EVIDENCE BASED POLICY AND PRACTICE

Review of syndromic surveillance: implications for waterborne disease detection

Magdalena Berger, Rita Shiau, June M Weintraub

J Epidemiol Community Health 2006;60:543–550. doi: 10.1136/jech.2005.038539

Syndromic surveillance is the gathering of data for public health purposes before laboratory or clinically confirmed information is available. Interest in syndromic surveillance has increased because of concerns about bioterrorism. In addition to bioterrorism detection, syndromic surveillance may be suited to detecting waterborne disease outbreaks. Theoretical benefits of syndromic surveillance include potential timeliness, increased response capacity, ability to establish baseline disease burdens, and ability to delineate the geographical reach of an outbreak. This review summarises the evidence gathered from retrospective, prospective, and simulation studies to assess the efficacy of syndromic surveillance for waterborne disease detection. There is little evidence that syndromic surveillance mitigates the effects of disease outbreaks through earlier detection and response. Syndromic surveillance should not be implemented at the expense of traditional disease surveillance, and should not be relied upon as a principal outbreak detection tool. The utility of syndromic surveillance is dependent on alarm thresholds that can be evaluated in practice. Syndromic data sources such as over the counter drug sales for detection of waterborne outbreaks should be further evaluated.

surveillance has the potential to effectively mitigate the extent of morbidity, mortality, and social and financial unrest resulting from natural or manmade outbreaks.

The effectiveness of syndromic surveillance in terms of informing timely and successful public health interventions has not been demonstrated in public health practice. Although the impetus for its development has been its potential use as a bioterrorism preparedness tool, it may be well suited to monitoring naturally occurring infectious and chronic disease. The utility of syndromic surveillance for monitoring waterborne or environmentally mediated disease incidence has not been assessed to date. However, any decision about the implementation of a syndromic surveillance system cannot focus on only one aspect of disease surveillance and response. This decision has to be made within the context of the existing public health system and its ability to respond to public health threats of all types. Additionally the design of an effective syndromic surveillance system must take into account the data needs and response capabilities of many different public health responders.

This review summarises the evidence gathered from retrospective, prospective, and simulation studies to assess the efficacy of syndromic surveillance for waterborne diseases detection. The original aim of this review was to determine the utility of syndromic surveillance systems for waterborne disease detection in the San Francisco Bay Area. All articles were selected from peer reviewed articles found on Medline under the search term “syndromic surveillance”.

DATA SOURCES

Three levels of data may be used for disease surveillance: pre-clinical data, clinical pre-diagnostic data, and diagnostic data.³ Syndromic surveillance usually uses two types of data sources: pre-clinical and clinical pre-diagnostic data. Traditional surveillance generally focuses on diagnostic data.

Pre-clinical data sources

Public behaviour may be an early indicator of an increase in disease incidence in the population. Information about school and work absenteeism,^{1 4} calls to nurse help lines,⁵ poison control centres,⁶ water utility complaint lines,⁷ and sales information of over the counter and prescription drugs^{1 4 7–14} may give information about the level of disease in the population. Other environmental data, such as drinking water turbidity levels, may also be useful indicators of disease incidence. While pre-clinical data sources tend to be

Syndromic surveillance is a tool for outbreak detection that has been used by public health departments since the mid-1990s.¹ The CDC has defined syndromic surveillance as “...surveillance using health-related data that precede diagnosis and signal a sufficient probability of a case or an outbreak to warrant further public health response.”² In the USA, interest in syndromic surveillance increased after September 11, 2001 because of concern about the possibility of bioterrorist attacks.

Theoretically, syndromic surveillance systems have the potential to supplement traditional infectious disease surveillance systems by providing information about the extent of an outbreak or seasonal increases in disease incidence. They may also provide reassurance that an outbreak is not happening. As syndromic data are gathered before diagnostic or laboratory information is available, health departments may be able to recognise and respond to increases in disease incidence before formal diagnoses are made, and to respond to outbreaks early in their course. In this way, syndromic

See end of article for authors' affiliations

Correspondence to: Ms M Berger, San Francisco Department of Public Health, Environmental Health Section, 1390 Market Street, Suite 910, San Francisco, CA 94102, USA; mberger@health.nyc.gov

Accepted for publication 19 December 2005

more timely than clinical data sources, they are much less specific and therefore provide a less solid basis for public health decision making.

Clinical pre-diagnostic data sources

Most clinical data sources currently used by syndromic surveillance systems consist of electronic data gathered for an independent purpose and adapted to the needs of the specific surveillance system. Sources of clinical data used for syndromic surveillance include health care utilisation records such as registers of emergency department chief complaints,^{1 4 15–24} discharge data,²⁰ ambulance dispatch data,²⁵ and ICD-9 or other procedure codes used in inpatient and outpatient settings.^{4 20 26} Electronically captured orders for diagnostic tests may also be used as a data source.

Different groupings of symptoms, ICD-9 codes, or drugs, as well as the most effective free text search algorithms are being researched to determine how these data are best analysed to accurately reflect disease activity in the population being studied.^{17 21 27 28}

DATA TIMELINESS AND ACCURACY

Usually, data sources that are chronologically distant from diagnostic information are timely but not as specific as data sources that are chronologically closer to diagnosis.³ Potential syndromic surveillance data sources along with each source's capacity for being timely and specific are described in figure 1. Comparing two data sources from this figure illustrates the trade off between data timeliness (the amount of lag time from an event to data about the event being available for interpretation by a public health department) and specificity/accuracy (the degree to which information with relevant patient characteristics is available). Over the counter (OTC) drug data have the benefit of timeliness; for example, people who are beginning to have symptoms of an intestinal illness may purchase OTC drugs a day or two before presenting at the emergency department (ED) or other health care facility. However, OTC sales data are not very specific; while a rise in sales of OTC antidiarrhoeal remedies may be the result of an increase in incidence of diarrhoeal illness, it may be unrelated to disease incidence. An increase in OTC sales may be explained by store specials, hoarding behaviours, or duplicate data transfers. It would be difficult to make a public health decision based solely on increased OTC sales because it does not contain any patient specific data. Involving pharmacists or pharmacy managers in initial models and categorisation of OTC data may increase the correlation of OTC sales data and possible human disease. Timeliness of data acquisition is only valuable if the signal is accurate and specific enough to inform public health decision making.

The timeliness and specificity of OTC data may be compared with ED chief complaint data. An increase in chief complaints of diarrhoea is an unambiguous signal that the incidence of intestinal illness has increased. It is less timely than the OTC data but yields much more detailed information. None the less this information does not distinguish between a natural and expected increase, a natural outbreak, a result of a bioterrorism event, or a coincidental, simultaneous increase in intestinal illnesses of varying aetiologies. It would still be difficult to make a public health decision based on these data alone without further investigation.

An increase in sales of OTC antidiarrhoeal drugs occurring concurrently with spikes in other data such as increases in worker absenteeism, and increases in ED chief complaints of diarrhoea would be a more reliable indicator of a true increase in disease incidence than any of these signals on their own. Developing a mechanism for evaluating data from disparate sources and of varying relevance to public health

response is one of the challenges in finding practical applications for syndromic surveillance information.³

The accuracy of aberration detection signals generated by a syndromic surveillance system is also greatly affected by the amount of baseline data that are available. Studies reporting the use of short term, or "drop-in" syndromic surveillance systems set up specifically for outbreak detection during a high profile event consistently cite the lack of baseline data as a factor that hampered the determination of an appropriate signal threshold for these systems, often leading to systems that were overly sensitive and did not account for seasonal or other temporal factors.^{15 29 30} Before a syndromic surveillance system is instituted for use as a regular surveillance tool, individual jurisdictions should ensure that at least one year of historical data are available for use as a reference for signal threshold determination and signal investigation.

SYNDROMIC SURVEILLANCE EVALUATIONS

Retrospective evaluations

Retrospective analyses show that syndromic surveillance systems may provide information that would allow public health departments to predict outbreaks earlier than by using traditional surveillance. For example, in the Milwaukee, Wisconsin cryptosporidiosis outbreak of 1993,³¹ newspaper reports of over the counter antidiarrhoeal drugs being sold out at pharmacies were one of the first clues that an unusual event had occurred.⁷ Calls to nurse hotlines showed a fourfold increase in the standard deviation of diarrhoea rates one day before observation of unusual activity by local pharmacists and five days before the local public health department was notified of the possible outbreak.⁵ Similarly, a retrospective review of pharmacy records showed that sales of over the counter drugs increased fivefold during an outbreak of cryptosporidiosis in North Battleford, Saskatchewan.^{7 32} During a 1983 outbreak of waterborne campylobacteriosis in Florida, pharmacy sales of antidiarrhoeals increased eightfold when compared with the same period in the previous year.¹⁶ Retrospective analysis of Canadian outbreaks of *Cryptosporidium*, *E coli* O157:H7, and *Campylobacter* also all showed increases in OTC sales that corresponded to increases in disease incidence.³³ In New York City, retrospective analyses of ambulance dispatch data and ED data showed that they were effective in predicting city wide respiratory, gastrointestinal, and influenza outbreaks.^{25 34}

Simulations

Simulated outbreaks can examine the ability of syndromic surveillance systems to accurately predict outbreaks while minimising false alarms.^{34 35} According to most simulations, detection algorithms used in syndromic surveillance can best identify large, geographically widespread increases in disease. Syndromic surveillance detection algorithms are less successful at identifying small counts or small increases in disease, showing that syndromic surveillance may augment traditional disease surveillance but may not improve timeliness or sensitivity. Simulations of syndromic surveillance suggest that it is best suited for detecting diseases that have a narrow incubation period distribution, a steep epidemic curve, a long prodromal phase, are not included on routine diagnostic tests, and do not have a specific disease identifying clinical feature.³⁶

To date no simulations of waterborne disease outbreaks have been conducted. Stoto *et al* tested a hypothetical syndromic surveillance system by simulating a "fast" outbreak (a disease that would have a steep epidemic curve) and a "slow" outbreak (a disease with a more gradual epidemic curve) of influenza-like illness (ILI) using three years of data from an ED.³⁷ Results of this exercise show that an episode of

Figure 1 Surveillance data sources and health seeking behaviour.

either “fast” or “slow” ILI during the annual influenza season was undetectable by syndromic surveillance when set to a 1% false positive rate. During the non-flu (RSI) season syndromic surveillance algorithms were successful at detecting a “fast” outbreak by day three. However, the “slow” outbreaks were much harder to detect. Only one algorithm had a 50% probability of detecting this outbreak by day nine. By day nine, such outbreaks would usually be detected through traditional surveillance systems. As Stoto *et al* point out, “The sudden appearance of large number of ILI cases...five times the daily average...especially in the summer- is clearly exceptional and ER physicians do not require any sophisticated statistical algorithm to tell them so”.³⁷

Prospective reports

Several prospective evaluations of long term surveillance systems^{18 23 34 38–43} have been published. Syndromic surveillance has been in use in New York City since 1995 to detect outbreaks of waterborne diarrhoeal disease. The waterborne disease surveillance system uses data from OTC antidiarrhoeal sales, ED chief complaint logs, reports from sentinel nursing homes, and reports of positive stool samples from clinical laboratories. Prospective studies evaluating this system between 2001 and 2004^{34–40} reviewed citywide signals and smaller spatial clusters to find out if signals represented real outbreaks, and whether outbreaks detected by traditional surveillance were also detected by the syndromic system. These studies found that, 75% of the citywide outbreak signals correlated with true citywide viral gastroenteritis outbreaks. However, there was no correlation between smaller spatial clusters detected by the syndromic surveillance system and gastrointestinal (GI) outbreaks detected by traditional surveillance. Additionally, there were 36 GI outbreaks detected by traditional surveillance, none of which were detected by the syndromic surveillance system. Despite the extensive use of staff resources, the availability of fully electronic and timely data, and the ability to quickly initiate responses to various alerts, the syndromic system did not appreciably improve timeliness or sensitivity to overall

disease surveillance.⁴⁰ A system in Connecticut that tracks hospital admissions shows similar discrepancies between GI outbreaks detected by traditional and syndromic surveillance.⁴² This system generated 35 GI illness alarms in the course of one year, only one of which turned out to be a true GI outbreak. Additionally, none of 15 GI illness clusters detected by laboratory surveillance were detected by the syndromic surveillance system.

The University of Maryland implemented a syndromic surveillance system in its university hospital and evaluated its system’s ability to detect actual clusters of patients with GI or respiratory symptoms presenting to the hospital from 2001 to 2002.³⁹ This system incorporated admission, discharge and transfer data as well as laboratory information system data. The evaluation showed that a peak in ED visits and an increase in stool test orders corresponded to a cluster of patients who later tested positive for infections with *Shigella spp.* The authors found the system to be timely and sensitive for their needs. A similar system is operated by the Department of Defense (DoD).²³ The DoD system detected three outbreaks of diarrhoeal diseases in three different locations, one of which was laboratory confirmed to be caused by rotavirus. A system based on ED chief complaint data administered by the Westchester County Department of Health was less specific. Over the course of 277 days, 59 signals were detected, none of which corresponded to an outbreak of a communicable disease.⁴³ A multi-jurisdictional system implemented for the 2002 Olympic games in Utah tracked encounters at urgent care centres and EDs using the real-time outbreak and disease surveillance system.¹⁸ During the two month Olympic event the system’s detection algorithms were exceeded twice; neither of these alarms corresponded to a real outbreak. The system was successfully maintained after the end of the Olympic event.

A study by Henry *et al* quantified the sensitivity, specificity, and positive predictive value (PPV) of the Kaiser Mid-Atlantic Region nurse hotline syndromic system by comparing the level of concordance between syndrome assignments based on the nurse hotline algorithm and the diagnosis made based on a subsequent Kaiser office visit. The authors found

Table 1 Characteristics of potentially waterborne disease agents

Agent	Incubation period ³⁰ mean (range)	Disease identifying feature ²⁰	Included on routine test?	Prodrome ²⁰
Parasitic				
<i>Cryptosporidium</i> *	7 days (1–12 days)	No symptoms often include watery diarrhoea and abdominal pain, and may include general malaise, fever anorexia and vomiting	No. Not included on routine O and P stool examination.	None in adults; children may experience a prodrome of vomiting and anorexia.
<i>Cyclospora</i> *	1 week	No. Symptoms include watery diarrhoea, and abdominal pain, nausea, anorexia, fatigue weight loss.	No. Not included on routine O and P stool examination	None.
<i>Giardia</i>	7–10 days (3–25 days)	No. Symptoms include watery diarrhoea, and abdominal pain, bloating, flatulence, fatigue weight loss, fatty bad smelling stool.	Yes. Included on standard O and P stool examination.	None.
<i>Entamoeba</i>	2–4 weeks (few days-years)	Sometimes. Symptoms range from mild diarrhoea to dysentery. Amebic granulomata, colitis, and liver lung, or brain abscess may occur.	Yes. Included on standard O and P stool examination.	None
Bacterial				
<i>Campylobacter</i>	2–5 days (1–10 days)	No. Symptoms include diarrhoea, abdominal pain, general malaise, fever, and vomiting.	Yes. Included on standard stool culture.	None
Enterohemorrhagic <i>E. coli</i> *	3–4 days (2–10 days)	Sometimes. Symptoms range from mild diarrhoea to dysentery. Fever is rare. Haemolytic uraemic syndrome and thrombotic thrombocytopenic purpura can develop.	No. Enterohemorrhagic strains require specific stool culture.	None
Enteropathogenic <i>E. coli</i> *	9–12 hours	Usually affects infants. Symptoms include watery diarrhoea and fever.	No. Requires specific stool culture.	None
<i>Legionella</i> ^{43, 52}	Legionnaires' disease: 5–6 days (2–10 days); Pontiac fever: 24–48 hours (5–66 hours)	No. Initial clinical presentation does not differ from other pneumonias. Symptoms include anorexia, malaise, myalgia, headache, fever, cough, abdominal pain, diarrhoea. In Legionnaires' disease these can proceed to respiratory failure and death.	No.	None.
<i>Shigella</i>	1–3 days (12 hours–1 week)	No. Symptoms depend on strain and range from mild to watery diarrhoea to dysentery and can include fever, nausea, and sometimes vomiting, and cramps.	Yes. Included on standard stool culture.	None.
<i>Vibrio cholerae</i>	2–3 days (Few hours–5 days)	Yes. Symptoms include sudden and profuse watery diarrhoea, nausea, vomiting, and rapid dehydration in some cases.	No. Special media required.	None.
<i>Vibrio parahaemolyticus</i> *	12–24 hours (4–30 hours)	No. Symptoms include watery diarrhoea, abdominal pain, nausea, fever, and vomiting. Dysentery sometimes develops.	No. Special media required.	None.
<i>Vibrio vulnificus</i> *	12–72 hours	Yes. Septicaemia, shock, hypotension, bullous skin lesions.	No. Special media required.	None.
Viral				
Hepatitis A*	28–30 days (15–50 days)	Yes. Symptoms include fever, anorexia, nausea, abdominal pain, jaundice.	Routine once hepatitis is suspected.	Yes. Prodrome before onset of jaundice usually lasts a few days.
Norovirus (Norwalk and Norwalk-like viruses)*	24–48 hours (10–50 hours)	No. Symptoms include nausea, vomiting, fever, diarrhoea, abdominal pain, fever.	No.	None.
Rotavirus*	24–72 hours	No. Symptoms include vomiting fever and watery diarrhoea, and most often affect children under 5 years old.	No.	None.

*May be good candidates for detection by syndromic surveillance. Criteria include: (1) narrow incubation period; (2) steep epidemic curve, long prodromal phase; (3) no clinical disease specific identifying feature; (4) disease not included in routine diagnostic tests.

Table 2 Potential syndromic surveillance data sources

	Lag time between exposure and information receipt	Data quality and sustainability			System sustainability		Other (pros and cons)
		Validity	Accuracy	Representativeness	Personnel	Equipment	
		Sensitivity					
		Specificity					
		PPV	Completeness				
Automated and electronically available	Very timely, data are available within ~24 hours of sale. Cases tend to self treat before visiting clinic.	No formal tests of validity have been done.	Data coverage in CA is 25%-33% of major pharmacies, this varies by county. Data source is not very detailed. Recent experience suggests that outbreaks picked up by traditional PH are missed (San Luis Obispo and Fresno counties).	May not be sustainable, data are gathered for an independent purpose, but the commitment of pharmacies may be variable, especially as time from last BT or other terrorist event increases.	Staff time used is minimal. Probably cannot be used for outbreak detection because the data are not detailed and personal identifiers are not available. Allows coverage of multi-jurisdictional area. Could potentially be used to measure extent of an outbreak once identified.		
Not automated and electronically available	This could be a timely data source.	Unknown	Most probably not very detailed. Could be sustainable, the information is gathered for an independent purpose. Data are often not available in electronic format. Personal identifiers/addresses are sometimes available.	Large amount of time would be needed to set up and maintain. Would place a burden on utilities many of which have limited staff.	This could strengthen the relationship between water utilities and public health departments may lead to cooperation on other issues.		
Sentinel nursing home surveillance	This could be a timely data source.	Unknown	Unclear if this population accurately represents surrounding population. Data based on clinical lab results would be accurate.	Difficult to implement across jurisdictions. Resource intensive in terms of DPH and facility staff time and stool testing materials, lab time.	Presents an added burden to nursing home residents.		
Nurse call lines	Data can be available within 24 hours for some systems. ⁴¹	According to Henry <i>et al</i> , Sensitivity = 72% Specificity = 96% PPV = 37% ³⁸	Data could potentially be linked to personal identifiers, and could be quite accurate. Data would probably be available in electronic format. Could be sustainable, the information is gathered for an independent purpose.	Large amount of time would be needed to set up and maintain.	This could strengthen the relationship between hospitals, insurance, and public health departments and may lead to cooperation on other issues.		
School/work absenteeism	This could be a timely data source.	Unknown	Most probably not very detailed, without names or reasons for absenteeism. If reasons are provided accuracy is very questionable.	Large amount of time would be needed to set up and maintain.	It is not clear which populations to target and how representative they would be. Many people commute to work or school.		

that the nurse hotline system achieved the highest sensitivity and PPV for respiratory and GI illness syndromes.³⁸

Staff and resources used

Performance evaluations often do not provide information about time and resources that are used in maintaining the system and investigating the alarms. In New York City, staff maintain the system, spending several hours a day, seven days per week to download and analyse data.¹ Investigation of false signals is a significant burden on staff resources.³⁷ An often mentioned solution is to investigate only those alarms that are maintained for two or more days and in two or more geographical areas. However, if the main goal of syndromic surveillance is to increase timeliness of outbreak detection and response then waiting two or more days to start an investigation would negate the timeliness benefit of the system.⁴⁴ Similarly, if, because of the low specificity of syndromic signals, local health departments require that syndromic signals are substantiated by specimen collection and laboratory confirmation before public health action may be taken, any potential advantages of timeliness are lost.⁴⁵ Decision makers must take into account the ability of the local public health system to respond to signals when considering the implementation of any syndromic surveillance system.

SYNDROMIC SURVEILLANCE USES FOR WATERBORNE DISEASE DETECTION

Criteria derived from simulations can be used to assess which waterborne diseases may be appropriate for identification by syndromic surveillance.³⁶

(1) The disease should have a narrow incubation period distribution.

Incubation periods of waterborne disease agents vary widely depending on dose, host susceptibility, and other factors.

(2) The disease should have a steep epidemic curve and a long prodromal phase.

Most waterborne diseases do not have a prodromal phase and an outbreak will not necessarily have a steep epidemic curve.

(3) The disease should not have a specific disease identifying clinical feature.

Initial symptoms of most potentially waterborne diseases are non-specific and generally include diarrhoea and other GI distress. Most of these illnesses do not have a disease identifying clinical or historical feature that allows clinicians to pinpoint the cause before performing laboratory tests.⁴⁶

(4) The disease should not be included in routine diagnostic tests.

Laboratory tests are not routinely ordered for many waterborne diseases,⁴⁶ and acute GI illnesses are generally under diagnosed and underreported.⁴⁷⁻⁴⁹

Table 1 lists characteristics of potentially waterborne pathogens. Because many waterborne diseases lack clinically identifying features and are not part of routine testing, they may be good candidates for detection by syndromic surveillance, based on the four criteria listed above. However, many

of the diseases that fit these criteria, while good candidates for syndromic surveillance, are often commonly occurring and nor necessarily of high public health importance.

Waterborne disease surveillance in the San Francisco Bay Area

In the San Francisco Bay Area four counties receive water from a common water utility. With the exception of a multi-county cryptosporidiosis surveillance project, surveillance for potentially waterborne disease in the San Francisco Bay Area is conducted by each county separately. There is no formal, timely coordination of waterborne disease surveillance across county lines. A system with multi-jurisdictional disease monitoring capabilities could provide public health benefit by permitting early detection of a multi-county waterborne outbreak. Surveillance data captured from multiple jurisdictions and interpreted centrally may lead to earlier outbreak detection than data gathered and interpreted by staff in separate counties who may not be aware of disease incidence in neighbouring jurisdictions.

Syndromic surveillance data sources may potentially provide cross-jurisdictional data, information about the geographical scope of an outbreak once one was identified, and additional reassurance that an outbreak was not occurring. However, the practical utility of syndromic data sources is uncertain; their outbreak detection benefits are currently theoretical and will remain so until accurate electronic capture of data and signal detection algorithms are refined or data sensitivity increases. The benefits of a regional waterborne disease surveillance system must be weighed against the resources required to set up such a system and the true risk of a waterborne disease outbreak in the San Francisco Bay Area. Based on the absence of any known waterborne outbreaks in the history of the water utility, extensive watershed protection measures, and a protected water source located in a national park it would seem that the risk of a waterborne disease outbreak occurring in the San Francisco Bay Area is quite small.

Potential waterborne disease syndromic surveillance data sources

Table 2 lists the potential waterborne disease syndromic surveillance data sources along with indicators of data quality and utility. Potential data sources in the table are divided into data that are currently available and accessible in electronic format in the San Francisco Bay Area and data that are not currently automated or electronically available but could be useful once they became automated and electronic.

Compared with other options, OTC surveillance for waterborne disease is currently the most feasible source of syndromic surveillance data available in the San Francisco metropolitan area because of the relative ease of implementing an existing, nationally funded system. Nursing home surveillance entails large inputs in terms of health department and on-site staff and fiscal resources for specimen testing, data monitoring, and signal investigation. Water utility complaint, nurse call line, and school and worker absenteeism logs are currently not compiled electronically in a central location; setting up a system to electronically capture these data would entail considerable commitments of

What this paper adds

To date, little attention has been paid to the possibility of using syndromic surveillance for monitoring waterborne disease incidence. This review of the benefits and shortcomings of syndromic surveillance may prove useful to public health practitioners and planners who are considering approaches to improve traditional surveillance systems.

Policy implications

This review will help policy makers weigh the costs and benefits of implementing a syndromic surveillance system and will clarify the drawbacks and advantages of potential data sources.

will and resources from the department of public health, other city agencies, and private and public partnerships with water utilities, insurance providers, hospitals and clinics, and large employers in the San Francisco Bay Area. Mechanisms for data storage, sharing, and retrieval would need to be established for each partnership. Finally and most importantly, dedicated public health staff who would compile, manage, and analyse syndromic data on a regular basis, and respond to syndromic surveillance signals would be needed. Signal verification and response activities may include: (1) determining data import and aberration detection algorithm problems that may lead to erroneous signals (for example, duplicate data, batch transfers from certain institutions, miscoding of information at the point of entry, text-string search algorithms that are too specific or not specific enough, etc); (2) verifying the validity of the signal by looking for the presence of signals in other data sources; (3) if the signal is deemed to be composed of possible true cases, hospital logs and charts may need to be manually reviewed by hospital or public health department staff and a line list compiled for clinical and/or laboratory based case verification; (4) traditional outbreak investigation activities and application of interventions. Timeliness provided by a syndromic surveillance system can only be useful if all of the above functions are supported by and integrated into the activities of the local health department on a sustained basis.

It is possible that outbreaks of cryptosporidiosis, cyclosporiasis, legionellosis, hepatitis A, and others may be detected through monitoring of OTC sales or that a suspected outbreak may be confirmed or better characterised through the use of these data. If OTC surveillance is to be used primarily as a back up data source to traditional surveillance, staff time for checking the web based interface would not exceed 15 minutes per day. Signal investigation would require additional resources. While the usefulness of OTC monitoring for waterborne diseases is only theoretical, given the availability of multi-county data and the low amount of staff time and effort needed to monitor the data, utilisation and prospective evaluation of these data may be recommended for two purposes: (1) reassurance of the absence of a waterborne disease outbreak and (2) establishing baseline familiarity that may prove helpful in the event of a waterborne disease outbreak. OTC data need to be correlated with known outbreaks in the geographical area where surveillance is occurring to clarify the validity and representativeness of these data before they can be used for prospective outbreak detection.

CONCLUSION

Because the efficacy of syndromic surveillance is not proved, it remains unclear whether it is worth the investment of personnel and financial resources to implement. In addition to the issues discussed in this review, it is challenging to develop sensible response protocols for syndromic surveillance systems because the likelihood of false alarms is so high, and because information is currently not specific enough to enable more timely outbreak detection or disease control activities. Although most syndromic surveillance systems do not use personally identifiable information, there are also considerations of personal privacy and public comfort with health information storage and analysis. The data needs of a syndromic surveillance system must be weighed against the needs and public demands for patient privacy.

There are many theoretical benefits of syndromic surveillance, including potential timeliness, increased response capacity, ability to establish baseline disease burdens, and ability to delineate the geographical reach of an outbreak. However, in the absence of empirical evidence of its efficacy to mitigate the effects of natural or intentional waterborne

disease outbreaks through earlier detection and response, syndromic surveillance should not be implemented at the expense of traditional public health activities, and it should not be relied upon as a principal waterborne disease outbreak detection tool at this time. For waterborne diseases, syndromic surveillance systems should continue to be assessed and implemented as a supplemental system to help develop better statistical methods and sensible alarm thresholds that can be applied and evaluated in practice.

Authors' affiliations

M Berger, J M Weintraub, San Francisco Department of Public Health, Environmental Health Section, USA

R Shiao, San Francisco Department of Public Health, Communicable Disease Control and Prevention Section

Funding: none.

Competing interests: none.

REFERENCES

1. **Heffernan R**, Mostashari F, Das D, *et al*. New York City syndromic surveillance system. *Syndromic surveillance: reports from a national conference*, 2003. *MMWR* 2004;**53**(suppl):25-7.
2. **Centers for Disease Control**. <http://www.cdc.gov/epo/dphsi/syndromic.htm> (accessed on 3 Oct 2005).
3. **Buckeridge D**, Graham J, O'Connor M, *et al*. Knowledge-based bioterrorism surveillance. *Proc AMIA Symp* 2002;76-80.
4. **Lombardo J**, Burkom H, Elbert E, *et al*. Systems overview of the electronic surveillance system for the early notification of community based epidemics (ESSENCE II). *J Urban Health* 2003;**80**(suppl):i32-41.
5. **Rodman J**, Frost F, Jakubowski W. Using nurse hotline calls for disease surveillance. *Emerg Infect Dis* 1998;**4**:329-32.
6. **Derby M**, McNally J, Ranger-Moore J, *et al*. A comparison of foodborne illness cases reported to a poison control center and public health department. *Syndromic surveillance: reports from a national conference*, 2004. *MMWR* 2005;**54**(suppl):35-40.
7. **Proctor ME**, Blair KA, Davis JP. Surveillance data for waterborne illness detection: an assessment following a massive waterborne outbreak of cryptosporidium infection. *Epidemiol Infect* 1998;**120**:43-54.
8. **Magruder S**, Lewis S, Najmi A, *et al*. Progress in understanding and using over-the-counter pharmaceuticals for syndromic surveillance. *Syndromic surveillance: reports from a national conference*, 2003. *MMWR* 2004;**53**(suppl):117-22.
9. **Foldy S**. Linking better surveillance to better outcomes. *Syndromic surveillance: reports from a national conference*, 2003. *MMWR* 2004;**53**(suppl):12-17.
10. **Beaudeau P**, Payment P, Bourderont D, *et al*. A time series study of anti-diarrheal drug sales and tap water quality. *Int J Environ Health Res* 1999;**9**:293-311.
11. **Hogan W**, Tsui F, Ivanov O, *et al*. Detection of pediatric respiratory and diarrheal outbreaks from sales of over-the-counter electrolyte products. *J Am Med Inform Assoc* 2003;**10**:555-62.
12. **Davies G**, Finch R. Sales of over-the-counter remedies as an early warning system for winter bed crises. *Clin Microbiol Infect* 2003;**9**:858-63.
13. **Sacks J**, Lieb S, Baldy L, *et al*. Epidemic campylobacteriosis associated with community water supply. *Am J Public Health* 1986;**76**:424-8.
14. **Wagner M**, Tsui F, Espino J, *et al*. National retail data monitor for public health surveillance. *Syndromic surveillance: reports from a national conference*, 2003. *MMWR* 2004;**53**(suppl):40-2.
15. **Foldy S**, Biedrzycki PA, Barthell EN, *et al*. Syndromic surveillance using regional emergency medicine internet. *Ann Emerg Med* 2004;**44**:242-6.
16. **Beitel AJ**, Olson KL, Reis BY, *et al*. Use of emergency department chief complaint and diagnostic codes for identifying respiratory illness in a pediatric population. *Pediatr Emerg Care* 2004;**20**:355-60.
17. **Begier EM**, Sockwell D, Branch LM, *et al*. The National Capitol Region's emergency department syndromic surveillance system: Do chief complaint and discharge diagnosis yield different results? *Emerg Infect Dis* 2003;**9**:393-6.
18. **Gesteland PH**, Gardner RM, Tsui FC, *et al*. Automated syndromic surveillance for the 2002 winter Olympics. *J Am Med Inform Assoc* 2003;**10**:547-54.
19. **Silka PA**, Geiderman JM, Goldberg JB, *et al*. Demand on ED resources during periods of widespread influenza activity. *Am J Emerg Med* 2003;**21**:534-9.
20. **Lober WB**, Trigg LJ, Karras BT, *et al*. Syndromic surveillance using automated collection of computerized discharge diagnoses. *J Urban Health* 2003;**80**(suppl 1):i97-106.
21. **Flaherty J**, Gillam M. Correlation of West Nile virus infection with emergency department chief complaints by using a passive syndromic surveillance model. (Abstract). *MMWR* 2004;**53**(suppl):237.
22. **Paladini M**. Daily emergency department surveillance system-Bergen County, New Jersey. *Syndromic surveillance: reports from a national conference*, 2003. *MMWR* 2004;**53**(suppl):47-9.
23. **Lewis M**, Pavlin J, Mansfield J, *et al*. Disease outbreak detection system using syndromic data in the greater Washington DC area. *Am J Prev Med* 2002;**23**:180-6.

- 24 **Irvin C**, Nouhan P, Rice K. Syndromic analysis of computerized emergency department patients' chief complaints: an opportunity for bioterrorism and influenza surveillance. *Ann Emerg Med* 2003;**41**:447–52.
- 25 **Mostashari F**, Fine A, Das D, *et al.* Use of ambulance dispatch data as an early warning system for communitywide influenzalike illness, New York City. *J Urban Health* 2003;**80**(suppl):i43–9.
- 26 **Lazarus R**, Kleinman K, Dashevsky I, *et al.* Use of automated ambulatory-care encounter records for detection of acute illness clusters, including potential bioterrorism events. *Emerg Infect Dis* 2002;**8**:753–60.
- 27 **Reis BY**, Mandl KD. Syndromic surveillance: the effects of syndrome grouping on model accuracy and outbreak detection. *Ann Emerg Med* 2004;**44**:235–41.
- 28 **Day FC**, Schriger DL, La M. Automated linking of free-text complaints to reason-for-visit categories and International Classification of Diseases diagnoses in emergency department patient record databases. *Ann Emerg Med* 2004;**43**:401–9.
- 29 **Osaka K**, Takahashi H, Ohya T. Testing a symptom-based surveillance system at high-profile gatherings as a preparatory measure for bioterrorism. *Epidemiol Infect* 2002;**129**:429–34.
- 30 **Cho J**, Kim J, Yoo I, *et al.* Syndromic surveillance based on the emergency department in Korea. *J Urban Health* 2003;**80**:i124.
- 31 **Mac Kenzie WR**, Hoxie NJ, Proctor ME, *et al.* A massive outbreak in Milwaukee of cryptosporidium infection transmitted through the public water supply. *N Engl J Med* 1994;**331**:161–7.
- 32 **Stirling R**, Aramini J, Ellis A, *et al.* Waterborne cryptosporidiosis outbreak, North Battleford, Saskatchewan, Spring 2001. *Can Commun Dis Rep* 2001;**27**:185–92.
- 33 **Edge V**, Pollari F, Lim G, *et al.* Syndromic surveillance of gastrointestinal illness using pharmacy over-the-counter sales. A retrospective study of waterborne outbreaks in Saskatchewan and Ontario. *Can J Public Health* 2004;**95**:446–50.
- 34 **Heffernan R**, Mostashari F, Das D, *et al.* Syndromic surveillance in public health practice, New York City. *Emerg Infect Dis* 2004;**10**:858–64.
- 35 **Kulldorff M**, Zhang Z, Hartman J, *et al.* Benchmark data and power calculations for evaluating disease outbreak detection methods. Syndromic surveillance: reports from a national conference, 2003. *MMWR* 2004;**53**(suppl):144–51.
- 36 **Buehler JW**, Berkman RL, Hartley DM, *et al.* Syndromic surveillance and bioterrorism-related epidemics. *Emerg Infect Dis* 2003;**9**:1197–04.
- 37 **Stato M**, Schonlau M, Mariano L. Syndromic surveillance: Is it worth the effort? *Chance* 2004;**17**:20–5.
- 38 **Henry J**, Magruder S, Snyder M. Comparison of office visit and nurse advice hotline data for syndromic surveillance—Baltimore-Washington, DC, Metropolitan Area, 2002. Syndromic surveillance: reports from a national conference, 2003. *MMWR* 2004;**53**(suppl):112–16.
- 39 **Peterson D**, Perencevich E, Harris A, *et al.* Using existing electronic hospital data for syndromic surveillance. *J Urban Health* 2003;**80**:i122–3.
- 40 **Balter S**, Weiss D, Hanson H, *et al.* Three years of emergency department gastrointestinal syndromic surveillance in New York City: What have we found? Syndromic surveillance: reports from a national conference, 2004. *MMWR* 2004;**54**(suppl):175–80.
- 41 **Cooper D**, Smith G, Baker M, *et al.* National symptom surveillance using calls to a telephone health advice service—United Kingdom, December 2001–February 2003. Syndromic surveillance: reports from a national conference, 2003. *MMWR* 2004;**53**(suppl):179–83.
- 42 **Dembek Z**, Carley K, Siniscalchi A, *et al.* Hospital admissions syndromic surveillance—Connecticut, September 2001–November 2003. Syndromic surveillance: reports from a national conference, 2003. *MMWR* 2004;**53**(suppl):50–2.
- 43 **Terry W**, Ostrowsky B, Huang A. Should we be worried? Investigation of signals generated by an electronic surveillance system—Westchester County, New York. Syndromic surveillance: reports from a national conference, 2003. *MMWR* 2004;**53**(suppl):190–5.
- 44 **Reingold A**. If syndromic surveillance is the answer, what is the question? *Biosecure Bioterror* 2003;**1**:77–81.
- 45 **Steiner-Sichel L**, Greenko J, Heffernan R, *et al.* Field investigations of emergency department syndromic surveillance signals—New York City. Syndromic surveillance: reports from a national conference, 2003. *MMWR* 2004;**53**(suppl):184–9.
- 46 **Centers For Disease Control and Prevention**. Diagnosis and management of food borne illness: a primer for physicians and other health care professionals. *MMWR* 2004;**53**:RR–4.
- 47 **Hellard M**, Sinclair M, Fairley C, *et al.* An outbreak of cryptosporidiosis in an urban swimming pool: why are such outbreaks difficult to detect? *Aust N Z J Public Health* 2000;**24**:272–5.
- 48 **Kasowski E**, Gackstetter G, Sharp T. Foodborne illness: new developments concerning an old problem. *Curr Gastroenterol Rep* 2002;**4**:308–18.
- 49 **Townes J**. Acute infectious gastroenteritis in adults: seven steps to management and prevention. *Postgrad Med* 2004;**115**:11–18.
- 50 **Chin J**. *Control of communicable diseases manual*. 17th ed. Washington, DC: American Public Health Association, 2000.
- 51 **Mulazimoglu L**, Yu V. Can Legionnaires disease be diagnosed by clinical criteria? *Chest* 2001;**120**:1049–53.
- 52 **Fiore A**, Butler J. Detecting nosocomial Legionnaires' disease. *Infect Med* 1998;**15**:625–35.