Chapter 3

Biosurveillance and Outbreak Detection

Paola Sebastiani, PhD[*], *Kenneth D. Mandl, MD, MPH* [×]
[*]Department of Biostatistics, Boston University, 715 Albany Street, Boston, MA 02118
[×]Informatics Program and Division of Emergency Medicine, Children's Hospital Boston, 300 Longwood Avenue, Boston, MA 02115

Abstract:
Faced with the very real threat of bioterrorism, the critical need for early detection of an outbreak has shortened the time frame for major enhancements to our public health infrastructure. The early detection of covert biological attacks requires real time data streams revealing of the health of the population, as well as novel methods to detect abnormalities.

Keywords: bioterrorism, spatial clustering, syndromic surveillance, temporal modeling.

3.1 Biological warfare agents

Our national security, threatened by asymmetrical warfare against military and civilian targets, has become increasingly dependent on quick acquisition, processing, integration and interpretation of massive amounts of data. In the mid 1990s, the White House and DOD identified bioterrorism directed against civilian populations as a substantial risk [Centers for Disease Control, 2000], and then stepped up efforts to prevent mass civilian casualties after the anthrax attacks in fall 2001 [Jernigan *et al.*, 2001]. Early detection of bioterrorism requires both real time data and real time interpretation. Toward this goal, the biomedical, public health, defense, law enforcement, and and intelligence communities all endeavor to develop new means of communication and novel sources of data. Public health surveillance is defined as “the ongoing, systematic collection, analysis, interpretation, and dissemination of data regarding a health-related event for use in public health action to reduce morbidity and mortality and to improve health” [Centers for Disease Control and Prevention, 2001].

In response to this need, a new information infrastructure to support active, real time surveillance is emerging, see [Kohane, 2002, Koplan, 2001, Lober *et al.*, 2002, O’Toole, 2001, Teich *et al.*, 2002, Yasnoff *et al.*, 2001]. In fact, real time surveillance systems are now in place in several cities including Boston [Reis, Pagano, & Mandl, 2003], Washington DC [Lombardo *et al.*, 2003], Pittsburgh [Tsui *et al.*, 2003], and New York [Greenko *et al.*, 2003]. These new systems, along with direct environmental monitoring [Miller, 2003] are expected to form the basis of a national surveillance network.

While developing technologies will support not only bioterrorism surveillance but also the tracking and understanding of natural outbreaks, it is the lethality and rapidity of effect of biological warfare agents that most strongly compel the need for a real time infrastructure. The Centers for Disease Control and Prevention (CDC) classifies biological warfare agents in terms of how easily they could be disseminated to large numbers of people with the potential for causing mass casualties. The most threatening agents are classified as Category A and include anthrax (bacillus anthracis), botulism, plague, smallpox, tularemia, and the viral hemorrhagic fevers.

The biological features of Category A agents and the clinical effects on their victims require that urgent efforts be focused on developing of methods for early detection and monitoring. The case of an anthrax attack is illustrative. Anthrax is a spore forming bacterium. Spores are dormant cells that can reactivate under the right conditions. From the time that a person is first exposed to anthrax, the incubation period is one to six days until symptoms appear. The first signs and symptoms resemble those of influenza and include fever, malaise and fatigue. Within two to three days, patients become severely ill and develop respiratory distress, overwhelming blood infection, and in about half of the cases, meningitis, or infection of the fluid and tissues surrounding the brain and spinal cord. If antibiotic treatment is not initiated prior to the onset of the acute, severe phase of the illness, survival is unlikely. However, antibiotic prophylaxis at the time of exposure or treatment during the early phase of the disease is very effective [U.S. Army Medical Research Institute of Infectious Diseases, 2001].

3.2 Detection

It is reasonable to assume that a biological attack on a civilian population will be carried out covertly. Hence, the first indication of an attack may be large numbers or smaller clusters of people developing an influenza-like illness between one and six days after the attack. Ideally, if this surge in people with influenza-like illness were detected at the early possible moment, a large scale public health response would be mobilized to initiate treatment of all those exposed. Epidemiological models presented in a Defense Department Jasons report project that fatalities resulting from a covert anthrax attack on the New York City subway system would be reduced by a factor of seven if an active surveillance system were in place. Such a system would reduce mortality by seven orders of magnitude for a communicable, highly contagious disease such as smallpox. An infrastructure to support fundamental change in the health care system must therefore include real time information about regional disease patterns, health care processes, and health related behaviors. Current health information systems fall far short of this capability [Fernandez, 1999], despite readily available information technology to process patient data. The Federal focus on preparedness for bioterrorism [Centers for Disease Control, 2000], which was sharpened after the 2001 anthrax mailings [Jernigan *et al.*, 2001], has thrown into greater relief the problems we face when providing health care and protecting the public health in the absence of real time information [Henderson, 1999].

The state of the art for detecting abnormalities in surveillance data is primitive. Analyses are typically manual, and even when the data are available, the methods often involve simply comparing a count to an expected value and alarming when the count is too high.

3.3 Approaches to surveillance

The goal of biosurveillance is to identify biological weapons before they are used or as soon as possible after release. The major targets for monitoring include the environment, animals, agriculture, the citizenry, or patient populations. Once an attack has occurred, surveillance of the environment provides the earliest possible warning of a biological release into, for example, the ventilation system of an office building. Detection systems such as those employing laser-induced fluorescence and immunologically based bioelectric sensors have been shown to be highly sensitive to biological releases [Primmerman, 2000]. Animals and agriculture are prime targets for bioterrorism. The extensive outbreak of foot and mouth disease in Great Britain [Ferguson, Donnelly, & Anderson, 2001] illustrates the devastating economic and social impact that such an attack could have. The citizenry may be observed, polled, or examined. Telemarketing and the emerging field of consumer informatics enable surveillance of the ever growing population of Internet users [Mandl *et al.*, 2000].

Behaviors of the citizenry, when their health is affected, may leave imprints on certain data sets. The principal underlying premise of these systems is that the first signs of a covert biological warfare attack will be clusters of victims who change their behavior because they begin to feel ill. When people become sick, they may make purchases such as facial tissue, orange juice, and over-the-counter cold remedies. [Goldenberg *et al.*, 2002], and rates of absenteeism. They may stay home from school or work. The next phase of detectable activity is likely to be encounters with the health care system. Patients may phone in to nurses or physicians. They may visit primary care settings [Lazarus *et al.*, 2002], and emergency departments [Lober *et al.*, 2002]. They may be hospitalized. Some may die. All of this activity may precede the first confirmed diagnosis of a biological warfare victim.

3.4 Syndromic surveillance

One recent approach to the problem of how to recognize an attack as soon as possible has been the development of syndromic surveillance systems. Syndromic surveillance relies on data that are available before the diagnosis of the individual patient and would precede recognition that there has been an attack. One approach to syndromic surveillance of emergency department populations has been drop-in surveillance as practiced by the CDC. Drop-in surveillance was implemented by the CDC at the World Trade Organization Ministerial (Seattle, Washington) in 1999, and at the Super Bowl (Tampa, Florida) and World Trade Center Attacks (New York, NY) in 2001 [Centers for Disease Control, 2002b]. Drop-in surveillance is accomplished by staffing the emergency departments around an area thought to be at high risk of an attack. Health care providers use a paper form to indicate whether each patient fits a particular syndrome of concern. At the World Trade Center, the syndromes tracked were anxiety, asthma, botulism-like, death, gastrointestinal, inhalation, neurological, rash, respiratory, sepsis, and trauma. There are limitations to drop-in surveillance as it has recently been practiced. The very nature of the method assumes that the time and location of an attack can be predicted. Furthermore, reliance on manual processes produces incomplete and at times inaccurate data. Designing automated systems and interpreting their output, however, poses substantial methodological challenges.

3.5 Outbreak Detection

3.5.1 Normal Periodicity

One approach to identify clusters of patients infected by biological warfare agents is to set up a detection system that distinguishes abnormal from normal patterns. The goal such a detection system is to signal an alarm whenever the observed pattern departures from the expected normal pattern.

The most straight forward approach builds a control chart, such as CuSUM described for example in [Basseville & Nikiforov, 1993], in which cumulative differences between observed and expected data in a time window are compared to a threshold. In traditional CuSUM, the expected value is a theoretical mean describing, for example, the average number of hospital visits, and it is constant over time. A suspicious increase in the observed data over the theoretical mean is potentially indicative of an emerging outbreak and, to allow for sampling variability, the threshold of the maximum difference between observed and expected values is typically some multiple of the standard error of the sample mean. Because many health care datasets show regular periodicities (annual, seasonal, monthly, daily) the theoretical mean needs to change over time [Lewis *et al.*, 2002, Reis, Pagano, & Mandl, 2003]. Figure 3.1, which shows the number of daily emergency department visits at Children's Hospital Boston between June 1992 and February 2002, demonstrates such periodicity. The CuSUM method

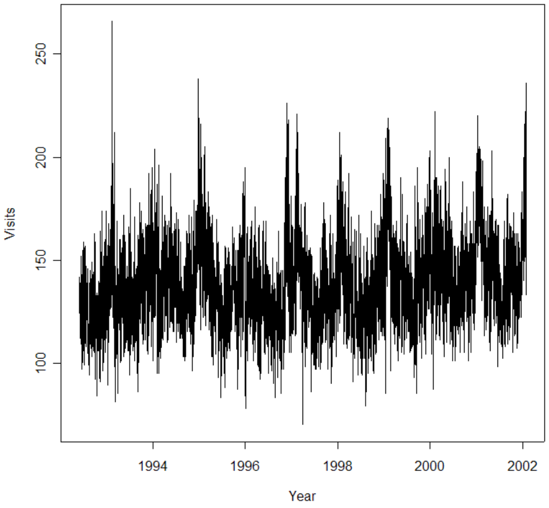


Figure 3.1: Number of daily visits at the emergency department of Children's Hospital, Boston, over the last 10 years.

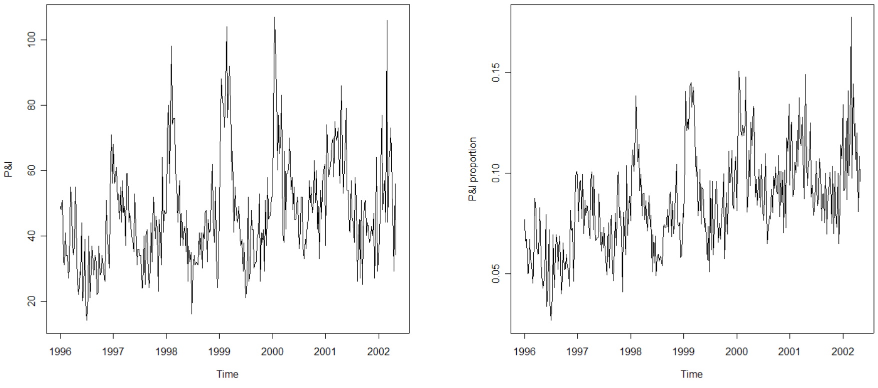
was corrected for seasonal and daily variations and is implemented in the CDC's Early Aberration Reporting Systems (EARS).

3.5.2 Temporal Modeling

Model-based detection approaches attempt to detect outbreaks by comparing the observation y_t with the predictions $\hat{y}_t$ generated by a model describing the normal pattern. More advanced multi-day temporal filters [Reis, Pagano, & Mandl, 2003] compare a weighted prediction of multiple days at once to a threshold. This, in effect lessens the effects of the large variability of hospital visit rates and improve both the timeliness and sensitivity of detection.

Crucial choices in model-building are how to characterize the normal patterns and at what threshold to signal an alarm. Threshold values may be a multiple κ of the standard error s_t of the prediction. Generally κ is set between 2 and 3.5 to ensure that the false alarm rate is below 5%. An alarm is signaled whenever the observation y_t is outside the interval $\hat{y} \pm \kappa s_t$. Adjustment to take into account multiple comparisons can be based on Bonferroni's correction, or randomization procedures as those implemented in the RODS system [Tsui *et al.*, 2002, Wagner *et al.*, 2001].

To establish normal patterns of temporal relationships among patients, it is helpful to have a few years of historical data at the surveillance sites that include regular recurrence of cyclic diseases, such as influenza. These historical data naturally contain local variations as well as trends in local population density and shifting referral patterns.



a) b)

Figure 3.2: Left: Weekly number of deaths for pneumonia and influenza in New England between January 1996 and April 2002. Right: Weekly proportion of deaths for pneumonia and influenza in New England between January 1996 and April 2002, after normalization with the number of deaths for all causes. Source: CDC, Morbidity and Mortality Weekly Report (http://wonder.cdc.gov/mmwr/mmwrmort.asp).

Data normalization can help to reduce the effect of some of these factors.

Figure 3.2 a) shows the weekly number of death for pneumonia and influenza in New England between January 1996 and April 2002 [Centers for Disease Control, 2002a]. The data are published by the CDC as part of its national influenza surveillance effort. Figure 3.2 b) shows the weekly proportion of deaths for pneumonia and influenza after normalization with the number of deaths for all causes. The apparent increasing trend may be part of a secular trend with cyclic episodes of severe influenza epidemics [Simenson *et al.*, 2000]. Modeling normal patterns of disease or visit rates may rely on regression type models [[Serfling, 1963], classical autoregressive moving average models ($ARIMA$) [Box & Jenkins, 1976] or a combination of both methods. Serfling's method monitors the normal pattern of susceptibility to death for pneumonia and influenza when there is not epidemic, and the objective is to determine an epidemic threshold. The method uses the cyclic regression equation

$$y_t = \mu_0 + \theta t + \alpha \sin(2\pi t/52) + \beta \cos(2\pi t/52) + \epsilon_t \tag{3.1}$$

to model the weekly number y_t of people susceptible to death from pneumonia and influenza when there is no epidemic. The model consists of the linear component θt describing the secular trend and the sine-wave component $\alpha \sin(2\pi t/52) + \beta \cos(2\pi t/52)$ describing the periodic recurrence of influenza epidemics. Assuming that the weekly numbers of deaths are normally distributed and uncorrelated, the standard least squares method is used to estimate the model parameters from non-epidemic data, and to compute the predicted values $\hat{y}_t$. This model can be easily expanded to monitor hospital visit data for influenza [Simenson *et al.*, 2000, Tsui *et al.*, 2002], to describe different secular trends by changing the linear component, and it can be adapted to include exogenous variables such as weather or pollution indicators and to model daily or monthly data, by changing the sine-wave component.

Serfling's method is the standard approach to detect epidemic thresholds for influenza and, in principle, it could be adapted to detect epidemic thresholds for other cyclic diseases. Its use requires a clear definition of the disease, the selection of data to identify a normal pattern of susceptible individuals, and the assumption that the normal pattern is periodical. However, to be able to detect an outbreak determined by a bioterrorism attack, recurrent incidence of cyclic diseases should be part of the "normal pattern of diseases" that underlies, for example, the dynamics of hospital visits rates. Furthermore, the assumption that the data are independent may be violated by hospital visits rates that may be dependent upon the past. Traditional time series models, such as $ARIMA$, seem to be more adequate to describe historical visits rates and can account for temporal dependency, trends corresponding to secular changes in the populations, and seasonal effects due, for example, to annual recurrence of influenza epidemics. An $ARIMA(p, d, q)$ model describes a non-stationary time series in which the series of differences $x_t = y_t - y_{t-d}$ is stationary and is described by an $ARMA(p, q)$ model. The $ARMA(p, q)$ model consists of two components: an autoregressive component of order p, and a moving average component of order q. The autoregressive component of order p describes explicitly the dependency of x_t on a linear combination of the past p points, whereas the moving average component uses an average of the past q uncorrelated errors to "shock" the process mean. Both the autoregressive and moving average orders can be chosen to get the best fitting from historical data, although large orders may compensate for recent trends and miss timely departures from normal patterns. Goodness of fit criteria such as the Akaike or the Schwarz information criteria [Shumway & Stoffer, 2000] can be used to select the best orders p, q by trading off goodness of fit and model complexity, thus limiting the risk of over fitting. $ARIMA$ models can be extended to account for seasonality, by either decomposing the time series into a trend, a seasonal component and the residual, and by fitting an $ARMA(p, q)$ model to the residual component, or by fitting all components simultaneously [Box & Jenkins, 1976]. $ARIMA$ models to fit time series of patients with respiratory syndromes are described in [Reis & Mandl, 2003].

$ARIMA$ models are standardly used to fit time series data that do not exhibit a clear pattern, although many historical data at surveillance sites show clear periodicities. The regression type model in Equation 3.1 describes periodicities of the mean by the sine wave component, but it does not take into account temporal correlation. A combination of the two models would give the advantage of an explicit description of important components such as periodicity and secular trends and, at the same time, it could account for the serial correlation.

More sophisticated approaches to model the normal pattern use hidden Markov models [LeStrat & Carrat, 1999, Rabiner, 1989, Rath, Carreras, & Sebastiani, 2003] by using a hidden state that describes the presence or absence of epidemic of a particular disease, and a model of the data conditional on the epidemic status. Closely related to hidden Markov models are change point algorithms to detect changes in a baseline model describing the normal pattern [Basseville & Nikiforov, 1993]. A potential use of change-point algorithms is to automatically identify the historical data most relevant for modeling the normal pattern. When massive data at the surveillance sites are available, it is important to decide whether all data should be used to formulate a model for the normal pattern. Often, changes in the data collection or in population demographics determine a change of the dynamics underlying the normal pattern and not all the historical data should be used to have the best prediction. For example, modeling data from influenza mortality surveillance has shown that more accurate forecasts are based on five years historical data, and the CDC syndromic surveillance practice is to use at least five years historical data [Farrington *et al.*, 1996, Hutwagner *et al.*, 1997, Stern & Lightfoot, 1999]. The choice of the amount of historical data is an open issue and the availability of such data will vary.

3.5.3 Spatial and Temporal Clustering Methods

Several traditional methods are available for cluster analysis in time, once a suspect location is identified. Methods as the scan test in [Naus, 1965] or the method in [Ederer, Myers, & Mantel, 1964] testing for temporal clustering using a cell-occupancy approach. In the scan test, the maximum number of cases observed in an interval of fixed length is found by scanning all intervals of the same length in a time period. The method in [Ederer, Myers, & Mantel, 1964] compares the distribution of cases into k disjoint subintervals with the distribution one would expect under the null hypothesis of no clustering. The difference between the two approaches is that the scan test uses overlapping intervals, compared to the disjoint intervals used in [Ederer, Myers, & Mantel, 1964].

Methods for spatial clustering can be grouped according to whether they use information about individuals or aggregate data. Our work has focused on detection of perturbations in the distribution of mutual distances among all individual cases in a geographical region to identify clusters. We find substantially increased power using this measure of geographical clustering compared with using relying on counts alone [Olson *et al.*, 2002].

The nearest-neighbor method examines the distribution of interpoint distances of infected individuals. The method of Ohno [Ohno, Aoki, & Aoki, 1978] is a simple test for spatial clustering that uses rates for geographic areas rather than data for individual cases. The test assesses whether the rates in adjacent areas are more similar than would be expected under the null hypothesis of no clustering. More advanced Bayesian models for detection of spatial clusters of disease - known as Bayesian disease mappings - are described in [Elliott *et al.*, 2000], and the extension of hidden Markov models to detect small clusters in space is in [Green & Richardson, 2002]. A potential limitation of Bayesian methods to disease mapping is their computational burden. This may make them unfeasible for monitoring rapidly shifting distributions of patients within an ill-defined window of hours to days. Furthermore, one of the fundamental steps in spatial analysis based on aggregate data is the definition of geographical areas to aggregate point data into groups. Geographical information systems are used to form spatial groups by aggregating small area units into larger, contiguous areas that can be given a particular classification. A common problem with spatial clustering procedures is that the scale of the area units chosen for the aggregation can have dramatic effects on the results of the classification. This effect is known as the Modifiable Area Unit Problem (MAUP) and a solution to this problem was proposed in [Hobbs, 1996]. Most of the methods described for spatial clustering work retrospectively, whereas on line surveillance system need to process observations sequentially. Statistical methods for on-line surveillance are described in [Rogerson, 1997]

Detection of clusters in space and time is a much harder problem, and fewer methods are available. Knox method [Knox, 1964] is a standard epidemiology approach to monitor the temporal-spatial spread of a disease by examining the distribution of geographical and temporal distances of affected individuals. The method works by classifying the pairs of cases as close in space and time, close in space only, close in time only, or close in neither space nor time. The limitation is the choice of time and space parameters to classify pairs as close or not, and it is overcome by temporal-spatial scan statistics [Kulldorf, 2001] that identify ”suspect clusters” by using a window that moves in time and space. In each window, the number of reported cases for a particular disease is compared to the number expected under the hypothesis that there is no cluster, and the window with largest evidence of cluster is identified. The K-nearest neighbor test of [Jacquez, 1996] uses a randomized test to assess the statistical significance of a potential space-time interaction process. Because most data at surveillance sites are count data, such as number of chief complaints for respiratory problems [Lazarus *et al.*, 2002], one can use generalized linear models to compute the expected number of cases of a particular syndrome in a particular time and location. The expected number is compared with the observed number and an unusually large incidence of a particular syndrome is used to detect a cluster of illness.

3.6 Challenges

The process of syndromic surveillance can be broken down into five stages. First, there is the Data Acquisition Stage, in which data are acquired from their sources. Next, in the Syndromic Grouping Stage, data are processed, according to a scheme that allows each the assignment of each patient to a syndromic group (such as influenza-like illness, or gastrointestinal illness). Next is the Modeling Stage, in which historical data, ideally reaching back from one to several years, are analyzed to establish a model of the normal temporal pattern. Fourth, in the Detection Stage, the expected values (for example, daily frequencies of patients presenting in each syndromic group) are compared against observed values collected in the field in order to determine abnormal activity is occurring. Finally, in the Alarms Stage, thresholds are set for evaluation of whether or not the unusual patterns warrant notification. Each stage has its own particular set of challenges.

3.6.1 Data Acquisition Stage

A problem that plagues health care information technology generally, and bioterrorism surveillance specifically, is a lack of universal standards for storage and communication of medical information [Kohane *et al.*, 1996]. Much medical data, even when stored electronically, exists as free text and requires natural language processing. (The

RODS Laboratory at the University of Pittsburgh has made a Bayesian language classifier available for this purpose at http://www.health.pitt.edu/rods/sw/default.htm). Even when the chief complaints are encoded, the classification schema have often been developed locally. Sharing information across systems requires a standardized method for describing data and their organization. Models for data exchange are being developed for health care [Beeler, 1998], emergency medicine [Pollock & Lowery, 2001, Barthell *et al.*, 2002], and public health. The Public Health Information Network (PHIN)[1] is a standards-based approach to connect public health and clinical medicine. As its implementation proceeds, PHIN will help standardize and facilitate the transfer of information needed for public health from clinical information systems.

Another issue is the timeliness of the data. To support real time detection, data must be available immediately. This requirement eliminates many otherwise useful data sources and sometimes precludes entire institutions from participation. The quality of the data used in automated systems varies. Establishing new manual data entry processes to collect data (effectively, relying on human intelligence) is difficult and costly. Therefore, the most successful surveillance systems to date rely on information already collected for other purposes. There are also important issues with the quality of geographical location data. Geographical information system (GIS) software can translate an address into precise latitude and longitude coordinates. In general, in the health care system, the only address reliably collected is the home address. Ideally, a surveillance system should have access to data about where each individual works, goes to school, and has been recently. If an attack were to occur at a baseball game, for example, the home address may not be the most salient geographical feature of the patients, but rather a cluster of patients who had attended the game might be identified. The home address may be entered inaccurately, making geocoding a challenging task. In hospital information systems, new patient addresses may permanently overwrite older ones, making the historical geographical patient distributions difficult to reconstruct. Also, the accuracy of geocoding results is far from perfect [Krieger *et al.*, 2001, Olson & Mandl, 2002].

3.6.2 Syndromic Grouping Stage

Since the data are automatically collected and were not originally designed to group patients into syndromes, they, may not be ideally suited to the purpose of surveillance. The mapping of each case onto a syndrome is imprecise. The degree of imprecision is, however, measurable against a gold standard such as a physician chart review of the case [Espino & Wagner, 2001, Beitel & Mandl, 2002].

3.6.3 Modeling Stage

A common problem when establishing a new surveillance system, is that data may only be available going forward, making the modeling of normality impossible before a few years have elapsed. If these data are available, then the long term and local periodicities must be known or discovered in order to model the baseline. Such modeling is

---

[1] http://www.cdc.gov/phin/index.htm

often based on an assumption that the data do not contain the signal of interest. While one can reasonably assume that a historical data set of, say, daily emergency department syndromic data, does not contain signal from a biological attack, it may well contain signal from non-periodically recurring events such as a food-borne outbreak of gastrointestinal disease.

3.6.4 Detection Stage

The detection performance of a syndromic surveillance system can be evaluated by measuring its ability to detect signal (disease outbreak) against a background of noise (normal variation in baseline disease rates in the region). To benchmark performance, training and validation data containing signal and noise are required. A major problem is the lack of training data. Fortunately, few people have been infected with biological warfare agents although there are notable exceptions such as the people of Sverdlovsk [Meselson *et al.*, 1994] exposed in 1979 during an accidental release of anthrax from a weapons plant, and those involved in the Florida, New York and Washington DC mail attacks in 2001. However, with few actual cases, syndromic surveillance systems cannot “learn” to detect bioterrorism with real world data. Nor can systems be benchmarked by their ability to detect actual attacks. These systems must be trained with nonbioterror related events or simulated events instead. Training and validation data sets can be samples of authentic regional data, synthetic data, or a combination of both (semisynthetic data).

Detecting abnormal clusters of syndromic cases in space/time poses substantial challenges. Consideration of the temporal-spatial distribution of syndrome cases may facilitate the detection of a bioterrorism attack, when cases are distributed over space and/or time in a different manner from the normal pattern. In classical cluster detection for public health and epidemiology, a specific cluster is often being investigated, and the goal is to identify a point source, such as a toxic waste dump causing an excess number of cases of leukemia [Waller, 2000]. In automated bioterrorism detection, the location of the cluster is not known in advance and the temporal window in which cases are grouped is constantly forward moving and may be only a few days wide.

Attacks may produce signals in the data that have different shapes. For example, in a syndromic surveillance system that tracked the daily emergency department rates illustrated in Figure 3.1, an attack might be characterized by a short high spike, a sustained low and flat signal, or an exponentially increasing high amplitude signal. Different methods, and different applications of methods, such as using varying temporal filters, may strongly influence detection capabilities [Reis, Pagano, & Mandl, 2003]. Therefore, a host of distinct methods may ultimately need to be simultaneously applied to the same data streams to enable the detection of all possible forms of attack.

3.6.5 Alarm Stage

Establishing appropriate thresholds is a nontrivial task. Obviously, a detection system with poor sensitivity, one that fails to detect most attacks, is not acceptable. However, in determining alarm thresholds, one must trade off sensitivity against specificity, ensuring that most alarms are real. Consequently, the more attacks one aims to detect, the more false alarms one must accept. The costs of false alarms are difficult to quantify but if a graded response is developed, news of a false alarm need not filter out to the public. In the case of bioterrorism, the cost of a missed signal might be astronomical. Decision analytic approaches are needed to estimate the human and economic implications across a range of alarm thresholds.

Using more than one signal stream may help to reduce false alarms. For example, if visits for influenza syndrome increase, but so do actual cases of influenza, there may not be a cause for alarm. If the two signal streams diverge, however, there may be more cause for concern. Ideally, a regional or national command center would act based upon an interpretation of multiple streams of data, whether they come from completely different sources (e.g. hospital visits, school absenteeism, sales of over-the-counter medication) or from different models of a single data source. Substantial work is needed to develop methods for such signal integration. Integration must be accomplished for: signals from same data using multiple methods, signals from separate but potentially correlated data streams, signals from overlapping geographical regions, and signals from remote geographical regions.

3.7 Conclusions

The precise role and efficacy of biosurveillance in public health has yet to be determined. Syndromic surveillance systems as they existed to date in the New York City and Washington DC areas failed to detect the anthrax attacks of 2001. While these attacks affected only a small number of people, they nonetheless are a cause for humility for anyone attempting to predict with certainty when, where, and how bioweapons will be used. Astute clinicians will always play a role in the accurate diagnosis and treatment of patients as well as in the identification of public health emergencies. However, biosurveillance is another modality that clearly has the ability to detect certain kinds of events. The work to be done over the coming months and years is to build our data integration infrastructure, develop and refine our methods, and estimate to the best of our ability, the promise and limits of our technology.

Acknowledgements

This work was supported by the Agency for Health Care Quality and Research, contract # 290-00-0020, and the National Institutes of Health through a grant from the National Library of Medicine R01LM07677-01.

Bibliography

[Barthell *et al.*, 2002] Barthell, E. N.; Cordell, W. H.; Moorhead, J. C.; Handler, J.; Feied, C.; Smith, M. S.; Cochrane, D. G.; Felton, C. W.; and Collins, M. A. 2002. The frontlines of medicine project: A proposal for the standardized communication of emergency department data for public health uses including syndromic surveillance for biological and chemical terrorism. *Annals of Emergency Medicine* 39:422–429.

[Basseville & Nikiforov, 1993] Basseville, M., and Nikiforov, I. 1993. *Detection of Abrupt Changes: Theory and Application.* Upper Saddle River, NJ: Prentice Hall.

[Beeler, 1998] Beeler, G. W. 1998. HL7 version 3–an object-oriented methodology for collaborative standards development. *International Journal of Medical Informatics* 48:151–161.

[Beitel & Mandl, 2002] Beitel, A. R. B., and Mandl, K. D. 2002. Use of emergency department chief complaint and ICD9 diagnostic codes for real time bioterrorism surveillance abstract. *Pediatric Research* 51:94.

[Box & Jenkins, 1976] Box, G. E. P., and Jenkins, G. M. 1976. *Time Series Analysis: Forecasting and Control*. San Francisco, CA: Holden-Day, 2nd edition.

[Centers for Disease Control and Prevention, 2001] Centers for Disease Control and Prevention. 2001. Updated guidelines for evaluating public health surveillance systems. In *MMWR Morbidity and Mortality Weekly Reports*. 1035.

[Centers for Disease Control, 2000] Centers for Disease Control. 2000. Biological and chemical terrorism: Strategic plan for preparedness and response. recommendations of the CDC strategic planning workgroup. In *MMWR Morbidity and Mortalilty Weekly Reports*, number 49(RR-4). 1–24.

[Centers for Disease Control, 2002a] Centers for Disease Control. 2002a. Morbidity and mortality tables. http://wonder.cdc.gov/mmwr/mmwrmort.asp.

[Centers for Disease Control, 2002b] Centers for Disease Control. 2002b. Syndromic surveillance for bioterrorism following the attacks on the world trade center–new york city, 2001. In *MMWR Morbidity and Mortality Weekly Reports*. 13–15.

[Ederer, Myers, & Mantel, 1964] Ederer, F.; Myers, M. H.; and Mantel, N. N. 1964. A statistical problem in space and time: do leukemia cases come in clusters? *Biometrika* 20:626–638.

[Elliott *et al.*, 2000] Elliott, P.; Wakefield, J. C.; Best, N. G.; and Briggseds., D. 2000. *Spatial Epidemiology: Methods and Applications.* Oxford: Oxford University Press.

[Espino & Wagner, 2001] Espino, J. U., and Wagner, M. M. 2001. Accuracy of ICD-9-coded chief complaints and diagnoses for the detection of acute respiratory illness. In *Proceedings of the Annual AMIA Fall Symposium*, 164–168.

[Farrington *et al.*, 1996] Farrington, C. P.; Andrews, N. J.; Beale, A. D.; and Catchpole, M. A. 1996. A statistical algorithm for the early detection of outbreaks of infectious disease. *Journal of the Royal Statistical Society A* 159:547–63.

[Ferguson, Donnelly, & Anderson, 2001] Ferguson, N. M.; Donnelly, C. A.; and Anderson, R. M. 2001. The foot-and-mouth epidemic in great britain: Pattern of spread and impact of interventions. *Science* 292:1155–1160.

[Fernandez, 1999] Fernandez, F. 1999. Statement before the subcommittee on emerging threats and capabilities, committee on armed services united states senate (DARPA).

[Goldenberg *et al.*, 2002] Goldenberg, A.; Shmueli, G.; Caruana, R. A.; and Fienberg, S. E. 2002. Early statistical detection of anthrax outbreaks by tracking over-the-counter medication sales. *Proceedings of the National Academy of Sciences of the USA* 99:5237—5240.

[Green & Richardson, 2002] Green, P. J., and Richardson, S. 2002. Hidden markov models for disease mapping. *Journal of the American Statistical Association* 97.

[Greenko *et al.*, 2003] Greenko, J.; Mostashari, F.; Fine, A.; and Layton, M. 2003. Clinical evaluation of the emergency medical services (EMS) ambulance dispatch-based syndromic surveillance system, New York City. *Journal of Urban Health* 80:i50–i56.

[Henderson, 1999] Henderson, D. A. 1999. The looming threat of bioterrorism. *Science* 283:1279–1282.

[Hobbs, 1996] Hobbs, M. H. 1996. Spatial clustering using a genetic algorithm. In Parker, D., ed., *Innovatins in GIS 3*. Taylor & Francis. chapter 8, 85–95.

[Hutwagner *et al.*, 1997] Hutwagner, L. C.; Maloney, E. K.; Bean, N. H.; Slutsker, L.; and Martin, S. M. 1997. Using laboratory-based surveillance data for prevention: An algorithm for detecting salmonella outbreaks. *Emerging Infectious Diseases* 3:395–400.

[Jacquez, 1996] Jacquez, G. M. 1996. A k nearest neighbour test for space-time interaction. *Statistics in Medicine* 15:1935–1949.

[Jernigan *et al.*, 2001] Jernigan, J. A.; Stephens, D. S.; Ashford, D. A.; Omenaca, C.; Topiel, M. S.; Galbraith, M.; M.Tapper; Fisk, T. L.; Zaki, S.; Popovic, T.; Meyer, R. F.; Quinn, C. P.; Harper, S. A.; Fridkin, S. K.; Sejvar, J. J.; Shepard, C. W.; Guarner, M. M. J.; Gerberding, W. J. S. J. M. M. J. L.; Hughes, J. M.; and and members of the Anthrax Bioterrorism Investigation Team, B. P. 2001. Bioterrorism-related inhalational anthrax: The first 10 cases reported in the united states. *Emerging Infectious Diseases* 7:933–944.

[Knox, 1964] Knox, E. G. 1964. The detection of space-time interactions. *Applied Statistics* 13:25–30.

[Kohane *et al.*, 1996] Kohane, I. S.; Fackler, F. J. V. W. J. C.; Cimino, C.; Kilbridge, P.; Murphy, S.; Rind, H. C. D.; Safran, C.; Barnett, O.; and Szolovits, P. 1996. Sharing electronic medical records across multiple heterogeneous and competing institutions. In *Proceedings of the Annual AMIA Fall Symposium*, 608–612.

[Kohane, 2002] Kohane, I. S. 2002. The contributions of biomedical informatics to the fight against bioterrorism. *Journal of the American Medical Informatics Association* 9:116–119.

[Koplan, 2001] Koplan, J. 2001. CDC's strategic plan for bioterrorism preparedness and response. *Public Health Reports* 116(Suppl):9–16.

[Krieger *et al.*, 2001] Krieger, N.; Waterman, P.; Lemieux, K.; Zierler, S.; and Hogan, J. W. 2001. On the wrong side of the tracts? Evaluating the accuracy of geocoding in public health research. *American Journal of Public Health* 91:1114–1116.

[Kulldorf, 2001] Kulldorf, M. 2001. Prospective time periodic geographical disease surveillance using scan statistics. *Journal of the Royal Statistical Society A* 164:61–72.

[Lazarus *et al.*, 2002] Lazarus, R.; Kleinman, K.; Dashevsky, I.; Adams, C.; Kludt, P.; DeMaria, A.; and Platt, R. 2002. Use of automated ambulatory-care encounter records for detection of acute illness clusters, including potential bioterrorism events. *Emergency Infectious Diseases* 8:753–760.

[LeStrat & Carrat, 1999] LeStrat, Y., and Carrat, F. 1999. Monitoring epidemiologic surveillance data using hidden Markov models. *Statistics in Medicine* 18:3463–78.

[Lewis *et al.*, 2002] Lewis, M.; Pavlin, J.; Mansfield, J.; OBrien, S.; Boomsma, L. G.; Elbert, Y.; and Kelley, P. W. 2002. Disease outbreak detection system using syndromic data in the greater Washington DC area. *American Journal of Preventive Medicine* 23:180–186.

[Lober *et al.*, 2002] Lober, W. B.; Karras, B. T.; Wagner, M. M.; Overhage, J. M.; Davidson, A. J.; Fraser, H.; Trigg, L. J.; Mandl, K. D.; Espino, J. U.; and C.Tsui, F. 2002. Roundtable on bioterrorism detection: Information system-based surveillance. *Journal of the American Medical Informatics Association* 9:105–115.

[Lombardo *et al.*, 2003] Lombardo, J.; Burkom, H.; Elbert, R.; Magruder, S.; Lewis, S. H.; Loschen, W.; Sari, J.; Sniegoski, C.; Wojcik, R.; and Pavlin, J. 2003. A systems overview of the electronic surveillance system for the early notification of community-based epidemics (ESSENCE II). *Journal of Urban Health* 80:i32–i42.

[Mandl *et al.*, 2000] Mandl, K. D.; Feit, S.; Pena, B. M. G.; and Kohane, I. S. 2000. Growth and determinants of access in patient e-mail and internet use. *Archives of Pediatric and Adolescent Medicine* 154:508–511.

[Meselson *et al.*, 1994] Meselson, M.; Guillemin, J.; Hugh-Jones, M.; Langmuir, A.; Popova, I.; Shelokov, A.; and Yampolskaya, O. 1994. The Sverdlovsk anthrax outbreak of 1979. *Science* 266:1202.

[Miller, 2003] Miller, J. 2003. U.S. is deploying a monitor system for germ attacks. *New York Times*. January 22, 2003.

[Naus, 1965] Naus, J. I. 1965. The distribution of the size of the maximum cluster of points on a line. *Journal of the American Statistical Association* 60:532–538.

[Ohno, Aoki, & Aoki, 1978] Ohno, Y.; Aoki, K.; and Aoki, N. 1978. A test of significance for geographic clusters of disease. *International Journal of Epidemiology* 8:273–281.

[Olson & Mandl, 2002] Olson, K., and Mandl, K. D. 2002. Use of geographical information from hospital databases for real time surveillance (abstract). *Pediatric Research* 51:94.

[Olson *et al.*, 2002] Olson, K.; Bonetti, M.; Pagano, M.; and Mandl, K. D. 2002. Enhanced power to detect bioterrorism with spatial clustering (abstract). *Pediatric Research* 51:132.

[O'Toole, 2001] O'Toole, T. 2001. The problem of biological weapons: Next steps for the nation. *Public Health Reports* 116(Suppl):108–111.

[Pollock & Lowery, 2001] Pollock, D., and Lowery, D. 2001. Emergency medicine and public health: New steps in old directions. *Annals of Emergency Medicine* 38:675–683.

[Primmerman, 2000] Primmerman, C. A. 2000. Detection of biological agents. *Lincoln Laboratory Journal* 12:3–31.

[Rabiner, 1989] Rabiner, L. R. 1989. A tutorial on Hidden Markov Models and selected applications in speech recognition. In *Proceedings of the IEEE*, volume 77, 257–285.

[Rath, Carreras, & Sebastiani, 2003] Rath, T.; Carreras, M.; and Sebastiani, P. 2003. Automated detection of influenza epidemics with Hidden Markov Models. In *Proceedings of the Fifth International Symposium on Intelligent Data Analysis*. New York: Springer. In press.

[Reis & Mandl, 2003] Reis, B. Y., and Mandl, K. D. 2003. Time series modeling for syndromic surveillance. *BMC Medical Informatics and Decision Making* 3:http://www.biomedcentral.com/1472–6947/1473/1472.

[Reis, Pagano, & Mandl, 2003] Reis, B. Y.; Pagano, M.; and Mandl, K. D. 2003. Using temporal context to improve biosurveillance. *Proceedings of the National Academy of Sciences of the USA* 100:1961–1965.

[Rogerson, 1997] Rogerson, P. A. 1997. Surveillance systems for monitoring the development of spatial patterns. *Statistics in Medicine* 16:2081.

[Serfling, 1963] Serfling, R. E. 1963. Methods for current statistical analysis of excess pneumonia-influenza deaths. *Public Health Reports* 78:494–506.

[Shumway & Stoffer, 2000] Shumway, R. H., and Stoffer, D. S. 2000. *Time Series Analysis and its Applications*. NY: Springer Verlag.

[Simenson *et al.*, 2000] Simenson, L.; Fukuda, K.; Schonberg, L. B.; and Cox, N. J. 2000. The impact of influenza epidemics on hospitalizations. *The Journal of Infectious Diseases* 181:831–837.

[Stern & Lightfoot, 1999] Stern, L., and Lightfoot, D. 1999. Automated outbreak detection: a quantitative retrospective analysis. *Epidemiology and Infection* 122:103–110.

[Teich *et al.*, 2002] Teich, J. M.; Wagner, M. M.; Mackenzie, C. F.; and Schafer, K. O. 2002. The informatics response in disaster, terrorism, and war. *Journal of the American Medical Informatics Association* 9:97–104.

[Tsui *et al.*, 2002] Tsui, F. C.; Espino, J. U.; Wagner, M. M.; Gesteland, P.; O., O. I.; Olszewski, R. T.; Liu, Z.; Zeng, X.; Chapman, W.; Wong, W. K.; and Moore, A. 2002. Data, network, and application: Technical description of the utah RODS winter olympic biosurveillance system. In *Proceedings of the Annual AMIA Fall Symposium*, 815–819.

[Tsui *et al.*, 2003] Tsui, F. C.; Espino, J. U.; Dato, V. M.; Gesteland, P. H.; Hutman, J.; and Wagner, M. M. 2003. Technical description of RODS: A real-time public health surveillance system. *Journal Journal of the American Medical Informatics Association*. To appear.

[U.S. Army Medical Research Institute of Infectious Diseases, 2001] U.S. Army Medical Research Institute of Infectious Diseases. 2001. *USAMRIID's Medical Management of Biological Casualties Handbook*. Fort Detrick, Maryland, USAMRIID.

[Wagner *et al.*, 2001] Wagner, M. M.; Tsui, F. C.; Espino, J. U.; Dato, V. M.; Sittig, D. F.; Caruana, R. A.; McGinnis, L. F.; Deerfield, D. W.; Druzdzel, M. J.; and Fridsma, D. B. 2001. The emerging science of very early detection of disease outbreaks. *Journal of Public Health Management and Practice* 7:51–59.

[Waller, 2000] Waller, L. A. 2000. A civil action and statistical assessments of the spatial pattern of disease: Do we have a cluster? *Regulatory Toxicology and Pharmacology* 32:174–183.

[Yasnoff *et al.*, 2001] Yasnoff, W. A.; Overhage, J. M.; Humphreys, B. L.; LaVenture, M.; Goodman, K. W.; Gatewood, L.; Ross, D. A.; Reid, J.; Hammond, W. E.; Dwyer, D.; Huff, S. M.; Gotham, I.; Kukafka, R.; Loonsk, J. W.; and Wagner, M. M. 2001. A national agenda for public health informatics. *Journal of Public Health Management and Practice.* 7:1–21.