wein and Kaplan). Although a safe and reliable anthrax vaccine exists, it requires a series of six shots over 18 months plus an annual booster, and hence is not viable for widespread civilian use (Inglesby et al.). A more practical vaccine may become available within the next several years.

Suppose a fraction $v$ of the population opts to be vaccinated prior to an attack, and assume that the vaccine’s prophylactic efficacy is $e_v$ (i.e., a fraction $e_v$ of vaccinees will be protected from an attack). Further suppose that pre-vaccinated people do not receive antibiotics after an attack; if pre-vaccinated people were to receive antibiotics after all of the unvaccinated people, then more lives would be saved by the vaccine, but because $e_v$ is likely to be high, we omit this for simplicity. Pre-attack vaccination changes our base-case result (70) in two ways: the number initially infected would shrink by the factor $(1 - ve_v)$ and the time to distribute antibiotics, $t_e - \tau$, would decrease by the factor $(1 - v)$. Hence, if we define $D_v$ to be the number dead in a service zone under pre-attack vaccination, then the relative reduction in deaths is (after substituting in base-case values from Table 1)

$$\frac{D - D_v}{D} = 1 - (1 - ve_v)\left(0.847 - 0.425v + \frac{0.545(1 - e^{-0.32(1-v)})}{1 - v}\right).$$ (119)

If the only effect of pre-attack vaccination was to reduce the number of initially infected people, then we would expect that $(D - D_v)/D = ve_v$. However, as shown in Figure 7 with $e_v$ set equal to 0.9, pre-attack vaccination has a secondary benefit by easing the logistics of antibiotic distribution: people that were not pre-vaccinated receive their post-attack antibiotics more quickly, and additional lives are saved. For example, if 50% of the population is prevaccinated with a vaccine that is 90% effective, then the relative reduction in deaths is 56% rather than 45%.
6 Spatial Aggregation of Service Zones

Because the estimated fraction of infected people who die, $D/I$, is independent of $I$, and because the great majority of infections occur in overcongested zones (see Fig. 2b in Wein et al.), we can estimate the total number of deaths over the entire region by replacing the $I$ in equation (69) with the total number infected over the whole region, a quantity we denote by $I_T$. For example, for a release of $10^{15}$ spores (Table 1), which is about 1 kg, the exact $D/I$ value in our model (found by spatially integrating out to 6000 km downwind and 40 km crosswind) is 0.253, compared to the approximated value of 0.29.

By equation (5), the total number of infected people is

$$I_T = \int_0^A \int_0^\infty \int_{-\infty}^{\infty} \min\left(\frac{N}{A}, \frac{k_1 x^{-2d} \exp\left(-k_2 y^2 x^{-2d} - k_3 x^{-2d}\right)}{c_1 - c_2 a}\right) \, dy \, dx \, da. \quad (120)$$

While we have not been able to integrate (120), for a release size less than about eight grams (roughly the amount in all the tainted envelopes in the 2001 postal attack), the first argument in the minimum function in (120) is not needed, and the total number of infected people

Figure 7: The synergistic effect of pre-attack vaccination.
people is given by

\[ I_T = \int_0^A \int_0^\infty \int_{-\infty}^{\infty} I_1(x, y, a, 0) \, dy \, dx \, da, \]  

(121)

\[ = \int_0^A \int_0^\infty \int_{-\infty}^{\infty} k_1 x^{-2d} e^{-k_3 x^{-2d}} \left[ \int_{-\infty}^{\infty} e^{-k_4 y^2} \, dy \right] \, dx \, da \quad \text{where} \quad k_4 = k_2 x^{-2d}, \]  

(122)

\[ = \int_0^A \int_0^\infty \frac{k_1 x^{-2d} e^{-k_3 x^{-2d}}}{(c_1 - c_2 a) \sqrt{k_4}} \, dx \, da \]  


\[ = \int_0^A \frac{k_1 \sqrt{\pi}}{\sqrt{k_2(c_1 - c_2 a)}} \left[ \int_0^{\infty} x^{-d} e^{-k_3 x^{-2d}} \, dx \right] \, da, \]  

(124)

\[ = \int_0^A \frac{k_1 \sqrt{\pi}}{\sqrt{k_2(c_1 - c_2 a)}} \left[ -\frac{1}{2d} \int_0^{\infty} w^{-(d+1)/2d} e^{-k_3 w} \, dw \right] \, da, \quad \text{where} \quad w = x^{-2d}, \]  

(125)

\[ = -\int_0^A \frac{k_1 \sqrt{\pi}}{2d \sqrt{k_2(c_1 - c_2 a)}} \left[ \Gamma \left( \frac{d-1}{2d} \right) \right] \frac{1}{k_3^{d/2d}} \, da, \]  

(126)

\[ = -\frac{k_1 \sqrt{\pi} \Gamma \left( \frac{d-1}{2d} \right) \log \left( \frac{c_1}{c_1 - c_2 a} \right)}{2\sqrt{k_2 c_2 k_3^{d/2d}}} \]  

(127)

Hence, for an attack under eight grams in size, the total number of dead is linear in the release size, and is concave for larger releases.

## 7 Concluding Remarks

This paper is the first to our knowledge to attempt to derive mathematical expressions for the number of deaths resulting from an aerosol bioterror attack with a noncontagious agent. As in Kaplan et al., we combine the traditional solution approach to differential equations with a probabilistic interpretation and subsequent analysis of these equations. This leads to relatively simple equations for the number of deaths (e.g., equation (70)), and reveals how this performance measure is impacted by various key parameter values; indeed, this probabilistic approach even works in the non-exponential case (see §5.1), where the model is an unwieldy system of partial differential equations (Wein et al.). In particular, we find that the number of deaths in a service zone is nearly linear in the number of infected
people in the service zone; i.e., the fraction of infected people that die is nearly a constant. This key ratio is linear in the length of time it takes to distribute antibiotics and in the efficacy of the antibiotics – both as a prophylactic and as treatment. This ratio is also increasing and concave in the intervention delay, i.e., the delay from the time of the attack until intervention begins. This delay itself depends on whether the attack is detected by biosensors or by early symptomatics: large attacks will have shorter delays and will be detected by biosensors, whereas small attacks will evade the biosensors and have longer delays. Our biosensor analysis shows that the minimal detectable attack size is linear in the biosensor detection limit and, in the sparse geographical density limit, varies with $w^{2d}$, where $w$ is the grid width and $d = 0.78$ for slightly unstable meteorological conditions and varies between 0.71 and 0.91 for other weather conditions (Hanna et al.). In §5, we also show how queue management policies, namely the geographical ring policy that places people in queue and the symptomatic priority policy that distributes antibiotics to symptomatic people first, impact the death toll. Finally, Figure 7 reveals the synergistic effect of voluntary pre-attack vaccination, which protects the people who are pre-vaccinated and allows the unvaccinated to receive their post-attack antibiotics more quickly.

The biggest shortcoming of our analysis is our failure to estimate the number of deaths in the moderately congested case, which would have allowed us to explicitly quantify how the number of deaths is reduced by adding servers to the medical care queue. However, our analyses of the uncongested and overcongested cases allow us estimate the difference in deaths between having no medical care providers and having an ample number of them.

Relative to the simulation study in Wein et al., one of the more important things we learned from this analysis is the big impact that the incubation period distribution has in our qualitative conclusions. Wein et al. assume a log-normal incubation period and find that rapid antibiotic distribution has a bigger impact than biosensors, that the number of deaths
is profoundly affected by the number of medical care providers, and that the symptomatic priority policy has only a modest impact. In contrast, with exponential incubation times, we find that biosensors prevent more deaths than rapid antibiotic distribution, medical care providers have only a modest impact on the number of deaths and the symptomatic priority policy is reasonably effective. The discrepancy in these results is due to the fatter left tail of the exponential distribution. However, we assume that the time to detect early symptomatics is 48 hours in both cases. Clearly, with an exponential incubation period, the time to detect early symptomatics would be significantly smaller than 48 hours, and the utility of biosensors would be smaller than suggested by our analysis. Moreover, the impact of the number of medical care providers is larger in the exponential incubation case for values of $\tau$ smaller than 48 hours.

This exponential vs. log-normal discrepancy is problematic because not that much is known about the incubation period distribution. Brookmeyer et al. (2001) fit intervention-censored data from the unintentional release from a former Soviet bioweapons factory (Guillemin 1999), which is the only relatively large human data set in existence, to a log-normal distribution that is independent of dose; Wein et al. could not identify any dose-dependence (or age-dependence) in the incubation period from an earlier version of the same data set (Meselson et al., 1994), perhaps because of the somewhat small size of the attack. Brookmeyer et al. (2003) develop a simple competing-risks mathematical model of inhalation anthrax, where inhaled spores either are cleared or germinate (eventually causing disease), which suggests that the incubation period distribution approaches an exponential distribution at very low doses, and that the incubation period may depend inversely on dose. However, the paucity and questionable integrity (despite heroic efforts by researchers to uncover the truth; see Meselson et al., Guillemin) of the underlying Sverdlovsk data (for over a decade, the Soviet government claimed that the cause of deaths was bad meat because the clandestine factory
was in defiance of a bioweapons treaty it signed), coupled with the impracticality of performing human studies, suggest that the precise nature of the incubation period distribution will remain unknown, at least until a large bioterror attack occurs. In the meantime, it would be prudent to hedge our bets and assume that either a log-normal or an exponential distribution, with or without dose-dependence, could transpire.
References


