Privacy and Security in Public Health: Maintaining the Delicate Balance between Personal Privacy and Population Safety

Dixie B. Baker, Ph.D.
Science Applications International Corporation
Dixie.B.Baker@saic.com

Abstract

Amidst threats of pandemic avian influenza and bioterrorist attack, public health surveillance and preparedness have never been more important. Early detection of biological events, electronic reporting of laboratory test results, efficient exchange of case reports across jurisdictions, and timely alerting of health threats are critical components of effective health protection. Essential to public health surveillance and preparedness is the timely availability of information relating to individuals' healthcare behaviors and clinical conditions – posing a threat to personal privacy. Public health is challenged to maintain an optimal balance between protecting the nation's health and respecting the personal privacy of its citizens.

1. Introduction and Update on PCASSO

I am honored to have been selected as Distinguished Practitioner for this year's Annual Computer Security Applications Conference (ACSAC). I had the pleasure of serving on the ACSAC Board during the early years, along with Marshall Abrams, Tom Haigh, Dan Faigin, Ron Gove, and others. I am equally honored that the Health Information Management and Systems Society (HIMSS) has requested permission to include this paper in the next release of its Privacy and Security Toolkit [1].

I last participated in the ACSAC in 1997, when my research team received the Best Paper Award for our paper about our Patient Centered Access to Secure Systems Online (PCASSO) project [2]. Sponsored by the National Library of Medicine, PCASSO was very successful and is widely regarded as seminal in both patient empowerment [3] and the use of high-assurance security methods and technology in healthcare systems [4]. PCASSO was the first experiment to enable patients to access and view their own clinical information over the Internet, including highly sensitive information relating to HIV/AIDS, abortion, adoption, and genetics. We achieved high assurance through the use of multi-level, label-based (a.k.a., mandatory) access control and an assortment of architectural and design approaches viewed by many at the time as "paranoid." To this day, PCASSO is the only project that has applied multi-level security to protect healthcare information – although enabling patients to access at least portions of their health records is becoming more commonplace. Unfortunately, even today, the healthcare provider community treats security as primarily a compliance issue, rather than an essential prerequisite to information confidentiality, data integrity, system stability, and continuity of operations. But that's another story.

My Distinguished Practitioner paper focuses on public health – specifically, the security and privacy challenges the public health community is facing as it seeks to protect the U.S. population from disease outbreaks and bioterrorist attacks. For the past three years, I have supported the National Center for Public Health Informatics (NCPHI) at the Centers for Disease Control and Prevention (CDC), primarily doing research and architecture studies relating to the Public Health Information Network (PHIN). Most of my work has focused on defining an architecture that will enable public health agencies to quickly construct and distribute event-specific data-collection instruments for epidemiologists and first-responders to use to collect information that is semantically consistent, computer consumable, and immediately analyzable. While my CDC work has not had a security focus, my experience has given me the opportunity to learn about the important role that public health plays in our lives and about the challenges that public health faces as it seeks to achieve and maintain an optimal balance between personal privacy and effective health surveillance, outbreak detection, preparedness, and response.

The following two sections of this paper provide a basic introduction to public health, its current challenges, and its future direction toward

Proceedings of the 22nd Annual Computer Security Applications Conference (ACSAC'06)

biosurveillance and nationwide interoperability. Section 4 establishes the legal and regulatory foundations for privacy and security in public health, including an explanation of the enabling requirements contained in the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule. Section 5 discusses security and privacy challenges and offers a risk model of the complex interplay between personal privacy and population health. Section 6 maps public health security safeguards and de-identification policy to the requirements the HIPAA Privacy Rule prescribes for healthcare providers and insurers.

2. A Primer on Public Health

The Institute of Medicine has defined public health as "what we as a society do collectively to assure the conditions in which people can be healthy" [5]. Although government agencies at the federal, state, tribal, and local levels bear explicit legal responsibility for protecting public health, many public and private organizations and institutions in our society contribute to both the science and practice of public health. These public health partners include clinical practitioners, health departments, laboratories, disease programs, researchers, and social services. The complex responsibilities and interactions among these partners require significant coordination of information technology and information exchange protocols to meet public health preparedness and response objectives.

A basic science of public health is epidemiology, the study of the occurrence and causes of health effects in human populations [6]. The work of public health is often invisible to the average person – very few people regularly visit their "family epidemiologist." An analogy attributed to a medical school epidemiology course is that a clinician (e.g., family physician) tries to decide what kind of disease a person has, whereas an epidemiologist tries to determine what kind of person has a disease or condition (and what factors can be manipulated to prevent that disease or condition) [7]. An important tool of public health professionals is public health informatics, the systematic application of information and computer science and technology to public health practice, research, and learning [8].

The basic responsibilities of public health are [9]:

1) To prevent epidemics and the spread of disease
2) To protect against environmental hazards
3) To prevent injuries
4) To promote and encourage healthy behaviors and mental health
5) To respond to disasters and assist communities in recovery
6) To assure the quality and accessibility of health services

Public health professionals at the federal level frequently specialize in a narrow field, such as lead poisoning or HIV/AIDS, whereas those working at the local level tend to be generalists with diverse responsibilities, and sometimes play the role of care provider for individuals unable to afford a family doctor. In general, federal and state public health agencies produce the research results, guidelines, and recommendations that local public health departments implement [7].

An important activity in disease prevention, detection, characterization, and eradication is public health surveillance, the ongoing systematic collection, analysis, and interpretation of health data for the purposes of improving the health and safety of a population [10]. Data are systematically collected and analyzed to determine what actions might need to be taken to prevent or control a disease or condition. Public health authorities generally rely on healthcare providers, laboratories, veterinarians, and others to report cases of reportable diseases and conditions when they are detected. Less commonly, health departments may contact or visit laboratories, hospitals, and providers to stimulate reporting of specific diseases and conditions.

Laws, regulations, and mandates for public health reporting (including the specific data items that are reported) fall under the authority of individual states and territories. Each state's health department and legislators decide which diseases and conditions to monitor within their state, and state legislatures may appropriate funding for conducting surveillance activities. Thus the diseases and conditions that are actively monitored will vary from state to state [11]. Healthcare providers and clinical laboratories report to their local, county, or state health departments all cases of diseases and conditions under surveillance in their state.

State public health officials, the Council of State and Territorial Epidemiologists (CSTE), and the CDC collaborate to determine which diseases and conditions should be "nationally notifiable" under the National Notifiable Diseases Surveillance System (NNDSS). The list of Nationally Notifiable Diseases and Conditions (NND) is reviewed annually and revised as new pathogens emerge or disease incidence declines. The 2006 NND list identifies 61 infectious diseases, with 31 subcategories [12]. However, states are not required by law to report NND cases to the CDC.

Local and state health departments monitor individuals ("cases") diagnosed with diseases and conditions under surveillance. Case reports generally include person-specific demographic information; clinical information, such as lab tests ordered and results; information about recent travels; and information about the entities (e.g., people, buildings, animals) with whom the case has come into contact. At the national level, each of the CDC's centers, institutes, and offices has its own surveillance mission and data requirements. Before a state sends case reports to the CDC, person identifiers are removed. (See Section 6.)

When the incidence of a disease under surveillance suddenly rises or when a new pattern of a set of symptoms emerges, an "outbreak" condition exists. Under outbreak conditions, timely investigation of reported cases and contacts is critical to effective containment and control. Local health departments spring into action to respond to the outbreak, perhaps requesting assistance from their state health department or the CDC. An outbreak response involves investigating known and suspected cases, tracing contacts, and implementing countermeasures such as vaccination, prophylaxis, or quarantine. Nearly as important as bringing the outbreak condition under control are managing the overall response and communicating with the community. If the outbreak is associated with a catastrophic natural disaster or is suspected to be the result of a bioterrorist act – either of which can cause extensive morbidity, mortality, economic loss, and social disruption – a "public health emergency" exists. When a public health emergency occurs, many roles and responsibilities, as well as data exchanges, are affected.

The sense of urgency that characterizes an outbreak can induce investigators to use whatever convenient means they may have to collect data. As a result, outbreak data may be collected using the investigators' own preferences for the questions asked, the terms and codes used to record responses, and an assortment of spreadsheets and databases to contain the data. This ad hoc approach can undermine efforts to control the outbreak, particularly when the outbreak crosses jurisdictional lines. Cases may be recorded in multiple databases (or spreadsheets) without recognizing duplicates; data may be added to some case records but not all; and laboratory data may not be linked to the patients who provided the specimens [13].

Historically, public health has used information technology tactically rather than strategically. As public health entered the twenty-first century, more than 100 different surveillance and health information systems were in use by the CDC's centers, institutes, and offices, and by state and local health departments. Many of these systems had been in place for several years, and were originally commissioned to detect simple disease and disability conditions. These systems were administered independently, used non-standardized formats for variable definition and grouping, and were unable to communicate with each another [14].

In some instances, computer technology proved to be a hindrance. Such was the case for the hantavirus outbreak in 1993. Data "locked" in local databases could not be analyzed or merged with data in other databases. Also, security measures designed to protect privacy and data availability have sometimes constrained responses and surveillance efforts by enforcing strict database design and handling requirements [13].

3. Public Health in Transition

The anthrax attack of 2001 caught the U.S. public health system off-guard and dramatically exposed the limitations of our public health infrastructure when confronted with a major public health crisis. Prior to our rude awakening to the reality of the bioterrorist threat, public health was already facing significant challenges, including cancer, obesity, violence, toxic environments, a large uninsured population, and health disparities. Rapid and dramatic changes were occurring in the scientific, social, cultural, technological, and global threat contexts of our nation's health. These challenges, coupled with unprecedented political and public scrutiny, prompted the public health system to critically examine its methods and priorities [15].

Out of this level of scrutiny and introspection came recognition of the need to realign priorities and to implement a public health infrastructure that could support disease surveillance with more effective coordination and collaboration among public health partners. Without such an infrastructure, bioterrorist events and outbreaks with potentially national impact could not be identified. Moreover, the use of multiple, non-integrated systems was contributing to an undesirable error rate in data records, an inefficient use of time and labor, a potential for under- and over-reporting, and a duplication of efforts [14].

Since 2001, the U.S. public health system has undergone some dramatic changes with respect to priorities, methods, and data sources. A realignment of priorities has resulted in an increased emphasis on and investment in preparedness and response capabilities. A cornerstone of the strategic plan for the future of public health is the Public Health Information Network (PHIN), which replaces the

"stovepipe" systems of the past with an interconnected web of interoperable information systems. The PHIN is aimed at encouraging and enabling the seamless sharing of electronic health information among public health partners, including local, state, and federal public health agencies as well as laboratories, hospitals, and other support organizations. The PHIN defines technology, data, vocabulary, and information security standards to enable the consistent exchange of health, disease tracking, and response data among public health partners; to protect the security of these data; and to ensure the network's reliability in times of national crisis [16].

The National Electronic Disease Surveillance System (NEDSS) program [17] is developing and implementing an integrated, standards-based solution to support on-going public health surveillance across the U.S. NEDSS functions support the identification and tracking of emerging infectious diseases, monitoring of disease trends, and response to outbreaks. The NEDSS program includes architecture, a set of terminology and messaging standards, and a reference implementation that will enable public health agencies to share information electronically, promoting timeliness and accuracy. Under federal funding, states are assessing their current systems and developing plans for either building upon the NEDSS reference implementation provided by the CDC or to acquire systems compatible with the NEDSS architecture and compliant with its standards.

To support the uniform collection and representation of information within the context of an outbreak, the CDC has developed an Outbreak Management System (OMS) available to state and local health departments. OMS is a complete application to support response to a public health emergency. The software includes a suite of tools for configuring data-collection instruments, collecting and analyzing data, managing controlled terminology, conducting case and contact investigations, and generating reports [18]. Because outbreak investigation often involves deploying teams of investigators into regions where network connectivity may be unavailable or intermittently available, the OMS is designed to run on a laptop computer with or without continuous network connectivity.

The availability of new sensor and data mining technologies, coupled with a recognition of the potential value of new and innovative data sources, have heightened interest in the use of technology to detect outbreaks and potential bioterrorist attacks earlier than conventional surveillance methods might allow. This interest has resulted in an increased use of biological sensors, information technologies, and public health informatics for the purposes of early detection and situational awareness – what has come to be known as "biosurveillance."

Biosurveillance is the automated monitoring of information sources of potential value in providing situational awareness and in detecting an emerging epidemic, whether naturally occurring, accidental, or the result of bioterrorism [19]. Some of the information sources that can be monitored by biosurveillance systems include clinical diagnostic data, consumer behaviors (e.g., purchases of non-prescription drugs), symptoms reported during ambulatory care, chief complaints reported to emergency-room staff, work or school absenteeism, lab orders, data collected by bio-sensors, as well as public health case information.

Biosurveillance systems leverage two major surveillance methods. First, well established public health surveillance methods used in epidemiologic investigations of infectious disease outbreaks and environmental conditions are used to provide baseline comparisons and to help confirm the accuracy and reliability of biosurveillance findings. Second, near real-time, automated analysis of cases and suspect cases, along with statistical analysis and data visualization of pre-diagnostic and diagnostic data, support the earliest possible detection of events that may signal a public health emergency, and can provide continuing situational awareness throughout a public health response [20]. Biosurveillance systems are being developed at the local, state, and national levels [19, 21, 22, 23, 24].

Figure 1 depicts a timeline from initial exposure to a biological agent to final diagnosis and treatment, and identifies some of the types of data available at various points [25]. Biosurveillance might use any number of the illustrative data types (and others) available from initial exposure to final diagnosis to detect patterns indicative of a potential outbreak or bioterrorist attack. Sensor technology can be used to detect biological events near the time of initial exposure to a biological agent. An example is the Department of Homeland Security's BioWatch, which uses sensors to detect trace amounts of biological materials in the air. These environmental data can assist public health analysts in determining the presence and geographic extent of a biological agent release, enabling federal, state, and local officials to more quickly detect and respond to biological events. BioWatch operates nationwide, focusing on major urban centers [26].

Syndromic surveillance is a type of biosurveillance used primarily for early detection. The principal data sources for syndromic surveillance are healthcare utilization patterns and pre-diagnostic clinical data – information indicative of a need for health care, beginning with the initial onset of symptoms, perhaps

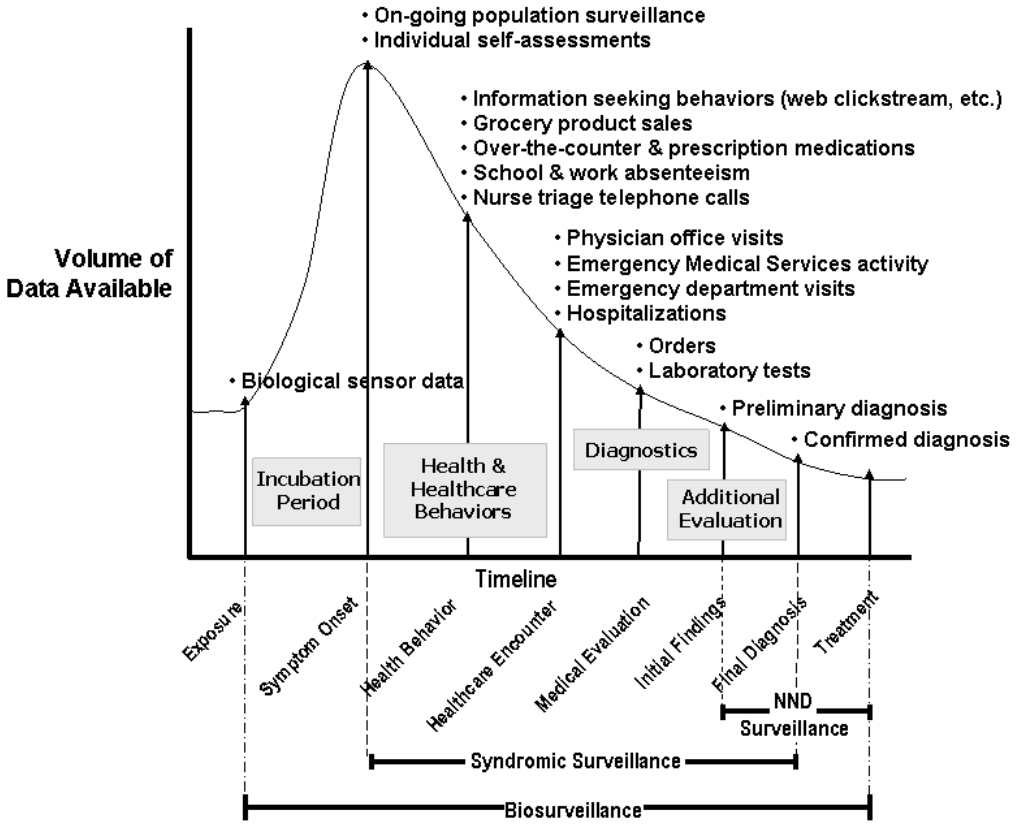

Figure 1. Public health surveillance uses data from many sources, enabling earlier detection and improved situational awareness – while also creating risks to personal privacy.

suggested by an increase in sales of consumer healthcare products (e.g., over-the-counter drugs), through confirmation of disease by lab test results. Such patterns are monitored in real time for the first signs of a covert biological attack or disease outbreak, which may appear as clusters of infected victims seeking health care [27]. A pattern of ill individuals exhibiting similar behavioral patterns, symptoms, signs, or preliminary laboratory findings could be an indicator of an emerging disease outbreak. Examples of syndromic surveillance systems in development and use today are the Department of Defense's prototype Electronic Surveillance System for the Early Notification of Community-based Epidemics (ESSENCE), designed for early detection of infectious disease outbreaks at military treatment facilities [28], and BioSense, being implemented in hospitals and health systems in major cities across the US [23].

The privacy sensitivity of information useful to public health surveillance increases as the information becomes increasingly specific and person-centric. For example, the air samplings collected by BioWatch contain no personal health information, while NND case reports sent to local and state health departments contain detailed personal and clinical information. A confirmed diagnosis or treatment plan that includes the patient's name is very personal and potentially damaging to the individual should it be disclosed. A significant challenge for public health is to achieve and maintain an optimal balance between assuring the health and safety of the U.S. population and protecting the privacy of individuals within that population.

4. Legal Framework for Privacy and Security Protection in Public Health

Arguably the most well-known and widely discussed federal law dealing with security and privacy for health information is the Health Insurance Portability and Accountability Act (HIPAA) of 1996,

which called for the adoption of a number of standards to enable and encourage the use and exchange of electronic health information [29]. Principal among these standards were the Privacy Rule [30] and the Security Rule [31]. The Privacy Rule specifies the actions that healthcare providers and health insurers, known as "covered entities," must take to safeguard Protected Health Information (PHI) and defines the rights of individuals with respect to their own PHI. The Security Rule specifies administrative, physical, and technical safeguards that covered entities must either implement or consider. Compliance with the Privacy Rule is enforced by the Department of Health and Human Services Office of Civil Rights (OCR).

The Privacy Rule explicitly enables covered entities to release PHI to public health authorities without individual authorization, but restricts such release to the "minimum necessary" for the purposes intended and requires an accounting of all such disclosures (See Figure 2). The CDC has published guidance on the interpretation of the Privacy Rule for public health [32], and the OCR has developed a Decision Tool to assist emergency responders in interpreting the Privacy Rule as it might apply to their planning and response activities [33].

However, HIPAA privacy requirements for public health are not as straight-forward as they may seem, and interpretations vary from state to state. One source of ambiguity is that some public health agencies and laboratories perform covered functions, such as diagnostic testing and providing patient care, making them "covered entities" subject to compliance with all HIPAA requirements. Also, although the Privacy Rule authorizes covered entities to release PHI to public health authorities, those authorities must comply with applicable federal and state privacy laws, which take precedence over regulation.

The U.S. Constitution establishes the legal foundation for privacy protection, and the Privacy Act of 1974 [34] defines privacy requirements for federal agencies. The Privacy Act prohibits the disclosure of individual records without the prior written consent of the individual, requires accountability of disclosures, and specifies exceptions. Among the exceptions is the protection of the health and safety of individuals and the public.

The E-Government Act of 2002 [35] defines the responsibilities of federal agencies in protecting the privacy of personal electronic information in government information systems. A key requirement is the conduct of a privacy impact assessment prior to developing or procuring any information system that collects, maintains, or disseminates identifiable information. Title III, the "Federal Information Security Management Act of 2002," specifies required security controls to protect confidentiality, integrity, and availability.

More specific to public health, Section 242m [36] (see Figure 3) of the Public Health and Welfare Act (Title 42) prohibits public health authorities from disclosing information identifying an individual or establishment, or using that information for any purpose other than the purposes intended, without the consent of that establishment or individual. Section 242m applies to research, evaluations, and demonstrations in health statistics, health services, and healthcare technology; the activities of the National Center for Health Statistics; and international public health cooperation. The Act separately imposes privacy and confidentiality protections for a number of specific types of information, including developmental disabilities, DNA, HIV/AIDS, and mental health.

Without individual authorization, a covered entity may disclose PHI to a public health authority* that is legally authorized to collect or receive the information for the purposes of preventing or controlling disease, injury, or disability including, but not limited to

- reporting of disease, injury, and vital events (e.g., birth or death); and
- conducting public health surveillance, investigations, and interventions.

PHI may also be disclosed without individual authorization to

- report child abuse or neglect to a public health or other government authority legally authorized to receive such reports;
- a person subject to jurisdiction of the Food and Drug Administration (FDA) concerning the quality, safety, or effectiveness of an FDA-related product or activity for which that person has responsibility;
- a person who may have been exposed to a communicable disease or may be at risk for contracting or spreading a disease or condition, when legally authorized to notify the person as necessary to conduct a public health intervention or investigation; and
- an individual's employer, under certain circumstances and conditions, as needed for the employer to meet the requirements of the Occupational Safety and Health Administration, Mine Safety and Health Administration, or a similar state law.

Source: Adapted from [45 CFR § 164.512(b)].

* Or to an entity working under a grant of authority from a public health authority, or when directed by a public health authority, to a foreign government agency that is acting in collaboration with a public health authority.

Figure 2. The HIPAA Privacy Rule allows covered entities to release PHI to public health authorities without individual authorization [32].

(d) Information; publication restrictions
No information, if an establishment or person supplying the information or described in it is identifiable, obtained in the course of activities undertaken or supported under section 242b [research, evaluations, and demonstrations in health statistics, health services, and health care technology], 242k [National Center for Health Statistics], or 242l [international cooperation] of this title may be used for any purpose other than the purpose for which it was supplied unless such establishment or person has consented (as determined under regulations of the Secretary) to its use for such other purpose; and ***in the case of information obtained in the course of health statistical or epidemiological activities*** under section 242b or 242k of this title, such information ***may not be published or released*** in other form if the particular establishment or person supplying the information or described in it is identifiable ***unless such establishment or person has consented*** (as determined under regulations of the Secretary) to its publication or release in other form.

Figure 3. Section 242m of the Public Health and Welfare Act limits the disclosure and use of information identifying individuals and establishments.

5. Security Challenges and Risk Dependencies

In seeking to achieve and effectively maintain an optimal balance between assuring the health and safety of the U.S. population and protecting the privacy of individuals within that population, public health must address a number of difficult ethical and political considerations at all levels. For example, how can a local health department that provides health care (i.e., a "covered entity" under HIPAA) effectively manage its dual roles with an individual who is both a patient and a case in an outbreak investigation? Anonymization methods can be used to protect personal privacy in aggregated data sets – but at what sacrifice of detection sensitivity and specificity? If an outbreak is detected within a de-identified data set, how is the expedience of intervention and containment impaired by having to trace back to re-identify affected individuals? How can an individual's DNA be effectively de-identified? If a massive outbreak occurs and additional people are recruited to help collect case information and trace contacts, how might the response be impeded by having to first create system accounts and issue individual X.509 digital certificates before allowing the recruits to use an outbreak management system for their case investigations? Alternatively, if several recruits are allowed to share a single account on an outbreak management system, how might the security risk measure up against the health risk of not having enough people to handle a response? The bioterrorist threat is only making public health decisions more difficult. For example, what security measures are strong enough to manage the risk that pathogens and toxins used in microbiology research might be misused as agents of bioterrorism?

Figure 4 depicts a risk model representing the complex interplay among concepts and relationships that must be considered in order to protect public health while respecting and preserving personal privacy. As shown in this model, both personal privacy and population health are subject to risks. Security countermeasures that provide confidentiality protection can reduce risk to personal privacy. Security countermeasures that protect data integrity and service availability can reduce risk to population health by helping assure that public health data are not corrupted and that critical systems and information are available when they are needed. However, when security measures reduce the sensitivity of a syndromic surveillance system or impede a response to an outbreak or bioterrorist attack, they can contribute to health risk.

On the other hand, disease surveillance systems and outbreak response systems can possess security vulnerabilities that increase risk to personal privacy. For example, a syndromic surveillance system that collects all data elements within an electronic health record, rather than a restricted, de-identified data set, increases risk to privacy. Security policy can serve as a countermeasure to reduce these risks. The least-privilege principle instantiated in the HIPAA "minimum necessary" requirement is such a policy. This policy can be applied not only to the release of PHI from a covered entity to a local health authority, but also to sharing information between local and state levels, and between state and federal levels. In other words, at each level, what is the minimum information public health officials need to know to effectively protect the health of their constituency?

Threats can exploit security vulnerabilities to increase risks to either personal privacy or population health. For example, a threat that spoofs a public health web site to capture PHI will increase privacy risk, while a threat that exploits a vulnerability in the access controls guarding entry into a research facility dedicated to the development of vaccines to protect against anthrax attack will increase the health risk to the surrounding population. Unlike identity theft, the economic drivers for PHI theft have not yet emerged, primarily because health information traditionally has

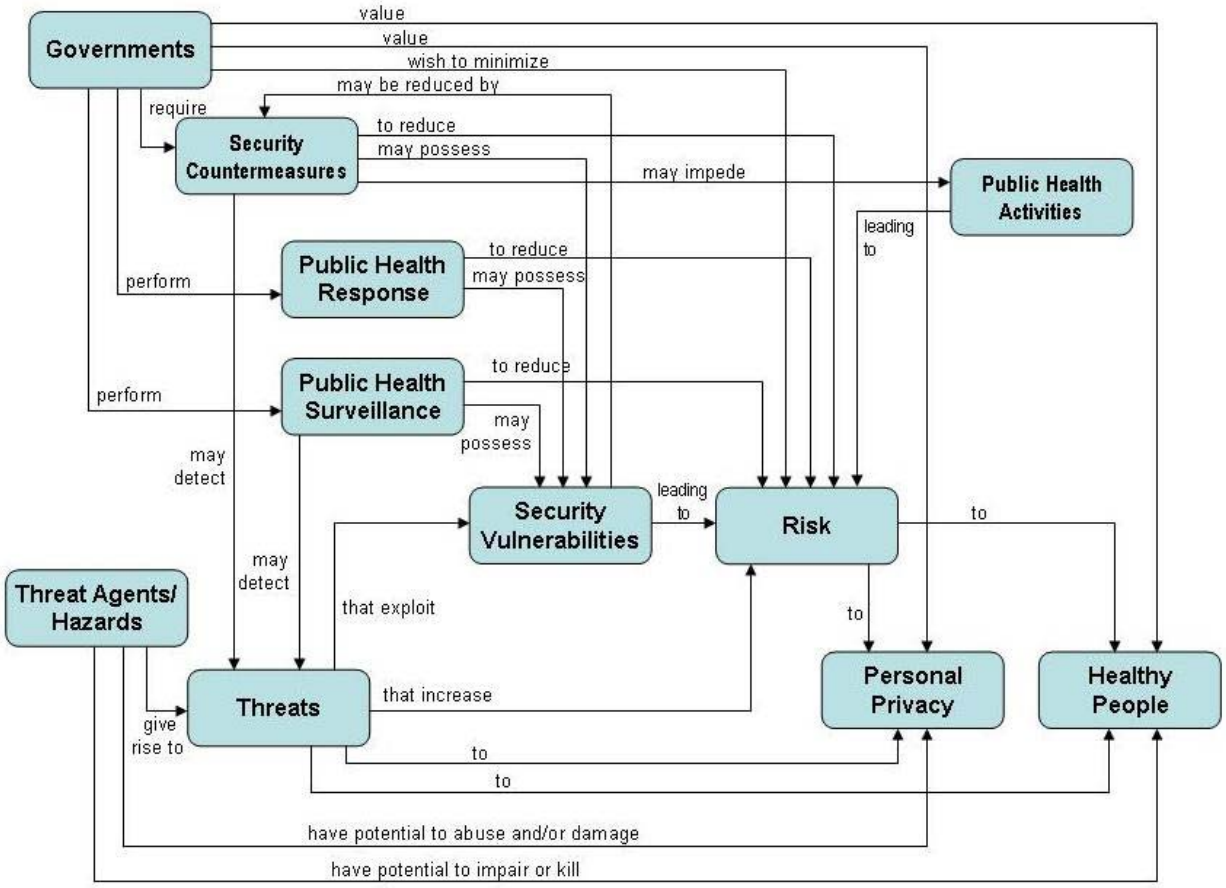

Figure 4. Public health must carefully consider the interdependencies among concepts and relationships in order to protect population health while preserving individual privacy.

been captured and retained on paper rather than in computers, and exchanged primarily by fax rather than over the Internet. As the retention and exchange of electronic health information become more commonplace, and as collaboration between healthcare providers and public health increases, security threats will become more pervasive and virulent.

Biological threat agents such as microbes, viruses, and toxins threaten the neurological, immunological, and endocrine systems of healthy people, posing health risks to large segments of the population, and creating national security risks because of potential social, economic, and political disruptions. To reduce these risks, governments enact regulations and laws, such as the Public Health Security and Bioterrorism Preparedness Response Act of 2002, which requires security measures for controlling biological agents and toxins, and for protecting the nation's food and drug supplies and drinking water [37].

6. Security and Privacy Requirements Analysis

The CDC has developed a PHIN certification process to establish the preparedness of public health partners to respond to a biological event that could have broad regional or national impact. The PHIN functional and technical requirements, based on industry data and systems standards, are intended to enable a secure, coordinated, nationwide network of public health IT systems capable of efficiently acquiring, managing, analyzing, and disseminating public health information. PHIN certification requirements are specified for the following nine IT functions [38]:

1) Automated exchange of data between public health partners
2) Use of electronic clinical data for event detection
3) Manual data entry for event detection and management

4) Specimen and lab result information management and exchange
5) Management of possible case, contacts and threat data
6) Analysis and visualization
7) Directories of public health and clinical personnel
8) Public health information dissemination and alerting
9) IT security and critical infrastructure protection

The objective articulated for the "IT security and critical infrastructure protection" function is to provide assurances that "access to sensitive or critical information and information systems is not lost, destroyed, misappropriated or corrupted by a internal or external malefactor or by systems failure or catastrophic event and that information is protected in ways that meet or exceed HIPAA standards" [39].

As discussed earlier, except under special circumstances, such as a local health department that provides healthcare services, public health organizations are not considered "covered entities" under HIPAA and therefore are not subject to compliance. However, as reflected in the PHIN functional objective for IT security and infrastructure protection, the public health community does regard the HIPAA Security and Privacy Rules as its benchmark for security and privacy protection of public health information and services. Thus the HIPAA Security and Privacy Rules can provide a useful framework for examining the protections needed and in place for public health. In this Section, we examine public health security safeguards and de-identification policy as they relate to the HIPAA Security and Privacy Rules.

6.1 Security Safeguards

In general, all of the administrative and physical safeguards specified in the HIPAA Security Rule apply to public health. Two notable exceptions are:

- Facility and equipment protections may be outside the control of public health authorities in some contexts, such as outbreak investigations and response.
- Public health authorities do not require Business Associate (BA) contracts with covered entities who release PHI to them. However, public health entities that are "covered entities" under HIPAA must establish BA contracts as required. Also, contract personnel who install software in covered entities for the purposes of extracting clinical data for public health surveillance may require BA contracts.

Table 1 identifies security requirements extracted from the nine PHIN certification areas and maps them to the technical safeguards required by the HIPAA Security Rule. A number of PHIN requirements address availability and continuity of operations – security objectives that are not well represented in the HIPAA technical safeguards, though to some extent included in the administrative safeguards (e.g., protection from malicious software, data back-up). Table 1 includes an additional column for these availability-protection measures.

The public health community has decided that despite the Internet's inherent security weaknesses, it is the best option for providing network connectivity among all public health partners. This decision is a good example of a risk decision represented by the model introduced earlier: while use of the Internet does increase risks to privacy, data integrity, and service continuity, its ubiquitous availability offers an immediate solution for enabling the connectivity necessary to support disease surveillance, preparedness, health alerting, outbreak response, and collaboration.

To protect system-to-system, bi-directional data exchanges over the Internet, secure ebXML messaging is required. The two systems involved in the exchange are mutually authenticated using X.509 digital certificates, and the payload is encrypted using the receiver's public key. Collaboration Protocol Agreements (CPAs) between messaging partners specify the transport protocol to use and the security constraints agreed upon by both parties. The CDC offers the PHIN Messaging System (PHINMS) as a reference implementation of secure ebXML messaging [40].

To assure that safety-critical clinical information and laboratory test results can be processed and acted upon as quickly as possible, every state and local health department must be able to electronically receive and immediately process clinical and laboratory information. Further, to assure that clinical and laboratory information with national implications is expeditiously reported to federal authorities, states must be able to immediately de-identify and forward the information to appropriate federal agencies, while retaining the ability to link the information back to an individual should that become necessary.

Lightweight Data Access Protocol (LDAP) directories are required to support authentication and authorization both within and across jurisdictions. Within the public health community, the distinction between authentication (i.e., proof of identity) and authorization (i.e., permission) is much less well understood than in the security community, and the terms are often misused. This misunderstanding is manifested in the PHIN requirement for "X.509 digital

certificates or comparable strong authentication for accessing sensitive or critical information…" This confusion can undermine the PHIN objective of facilitating information exchange. For example, X.509 certificates are issued by the CDC Certificate Authority for "authentication" to support PHIN ebXML messaging authentication. However, determining access policies and assigning "authorizations" to access information within a particular local, state, or federal information system are the responsibilities of the entity that controls that system.

At the federal level, the implementation of electronic authentication is driven by guidance issued by the Office of Management and Budget (OMB) and the National Institute of Standards and Technology (NIST) [41, 42]. As the National Health Information Network (NHIN) is implemented, these guidelines are likely to drive how electronic authentication is implemented throughout health care and public health.

Requirements for assurance measures, such as architectural assurances, periodic risk assessment, and certification and accreditation, are generally absent. Also, note that the PHIN requirement for "role-based, mandatory access control protocols" does not refer to the label-based controls commonly known to the security community as "mandatory access control," but rather role-based control that is "mandatorily" enforced.

6.2 Privacy Protection

As discussed in Section 4 above, covered entities under HIPAA may release PHI to public health authorities without the explicit consent of the individual to whom the information applies. However, the information released must be the "minimum necessary" for the intended purposes, and the covered entity must keep a record of all such releases. The Privacy Rule identifies 18 data elements that must be removed for PHI to be considered "de-identified." However, it does *not* require that information released to public health authorities be "de-identified."

In practice, identifiable health information is routinely released only to local and state public health authorities. PHIN certification requires that information provided to appropriate federal agencies be "linked but de-identified." That is, the states must remove data elements that identify the individual before sending the information to appropriate federal agencies, and they must be able to link the information back to the individual should they need to do so.

However, the PHIN certification requirements do not define "de-identified" or specify which data elements may be included, or which must be excluded, for any given data exchange. Selecting which potentially identifiable data elements to include in any data-collection scenario or data exchange is a risk-management decision. That is, deciding which data elements must be removed for any given data collection or exchange involves weighing the safety risk that excluding the data element poses for the population, against the risk that including the element poses to individuals whose personal information may be disclosed. Not only immediate, but long-term impact to the individual must be considered. For example, attributing a diagnosis of sinusitis, which can be cured within 14 days, is quite different from disclosing DNA data that are permanently attributable to an individual.

Table 2 lists the 18 data elements identified in the Privacy Rule and identifies whether and how these data elements are included in data exchanges for the following three contexts:

1) Electronic Laboratory Reporting (ELR) of lab results to the states
2) NEDSS reporting of case data from the states to the CDC
3) BioSense capture and reporting of clinical data from hospitals to the CDC

Several observations can be made. First, individual names are kept local; they are not sent to the CDC nor are they captured by BioSense. Second, more geographic information is captured than the Privacy Rule allows for de-identification. Whereas the Privacy Rule allows no geographic subdivisions smaller than a state and only the initial three digits of the ZIP code, full 5-digit ZIP codes are included in case reports sent to the CDC and in the clinical information captured by BioSense. Lab reporting includes the full 9-digit ZIP code. The inclusion of geographic detail reflects the importance of geographic information analysis to public-health surveillance. Epidemiologists rely heavily on geographic information systems (GIS) to help them analyze diseases within the context of their geographic areas, social and health services that might be available within a particular area, and the natural environment itself. Analyzing epidemiological data within a geographic context can reveal trends and interrelationships that would be more difficult to discover outside this context.

A third observation is the inclusion of case identifiers, specimen identifiers, order numbers, and result tracking numbers in reports that NEDSS systems send to the CDC. All of these data elements are important in linking back to individuals should that be necessary, and in maintaining the integrity of case reports.

7. Summary and Conclusions

Amidst threats of pandemic avian influenza and bioterrorist attack, public health surveillance and preparedness have never been more important. The capability to detect biological events, including emerging disease outbreaks and bioterrorist attacks, as near to the time of initial exposure as possible can significantly reduce the health risk to the public. The ability for laboratories to electronically report test results suggesting potential health risks, for public health departments to efficiently exchange case reports across jurisdictional lines, and for public health authorities to disseminate guidance and alerts are all necessary to effectively manage health risks.

Against this backdrop is recognition of the very private nature of health information and the significant harm that could result from its unauthorized disclosure. This recognition motivated the rigorous requirements contained in the HIPAA Privacy Rule. Yet, in an effort to effectively balance personal privacy against population safety, the Privacy Rule explicitly allows covered entities to release information to public health authorities without consent – balancing the risk by restricting the release to the "minimum necessary" and requiring accountability of all such releases.

The public health community recognizes the awesome responsibility this HIPAA provision carries with it to respect and protect the public's trust. Not only must public health agencies comply with federal and state laws and regulations regarding the protection of private and confidential information, but the public demands and expects that their personal privacy be respected and their confidential information be protected. As electronic health information flows from laboratories to hospitals, clinics, and health departments, and from local to state to national health agencies, public health is challenged to achieve and maintain an optimal balance between protecting the health and safety of the nation's people, and respecting the privacy of individuals. The decision to use the Internet as the backbone for the PHIN requires security features and architectural assurances to counter inherent vulnerabilities not only to protect sensitive information against unauthorized disclosure, but also to assure the integrity of data and the availability of critical services.

Public health is continually addressing this challenge at all jurisdictional levels and in all functional areas.

8. Acknowledgements

The author wishes to thank the reviewers for their time, insights, and helpful suggestions, and to acknowledge and thank Margaret Marshburn for her help in the data-element analysis given in Table 2.

Table 1. The PHIN Certification Requirements address all of the technical safeguards prescribed by the HIPAA Security Rule.

Requirements	Control	Audit	Integrity	cation		
Continuous connectivity to the Internet						✓
Internet-based, bidirectional, system-to-system data exchanges over mutually authenticated, encrypted ebXML connections	✓		✓	✓	✓	
Encryption of sensitive data prior to transmission over a secured HTTPS connection	✓		✓		✓	
Capability for any jurisdiction in the U.S. and its territories to receive a secured transmission from the CDC					✓	✓
Capability for state and local health departments to electronically receive and immediately process clinical and laboratory information, and to immediately send a linked but de-identified form of this information to appropriate federal agencies	✓					✓
Registry de-duplication and automatic data-linking to preserve the integrity of case-investigation data			✓			
Lightweight Data Access Protocol (LDAP) directories of public health personnel, including name, role, affiliation, and geographical location, to support authentication and access authorizations within and across jurisdictions	✓			✓		
Secure messaging and alerting (via multiple channels), including secure archival and authenticated electronic discussions				✓	✓	✓
Assurance that access to sensitive or critical information and systems is not lost, destroyed, misappropriated or corrupted by a internal or external malefactor or by systems failure or catastrophic event	✓		✓			✓
Information protection that meets or exceeds HIPAA standards	✓	✓	✓	✓	✓	
Assurance that processes cannot be initiated or controlled by	✓		✓	✓		✓

PHIN Certification Requirements	Technical Safeguards					Availability/ Continuity of Operations
	Access Control	Audit	Integrity	Authenti-cation	Trans-mission Security	
unauthorized individuals and that continuity of operations can be maintained subsequent to a catastrophic event						
Client and server X.509 digital certificates or comparable strong authentication for accessing sensitive or critical information over the Internet			✓	✓	✓	
Role-based, mandatory access control protocols	✓					
Realistic and effective policies for use and administration of information technology resources	✓					
Prompt application of security patches and configuration corrections			✓			✓
Desktop and server-based virus scanning			✓			✓
Intrusion detection	✓	✓				
Network vulnerability analysis, regular penetration testing, and active threat intelligence monitoring		✓			✓	
Security policy monitoring		✓				
Continuity of operations planning and procedure implementation, including man-made and natural catastrophic event management						✓
Routine offsite back-ups			✓			✓
Security policies	✓	✓	✓	✓	✓	✓
Authentication based on industry standard X.509 certificates, secure tokens, or comparable means				✓		
Access and control of data via selective integrated repository authorization	✓		✓			
Encryption	✓		✓		✓	
Access control through a firewall, including secure access to ebXML receiver and to restricted web sites	✓		✓		✓	✓

Table 2. Personal identifiers are removed as necessary and appropriate for data exchanges among public health partners. (YES = data element is included; NO = data element is excluded)

HIPAA De-Identification Data Elements		CDC)	
(A) Names;			
Patient Name	Yes	No	No
Next of Kin Name	Yes	No	No
Provider Names	Yes	No	No
(B) All geographic subdivisions smaller than a State, including street address, city, county, precinct, zip code, and their equivalent geocodes, except for the initial three digits of a zip code if, according to the current publicly available data from the Bureau of the Census: *(1) The geographic unit formed by combining all zip codes with the same three initial digits contains more than 20,000 people; and* *(2) The initial three digits of a zip code for all such geographic units containing 20,000 or fewer people is*			
changed to 000.			
Street Address	Yes	No	No
City	Yes	No	No
County	Yes	Yes	Yes
Precinct	No	No	No
Zip Code	ZIP plus 4	5 digit ZIP	5 digit ZIP
Census Tract	No	Yes	No
State	Yes	Yes	Yes
Equivalent Geocode	No	No	No
(C) All elements of dates (except year) for dates directly related to an individual, including birth date, admission date, discharge date, date of death; and all ages over 89 and all elements of dates (including year) indicative of such age, except that such ages and elements may be aggregated			
into a single category of age 90 or older;			
Date of Birth	YYYYMMDD	YYYYMMDD	YYYYMM only
Admit Date	N/A	Full date	Full date
Discharge Date	N/A	Full date	Full date
Deceased Date	Full date	Full date	Full date
Age	Yes	Yes (without aggregation of >89)	Yes (without aggregation of >89)
(D) Telephone numbers; *(E) Fax numbers;*			
(F) Electronic mail addresses;			
Patient/Next of Kin	Yes	No	No
Providers	Yes	No	No

HIPAA De-Identification Data Elements	ELR (Lab Reporting to State)	NEDSS (State to CDC)	BioSense (Hospital to CDC)
Organizations	Yes	Yes	No
(G) Social security numbers; *(H) Medical record numbers;* *(I) Health plan beneficiary numbers;* *(J) Account numbers;* *(K) Certificate/license numbers;* *(L) Vehicle identifiers and serial numbers, including license plate numbers;* *(M) Device identifiers and serial numbers;* *(N) Web Universal Resource Locators (URLs);* *(O) Internet Protocol (IP) address numbers;* *(P) Biometric identifiers, including finger and voice prints;* *(Q) Full face photographic images and any comparable images; and* *(R) Any other unique identifying number, characteristic, or code*			
Social security numbers	Yes	No	No
Medical record numbers	Yes	No	No (but used to generate and manage the BioSense Patient ID)
Health plan beneficiary numbers	No	No	No
Account numbers	Yes	No	No (but used to generate and manage the BioSense Visit ID)
Certificate/license numbers	No	No	No
Vehicle identifiers and serial numbers, including license plate numbers	No	No	No
Device identifiers and serial numbers	No	No	No
Web Universal Resource Locators (URLs)	No	No	No
Internet Protocol (IP) address numbers	No	No	No
Biometric identifiers, including finger and voice prints	No	No	No
Full face photographic images and any comparable images	No	No	No
Any other unique identifying number, characteristic, or code:			
State and Local Case IDs	No	Yes	No
Laboratory account numbers	Yes	No	No
Specimen identifiers	Yes	Yes	Yes
Order numbers	No	Yes	Yes
Result tracking numbers	Yes	Yes	Yes

Proceedings of the 22nd Annual Computer Security Applications Conference (ACSAC'06)
0-7695-2716-7/06 $20.00 © 2006

9. References

[1] Health Information Management and Systems Society. *HIMSS CPRI Toolkit: Managing Information Privacy and Security in Healthcare*. Available from http://www.himss.org/ASP/topics_cpriToolkit.asp?faid=78&tid=4. Last accessed Aug 15, 2006.

[2] Baker DB, Barnhart R, and Buss T, PCASSO: Applying and extending state-of-the-art security in the healthcare domain. *Proceedings of the Annual Computer Security Applications Conference*. San Diego CA. Dec 1997.

[3] Baker DB and Masys DR. PCASSO: Vanguard in patient empowerment. In *Consumer Informatics: Applications and Strategies in Cyber Health Care*. R. Nelson and MJ Ball, Eds. Springer-Verlag New York Inc. Jan 2004, pp. 63-74.

[4] Masys D, Baker D, Butros A, and Cowles KE. Giving patients access to their medical records via the Internet: the PCASSO experience. *Journal of the American Medical Informatics Association*. 9:2: 181-191. Mar/Apr 2002.

[5] Institute of Medicine, Committee for the Study of the Future of Public Health. *The Future of Public Health*. National Academy Press, Washington DC. 1988.

[6] National Safety Council. Environmental Health Center: Glossary. Available from http://www.nsc.org/ehc/glossary.htm. Last accessed Aug 15, 2006.

[7] Koo D, O'Carroll P, and LaVenture M. Public health 101 for informaticians. *Journal of the American Medical Informatics Association,* 8:6:585-597. Nov/Dec 2001.

[8] Wikipedia. Available from http://en.wikipedia.org/wiki/Public_health_informatics. Last accessed Aug 15, 2006. Attributed to Ross DA et al. *Public Health Informatics and Information Systems*. Springer-Verlag. New York. 2003.

[9] Essential Services Workgroup, Public Health Functions Steering Committee. The essential services of public health. 1994. Available from http://www.apha.org/ppp/science/10ES.htm#essnserv. Last accessed Aug 15, 2006.

[10] Centers for Disease Control and Prevention. National Institute for Occupational Safety and Health. Available from http://www.cdc.gov/niosh/topics/surveillance/. Last accessed Aug 22, 2006.

[11] Centers for Disease Control and Prevention. National notifiable diseases surveillance system. Available at http://www.cdc.gov/epo/dphsi/nndsshis.htm. Last accessed Aug 15, 2006.

[12] Centers for Disease Control and Prevention. Nationally notifiable infectious diseases. Available from http://www.cdc.gov/epo/dphsi/phs/infdis2006.htm. Last accessed Aug 16, 2006.

[13] Martin SM and Bean NH. Data management issues for emerging diseases and new tools for managing surveillance and laboratory data. *Emerging Infectious Diseases*. 1:4. Oct-Dec 1995.

[14] Pezzini G, Ed. *A Guide to the Implementation of the National Electronic Disease Surveillance System (NEDSS) in State Public Health Agencies.* Council of State and Territorial Epidemiologists. Apr 2001. Available from http://www.cdc.gov/nedss/Archive/Stakeholder2/Appendix_B_CSTE_Guidance_Document.pdf. Last accessed Aug 24, 2006.

[15] Institute of Medicine, Committee on Assuring the Health of the Public in the 21st Century. *The Future of the Public's Health in the 21st Century.* National Academy Press, Washington DC. 2002.

[16] Broome CV. Federal role in early detection preparedness systems. Syndromic Surveillance: Reports from a National Conference, 2004. *Morbidity and Mortality Weekly Report Special Supplement.* Aug 26, 2005.

[17] Centers for Disease Control and Prevention. *NEDSS and NEDSS PAMS Business Discovery Statement*. V1.2. Feb 12, 2002. Available from http://www.cdc.gov/nedss/BaseSystem/NEDSSBusinessDiscoveryStatement1_2.pdf#search=%22NCID%20NEDSS%22. Last accessed Aug 22, 2006.

[18] Centers for Disease Control and Prevention. PHIN: Outbreak management system (OMS). Available from http://www.cdc.gov/phin/software-solutions/oms/index.html. Last accessed Aug 22, 2006.

[19] Hoffman, MA, et al. Multijurisdictional approach to biosurveillance, Kansas City. *Emerging Infectious Diseases*. 9:10:1281-1286. October 2003. Available from http://www.cdc.gov/ncidod/EID/vol9no10/03-0060.htm. Last accessed Aug 22, 2006.

[20] Association of State and Territorial Health Officials. *Position Statement: Biosurveillance.* Mar 7, 2006. Available from http://www.astho.org/pubs/BiosurveillancePositionStatementFINAL030706.pdf#search=%22biosurveillance%20definition%22. Last accessed Aug 22, 2006.

[21] Centers for Disease Control and Prevention. Syndromic surveillance for bioterrorism following the attacks on the World Trade Center – New York City 2001. *MMWR Weekly Special Issue.* 51:13-15. Sept 11, 2002. Available from http://www.cdc.gov/mmwr/preview/mmwrhtml/mm51SPa5.html. Last accessed Aug 22, 2006.

[22] RODS Laboratory. Available at http://rods.health.pitt.edu/default.htm. Last accessed Aug 22, 2006.

[23] Centers for Disease Control and Prevention. PHIN: Biosense. Available from http://www.cdc.gov/phin/component-initiatives/biosense/index.html. Last accessed Aug 16, 2006.

[24] Department of Homeland Security. Fact Sheet: BioWatch Early Detection, Early Response. Available from http://www.milnet.com/wh/DoHS/BioWatchFactSheetFINAL.pdf. Last accessed Aug 16, 2006.

[25] Figure adapted from Mandl KD, et al. Implementing syndromic surveillance: A practical guide informed by the early experience. *Journal of the American Medical Informatics Association*. 11:2:143, Fig 1. Mar/Apr 2004.

[26] Department of Homeland Security. Fact Sheet: BioWatch Early Detection, Early Response. Available from http://www.milnet.com/wh/DoHS/BioWatchFactSheetFINAL.pdf. Last accessed Aug 16, 2006.

[27] Reis BY, Pagano M, and Mandl KD. Using temporal context to improve biosurveillance. *Proceedings of the National Academy of Science*. 100:4:1961-1965. Feb 18, 2003. Available from http://www.pnas.org/cgi/reprint/0335026100v1.pdf#search=%22syndromic%20biosurveillance%20difference%22. Last accessed Aug 22, 2006.

[28] Department of Defense, Global Emerging Infections System. ESSENCE: Electronic surveillance system for the early notification of community-based epidemics. Available from http://www.geis.fhp.osd.mil/geis/surveillanceactivities/essence/essence.asp. Last accessed Aug 16, 2006.

[29] 104th U.S. Congress. P.L. 104-191. *Health Insurance Portability and Accountability Act of 1996*. Aug 21, 1996. Available from http://aspe.hhs.gov/admnsimp/pl104191.htm. Last accessed Aug 16, 2006.

[30] Department of Health and Human Services. 45 CFR Parts 160 and 164. Standards for Privacy of Individually Identifiable Health Information; Final Rule. *Federal Register*. Aug 14, 2002. Available from http://www.hhs.gov/ocr/hipaa/privrulepd.pdf. Last accessed Aug 16, 2006.

[31] Department of Health and Human Services. 45 CFR Parts 160, 162, and 164. Health Insurance Reform: Security Standards; Final Rule. *Federal Register*. Feb 20, 2003. Available from http://www.cms.hhs.gov/SecurityStandard/Downloads/securityfinalrule.pdf. Last accessed Aug 16, 2006.

[32] Centers for Disease Control and Prevention. HIPAA privacy rule and public health. *MMWR Early Release*. Apr 1, 2003. Available from http://www.cdc.gov/mmwr/preview/mmwrhtml/m2e411a1.htm. Last accessed Aug 16, 2006.

[33] Department of Health and Human Services, Office of Civil Rights. HIPAA Privacy Rule: Disclosures for emergency preparedness – a decision tool. Available from http://www.hhs.gov/ocr/hipaa/decisiontool/. Last accessed Aug 16, 2006.

[34] 5 USC Sec. 552a. The Privacy Act of 1974.

[35] 107th US Congress. PL 107-347. E-Government Act of 2002. *Federal Register*. Dec 17, 2002.

[36] 42 USC Sec 242m. The Public Health and Welfare. Chapter 6a, Public Health Service. Subchapter II, General Powers and Duties. Part A, Research and Investigations. Section 242m, General provisions respecting effectiveness, efficiency, and quality of health services. Jan 19, 2004.

[37] 107th U.S. Congress. PL 107-188. *Public Health Security and bioterrorism Preparedness Response Act of 2002*. Jun 12, 2002. Available from http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=107_cong_public_laws&docid=f:publ188.107.pdf. Last accessed Sep 7, 2006.

[38] Centers for Disease Control and Prevention. Public Health Information Network functions and specifications. V1.2. Dec 18, 2002. Available from http://www.cdc.gov/phin/governance/PHIN_Functions_Specifications_121802.pdf. Last accessed Aug 23, 2006.

[39] Centers for Disease Control and Prevention. PHIN: IT security and infrastructure protection. Available at http://www.cdc.gov/phin/architecture/standards/IT_Security.html. Last accessed Aug 23, 2006.

[40] Rhodes B and Kailar R. On security and public health information network messaging system. *Proceedings of the 4th PKI R&D Workshop*. Internet2. April 2005. Available from http://middleware.internet2.edu/pki05/proceedings/kailar-phinms.pdf. Last accessed Sep 7, 2006.

[41] Office of Management and Budget. Memorandum 04-04. E-authentication guidance for federal agencies. Dec 16, 2003. Available from http://www.whitehouse.gov/OMB/memoranda/fy04/m04-04.pdf#search=%22e-authentication%20OMB%22. Last accessed Sept 15, 2006.

[42] Burr WE, Dodson DF, and Polk WT. *Electronic Authentication Guideline*. National Institute of Standards and Technology. NIST Special Publication 800-63. Version 1.0.2. April 2006. Available from http://csrc.nist.gov/publications/nistpubs/800-63/SP800-63V1_0_2.pdf#search=%22NIST%20800-63%22. Last accessed Sept 15, 2006.

Proceedings of the 22nd Annual Computer Security Applications Conference (ACSAC'06)
0-7695-2716-7/06 $20.00 © 2006