Monitoring Infectious Diseases with Automated Systems

U.S. experts debate the merits of algorithm-based systems for monitoring outbreaks, including bioterrorist attacks

William Check

Detecting a cholera epidemic in London during the mid-19th century consisted rather simply of counting bodies. Today, public health investigators work at a much finer level, watching at the early stage for outbreaks, with the intent of quickly stopping their spread with treatments, vaccinations, or other measures.

More ambitious goals call for applying evermore-sophisticated tools to detect outbreaks, and public health and infection control experts are increasingly resorting to computer software to enhance these efforts. Some observers believe that automated surveillance systems can detect not only natural outbreaks but also agents of bioterrorism, although testing this possibility remains problematic.

Automated Surveillance Is Coming into Wider Use

Nonetheless, two automated syndromic surveillance algorithms proved accurate enough in a competition for officials at the U.S. Department of Defense to deploy one of them worldwide to protect military personnel. Even so, there is continuing disagreement about what will improve public health more—installing automated surveillance systems or, instead, hiring more people to cope with outbreaks once they are detected.

“We’re at a point where we have proven the concept [of automated surveillance],” says Nebraska State Epidemiologist Tom Safranek. “We have the highways built, now we need to have more cars on these highways.” When computer algorithms are more widely implemented, he says, “We will have a much more accurate picture of public health than we do now.”

Automated surveillance systems are costly, but advocates argue that they pay for themselves. “I wouldn’t classify any of these products as cheap,” says Michael Edmond, Associate Chair for Education in the Department of Internal Medicine and Hospital Epidemiologist and Medical Director of Performance Improvement at Virginia Commonwealth University (VCU) Medical Center in Richmond. “But you don’t have to prevent a lot of [hospital-acquired] infections to save enough money to pay for them.” He expected administrators at his hospital to oppose automated surveillance systems, but they asked him to look into them, citing potential financial savings, particularly at the hospital pharmacy.

“My impression is that these systems will be necessary adjuncts to our ability to recognize nascent public health events because there is so much information of potential value, only a small amount of which can be processed by individuals,” says David Relman of Stanford University and the Veterans Administration Palo Alto Health Care

- Automated surveillance systems can help to speed efforts to identify infectious disease outbreaks, both natural and bioterrorism-related.
- Automated case reporting could improve both the timeliness and completeness of identifying outbreaks of public health concern such as tuberculosis.
- Data-mining algorithms set to detect clinically interesting and statistically significant changes over set periods may help to uncover disease outbreaks and also to identify their causes.

William Check is a freelance writer in Wilmette, Ill.
System. “Automated methods that make use of all information and look for patterns that would not otherwise be apparent are the way of the future.

“In my mind,” Relman continues, “one of the big remaining deficits is the lack of validation.” He notes that it is difficult to test proposed patterns or predictors because of insufficient numbers of community public health events similar to those that epidemiologists are interested in detecting. While Relman sees computerized algorithms as being valuable to track community outbreaks, he calls the need for them within hospitals “moderate” compared to the need for better resources to respond to identified carriers of infectious disease agents. “Sometimes there are not sufficient numbers of trained infection control personnel to maintain the recommended ratio of providers to patients,” he says.

But Lance Peterson, epidemiologist and professor of pathology and medicine at Evanston (Ill.) Northwestern Healthcare (ENH), which has been an early adopter of automated surveillance systems, believes that they have value for individual institutions. “Most hospitals don’t have a clue as to how much money they lose to hospital-acquired infections [HAIs],” Peterson says. At ENH with 40,000 admissions per year, Peterson estimates that annual losses to HAIs are $10–15 million, which he considers lower than in many hospitals of comparable size. “Providers won’t pay for care related to proven nosocomial infections,” he says. And, as more states mandate reporting of HAI rates, large hospitals may see automatic surveillance systems as necessary, if expensive, support technology.

Some Skeptics Continue To Question Benefits of Automated Surveillance

Some skeptics question the value of automated surveillance when dealing with community outbreaks, offering only tepid acceptance of its worth. “Computer monitoring has a lot to be said for it” when tracking infections in hospitals, says Arthur Reingold, an epidemiologist in the School of Public Health at the University of California, Berkeley. But he questions whether complex algorithms will make a difference when dealing with community outbreaks. “My issue is whether you accomplish anything by doing that, and whether we have adequate human resources to follow up,” he says. “I’m not convinced that we will have an impact on morbidity and mortality by syndromic surveillance.”

A typical evaluation of automated surveillance for HAIs was reported in 2004 by William Trick at Stroger Hospital of Cook County, Ill., and collaborators. For the computer algorithm, which was written in house, they pooled data from various departments, including laboratory and pharmacy databases. The top standard for assigning outbreaks of bloodstream infections was retrospective review by investigators. Four other methods were compared: prospective review by professionals; a positive blood culture plus manual determination of a central venous catheter (CVC); computer algorithm; and computer algorithm plus manual CVC determination, which correlated best with their top standard. “Humans reviewing the same potential bloodstream infections came to different conclusions more often than the computer,” Trick says.

In an evaluation done at ENH, medical record review (MRR) by infection control professionals was compared to results from a commercial algorithm from MedMined, Inc. (MM) that monitors microbiology laboratory data and presents daily alerts. Over a four-day period in April 2004, the sensitivity and specificity of MRR (92% and 100%) were slightly higher than those of the MM program (86% and 98.4%). However, MRR took 17 minutes per admission, while MM took 10 minutes of personnel time per week.

A commercial automated surveillance system from Cereplex, Inc., was recently installed at VCU. “To do efficient epidemiology,” Edmond says, “you need to be able to do real-time queries of your microbiology database. Many hospital and laboratory information systems don’t have that capability or are relatively primitive.” He believes that automatic notification and real-time monitoring of microbiology data will make it possible to catch outbreaks more quickly. In addition, they could allow infection control personnel to do more targeted surveillance. “We realized how much time our infection control nurses are using to do manual tasks,” Edmond says. “If we could free up their time, they could survey more surgery and more units in the hospital.” He sees automated surveillance being most cost-effective in larger hospitals, especially because these hospitals see more complicated
Using Automated Surveillance on a Broader Basis

On a larger playing field, Nebraska State Epidemiologist Safranek is directing a pilot project that could lead to automated, standardized reporting by laboratories of communicable diseases, such as tuberculosis or anthrax and salmonella infections, with public health implications. Traditionally, the reporting of such cases to state public health departments has been manual and labor intensive, requiring someone at each lab to recognize, record, and later transmit positive test results, typically on a weekly basis except in the case of an emergency.

In a joint project with Cereplex, labs participating in this Nebraska pilot project will automate mandatory reporting by inserting filters into information systems that flag positive tests for reportable diseases. For instance, this filtered data will be sent nightly from Methodist Hospital in Omaha to Cereplex’s main computers in Maryland for standardized formatting, which is essential when handling data from different hospitals. Once analyzed, reports will be forwarded to CDC.

Safranek calls this formatting and consolidation of data “a major challenge.” For instance, many different tests are being used to identify hepatitis infections, and they often are identified by different names. Yet, the software has to recognize all those possibilities and then group them within a single rubric. “Although this sounds simple,” he says, “It is not. We are dealing with legacy systems and trying to force them to accomplish something they were not designed to do.”

However, if handled successfully, such automated reports could improve both the timeliness and completeness of reporting, which can be as low as one-half to two-thirds when labs rely on traditional manual systems. “We have been setting up for about a year,” Safranek says. “We are on the verge of going live.”

“Laboratory-based reporting of infectious diseases is good public health practice,” says infectious disease physician Dan Peterson, who is now CEO of Cereplex. “Automated reporting of reportable diseases from the laboratory to the public health department would be the single best way to detect bioterrorism events as well as clusters other than bioterrorism.” The reporting of salmonella infections, for instance, is faster from a hospital microbiology laboratory than from physicians’ offices. But no one yet knows whether automated reporting from a microbiology laboratory is faster than manual reporting.

Data-Mining Programs Probe Surveillance Data for Patterns

Data-mining algorithms offer a more sophisticated approach for detecting outbreaks—going beyond merely speedier reporting of cases by using artificial intelligence rules to look for unusual patterns that could signify outbreaks. Data mining was first applied to marketing, such as telling appropriate retail outlets that dolls and chocolate candy are often purchased in the same market basket.

Such “market-basket” problems are a subset of the “frequent-set” problem, a staple of data mining, says former infectious diseases physician and computer scientist Steven Brossette, president and founder of MedMined. “From frequency sets you can derive association rules, which are probability statements about the likelihood of finding one item with another two things in same market basket,” Brossette says.

Frequent-set analysis can be applied to community epidemics or outbreaks of HAIs by replacing items in the typical market-basket algorithm with patient demographics or microbiology laboratory data. Those algorithms can be set to detect clinically interesting and statistically significant changes over set time periods, thus uncovering outbreaks and also helping to identify their causes. In general, however, Brossette says, solving such problems proves very challenging. “Most data-mining projects fail because they start with a pile of data, which people put into canned algorithms,” he explains. “That generates a pile of patterns, but many are uninteresting, leaving the meaningful ones buried in a pile of dreck.” At ENH, Lance Peterson and colleagues are validating the MedMined data-mining system.

An earlier, large-scale evaluation of a syndromic surveillance system took place in 2003 as part of a project sponsored by the Defense Advanced Research Projects Agency (DARPA) in its Bio-ALIRT (Bio-Event Advanced Leading Indicator Recognition Technology) surveillance...
program that is intended to detect terrorist releases of pathogens or toxins. A military syndromic surveillance system, ESSENCE (Electronic Surveillance System for the Early Notification of Community-Based Epidemics), which had been developed jointly by the Department of Defense (DoD), Walter Reed Army Institute of Research, and Johns Hopkins University Applied Physics Laboratory, was instituted, but on a limited basis, even earlier. However, the events of September 11, 2001, and subsequent incidents involving anthrax were “a clear trigger to expand” the scope of ESSENCE, according to Colonel Kenneth L. Cox of the Force Health Readiness program of the U.S. Air Force in Falls Church, Va. Since November 2001, ESSENCE has proved able to capture a number of disease outbreaks “of mundane, routine character,” such as influenza and diarrhea from Norwalk virus, he says. It was operating during the initial stages of the SARS outbreak in 2003 and was soon used to verify the absence of that virus among military personnel in Asia.

In 2003, the performances of ESSENCE and five other syndromic surveillance systems were compared to one another in a competitive fashion, using military and civilian data for respiratory and gastrointestinal (GI) illnesses from five metropolitan areas. All six teams were provided with illness data covering 18 months with which to “train” their algorithms, says David Siegrist, senior research fellow at the Potomac Institute in Arlington, Va., and principal investigator for the DARPA Bio-ALIRT evaluation.

Later, during the competitive phase, teams were given 6 months of data to determine whether their “trained” algorithms could detect outbreaks accurately. Meanwhile, public health professionals were organized into a separate outbreak detection group, whose task was to identify every outbreak within the 6-month dataset and to profile it—when it started and peaked, when they estimated that public health professionals would make the call. “This made me believe that these systems would be very good to deploy for public health purposes,” he says.

ESSENCE IVB is now in use by civilian public health departments in Washington, D.C., Virginia, and San Diego. RODS is being used in Pennsylvania, Ohio, Utah, and Michigan. ESSENCE will begin running on DoD computers next year to monitor military personnel worldwide.

Data Mining versus Shoes on the Ground, Eyes on the Clinic

Whether automated surveillance systems will routinely detect community or regional outbreaks before public health professionals can do so remains to be determined. “What’s most important is shoes on the ground and eyes on the clinic,” Cox says. He notes that when power went out in New York City, experts knew that the ensuing lack of refrigeration would likely soon become a public health issue. As expected, there was a major increase in the number of people coming in with GI-related complaints.

“Public health departments identify foodborne outbreaks, which are usually smaller and community-focused, that don’t show up in the syndromic surveillance system,” Cox says. “Preventive medicine staff have always known about the outbreaks we find in ESSENCE before we call.” What the automated system does is to allow officials to monitor such events more systematically. “As we add more data, we hope that we can predict things,” he says. “The more [that automated systems] involve local historical data rather than national averages, the more likely the system reflects reality.”

Reingold questions the ability of syndromic surveillance systems to detect either natural or deliberate community outbreaks. He notes that politicians have become interested in these algorithms because of concerns generated by the deliberately inflicted cases of anthrax in 2001 and of bioterrorism in general. “Most people
who know anything about bioterrorism share a deep skepticism about whether syndromic surveillance as it is being done will be useful for detecting bioterrorist events,” he asserts. “It is highly unlikely ever to be useful early in the course of a bioterrorist attack, and inconceivable that anyone will ever be able to prove it has been effective in that setting.” Moreover, he notes that, of the three documented events in which biological agents were used for criminal purposes in the U.S. over the last 20 years, “All have been domestic and, except for anthrax, there have been no fatalities.”

As for detecting natural community outbreaks through syndromic surveillance, Reingold says, “I suggest that we already learn about most important outbreaks. You could say we will detect them sooner, but I would argue that most public health departments already detect more outbreaks than they can do anything about and end up triaging many that are not so important to public health.” In such circumstances, augmented resources could serve public health needs better than would automated systems whose uses are limited to quickly detecting those outbreaks.

SUGGESTED READING