Web-Based Surveillance Systems for Human, Animal, and Plant Diseases

LAWRENCE C. MADOFF1 and ANNIE LI2

1ProMED-mail, University of Massachusetts Medical School, Massachusetts Department of Public Health, Jamaica Plain, MA 02130; 2City University of Hong Kong, Department of Biology and Chemistry, Kowloon Tong, Kowloon, Hong Kong

ABSTRACT The emergence of infectious diseases, caused by novel pathogens or the spread of existing ones to new populations and regions, represents a continuous threat to humans and other species. The early detection of emerging human, animal, and plant diseases is critical to preventing the spread of infection and protecting the health of our species and environment. Today, more than 75% of emerging infectious diseases are estimated to be zoonotic and capable of crossing species barriers and diminishing food supplies. Traditionally, surveillance of diseases has relied on a hierarchy of health professionals that can be costly to build and maintain, leading to a delay or interruption in reporting. However, Internet-based surveillance systems bring another dimension to epidemiology by utilizing technology to collect, organize, and disseminate information in a more timely manner. Partially and fully automated systems allow for earlier detection of disease outbreaks by searching for information from both formal sources (e.g., World Health Organization and government ministry reports) and informal sources (e.g., blogs, online media sources, and social networks). Web-based applications display disparate information online or disperse it through e-mail to subscribers or the general public. Web-based early warning systems, such as ProMED-mail, the Global Public Health Intelligence Network (GPHIN), and Health Map, have been able to recognize emerging infectious diseases earlier than traditional surveillance systems. These systems, which are continuing to evolve, are now widely utilized by individuals, humanitarian organizations, and government health ministries.

INTRODUCTION

The early detection of infectious diseases is critical for early response and control of disease outbreaks. If outbreaks can be detected early, public health interventions may be able to reduce the size of the outbreak and mitigate its consequences. With more than 75% of emerging infectious diseases estimated to be zoonotic (1), the One Health concept can be applied to the surveillance of disease outbreaks that affect human, animal, and plant species. Traditional surveillance systems often rely on time-consuming, labor-intensive, and expensive methods (such as laboratory data or clinician case-report forms) to collect outbreak data, which then pass through multiple levels of health professionals before being confirmed and announced. This conventional system can delay public health responses that would otherwise prevent the spread of disease. However, current information technology allows the use of computers, mobile phones, remote sensing, and Internet searches to globally communicate and share information on disease outbreaks (2, 3). This information is often free to the general public, with access to valuable information from websites and Internet-based applications.

In comparison with traditional surveillance systems, Web-based surveillance systems enable health professionals to identify and disseminate disease outbreak information more rapidly. Many of these systems
are automated and search for outbreak information from Web-accessible resources that include official reports (e.g., World Health Organization [WHO], World Organisation for Animal Health [Office International des Epizooties, or OIE], Food and Agriculture Organization [FAO], and Centers for Disease Control and Prevention [CDC] reports) and unofficial reports (e.g., news outlets, blogs, social networks, websites, mailing lists, and discussion sites). By aggregating a plethora of disparate outbreak information, users can visit a Web-based surveillance site to seek information instead of searching through multiple websites. They may also instantly receive alerts through e-mail via unrestricted subscriptions (4, 5).

One of the first examples of early recognition of a disease outbreak occurred in 2002 when a severe acute respiratory syndrome outbreak occurred in Guangdong Province, China. The International Society for Infectious Diseases’ (ISID) Program for Monitoring Emerging Diseases (ProMED) and the Global Public Health Intelligence Network (GPHIN) played an important role in detecting the outbreak at an early stage (6, 7). Cases could be found as early as November 2002 from Guangdong Province (2, 8), at least 2 months prior to the WHO’s officially declaring cases of a novel respiratory disease.

ProMED-mail, which began operation in 1994 as a mailing list serving about 40 members, explored the ability of informal Internet-based reports from clinician reports, news media, and other sources distributed to a rapidly expanding user base. ProMED stresses transparency and in many ways functioned as a social network, since users interacted with one another, seeking clarification regarding outbreaks and finding expertise among its participants. ProMED has always been available to any interested user and free of charge. Subject area and regional experts screen each report and provide commentary and context (6).

GPHIN, which debuted in 1997, pioneered the use of automatic Web crawling of media sites to discover potential outbreaks. The software searches and reads numerous media sites in multiple languages for stories that might relate to infectious disease outbreaks. It distributes reports to a network comprising largely official public health agencies that subscribe to the service. A team of specialists analyzes reports for subscribers (7).

Since the development of ProMED-mail in 1994 and GPHIN in 1997, many other Internet-based surveillance systems have been developed. HealthMap, for example, which began in 2006, automatically mines a wide variety of public sources and aggregates outbreak reports geospatially. Other systems include Argus, MedISys, and EpiSPIDER. Internet users of these surveillance systems include public health officials, clinicians, and international travelers (4). Public health agencies, such as the WHO, also rely on these sources of information as early evidence of an outbreak and communicate findings via the WHO’s Global Outbreak Alert and Response Network (GOARN) (9).

The use of Web-based surveillance systems offers a technologically advanced means of recognizing disease outbreaks in reduced time (2) so that actions can be taken to control diseases that may affect multiple species, food resources, and the environment. Health affects not only an individual or even an individual species, but also other species in a global population. Zoonotic diseases, those that can be transmitted from animals to humans, continue to emerge. Plant diseases plague crops that animal and human populations depend on for their nutrition and health. Plant diseases may have profound effects on human and animal populations (a frequently cited example is the potato blight that affected Europe in the 19th century and led to mass mortality and population movements, particularly in Ireland). In addition, food crops have been considered a target of bioterrorism, since crop loss could lead to human and animal mortality and social disruption. Moreover, transboundary diseases are capable of crossing geographic barriers and political borders through movements, trade, and transportation. These examples are evidence that it is necessary to improve communication, build collaborations, and apply epidemiological intelligence for the promotion of global health. In this article, different digital surveillance systems and their functional capabilities are discussed.

OVERVIEW OF SURVEILLANCE SYSTEMS

Surveillance can be defined as the “ongoing systematic and timely collection, analysis, interpretation, and dissemination of information about the occurrence, distribution, and determinants of diseases” (2). By providing an effective surveillance system, one can monitor diseases to recognize outbreaks so that public health measures can be taken in a timely manner.

In a traditional surveillance system, official reports are often used to count the number of confirmed cases. The process involves a person becoming ill, seeking care, having proper diagnostic testing, and being reported as a confirmed case to the health department or government agency. However, official reports can be slow to receive if reported at all, with the possibility of every line
within the surveillance system being breached during the process. For example, official reports depend highly on people—the patient and/or health care provider—to recognize the illness and report it to the appropriate entity. If the patient fails to seek medical care or the health care provider fails to make the correct diagnosis or properly submit a report, the case can be missed by the existing surveillance system. In addition, if a laboratory improperly runs a diagnostic test to confirm the etiological agent or fails to report confirmed results to the appropriate authorities, the case is also lost before it reaches the health ministry. In the traditional surveillance system, underreporting of cases can be problematic because of the possibility of losing cases within the traditional process.

This is especially true for novel infectious diseases, which are difficult to detect because the etiological agent is often unknown and the diagnostic test and health expertise not present. Moreover, resource-limited settings especially face difficulties in detecting newly emerged infectious diseases due to the lack of proper expertise and facilities. Yet they are more threatened by infectious diseases (10). Detecting diseases may therefore depend on other sources of information, since their surveillance systems are limited in scope if available at all.

Over the years, surveillance systems have grown in capacity as early warning systems and have evolved by utilizing today’s advancements in technology from computers to phones to the Internet (2). Web-based surveillance systems are becoming more powerful tools for reaching Internet users globally. Systems can be automated to collect articles on a 24/7 basis by using search queries to aggregate data. Alternatively, users who may be eyewitnesses to an outbreak can report cases to the Web-based surveillance system via e-mail, an online submission, or a smartphone application. Analysis and organization of data may also be made by machine learning and verified by health experts. Finally, information may be displayed on the Internet through a website or periodic reports or e-mailed to subscribers (4, 5).

**TYPES OF INFORMATION SOURCES FOR DATA COLLECTION**

Two main types of data sources exist for Web-based surveillance systems: formal sources of information and informal sources of information (Table 1).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Formal information</th>
<th>Informal information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Types of surveillance</td>
<td>Passive and active surveillance, traditional and Web-based surveillance</td>
<td>Passive and active surveillance, Web-based surveillance</td>
</tr>
<tr>
<td>Examples of resources</td>
<td>Health ministry reports, WHO reports, laboratory data, clinical data</td>
<td>Blogs, discussion forums, mailing lists, media outlets, social networks, Internet articles, eyewitness reports</td>
</tr>
<tr>
<td>Reliability of information</td>
<td>More reliable because of confirmation</td>
<td>Less reliable due to false reports</td>
</tr>
<tr>
<td>Timeliness</td>
<td>Reports may be delayed because of need for official confirmation or approval</td>
<td>Often earlier reporting</td>
</tr>
<tr>
<td>Case definition</td>
<td>Often based on laboratory confirmation</td>
<td>Not necessarily laboratory confirmed, occasionally based on uninformed or misleading information</td>
</tr>
<tr>
<td>Outbreak investigation</td>
<td>Useful for serious diseases, such as reportable diseases or diseases of significant socioeconomic impact</td>
<td>Useful for emerging infectious diseases or diseases not reported by official sources</td>
</tr>
<tr>
<td>Limitations</td>
<td>Labor-intensive, requires passing through multiple levels of health professionals, underreporting of cases, may not include diseases that are not reportable, may be more costly to obtain formal information</td>
<td>Higher risk of false reports (mis- or disinformation), reporting bias</td>
</tr>
</tbody>
</table>

**TABLE 1** Formal versus informal information sources
disease outbreaks. In addition, biological samples suspected of containing pathogens can be collected and submitted to public health, academic, or military labs for testing. More detailed information can further be sought in regards to these samples to determine clues about their origin. Samples may be genotyped and phenotyped and the information then stored in a reference databank. Developed countries often access these databanks to compare past genotypes and phenotypes of diseases to current strains to determine where the outbreak originated. The CDC’s PulseNet is an example of this system and focuses on surveillance of food-borne and waterborne bacterial diseases using DNA electrophoretic fingerprints of pathogens, which are stored in the system (2).

In contrast, informal sources of data may provide up-to-date, local information for early recognition of disease outbreaks, even in locations that are resource limited and lack a traditional public health infrastructure. Informal channels of information include news reports, blogs, discussion rooms, social networks, and mailing lists (3, 4). Nongovernmental organizations can also contribute to disease reporting. Examples include the Red Cross and Red Crescent, Médecins Sans Frontières, Medical Emergency Relief International (Merlin), and religious organizations (3). Although informal sources can overload a surveillance system and may result in an increase in reporting bias or greater numbers of false reports, history has proven that informal sources are advantageous in detecting disease outbreaks earlier than with traditional sources. Informal reports have also discouraged countries from hiding outbreak information and encouraged surveillance systems to utilize different forms of data (4).

Today, many Web-based surveillance systems use informal sources of data to gain early knowledge on disease outbreaks. An example is the WHO’s GOARN, which currently detects its verified outbreaks mostly through nontraditional information sources (9). The use of informal sources of information is further supported by the WHO’s updated International Health Regulations of 2005 (became effective in 2007), which stipulate that the WHO is permitted to take preventative actions in response to these types of data (11). Surveillance systems that currently utilize nontraditional sources of epidemiological data include HealthMap, ProMED-mail, Emergency Prevention System Global Animal Disease Information System (EMPRES-i), and GPHIN, among others.

WEB-BASED SURVEILLANCE SYSTEMS

A number of Web-based surveillance systems have been developed for surveillance of infectious diseases of humans, animals, and plants (Table 2).

TABLE 2 Examples of digital surveillance systems

<table>
<thead>
<tr>
<th>Surveillance system</th>
<th>Website</th>
<th>Animal cases</th>
<th>Human cases</th>
<th>Plant cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>BioCaster Global Health Monitor</td>
<td><a href="http://biocaster.nii.ac.jp">http://biocaster.nii.ac.jp</a></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Emergency Prevention System</td>
<td><a href="http://empres-i.fao.org">http://empres-i.fao.org</a></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Global Animal Disease Information System (EMPRES-i)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EpiSPIDER</td>
<td><a href="http://www.epispider.org/">http://www.epispider.org/</a></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>European and Mediterranean Plant Protection Organization (EPPO)</td>
<td><a href="http://www.eppo.int/QUARANTINE/Alert_List/alert_list.htm">http://www.eppo.int/QUARANTINE/Alert_List/alert_list.htm</a></td>
<td>X</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>GeoSentinel</td>
<td><a href="http://www.istm.org/geosentinel/main.html">http://www.istm.org/geosentinel/main.html</a></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GermTraX</td>
<td><a href="http://www.germtrax.com/">http://www.germtrax.com/</a></td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Google Flu Trends</td>
<td><a href="http://www.google.org/flutrends/">http://www.google.org/flutrends/</a></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global Public Health Intelligence Network (GPHIN)</td>
<td><a href="https://www.gphin3.net/">https://www.gphin3.net/</a></td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>International Plant Protection Convention (IPCC)</td>
<td><a href="https://www.ippc.int/">https://www.ippc.int/</a></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MedISys</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>North American Plant Protection Organization (NAPPO)</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ProMED-mail</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Wildlife Data Integration Network (WDIN)</td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>World Animal Health Information Database (WAHID)</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Last updated April 2013.
**ProMED-mail**

ProMED-mail ([http://www.promedmail.org/](http://www.promedmail.org/)) was founded as an Internet-based early warning system for emerging infectious diseases in 1994, and is now a program of the ISID (Fig. 1). The system was one of the earliest Internet-based reporting systems to be utilized for disseminating information on infectious diseases and acute, toxic exposures that threaten human, animal, and food plant health. It operates 7 days a week, and reports are collected from both official and unofficial sources, including alerts from subscribers. Information is sought, analyzed, and commented upon by subject area experts in order to be posted on the ProMED website and e-mailed directly to subscribers through a freely available and open mailing list. Reports are also available via Facebook and Twitter. As ProMED seeks to promote communication among different health professionals, subscribers can discuss their views and work with others; ProMED has been cited as an early form of social networking. Currently, ProMED-mail is available in multiple languages and has a total of more than 60,000 subscribers from at least 185 countries. Regional networks in the Mekong

---

**Figure 1** Website of ProMED-mail ([http://www.promedmail.org/](http://www.promedmail.org/)). ProMED is a service of the ISID and provides reports, moderated by a panel of experts, on outbreaks of emerging diseases in humans, animals, and plants. [doi:10.1128/microbiolspec.OH-0015-2012.f1](https://doi.org/10.1128/microbiolspec.OH-0015-2012.f1)
area of Southeast Asia, the former Soviet Union (in Russian), Latin America (in Spanish and Portuguese), and Africa (in French and English) help less information-rich countries benefit from informal information exchange, and also enhance detection of diseases in areas that are potential hot spots of disease emergence. ProMED is unusual among disease surveillance systems in its coverage of plant diseases affecting food crops and includes a plant pathologist on its staff. These reports are picked up and monitored by some other systems such as HealthMap.

**GPHIN**

In 1997, the GPHIN (home page, https://gphin3.net/; description at http://biosurveillance.typepad.com/files/gphin-manuscript.pdf) was developed by the Public Health Agency of Canada in collaboration with the WHO. It is managed by the Agency’s Centre for Emergency Preparedness and Response. GPHIN pioneered the use of automated Web crawling and serves as an important early warning system for human, animal, and plant disease outbreaks; contaminated food and water; bioterrorist agents; chemical and radionuclear agent exposures; and natural disasters. It also covers drug and medical device safety problems. As one of the earliest digital surveillance networks, GPHIN, along with ProMED-mail, has been credited with detecting severe acute respiratory syndrome before official reports were published. The Microsoft/Java application runs 24/7 real-time in seven languages (Arabic, English, French, Russian, Farsi, Chinese, and Spanish), and data are interpreted by a team of GPHIN analysts based at the Public Health Agency of Canada. A subscription fee is required to access the secure website, with the cost depending on the type of organization.

**HealthMap**

HealthMap (http://www.healthmap.org/en) is a non-profit organization based at Children’s Hospital Boston and Harvard Medical School. Founded in 2006 (Fig. 2), this digital surveillance system monitors human, animal, and plant diseases and is freely available online without a subscription fee. The automated system operates 24/7 to monitor, organize, integrate, filter, visualize, and disseminate information in near real time. Disparate sources include both informal and formal resources.

![Figure 2](https://www.healthmap.org/en/) Website of HealthMap (http://www.healthmap.org/en/), based at Children’s Hospital Boston, which shows, on an interactive map, infectious disease outbreaks automatically derived from numerous sources and curated by a human team. doi:10.1128/microbiolspec.OH-0015-2012.f2
in nine languages. The Linux/Apache/MySQL/PHP
application searches from more than 50,000 sources,
including news aggregators, eyewitness reports, and
verified official reports, for disease names, symptoms,
and key words and phrases. Alert sources are organized
and filtered using Fisher-Robinson Bayesian filtering and
then analyzed by trained personnel to ensure the proper
organization of collected data. Users may browse and
search for outbreaks by specific location or disease. The
original source of information is also linked and avail-
able for users to peruse online. A new addition to the
website is “The Disease Daily,” which highlights out-
break news for the general public. Individual users may
also participate in submitting their own outbreak alert
online or through a mobile application, “Outbreaks
Near Me.” This smartphone application allows users to
view and search for reports on an interactive map, with
the option to receive alerts of local outbreaks.

**EMPRES-i**
The United Nations FAO created EMPRES-i (http://
empres-i.fao.org) to monitor transboundary animal dis-
eases and animal diseases of high impact (Fig. 3). This
Web-based application provides an early warning sys-
tem by utilizing both formal and informal resources,
including project and field mission reports, institutions,
nongovernmental organizations, government ministries
of agriculture and health, FAO or United Nations
representatives, public domains, media, and other
Web-based surveillance systems. It provides advanced
technology to collect sources, analyze animal disease
information, and provide information to users. Inform-
ation is displayed on outbreak maps with the option of
viewing additional geospatial layers on livestock
populations, biophysical characteristics, socioeconomics
of an area, etc., which are provided by the Global
Livestock Production and Health Atlas (GLiPHA). Data
and graphs can also be created in different formats (e.g.,
PDF, CSV, or Excel) to be exported for personal use.

**World Animal Health Information System (WAHIS)**
The OIE, established in 1924, consists of member
countries and territories that are legally obliged to report
exceptional disease events of OIE-listed animal diseases,
including zoonotic diseases and emerging animal dis-
cases. Information is freely available to the general
public through the World Animal Health Information
Database (WAHID), which provides data from the WAHIS (http://www.oie.int/wahis_2/public/wahid.php/Wahidhome/Home). A more secure site is also available to authorized users, primarily OIE delegates and authorized focal points, that are obligated to notify the OIE within 24 hours of confirming a significant animal disease. Information that is collected is verified and validated by the OIE before being published in English, French, and Spanish and sent through the OIE-Info list to delegates. WAHID differs from EMPRES-i by publishing only confirmed disease reports (formal sources of information) to the general public and sharing unconfirmed information (informal sources) to OIE members. The system allows users to search for outbreaks by specific animal diseases or countries/regions. The website provides information on disease characteristics and timelines, veterinary and laboratory services, animal populations, human cases of zoonoses, sanitary conditions of countries, and control measures. The interface also displays maps of disease outbreaks to view disease distribution within a specific country or the world. Older data from 1996 to 2004 may also be sought online from the former system, HandiStatus II, which was replaced by WAHID in 2005. However, a newer version of WAHIS was released in August 2012, with improved reports of wildlife diseases. In addition to routine monitoring of OIE-listed wildlife diseases, the updated system allows voluntary reporting of wildlife diseases that are not officially OIE notifiable.

SEARCH TERM SURVEILLANCE

Syndromic surveillance is another method used for early detection of diseases before they are diagnostically confirmed. By detecting individual and population health indicators, one can potentially recognize an outbreak before traditional reports. Examples of indicators may include an abnormal increase in over-the-counter medication purchases, school and work absenteeism, emergency hospital visits, or data aggregated from Web-based clickstream and key word searching (12).

Online search engines are commonly used by the general public to seek health information. During large-scale outbreaks, an increase in searches for specific diseases may mirror or predict disease outbreaks. For instance, past studies using search queries from Google and Yahoo have shown patterns mimicking disease activity during the H1N1 pandemic (13, 14). Other studies have also used key terms aggregated from social networks, such as Twitter, to show real-time estimates of disease activity (15).

Google Trends

Google Trends (http://www.google.com/trends/) is an automated tool used to monitor search queries in near real time. Although it is available to be used by all Internet users including those seeking non-health-related information, search queries for specific diseases may be aggregated and analyzed to estimate disease activity and geospatial patterns. The more widely studied version is Google Flu Trends (www.google.com/flutrends/), which monitors health-seeking behavior by users who potentially show influenza symptoms. Data on their search queries for influenza-like illnesses may be viewed through a search volume index graph, which is updated on a daily basis. Past studies have shown a strong correlation between Google Flu Trends and the U.S. CDC data from 2003 to 2008. The study also suggests earlier detection by 1 to 2 weeks before the CDC influenza-like illness surveillance reports, which could potentially be followed by an earlier public health response. Although a freely available tool, Google Trends is best used only for highly prevalent diseases, such as influenza, and in developed countries, where a larger population of Internet users are located (13).

TWITTER

Social networks are also powerful tools in connecting people in ways that allow them to share information. Twitter is an example of a free social network that can be used to track health information. The microblog has millions of users who “tweet” or submit messages of up to 140 characters. Most tweets embedded within the Twitter stream contain general conversations, sharing of information, and additional links to items of interest. In one study, researchers used Twitter’s streaming application programmer’s interface to collect a subset of Twitter messages that contained key words related to influenza-like illness. Using machine learning methods, they were able to aggregate information to measure H1N1 activity and public interest/concern in the United States. However, there is still difficulty with using this type of information because of the lack of specificity, as some tweets may be due to an interest in H1N1 while others may be tweeted because of an onset of symptoms (15).

CONCLUSION

Digital, Internet-based surveillance systems have changed the way we detect disease outbreaks and have created opportunities to conduct surveillance globally at the local, regional, and international level. Automated systems
Web-Based Surveillance Systems for Human, Animal, and Plant Diseases

that use machine learning technology function 24/7 and are capable of early monitoring, early warning, and promoting early responses to diseases of public health significance. In addition, information can instantly be displayed on websites or disseminated to subscribers of Web-based surveillance systems.

Currently, Internet-based surveillance systems are better able to detect outbreak information in countries with high levels of Internet access. However, developing countries, while often facing the greatest burden of diseases (10), often lack the public health infrastructure to detect and prevent the spread of infection. Unlike industrialized countries, which may seek to prevent pathogens from entering and reemerging, resource-limited countries are challenged with a more serious problem. Infectious diseases are often endemic to these countries, which must concentrate their efforts on detecting outbreaks to decrease fatalities, reduce spread, and prevent harm to trade and tourism (3). One method of improving disease surveillance in areas that lack resources and Internet penetration is through the use of mobile phones and other hand-held devices that can be used for reporting cases through Short Message Services. The use of such technology has already been rising in resource-limited settings (16) and is likely to continue in the near future to help with disease reporting.

Other challenges for Internet-based disease detection include the obligation to respect privacy. For example, search queries can contain specific geographic information, such as a patient’s address. Therefore, care must be taken to use appropriate spatial aggregation according to the public health need, such as mapping administrative units or using spatial skewing (17).

More research will also be needed to solve the problem of overload of information that is often unstructured and difficult to understand. Because the sensitivity and specificity of Internet-based surveillance is often unclear (5), information should be verified by trained personnel, who should also utilize advanced techniques such as machine learning to help organize data. False reports and reporting bias may occur with nontraditional sources of information (4), and media tend to focus their reports on those that can make headlines but fail to accurately communicate risks. An example of this phenomenon occurred in 2011 when Spanish cucumbers were initially and mistakenly blamed for an Escherichia coli 0104:H4 outbreak causing gastrointestinal disease and hemolytic-uremic syndrome in Europe. These false reports dominated the media, causing a scare that resulted in a drop in cucumber sales. Political and economic tensions grew within the Spanish government, and the culprit was later confirmed as fenugreek sprouts that came from seeds imported from Egypt (18). This example shows that although digital surveillance can help speed the detection of outbreaks, it can only be as good as its data.

Another challenge is harmonizing data from parallel surveillance programs that overlap in function. Systems are currently differentiated by their functions, such as concentrating on specific regions, species types, and types and sources of data. The situation is further complicated by diseases that cross national borders and/or affect more than one species. For instance, FAO’s EMPRES-i and OIE’s WAHID focus their surveillance efforts on animal health, but many of these infectious agents also affect humans. Although these sources provide information on zoonotic diseases, the direct surveillance of human cases is typically found in other epidemiological systems. Collaborative efforts, therefore, need to be made to bring together resources to build integrated applications that fulfill a variety of needs and communicate disease risks to different populations.

We live in a highly globalized world, where international travel and trade of food, biological products, and live species can serve as potential sources of infectious disease outbreaks and the emergence and reemergence of pathogens (3). By bringing together different types of professionals, including epidemiologists, computer scientists, physicians, veterinarians, and public health professionals from both the private and public sectors, better surveillance systems can be developed to represent animal, human, and plant infections. Furthermore, an integrated system could provide insight into antimicrobial resistance, water quality, animal reservoirs, and insect vectors (8). The need for a One Health approach is evident in the 21st century, as emerging diseases continue to cross species barriers as well as geographic ones.

ACKNOWLEDGMENTS

We thank our colleagues at ProMED and HealthMap for their dedication. Dagmar Hanold, ProMED plant disease moderator, provided information on plant disease resources on the Web.

The authors have no commercial affiliations, consultancies, stock or equity interests, or patent-licensing arrangements that could be considered to pose a conflict of interest regarding the manuscript.

REFERENCES


