Evaluation Challenges for Syndromic Surveillance — Making Incremental Progress

Daniel M. Sosin, J. DeThomasis
Epidemiology Program Office, CDC

Corresponding author: Daniel M. Sosin, Associate Director for Science, Office of Terrorism Preparedness and Emergency Response, CDC, 1600 Clifton Road, Mailstop D-44, Atlanta, GA 30333. Telephone: 404-639-1528; Fax: 404-639-7977; E-mail: dms8@cdc.gov.

Abstract

Introduction: *The 2003 National Syndromic Surveillance Conference provided an opportunity to examine challenges and progress in evaluating syndromic surveillance systems.*

Objectives: *Using the conference abstracts as a focus, this paper describes the status of performance measurement of syndromic surveillance systems and ongoing challenges in system evaluation.*

Methods: *Ninety-nine original abstracts were reviewed and classified descriptively and according to their presentation of evaluation attributes.*

Results: *System evaluation was the primary focus of 35% of the abstracts submitted. Of those abstracts, 63% referenced prospective evaluation methods and 57% reported on outbreak detection. However, no data were provided in 34% of the evaluation abstracts, and only 37% referred to system signals, 20% to investigation of system signals, and 20% to timeliness.*

Conclusions: *Although this abstract review is not representative of all current syndromic surveillance efforts, it highlights recent attention to evaluation and the need for a basic set of system performance measures. It also proposes questions to be answered of all public health systems used for outbreak detection.*

Introduction

Interest in syndromic surveillance remains high in the United States, with approximately 100 state and local health jurisdictions conducting a form of syndromic surveillance in 2003 (*1*). However, skepticism about the efficacy of syndromic surveillance for early detection of terrorism-related illness has increased (*1–4*).

At the 2002 National Syndromic Surveillance Conference, an evaluation framework (*5*) was presented that closely followed CDC's Updated Guidelines for Evaluation of Public Health Surveillance Systems (*6*). That evaluation framework described the system attributes that should be measured but provided limited guidance on how to measure those attributes consistently.

In 2003, CDC convened a national working group on outbreak-detection surveillance.* The working group clarified terminology and revised earlier frameworks to emphasize early outbreak detection, putting syndromic surveillance into context as a specialized surveillance tool. The resulting Framework for Evaluating Public Health Surveillance Systems for Early Detection of Outbreaks (*7*) provides a structure for evaluating syndromic surveillance systems and reporting the results. The revised framework offers a task list for describing a surveillance system (Box 1) and provides visual aids to improve standard collection and reporting of evaluation information. The framework also provides a timeline with milestones in outbreak development and detection, from exposure to a pathogen to the initiation of a public health intervention. Although this timeline does not specify a single, reproducible measure to reflect the timeliness of detection, it does provide more consistent specification of intervals for comparing performance among different systems and different settings. The framework also describes two approaches, encompassing sensitivity, predictive value negative, and predictive value positive, to evaluate system validity for outbreak detection: 1) the systematic description and accumulation of experiences with outbreak detection, and 2) simulation-based methods.

The importance of evaluating syndromic surveillance systems is widely recognized (*1,3–5,8–11*), but a common set of measures have not yet been defined that will establish the added value of syndromic surveillance compared with current surveillance tools. Nonetheless, progress has been made toward uniform guidance on evaluating syndromic surveillance systems (*7*). This paper summarizes progress during 2003 and describes steps for the future.

* **Working group members:** Daniel M. Sosin, M.D., Claire Broome, M.D., Richard Hopkins, M.D., Henry Rolka, M.S., Van Tong, M.P.H., James W. Buehler, M.D., Louise Gresham, Ph.D., Ken Kleinman, Sc.D., Farzad Mostashari, M.D., J. Marc Overhage, M.D., Julie Pavlin, M.D., Robert Rolfs, M.D., David Siegrist, M.S.

BOX 1. Tasks for evaluating public health surveillance systems for early detection of outbreaks*

Task A. Describe the system

1. Purpose: What is the system designed to accomplish?
2. Stakeholders: Whom does the system serve?
3. Operation: How does the system work?
   a. Systemwide processes
   b. Data sources
   c. Data preprocessing
   d. Statistical analysis
   e. Epidemiologic analysis, interpretation, and investigation

Task B. Provide data demonstrating outbreak detection attributes

1. Timeliness: How early in the outbreak is the event detected?
2. Validity: How well does the system perform in distinguishing outbreak detection of public health significance from less important events or random variations in disease trends?
   a. Sensitivity and predictive value: What percentage of true outbreaks are detected by the system? What percentage of signals by the system are relevant (true positives)? What percentage of negative results are truly negative?
   b. Data quality: How does data quality affect validity of outbreak detection?
      i. Representativeness: How well does the system reflect the population of interest?
      ii. Completeness: What percentage of data are present for each record?

Task C. Describe the system experience

1. System usefulness: In what ways has the system demonstrated value relevant to public health?
2. Flexibility: How adaptable is the system to changing needs and risk thresholds?
3. System acceptability: Have stakeholders been willing to contribute to and use the system?
4. Portability: How readily can the system be duplicated at another location?
5. System stability: How consistent has the system been in providing access to reproducible results?
6. System costs: What are the resource requirements to deploy and maintain the system?

Task D. Summarize conclusions and make recommendations for use and improvement of systems for early outbreak detection

* **Source:** CDC. Framework for evaluating public health surveillance systems for early detection of outbreaks: recommendations from the CDC working group. MMWR 2004;53(No. RR-5).

Methods

The authors reviewed the original 99 abstracts submitted to the 2003 National Syndromic Surveillance Conference and divided them into two categories: 1) surveillance systems and 2) analytic methods. Abstracts about surveillance systems were subcategorized into 1) system descriptions, 2) implementations, and 3) evaluations. Analytic methods abstracts included those addressing detection algorithms, data modeling, and case definitions. For each abstract, the reviewers identified the geographic location of the surveillance system or primary author and the responsible entity for the system or study being described (e.g., local health department or university). Information was also gathered about the data-collection method used, the purpose of the system, and the type of data used. An abstract was classified as pertaining to system evaluation if the author indicated intent to present a system evaluation or if the abstract provided results of the system's experience in detecting outbreaks. Evaluation variables abstracted were frequency of system signals, investigations, outbreaks detected and missed, estimation of timeliness, and the effect of early detection.

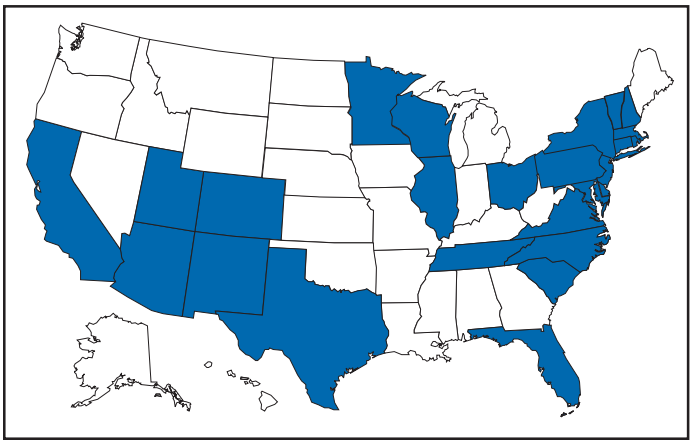
Each abstract was reviewed by both authors of this paper and results were reconciled in a meeting. Abstract forms were entered into Epi Info 2002 (http://www.cdc.gov/epiinfo/) for analysis.

Results

The 99 abstracts were submitted by authors from 23 states, the District of Columbia, and seven countries outside the United States (Figure). The bulk of the syndromic surveillance work, as reflected in these abstracts, is occurring in state and local health departments and within U.S. academic institutions. Abstract authors were based in state and local health departments (40%), universities (32%), federal government agencies (13%), health-care organizations (11%), and businesses (4%). Abstracts focused on system evaluation (35%), description of systems or their implementation (26%), data management, modeling, and detection algorithms (28%), and case definition (11%).

Of the 60 abstracts that described a full syndromic surveillance system, 30% indicated use of manual data collection

FIGURE. Location of U.S.-based syndromic surveillance systems described in 99 abstracts submitted for the 2003 National Syndromic Surveillance Conference



outside the typical workflow of the data provider. Ninety-five percent described systems designed to detect outbreak patterns in the data, with only 5% using syndromic surveillance for individual case detection (e.g., severe acute respiratory syndrome or West Nile encephalitis). Of the 35 abstracts that described system evaluation, 34% provided no data in the abstract, only describing the intent to present evaluation data. Nonetheless, 63% addressed the outbreak-detection experience in a prospective direction. Of the 35 abstracts that described system evaluation, 37% reported on the signaling of a system; 20% referred to one or more investigations; 57% addressed one or more outbreaks detected or missed; and 20% addressed timeliness in any fashion. None of the abstracts estimated the public health effect of early detection.

Discussion

The systems described in these conference abstracts are not a representative sample of jurisdictions conducting syndromic surveillance; rather, they are a synopsis from those jurisdictions willing to share their experiences at a national conference. Furthermore, certain presentations were invited talks for which abstracts were not submitted.

The diversity of data sources being used reflects the early stage of development of syndromic surveillance and the exploration of novel data sources (Table). The predominant focus, consistent with recommendations from the 2002 National Syndromic Surveillance Conference (*9*), is on data from emergency departments and other clinical sources. A substantial number of systems (30%) continue to rely on manual data collection at the data source. The sustainability of such a system has been questioned (*3,8–10,12*). Whether for routine data collection or for innovative surveillance

TABLE. Data sources for 60 abstracts on syndromic surveillance submitted to the 2003 National Syndromic Surveillance Conference

Data source	No. of abstracts*	%*
Emergency departments	29	48
Office or clinic visits	13	22
Hospital diagnoses	7	12
School absences	7	12
911 calls/EMS runs	6	10
Over-the-counter drug purchases	5	8
Poison control centers	5	8
Nurse advice lines	4	7
Veterinary clinics	3	5
Medical examiners	2	3
Pharmacy prescriptions	2	3
Laboratory results	2	3
Laboratory orders	1	2
Medical center parking-lot volume	1	2
Online obituaries	1	2
Subway-worker absences	1	2
School perception of an outbreak	1	2

* Certain systems used multiple data sources.

systems, automated data captured during the usual course of care (or business) is preferred to manual data collection when continuous, complete reporting is the goal. Manual data collection will continue to play a role in actual or threatened outbreak settings that have special data needs that cannot be filled by using existing electronic data (*3,7,9,10,12*).

A substantial number of abstracts (35%) focused on the evaluation of a system, although the rigor and methods of evaluation varied considerably. One third of abstracts that stated intent to present a system evaluation provided no data at all in the abstract regarding how effectively the system was working. However, approximately two thirds of the evaluation abstracts referred to tracking performance prospectively rather than simply analyzing historical data to identify known events. Not only is prospective identification of an outbreak a more substantial indicator of success, but it also offers benefits beyond identifying specific events (e.g., stronger relationships between clinicians and public health practitioners and higher quality surveillance data) (*4,13–15*).

To better understand the performance of outbreak-detection systems, basic measures of performance need to be counted. How often a system signals (i.e., how often it indicates that something worthy of further investigation is occurring) also needs to be reported. This applies to all the ways that health departments detect outbreaks (e.g., phone calls from the public), not just to syndromic surveillance. Every surveillance system should be able to report how many times in a given period (e.g., 1 month) it has triggered a follow-up investigation, yet only 37% of the evaluation abstracts gave any indication of system signals, much less a rate of signaling.

More information is needed about different responses to signals and the results of those responses. When a system signals, multiple responses can be made, from deciding not to act on the signal to launching a full investigation with staff participation and new data collection. Intermediate steps might include reviewing the data for errors, reviewing records manually within syndrome categories to search for patterns, conducting manual epidemiologic analysis for subgroup associations with the signal, examining data from other sources, and ensuring early submission of the next cycle of reports from affected locations. Although certain systems are potentially not signaling and therefore not instigating investigations, that only 20% of the systems presented in the evaluation abstracts have initiated investigations seems unlikely. Routine reporting of how often signals elicit a response and what those responses entail is essential.

Jurisdictions should report routinely both on outbreaks detected through syndromic surveillance and outbreaks missed. Practitioners should also report outbreaks detected through other methods to understand the relative value of syndromic surveillance. Of the 2003 evaluation abstracts, >50% addressed the detection or nondetection of outbreaks, but room for improvement remains.

Lastly, early detection is essential in syndromic surveillance, yet only 20% of the evaluation abstracts addressed timeliness. Measuring timeliness should be a routine part of reporting. The evaluation timeline in the Framework for Evaluating Public Health Surveillance Systems for Early Detection of Outbreaks (*7*) provides milestones that should aid in the reporting of timeliness.

Conclusion and Next Steps

Evaluation requirements should be simplified and standardized to allow comparisons across systems and across outbreak-detection approaches. Simulations offer promise for testing and improving systems designed to detect rare events. The abstracts submitted to the 2003 conference reflect initial efforts to evaluate analytic methods in isolation with simulation exercises. Testing intact systems is needed to verify how well they might perform in practice at providing early warning of public health emergencies. Additional research is needed to validate the assumptions necessary for modeling disease outbreaks (e.g., the spread of disease in various scenarios, or the individual and community behavior patterns after onset of illness that might serve as early outbreak indicators).

Although detailed descriptions of systems would be a helpful step forward, the reporting burden could be heavy and additional experience is needed to determine the required system attributes and to standardize the descriptions. An interim approach might be to prioritize a limited number of measures of likely value now until experience is gained with other measures. A simplified version of the Framework for Evaluating Public Health Surveillance Systems for Early Detection of Outbreaks (*7*) might focus on questions regarding timeliness, validity, and usefulness of an outbreak-detection system (Box 2). Such a framework could help standardize reporting of the different methods used by public health departments to detect outbreaks. Ultimately, the goal is to measure the effect of detection methods — how public health is improved by detection, and at what cost. The proposed framework could move the field forward incrementally by using readily available information and measures until additional information on metrics for outcomes and costs becomes available.

BOX 2. Priority evaluation questions for early outbreak-detection systems

1. How often does the system signal an event for further epidemiologic attention?
   a. What was the time period (e.g., 1 month)?
   b. What was the statistical threshold (e.g., p-value)?
   c. If the threshold has changed, explain why.
2. How were signals responded to?
   a. What percentage of signals were investigated through new data collection?
   b. What percentage caused increased reporting frequency from affected sites?
   c. What percentage conducted detailed manual analysis of any data available to the jurisdiction?
   d. What percentage conducted manual analysis of data from the system?
   e. What percentage were reviewed for data errors?
   f. What percentage of signals were ignored?
   g. What resources were directed to follow-up?
3. How many outbreaks were detected through the system?
   a. How timely was detection relative to other systems?
   b. How timely was detection relative to the stage of the outbreak?
   c. What were the agent, host population, and environmental conditions of the outbreak?
4. How many outbreaks were missed by the system?
   a. What were the agent, host, and environmental conditions?
   b. How was the outbreak detected?
5. What was the public health response to detection (e.g., no response, urgent communication to clinicians, or vaccination campaign)?

References

1. Buehler JW, Berkelman RL, Hartley DM, Peters CJ. Syndromic surveillance and bioterrorism-related epidemics. Emerg Infect Dis 2003. Available at http://www.cdc.gov/ncidod/EID/vol9no10/03-0231.htm.
2. Becker J. High-tech health watch draws cash, questions. Washington Post, November 23, 2003:A17.
3. Reingold A. If syndromic surveillance is the answer, what is the question? Biosecur Bioterr 2003;1:1–5.
4. Sosin DM. Syndromic surveillance: the case for skillful investment. Biosecur Bioterr 2003;1:1–7.
5. Sosin DM. Draft framework for evaluating syndromic surveillance systems. J Urban Health 2003;80(2 Suppl 1):i8–13.
6. CDC. Updated guidelines for evaluating public health surveillance systems: recommendations from the guidelines working group. MMWR 2001;50(No. RR-13).
7. CDC. Framework for evaluating public health surveillance systems for early detection of outbreaks: recommendations from the CDC working group. MMWR 2004;53(No. RR-5).
8. Henning KJ. Appendix B: syndromic surveillance. In: Smolinski MS, Hamburg MA, Lederberg J, eds. Microbial threats to health: emergence, detection, and response. Washington, DC: National Academies Press, 2003:281–312. Available at http://books.nap.edu/books/030908864X/html/281.html#pagetop.
9. Mostashari F, Hartman J. Syndromic surveillance: a local perspective. J Urban Health 2003;80(2 Suppl 1):i1–7.
10. Pavlin JA, Mostashari F, Kortepeter MG, et al. Innovative surveillance methods for rapid detection of disease outbreaks and bioterrorism: results of an interagency workshop on health indicator surveillance. Am J Public Health 2003;93:1230–5.
11. Smolinski MS, Hamburg MA, Lederberg J, eds. Microbial threats to health: emergence, detection, and response. Washington, DC: National Academies Press, 2003. Available at http://books.nap.edu/books/030908864X/html/index.html.
12. Das D, Weiss D, Mostashari F, et al. Enhanced drop-in syndromic surveillance in New York City following September 11, 2001. J Urban Health 2003;80(2 Suppl 1):i76–88.
13. Brown K, Pavlin J, Mansfield J, Elbert E, Foster V, Kelley P. Identification and investigation of disease outbreaks by ESSENCE [Abstract]. J Urban Health 2003;80(2 Suppl 1):i119.
14. Johnson J, McClean C, Poggemeyer K, Ginsberg M. Application of bioterrorism surveillance methods in San Diego County [Abstract]. J Urban Health 2003;80(2 Suppl 1):i137.
15. Schumacher M, Nohre L, Santana S. Partial evaluation of a drop-in bioterrorism surveillance system in Phoenix, Arizona [Abstract]. J Urban Health 2003;80(2 Suppl 1):i118.