



MMWRTM

Morbidity and Mortality Weekly Report

Recommendations and Reports

May 7, 2004 / Vol. 53 / No. RR-5

Framework for Evaluating Public Health Surveillance Systems for Early Detection of Outbreaks

Recommendations from the CDC Working Group

The *MMWR* series of publications is published by the Epidemiology Program Office, Centers for Disease Control and Prevention (CDC), U.S. Department of Health and Human Services, Atlanta, GA 30333.

SUGGESTED CITATION

Centers for Disease Control and Prevention. Framework for evaluating public health surveillance systems for early detection of outbreaks; recommendations from the CDC Working Group. *MMWR* 2004;53(No. RR-5):[inclusive page numbers].

Centers for Disease Control and Prevention

Julie L. Gerberding, M.D., M.P.H.
Director

Dixie E. Snider, Jr., M.D., M.P.H.
(Acting) Deputy Director for Public Health Science

Tanja Popovic, M.D., Ph.D.
(Acting) Associate Director for Science

Epidemiology Program Office

Stephen B. Thacker, M.D., M.Sc.
Director

Office of Scientific and Health Communications

John W. Ward, M.D.
Director
Editor, MMWR Series

Suzanne M. Hewitt, M.P.A.
Managing Editor, MMWR Series

C. Kay Smith-Akin, M.Ed.
Lead Technical Writer/Editor

David C. Johnson
Project Editor

Beverly J. Holland
Lead Visual Information Specialist

Lynda G. Cupell
Malbea A. LaPete
Visual Information Specialists

Kim L. Bright, M.B.A.
Quang M. Doan, M.B.A.
Erica R. Shaver
Information Technology Specialists

CONTENTS

Introduction	1
Background	2
Disease Case Reports	2
Pattern Recognition	2
New Data Types	2
Framework	3
A. System Description	3
B. Outbreak Detection	4
C. System Experience	8
D. Conclusions and Recommendations for Use and Improvement of Systems for Early Outbreak Detection ...	10
Acknowledgements	10
References	10
Appendix	12

Framework for Evaluating Public Health Surveillance Systems for Early Detection of Outbreaks

Recommendations from the CDC Working Group

Prepared by
James W. Buehler, M.D.,¹
Richard S. Hopkins, M.D.,²
J. Marc Overhage, M.D.,³
Daniel M. Sosin, M.D.,²
Van Tong, M.P.H.²

¹ Department of Epidemiology, Rollins School of Public Health, Emory University

² Division of Public Health Surveillance and Informatics, Epidemiology Program Office, CDC

³ Indiana University School of Medicine

Summary

The threat of terrorism and high-profile disease outbreaks has drawn attention to public health surveillance systems for early detection of outbreaks. State and local health departments are enhancing existing surveillance systems and developing new systems to better detect outbreaks through public health surveillance. However, information is limited about the usefulness of surveillance systems for outbreak detection or the best ways to support this function. This report supplements previous guidelines for evaluating public health surveillance systems. Use of this framework is intended to improve decision-making regarding the implementation of surveillance for outbreak detection. Use of a standardized evaluation methodology, including description of system design and operation, also will enhance the exchange of information regarding methods to improve early detection of outbreaks. The framework directs particular attention to the measurement of timeliness and validity for outbreak detection. The evaluation framework is designed to support assessment and description of all surveillance approaches to early detection, whether through traditional disease reporting, specialized analytic routines for aberration detection, or surveillance using early indicators of disease outbreaks, such as syndromic surveillance.

Introduction

Public health surveillance is the ongoing, systematic collection, analysis, interpretation, and dissemination of data about a health-related event for use in public health action to reduce morbidity and mortality and to improve health (1). Surveillance serves at least eight public health functions. These include supporting case detection and public health interventions, estimating the impact of a disease or injury, portraying the natural history of a health condition, determining the distribution and spread of illness, generating hypotheses and stimulating research, evaluating prevention and control measures, and facilitating planning (2). Another important public health function of surveillance is outbreak detection

(i.e., identifying an increase in frequency of disease above the background occurrence of the disease).

Outbreaks typically have been recognized either based on accumulated case reports of reportable diseases or by clinicians and laboratorians who alert public health officials about clusters of diseases. Because of the threat of terrorism and the increasing availability of electronic health data, enhancements are being made to existing surveillance systems, and new surveillance systems have been developed and implemented in public health jurisdictions with the goal of early and complete detection of outbreaks (3). The usefulness of surveillance systems for early detection and response to outbreaks has not been established, and substantial costs can be incurred in developing or enhancing and managing these surveillance systems and investigating false alarms (4). The measurement of the performance of public health surveillance systems for outbreak detection is needed to establish the relative value of different approaches and to provide information needed to improve their efficacy for detection of outbreaks at the earliest stages.

The material in this report originated in the Epidemiology Program Office, Stephen B. Thacker, M.D., Director, and the Division of Public Health Surveillance and Informatics, Daniel M. Sosin, M.D., Director.

This report supplements existing CDC guidelines for evaluating public health surveillance systems (1). Specifically, the report provides a framework to evaluate timeliness for outbreak detection and the balance among sensitivity, predictive value positive (PVP), and predictive value negative (PVN) for detecting outbreaks. This framework also encourages detailed description of system design and operations and of their experience with outbreak detection.

The framework is best applied to systems that have data to demonstrate the attributes of the system under consideration. Nonetheless, this framework also can be applied to systems that are in early stages of development or in the planning phase by using citations from the published literature to support conclusions. Ideally, the evaluation should compare the performance of the surveillance system under scrutiny to alternative surveillance systems and produce an assessment of the relative usefulness for early detection of outbreaks.

Background

Early detection of outbreaks can be achieved in three ways: 1) by timely and complete receipt, review, and investigation of disease case reports, including the prompt recognition and reporting to or consultation with health departments by physicians, health-care facilities, and laboratories consistent with disease reporting laws or regulations; 2) by improving the ability to recognize patterns indicative of a possible outbreak early in its course, such as through analytic tools that improve the predictive value of data at an early stage of an outbreak or by lowering the threshold for investigating possible outbreaks; and 3) through receipt of new types of data that can signify an outbreak earlier in its course. These new types of data might include health-care product purchases, absences from work or school, presenting symptoms to a health-care provider, or laboratory test orders (5).

Disease Case Reports

The foundation of communicable disease surveillance in the United States is the state and local application of the reportable disease surveillance system known as the National Notifiable Disease Surveillance System (NNDSS), which includes the listing of diseases and laboratory findings of public health interest, the publication of case definitions for their surveillance, and a system for passing case reports from local to state to CDC. This process occurs best where two-way communication occurs between public health agencies and the clinical community: clinicians and laboratories report cases and clusters of reportable and unusual diseases, and health departments consult on case diagnosis and management, alerts,

surveillance summaries, and clinical and public health recommendations and policies. Faster, more specific and affordable diagnostic methods and decision-support tools for diseases with substantial outbreak potential could improve the timely recognition of reportable diseases. On-going health-care provider and laboratory outreach, education, and 24-hour access to public health professionals are needed to enhance reporting of urgent health threats. Electronic laboratory reporting (i.e., the automated transfer of designated data from a laboratory database to a public health data repository using a defined message structure) also will improve the timeliness and completeness of reporting notifiable conditions (6–8) and can serve as a model for electronic reporting of a wider range of clinical information. A comprehensive surveillance effort supports timely investigation (i.e., tracking of cases once an outbreak has been recognized) and data needs for managing the public health response to an outbreak or terrorist event.

Pattern Recognition

Statistical tools for pattern recognition and aberration detection can be applied to screen data for patterns warranting further public health investigation and to enhance recognition of subtle or obscure outbreak patterns (9). Automated analysis and visualization tools can lessen the need for frequent and intensive manual analysis of surveillance data.

New Data Types

Many new surveillance systems, loosely termed syndromic surveillance systems, use data that are not diagnostic of a disease but that might indicate the early stages of an outbreak. The scope of this framework is broader than these novel systems, yet the wide-ranging definitions and expectations of syndromic surveillance require clarification. Syndromic surveillance for early outbreak detection is an investigational approach where health department staff, assisted by automated data acquisition and generation of statistical signals, monitor disease indicators continually (real-time) or at least daily (near real-time) to detect outbreaks of diseases earlier and more completely than might otherwise be possible with traditional public health methods (e.g., by reportable disease surveillance and telephone consultation). The distinguishing characteristic of syndromic surveillance is the use of indicator data types. For example, a laboratory is a data source that can support traditional disease case reporting by submitting reports of confirmatory laboratory results for notifiable conditions; however, test requests are a type of laboratory data that might be used as an outbreak indicator by tracking excess volume of test requests for diseases that typically cause outbreaks. New data types have been used by public health to enhance

surveillance, reflecting events that might precede a clinical diagnosis (e.g., patient's chief complaints in emergency departments, clinical impressions on ambulance log sheets, prescriptions filled, retail drug and product purchases, school or work absenteeism, and constellations of medical signs and symptoms in persons seen in various clinical settings).

Outbreak detection is the overriding purpose of syndromic surveillance for terrorism preparedness. Enhanced case-finding and monitoring the course and population characteristics of a recognized outbreak also are potential benefits of syndromic surveillance (4). A manual syndromic surveillance system was used to detect additional anthrax cases in the fall of 2001 when the outbreak was recognized (10). Complicating the understanding of syndromic surveillance is that syndromes have been used for case detection and management of diseases when the condition is infrequent and the syndrome is relatively specific for the condition of interest. Acute flaccid paralysis is a syndromic marker for poliomyelitis and is used to detect single cases of suspected polio in a timely way to initiate investigation and control measures. In this case, the syndrome is relatively uncommon and serious and serves as a proxy for polio (11). Syndromes also have been used effectively for surveillance in resource-poor settings for sexually transmitted disease detection and control where laboratory confirmation is not possible or practical (12). However, syndromic surveillance for terrorism is not intended for early detection of single cases or limited outbreaks because the early clinical manifestations of diseases that might be caused by terrorism are common and nonspecific (13).

Framework

This framework is intended to support the evaluation of all public health surveillance systems for the timely detection of outbreaks. The framework is organized into four categories: system description, outbreak detection, experience, and conclusions and recommendations. A comprehensive evaluation will address all four categories.

A. System Description

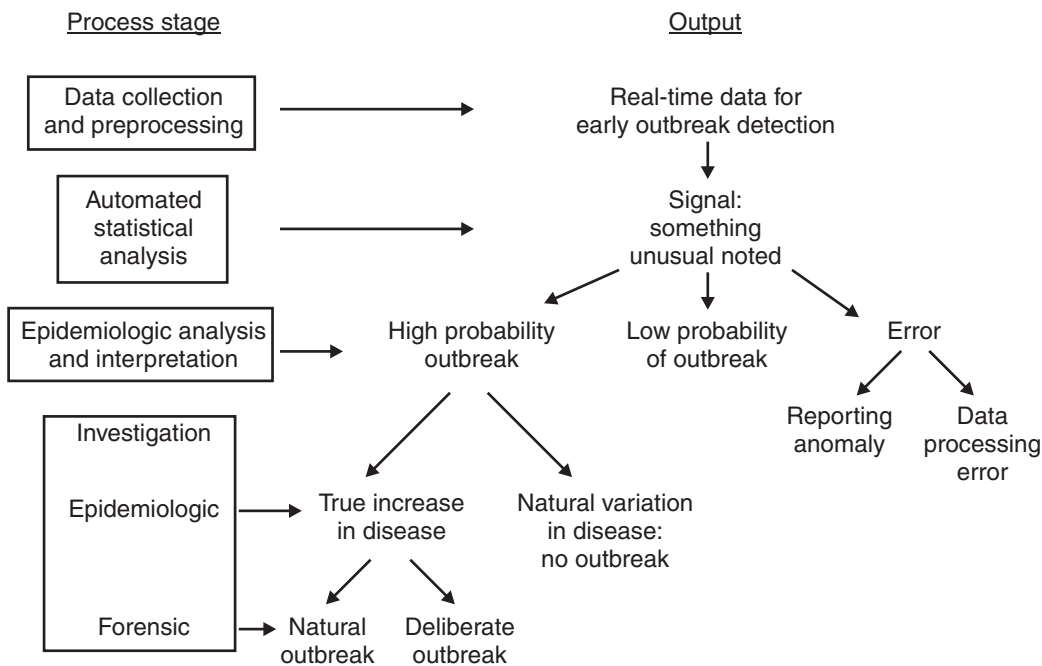
1. Purpose. The purpose(s) of the system should be explicitly and clearly described and should include the intended uses of the system. The evaluation methods might be prioritized differently for different purposes. For example, if terrorism is expected to be rare, reassurance might be the primary purpose of the terrorism surveillance system. However, for reassurance to be credible, negative results must be accurate and the system should have a demonstrated ability to detect outbreaks of the kind and size being dismissed.

The description of purpose should include the indications for implementing the system; whether the system is designed for short-term, high-risk situations or long-term, continuous use; the context in which the system operates (whether it stands alone or augments data from other surveillance systems); what type of outbreaks the system is intended to detect; and what secondary functional value is desired. Designers of the system should specify the desired sensitivity and specificity of the system and whether it is intended to capture small or large events.

2. Stakeholders. The stakeholders of the system should be listed. Stakeholders include those who provide data for the system and those who use the information generated by the system (e.g., public health practitioners; health-care providers; other health-related data providers; public safety officials; government officials at local, state, and federal levels; community residents; nongovernmental organizations; and commercial systems developers). The stakeholders might vary among different systems and might change as conditions change. Listing stakeholders helps define who the system is intended to serve and provides context for the evaluation results.

3. Operation. All aspects of the operation of the syndromic surveillance system should be described in detail to allow stakeholders to validate the description of the system and for other interested parties to understand the complexity and resources needed to operate such a system. Detailed system description also will facilitate evaluation by highlighting variations in system operation that are relevant to variations in system performance (Figure 1). Such a conceptual model can facilitate the description of the system. The description of the surveillance process should address 1) systemwide characteristics (data flow [Figure 2]), including data and transmission standards to facilitate interoperability and data sharing between information systems, security, privacy, and confidentiality; 2) data sources (used broadly in this framework to include the data-producing facility [i.e., the entity sharing data with the public health surveillance system], the data type [e.g., chief complaint, discharge diagnosis, laboratory test order], and the data format [e.g., electronic or paper, text descriptions of events or illnesses, or structured data reworded or stored in standardized format]); 3) data processing before analysis (the data collation, filtering, transformation, and routing functions required for public health to use the data, including the classification and assigning of syndromes); 4) statistical analysis (tools for automated screening of data for potential outbreaks); and 5) epidemiologic analysis, interpretation, and investigation (the rules, procedures, and tools that support decision-making in response to a system signal, including adequate staffing with trained epidemiologists who can review, explore, and interpret the data in a timely manner).

FIGURE 1. Process model for early outbreak detection



B. Outbreak Detection

The ability of a system to reliably detect an outbreak at the earliest possible stage depends on the timely capture and processing of the data produced by transactions of health behaviors (e.g., over-the-counter pharmaceutical sales, emergency department visits, and nurse call-line volume) or health-care activities (e.g., laboratory test volume and triage categorization of chief complaint) that might indicate an outbreak; the validity of the data for measuring the conditions of interest at the earliest stage of illness and the quality of those data; and the detection methods applied to these processed surveillance data to distinguish expected events from those indicative of an outbreak.

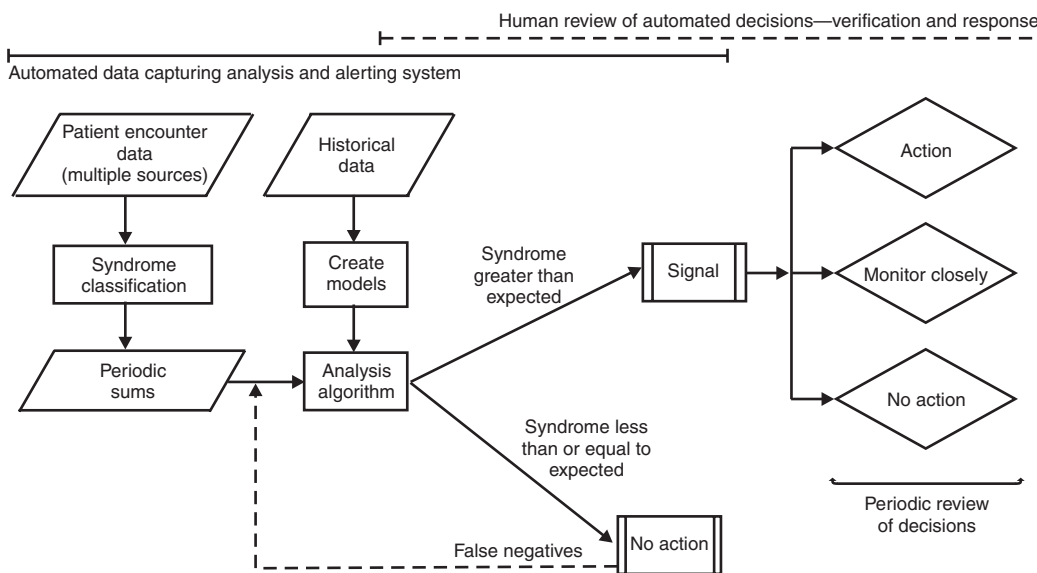
1. Timeliness. The timeliness of surveillance approaches for outbreak detection is measured by the lapse of time from exposure to the disease agent to the initiation of a public health intervention. A timeline with interim milestones is proposed to improve the specificity of timeliness measures (Figure 3). Although measuring all of the time points that define the intervals might be impractical or inexact in an applied outbreak setting, measuring intervals in a consistent way can be used to compare alternative outbreak-detection approaches and specific surveillance systems.

- **Onset of exposure:** By anchoring the timeline on exposure, the timeliness advantage of different data sources can be assessed and compared. Exposure can most easily be estimated in a point-source outbreak. Time of exposure is often inferred from knowledge of the agent (e.g., incubation period) and the epidemiology of the outbreak.

- **Onset of symptoms:** The interval to symptom onset in each case is defined by the incubation period for the agent. Time of symptom onset might be estimated using case interviews or existing knowledge of the agent and the time of exposure. The incubation period might vary according to host factors and the route and dose of the exposure.

- **Onset of behavior:** Following symptom onset, several health behaviors can occur (e.g., purchasing over-the-counter medication from a store, calling in sick to work, or visiting an urgent-care center). When an affected person interacts with the health-care system, a variety of health-care provider behaviors might be performed (e.g., order of a laboratory test and admission to a hospital). The selection of data sources for a system has a strong influence on timeliness. Some of those experiencing symptoms will initiate a health behavior or stimulate a health-care provider behavior that is a necessary step to being captured in the surveillance system.
- **Capture of data:** The timing of the capture of a behavior by the data-providing facility varies by data type and can be influenced by system design. A retail purchase might be entered in an electronic database at the moment the transaction is completed, or a record might not be generated in a clinical setting until hours after health care was sought.
- **Completion of data processing:** Time is required for the facility providing the data to process the data and produce the files needed for public health. Records might be transmitted to a central repository only periodically (e.g., weekly). Data form can influence processing time (e.g., transcription from paper to electronic form and coding text-based data), and data manipulations needed to de-identify data and prepare necessary files can affect processing time.

FIGURE 2. Prototypical surveillance data flow chart for emergency department encounters



• **Initiation of public health intervention:** When an outbreak of public health significance is confirmed, interventions can be implemented to control the severity of disease and prevent further spread. Interventions might be of a general nature directed to the recognition of an outbreak (e.g., apply respiratory infection precautions and obtain clinical specimens for diagnosis) or can be specific to the diagnosis (e.g., antibiotic prophylaxis or vaccination).

2. Validity. Measuring the validity of a system for outbreak detection requires an operational definition of an outbreak. Although a statistical deviation

• **Capture of data in public health surveillance system:**

The time required to transfer data from the data providing facility to the public health entity varies according to the frequency established for routine data transmission and by the data transmission method (e.g., Internet, mail, or courier).

• **Application of pattern recognition tools/algorithms:**

Before analytic tools can be applied to the data in the surveillance system, certain processing steps are necessary (e.g., categorization into syndrome categories, application of case definition, and data transformations).

• **Generation of automated alert:** The detection algorithm's alerting interval is a product of how often the algorithm is run and a report generated and the capacity of the algorithm to filter noise and detect an aberration as early as possible in the course of the outbreak.

• **Initiation of public health investigation:** The initiation of a public health investigation occurs when a decision is made to acquire additional data. Analysis and judgment are applied by public health-care providers to the processed surveillance data and other available information to decide whether new data collection is warranted to confirm the existence of an outbreak. The challenge of interpreting data from multiple surveillance systems could diminish potential advantages in timeliness. The focus on outbreak detection allows for investigations of potential outbreaks to proceed before a specific clinical diagnosis is obtained.

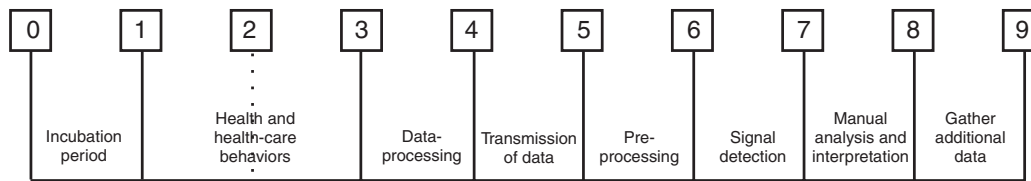
from a baseline rate can be useful for triggering further investigation, it is not sufficient for defining an outbreak. In practice, the confirmation of an outbreak is a judgment that depends on past experience with the condition, the severity of the condition, the communicability of the condition, confidence in the diagnosis of the condition, public health concern about outbreaks at the time, having options for effective prevention or control, and the resources required and available to respond. Operationally, an outbreak is defined by the affected public health jurisdiction when the occurrence of a condition has changed sufficiently to warrant public health attention.

The validity of a surveillance system for outbreak detection varies according to the outbreak scenario and surveillance system factors. These factors can confound the comparison of systems and must be carefully described in the evaluation. For example, the minimum size of an outbreak that can be detected by a system cannot be objectively compared among systems unless they are identical or differences are accounted for in several ways.

• **Case definitions:** Establish the specificity and sensitivity for the condition of interest on the basis of the data source, data type, and response criteria.

• **Baseline estimation:** Determine the stability of the background occurrence of cases. Estimations are affected by factors such as population size and geographic distribution. The performance of detection algorithms will vary by the quality and duration and inherent variability of baseline data.

FIGURE 3. Timeline milestones for early outbreak detection



0. Onset of exposure
1. Onset of symptoms
2. Onset of behavior
3. Capture of data
4. Completion of data processing
5. Capture of data in surveillance system
6. Application of pattern recognition tools/algorithms
7. Generation of automated alert
8. Initiation of public health investigation
9. Initiation of public health intervention

- **Reporting delays:** Result in incomplete data, introducing bias that will diminish the performance of detection algorithms.
- **Data characteristics:** Includes underlying patterns in the data (e.g., seasonal variation) and systematic errors inherent in the data (e.g., product sales that influence purchasing behaviors unrelated to illness).
- **Outbreak characteristics:** Results from agent, host, and environmental factors that affect the epidemiology of the outbreak. For example, a large aerosol exposure with an agent causing serious disease in a highly susceptible population will have different detection potential than an outbreak of similar size spread person-to-person over a longer time and dispersed distribution.
- **Statistical analysis:** Defines how data are screened for outbreak detection. Detection algorithms have different performance characteristics under different outbreak conditions.
- **Epidemiologic analysis, interpretation, and investigation:** The procedures, resources, and tools for analysis, interpretation, and response that can substantially affect the ability to detect and respond to outbreaks.

Validation Approaches

Different approaches to outbreak detection need to be evaluated under the same conditions to isolate the unique features of the system (e.g., data type) from the outbreak characteristics and the health department capacity. The data needed to evaluate and compare the performance of surveillance systems for early outbreak detection can be obtained from naturally occurring outbreaks or through simulation.

Controlled comparisons of surveillance systems for detection of deliberately induced outbreaks will be difficult because of the infrequency of such outbreaks and the diversity of systems and outbreak settings. However, understanding the value of different surveillance approaches to early detection will increase as descriptions of their experience with detecting and missing naturally occurring outbreaks accumulate. Accumulation of experience descriptions is made more difficult by not having standard methods for measuring outbreak detection successes and

failures across systems and by the diversity of surveillance system and outbreak factors that influence performance. Standardized classification of system and outbreak factors will enable comparison of experiences across systems. Pending the development of classification standards, descriptive evaluation should include as much detail as possible. Proxy outbreak scenarios reflect the types of naturally occurring outbreaks that should not be missed to instill confidence in the ability of these systems to detect outbreaks caused by terrorism. Examples of proxy events or outbreaks include seasonal events (e.g., increases in influenza, norovirus gastroenteritis, and other infectious respiratory agents) and community outbreaks (e.g., foodborne, waterborne, hepatitis A, child-care-associated shigellosis, legionellosis, and coccidioidomycosis and histoplasmosis in areas where the diseases are endemic).

The measurement of outbreaks detected, false alarms, and outbreaks missed or detected late should be designed as a routine part any system workflow and conducted with minimal effort or complexity. Routine reporting should be automated where possible. Relevant information needs include: the number of statistical aberrations detected at a set threshold in a defined period of time (e.g., frequency per month at a given p-value); the action taken as a result of the signals (e.g., review for data errors, in-depth follow-up analysis of the specific conditions within the syndrome category, manual epidemiologic analysis to characterize a signal, examining data from other systems, and increasing the frequency of reporting from affected sites); resources directed to the follow-up of the alert; public health response that resulted (e.g., an alert to clinicians, timely dissemination of information to other health entities, a vaccination campaign, or no further

response); documentation of how every recognized outbreak in the jurisdiction was detected; an assessment of the value of the follow-up effort (e.g., the effort was an appropriate application of public health resources); a detailed description of the agent, host, and environmental conditions of the outbreak; and the number of outbreaks detected only late in their course or in retrospect.

To evaluate the relative value of different methods for outbreak detection, a direct comparison approach is needed. For example, if a health department detects a substantial number of its outbreaks through telephone consultations, then a phone call tracking system might produce the data needed to compare telephone consults with other approaches for early detection of outbreaks.

As an alternative to naturally occurring outbreaks, simulations can allow for the control and modification of agent, host, and environmental factors to study system performance across a range of common scenarios. However, simulations are limited in their ability to mimic the diversity and unpredictability of real-life events. Whenever possible, simulated outbreaks should be superimposed on historical trend data. To evaluate detection algorithms comparatively, a shared challenge problem and data set would be helpful. Simulation is limited by the availability of well-documented outbreak scenarios (e.g., organism or agent characteristics, transmission characteristics, and population characteristics). Simulations should incorporate data for each of the factors described previously. Multiple simulation runs should be used to test algorithm performance in different outbreak scenarios, allowing for generation of operating characteristic curves that reflect performance in a range of conditions.

Focused studies to validate the performance of limited aspects of systems (e.g., data sources, case definitions, statistical methods, and timeliness of reporting) can provide indirect evidence of system performance. Component studies also can test assumptions about outbreak scenarios and support better data simulation. Syndrome case definitions for certain specific data sources need to be validated. Component validation studies should emphasize outbreak detection over case detection. These studies contain explicit hypotheses and research questions and should be shared in a manner to advance the development of outbreak detection systems without unnecessary duplication.

Statistical Assessment of Validity

Surveillance systems must balance the risk for an outbreak, the value of early intervention, and the finite resources for investigation. Perceived high risk and high value of timely detection support high sensitivity and low thresholds for investigation. A low threshold can prompt resource-intensive

investigations and occupy vital staff, and a high threshold might delay detection and intervention. The perceived threat of an outbreak, the community value attached to early detection, and the investigation resources available might vary over time. As a result, specifying a fixed relation between optimal sensitivity and predictive value for purposes of evaluation might be difficult.

The sensitivity and PVP and PVN are closely linked and considered together in this framework. Sensitivity is the percentage of outbreaks occurring in the jurisdiction detected by the system. PVP reflects the probability of a system signal being an outbreak. PVN reflects the probability that no outbreak is occurring when the system does not yield a signal. The calculation of sensitivity and predictive value is described in detail in the updated guidelines for evaluating public health surveillance systems (1). Measurement of sensitivity requires an alternative data source of high quality (e.g., “gold” standard) to confirm outbreaks in the population that were missed by the surveillance system. Sensitivity for outbreak detection could be assessed through capture-recapture techniques with two independent data sources (14). The high costs associated with responding to false alarms and with delayed response to outbreaks demand efforts to quantify and limit the impact of both. As long as the likelihood of terrorism is extremely low, PVP will remain near zero and a certain level of nonterrorism signals will be a necessary part of conducting surveillance for the detection of terrorism. Better performance can be achieved in one attribute (e.g., sensitivity) without a performance decrement in another (e.g., PVP) by changing the system (e.g., adding a data type or applying a better detection algorithm). Improving sensitivity by lowering the cut-off for signaling an outbreak will reduce PVP. Sensitivity and PVP for these surveillance systems will ultimately be calibrated in each system to balance the secondary benefits (e.g., detection of naturally occurring outbreaks, disease case finding and management, reassurance of no outbreak during periods of heightened risk, and a stronger reporting and consultation relation between public health and clinical medicine) with the locally acceptable level of false alarms.

Data Quality

The validity of syndromic surveillance system data is dependent on data quality. Error-prone systems and data prone to inaccurate measurement can negatively affect detection of unusual trends. Although data quality might be a less critical problem for screening common, nonspecific indicators for statistical aberrations, quality should be evaluated and improved to the extent possible. Measuring data quality is dependent on a standard (e.g., medical record review or fabricated test data with values known to the evaluator). The updated

guidelines for evaluating public health surveillance systems (1) describe data quality in additional detail.

- **Representativeness:** When case ascertainment within a population is incomplete (e.g., in a sentinel system or a statistically based sample), representativeness reflects whether a system accurately describes the distribution of cases by time, place, and person. Geographic representativeness is particularly important for detecting outbreaks of infectious diseases.
- **Completeness of data:** The frequency of unknown or blank responses to data items in the system can be used to measure the level of completeness. For systems that update data from previous transmissions, time should be factored into measurement by indicating the percentage of records that are complete (i.e., all variables are captured for a record) on initial report and within an appropriate interval (e.g., 48 hours) of submission. Sites with substantial reporting delays can be flagged for reliability concerns and targeted for improvement. Incomplete data can require follow-up before analysis, with associated decreases in timeliness and increase in cost. When multiple data providers contribute to a common data store for statistical analysis, the percentage of reporting sources that submit their data on a routine interval (e.g., every 24 hours) conveys the completeness of the aggregate database for routine analysis. Evaluation of completeness should include a description of the problems experienced with manual data management (e.g., coding errors or loss of data) and the problems with automated data management (e.g., programming errors or inappropriate filtering of data).

C. System Experience

The performance attributes described in this section convey the experience that has accrued in using the system.

1. System usefulness. A surveillance system is useful for outbreak detection depending on its contribution to the early detection of outbreaks of public health significance that leads to an effective intervention. An assessment of usefulness goes beyond detection to address the impact or value added by its application. Measurement of usefulness is inexact. As with validity, measurement will benefit from common terminology and standard data elements. In the interim, detailed efforts to describe and illustrate the consequences of early detection efforts will improve understanding of their usefulness.

Evaluation should begin with a review of the objectives of the system and should consider the priorities. To the extent possible, usefulness should be described by the disease

prevention and control actions taken as a result of the analysis and interpretation of the data from the system.

The impact of the surveillance system should be contrasted with other mechanisms available for outbreak detection. An assessment of usefulness should list the outbreaks detected and the role that different methods played in the identification of each one. Examples of how the system has been used to detect or track health problems other than outbreaks in the community should be included. The public health response to the outbreaks and health problems detected should be described as well as how data from new or modified surveillance systems support inferences about disease patterns that would not be possible without them.

Surveillance systems for early outbreak detection are sometimes justified for the reassurance they provide when aberrant patterns are not apparent during a heightened risk period or when the incidence of cases declines during an outbreak. When community reassurance is claimed as a benefit of the surveillance system, reassurance should be defined and the measurement quantified (e.g., number of phone calls from the public on a health department hotline, successful press conferences, satisfaction of public health decision-makers, or resources to institutionalize the new surveillance system). A description should include who is reassured and of what they are reassured, and reassurance should be evaluated for validity by estimating the PVN.

2. Flexibility. The flexibility of a surveillance system refers to the system's ability to change as needs change. The adaptation to changing detection needs or operating conditions should occur with minimal additional time, personnel, or other resources. Flexibility generally improves the more data processing is handled centrally rather than distributed to individual data-providing facilities because fewer system and operator behavior changes are needed. Flexibility should address the ability of the system to apply evolving data standards and code sets as reflected in Public Health Information Network (PHIN) standards (<http://www.cdc.gov/phn>). Flexibility includes the adaptability of the system to shift from outbreak detection to outbreak management. The flexibility of the system to meet changing detection needs can include the ability to add unique data to refine signal detection, to capture exposure and other data relevant to managing an outbreak, to add data providers to increase population coverage and detect or track low frequency events, to modify case definitions (the aggregation of codes into syndrome groupings), to improve the detection algorithm to filter random variations in trends more efficiently, and to adjust the detection threshold. Flexibility also can be reflected by the ability of the system to detect and monitor naturally occurring outbreaks in the absence of terrorism. System flexibility is needed to

balance the risk for an outbreak, the value of early intervention, and the resources for investigation as understanding of these factors changes.

3. System acceptability. As with the routine evaluation of public health surveillance systems (1), the acceptability of a surveillance system for early outbreak detection is reflected by the willingness of participants and stakeholders to contribute to the data collection and analysis. This concept includes the authority and willingness to share electronic health data and should include an assessment of the legal basis for the collection of prediagnosis data and the implications of privacy laws (e.g., Health Insurance Portability and Accountability Act Privacy Rule) (15). All states have broad disease-reporting laws that require reporting of diseases of public health importance, and many of these laws appear compatible with the authority to receive syndromic surveillance data (16). The authority to require reporting of indicator data for persons who lack evidence of a reportable condition and in the absence of an emergency is less clear and needs to be verified by jurisdictions. Acceptability can vary over time as the threat level, perceived value of early detection, support for the methods of surveillance, and resources fluctuate.

Acceptability of a system can be inferred from the extent of its adoption. Acceptability is reflected by the participation rate of potential reporting sources, by the completeness of data reporting, and by the timeliness of person-dependent steps in the system (e.g., manual data entry from emergency department logs as distinguished from electronic data from the normal clinical workflow).

4. Portability. The portability of a surveillance system addresses how well the system could be duplicated in another setting. Adherence to the PHIN standards can enhance portability by reducing variability in the application of information technology between sites. Reliance on person-dependent steps, including judgment and action criteria (e.g., for analysis and interpretation) should be fully documented to improve system portability. Portability also is influenced by the simplicity of the system. Examples should be provided of the deployment of similar systems in other settings, and the experience of those efforts should be described. In the absence of examples, features of the system that might support or detract from portability should be described.

5. System stability. The stability of a surveillance system refers to its resilience to system changes (e.g., change in coding from *International Classifications of Disease, Ninth Revision* [ICD-9] to ICD-10). Stability can be demonstrated by the duration and consistent operation of the system. System stability is distinguished from the reliability of data elements within the system. The consistent representation of the

condition under surveillance (reliability) is an aspect of data quality. Stability can be measured by the frequency of system outages or downtime for servicing during periods of need, including downtime of data providers, the frequency of personnel deficiencies from staff turnover, and budget constraints. Ongoing support by system designers and evolving software updates might improve system stability. Stability also can be reflected in the extent of control over costs and system changes that the sponsoring agency maintains.

6. System costs. Cost is a vital factor in assessing the relative value of surveillance for terrorism preparedness. Cost-effectiveness analyses and data modeling are needed under a range of scenarios to estimate the value of innovations in surveillance for outbreak detection and terrorism preparedness (17). Improved methods of measuring cost and impact are needed. Costs borne by data providers should be noted; however, the cost perspective should be that of the community (societal perspective) to account for costs of prevention and treatment born by the community.

Direct costs include the fees paid for software and data, the personnel salary and support expenses (e.g., training, equipment support, and travel), and other resources needed to operate the system and produce information for public health decisions (e.g. office supplies, Internet and telephone lines, and other communication equipment). Fixed costs for running the system should be differentiated from the variable costs of responding to system alarms. Variable costs include the cost of follow-up activities (e.g., for diagnosis, case-management, or community interventions). The cost of responding to false alarms represents a variable but inherent inefficiency of an early detection system that should be accounted for in the evaluation. Similarly, variable costs include the financial and public health costs of missing outbreaks entirely or recognizing them late. Costs vary because the sensitivity and timeliness of the detection methods can be modified according to changes in tolerance for missing outbreaks and for responding to false alarms. Similarly, the threshold and methods for investigating system alarms can vary with the perceived risk and need to respond. Costs from public health response to false alarms with traditional surveillance systems need to be measured in a comparable way when assessing the relative value of new surveillance methods. Cost savings should be estimated by assessing the impact of prevention and control efforts (e.g., health-care costs and productivity losses averted) Questions to answer include the following:

- How many investigations were initiated as a result of these data?
- What response was made and what cost was incurred through follow-up of flagged events?

- What were the indications for responding?
- How much staff time was required for follow-up?
- Was anxiety raised unnecessarily by false alarms?
- Was benefit obtained (e.g., through improved communication and confidence in the responsibility and performance of public health) when false alarms were investigated?
- Who was affected?
- What costs did partners incur in follow-up of signals (e.g., medical record staff work and clinical staff efforts)? Follow-up costs for false alarms should be distinguished from costs related to investigations that uncover real outbreaks that warrant a public health response.
- Did the health department fail to respond to a true event because of complacency or the response burden resulting from false alarms?
- Did late recognition of an outbreak result in unnecessary morbidity?
- Have lessons learned from earlier events reduced costs as the system has continued to operate?

D. Conclusions and Recommendations for Use and Improvement of Systems for Early Outbreak Detection

The evaluation should be summarized to convey the strengths and weaknesses of the system under scrutiny. Summarizing and reporting evaluation findings should facilitate the comparison of systems for those making decisions about new or existing surveillance methods. These conclusions should be validated among stakeholders of the system and modified accordingly. Recommendations should address adoption, continuation, or modification of the surveillance system so that it can better achieve its intended purposes. Recommendations should be disseminated widely and actively interpreted for all appropriate audiences.

An Institute of Medicine study concluded that although innovative surveillance methods might be increasingly helpful in the detection and monitoring of outbreaks, a balance is needed between strengthening proven approaches (e.g., diagnosis of infectious illness and strengthening the liaison between clinical-care providers and health departments) and the exploration and evaluation of new approaches (17). Guidance for the evaluation of surveillance systems for outbreak detection is on-going. Many advances are needed in understanding of systems and outbreak characteristics to improve performance metrics. For example, research is needed to understand the personal health and clinical health care

behaviors that might serve as early indicators of priority diseases; analytic methods are needed to improve pattern recognition and to integrate multiple streams of data; a shared vocabulary is needed for describing outbreak conditions, managing text-based information, and supporting case definitions; and evaluation research is needed, including cost-effectiveness of different surveillance models for early detection, both in real-life comparisons and in simulated data environments, to characterize the size and nature of epidemics that can be detected through innovative surveillance approaches. Pending more robust measures of system performance, the goal of this framework is to improve public health surveillance systems for early outbreak detection by providing practical guidance for evaluation.

Acknowledgements

This report includes contributions by Daniel S. Budnitz, M.D., National Center for Injury Prevention and Control; Richard L. Ehrenberg, M.D., National Institute for Occupational Safety and Health; Timothy Doyle, Robert R. German, Dr.P.H., Timothy A. Green, Ph.D., Samuel L. Groseclose, D.V.M., Division of Public Health Surveillance and Informatics, Denise Koo, M.D., Division of Applied Public Health Training, Carol A. Pertowski, M.D., Stephen B. Thacker M.D., Epidemiology Program Office; José G. Rigau-Pérez, M.D., National Center for Infectious Diseases, CDC, Atlanta, Georgia. Melvin A. Kohn, M.D., State Epidemiologist, Oregon Department of Human Services, Portland. Steven C. Macdonald, Ph.D., Office of Epidemiology, Washington State Department of Health, Olympia. Nkuchia M. M'ikanatha, Dr.P.H., Pennsylvania Department of Health and Department, Harrisburg. Kelly Henning, M.D., New York City Department of Health and Mental Hygiene. Dan E. Peterson, M.D., Cereplex, Inc., Gaithersburg, Maryland. Michael L. Popovich, Scientific Technologies Corporation, Tucson, Arizona. Scott F. Wetterhall, M.D., DeKalb County Board of Health, Decatur, Georgia. Christopher W. Woods, M.D., Division of Infectious Diseases, Duke University Medical Center, Durham, North Carolina.

References

1. CDC. Updated guidelines for evaluating public health surveillance systems: recommendations from the guidelines working group. *MMWR* 2001;50(No. RR-13).
2. Teutsch SM, Churchill RE. Principles and practice of public health surveillance. 2nd ed. Oxford, New York: Oxford University Press, 2000.
3. Lober WB, Karras BT, Wagner MM, et al. Roundtable on bioterrorism detection: information system-based surveillance. *J Am Med Inform Assoc* 2002;9:105-15.
4. Reingold A. If syndromic surveillance is the answer, what is the question? *Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science* 2003;1:1-5.
5. Wagner MM, Tsui FC, Espino JU, et al. The emerging science of very early detection of disease outbreaks. *J Pub Health Mgmt Pract* 2001;6:51-9.

6. Effler P, Ching-Lee M, Bogard A, Jeong MC, Nekomoto T, Jernigan D. Statewide system of electronic notifiable disease reporting from clinical laboratories: comparing automated reporting with conventional methods. *JAMA* 1999;282:1845–50.
7. Panackal AA, M'ikanatha NM, Tsui FC, et al. Automatic electronic laboratory-based reporting of notifiable infectious diseases at a large health system. *Emerg Infect Dis.* 2002;8:685–91.
8. Hoffman MA, Wilkinson TH, Bush A, et al. Multijurisdictional approach to biosurveillance, Kansas City. *Emerg Infect Dis.* 2003;9:1281–6.
9. Hutwagner L, Thompson W, Seeman GM, Treadwell, T. The bioterrorism preparedness and response early aberration reporting system (EARS). *J Urban Health* 2003;80:89–96.
10. Tan CG, Sandhu HS, Crawford DC, et al. Surveillance for anthrax cases associated with contaminated letters, New Jersey, Delaware, and Pennsylvania, 2001. *Emerg Infect Dis.* 2002;8:1073–6.
11. Robertson SE, Suleiman AJM, Mehta FR, Al-Dahry SHS, El-Bualy MS. Poliomyelitis in Oman: acute flaccid paralysis surveillance leading to early detection and rapid response to a type 3 outbreak. *Bull WHO* 1994;72:907–14.
12. Grosskurth H, Mosha F, Todd J, et al. Impact of improved treatment of sexually transmitted diseases on HIV infection in rural Tanzania: randomized controlled trial. *Lancet* 1995;346:530–6.
13. Zeng X, Wagner M. Modeling the effects of epidemics on routinely collected data. *Proc AMIA Annu Symp* 2001:781–5.
14. Bishop YNM, Fienberg SE, Holland PW. Estimating the size of a closed population. In: *Discrete multivariate analysis: theory and practice.* Cambridge, MA: MIT Press, 1975.
15. CDC. HIPAA privacy rule and public health: guidance from CDC and the U.S. Department of Health and Human Services. *MMWR Supplement* 2003;52:1–12.
16. Broome CV, Horton HH, Tress D, Lucido SJ, Koo D. Statutory basis for public health reporting beyond specific diseases. *J Urban Health* 2003;80:14–22.
17. Smolinski MS, Hamburg MA, Lederberg J. *Microbial threats to health: emergence, detection, and response.* Washington, DC: National Academies Press, 2003.

CDC Evaluation Working Group on Public Health Surveillance Systems For Early Detection of Outbreaks

Chair: Dan Sosin, M.D., Division of Public Health Surveillance and Informatics, Epidemiology Program Office, CDC.

Members: Claire Broome, M.D., Office of the Director; Richard Hopkins, M.D., Henry Rolka, M.S., Division of Public Health Surveillance and Informatics; Van Tong, M.P.H., Epidemiology Program Office, CDC Atlanta, Georgia. James W. Buehler, M.D., Dept of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, Georgia. Louise Gresham, Ph.D., San Diego Health and Human Services, Public Health Services, San Deigo, California. Ken Kleinman, Sc.D., Harvard Pilgrim Health Care, Cambridge, Massachusetts. Farzad Mostashari, M.D., New York City Department of Health and Mental Hygiene, New York. J. Marc Overhage, M.D., Indiana University School of Medicine, Indianapolis. Julie Pavlin, M.D., Division of Preventive Medicine, Walter Reed Army Institute of Research, Silver Springs, Maryland. Robert Rolfs, M.D., Utah Department of Health, Salt Lake City. David Siegrist, M.S., Potomac Institute for Policy Studies, Arlington, Virginia.

Appendix. Operations Checklist

A. System-wide Issues

- Describe the political, administrative, and geographic context for the system
- Provide a process model that describes the data flow of the system:
 - Who inputs the data into the system
 - Who can view the data
 - Who can manipulate the data
 - Indicate where processing occurs centrally and where at distributed sites
 - Indicate where steps are automated and where manual
 - Indicate which steps are managed on-site and which can be done remotely
 - Estimate the time required for each step of the data flow
 - Indicate whether source data are produced in the course of routine workflow or specifically for the purpose of syndromic surveillance
- Describe data and messaging standards:
 - Identify standards used to facilitate interoperability
 - Identify standards used to facilitate data sharing
 - Describe how the system interfaces with other surveillance systems from the same sites to limit reporting burden
 - Cite relevant PHIN standards and ability to meet them (<http://www.cdc.gov/phin/>)
 - Provide legal documentation allowing data sharing
- Describe procedures to maintain security:
 - Indicate security procedures employed for transmission of data between sites and for data management at the central repository
 - Describe security measures to protect data integrity at the central repository
 - Cite relevant PHIN standards and ability to meet them (<http://www.cdc.gov/phin/>)
- Describe procedures to assure privacy and confidentiality:
 - Identify the legal authority under which the surveillance activity is being conducted
 - Indicate the rules, procedures, and tools used to assure privacy and confidentiality, including methods for de-identification and re-identification, if used, and the points in the data flow where statistical disclosure limitation methods are applied

B. Data Sources

- Describe the following:
 - Data producing facility
 - Data type
 - Data format
 - Data element definitions
 - Code sets (e.g., International Classification of Diseases (ICD) codes) used to describe the response categories
 - Data captured for geographic location (e.g., zipcode, geocode)
- Provide a data model describing the relationship between data elements and the code sets (The architecture of the National Electronic Disease Surveillance System (NEDSS <http://www.cdc.gov/nedss/>) and the Public Health Conceptual Data Model (<http://www.cdc.gov/nedss/DataModels/phcdm.pdf>) can serve as illustrations of comprehensive data models.)
 - Indicate which data standards are used and whether they are proprietary
 - Identify the standards used for assembling data documentation (i.e., metadata)

C. Data Preprocessing

- Indicate the steps taken to share data between information systems and indicate the responsible organization for assuring each step (e.g., clinical facility, data clearinghouse, local health department, state health department)
- Indicate the frequency of data collection
- Indicate the volume of data (e.g., average number of records per day)
- Indicate how the accumulation of data is handled
- Describe how different data streams or data elements are assembled, subset, and manipulated to prepare them for analysis
- Indicate whether a relational database is formed to link datasets and the unique identifier(s) used for linkage
- Indicate the health-related events, syndromes, or constellation of findings under surveillance, including the derivation of the case-definitions
- Identify who has authority to determine the criteria for case definitions and how case criteria are applied to the data
- Provide a description of any algorithms used to establish the status of a potential case

- Indicate the frequency of editing and updating the electronic file
- Indicate how incomplete records are handled in analysis and reports
- Describe how data archiving and disposal is managed
- Describe how new data sources or necessary changes in data sources are identified and incorporated in the system.

D. Statistical Analysis

- Describe how the health outcome baseline is established:
 - Describe the population under surveillance
 - Describe the source, the criteria, and the methods for establishing the background frequencies used to detect aberrations
 - How much baseline data are managed in the analysis database
- Describe analytic methods used in automated analyses (i.e., aberration detection):
 - Describe in mathematical and statistical detail the algorithms intended to signal an event requiring further investigation
 - Describe adaptations in analytic methods to account for different outbreak patterns that might be anticipated in different data sources and types and for different outbreak scenarios
 - Indicate how reporting delays are corrected for in the analysis.
 - Describe the method of adjusting results for potential confounding factors
 - Describe how the system adapts over time and the empirical basis for modifications in the methods
- Describe the detection process:
 - The frequency of data analysis
 - How an alarm is generated
 - Where the alarm goes
 - The type of alarms generated by the system
 - What is done to ensure that signals are not being missed
- Describe the report generation process:
 - What routine reports are generated
 - Whether data are presented graphically or in tables
 - Whether data can be manipulated to get a specific table/chart
 - How often charts and tables are refreshed with new data
- Indicate training level of personnel needed to manage the detection methods.

E. Epidemiologic Analysis, Interpretation, and Investigation

- Describe the process for managing system alarms:
 - Describe the special procedures instituted when the alarm is generated (e.g., review for data errors, in-depth manual analysis of the specific conditions within the syndrome category, manual epidemiological analysis to identify subgroups responsible for an alarm, examining data from other systems, increasing the frequency of reporting from affected sites)
 - Estimate the person-hours that are devoted to review and analysis each day and the interval at which data are analyzed
 - Indicate documented procedures for managing system alarms.
 - Indicate communication method that staff is alerted of alarms (e.g., whether they get paged at home, receive an automated e-mail, etc.?)
 - Indicate the expectations and schedule of staff to actively check the system and schedule, including nights and weekends
 - Indicate the response options to an alarm and the factors that influence the choice (e.g., wait for an alarm in another system, initiate an onsite investigation, alert clinicians to gather information)
- Describe the process for identifying cases for investigation when the data analyzed routinely are unidentified
- Describe how independent data types are integrated in the analysis for improved decision making
- Describe the rules, procedures, and tools for communication
 - Indicate the mechanisms used and content guidance provided for sharing results with 1) reporting sources, 2) response community, and 3) the public;
 - Describe how decisions are made for sending urgent communications and the methods for sending urgent communications
 - Indicate whether receipt of a communication is acknowledged and how unacknowledged receipt is managed
 - Indicate how often urgent communications and routine reports are sent
- Describe the protocol for conducting surveillance during outbreak management, if one exists
 - Indicate how often data will be updated and analyzed
 - Describe how the system can be modified or customized to meet special data needs
- Describe how the system will monitor the impact of prevention and control measures
 - Describe how and how often system components are tested for operational readiness (e.g., 'spiked' data or modeling exercises)

The *Morbidity and Mortality Weekly Report (MMWR)* Series is prepared by the Centers for Disease Control and Prevention (CDC) and is available free of charge in electronic format and on a paid subscription basis for paper copy. To receive an electronic copy each week, send an e-mail message to listserv@listserv.cdc.gov. The body content should read *SUBscribe mmwr-toc*. Electronic copy also is available from CDC's World-Wide Web server at <http://www.cdc.gov/mmwr> or from CDC's file transfer protocol server at <ftp://ftp.cdc.gov/pub/publications/mmwr>. To subscribe for paper copy, contact Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402; telephone 202-512-1800.

Data in the weekly *MMWR* are provisional, based on weekly reports to CDC by state health departments. The reporting week concludes at close of business on Friday; compiled data on a national basis are officially released to the public on the following Friday. Address inquiries about the *MMWR* Series, including material to be considered for publication, to Editor, *MMWR* Series, Mailstop C-08, CDC, 1600 Clifton Rd., N.E., Atlanta, GA 30333; telephone 888-232-3228.

All material in the *MMWR* Series is in the public domain and may be used and reprinted without permission; citation as to source, however, is appreciated.

All *MMWR* references are available on the Internet at <http://www.cdc.gov/mmwr>. Use the search function to find specific articles.

Use of trade names and commercial sources is for identification only and does not imply endorsement by the U.S. Department of Health and Human Services.

References to non-CDC sites on the Internet are provided as a service to *MMWR* readers and do not constitute or imply endorsement of these organizations or their programs by CDC or the U.S. Department of Health and Human Services. CDC is not responsible for the content of these sites. URL addresses listed in *MMWR* were current as of the date of publication.