STATISTICS IN MEDICINE
Statist. Med. 2007; **26**:5225–5252
Published online 22 October 2007 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/sim.3093

The Bayesian aerosol release detector: An algorithm for detecting and characterizing outbreaks caused by an atmospheric release of *Bacillus anthracis*

William R. Hogan*,†, Gregory F. Cooper, Garrick L. Wallstrom, Michael M. Wagner and Jean-Marc Depinay

The RODS Laboratory, Department of Biomedical Informatics, University of Pittsburgh, Pittsburgh, PA, U.S.A.

SUMMARY

Early detection and characterization of outdoor aerosol releases of *Bacillus anthracis* is an important problem. As health departments and other government agencies address this problem with newer methods of surveillance such as environmental surveillance through the BioWatch program and enhanced medical surveillance, they increasingly have newer types of surveillance data available. However, existing methods for the statistical analysis of surveillance data do not account for recent meteorological conditions, as human analysts did in the case of the Sverdlovsk anthrax outbreak of 1979 to determine whether the locations of victims were consistent with meteorological conditions in the week preceding their onset of illness.

This paper describes the Bayesian aerosol release detector (BARD), an algorithm that analyzes both medical surveillance data and meteorological data for early detection and characterization of outdoor releases of *B. anthracis*. It estimates a posterior distribution over the location, quantity, and date and time conditioned on a release having occurred. We report a proof-of-concept evaluation of BARD, which demonstrates that the approach shows promise and warrants further development and evaluation. Copyright 2007 John Wiley & Sons, Ltd.

KEY WORDS: biosurveillance; outbreak-detection algorithms; Bayesian analysis; atmospheric dispersion models; statistical surveillance; inhalational anthrax

*Correspondence to: William R. Hogan, Department of Biomedical Informatics, Suite M-183 Parkvale Bldg, 200 Meyran Ave., Pittsburgh, PA 15260, U.S.A.
†E-mail: wrh9@pitt.edu, wrh@cbmi.pitt.edu

Contract/grant sponsor: Centers for Disease Control and Prevention; contract/grant number: 1 RO1 PH000026-01
Contract/grant sponsor: Defense Advanced Research Projects Agency; contract/grant number: F30602-01-2-0550
Contract/grant sponsor: National Science Foundation; contract/grant number: 0325581
Contract/grant sponsor: Pennsylvania Department of Health; contract/grant number: ME-01-727

Received 12 October 2006
Accepted 29 August 2007

1. INTRODUCTION

Early detection of a surreptitious aerosol release of *Bacillus anthracis* is an important problem [1, 2]. A release could infect hundreds of thousands of individuals. Without early detection, mortality could be as high as 30 000 [1] to 3 million [3].

There is a narrow window of time within which infected individuals must receive antibiotics and vaccines. Statistical models of release scenarios suggest that the ability of antibiotics and vaccines to reduce mortality falls rapidly over the week following a release [1, 4–7]. If these treatments could be administered to all exposed immediately, mortality could be limited to less than 1 per cent [4]. A response beginning on day 2 results in mortality of approximately 8 per cent [7] to 42 per cent [1], whereas a response on day 7 has mortality of 25–95 per cent. In the model of Wagner *et al.*, a delay of just 1 h results in additional economic costs (primarily from additional deaths) of as much as $250 million [2].

To enable treatment of the maximum number of exposed people within the short window of opportunity, early detection must be accompanied by rapid characterization of the release. In particular, identification of the people exposed, or at least rapid and accurate risk stratification of people according to their probability of exposure, is important. Even the most highly developed emergency response system has a finite rate at which it can administer antibiotics and vaccinations to people. At a treatment rate of 10 000 people/h, it would take just over 4 days to treat one million people. In the event of a large-scale release, responders will face a triage problem, whose optimal solution with respect to minimizing mortality and morbidity depends on accurate estimation of exposure levels.

In response to these requirements for early detection and identification of exposed individuals, health departments and other government agencies have been augmenting notifiable disease surveillance with two new methods of surveillance: (1) monitoring of the air (the BioWatch program) and (2) more intensive monitoring of human health through collection of data from emergency departments (EDs) and other sources such as the retail industry. The latter method of enhanced medical surveillance is sometimes referred to as *syndromic surveillance* [8–10]. Both methods attempt to detect releases before the healthcare system confirms the diagnosis in individual cases (the basis of the notifiable disease reporting system), a process that took 9 days after the release in Sverdlovsk [11] and at least 8 days after terrorists mailed spore-laden letters in 2001 [12].

A consequence of these new surveillance methods is that health departments have a broader range of data to analyze than in the past. The data may include results of tests from BioWatch that indicate the presence or absence of selected biological agents in the air. They may include meteorological data such as wind direction and speed needed to interpret the results from BioWatch devices. The data may also include counts of ED visits for particular syndromes (such as generic respiratory and gastrointestinal syndromes), sales of selected over-the-counter healthcare products, orders for laboratory tests, and school absenteeism.

At present, health departments analyze these data primarily using univariate time series algorithms and spatial scan statistics [9, 13–17]. They may also utilize tools provided by the BioWatch program for the integrated analysis of air-sampler data and meteorological data to determine likely release locations, release dates, and locations where people are likely to have been exposed [18]. Algorithms for multivariate analysis exist [19–22], but their use in practice is less frequent.

DOI: 10.1002/sim

A limitation of these methods is that they do not analyze both meteorological and medical surveillance data to detect and characterize releases. Weather conditions at the time of the release and the release location, primarily, will determine where exposure occurs and at what levels. An analyst considering the possibility of a release would therefore check the spatial distribution of cases against recent meteorological data, as did Meselson *et al.* when analyzing data from the Sverdlovsk anthrax outbreak of 1979 [23]. However, Meselson *et al.* did not formalize their analysis as an algorithm, and no prior work on such an algorithm exists, to our knowledge. The existence of BioWatch tools for integrated analysis of air-sampler results and meteorological data also confirms the need for tools for similar analysis of medical surveillance and meteorological data.

This paper describes the Bayesian aerosol release detector (BARD), an algorithm that analyzes both medical surveillance data and meteorological data for early detection and characterization of outdoor releases of *B. anthracis*. BARD computes the posterior probability of a release given these data. It estimates a posterior distribution over the location, quantity, and date and time conditioned on a release having occurred. BARD can also estimate, at various locations within the area under surveillance, the probability of infection given assumptions about the infectious dose and other relevant model parameters. This information may help guide prioritization of antibiotic and vaccine administration. The current version of BARD is specialized to detect anthrax releases and analyzes only data about ED visits for respiratory complaints (medical surveillance) and meteorological data. However, the approach is general and could be extended to include additional data and other biological agents that can be disseminated by outdoor aerosol release.

2. BACKGROUND

In this section, we review the three current methods of anthrax surveillance. Additionally, both to illustrate the need for integrated analysis of surveillance and meteorological data and to motivate our fundamental approach in BARD, we review how Meselson *et al.* performed an integrated analysis of human health, animal health, and meteorological data to characterize the 1979 Sverdlovsk anthrax outbreak.

2.1. Notifiable disease surveillance

The traditional method of disease surveillance is called the *notifiable disease reporting system*. Laws in all 50 states require physicians, hospitals, and laboratories to report to health departments the existence of individuals with any of approximately 50 diseases, including anthrax. The diseases for which reporting is mandatory are called *notifiable diseases*.

The strength of the notifiable disease reporting system is that it is in place nationwide and therefore covers 100 per cent of the population. Its chief limitation is the delay from infection to suspected or confirmed diagnosis, which is the event that triggers reporting. The delay between infection and reporting of the first case in the Sverdlovsk anthrax outbreak was 9 days [11]. The delay in the 2001 mail contaminations was at least 8 days [12]. Although the latter incident was not an outdoor aerosol release, the time to detection was similar and it is one of only two data points available on how quickly the notifiable disease reporting system detected terrorist incidents involving anthrax.

DOI: 10.1002/sim

2.2. BioWatch

As of 2003, BioWatch—a program of the Department of Homeland Security—had deployed approximately 500 air-sampling devices in 31 cities [24]. These devices capture particulate matter from the air on filter paper, which are collected periodically and sent to laboratories for testing. The testing identifies the presence of the DNA of several biological agents, including *B. anthracis* [24]. At present, the filters are retrieved and analyzed daily, unless there is heightened suspicion or threat.

Although much information about BioWatch is classified, there is enough information publicly available to identify its strengths and limitations with reasonable certainty. The strengths of the BioWatch program are that it can detect an aerosol release quickly—as early as 24 h after it occurs. It generates few false alarms, identifies the organism released, is unlikely to miss a large release, can roughly estimate the concentration of spores that were in the air, and—if there are enough devices in a region—it may be able to rapidly and accurately characterize the geographic extent of contamination. The limitations of BioWatch are that it is not universally deployed (approximately 30 cities) and it cannot distinguish between viable (infectious) organisms and those that are dead. Even in those cities in which it is deployed, the limited number of air collectors—approximately 15 per city—allows for the possibility that a 'significant' release will not be detected if it passes between detectors [25]. Additionally, if only one or two collectors register the presence of *B. anthracis*, then characterization of the source (location, time, quantity, etc.) may be difficult because several recent wind directions, speeds, etc. could be consistent with the observed results.

BioWatch also provides health departments with a system called the Biological Warning and Incident Characterization (BWIC) system, which estimates the exposed area resulting from a release by analyzing BioWatch results and meteorological data [18]. BWIC also has capabilities for displaying medical surveillance data on a map but does not use these data in its analysis.

2.3. Enhanced medical surveillance

Enhanced medical surveillance is the monitoring of data that may reflect early stages of illness in humans. The data may come from both the healthcare system and entities outside the healthcare system such as retail chains.

The strength of enhanced medical surveillance is its potential for earlier detection than the notifiable disease reporting system. It also has the potential to detect novel pathogens, whereas BioWatch and the notifiable disease reporting system are limited to known pathogens (however, our focus in this paper is the known pathogen anthrax). Its limitations are that it is not universally deployed throughout the United States, it cannot identify the biological agent, and it produces frequent false alarms in most settings (although the frequency of false alarms is under the control of the operators and therefore could be reduced). Whether enhanced medical surveillance ultimately can provide earlier warning than notifiable disease surveillance is controversial [26–28], but the work we report here represents an effort in that regard.

Statistical analysis is important in enhanced medical surveillance because the data have high variability in the absence of outbreaks; thus, extracting an outbreak 'signal' from background 'noise' can be difficult [28–30]. In addition, researchers typically automate the statistical analyses because the volume of data is often too high for routine, manual review. Common statistical approaches include time series analysis, control charts, and spatial scan statistics [9, 15, 27–29, 31].

DOI: 10.1002/sim

2.4. Meslson's analysis of the 1979 Sverdlovsk outbreak

Meselson *et al.* analyzed human, animal, and meteorological data in their analysis of whether an outbreak of anthrax in Sverdlovsk, the U.S.S.R., in 1979 was caused by contaminated meat—the official Soviet explanation—or an aerosol release of *B. anthracis* from a military microbiology facility, as was suspected by the U.S. intelligence experts [23]. The question was important because the latter explanation would be relevant to whether the Soviet Union was in violation of international treaties to which it was a party.

Meselson *et al.* obtained data about the victims of the outbreak, outbreaks of anthrax in animals that occurred at the same time, and meteorological conditions. In particular, they interviewed victims and their families and acquaintances and reviewed available records to determine time of symptom onset, time of hospital admission, and time of death (if applicable). They collected detailed information about the locations of victims prior to the onset of illness. They obtained records of outbreaks of anthrax in animals in nearby villages. They obtained records of three-hourly surface weather observations from local airports.

They created several maps of the locations of victims at different times. They compared these maps with the locations of animal outbreaks, meteorological conditions, and the most likely source of a possible release, a military microbiology facility.

They found that the area encompassing the daytime locations on 2 April 1979 of 57 victims was a narrow, 4-km long zone extending south–southeast from the microbiology facility (Figure 1(a)). Of the remaining nine victims, eight were plausibly inside the zone and for one there was insufficient information. They found that the six villages that had experienced outbreaks of anthrax in animals were located along the axis (when extended to 50 km) of the zone that encompassed the daytime locations of humans on 2 April (Figure 1(b)).

The meteorological data showed that the wind direction had a compass bearing that ranged from 320 to 350° from 1:00 AM to 7:00 PM local time (Figure 1(c)) on 2 April, consistent with the compass bearing of 330° of the high-risk zone for humans and animals (meteorological convention is to report the direction from which the wind is blowing). A postulated release date of 2 April was also consistent with available information about the incubation period of inhalational anthrax and the dates of symptom onset in victims.

Meselson *et al.* also estimated that a few milligrams to a gram of spores escaped into the air from the facility, depending on assumptions about the lethal dose of spores and the turbulence of the atmosphere [32]. This quantity of release is consistent with the explanation for the release of two former Soviet officials, who claim that workers in the military microbiology facility mismanaged a ventilation system with subsequent discharge of spores into outside air [33]. Because it estimated a small size of the release, the study by Meselson *et al.* was insufficient evidence *on its own* that the Soviets were in violation of a treaty that banned the production of biological weapons (the treaty permitted production of small amounts for defensive purposes). When considered in the light of other evidence, however, the study supported the conclusion of a violation.

The analysis conducted by Meselson *et al.* was labor and time intensive. It required manually searching through the varying physical locations of victims during the week preceding the outbreak for spatial patterns consistent with the weather and a model of atmospheric dispersion. A motivation for developing BARD was to automate this type of analysis so that it can be performed on a routine basis and in real time, with notification to relevant personnel (i.e. an alarm) when the surveillance data are sufficiently suggestive of a release.

DOI: 10.1002/sim

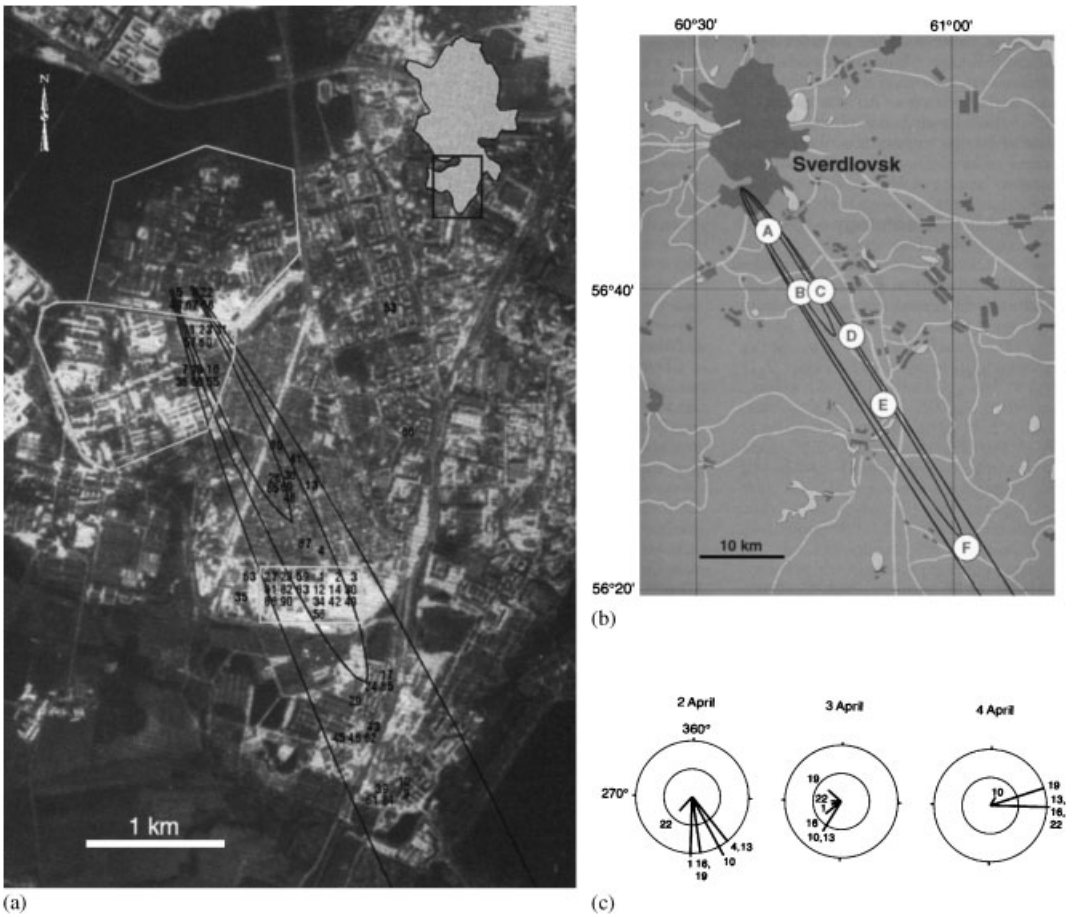

Figure 1. (a) The high-risk zone for humans. Numbers indicate daytime locations of victims on 2nd April. The ellipse-like lines are isopleths of the airborne concentration of spores produced by the Gaussian plume model. The military microbiology facility resides just northwest of the zone. (b) The high-risk zone for animals. The six villages where outbreaks of anthrax occurred in animals are shown. (c) Meteorological data from 2nd to 4th April 1979. The only times the wind was blowing towards the south-southeast was from 1 AM to 7 PM on April 2nd. All three figures are reprinted with permission from Meselson M *et al*. The Sverdlovsk anthrax outbreak of 1979 (*Science* 1994; **266**(5188):1202–1208). Copyright 1994 AAAS.

3. DESCRIPTION OF THE BARD ALGORITHM FOR DETECTING WIND-BORNE OUTBREAKS

BARD uses Bayes' rule to compute the posterior probability of an outbreak hypothesis:

$$P(H_i|\mathbf{B},\mathbf{G},\mathbf{M}) = \frac{P(\mathbf{B}|H_i,\mathbf{G},\mathbf{M})\,P(H_i|\mathbf{G},\mathbf{M})}{\sum_j P(\mathbf{B}|H_j,\mathbf{G},\mathbf{M})\,P(H_j|\mathbf{G},\mathbf{M})} \tag{1}$$

DOI: 10.1002/sim

where the H_i and H_j belong to a set H of mutually exclusive and exhaustive outbreak hypotheses, **B** is a matrix of biosurveillance data, **G** is a matrix that contains population and location data for the geographical sub regions (such as zip codes) of the region (city, county, or metropolitan statistical area) from which the surveillance data were taken, and **M** is a matrix of meteorological data. We describe these matrices and how we compute the likelihood of the surveillance data under H_i, $P(\mathbf{B}|H_i, \mathbf{G}, \mathbf{M})$ below.

At present, BARD considers two outbreak hypotheses and makes assumptions to simplify the derivation of the necessary prior and conditional probabilities. The first hypothesis H_0 is that 'background' respiratory disease that we have seen historically is the only 'respiratory disease' causing illness in the geographic region. The second hypothesis H_1 is that both background respiratory disease and an outbreak of inhalational anthrax due to a release of *B. anthracis* are causing illness in the most recent 24-h period. At present, we assume that (1) H_0 and H_1 are mutually exclusive and exhaustive, (2) the prior probabilities of the hypotheses are independent of **G** and **M**, and (3) the probability of **B** given H_0 is independent of **M**. As we discuss below, we condition background respiratory disease instead on month of year and day of week to account for seasonal and day-of-week effects in historical surveillance data. Assuming exclusivity of H_0 and H_1 could cause BARD to signal false alarms during other types of outbreaks such as influenza, and thus worsen its overall measured outbreak-detection performance in the evaluation we discuss below. Conditioning background disease on weather—either in addition to or instead of—month of year and day of week could improve BARD's ability to detect releases, but we did not study it here.

With the above assumptions and the fact that at present **B** contains one column (counts of ED visits for respiratory complaints) and thus is the vector **b**, equation (1) simplifies to

$$P(H_1|\mathbf{b}, \mathbf{G}, \mathbf{M}) = \frac{P(\mathbf{b}|H_1, \mathbf{G}, \mathbf{M})P(H_1)}{P(\mathbf{b}|H_1, \mathbf{G}, \mathbf{M})P(H_1) + P(\mathbf{b}|H_0, \mathbf{G})P(H_0)} \tag{2}$$

We must therefore specify the two prior probabilities—$P(H_0)$ and $P(H_1)$—and the two conditional probabilities—$P(\mathbf{b}|H_0, \mathbf{G})$ and $P(\mathbf{b}|H_1, \mathbf{G}, \mathbf{M})$. Because we assume that H_0 and H_1 are mutually exclusive and exhaustive, $P(H_0) = 1 - P(H_1)$. We use $P(H_1) = 10^{-5}$ (Table I); this value is a subjective estimate that roughly one release of *B. anthracis* will occur every 10 years in any one of the 30 largest cities in the United States (i.e. approximately $1/(10 \times 365 \times 30)$). Other researchers have used a daily probability of approximately 2.7×10^{-5} [34].

To allow decision makers to substitute an alternative prior probability for H_1, BARD also outputs the following likelihood ratio, or Bayes' Factor [35]:

$$\frac{P(\mathbf{b}|H_1, \mathbf{G}, \mathbf{M})}{P(\mathbf{b}|H_0, \mathbf{G})} \tag{3}$$

In the remainder of this section, we describe in more detail the input and output of the BARD algorithm and how it computes the conditional probabilities $P(\mathbf{b}|H_0, \mathbf{G})$ and $P(\mathbf{b}|H_1, \mathbf{G}, \mathbf{M})$.

3.1. Input and output

BARD requires three inputs: (1) the vector **b** of counts of ED visits for respiratory complaints; (2) the matrix **G** of data about geographical areas; and (3) the matrix **M** of meteorological data. The vector **b** comprises the most recent 24-h counts of ED visits for respiratory complaints for each zip code in the surveillance region.

DOI: 10.1002/sim

Table I. Notation and description of parameters in BARD's model.

Variable	Default value	Description
h		Height in meters above ground at which *B. anthracis* spores were released
Q		Quantity of anthrax spores released in kg
$\dot{V}_E$	2.23×10^{-4}	Minute ventilation (m^3/s)
x, y		x, y coordinate of release location on surface of the Earth*
u		Wind speed at time of release in m/s
s		Atmospheric stability class
t		Number of hours from time of release to the end of a 24-h interval for which we counted ED visits
d		Number of anthrax spores an individual inhales
ID_{50}	8625	Dose of spores infectious for 50 per cent of population
Probit slope	0.669	Slope of linear relationship between probit and $\log_{10}(d)$
zip		Home zip code
moy		Month of year of the 24-h interval of counts of ED visits
dow		Day of week of the 24-h interval of counts of ED visits†
θ_1^+		P (individual presents to ED with respiratory chief complaint between $t-1$ and $t \mid d, t$)
θ_0		P (individual presents with respiratory chief complaint $\mid moy, dow$)
$P(H_1)$	10^{-5}	Prior probability of a release

*One must convert from longitude and latitude to meters using the map projection functions of geographical information system software.

†We use the day of week in which the majority of the 24-h period resides. For example, if the 24-h interval ends at 12:00 PM on Thursday, we use dow = 5 (Thursday), but if it ends at 8:00 AM on Thursday, we use dow = 4 (Wednesday).

Similarly each zip code has one row in matrix **G** with five columns of data. The five columns are (1) population, (2) the x coordinate of a central point in the zip code, (3) the y coordinate of the same point, and the historical (4) mean and (5) standard deviation of counts of ED visits for the zip code and the month of year and day of week that correspond to the date/time of the counts of ED visits in **b** (we discuss these columns further in Section 3.2.1).

The rows of matrix **M** correspond to times at which meteorological observations were made and the columns to the particular variables observed. Specifically, the columns are wind speed, wind direction, and atmospheric stability class. Atmospheric stability class is a measurement of atmospheric turbulence, a key determinant of atmospheric dispersion of substances [36].

In addition to the posterior probability $P(H_1 \mid \mathbf{b}, \mathbf{G}, \mathbf{M})$, BARD outputs (1) the posterior distribution over the release location, quantity, and date/time conditional on a release having occurred and the expectation of these quantities, (2) the zip code whose boundary contains the point corresponding to the posterior expectation of the release location, and (3) a listing of the zip codes in descending order of the probability that some individuals in the zip code are infected.

3.2. Computation of the conditional probabilities

BARD assumes that the counts of ED visits in the zip codes, $\mathbf{b} = [c_1, c_2, \ldots, c_m]'$, are independent given either hypothesis (relaxing this assumption is future work); thus, if there are m zip codes in

DOI: 10.1002/sim

the region BARD is monitoring:

$$P(\mathbf{b}|H_0,\mathbf{G}) = P(c_1,c_2,\ldots,c_m|H_0,\mathbf{G}) = \prod_{i=1}^{m} P(c_i|H_0,\mathbf{g}_i) \tag{4a}$$

$$P(\mathbf{b}|H_1,\mathbf{G},\mathbf{M}) = P(c_1,c_2,\ldots,c_m|H_1,\mathbf{G},\mathbf{M}) = \int_{\mathbf{r}} \left[\prod_{i=1}^{m} P(c_i|H_1,\mathbf{g}_i,\mathbf{m}_t,\mathbf{r})\right] P(\mathbf{r}|H_1,\mathbf{g}_i,\mathbf{M})\,\mathrm{d}\mathbf{r} \tag{4b}$$

where $\mathbf{g}_i$ is the row vector of $\mathbf{G}$ for zip code i, $\mathbf{r}$ is a vector of values for a particular location and quantity and time of release, and $\mathbf{m}_t$ is the row vector of $\mathbf{M}$ that specifies the meteorological data corresponding to the hypothesized time of release in $\mathbf{r}$. We discuss $\mathbf{r}$, its dependency relationship with c_i, its prior distribution, and the procedure we use to integrate over it to compute $P(\mathbf{b}|H_1,\mathbf{G},\mathbf{M})$ in Section 3.2.2.

BARD represents each c_i as a binomial process:

$$P(c_i|H_0,\mathbf{g}_i) = \int_0^1 \binom{n_i}{c_i} \theta_{0,i}^{c_i}(1-\theta_{0,i})^{n_i-c_i} f_0(\theta_{0,i}|\mathbf{g}_i)\,\mathrm{d}\theta_{0,i} \tag{5a}$$

$$P(c_i|H_1,\mathbf{g}_i,\mathbf{m}_t,\mathbf{r}) = \int_0^1 \binom{n_i}{c_i} \theta_{1,i}^{c_i}(1-\theta_{1,i})^{n_i-c_i} f_1(\theta_{1,i}|\mathbf{g}_i,\mathbf{m}_t,\mathbf{r})\,\mathrm{d}\theta_{1,i} \tag{5b}$$

where $\theta_{j,i}$ is the probability under hypothesis H_j that an individual in zip code i will visit an ED with a respiratory complaint in a given 24-h period, $f_j(\cdot)$ is the density function over $\theta_{j,i}$, and n_i is the population of zip code i. Although researchers more typically represent count data using the Poisson distribution, we assume that the number of ED visits from a zip code is bounded by the number of people who live there (the ED data are given by patient home zip code); thus, the binomial distribution is appropriate. Nevertheless, given that typically we will have small values of $\theta_{j,i}$ and c_i, and large values of n_i, our use of the binomial typically will closely approximate the Poisson.

We discuss next how BARD estimates $f_0(\theta_{0,i}|\mathbf{g}_i)$ and $f_1(\theta_{1,i}|\mathbf{g}_i,\mathbf{m}_t,\mathbf{r})$ and uses them to compute the conditional probabilities $P(c_1,c_2,\ldots,c_m|H_0,\mathbf{G})$ and $P(c_1,c_2,\ldots,c_m|H_1,\mathbf{G},\mathbf{M})$.

3.2.1. Computing $P(c_1,c_2,\ldots,c_m|H_0,\mathbf{G})$. We condition $\theta_{0,i}$ on month of year and day of week (*moy* and *dow*, respectively, in Table I) to account for seasonal and day of week effects in the counts of ED visits for respiratory complaints that we have observed historically. We assume that no releases of *B. anthracis* occurred during the time period covered by the historical data that we use to parameterize $f_0(\theta_{0,i}|\mathbf{g}_i)$. Specifically, for each zip code i, we use a beta probability density function

$$\theta_{0,i} \sim \mathrm{beta}(\hat{\alpha}_i,\hat{\beta}_i) \tag{6}$$

where $\hat{\alpha}_i$ and $\hat{\beta}_i$ are parameters of the beta distribution, which we estimate from historical data as discussed below.

DOI: 10.1002/sim

The value of $P(c_i|H_0,\mathbf{g}_i)$ is then a beta binomial probability:

$$P(c_i|H_0,\mathbf{g}_i)=\binom{n_i}{c_i}\frac{\Gamma(\hat{\alpha}_i+\hat{\beta}_i)}{\Gamma(\hat{\alpha}_i)\Gamma(\hat{\beta}_i)}\frac{\Gamma(\hat{\alpha}_i+c_i)\Gamma(\hat{\beta}_i+n_i-c_i)}{\Gamma(\hat{\alpha}_i+\hat{\beta}_i+n_i)} \tag{7}$$

and we compute $P(c_1,c_2,\ldots,c_m|H_0,\mathbf{G})$ as the product of the $P(c_i|H_0,\mathbf{g}_i)$ as in equation (4a).

We use a method of moments approach to compute the estimates $\hat{\alpha}_i$ and $\hat{\beta}_i$ from historical data. For example, if zip code 15219 on Mondays ($dow=2$) in March ($moy=3$) had the following historical counts: 5, 7, 13, 7, 8, 5, 12, 8, then the sample mean ($\overline{p}$) and variance ($\hat{\sigma}^2$) of the proportion (of the residents of the zip code) that visited the ED for a respiratory complaint are $8.125/n$ and $8.70/n^2$, respectively, where n is the population of the zip code. Zip code 15219 has a population of 19 531 according to the United States census of 2000, leading to $\overline{p}=4.16\times10^{-4}$ and $\hat{\sigma}^2=2.28\times10^{-8}$. These values for $\overline{p}$ and $\hat{\sigma}^2$ are in the row of $\mathbf{G}$ that corresponds to zip code 15219.

We then derive $\hat{\alpha}_{15219}$ and $\hat{\beta}_{15219}$ using the following moment-matching equations [37]:

$$\hat{\alpha}+\hat{\beta}=\frac{\overline{p}\cdot(1-\overline{p})}{\hat{\sigma}^2}-1$$

$$\hat{\alpha}=\overline{p}\cdot(\hat{\alpha}+\hat{\beta})$$

$$\hat{\beta}=(1-\overline{p})\cdot(\hat{\alpha}+\hat{\beta})$$

In this example, $\hat{\alpha}_{15219}=7.59$ and $\hat{\beta}_{15219}=18\,231.47$.

For combinations of zip code, month of year, and day of week for which the mean of the historical data was zero, we chose values for $\overline{p}$ and $\hat{\sigma}^2$ such that the probability of observing more than one ED visit (under H_0) was approximately 0.025. The reason for this arbitrary choice was that we wanted to minimize the effect on the likelihood of H_0 of a single visit from such zip codes to avoid false alarms due to occasional ED visits due to background respiratory disease, but we wanted two visits from such zip codes to have a more substantial effect. It turns out that—for all zip-code populations ranging from 100 to 150 000—$\overline{p}\approx0.125/n$ and $\hat{\sigma}^2=\overline{p}(1-\overline{p})/n$ result in a probability of observing two or more visits of approximately 0.025. To study the sensitivity of BARD to this choice, we also tried $\overline{p}=1/n\,(P(>1\text{ visit})=0.25)$ and $\overline{p}=10^{-6}/n\,(P(>1\text{ visit})=2\times10^{-7})$ in the evaluation we discuss below.

3.2.2. Computing $P(c_1,c_2,\ldots,c_m|H_1,\mathbf{G},\mathbf{M})$. Because H_1 is the hypothesis that both background respiratory disease and an outbreak of inhalational anthrax are occurring, we condition $\theta_{1,i}$ on these two causes of individuals visiting EDs for respiratory complaints. The probability that an individual presents to an ED with a respiratory complaint given that he or she has background respiratory disease is $\theta_{0,i}$ as above. We denote the probability that an individual presents to an ED with a respiratory chief complaint given that he or she has inhalational anthrax as $\theta^+_{1,i}$.

We assume that these causes are independent; that is, a person with inhalational anthrax will not be less or more likely to present with background respiratory disease or *vice versa*. Thus, ignoring for now the density function over θ_0, we can compute θ_1 from θ_0 and θ^+_1 as [38]

$$\theta_{1,i}=1-(1-\theta_{0,i})(1-\theta^+_{1,i}) \tag{8}$$

DOI: 10.1002/sim

Following the method used by other researchers [4, 39], we model $\theta^{+}_{1,i}$ as a function of the number of spores that an individual inhales—d—and the number of units of time that has elapsed since he or she inhaled the spores—t. BARD derives d from the Gaussian plume model of dispersion and an estimate of minute ventilation (the volume of air that an individual breathes per minute).

The Gaussian plume model of atmospheric dispersion computes the time-integrated concentration at an arbitrary downwind location due to a near-instantaneous release of substance:

$$\mathrm{TIC} = \frac{Q}{2\pi\sigma_y(s, x-x_i)\sigma_z(s, x-x_i)u} \mathrm{e}^{-(y-y_i)^2/2\sigma_y(s,x-x_i)^2} \times [\mathrm{e}^{-(h_i-h)^2/2\sigma_z(s,x-x_i)^2} + \mathrm{e}^{-(h_i+h)^2/2\sigma_z(s,x-x_i)^2}] \tag{9}$$

where TIC is the time-integrated concentration in units of mass × time/volume; Q is the mass of substance released; (x, y, h) is the coordinate[‡] of the release location where x and y specify the location on the surface of the earth and h specifies height above ground; (x_i, y_i, h_i) is similarly the coordinate of the location for which we are computing TIC (i.e. the x, y coordinate from row i of $\mathbf{G}$, and we assume $h_i = 0$); σ_y and σ_z (in units of length) are the distributions of spores in the crosswind direction (parallel to ground and perpendicular to wind direction) and vertical direction (perpendicular both to ground and wind direction), respectively, as a function of downwind distance from the release location (i.e. x–x_i) and atmospheric turbulence, which we denote by s to indicate the fact that we use stability classes as a measurement of turbulence; and u is the wind speed (in units of length/time). We note that Meselson *et al.* [23] and other researchers [4, 39] have used the Gaussian plume model in the same manner for modeling aerosol releases of *B. anthracis*.

The variables x, y, h, Q, and t constitute the vector $\mathbf{r}$ of release parameters. We will assume they are fixed for the moment, but they are unknown. We discuss below how we integrate over them to derive the posterior probability of a wind-borne outbreak.

We specify Q in terms of the number of spores released—and thus TIC has units of spores × time/volume—and multiply TIC by minute ventilation or $\dot{V}_E$ to compute the total dose of spores inhaled at the central point of each zip code:

$$d = \mathrm{TIC} \cdot V_E \tag{10}$$

We use a value of 13.38 l/min for minute ventilation (see Appendix A.1 for derivation and rationale). To match the units of time and distance we use for variables in the Gaussian plume model, the appropriate value is $2.23 \times 10^{-4}\,\mathrm{m}^3/\mathrm{s}$ (Table I). This estimate for minute ventilation is lower than the value other researchers use; thus with our model individuals will inhale fewer spores than with other models, all other things being equal.

The next step in computing $\theta^{+}_{1,i}$ is to estimate the probability that an individual who inhaled d spores visits an ED with a respiratory complaint. We denote this probability as $P(\text{visit ED}|d)$.

We use the fraction of people infected as an approximation of $P(\text{visit ED}|d)$. Our rationale was that during the 2001 mail-borne anthrax attack, at least nine out of the first 10 victims visited an ED [40]. We did not adjust for the fact that not all victims will have a respiratory chief complaint or that classifiers of chief complaints are imperfect and thus they fail to properly classify some

[‡]For analytic convenience, we assume that all coordinates are in meters and are in a coordinate system that has its x-axis pointing in the direction to which the wind is blowing.

DOI: 10.1002/sim

individuals. The net effect is that we use an upper bound for $P(\text{visit ED}|d)$. We plan to relax these assumptions in future work.

Data about the lethal dose of anthrax spores (we adjust what follows to an infectious dose below) in monkeys have shown there is a linear relationship between probits and $\log_{10}(d)$ [41]. A probit is the inverse cumulative density function of the standard-normal distribution computed at the fraction of exposed individuals who die: that is, $\text{probit} = \Phi^{-1}$ (fraction killed). Glassman derived a probit slope of 0.669, and using his estimate of a median lethal dose (LD_{50}, probit = 0) of 4130 inhaled spores, we obtain

$$\text{probit} = 0.669 \log_{10}(d) - 2.42 \tag{11}$$

In a recent paper, Wilkening found that three models of anthrax—which he labeled as $A1$, $A2$, and D—provided the best fit to Sverdlovsk and animal model data [7]. Of those models, we chose values for probit slope and ID50 (dose infectious for 50 per cent of the population) from model $A1$. Our rationale was that it was the only model that had a linear relationship between median incubation period and $\log_{10}(d)$, which is consistent with the relationship seen with at least one other disease caused by inhalation of pathogens—Q fever [42]. In model $A1$, ID50 = 8625 spores and probit slope = 0.67. Changing the ID_{50} only changes the y-intercept of equation (11); for 8625 spores, the y-intercept is -2.64. Anyone using BARD may vary the values of ID_{50} and probit slope.

BARD thus uses the following equation:

$$P(\text{visit ED}|d) = \text{fraction infected}|d = \Phi(0.669 \log_{10}(d) - 2.64) \tag{12}$$

The next step in computing $\theta^{+}_{1,i}$ is to estimate the probability that an individual who visits the ED will do so during a 24-h interval $t-1$ to t days after inhaling spores. This computation requires that we estimate two time intervals: the incubation period (exposure to symptom onset) and visit delay (symptom onset to presentation to an ED). In Wilkening's $A1$ model of anthrax, the parameters of the lognormal distribution for incubation period vary with $\log_{10}(d)$. We have previously described how we combined incubation period parameters from model $A1$ and data about visit delay in inhalational anthrax into a new set of functions for computing—as a function of $\log_{10}(d)$—parameters for a lognormal distribution that closely approximates the sum of the incubation period and visit delay [43]. This method models visit delay as a lognormal distribution and approximates the sum of the lognormal distributions for incubation period and visit delay as a single lognormal distribution (using the method of Wu [44]).

BARD then estimates the probability of presenting in $(t-1, t]$ as

$$P(\text{visit ED during } (t-1,t]|\text{visit ED}, d) = F(t|\mu_d, \sigma_d) - F(t-1|\mu_d, \sigma_d) \tag{13}$$

where $F(\cdot)$ is the cumulative density function of the lognormal distribution that approximates the sum of incubation period and visit delay, and μ_d, σ_d are calculated as a function of $\log_{10}(d)$ as per Hogan and Wallstrom [43].

We set the interval to $(t-1, t]$ or 1 day because biosurveillance systems typically make counts of ED visits (and other data) available as 24-h counts and because it simplifies the model to exclude time-of-day effects in the ED visit data.[§]

[§] Taking into account time-of-day effects may lead to earlier detection, and we plan to explore this possibility in future research.

DOI: 10.1002/sim

We assume that the time of release and time that individuals inhale spores are the same, although they may differ by as much as a few hours. Given that the variances of the lognormal distributions are on the order of >1 day, this difference is not likely to be of much consequence.

Finally, BARD computes $\theta^{+}_{1,i}$ as

$$\theta^{+}_{1,i} = P(\text{visit ED during } (t-1,t] | \text{visit ED}, d) \cdot P(\text{visit ED} | d) \tag{14}$$

As stated before, this computation of $\theta^{+}_{1,i}$ requires particular values of the release parameters $\mathbf{r}=[x, y, h, Q, t]$; thus by extension, $\theta_{1,i}$ is also dependent on $\mathbf{r}$. Therefore, we first discuss how we compute each $P(c_i | H_1, \mathbf{g}_i, \mathbf{m}_t, \mathbf{r})$, where $\mathbf{m}_t$ is the row vector of $\mathbf{M}$ corresponding to the value of t in $\mathbf{r}$, and then how we integrate over $\mathbf{r}$ to obtain $P(c_1, c_2, \ldots, c_m | H_1, \mathbf{G}, \mathbf{M})$.

We do not directly derive the expression for $f_1(\theta_{1,i} | \mathbf{g}_i, \mathbf{m}_t, \mathbf{r})$, but instead substitute the right-hand side of equation (8) for $\theta_{1,i}$ in equation (5b):

$$P(c_i | H_1, \mathbf{g}_i, \mathbf{m}_t, \mathbf{r}) = \int_0^1 \binom{n_i}{c_i} (1-(1-\theta^{+}_{1,i})(1-\theta_{0,i}))^{c_i} \times ((1-\theta^{+}_{1,i})(1-\theta_{0,i}))^{n_i - c_i} \operatorname{beta}(\theta_{0,i} | \hat{\alpha}_i, , \hat{\beta}_i) \, \mathrm{d}\theta_{0,i} \tag{15}$$

This integral has the following closed-form solution, with the i subscript omitted for notational convenience (see Appendix A.2 for proof):

$$P(c | H_1, \mathbf{g}, \mathbf{m}_t, \mathbf{r}) = \binom{n}{c} (1-\theta^{+}_{1})^{n-c} \frac{\Gamma(\hat{\alpha}+\hat{\beta})\Gamma(\hat{\beta}+n-c)}{\Gamma(\hat{\alpha})\Gamma(\hat{\beta})} \sum_{q=0}^{c} \binom{c}{q} (\theta^{+}_{1})^{c-q} (1-\theta^{+}_{1})^{q} \times \frac{\Gamma(\hat{\alpha}+q)}{\Gamma(\hat{\alpha}+q+\hat{\beta}+n-c)} \tag{16}$$

Because the release parameters $\mathbf{r}=[x, y, h, Q, t]$ are unknown, BARD integrates over these variables to obtain $P(c_1, c_2, \ldots, c_m | H_1, \mathbf{G}, \mathbf{M})$ as shown in equation (4b).

Equation (5b) requires that we assign a prior distribution to the vector of release parameters $\mathbf{r}$. We assume that the release parameters are conditionally independent, given H_1 and independent of $\mathbf{G}$ and $\mathbf{m}_t$ to obtain

$$P(\mathbf{r} | H_1, \mathbf{G}, \mathbf{m}_t) = P(x, y | H_1) \cdot P(h | H_1) \cdot P(Q | H_1) \cdot P(t | H_1) \tag{17}$$

We assigned a uniform prior distribution to all but $P(h|H_1)$, for which we used a probability function that decreases as values for h increase (see Appendix A.3).

BARD uses a Monte Carlo integration method called *likelihood weighting* [45, 46] to numerically solve equation (5b) (Appendix A.4), because it is not solvable in closed-form and simple iterative numerical methods can be computationally intractable. A desired by-product of this procedure is that—with minimal additional computational effort—we can also estimate a posterior distribution for the elements of $\mathbf{r}$. The expectations of these posterior distributions are the characterization of the release location, quantity, and date/time that BARD provides at present. In the future, we also plan to derive additional features, such as credible intervals, of these distributions as characterizations.

DOI: 10.1002/sim

4. A PROOF-OF-CONCEPT EVALUATION OF BARD

To determine whether BARD shows promise as a method for integrated analysis of data for early detection, we conducted a proof-of-concept evaluation. Specifically, we measured its ability to detect aerosol releases created by a simulator that uses the same assumptions as BARD. We also conducted one-way sensitivity analyses of BARD's performance over several key parameters in the model to determine the effect of varying the values for these parameters. If BARD has acceptable outbreak-detection performance and its performance is robust to variations in its key model parameters, then additional development and evaluation of BARD would be warranted.

4.1. Methods

We measured BARD's performance for the detection of simulated aerosol releases. We added ED visits from simulated outbreaks to actual, historical ED visit data, a common approach to the evaluation of outbreak-detection algorithms [19, 22, 47–49]. The Institutional Review Board of the University of Pittsburgh approved this study.

4.1.1. Setting. We evaluated BARD using ED, meteorological, and zip code data from the six-county Pittsburgh Metropolitan Statistical Area (MSA), consisting of Allegheny, Beaver, Butler, Fayette, Washington, and Westmoreland counties.

4.1.2. Data sets. The historical ED data we used in this study were actual ED visits to 10 EDs operated by one health system. These EDs represent approximately 30 per cent of all ED visits for the Pittsburgh MSA.

We divided the actual historical ED visit data into training and test sets. The training set—which we used to parameterize BARD's beta distributions for θ_0—spanned a four-year time period from 1 January 1999 to 31 December 2002. The test set—into which we injected simulated releases—spanned three years from 1 January 2003 to 31 December 2005.

We created ED outbreak data for 225 simulated releases of *B. anthracis* according to the procedure we describe in Appendix A.5. Specifically, we selected 25 random release locations (x, y, h) and dates/times. We then simulated nine releases from each of the 25 sets of locations and dates using different values for various parameters: three different release quantities (values of Q), two different values of ID50, two different values of probit slope, and two different y-intercepts and slopes for the linear relationship between median incubation period and $\log_{10}(d)$.

We varied the parameter values used for simulation as opposed to the values for detection because we wanted to evaluate the effect of a discrepancy between actual outbreak parameters and the typical values we use operationally for detection. Specifically, to measure the effect of the quantity of release—and by extension size of outbreak—on BARD's performance, we used $Q = 10^{12}, 10^{13}$, and 10^{14}; values that overlap the range of quantities of spores assumed by other researchers [4, 50, 51]. If one assumes 10^{14} spores/kg (approximately the value for spores mailed to Senator Tom Daschle [52]), then these values are 1, 0.1, and 0.01 kg released, respectively.

For the other parameters, we chose the minimum and maximum extremes of values to maximally stress the assumptions of the model. We used a fixed release quantity of 10^{13} spores for these analyses. For ID50, we used values of 2000 and 55 000 spores [53]; for probit slope, 0.3 (less than one half of our baseline value) and 1.43 [51]; and for y-intercept and slope of $\log_{10}(d)$ *vs* median incubation, 6.33 and -0.725 days and 12 and -1.618 days, respectively. The former choice for

DOI: 10.1002/sim

the relationship between $\log_{10}(d)$ and median incubation is from Wilkening's model B, which Wilkening rejected as not being consistent with Sverdlovsk data [7], but which Rickmeier *et al.* derived from expert consensus [51]. The latter choice for the relationship between $\log_{10}(d)$ and median incubation period we derived by choosing a median of 12 days at the lowest value of d (one spore)—1 day longer than the median incubation estimated by Brookmeyer [54])—and then choosing the slope such that the linear relationship between $\log_{10}(d)$ and median incubation period intersects the same point at 10^6 spores as does the line for models $A1$ and B.

Because the market share of the historical data was 30 per cent, we ensured that only 30 per cent of cases from the simulations were included in the data input into BARD.

The meteorological data we used are actual data from the National Weather Service, which included wind speed and wind direction, but not stability class. We computed stability class with the available data using Turner's method [55]. It is not a requirement to use these meteorological data with BARD.

The populations and central zip code points that we used in this study are from the ESRI® ArcGIS™ Desktop product.

4.1.3. Measurements. We measured outbreak-detection performance with standard metrics: false-alarm rate, sensitivity, and mean time to detection. We defined a false alarm as a posterior probability that exceeds the alarm threshold in the absence of an outbreak. The false-alarm rate is then the number of false alarms that occur per unit of time. To estimate the false-alarm rate, we ran BARD on the test set with no simulated ED visits added. We computed the sensitivity of outbreak detection as the percentage of outbreaks during which the posterior probability exceeds the alarm threshold at least once. If BARD did not detect the outbreak before day 7, then we counted that outbreak as a false negative. Time to detection is the duration of the interval between the time of the release and the time the posterior probability first exceeds the alarm threshold.

We also measured time to detection as the duration of the interval between the time of the first simulated ED visit and the time of detection because it allows a comparison of BARD's time to detection with estimates of the time to detection of the notifiable disease reporting system. The latter are derived from case reports in the literature about the interval from the time of presentation for healthcare to the time of a confirmed diagnosis. One advantage of using this interval is that it is not dependent on assumptions about the incubation period used in simulations.

Throughout the evaluation we ran BARD as if it were running every 4 h on the production real-time outbreak and disease surveillance (RODS) system using counts of ED visits from the previous 24 h. The reason is that at the time of this study, most outbreak-detection algorithms in RODS evaluate data with this frequency, including BARD. Running BARD every 4 h on a real-time system such as RODS has the potential to improve the timeliness of detection relative to running just once every 24 h on a daily batch-feed system. However, running BARD multiple times per day also has the potential to increase BARD's false-alarm rate. Here we did not study alternative frequencies of running BARD on surveillance data.

Because there is a fundamental trade-off between the false-alarm rate and time to detection, it is common when evaluating outbreak-detection algorithms to perform an activity monitoring operating characteristic or AMOC analysis [19, 22, 29, 47, 49, 56–59]. AMOC analysis (first described by Fawcett and Provost [60]) resembles ROC analysis, but plots the false-alarm rate *vs* time to detection. It allows users of biosurveillance systems to set the alarm threshold based on the trade-off between the costs of false alarms and the benefits of earlier detection. In this study, we plotted

DOI: 10.1002/sim

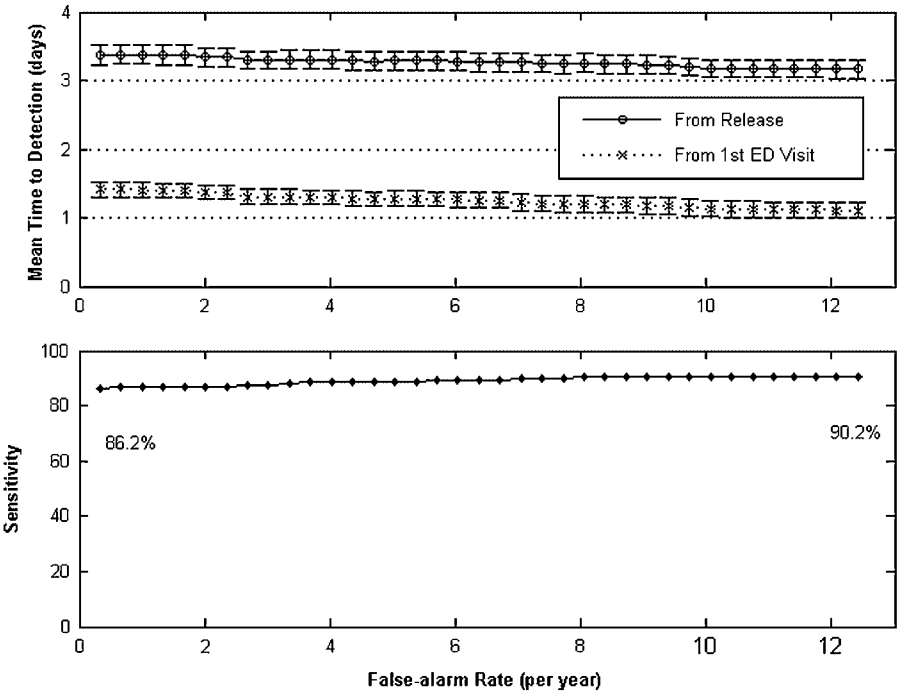

Figure 2. Mean time to detection (measured both from release date/time and from date/time of first anthrax case that visited the ED) and sensitivity *versus* false alarm rate, illustrating BARD's performance over a range of alarm thresholds (error bars are 95 per cent CI on mean).

false-alarm rate *vs* mean time to detection and sensitivity. We constructed AMOC plots for each sensitivity analysis.

We also analyzed BARD's performance as a function of outbreak size, which we measured as the total number of cases produced by the simulator. We plotted AMOC curves by quintiles of outbreak size (i.e. 0–20th percentile, 20th–40th percentile, etc.) Note that because the market share of the historical data was approximately 30 per cent, only approximately 30 per cent of the cases were available to BARD for detection. We also determined the approximate size of the smallest outbreak BARD could detect with 100 per cent sensitivity at a false-alarm rate of one per three years (the lowest false-alarm rate we could measure with our test data set) by manually reviewing the data about BARD's performance as a function of outbreak size.

To determine whether BARD could detect simulated releases before they become obvious on univariate time series, we made several measurements. First, we computed the mean number of cases who visited the ED in the 24 h prior to detection (only those cases influence BARD's posterior probability at present). We also computed the mean and standard deviation of 24-h counts in the test set of historical data. Finally, we standardized the counts of ED visits at detection relative to the mean and standard deviation of counts in the month prior to the release and computed the mean number of standardized counts at detection.

We measured the mean running time of BARD to determine whether its inference is sufficiently fast that it is practicable to include it in a functioning biosurveillance system. We ran BARD on

DOI: 10.1002/sim

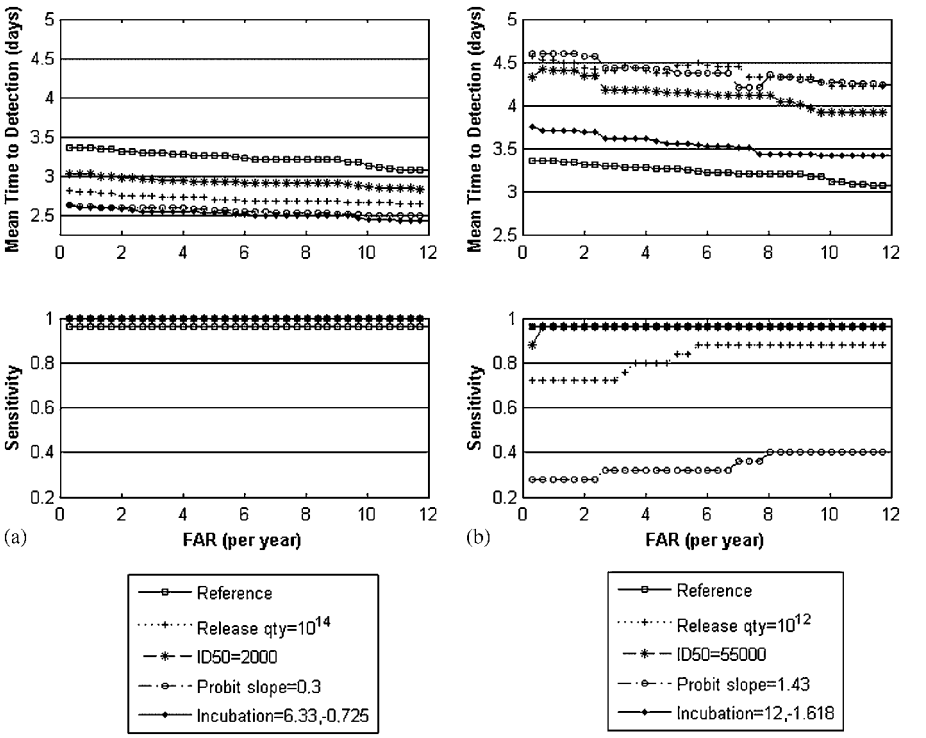

Figure 3. Results of sensitivity analyses over parameters used in simulation. Variations in parameter values that increased outbreak size (a) and variations in parameter values that decreased outbreak size (b). In both plots, BARD's performance for simulations using the baseline parameter values with a release quantity of 10^{13} spores (Reference) is designated with square markers. FAR, false-alarm rate.

a dual-processor server from Dell that cost less than $3500. BARD is a single-threaded Java™ application and thus did not take advantage of the additional processor.

4.2. Results

Of the 225 simulated outbreaks, 224 ranged in size from 8 to 168 190 cases of inhalational anthrax. One simulation with a probit slope of 1.43 failed to generate any case, but we did not exclude it from further analysis. The mean outbreak size was 19 386 cases, with a standard deviation of 27 623. The median outbreak size was 8316 cases; the 90th percentile was just under 51 000. Because the historical data had a market share of approximately 30 per cent, we limited the number of simulated cases available to BARD to approximately 30 per cent of these numbers.

At a sensitivity of 86.2 per cent and a false-alarm rate of one per three years, BARD's mean time to detection was 3.37 days (95 per cent CI 3.23–3.51) (Figure 2). The alarm threshold that produced these results was a posterior probability of 0.0014. As the false-alarm rate increased, time to detection decreased (Figure 2): at a false-alarm rate of 12 per year (or 1 per month), time to detection was 3.16 days (95 per cent CI 3.04–3.29) and sensitivity increased to 90.2 per cent. When measured as the duration of the interval from the time that the first case presented to an ED

DOI: 10.1002/sim

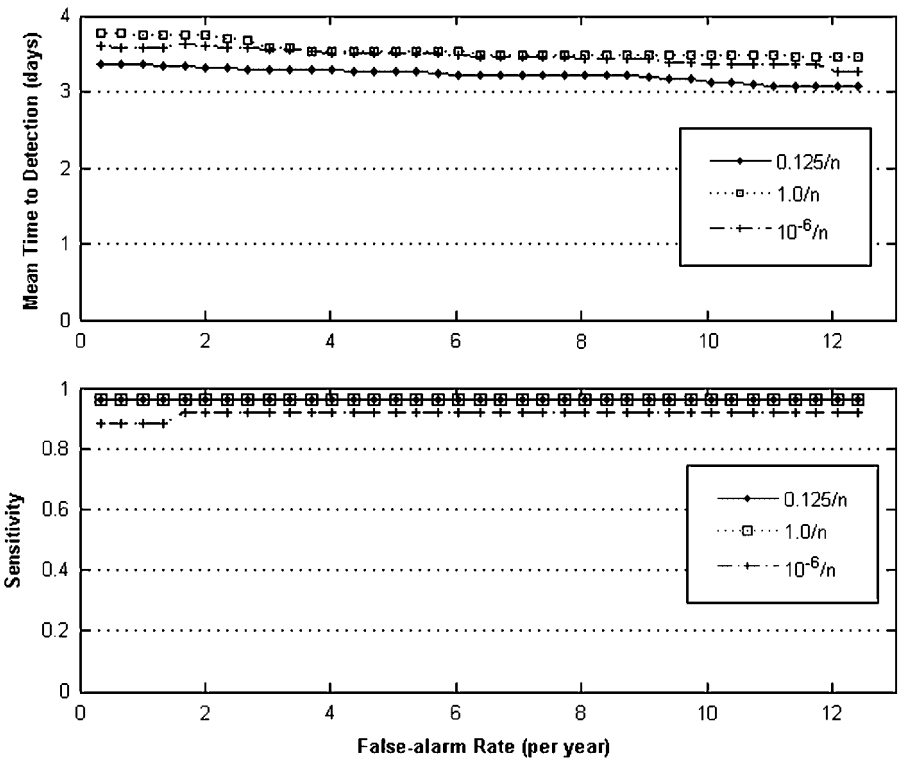

Figure 4. Results of sensitivity analysis over default value for $\overline{p}$.

to time of detection, time to detection varied with false-alarm rate from 1.11 days (95 per cent CI 0.99–1.22) at a false-alarm rate of one per month to 1.40 days (95 per cent CI 1.31–1.52) at a false-alarm rate of one per three years (Figure 2).

In the sensitivity analyses, variations of parameters in the simulation that increased outbreak size or shortened the incubation period resulted in improved outbreak-detection performance (Figure 3(a)). Similarly, variations of parameters that decreased outbreak size or lengthened the incubation period worsened outbreak-detection performance (Figure 3(b)). In no case did a parameter change that increased outbreak size worsen outbreak-detection performance; neither did a parameter change that decreased outbreak size improve outbreak-detection performance. The parameter changes that led to the largest improvement in outbreak-detection performance were shortening the median incubation period and lowering probit slope (Figure 3(a)). The parameter changes that led to the largest reduction in outbreak-detection performance were increasing probit slope and lowering the release quantity to 10^{12} spores (Figure 3(b)).

Varying the default value for $\overline{p}$ in the model of background respiratory disease (when the mean and/or variance of historical counts were zero) led to later and/or less sensitive outbreak detection at all measured false-alarm rates (Figure 4). This result held for both values tried: $1/n$ and $10^{-6}/n$.

In the analysis of BARD's outbreak-detection performance by outbreak size, BARD's performance improved with each quintile at all measured false-alarm rates (Figure 5(a)). We found that

DOI: 10.1002/sim

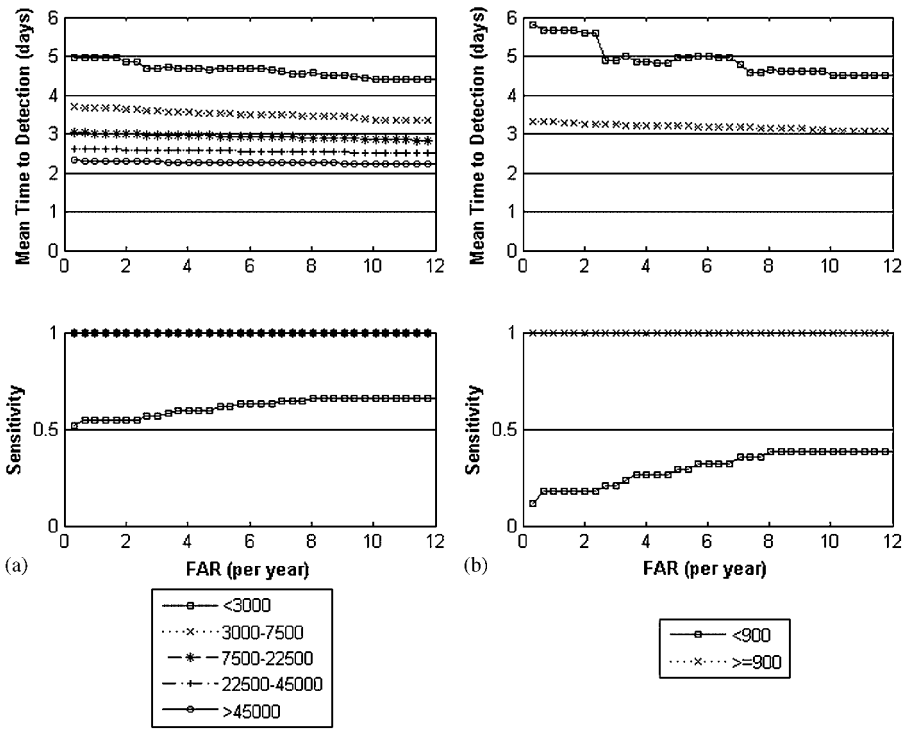

Figure 5. Analysis of BARD's outbreak detection performance by outbreak size. Outbreak detection performance by quintiles of outbreak size (a) and by threshold value of 900 cases (b). FAR, false-alarm rate.

above an outbreak size of 900 cases, BARD detected 100 per cent of outbreaks at a false-alarm rate of one per three years by 3.32 days on average (Figure 5(a)). Conversely, BARD detected 12 per cent of outbreaks of size $\leqslant$900 cases at a false-alarm rate of one per three years by 5.79 days on average (Figure 5(b)).

At a false-alarm rate of one per three years, there was a mean of 22.1 simulated ED visits in the 24 h period prior to the time of detection. The mean count of 24-h ED visits in the historical data was 38.3, with a standard deviation of 9.9. The count of ED visits at detection was an average of 2.9 standard deviations (median 2.7) above the mean 24-h count of visits in the month prior to the release. The correlation between outbreak size and this number of standard deviations was 0.11.

BARD's mean running time—for one analysis of one set of 24-h ED visit counts for all 271 zip codes—was 2 min and 24 s on our server.

5. DISCUSSION

We developed an algorithm—the BARD—for the integrated analysis of surveillance and meteorological data for early detection and characterization of aerosol releases of *B. anthracis*. BARD is the first algorithm to integrate meteorological data and a model of atmospheric dispersion

DOI: 10.1002/sim

into the analysis of medical surveillance data. It computes the posterior probability of a release of *B. anthracis* from two types of data readily available to biosurveillance systems: counts of ED visits and meteorological observations. BARD also computes a posterior distribution over the release location, quantity, and time. The model is flexible and allows for the substitution of alternative dispersion models and for varying a number of parameters such as minute ventilation, the number of spores per kilogram, the lethal dose of spores, the incubation period of inhalational anthrax, and the dose–response relationship between spores inhaled and incubation period.

In our idealized 'proof-of-concept' evaluation, BARD detected relatively small simulated releases early with high sensitivity and a low false-alarm rate. At the lowest false-alarm rate we measured, one per three years, BARD detected half the outbreaks when the count of ED visits was just 2.7 or fewer standard deviations above the mean count in the preceding month.

BARD proved sensitive and timely for all but the smallest releases. If one assumes 10^{14} spores/kg, 200 of the 225 simulations used a release quantity of 0.1 kg or smaller. Furthermore, our reanalysis of the data by outbreak size showed that BARD could detect 100 per cent of outbreaks with >900 cases at a low false-alarm rate of one per three years. Because the market share of our data set was 30 per cent, only 30 per cent of the simulated cases (>300) were available to BARD. Given higher market share, BARD may be able to detect even smaller outbreaks reliably.

Besides small release quantities and the relatively low market share of our historical data set (which we matched in the simulations), other factors in our evaluation served to make the detection problem more challenging for BARD. First, we used a more appropriate estimate for minute ventilation than other researchers; thus, our simulated releases may have fewer cases per equivalent release quantity than other researchers. Second, the assumption of mutually exclusive and exhaustive hypotheses H_0 and H_1 led to higher posterior probabilities during influenza outbreaks when establishing the false-alarm rate, leading to a higher threshold for detection of simulated releases at equivalent false-alarm rates—and thus potentially leading to delayed detection.

On the other hand, several assumptions in our model made the detection task less challenging for BARD. The assumption that all individuals are exposed in their home zip code is strong and unlikely to hold in the event of an actual release. Thus, the spatial pattern of cases that BARD recognizes may have been stronger in our simulations than they would be in reality. Second, we assumed that every case of anthrax had a respiratory chief complaint, but many victims are likely to have non-respiratory chief complaints (generally, only one or two symptoms are listed in a chief complaint). Also, we assumed that a syndromic surveillance system will classify all respiratory chief complaints as respiratory when in fact the accuracy of classifying chief complaints is less than 100 per cent.

Even if the net effect of these assumptions were to reduce BARD's performance on 1 kg releases to that of 0.1 kg releases, BARD could still detect a 1 kg (or approx. 10^{14} spores) release by 3.36 days with a false-alarm rate of one per three years and sensitivity of 100 per cent (as opposed to 2.81 days at the same false-alarm rate and sensitivity, see Figure 3).

The sensitivity analyses demonstrated that BARD was more sensitive to outbreak size and incubation period than to extreme variations in parameters used in simulation or to the choice of default parameters in the model of background respiratory disease. BARD never performed worse when the parameter change increased outbreak size or shortened incubation period; it never performed better when the parameter change decreased outbreak size or lengthened incubation

DOI: 10.1002/sim

period. Note that the sensitivity analyses were designed to explore what would happen if an actual outbreak had a parameter value at the extreme values of what is considered possible. Thus, the sensitivity analyses were designed to maximally stress the BARD model.

One should not interpret our time-to-detection results as realistic estimates of when BARD would detect a release in the real world. Our study demonstrates that shorter incubation periods of disease lead to shorter detection times. Our time-to-detection results should, therefore, be considered in light of the lognormal distributions we used for the incubation period of inhalational anthrax in our simulated outbreaks. Even at a dose of one million inhaled spores, a time to detection of 3.37 days corresponds to a cumulative distribution function value of 0.15 on the lognormal incubation period; thus, detection occurred on average before 15 per cent of cases presented to the ED.

Additionally, our alternative definition of time to detection may assist with the interpretation of our results. It is likely that the first case of anthrax for which the healthcare system confirms the diagnosis will also be one of the first cases to present to the healthcare system. The history of the Sverdlovsk incident indicates that laboratory confirmation of anthrax was made 3 days after the first victim was admitted to the hospital [33]. During the anthrax attacks of 2001, the first case that astute clinicians diagnosed was the second case to present to the healthcare system, approximately 1 day after the first case presented [40]. These two events are likely the upper and lower bounds of the time to detection of the healthcare system for inhalational anthrax, as the first event was not suspected *a priori* but the latter event, coming on the heels of 11 September 2001, was anticipated to at least some degree. Our results, although somewhat idealized, compare favorably with the lower bound of 1 day seen in 2001, especially as the release quantity becomes larger. Additionally, whether either of these outbreaks was wind-borne was not known until a long time after the first cases were diagnosed.

As a recent case of inhalational anthrax in a New York City man [61] illustrates, a diagnosis of a single case of inhalational anthrax does not indicate a terrorist event, let alone wind-borne dissemination of spores. BARD's ability to characterize the outbreak as windborne at the time of detection is therefore potentially useful even if the healthcare system confirms a diagnosis prior to BARD sounding an alarm.

Future work on BARD includes several enhancements to its model. Incorporating a model of population mobility is perhaps the most critical. We also plan to relax the assumptions that every case has a respiratory chief complaint and that chief-complaint classification has perfect accuracy. Lastly, the Gaussian plume model makes some assumptions that may not hold, despite the success that Meselson *et al*. had when using it to characterize the Sverdlovsk outbreak [23]. These assumptions include (1) the terrain is flat; (2) wind speed and direction are constant over space and time; (3) the substance released behaves like a gas with the same density as air; (4) there is no settling or deposition of particles; (5) the substance that is dispersing does not react with air or sunlight to become inactive; and (6) the release occurs at a single point. As discussed by Meselson *et al.*, infectious *B. anthracis* spores are on the order of 5–10 μm in diameter and particles of this size behave like a gas with the same density as air [23]. In addition, any decay in the infectivity of spores over time is likely to be negligible [62]. Thus, assumptions (3) and (5) are reasonable. The other assumptions were similarly reasonable in the case of the Sverdlovsk outbreak, but in general they are strong and not likely to hold. Future work thus includes incorporating a more realistic dispersion model into BARD, especially ones that can handle varying meteorological conditions and releases that occur at multiple points and/or line releases (such as a release from a moving car or airplane).

DOI: 10.1002/sim

Extending BARD to relax the assumptions it makes is likely to increase the complexity of inference. However, the running time of BARD was <3 min and we did not employ advanced Markov chain Monte Carlo (MCMC) inference techniques. Use of MCMC could help maintain tractable and accurate inference as we extend the model.

Overall, our evaluation of BARD demonstrates that BARD is a promising approach to integrated analysis of medical surveillance and meteorological data and that it is suitable for further development and evaluation. Future work includes enhancements to its model, evaluating its ability to characterize as well as detect releases, and comparing its performance with that of other algorithms.

APPENDIX A

A.1. Estimation of minute ventilation

We derived our estimate for minute ventilation from data tabulated in Altman *et al*. [63] (see Table AI for calculation). We arrived at a value of 13.68 l/min, which is $2.23 \times 10^{-4}\,\mathrm{m}^3/\mathrm{s}$ when converted to match the units of distance and time used in the Gaussian plume model. We used our own estimates of the proportions of the population of particular genders who are engaged in particular levels of physical activity, which are shown in parentheses in Table AI. We did not explicitly break out activity level by time of day, but included time of day considerations when forming our estimates. Taking into account (1) children and infants and (2) census data about the proportion of the population that are male *vs* female and adult *vs* child *vs* infant lowered the value only slightly to $2.02 \times 10^{-4}\,\mathrm{m}^3/\mathrm{s}$ (calculation not shown). Given the uncertainty in both estimates, we kept the first one.

By contrast, Rickmeier *et al*. used a value of $2.5 \times 10^{-4}\,\mathrm{m}^3/\mathrm{s}$ [51] and Wein *et al*. [4] used a value of $5.0 \times 10^{-4}\,\mathrm{m}^3/\mathrm{s}$, the same minute ventilation that Meselson *et al*. used [23]. Meselson *et al*. used the minute ventilation of a man engaged in light physical work, in the context of a shop for the manufacture of ceramic pipes where several male victims were working at the time they were exposed. Since we are concerned in this work with the entire population (not just males), and given that (1) women and children on average have a lower minute ventilation than men at a particular level of physical activity [63] and (2) many people will not be performing physical work but will instead be at rest, we arrived at our lower estimate of minute ventilation.

Table AI. Calculation of minute ventilation.

	Activity level*		
Gender	Rest†	Light work†	Heavy work†
Male	7.43 (0.3)	28.6 (0.15)	42.9 (0.05)
Female	4.50 (0.3)	16.3 (0.15)	24.5 (0.05)

*Numbers in parentheses are our estimates of the proportion of the population of a particular gender and engaged in a particular level of activity.

†Expectation: 13.68. Values are in units of l/min.

DOI: 10.1002/sim

A.2. Proof of equation (16)

The solution to the integral begins with expanding the terms of the following polynomial factor of the integrand of equation (15) (in what follows, we drop the subscripts for the zip code, month of year, and day of week for notational convenience):

$$(1-(1-\theta_1^+)(1-\theta_0))^c \tag{A1}$$

We expanded this polynomial in this form as a binomial series, but doing so led to a numerically unstable solution to the integral that—when implemented—occasionally produced negative probabilities. This problem is likely due to the alternating positive and negative terms in the series. We therefore rewrote the polynomial as

$$((1-\theta_1^+)\theta_0+\theta_1^+)^c \tag{A2}$$

The binomial series expansion of this form of the polynomial is

$$\sum_{q=0}^{c}\binom{c}{q}(1-\theta_1^+)^q\theta_0^q(\theta_1^+)^{c-q} \tag{A3}$$

Substituting Expression (A3) for Expression (A1) where it appears in equation (15) and inserting the definition of the probability density function for the beta distribution yields

$$\int_0^1\binom{n}{c}((1-\theta_1^+)(1-\theta_0))^{n-c}\left[\sum_{q=0}^{c}\binom{c}{q}(1-\theta_1^+)^q\theta_0^q(\theta_1^+)^{c-q}\right]\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta_0^{\alpha-1}(1-\theta_0)^{\beta-1}\,\mathrm{d}\theta_0$$

After moving constants outside the integral, distributing the remaining terms into the sum, and collecting exponentiations with the same base, we obtain

$$\binom{n}{c}(1-\theta_1^+)^{n-c}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\int_0^1\sum_{q=0}^{c}\binom{c}{q}(\theta_1^+)^{c-q}(1-\theta_1^+)^q\theta_0^{\alpha+q-1}(1-\theta_0)^{\beta+n-c-1}\,\mathrm{d}\theta_0$$

Moving the integral inside the summation such that no constants remain inside the integral, we obtain

$$\binom{n}{c}(1-\theta_1^+)^{n-c}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\sum_{q=0}^{c}\binom{c}{q}(\theta_1^+)^{c-q}(1-\theta_1^+)^q\int_0^1\theta_0^{\alpha+q-1}(1-\theta_0)^{\beta+n-c-1}\,\mathrm{d}\theta_0$$

The solution to the integral is

$$\frac{\Gamma(\alpha+q)\Gamma(\beta+n-c)}{\Gamma(\alpha+q+\beta+n-c)}$$

After substituting the solution to the integral and moving a term not dependent on q outside the summation, we obtain equation (16):

$$P(c|H_1,\mathbf{g},\mathbf{m}_t,\mathbf{r})=\binom{n}{c}(1-\theta_1^+)^{n-c}\frac{\Gamma(\alpha+\beta)\Gamma(\beta+n-c)}{\Gamma(\alpha)\Gamma(\beta)}$$
$$\times\sum_{q=0}^{c}\binom{c}{q}(\theta_1^+)^{c-q}(1-\theta_1^+)^q\frac{\Gamma(\alpha+q)}{\Gamma(\alpha+q+\beta+n-c)}$$

DOI: 10.1002/sim

We did not represent our uncertainty over θ_1^+ in this work, but note that if we were to do so using a beta distribution, we would still have a closed-form solution to the integral, where (α_0, β_0) are the parameters of the beta distribution over θ_0 and (α_1, β_1) are the parameters of the beta distribution over θ_1^+:

$$\binom{n}{c} \frac{\Gamma(\alpha_0+\beta_0)\Gamma(\alpha_1+\beta_1)\Gamma(\beta_0+n-c)}{\Gamma(\alpha_0)\Gamma(\beta_0)\Gamma(\alpha_1)\Gamma(\beta_1)\Gamma(\alpha_1+\beta_1+n)} \sum_{q=0}^{c} \binom{c}{q} \frac{\Gamma(\alpha_0+q)\Gamma(\alpha_1+c-q)\Gamma(\beta_1+n-c+q)}{\Gamma(\alpha_0+q+\beta_0+n-c)}$$

A.3. Prior distributions on release parameters

Table AII shows the ranges of values we used for each of the five release variables, whether we represented the parameter as a discrete or continuous variable and whether the distribution was uniform. For h (the only variable for which we assigned a non-uniform prior), we assigned an informative prior by evaluating the following function for all values of h between 0 and 3350 m that are also multiples of 10 m and then normalizing the resulting values to 1.0:

$$\frac{1}{1+e^{h/200-6}}$$

We chose this function to generate a distribution function (df) of 0.5 at approximately 330 m, a df of 0.95 at approximately 1000 m, and a df of 0.999 at approximately 1650 m. Our rationale for this distribution is that higher release locations are less likely because fewer spores will reach ground level; thus, terrorists will prefer lower release heights to maximize morbidity and mortality. This distribution may still overestimate the probability of releases from higher than 100 m (or approximately 300 feet) above ground.

When creating simulated outbreaks, we compute a prior distribution for x, y that assigns higher probabilities to release locations that are more likely to result in higher numbers of infected individuals. Thus, this prior distribution favors release locations desirable for their impact in terms

Table AII. Prior probability distributions for release location, quantity, and time.

Variable	Discrete *vs* continuous	Minimum	Maximum	Uniform?
x (m)*	Continuous	$x_{\min} - 10000\,\text{m}^\dagger$	$x_{\max}$	Yes
y (m)‡	Continuous	$y_{\min}$	$y_{\max}$	Yes
Q (spores)	Discrete	10^6 spores	10^{15} spores	No
h (m)	Discrete (every 10 m)	0	3350	No
t (h)§	Discrete (every 4 h)	28	168	Yes

*The values for $x_{\min}$ and $x_{\max}$ are the minimum and maximum x-coordinates over all zip code centroids after rotating the x-axis so that it points in the direction of the wind.

†We include an additional 10 km of distance upwind of the minimum in case the release location is upwind of the region BARD is monitoring.

‡The values for $y_{\min}$ and $y_{\max}$ are the minimum and maximum y-coordinates over all zip code centroids after rotating the x-axis so that it points in the direction of the wind.

§The date/time of the release can be obtained by subtracting t hours from the date/time of the end of the 24-h interval for which the counts of ED visits were made.

DOI: 10.1002/sim

of number of individuals infected, but it does not account for other strategic factors such as the likelihood that an attack would go unnoticed if carried out from a particular location or ease of access (e.g. tops of buildings, highways, airports, etc). The procedure is as follows:

1. Divide the region into a 100 by 100 grid of potential x, y release coordinates.
2. Simulate multiple outbreaks from the center of each of the 10 000 grid cells at varying release heights.
3. For each grid cell center, sum the total number of ED visits generated by the outbreaks in step 2.
4. Normalize the counts for each center such that the sum over all centers equals one.

A.4. The likelihood weighting procedure

The likelihood weighting procedure for computing $P(c_1, c_2, \ldots, c_m | H_1, \mathbf{G}, \mathbf{M})$ is as follows:

1. Let sum := 0.
2. For $j := 1$–N do the following (each iteration over j is a *cycle*):
   (a) For variables with known values (e.g. *moy*, *dow*, zip) set them to their values.
   (b) For each variable with unknown values (i.e. x, y, Q, h, and t), sample a value randomly according to its prior probability distribution. Let $\mathbf{r}(j)$ collectively denote the sampled values $x(j), y(j), h(j), Q(j)$, and $t(j)$.
   (c) For u(wind speed), s (stability class), and wind direction, just use the row of $\mathbf{M}$ corresponding to the value for $t(j)$ obtained in step 1b. Let $\mathbf{m}_t(j)$ denote this data.
   (d) Let score := 1.
   (e) For each zip code $i = 1$–m:
      (i) compute d using equation (10);
      (ii) compute $\theta^+_{1,i}$ from d and $t(j)$ using equation (14);
      (iii) compute $P(c_i | H_1, \mathbf{g}_i, \mathbf{m}_t(j), \mathbf{r}(j))$ using equation (16);
      (iv) score := score $\times P(c_i | H_1, \mathbf{g}_i, \mathbf{m}_t(j), \mathbf{r}(j))$.
   (f) sum := sum + score. Note that score $= P(c_1, c_2, \ldots, c_m | H_1, \mathbf{G}, \mathbf{m}_t(j), \mathbf{r}(j))$.
3. Estimate $P(c_1, c_2, \ldots, c_m | H_1, \mathbf{G}, \mathbf{M})$ as sum$/N$.
4. We can also estimate the posterior distribution for x, y, Q, h, and t, but do not show it here.

In the evaluation and in practice, we use $N = 200\,000$ cycles.

A.5. Simulating releases of B. anthracis

Because BARD is a Bayesian model, it is also possible to use it to simulate releases of *B. anthracis* spores and the resulting outbreak. The procedure we used in this study is as follows:

1. Choose an arbitrary date/time of release of spores and obtain the meteorological data for that date/time. We chose an arbitrary date/time of release from the interval covered by the test set of historical data (1 January 2003 to 31 December 2005).
2. Sample values for x, y, and h from their prior probability distributions (as discussed above, we deterministically chose Q). For simulated outbreaks, we use a different prior over x and y than uniform (Appendix A.3).

DOI: 10.1002/sim

3. For each zip code i in the region:
   (a) use equation (10) to compute d.
   (b) Use equation (12) to derive $P(\text{visit ED} \mid d)$. Call this probability p_i.
   (c) For the total number of cases from zip code i, use a random variate c_i from $\text{Bin}(p_i, n_i)$, where n_i is the population of the zip code.
   (d) For each of the c_i cases, assign a date/time of presentation to the ED:
      (i) Generate a random variate from the lognormal distribution corresponding to the value of d obtained in step 3a.
      (ii) Generate a random variate from the lognormal distribution for visit delay ($\mu = 1.015$ and $\sigma = 0.737$, see Hogan and Wallstrom [43]).
      (iii) Add these values to the release date and time to obtain ED presentation date/time.

Because (1) not every case of inhalational anthrax would visit one of the 10 EDs in our historical data set and (2) our historical data set contains data on approximately 30 per cent of ED visits in the Pittsburgh MSA, we multiplied p_i in step 3b by a constant multiple of 0.30. In reality, each zip code likely has a different fraction of sick individuals that would visit one of the EDs when ill, but we did not have sufficient information to estimate this fraction.

ACKNOWLEDGEMENTS

This material is based upon work supported by the Centers for Disease Control and Prevention under grant number 1 RO1 PH000026-01, the Defense Advanced Research Projects Agency under contract F30602-01-2-0550, the National Science Foundation under grant no. 0325581, and the Pennsylvania Department of Health under award number ME-01-727.

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation, the official views of the Centers for Disease Control and Prevention, or the official policies, either expressed or implied, of the Defense Advanced Research Projects Agency, Rome Laboratory, or the United States Government.

REFERENCES

1. Kaufmann A, Meltzer M, Schmid G. The economic impact of a bioterrorist attack: are prevention and postattack intervention programs justifiable? *Emerging Infectious Diseases* 1997; **3**(2):83–94.
2. Wagner MM, Tsui FC, Espino JU, Dato VM, Sittig DF, Caruana RA *et al*. The emerging science of very early detection of disease outbreaks. *Journal of Public Health Management Practice* 2001; **7**(6):51–59.
3. Office of Technology Assessment. Proliferation of weapons of mass destruction: assessing the risks. Washington, DC. *Report No. OTA-ISC-559*, US Congress, August 1993. Available at: http://www.wws.princeton.edu/ota/disk1/1993/9341_n.html (Accessed on 21 August 2006).
4. Wein LM, Craft DL, Kaplan EH. Emergency response to an anthrax attack. *Proceedings of the National Academy of Sciences USA* 2003; **100**(7):4346–4351.
5. Brookmeyer R, Johnson E, Bollinger R. Public health vaccination policies for containing an anthrax outbreak. *Nature* 2004; **432**(7019):901–904.
6. Craft D, Wein L, Wilkins A. Analyzing bioterror response logistics: the case of anthrax. *Management Science* 2005; **51**(5):679–694.
7. Wilkening DA. Sverdlovsk revisited: modeling human inhalation anthrax. *Proceedings of the National Academy of Sciences USA* 2006; **103**(20):7589–7894.
8. Mandl KD, Overhage JM, Wagner MM, Lober WB, Sebastiani P, Mostashari F *et al*. Implementing syndromic surveillance: a practical guide informed by the early experience. *Journal of American Medical Informatics Association* 2004; **11**(2):141–150.
9. Heffernan R, Mostashari F, Das D, Karpati A, Kuldorff M, Weiss D. Syndromic surveillance in public health practice, New York City. *Emerging Infectious Diseases* 2004; **10**(5):858–864.

DOI: 10.1002/sim

10. Henning KJ. What is syndromic surveillance? *Morbidity and Mortality Weekly Report* 2004; **53**(Suppl.):5–11.
11. Guillemin J. Detecting anthrax: what we learned from the 1979 Sverdlovsk outbreak. In *Scientific and Technical Means of Distinguishing Between Natural and Other Outbreaks of Disease*, Dando MR, Pearson GS, Kriz B (eds). Kluwer: Dordrecht, Boston, 2001.
12. Update: Investigation of anthrax associated with intentional exposure and interim public health guidelines, October 2001. *Morbidity and Mortality Weekly Report* 2001; **50**(41):889–893.
13. Lewis MD, Pavlin JA, Mansfield JL, O'Brien S, Boomsma LG, Elbert Y *et al*. Disease outbreak detection system using syndromic data in the greater Washington DC area. *American Journal of Preventive Medicine* 2002; **23**(3):180–186.
14. Lombardo J. The ESSENCE disease surveillance test bed for the National Capital Area. *Johns Hopkins APL Technical Digest* 2003; **24**(4):327–334.
15. Tsui F-C, Espino JU, Dato VM, Gesteland PH, Hutman J, Wagner MM. Technical description of RODS: a real-time public health surveillance system. *Journal of American Medical Informatics Association* 2003; **10**(5):399–408.
16. Wagner MM, Robinson JM, Tsui FC, Espino JU, Hogan WR. Design of a national retail data monitor for public health surveillance. *Journal of American Medical Informatics Association* 2003; **10**(5):409–418.
17. Bradley CA, Rolka H, Walker D, Loonsk J. BioSense: implementation of a national early event detection and situational awareness system. *Morbidity and Mortality Weekly Report* 2005; **54**(Suppl.):11–19.
18. Biological warning and incident characterization. *Report No. 2006-4491P*, Department of Homeland Security, 2006. Available at: http://www.sandia.gov/mission/homeland/factsheets/new06/BWIC_factsheet.pdf (Accessed on 16 August 2006).
19. Wong WK, Moore A, Cooper G, Wagner M. WSARE: what's strange about recent events? *Journal of Urban Health* 2003; **80**(Suppl. 1):l66–l75.
20. Ozonoff A, Forsberg L, Bonetti M, Pagano M. Bivariate method for spatio-temporal syndromic surveillance. *Morbidity and Mortality Weekly Report* 2004; **53**(Suppl.):59–66.
21. Burkom HS, Murphy S, Coberly J, Hurt-Mullen K. Public health monitoring tools for multiple data streams. *Morbidity and Mortality Weekly Report* 2005; **54**(Suppl.):55–62.
22. Wong WK, Cooper G, Dash D, Levander J, Dowling J, Hogan W *et al*. Use of multiple data streams to conduct Bayesian biologic surveillance. *Morbidity and Mortality Weekly Report* 2005; **54**(Suppl.):63–69.
23. Meselson M, Guillemin J, Hugh-Jones M, Langmuir A, Popova I, Shelokov A *et al*. The Sverdlovsk anthrax outbreak of 1979. *Science* 1994; **266**(5188):1202–1208.
24. Bridis T. Government previews bioterrorism sensors. *Miami Herald*, 15 November 2003.
25. Office of Inspector General. EPA needs to fulfill its designated responsibilities to ensure effective BioWatch program. *Report No. 2005-P-00012*, Environmental Protection Agency, 2005. Available at: www.epa.gov/oig/reports/2005/20050323-2005-P-00012.pdf (Accessed on 20 September 2006).
26. Broome C, Pinner R, Sosin D, Treadwell T. On the threshold. *American Journal of Preventive Medicine* 2002; **23**(3):229.
27. Buckeridge D, Owens D, Switzer P, Frank J, Musen M. Evaluating detection of an inhalational anthrax outbreak. *Emerging Infectious Diseases* 2006; **12**(12):1942–1949.
28. Stoto M, Schonlau M, Mariano L. Syndromic surveillance: is it worth the effort? *Chance* 2004; **17**(1):19–24.
29. Buckeridge DL, Burkom H, Campbell M, Hogan WR, Moore AW. Algorithms for rapid outbreak detection: a research synthesis. *Journal of Biomedical Informatics* 2005; **38**(2):99–113.
30. Fienberg SE, Shmueli G. Statistical issues and challenges associated with rapid detection of bio-terrorist attacks. *Statistics in Medicine* 2005; **24**(4):513–529.
31. Burkom HS. Biosurveillance applying scan statistics with multiple, disparate data sources. *Journal of Urban Health* 2003; **80**(2 Suppl. 1):i57–i65.
32. Meselson M. Note regarding source strength. *The ASA Newsletter* 2001; **87**:1, 10–12.
33. Guillemin J. *Anthrax*: *The Investigation of a Deadly Outbreak*. University of California Press: Berkeley, 1999.
34. Fowler RA, Sanders GD, Bravata DM, Nouri B, Gastwirth JM, Peterson D *et al*. Cost-effectiveness of defending against bioterrorism: a comparison of vaccination and antibiotic prophylaxis against anthrax. *Annals of Internal Medicine* 2005; **142**(8):601–610.
35. Kass RE, Raftery AE. Bayes factors. *Journal of the American Statistical Association* 1995; **90**(430):773–795.
36. Hanna S, Briggs G, Hosker R. Handbook on atmospheric diffusion. *Report No. DOE/TIC-11223*. Technical Information Center, Department of Energy, Washington, DC, 1982.
37. Gelman A, Carlin JB, Stern HS, Rubin DB. *Bayesian Data Analysis*. Chapman & Hall: London, 1995.
38. Heckerman D, Breese J. A new look at causal independence. In *Uncertainty in Artificial Intelligence*, Poole D, Mantaras RLd (eds), vol. 10. Morgan Kaufmann: Los Altos, CA, 1994; 286–292.

DOI: 10.1002/sim

39. Buckeridge DL, Burkom H, Moore A, Pavlin J, Cutchis P, Hogan W. Evaluation of syndromic surveillance systems—design of an epidemic simulation model. *Morbidity and Mortality Weekly Report* 2004; **53**(Suppl.): 137–143.
40. Jernigan JA, Stephens DS, Ashford DA, Omenaca C, Topiel MS, Galbraith M *et al*. Bioterrorism-related inhalational anthrax: the first 10 cases reported in the United States. *Emerging Infectious Diseases* 2001; **7**(6):933–944.
41. Glassman H. Industrial inhalation anthrax: discussion. *Bacteriological Reviews* 1966; **30**(3):657–659.
42. Tigertt WD, Benenson AS, Gochenour WS. Airborne Q fever. *Bacteriological Reviews* 1961; **25**:285–293.
43. Hogan WR, Wallstrom GL. Approximating the sum of lognormal distributions to enhance models of inhalational anthrax. *Quantitative Methods in Defense and National Security*. Fairfax, VA, 30 March 2007; 32–34. Available at: http://www.galaxy.gmu.edu/QMDNS2007/QMDNS2007-booklet.pdf (Accessed on 30 March 2007).
44. Wu J, Mehta NB, Zhang J. A flexible lognormal sum approximation method. *IEEE Global Telecommunications Conference* (*GLOBECOM*), St. Louis, MO, 2005; 3413–3417.
45. Shachter RD, Peot MA. Simulation approaches to general probabilistic inference on belief networks. In *Uncertainty in Artificial Intelligence*, Henrion M, Shachter RD, Kanal LN, Lemmer JF (eds), vol. 5. North-Holland: Amsterdam, 1990; 221–231.
46. Fung R, Chang KC. Weighting and integrating evidence for stochastic simulation in Bayesian networks. *Proceedings of the Workshop on Uncertainty in Artificial Intelligence*, Windsor, Ont., 1989; 112–117.
47. Zhang J, Tsui FC, Wagner MM, Hogan WR. Detection of outbreaks from time series data using wavelet transform. *AMIA Annual Symposium Proceedings*, Washington, DC, 2003; 748–752.
48. Reis B, Pagano M, Mandl K. Using temporal context to improve biosurveillance. *Proceedings of the National Academy of Sciences USA* 2003; **100**(4):1961–1965.
49. Cooper G, Dash DH, Levander JD, Wong WK, Hogan WR, Wagner MM. Bayesian biosurveillance of disease outbreaks. *Proceedings of the Conference on Uncertainty in Artificial Intelligence*, Banff, Alberta, 2004; 94–103.
50. Buckeridge D. A method for evaluating outbreak detection in public health surveillance systems that use administrative data. *Ph.D. Thesis*, Stanford, 2005.
51. Rickmeier G, McClellan G, Anno G. Biological warfare human response modeling. *Military Operations Research* 2001; **6**(3):35–57.
52. Kennedy H. Daschle letter bombshell billions of anthrax spores. *New York Daily News* 31 October 2001; Section 5.
53. Inglesby TV, O'Toole T, Henderson DA, Bartlett JG, Ascher MS, Eitzen E *et al*. Anthrax as a biological weapon, 2002: updated recommendations for management. *Journal of the American Medical Association* 2002; **287**(17):2236–2252.
54. Brookmeyer R, Blades N, Hugh-Jones M, Henderson DA. The statistical analysis of truncated data: application to the Sverdlovsk anthrax outbreak. *Biostatistics* 2001; **2**(2):233–247.
55. WebMET. *Turner's Method*. 2002. Available at: http://www.webmet.com/met_monitoring/641.html (Accessed on 15 March 2005).
56. Wong W, Moore A, Cooper G, Wagner M. Bayesian network anomaly pattern detection for disease outbreaks. In *Proceedings of the 20th International Conference on Machine Learning*, Fawcett T, Mishra N (eds). AAAI Press: Menlo Park, CA, 2003; 808–815.
57. Wong W, Moore A, Cooper G, Wagner M. Rule-based anomaly pattern detection for detecting disease outbreaks. *Proceedings of the Conference of the American Association of Artificial Intelligence* (AAAI), Edmonton, Alberta, 2002; 217–223.
58. Lombardo JS, Burkom H, Pavlin J. ESSENCE II and the framework for evaluating syndromic surveillance systems. *Morbidity and Mortality Weekly Report* 2004; **53**(Suppl.):159–165.
59. Siegrist D, Pavlin J. Bio-ALIRT biosurveillance detection algorithm evaluation. *Morbidity and Mortality Weekly Report* 2004; **53**(Suppl.):152–158.
60. Fawcett T, Provost F. Activity monitoring: noticing interesting changes in behavior. *Proceedings of the Fifth International Conference on Knowledge Discovery and Data Mining*, San Diego, CA, 1999; 53–62.
61. Inhalation anthrax associated with dried animal hides—Pennsylvania and New York City, 2006. *Morbidity and Mortality Weekly Report* 2006 **55**(2):280–282.
62. World Health Organization. Health Aspects of Chemical and Biological Weapons. World Health Organization, Geneva, 1970.
63. Altman P, Gibson J, Wang C. *Handbook of Respiration*. WB Saunders: Philadelphia, 1958.

DOI: 10.1002/sim