Second Interim Report

Availability and Comparative Value of Data Elements Required for an Effective Bioterrorism Detection System

November 28, 2001

Funded by a grant from the
Agency for Healthcare Research and Quality
"Using Information Technology to Improve Clinical Preparedness for Bioterrorism"
(Principal Investigator: Michael M. Wagner, M.D., Ph.D., Contract Number 290-00-0009).

Independent Report Reviewers

Reviewers chosen for their diverse perspectives and technical expertise were invited to review this report. The purpose of this independent review is to provide candid and critical comments that will assist the institution in making its published report as sound as possible and to ensure that the report meets institutional standards of objectivity, evidence, and responsiveness to the study charges.

The members of the Scientific Advisory Committee for this research listed in Appendix D were invited to review this report as well as additional reviewers suggested by the agency.

AUTHORS

Michael M. Wagner, M.D., Ph.D.
Assistant Professor of Medicine and Intelligent Systems
Director, Laboratory for Computational Epidemiology
Center for Biomedical Informatics
Department of Medicine
University of Pittsburgh
Scientific Co-Director, Biomedical Security Institute
University of Pittsburgh and Carnegie Mellon University

Ron Aryel, M.D., M.B.A.
Principal Consultant, Healthcare
Del Rey Technologies, LLC
Philadelphia, PA

Virginia Dato, M.D., M.P.H.
Assistant Professor of Health Services Administration
Senior Public Health Physician
Center for Public Health Practice
Graduate School of Public Health
University of Pittsburgh

Contributors

Abi Fapohunda Dr.P.H., M.P.H, M.S.
Consultant
FOB Group, LLC
Monroeville, PA

Jeremy Espino, MD
National Library of Medicine Fellow in Medical Informatics
RODS Laboratory
Center for Biomedical Informatics
University of Pittsburgh

Ed Krenzelok, Ph.D., PharmD, FAACT, ABAT
Director
Pittsburgh Poison Center
Professor of Pharmacy and Pediatrics
University of Pittsburgh

Ravi K. Sharma, Ph.D.
Division of Behavioral and Community Health Sciences
Department of Health Service Administration
University of Pittsburgh

Fu Chiang Tsui, PhD
Associate Director, RODS Laboratory
Center for Biomedical Informatics
University of Pittsburgh

Xiaoming Zeng, M.D.
Graduate Student Researcher
Center for Biomedical Informatics
Department of Medicine
University of Pittsburgh

Executive Summary

This report commissioned by the Agency for Healthcare Research and Quality addresses three related questions: What data are required for an effective bioterrorism detection system? What are their comparative values? And what are their availability?

For purposes of this report, we assumed that the set of pathogens of concern is a set formed by taking the union of the CDCs' threat list, the NATO threat list, the set of nationally notifiable diseases, and other lists. We assumed that the scope for the detection system would be territorial US, not international.

Recognizing that the large number of pathogens present different detection challenges, and, we nevertheless found useful a simple dichotomy between outbreaks that present initially with many cases (as in the case of some food poisonings or bioaerosol releases of pathogens), and outbreaks in which cases accumulate more slowly. Detection of the first phenomena can potentially be accomplished through noticing a large number of individuals with early symptoms (prodromes) and perhaps exploiting the size and rapidity of the outbreak (or other commonalities amongst the cases such as downwind location) to narrow the diagnostic possibilities. This observation suggests that effective detection must use data that are available early in the course of illness, such as the fact that a person is experiencing aches and chills. In such outbreaks, sensitivity, accuracy, and confirmation of diagnosis in any particular individual are not critical to detection of the outbreak or even to its first-order characterization. With small or slowly growing outbreaks, however, detection and accurate characterization of one of the early cases is crucial. This requirement implies that effective bioterrorism detection must also use data that provide high sensitivity--not missing cases--and there must also be data collected that are sufficient to achieve a precise diagnosis of a single case.

Chapter Two identifies required data using CDC case definitions; analysis of data actually collected by public health; analysis of data used in recognition and characterization of 57 recent outbreaks; review of the literature on health-psychology, especially the sub literature relevant to behaviors of ill individuals between the onset of symptoms and presentation (if ever) for medical care; and first principle analysis of various detection strategies. It begins to address the question of relative importance by laying out data on a timeline.

Chapters 3-16 discuss the availability of the data elements identified in Chapter Two. Each chapter describes one or more "industries" that collect data that have potential value in bioterrorism detection. There are examples of industries that are organizing to provide (for a fee) real-time feeds of their data for homeland defense/public health surveillance purposes. Since these industries have technologists who understand their data, and their systems, this trend represents a substantial resource that should be encouraged and managed to assist in the national effort to improve detection capability.

Chapter 17 is a detailed study of the availability of data discussed in this report for one large metropolitan area. Our understanding of data availability in this region is not theoretical: It is the result of a 2-year-long project, described in the chapter, developing an advanced public health surveillance capability for this region. Early warning systems for bioterrorism that are under development at present are city-centric; therefore, such an analysis is expected to be informative to decision makers, especially when coupled with a shallower but broader analysis of the issues at a national level provided by the other chapters.

The conclusions from this study are that near-term improvement in detection of outbreaks that presents with many cases (e.g., bioaerosol releases) should be pursued by developing real-time access to clinical and veterinary information systems, call center data, poison center data, and over-the-counter drug sales data. Research and development on ways to detect health-information seeking behavior, or to encourage ill individuals to self-report could provide results that could quickly translate into improved capabilities. Physiological monitoring and absenteeism data (direct or indirect) may provide improved detection in limited settings such as facility protection.

Longer-term improvement in the ability for early detection of outbreaks that presents with many cases will be facilitated by deployment of point-of-care clinical information systems, veterinary electronic medical records, and physiological monitoring gear. Additionally, research on methods to obtain epidemiological information more quickly can be expected to result in additional improvements in detection. Deployment of interoperable systems adherent to NEDSS standards will also improve detection of widely distributed outbreaks that cross jurisdictions.

Near-term improvement in the detection of small, or slowly growing outbreaks can be attained by developing better real-time access to pathology systems, claims data, death records, and autopsy results.

Longer-term improvement in the detection of small, or slowly growing outbreaks can be facilitated by research in natural language processing of imaging reports, and obituaries.

Table of Contents

PART I: Required Data

Chapter 1: Introduction

This report, commissioned by the Agency for Healthcare Research and Quality (AHRQ), addresses three related questions: What data are required for an effective bioterrorism detection system? Which data are more useful for bioterrorism detection? What are the availability of these data?

The importance of detection of bioterrorism does not need elaboration in the wake of 17 cases of cutaneous and inhalational anthrax during the months of September and October 2001 that led to considerable morbidity, four deaths, treatment of an estimated 32,000 individuals with antibiotics, and closing of postal and congressional facilities. These consequences are small in contrast to the 20-billion-dollar impact estimated by Kaufmann and colleagues of a large-scale bioaerosol release in a suburban area.[3] Such an event might accrue economic costs at rates as high as $200M per hour during the peak of the epidemic, suggesting that large investments in methods capable of improving detection and response by just an hour would be warranted.

What does require discussion is the subject of which detection strategies can address such a threat. The current national opinion ranges diversely over proposals that include the idea of enhanced human detection via increases in the size of the public health workforce; providing Web consoles to enable physicians to report suspicious cases more rapidly; computer analysis of data collected routinely for other purposes; more reliance on molecular diagnostics; and use of physiological monitors and biosensors. These strategies are not mutually exclusive, although the proponents of each sometimes imply as much. A variety of pathogenic organisms must be dealt with, and the particulars of any outbreak vary such that it is likely that no single approach will cover all possible threats. Therefore, until proven otherwise, it might be best to assume that nearly every proposed strategy has a role to play. Moreover, all of these approaches fundamentally rely on availability of data, thus the basic results of this report are not highly dependent on assumptions about which detection strategies are the most effective either overall, or for specific threats.

Nevertheless, in this report, we admittedly focus on computer-based detection strategies that rely on the analysis of data collected routinely for other purposes. Our rationale is that time has always been of the essence in public health surveillance due to the effects of disease on health, however, the time latencies inherent in the current public health system have not produced sufficient consequences on health to attract society's attention, at least not until recently. With the advent of bioterrorism, there is a new requirement for very early detection, and the new tight time frame for operations does not allow for new collections of data, and it does not allow for manual processing of data [4].

Even the relatively small outbreaks of anthrax recently experienced expose critical time latencies in the current system, even when public health is on a heightened alert status. Exposed individuals in Washington, D.C. who subsequently died were not contacted at the time that building environmental investigations were initiated for reasons that are not clear but include lack of ability to assess relevant information in a sufficiently rapid time frame.

Definitions and Scope

The goal of this report is to identify data needed for an effective bioterrorism detection system. The meaning of the phrase *effective bioterrorism detection* is suggested by the previous discussion, but must be clarified and operationalized for the purposes of the present study. An earlier study conducted by the same authors, entitled *The Nation's Current Capacity for the Early Detection of Public Health Threats including Bioterrorism*, had similar requirements for clarity about scope and definitions. In the present report, we utilize the same scope and definitions, and reproduce them in this section.

We use the term **detection** to refer to a set of public-health surveillance processes that lead to initial **recognition** of an outbreak and to its **characterization** (e.g., causative organism, source, route of transmission, host characteristics, relevant environmental factors). Thus, the problem of detection includes recognition of some anomaly worthy of further investigation as well as the investigation and subsequent characterization of the anomaly.[1]

The term *bioterrorism* refers to the criminal use of biological agents to achieve some objective other than simply financial gain or murder. Almost any biological agent that has virulence against people, crops, livestock or other living entities can be used in a by a terrorist; thus, in this report, we consider a large set of biological agents that have been identified as public health threats by government agencies, or that are monitored routinely by health departments. We discuss this list in Chapter Two.

To be considered *effective*, a bioterrorism detection system must provide sufficiently good detection for the purposes of response, mitigation, or attribution. Detection must be sufficiently sensitive, specific and timely, where sufficiency is determined by characteristics of each disease process, such as treatability, as well as characteristics of the response system such as capacity, access, and time latencies. Although detailed quantitative characterization of the requirements for each disease process are not available, and would improve the accuracy of an analysis, it is

[1] In bioterrorism investigations, *forensic attribution* is also a key process that must be supported. **Forensic attribution** is the use of technology to assign responsibility (but not necessarily to the a level required by a specific court as connoted by the term *forensic science*). An effective bioterrorism detection system probably must address all three processes, but to reduce complexity in this report, we will consider data needed for initial detection of outbreaks and their characterization.

still possible to draw conclusions based on current levels of understanding of diseases and response systems.

In our previous study, we recognized that our analysis was affected by whether we assumed a national or international scope. In the present report, as in the earlier one, we use the territorial United States.

Although it is generally acknowledged that the earliest possible detection will eventually be accomplished by biosensors that detect bacteria and viruses when they first appear in our environment, such technology is a ways off especially in a sufficiently cheap form that every person can carry a personal monitor, or every building and open space can be monitored continuously. In the meantime, early detection of a surreptitious release will depend on monitoring people and animals for the early effects of that release, and through detailed analysis of the epidemiological characteristics of sick individuals. For this reason, the scope of this report excludes biosensors.

In summary, our operational definition of *effective bioterrorism detection* involves:

- Defining *detection* as encompassing initial recognition and subsequent characterization of an outbreak.
- A set of threats that is large
- *Efficacy* defined as sufficient sensitivity and timeliness to at least partially improve the outcomes of an epidemics
- Capability to be effective in a surveillance scope as large as, but no larger than the territorial US

Additional definitions are given in the body of the report at the time of first use, and all are provided for the reader's easy reference in a glossary of abbreviations and terms.

Organization of the Report

As we discuss in Chapter 2, it is too early in the development of the field of biodefense to be able to say, definitively, which data will contribute most to the effective detection of each threat and to, moreover, quantify the contributions of different data to detection. The value of data depends on many factors including availability, accuracy, and the existence of analytic methods for their processing. Therefore, the objectives of this report are to identify types of data and the factors that determine their value. We attempt to prioritize types of data for further evaluation and to reach provisional conclusions that should be viewed as testable hypotheses, rather than definitive conclusions.

Since the value of data is determined by both their inherent value (if fully available) as well as pragmatic issues related to their availability, these two characteristics both should guide researchers in prioritizing their experimentation. Researchers can select data for further analysis based on conjectures about their inherent value, and or they can select data based on pragmatic issues such as how difficult would it be to collect the data in real-time for a geographic region as large as a metropolitan statistical area, a state, or a nation.

In this report, we provide information about which data are required, their inherent timeliness, and their availability. We identify broad classes of data that are likely to be of value. We take

this approach because our ability to be definitive about *lack of* inherent value for many types of data is limited; therefore, we err on the side of over-inclusiveness. We assume that there are many ways to detect an outbreak from data and that the optimal engineering solution to early detection will be highly constrained by pragmatic availability issues. When the filter of availability is applied to the large set of potential types of data, the set of data to investigate may be quite focused, small, and much more manageable as a research agenda than the unfocused set.

Thus, we have organized the present report into three major parts.

Part I: Required Data consists of this introductory chapter and Chapter Two, which addresses the questions of which data are required and which data are more valuable for bioterrorism detection, excluding considerations of data availability. Obviously if data are not available they are of no value, but by excluding availability from the analysis, we achieve a pure analysis of the information content of data. We believe this perspective is useful for longer-term design and planning for a detection approach that will produce earliest possible detection, once the data are available.

Thus in Chapter 2, our analytical approach is to identify a large set of data potentially useful based on multiple heuristics that include:

- Data currently collected by public health
- Data that have played a role in recent public health investigations
- Data identified by first principle analyses of the problem of early detection
- Data identified by the literature about health psychology and the behavior of sick individuals up to and including the time that they seek medical care.

Part II: Data Availability comprises 14 chapters (Chapters 3-16) that address the availability of different types of data. Each of the chapters describes one or more "industries" that collect data with potential value in public health surveillance. We think this organization will make sense to those working in the field of early detection. Included are examples of industries that are organizing themselves to provide (for a fee) real-time feeds of their data for homeland defense/public health surveillance purposes. Since these industries employ technologists who understand their data, and their systems, this trend represents a substantial resource that should be encouraged and managed to assist in the national effort to improve detection capability.

Part III: Case Study and Conclusions comprises two chapters. Chapter 17--a case study in the Pittsburgh region--focuses on data availability. This analysis differs from the analysis in Chapters 3-16 in its attempt to provide a more detailed inventory of the sources of data and availability factors in one, city-centered region. Emerging 'early warning systems' for bioterrorism are city-centric and therefore such an analysis is expected to be informative to system designers and researchers, especially when coupled with our shallower, but broader analysis of the issues at a national level provided by Chapters 3-16. The last chapter of the report, Chapter 18, is a summary chapter that provides overall recommendations about priorities for data integration for public health surveillance and the development of better early warning capability. By factoring in availability, we obtain an analysis useful for near-term planning and prioritization.

Chapter 2. Identifying Required Data Elements

There are many potential ways to identify data required for effective bioterrorism detection, each having its limitations. We can list, for example, data used by human experts to detect actual outbreaks. Some limitations of this approach are that it can only identify data that were available to the experts (i.e., it mixes in availability issues), and it also can only identify data that human experts can process cognitively (i.e., it confounds the analysis of inherent value of data with limitations of the detection scheme). As a second approach, we can list data used by computer algorithms capable of detecting outbreaks. A limitation of this approach is that the field of computer-based detection is young and that many types of data have not yet been tried.

For these reasons, we use multiple methods so as not to overlook 'required' data.

Potential Methodologies

Ideally, an analysis of the relative value of different data elements would be rooted in value-of-information measurements; that is, for a set of public-health threats of concern, we would examine the ability of either a computer or a human expert to detect the outbreak with and without the data. The value of the data can actually be quantified, if we have an economic model of the consequences of the two scenarios. Value-of-information is calculated by subtracting the economic impact of the epidemic under a public health surveillance and response scenario with and without the data in question. To obtain a quantitative estimate, researchers usually employ a model of the way in which the data are used for detection and for response decision-making.

Because such models are only beginning to be developed in the field of public health surveillance, value-of-information measurements are not yet feasible. The actual state-of-the-art is that for a few types of data and for the single epidemic disease *Influenza*, the detection performance achievable using automatic detection methods and available data has been

measured (see Table 2.1). The economic models also needed for value-of-information analyses have not yet been developed for public health threats with the exception of large-scale releases of Anthrax and smaller outbreaks of Smallpox.[3, 5]

We can also examine the inherent correlation between various types of data and real outbreaks. Figure 2.1 shows the footprint of Influenza outbreak on 10 types of secondary data. It is possible to measure the time lag between the start or the peaks of these different signals using a variety of mathematical functions. However, these types of measures have only been reported for a few types of outbreaks, such as Influenza and for a limited number of data such as the types of information showed in this figure.

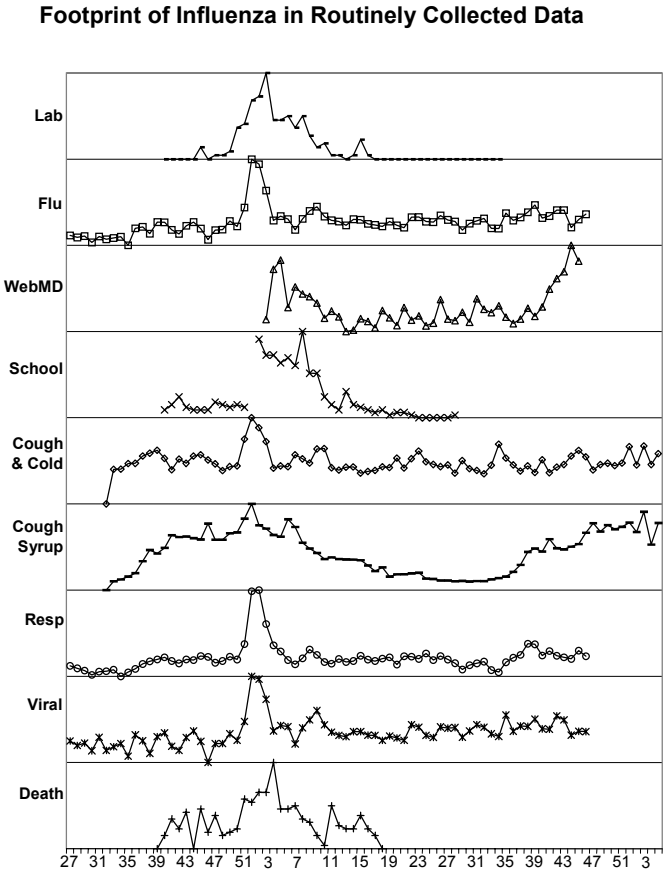


Figure 2.1: Weekly counts for several types of routinely collected data for different time periods around the December 1999 Influenza outbreak in Pittsburgh. Each data type is plotted on a normalized 0-1 scale. Legend: Lab, influenza cultures from the UPMC Health System; WebMD, counts of queries to a national web health site using words such as cold and flu*; Cough and cold and cough syrup, grocery chain point of purchase counts,* School,school nurse influenza reporting;, Resp. and Viral, Categories of emergency department ICD-9-coded chief complaints.

For these reasons, we use *earliness* of the data themselves as a heuristic measure of relatively value.

Table 2.1 Data Useful for Influenza Detection

Signal	Sens.	Spec./*PPV	Timeliness (weeks)
Emergency home visits	81	75	-1.6
Sick-leave reported to national health service	29	74	-1.4
Sick-leave reported to GP	76	65	-1.2
Sick-leave reported by companies	74	67	0.0
Sentinel GP visits	67	72	-1.2
Sentinel GP visits due to ILI	69	69	-0.4
Sentinel pediatrician visits	64	65	-1.7
Hospital fatality	47	82	-1.0
Influenza-related drug consumption	57	65	-1.2
Sentinel GP overall activity	57	63	-1.2
Sentinel pediatrician overall activitiy	47	67	1.3
Emergency room respiratory chief complaints	100	50	-1.0
Emergency room viral chief complaints	100	25	-1.0

First 11 rows are from Quenel et al. 1994 [1]
Last two rows taken from Tsui et al. 2001 [2]

Methods

In this report, we identified 'required' data by:

- First principle analysis of strategies for early detection
- Review of literature about the behavior of sick individuals up to and including the time that they seek medical care.
- Identifying data collected routinely by public health surveillance systems
- Identifying data potentially useful in testing CDC case definitions for notifiable diseases
- Identifying data used in recent public health investigations

We use these methods to identify an exhaustive set of data elements of potential value for detection of bioterrorism (Figure 2.2). We then lay them out on an *earliness* timeline. We identify existing information systems that collect these data. These information systems are then the focus of discussion in the Chapters in PART II of this report.

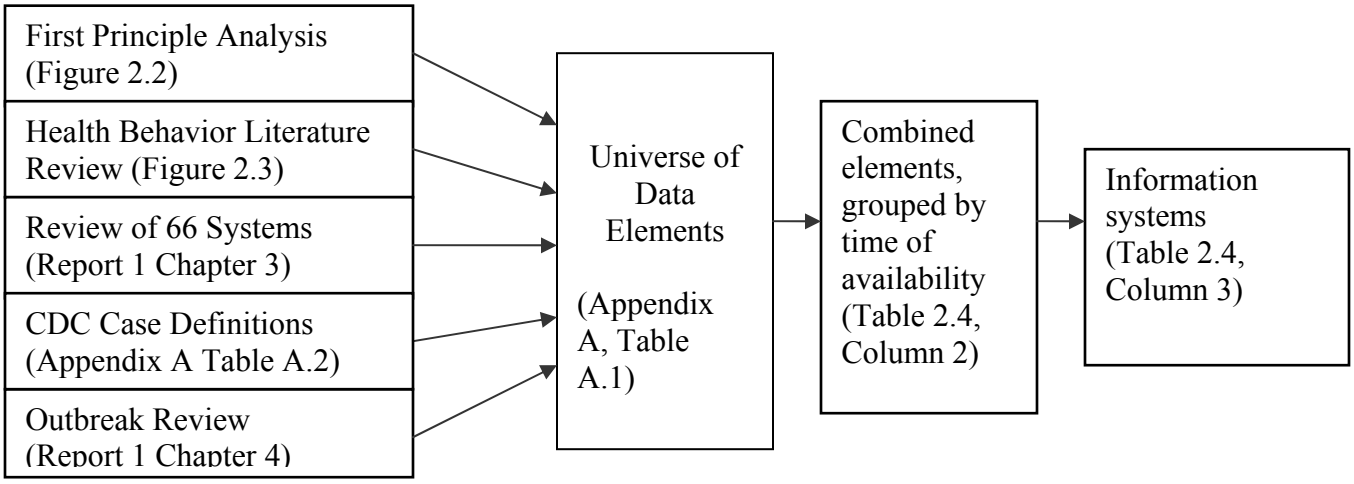


Figure 2.2 Framework of Analysis

Data Elements Identified by First-Principle Analyses

The importance of early detection suggests that data required by an effective bioterrorism detection system should ideally include *any and all* data that are available during the earliest stages of an outbreak. The set of such data is no doubt larger than the set of data typically collected and used by surveillance systems.

Figure 2.3 illustrates a first principles analysis of the data that can support early detection. In this example, the detection problem is one involving a biological warfare agent that has been adapted for aerosol release. That delivery mechanism is preferred in warfare because it allows an attacker to simultaneously infect as many individuals as possible, and because the atmosphere is much harder to protect than is the intended victims' water or food supply. Anthrax, plague, tularemia, glanders, and smallpox are examples of organisms that can be delivered by this route.

This figure is based on the economic analysis conducted by Kaufmann and colleagues of the hypothetical effects of a large-scale bioaerosol release of *B. anthracis* on a suburban population. In this model, a population of 50,000 individuals is exposed. The three curves show the time courses of onset of non-specific symptoms, presentations to emergency departments with signs symptoms sufficiently specific to *B. anthracis* that presumptive treatment would be initiated (e.g., radiographic findings in the lung on plain films or computerized tomography, coupled with a compatible clinical picture), and deaths. The estimates for onset of symptom presentation and percent mortality are based on Kaufmann's model, and are also influenced by recent American data[6]

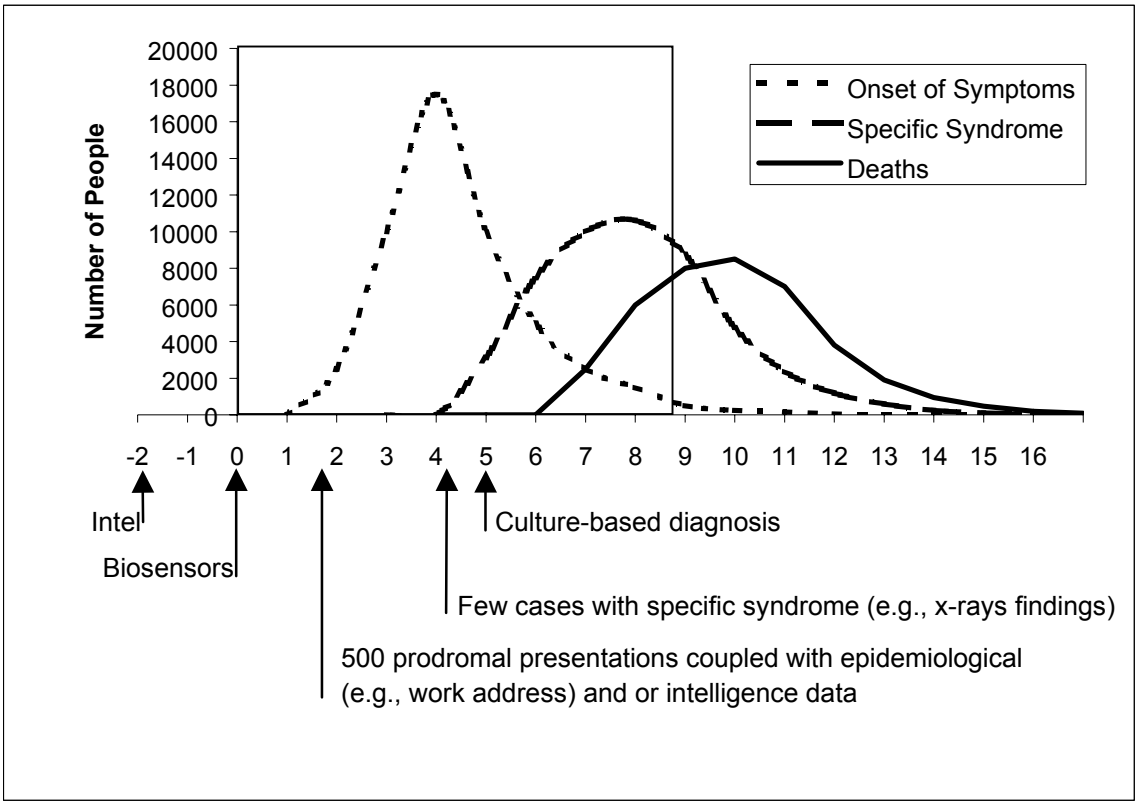


Figure 2.3. Epidemic curves and detection strategies for an outbreak of inhalational *B. anthracis* due to a large bioaerosol release. The release occurs on day zero. The boxed area represents a hypothetical window of opportunity for response measures to partially mitigate the outbreak. [Courtesy JAMIA].

The boxed area in the figure indicates the relatively brief window of opportunity for mitigation, post release. Anthrax is most treatable early, ideally pre-symptomatically; this is the rationale for prophylactic antibiotic treatment. After day eight, treatment is unlikely to significantly reduce further morbidity and mortality. Therefore, both detection and response must occur within this narrow time frame. Generating a response (treatment and prophylaxis) for hundreds of thousands of individuals requires time, so the window for detection is even narrower.

Figure 2.3 suggests several possible strategies to improve the promptness of detection over current capabilities. Early identification could possibly be achieved based on a small number of patients with findings strongly suggestive or pathognomonic of anthrax, if the set of such patients happened to present to a single clinicians, or if the existence of all such cases could be known in real-time through a detection and communications strategy.

An even earlier diagnosis could be achieved based on an abnormally high number of patients with non-specific symptoms consistent with anthrax. This detection method has lower specificity, as the unusual spike in presentations could derive from any of a number of agents. The detection of the spike, however, could prompt further investigation, such as probing for an epidemiologic association among the affected patients, and/or increased, more specific testing of those patients. Like the previous method, this strategy would also be greatly enhanced by a real-time detection and communications system, especially one that includes epidemiological data

such as work address and home address, along with symptom data. This is a strategy being pursued by many research groups at present, and those projects are described in Lober [7].

Figure 2.3 also suggests that physiological monitors and environmental biosensors that detect bacteria and viruses during the pre-symptomatic periods will eventually accomplish the earliest possible detection. However, such technology is currently expensive and uncommon. In the meantime, early detection of a surreptitious release will depend on monitoring people and animals for the early effects of that release, and through detailed analysis of the epidemiological characteristics of sick individuals. Moreover, such detection strategies are outside of the scope of this report.

Using the framework we identified the following categories of data of use for early detection:

- *Pre-outbreak data:* This category refers to data obtained during the period prior to the release of a biologic agent. Data obtained during this period may contribute to prevention of outbreaks, which is not the focus of this report, or may contribute to detection of an outbreak. Examples of data that might contribute to detection include intelligence that heightens suspicion or host factors such as vaccinations that determine susceptibility.
- *Attack, release/or exposure data* refers to data obtained at or very near the time of release. Data from this time might come from biosensor arrays, police reports of observed explosions, unauthorized airplane flights, or other activities. These types of data are beyond the scope of this report
- *Pre-symptomatic data (incubation period data)* refers to data obtained between the time of release of an agent until the recognition of first symptoms in people. Examples of pre-symptomatic data are serology or cultures from pre-symptomatic individuals that are obtained either serendipitously, through routine screening, or through enhanced screening because pre outbreak conditions are favorable.
- *Early symptom data* refers to data from the period between the onset of symptoms and when the illness becomes more fully developed and distinguishable from other illnesses. Examples of data include diarrheal symptoms or upper respiratory symptoms. Examples of indirect data of potential value for the detection of individuals experiencing early symptoms include sales of over-the-counter cold medicines, vital signs, physical findings, and absenteeism.
- *Specific syndrome data* are data that either singly or in combination strongly suggests a specific agent. Examples include selected symptoms, histories of exposures, vital signs, physical findings, laboratory results, radiology results, and preliminary results from microbiology laboratories (e.g., gram stains).
- *Definitive data* include data that are sufficient on their own to conclude that a patient has a disease. Examples of such data include microbiology culture or autopsy reports. Definitive agent data usually can be obtained during the specific syndrome period but can also be found during other periods through screening, routine testing, or testing of the environment.

Three additional classes of data are suggested by this analysis of potential detection strategies. They do not fit neatly on a timeline. These include:

- *Epidemiologic data,* which refer to the whereabouts of patients prior to onset of disease; food and water consumption; contacts with affected individuals; and work and home addresses.

- *Zoonotic data*, which refer to data from veterinary and public health sources. Examples include small animal deaths, positive mosquito pools for malaria, animal vaccination status, and sentinel chicken serology.
- *Environmental data* refers to data that are the results of environmental investigations. Examples include refrigeration temperature, ventilation plans, and water supply areas.

To identify specific data using this framework, we analyzed 10 types of outbreaks. These types were:

1. Large-scale bioaerosol
2. Premonitory release of bioaerosol
3. Continuous release of bioaerosol
4. Building contamination
5. Contagious person-to-person aerosol
6. Food borne
7. Waterborne
8. Sexually transmitted
9. Vector borne
10. Recreational Water

The source of these categories, and a description of the method by which they were created, can be found in our first report to AHRQ entitled *The Nation's Current Capacity for the Early Detection of Public Health Threats including Bioterrorism* (Chapter Two[8]) Briefly, we reviewed 87 disease threats that appeared on lists developed by the CDC, DTRA, National Notifiable Diseases, NATO, interviews with soviet experts, and USAMRIID. We grouped these diseases into ten categories based on similarities of the threats from an early detection perspective. We believe that these categories cover the problem space of detection.

Literature on Care-seeking Behavior and Health Psychology

We examined the literature on health-seeking behavior since it provides insights into the behaviors of individuals before they seek medical care, and these insights can be used to suggest novel types of data and patterns in data resulting from those behaviors. Such data may support detection of sick individuals earlier than is possible using clinical data alone.

Figure 2.4 summarizes the results of this study. It combines a health psychological model of patient care-seeking behavior after the onset of symptoms with a data driven early detection model. The backbone health psycho-behavioral model is ordered by time, so that we can see which data sources are inherently earlier.

Our analysis of the health psychology literature suggested many unconventional types of data (for public health) with potential value for early detection of epidemics.

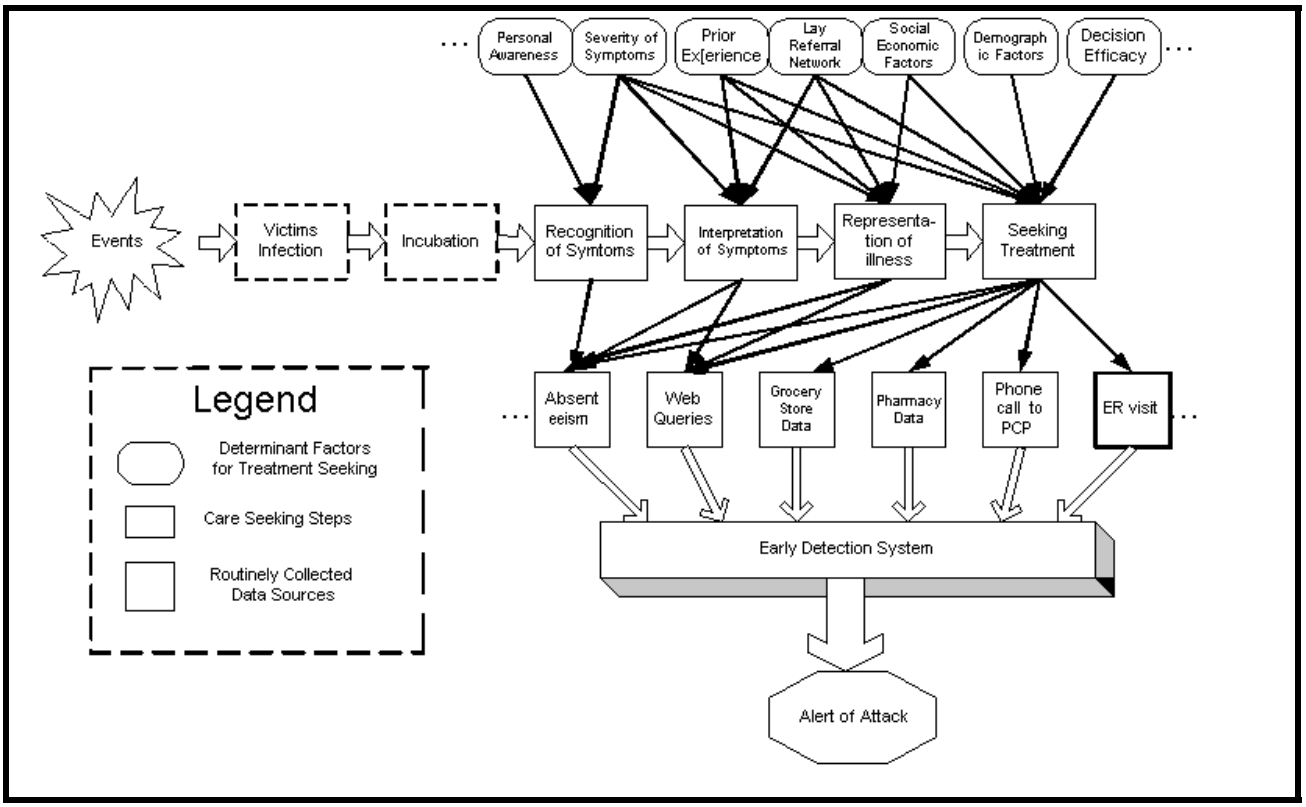


Figure 2.4. An existing psycho behavioral model coupled with an early detection system model. The psycho behavioral model is derived from observational and survey studies of the information and health-care seeking behavior of individuals with different illnesses [9].

Data Currently Collected by Public Health Surveillance Systems

In our first AHRQ-sponsored report on current capacity[10], we reported research that identified 66 public health surveillance systems that collectively comprise the National Public Health Surveillance System. For this report, we reviewed publicly available descriptive information to identify data collected by those systems.

Data Used During Outbreak Investigations

We identified 57 US outbreaks of diseases that appeared on our merged list of public health threats. We reviewed published information to determine the data that were used to recognize and characterize those outbreaks. Table 2.2 lists pathogens associated with those outbreaks.

Table 2.2 - Agents Responsible for Outbreaks by Detection Category

Category	Agent Causing Outbreak
Large-scale bioaerosol	No outbreaks
Premonitory release of agent	Carbon monoxide, *Legionella*, *Salmonella*, phage-group 1 staphylococci, *Burkholderia mallei*, *Rickettsia rickettsii*, methicillin-resistant *Staphylococcus aureus*
Building contamination	Carbon monoxide poisoning (2), *Legionella*
Continuous release of bioaerosol	*Blastomyces dermatitidis*, *Cocciodioides immitis*, hantavirus infection, *Legionella*
Contagious person-to-person aerosol	Influenza A, rubella (2), varicella, *Neisseria meningitides*, toxic shock syndrome, tuberculosis (3)
Foodborne	Anthrax, *E. coli* 0157 (2), *E. coli* O111:H8, group A rotavirus, histamine from tuna, *Listeria monocytogenes* (LM), new unknown agent, Norwalk-like virus (2), *Shigella* (2), *Salmonella* (2), *Vibrio parahaemolyticus*
Water supply	*E.coli 0157* and *Campylobacter* (same outbreak), *Cryptosporidium*
Vector/Host borne	Malaria, *Salmonella*, West Nile viral encephalitis, Lyme disease
Sexually transmitted	HIV (2), *N. gonorrhoeae,* hepatitis C
Recreational water	*Pseudomonas Dermatitis*, *Shigella* and *Cryptosporidium* (same outbreak)

Data Mentioned in CDC Case Definitions

We analyzed the 69 case definitions for the mandatory reportable diseases listed on the CDC Website http://aspe.os.dhhs.gov/datacncl/datadir/cdc4.htm. Examples of case definitions include: Acquired Immunodeficiency Syndrome (AIDS); Anthrax; Botulism; Botulism, foodborne; Botulism, infant; Botulism, other (includes wound); Brucellosis; Chancroid; Chlamydia trachomatis, genital infections; and Cholera.

We grouped these case definitions into categories as shown in Table 2.3 to demonstrate that they cover well the range of threats of concern for bioterrorism. As stated earlier, the categories are described in detail in our first AHRQ report. A disease may appear more than once in Table 2.3. We analyzed the case definitions to identify specific data elements within the definitions.

Table 2.3 Categories of Threats

Category	Threats
Non contagious airborne, single release	Weaponized anthrax, weaponized staphylococcus enterotoxin B, weaponized tularemia, weaponized botulism, weaponized *Coxiella burnetti*: Q Fever, weaponized *Pseudomonas mallei,* glanders, weaponized *Clostridia perfringens* toxin, weaponized *Brucellae sp*., weaponized ricin aerosol, T2 Mycotoxin aerosol, Histo-Coccidiomycosis
Small premonitory release or contamination	Weaponized anthrax, weaponized staphlycococcus enterotoxin B, weaponized tularemia, weaponized botulism, weaponized *Coxiella Burnetti,* Q Fever, weaponized *Pseudomonas Mallei*: glanders, weaponized clostridia perfringens toxin
Enclosure (building, ship) Contamination	CO poisoning,any bioaerosol, marine toxin: saxitoxin, ciguatoxin, tetrodotoxin, palytoxin
Non contagious airborne, continuous release	*Legionella pneumophila*,Histo-*Coccidiomycosis,Chlamydia Psittici*: Psitticosis, weaponized anthrax, weaponized staphylococcus enterotoxin B, weaponized tularemia, weaponized botulism, weaponized *Coxiella burnetti*: Q Fever, weaponized *Pseudomonas mallei,* glanders, weaponized *Clostridia perfringens* toxin, weaponized *Brucellae sp*., weaponized Ricin aerosol, T2 mycotoxin aerosol, histo-coccidiomycosis
Contagious person-to-person	Influenza,*Variola*, rubeola, rubella, *Mycobacterium tuberculosis*, mumps:paramyxovirus, smallpox, *Corynebacterium diptheria*, *Hemophilus influenza*, *Mycobacterium leprae*: Leprosy, *Neisseria meningititus*, contagious-person-to-person, "group A strep: rheumatic fever, toxic shock syndrome, necrotizing fascitis"
Foodborne	Salmonellosis, *Shigella* sp., *E coli* 0157, *Brucella* sp., staph enterortoxin *B.Vibrio,* Cholera, *B. Anthracis*, toxic alimentary aleukia: T2 mycotoxin, *Clostridium botulinum*: Botulism, hepatitis A, *C. Perfringins* E toxin, ricin toxin,"heavy metals: Pb, Hg, As,"Nipah virus, "marine toxin: saxitoxin, ciguatoxin, tetrodotoxin, palytoxin", *Trichinella*: trichinosis, Norwalk virus
Waterborne	Cryptosporoidosis, *Shigella* sp, *Camphylobacter*, giardiasis, staph enterotoxin B, *E.coli* 0157:, "botulism, bioterroristic",ricin toxin, *Entamoeba histolytica*
Vector/Host –borne	*Plasmodium sp*: Malaria,West Nile,"Flavaviridae: Yellow Fever, Dengue",*Yersinia Pestis*-bubonic,*Francisella tularensis*: tularemia, "Ebola, Marburg", hantaviruses, *Coxiella*: Q fever, *Pseudomonas Mallei*: glanders meloidiosis, "Arenaviridae: Lassa, Machupo, Junin", "Bunyaviridae: Rift Valley, CCHF, Hantaan", "Alphaviridae: VEE, EEE, WEE, Chikungunya", "Flaviviridea encephalitis: (Russian spring summer, eastern equine, St. Louis, West Nile, Venezuelan…", Nipah Virus, rhabdovirus: rabies, *Borelia burgdorfi*: Lyme, "*Rickettsia* sp.: Rocky Mountain spotted fever, typhus"
Sexually transmission/parenteral	HIV,*Neisseria gonorrhea*,*Hemophillus ducrei*: chancroid,*Treponema pallidum*: Sypillus, *Chlamydia trachomatis*, hepatitis B & C
Recreational Water	*Shigella* sp, *E. coli* 0157, cryptosporodiosis, leptospirosis, giardiasis, *Enterovirus*: poliomyelitits

Results

Universe of Data Elements

Table A.1 in Appendix A lists all data elements identified by our five methods. This list is comprehensive and contains references to sources for the data elements, where relevant. Table A.2 contains the data elements used in actual outbreak detection and in the CDC Case Definitions classified by type of outbreak (water borne, food borne etc.). This analysis is intended to illustrate a way of clustering the data from the perspective of specialized detection system development.

Data and Data Systems for Early Detection

Table 2.4 presents a distillation of the data provided in Appendix A. Table 2.4 gives selected examples of data, spreads them out across the outbreak timeline, and provides sub categorizations that are sensible for each time period. For example, pre-release phase data are subcategorized into biologic agent factors, host susceptibility factors, and environmental factors--the classic epidemiological triad of outbreak preconditions. For the illness phases, we organize the information based upon the sequence of behaviors expected from individuals in response to illness.

Table 2.4 also identifies data systems that routinely collect such data either for the purposes of public health, or more typically for other purposes. Part III of this report examines these systems in detail.

Table 2.4. Data and Data Systems for Early Detection

Category	Data 'Required' for Detection	Data Systems
Pre-outbreak	***Environmental conditions favorable for outbreaks*** Vegetation, climate, sea surface temperature, cloudiness, rainfall ***Information about susceptibility of population (host)*** Immunization information ***Information about outbreaks in other regions*** Outbreak report from WHO Emergence of new infections in other regions Keywords in the Internet and electronic reports related to or indicative of outbreaks in other areas ***Information about pathogens and their occurrence in the environment*** Antimicrobial resistance patterns Routine testing of food and water supplies Monitoring temporal and geographic patterns of specific viruses in order to tract conditions likely to be cause future outbreaks (e.g. looking for serotypes of Infuenza likely to spread during the next year) ***Information about animals*** Avian morbidity and mortality Captive or free ranging sentinel animals Monitoring diseases that animals can transmit to humans. including animals which will eventually be distributed to consumers in the form of pets or food	Satellite systems Meteorological data systems Immunization registries Public health systems Veterinary systems Food and water monitoring Information retrieval systems
Agent release or exposure	***Detection of release of pathogen*** Monitoring of emergency services radio frequencies for activities that might be indicative of a release 911 call data (e.g., report envelop with white power) Poison control center data (e.g., report of exposure to agent) ***Detection of pathogen in environment*** Routine testing of water, food and air for pathogens	Public health systems Food, water, engineering systems Call systems/ Emergency service Environmental sensors

Category	Data 'Required' for Detection	Data Systems
Pre symptomatic period	***Detection of pathogen in animals or crops*** Dog bite reports (or reports of the exposures to organisms in host animals) Horse or other terminal host surveillance (arboviruses) to observe for the presence of organisms in the environment which also infect humans Institutional veterinary practice diseases that animals and humans both get but which animals may be come ill faster ***Detection of pathogen in humans prior to symptoms*** Tuberculin skin testing Contact tracing and testing for person-to-person diseases ***Detection of physiologic changes in humans*** Biometric data (temperature, increase heart rate)	Public health systems Veterinary systems Physiologic monitoring Clinical systems
Early symptom period	***Behaviors following recognition of symptoms*** Increases in sales of non prescription products related to infectious diseases ***Sentinel population monitoring*** Nursing home respiratory illness rates, absenteeism ***Health-information seeking*** Web-related health sites (queries related to words like cough) Poison control system usage (reports of illness) Phone calls to physicians ***Seeking treatment*** Emergency medical services dispatching , Increased HMO usage Increased outpatient volumes, emergency room volume, appointments ***Initial clinical care and testing*** Chief complaints. Medical history. Physical exams, imaging studies, Clinical laboratory results. Diagnosis (lymphadenopathy) adult respiratory distress syndrome), Medications,Unexplained critical illness ***Increase in deaths of nonspecific cause*** Unexplained deaths, sudden deaths, newspaper reports (obituaries, articles of unusual deaths of animals and humans, clustering of victims)	Pharmaceutical sales/marketing systems Clinical data systems Absentee-recording Call center/Emergency service Public health systems

Category	Data 'Required' for Detection	Data Systems
Specific syndrome data	***Clinical care or new interpretation of symptoms*** Specific, but not definitive test results (e.g., CD4 counts) Specific combinations of test results (e.g., findings consistent with anthrax and gram positive rods in blood) Medical diagnosis meeting specific diagnostic criteria Emergency room or outpatient discharge diagnosis Hospital diagnosis such as necrotizing fasciitis, and streptococcal toxic shock syndrome (STSS). Reportable disease submissions Treatment driven data Medication usage (usage of medications reserved for unusual diseases) Physician data base searches (web-based medical queries) Deaths from known cause Cause of death on death certificate (specific diagnosis related to bioterrorism)	Clinical systems Public health systems Pharmaceutical Sales/marketing systems County coroner's office/medical examiner information
Definitive agent data	***Culture and sensitivity results*** Bacterial, viral and other cultures (sputum, stool, blood) Resistance patterns ***Other definitive tests*** Direct antigen, EIA, fluorescent antibodies Direct fluorescent antibody test, rapid influenza A antigen-detection tests, serologic titers ***Post-mortem examination*** Trans bronchial biopsy, microscopic examination, lung biopsy specimens, skin biopsy) Postmortem COHB levels, toxicologic and autopsy	Clinical laboratory systems Public health systems County coroner's office/medical examiner information

Category	Data 'Required' for Detection	Data Systems
Epidemiologic data	***Demographic information*** Age, sex, race, ethnicity and so on Descriptive epidemiologic information such as onset of illness, incubation period, duration of illness, symptoms of illness, number of ill, deaths, hospitalizations, submitting physician, outcome (died, alive), hospital transfer status, county, travel history.. ***Risk factors/ information about possible common exposures*** Water sources, food sources, attendance at large indoor and outdoor gatherings, geographic location of home , work and or entertainment/ leisure, food preferences, food histories, food handler interviews, travel, work histories, pet histories exposures to animals, vaccination status, illness in family members, recent illness, religion, employer. geographic dispersion of victims ***Analytic outbreak information to determine common exposure*** Odds ratios and relative risk of common exposures when compared to cohort or controls, dose effects, infection rates, incidence and prevalence, sensitivity, specificity, positive predictive values ***Contact tracing*** Sexual partners, household or other appropriate contacts, contact information – address, phone number ***Response-related information*** Treatment history ***Laboratory-related information*** Type and method of laboratory test, specimen collection date, submitting laboratory, source of specimen, specimen id number)	Public health systems Vital statistic Clinical systems Call center/Emergency service
Zoonotic data	***Reportable zoonotic diseases*** (See also syndromic, incubation and pre-outbreak sections)	Veterinary systems

Category	Data 'Required' for Detection	Data Systems
Environmental data	***Food distribution*** Source, marketing, processing, preparation, sale, packaging, refrigeration, storage, cooking, culture, poor cooking practices, food cultures, food handling methods, sanitary inspections, food delivery schedule, check of water temperature at time of seafood harvest, farm inspection, shell fish tags, transportation and so on ***Water distribution*** Wells, surface type, chlorination, contaminated, inadequate pool or hot tub filtration system ***Air distribution*** Heating, ventilation and air conditioning system records, lack of personal protective equipment, poor plant ventilation, air-conditioner system and exhaust fan	Food, Water, Engineering Systems

Discussion

The use of multiple methods helped to identify data elements that would not have been identified otherwise. This approach does not provide a precise ranking of data by relative utility. Nevertheless, the time of appearance of data is a useful characteristic that will in Chapter 18 contribute to an overall assessment of value.

The lists in this Chapter and Appendix A may be of value to designers of public health surveillance systems. Information-system design should include a specification of data to be collected, stored, and processed; thus, it is surprising that a data inventory does not exist for public health surveillance. The tables provided here are a starting point for developing such specifications for a comprehensive, dual purpose, national disease surveillance system with both bioterrorism and routine public health surveillance roles.

Summary

This chapter identified a large set of data elements that are either perceived to be of value in public health (because they are collected), are known to be useful because they have helped in the recognition or characterization of real outbreaks, or can be identified as potentially useful based on first principle information about how an epidemic will develop over time and impact data sources, and available information about how people behave when they are ill. The data types fall into several broad categories: data useful for inferring that a specific patient has a particular disease or prodrome, data that can help identify someone who is sick, data that suggest that a person or community is at increased risk, data that can be used to establish epidemiological ties between sick individuals and each other or external factors, and data that can be used to infer illness in animal populations. It is noteworthy that for all categories, data can provide either direct evidence (e.g., the patient has cough) or indirect evidence (a person purchased cough syrup therefore he may have cough).

In Part II, we discuss the availability of these data as secondary information--from the public health perspective--from a large number of information systems that collect and store such data, and issues related to their availability to public health.

Questions for Further Study

1. What are the economic consequences of various epidemics with and without intervention?
2. What are the timeliness and specificity requirements of detection for various epidemics?
3. What are the values-of-information of various data for various epidemics?
4. What is the relatively timeliness of various types for real outbreaks other than Influenza?
5. What are the early behaviors of sick individuals for all threats? How do those behaviors change under epidemic conditions (e.g., when people are aware of Anthrax laced letters, or when public-health authorities announce prescribed behaviors)? Method may include analysis of detailed case reports from epidemic investigations. What can be learned about behaviors of individuals during early symptomatic period from secondary data sources?

PART II: Data Availability

Chapter 3. Overview of Availability

Data availability is relative, not absolute: It is almost always possible to obtain needed data for public health by *asking* the patient or family member, or through manual review of clinical or other records. However, when there is a requirement for real-time detection of bioterrorism, availability implies availability within some fairly tight time constraints and it is a certainty that key data needed for epidemic investigations are difficult to assemble at present in a timely manner.

The main focus of this report, therefore, is on sources of **secondary data**, which are data collected for other purposes such as clinical care that can be re-purposed for public health surveillance. These data contain information (as well as noise), often provide indirect evidence of the desired quantity, but ideally more than compensate for these limitations by being available in real-time (or could with information system improvements, become available in real-time). We emphasize that we are not claiming that additional primary collection of data is not a feasible approach, or even a key element in an optimal detection system for bioterrorism; only that it cannot be the mainstay of early detection, especially those focused on initial recognition of an anomaly, because there is no time for new data collection.

Accordingly, Chapters 4-16 address the following questions: To what extent are existing electronic data collection systems already collecting needed data (or data which are approximations of needed data) or do existing systems need to be extended, or entirely new systems developed? What are the architectural characteristics of existing data systems that influence availability? Are the data accurate and complete? What is the level of deployment of each type of system in the nation? How many different vendors exist? Are there any standards? Are the systems reliable? Are there legal or administrative barriers to their use in public health? Are there technical barriers to their use in public health? Are there any leverage points that can be exploited to reduce the numbers of interfaces to data systems that may have to be developed?

The method used in these chapters involved telephone interviews and literature reviews that:

- Identified existing data systems for study
- Characterized the data collected, as well as the method and frequency of collection

- Determined the relationship, if any, of those data to the public health threats we seek to detect
- Assessed their utility according to relevant criteria, such as reliability of the databases, sensitivity, specificity, positive predictive value, timeliness, and cost.

The existing sources of data that we analyzed were not limited to special-purpose public health surveillance systems. We also analyzed clinical information systems, databases maintained by supermarkets, drugstores, or market research firms, and other entities as described below.

In each chapter, we discuss the following attributes of data systems that determine their utility for public health:

- Availability of data in data system
- Data and system integrity
    - Reliability of data systems
    - Data coverage (what percent of country)
    - Accuracy of data in data systems
- Legal or administrative barriers
    - Confidentiality
    - Jurisdictional
    - Cost of access
- Technical barriers
    - Time delays due to batch processing
    - Data coding differences

The systems and databases for analysis were selected to be relatively exhaustive and inclusive and were suggested by (1) systems suggested by the data elements identified in Chapter 2, (2) systems for which the research literature suggested some potential, and (3) our knowledge of data being used in systems under development (unpublished). We excluded some systems that were defined as outside of scope by AHRQ (e.g., credit card databases).

Table 3.1 Selected Electronic Data Systems

Clinical Information Systems

1. Registration
2. Scheduling
3. Billing
4. Ordering
5. Point of care systems including ED, hospital, clinic, long term care
6. Departmental Systems:
   Radiology
   Chemistry/Pathology Laboratory
   Commercial labs
   Emergency Department

Call Center/Emergency Service Data Systems

1. Enhanced 911 Electronic Records
2. EMS/Fire/Police/Military Dispatch computer systems
3. Poison Information Centers

Public Health and Related Systems

1. State/Local Public Health electronic systems
2. Municipal Water Supply/Water Co. Surveillance Reporting
3. Food Processing/Manufacturing Plant Surveillance, incl. Pasteurization/Irradiation Records

Pharmaceutical Sales/Marketing Systems

1. Post-marketing Surveillance Systems
2. Pharmacy Sales Tracking
3. Supermarket Sales tracking
4. Electronic prescription clearinghouses

Absenteeism Reporting Systems

1. Public/private School Attendance
2. Employer-run Attendance Reporting
3. Military Attendance/Compliance (perhaps out of scope)
4. National-security related Attendance
5. Other sign-in sign-out systems (example: nuclear power stations)

Satellite Systems/Weather

1. Satellite Imaging
2. Weather forecasting

Claims data/billing

Sentinel population/biometrics data

Vital statistics

Veterinary and other animal monitoring systems

Chapter 4. Clinical Data Systems

A modern healthcare enterprise operate scores of computer systems for scheduling, registration, billing, radiology, laboratory, pharmacy services, dictation, ordering of tests, recording of clinical observations, emergency room management, and intensive care unit operations. It is therefore surprising that, although public health practice is highly dependent on data from healthcare systems for surveillance, there are few if any direct links between clinical information systems and public health. Instead, data are typically transferred back and forth by fax, mail, or phone between the two domains of practice. The untapped potential of data collected by existing computer systems (and the untapped potential of point-of-care systems to support real-time, decision-supported interactions between healthcare providers and public health) is enormous.

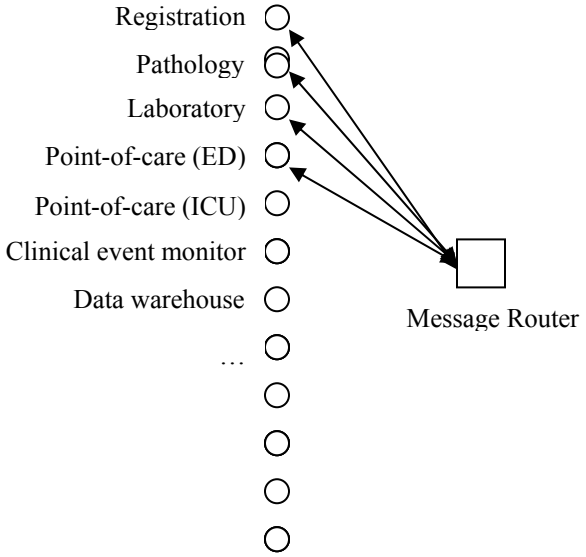


Figure 4.1 Representative architecture of a modern healthcare information environment.

HL7 Message Routers

Figure 4.1 illustrates typical health-system information architecture. HL7 is a dominant messaging standard in health-care computing that allows systems built by different vendors to more readily exchange data. Although we do not know the exact national market penetration of HL7 as a deployed standard, in our work in the Pittsburgh region described in Chapter 17, we found it to be ubiquitous. HL7 message routers are recognized as a key leverage point for bi-directional, real-time communications among existing clinical systems described in this report and public health information systems.

Registration, Scheduling, and Billing Systems

Most healthcare systems handle patient registration, scheduling, and billing automatically. There are a few vendors that supply the systems and most provide HL7 outbound interfaces. Data useful for public health surveillance include chief complaints and reasons for a health visit, patient age, gender, and home address. Many research groups in bioterrorism detection are investigating the use of such systems and data to provide an earlier signal than current methods [2, 4, 11, 12].

Electronic registration is almost ubiquitous in emergency departments and hospital based or associated practices, as well as large HMOs. Smaller practices may not utilize electronic systems to as great an extent, however.

Timeliness

The inherent timeliness of registration data for public health surveillance is certainly superior to the mainstay types of data used in public health surveillance, but they are not ideal. Sick individuals must reach a stage that they are willing to seek health care before a registration or scheduling record will appear. Artificial time latencies are not present since these data are recorded directly when appointments are made or patients register for the services. Billing data are generated during and after the time of service, so are inherently later.

Reliability

The reliability of these data systems is high. The accuracy of diagnostic coding is known to be poor, but for certain public health purposes the accuracy of detecting sick from non sick people and to achieve a symptom class of categorization may be adequate [2, 11].

Accessibility

Billing, scheduling, and registration data are collected at the health system level; thus to achieve regional surveillance, multiple health system's data must be integrated. There is heterogeneity in coding practices and schemes used that represent moderate obstacles to data integration.

Table 4.1. Clinical systems, data, and market penetration (estimated)

Clinical System	Data	Hospital/ Health system	ED	Office/ Home health	LTCF	Potential Uses
Registration	Chief complaints, addresses, age, gender	High	High	?	-	Prodrome-based detection strategies
Scheduling	Appts. and reasons	High	High	?	-	"
Billing	CPT-4 codes	High	High	High		Indirect evidence
Laboratory	Cultures, tests	High	High	High	High	Culture- and test-based detection strategies
Radiology	Chest radiographs	High	High	High		Test-based strategies
Pathology	Biopsies, autopsies	Mod	Mod.	Mod		Diagnosis-based strategies
Dictation	Symptoms and signs	High	High			Symptom and sign data
Pharmacy						Indirect evidence
Orders	Tests, drugs ordered	Low	Low	Low	Low	Indirect evidence
Data warehouse	All clinical data	Low	-	-	-	All of the above
Event monitor	-	Low	-	-	-	-
Point-of-care systems	Coded symptoms, vital signs, signs, diagnoses, orders, epidemiological data	Low	Low	Low		Data needed to satisfy case definitions; potential for decision support for physicians
Patient Web portals and call centers	Symptoms, referrals, appts.	Low	-	-		Collect early symptom information, potential for decision support for patients and doctors

Legend: *ED,* emergency department; *LTCF*, long term care facility; -, not applicable; *?*, unknown

Clinical Laboratory Systems

The data supplied by laboratories is important to surveillance of virtually every public health threat. Happily, the vast majority of clinical laboratories in the US are highly automated, utilizing computers to run tests, store results, and communicate results. Clinical laboratories perform tests on a variety of human tissues and fluids: blood, urine, stool, cerebral-spinal fluid (CSF), saliva, mucus, semen and fluid aspirated from joints. Tests performed include examination of the fluids and their components, measurement of electrolytes, levels of chemicals, and pH, and the incubation of fluid to provoke multiplication of microbes present within. Laboratory test results are often important is establishing diagnoses, and are generally available on-line and very soon after they are run, making the laboratory's time lag among the shortest of all data systems. Results are available not only directly from the laboratory systems, but also through, where deployed, the laboratory reporting interfaces of point-of-care systems. Thus, because of ubiquitous availability, public health authorities would be justified in closely examining methods of integrating laboratory data into an early-warning system for threats to public health, as they have been doing [13, 14].

Timeliness of Laboratory Data

Laboratory services in the US are delivered by either "in-house" laboratories at health-care facilities, or by commercial laboratories under contract to providers. The vast majority of laboratory work in the US is highly automated, and its reliability is enhanced by a combination of regular, manufacturer-specified calibration, automatic data recording directly from

instruments, and straightforward interfaces designed for laboratory technicians. Most data are entered into databases in real-time as tests are performed, though some tests, such as cultures, or tests run only periodically are entered as batches on a schedule. For example, culture dishes placed in an incubator will typically be read at one specified time in the morning or afternoon; all cultures will be read at once, and not again, for another 24 hours. This practice still provides very useful, relevant information with short latencies.

Data availability may be delayed due to the need to ship a test sample by courier, or airfreight, to a laboratory. This could add from a few hours to a day or more to the latency between collection of sample from a patient, and the posting of test results.

The efficiency of laboratory-result reporting enhances its value in the early detection of public health threats. However, an underlying assumption is that the sample is obtained at the same location as the laboratory, and is submitted immediately for processing. Although this assumption is valid in the majority of cases, in some cases, however, especially those involving highly specialized, infrequently requested or very expensive tests, the healthcare provider may submit samples knowing that tests will be "batched" and run only once per week, or shipped to a laboratory some distance away, adding a latency of hours, days, or even weeks. Whether these delays would significantly impact an early-warning system performance obviously depends on whether a given public-health threat's detection depends on a test performed only infrequently, or by special request at an off-site location. For example, carbon monoxide poisoning or ethanol poisoning is readily diagnosed with the aid of lab tests whose results are available almost immediately. On the other hand, certain viral cultures are handled as a "send-out." If that posed a problem regarding an imminent threat, public health authorities might have to consider subsidizing the cost of running the tests more frequently, and/or provide financial resources to allow on-site testing.

Reliability of Data

Laboratory data are, in general, trustworthy due to a number of factors. Most tests performed by clinical laboratories are almost entirely automated; the instrument performing a measurement sends this result, immediately upon processing the sample, directly into a computer database with no human intervention. Equipment capable of doing this can perform tests such as blood counts, serum chemistries, drug levels, toxicology, and those of physical and chemical characteristics such as pH and specific gravity. Laboratory technicians perform other tests manually then, enter results into a database by means of a user interface. Both equipment calibration and data entry are integral to a typical laboratory technician's daily duties. The consequences of incorrect reporting are serious, since decisions about patient care often depend on laboratory results. Hence, there is a strong incentive to ensure that test results are reported accurately. This is not to say that errors do not occur. They do not occur often enough to negate the value of these data to public health authorities.

Accessibility of Data

Clinical laboratory systems are most often commercial systems supplied by vendors, such as Sunquest, which has the largest market share, or Clinsoft, which licenses its software to small- or medium-sized laboratories. Some large commercial laboratory companies, which receive specimens on contract from health-care providers, such as Quest Diagnostics, may develop software platforms in house. About 80% of all clinical laboratories utilize database software packages released within the last two to three years, and these are powerful and flexible

relational databases. In some cases the software is an "add-on" package; for example, Sunquest's Flexilab ® system resides on a hierarchical database but is mapped to an optional relational database also marketed by the company. The amount of real-time data that can be stored depends on the storage capacity of a given laboratory's computer system; many clients can hold up to a few years' worth of data on-line and archive the rest.

Data from laboratory systems can be retrieved directly from the laboratory system's database, or sent to electronic medical record systems and retrieved there, as part of a patient's complete medical record.

Although laboratory vendors do not favor HL-7, many software vendors have developed HL-7 messaging applications to allow laboratory results to be transmitted to other systems, including point-of-care systems. This facilitates the integration of data from different sources into a patient's record.

Commercial laboratories can offer public health authorities access to their systems. Commercial software vendors point out that it is the client laboratory's prerogative to grant access to such data; interviews with laboratory officials contacted would have no inherent problem doing so, so long as patient privacy safeguards (and compliance with federal law) were in place. Another absolute requirement, however, is that public health would have to pay for access. One commercial laboratory complained, during an interview, that it had actually been approached by officials from CDC and asked to modify its systems without charge to the government. It refused to do so, stating that such modifications, while feasible, would cost millions of dollars to accomplish.

Radiology Systems

Radiology departments were also relatively early conversions in healthcare to computer processing. Most healthcare systems today handle radiology scheduling and resulting automatically. There are a few vendors that supply the systems and most provide HL7 outbound interfaces. Data useful for public-health surveillance include results of imaging, especially of chest and lungs. Many research groups in medical informatics have investigated methods for processing radiology imaging reports to recover information about patient characteristics such as presence or absence of pneumonia on chest radiograph [15-19]. Technical advances have made the digital storage and presentation of images at the point-of-care feasible although the market penetration of this functionality is low at present. Such functionality would be of limited use for public health, mainly reserved for detailed investigation of new disease outbreaks that are eluding characterization and requiring expert analysis of source materials.

Radiology information systems are almost ubiquitous in emergency departments and hospital-based practices. Reliability is about the same as laboratory systems. Most reports are dictated so that the results can be time lagged by a day or more by the transcription process.

Pathology

Pathology information systems are more recent additions to healthcare computing because of the image-intensive nature of pathology practice (gross and microscopic examinations). Thus, penetration into the market is less than for laboratory or radiology systems. Data useful for public health surveillance include orders for tests and results. Data accuracy would be expected to be very high. Timeliness and availability are not as good as laboratory systems due to the

nature of pathology practice and that these systems in many places are free-standing (used to generate printed reports) and integration has not been as critical of a design factor as in laboratory or radiology resulting.

Dictation

Dictation systems are a mainstay of clinical data recording in the hospital, ED, and outpatient settings. Dictation services can range from single part-time transcriptions, working with word-processing software, to pools of transcriptions producing reports in dedicated systems produced by companies such as Lanier. Although dictations are rich in clinical detail including the patient's presenting complaint, the history of the illness, exposure information, vaccinations, vitals signs and physical findings, and diagnostic impressions, the data are in English and are difficult to process computationally. There is also a time delay due to transcription that further limits data utility. Nevertheless, the value of the information is sufficiently high that many researchers in medical informatics have developed approaches to processing these data. The accessibility of these data is lower than laboratory systems because dictations are less frequently routed through a message router.

Pharmacy

Pharmacy information systems receive and process orders for medications. These orders may include antibiotics, antidotes for toxins that provide indirect evidence of the patient's symptoms or diagnoses. In the vast majority of implementations, the orders are received on paper and transcribed by the pharmacist into the pharmacy information system; thus, there is a delay from the time that the physician expresses his or her understanding of the clinical problem in the form of orders to the time that information is available electronically. The reliability of the information is extremely good, however, since pharmacists use expert knowledge and contextual information (available in the orders themselves, the pharmacy information system, and sometimes from review of the patient chart or contact with the physician) to validate the order prior to dispensing a medication. Pharmacy data is not highly accessible to other computer systems in most health systems, however.

Orders

Many health-care systems have computer applications that allow ward clerks to communicate physician orders to the laboratory and other departments in the health care system. Orders, similar to pharmacy orders, contain indirect information about a patient's condition. For example, if a cerebrospinal fluid examination or blood cultures are ordered, this action suggests that the patient may have an infectious condition. If respiratory isolation is ordered, this action may indicate that the patient should be seen by hospital infection control. The only time delay represented by orders is the time lag from when the clinician writes the order to the time when a ward clerk transcribes the order, which can be significant but is usually a delay measures in hours at most. The reliability and accuracy of this information is high. In a very small number of health care institutions, physicians enter orders directly into computers, eliminating the time delay and creating an opportunity for direct interaction and decision support of the clinician [20-23]. The potential of this type of data and activity is discussed under point-of-care systems below.

Data Warehouses

Healthcare systems often operate data warehouses to integrate data from multiple information systems and multiple hospitals. Data warehouses have become increasingly prevalent as a result of the mergers and acquisitions that occurred in healthcare in the 1990's. Although data

warehouses do not represent a primary source of data like systems heretofore discussed, they represent a point of integration of data that can be considered as a leverage point for public health surveillance. Data warehouses are not always real-time, although if they are being used by a healthcare system to achieve a useful consolidated view of data for clinical purposes (referred to as "results review"), they will have the real-time characteristic. The availability of real-time data warehouses in health care is low to moderate at present.

Clinical Event Monitors

A clinical event monitor is a type of expert system specialized for real-time operation in a clinical setting [24-26]. Its purpose is improvement in patient outcomes through automatic detection of individuals matching pre-specified criteria (e.g., patients with suspected tuberculosis who are not in respiratory isolation). The potential relevance of clinical event monitors to PHS is suggested by their feed-forward data-driven architectures and their core capabilities, which include (1) access to relevant data collected routinely by other information systems, and (2) facilities for operators to define logical patterns that can be compared against incoming patient data. Like data warehouses, clinical event monitors are not primary data collection systems but rather represent data analysis and communication functionality in a healthcare system. Their relevance to public health surveillance is--in addition to representing a point of data integration--they represent capability for data-driven real-time analysis. In particular, they are well suited for implementing case-detection logic and for supporting real-time interactions with clinicians to influence clinicians' actions. The potential role of clinical event monitors in public health surveillance is discussed in [27].

Point-of-Care Systems

Point-of-care systems are sometimes referred to as electronic medical record systems, but we find that term ambiguous and confusing in practice; thus, in this report, we use the term *point-of-care systems* to refer to systems that are used by clinicians to record directly details of patient encounters, to review information, and to order tests and other services. Vendors sell point-of-care systems specialized for diverse settings including ED, Office, Hospital, ICU, long term care facilities, and home health care. Point-of-care systems can find uses even prehospital care settings. "Ruggedized" handheld computers have been deployed in the field to aid emergency medical units in the delivery of care to patients.

In a point-of-care system, a clinician enters a patient's complaint, medical history, physical examination, orders for tests, surgical notes and or diagnoses into a computer in place of using a pen and paper, or other manual method. Advantages of point-of-care systems include quicker access to such information, simultaneous access by different healthcare workers, the ability to communicate orders more quickly, elimination of the difficulties involved in reading the products of poor penmanship, and the ability to harness integrated decision-support tools such as electronic formularies, drug interaction warning databases, and electronic implementations of practice guidelines.

Timeliness and Accessibility of Data

Point-of-care systems record data in real-time or near-real time; those containing dictation subsystems will receive data regarding a patient visit only after the provider has dictated the note and it has been transcribed. In some cases the appearance of the note in a point-of-care system may be delayed due to a provider's "batch" entering dictations at the end of a workday's visit. A point-of-care system may also be set up to receive laboratory data directly from a laboratory's

own system; a similar arrangement may exist regarding radiology reports. Thus, from the timeliness-of-data-availability viewpoint, point-of-care systems have excellent potential to contribute to public health surveillance.

Point-of-care systems use a variety of hierarchical and relational database storage technologies; many systems are based on MUMPS, a database with an associated language invented in the 1960's. Newer ones use relational technology. Some reside on mainframes, some on client-server systems, and the newest are web-based. The point-of-care systems relying on hierarchical database technology are more difficult to adapt to new uses, but mapping software has been developed to ease the placement of data in relational databases. A variety of views through the database can be established, and many systems offer access through both traditional interfaces and Web-based interfaces.

Reliability and Integrity of Data

Data are entered into point-of-care systems via a number of pathways. Clerks may enter basic patient demographic information; doctors and nurses may enter vital signs, chief compliant, medical history and physical examination information. Point-of-care systems do not all look the same. Commercial vendors sell some, while others are designed and maintained by successions of post-doctoral fellows. Some are general purpose, while others are optimized for certain specialty services (for example, a prototype project sponsored in 1996 by the American College of Obstetrics and Gynecology). Some require explicit entry of numeric values into fields, while others use menu bars, buttons or check-off lists. Each requires some degree of familiarity to ensure efficient, reliable data entry.

The greater structure of data entry characterizing point-of-care systems, when compared to a pen and paper, can, on the one hand, improve the reliability of patient encounter notes by forcing providers to consider and address many explicitly presented tagged fields. On the other hand, any point-of-care system, especially if its interface is poorly designed, can sometimes confuse a provider not prepared to deal with it. This can result in errors or omissions in data entry; so far, however, this has not led to decisions to abandon point-of-care systems as a tool for patient care [28, 29].

Limitations of Point-of-Care Systems

Unfortunately, in the US, point-of-care systems are deployed in fewer than a quarter of hospitals nationwide, and even fewer physicians' offices; even this number can overstate the extent of their actual availability, since for example, in this statistic, a point-of-care system present in but one department within a hospital is sufficient to be counted. A small minority of agencies responsible for prehospital care, such as fire and emergency medical services, utilize point-of-care systems. The Surgeons General offices of the nation's military services, when interviewed, were not aware of wide-spread use of point-of-care systems in military facilities; if they are deployed, such deployments may be scattered in specific facilities. The Veterans Administration, on the other hand, has high level of deployment of point-of-care systems. Reasons for low utilization include cost and reluctance by physicians and other providers to adopt these systems. On the other hand, a large multi-hospital organization may make a strategic decision to deploy point-of-care system, and, over time, those systems may become available to all member hospitals – if initial "rollout" is considered successful. That is not always the case.

In the United Kingdom, by contrast, point-of-care systems are ubiquitous. The value of these systems for public health surveillance is illustrated by the rapidity with which the UK will be able to implement an Anthrax surveillance strategy. By changing the decision support logic only, once, in a central location, the ability to detect postal workers presenting with Influenza symptoms at the time of phone or physical presentation to any primary care physician in the country, with central reporting will exist. Moreover, the ability to provide advice at that very moment to the physician seeing the patient will also exist.

The relative paucity of those systems in the US limits their usefulness in public health, as well as the lack of standard approaches among vendors that would allow single point changes to produce immediate front-line capabilities. If there are many more healthcare facilities a patient can present to with no point-of-care system, than facilities with point-of-care systems, a point-of-care-system-centric electronic surveillance strategy may not work well, especially if the goal is detecting a small outbreak.

Until the number of point-of-care systems deployed increases substantially, the role of these systems in an early-warning network will be limited. The immense benefits for public health surveillance that accrue from having a real-time interaction with clinicians actually argue, in spite of the limited deployment of such systems at present, that the federal government take steps to accelerate deployment.

Patient Web Portals and Call Centers

Two additional types of information systems are beginning to be deployed in health care. Patient Web portals are designed to provide scheduling and triage information for patients. Call centers are designed to provide similar functionality for those patients that do not have access to or the inclination to use Web services. Call centers also support communication among physicians as in the case of physician-to-physician consultation. The types of data collected by these systems justify their inclusion in this discussion, despite their very low market presence. Patient Web portals have the potential to collect symptom level data as early as the day of onset of illness. Call centers, if patients are encouraged to use them early, rather than waiting for illness to progress, have similar potential. The reliability and availability of such data have potential to be very high, especially if such services are designed from the ground up with the needs of regional integration of data for public health surveillance purposes in mind.

Application Service Providers

A last important feature of the clinical computing landscape is a trend towards the application service provider model (ASP). ASPs are companies that host applications in central location allowing health systems to outsource their data processing. Clinicians interact with the server-based applications over private or public networks. An example of a company that provide such services is Siemens Medical Systems, which hosts over 1,000 hospitals (and for each hospital, scores of information systems). The interesting thing from a public health surveillance perspective is that an ASP has ability to, after obtaining appropriate legal and administrative permission, access data collected by many health systems, and, moreover, to start to address the difficult data integration issues for public health surveillance purposes. Thus, the physical co-location of thousands of clinical information systems in a single location is helpful, but it represents a 20% solution with the residual 80% comprising unaddressed confidentiality, organizational, vocabulary, and other data integration issues.

Summary

Clinical data are highly relevant to public health-surveillance. Clinicians and health systems are a primary point of data collection about the sick, including data about demographics, risk factors, symptoms, signs, special testing, and diagnoses.

Clinical data in the US at present, however, are not highly available to public health. Although health care has long used information systems to perform *administrative* tasks (and market penetration is high for such systems), the use of computers to record clinical information has lagged administrative use and market penetration is variable depending on the type of system. Clinical information systems are widely deployed in clinical laboratories, radiology departments, and for registration; less so in pathology departments; and least used as point-of-care systems. Some administrative systems—like registration, scheduling and billing—have data that are of value for public-health surveillance and developers of new strategies for early detection of outbreaks are using them.

Specific data that are highly available include chief complaints, demographic information, and laboratory results. Key gaps are symptom and sign data, which are often recorded using English, not computer encoding.

Point-of-care systems represent a unique type of system not only within the clinical chapter of this report, but for the entire report. The reason is that point-of-care systems afford public health an opportunity to interact directly with front-line clinicians in real-time. Thus, if a patient is identified automatically by public health surveillance to require special treatment or data collection, a point-of-care system provides a channel into the process of care to affect those goals. In addition, point-of-care systems are capable of collecting symptom and sign data needed for public health surveillance.

Clinical data systems tend to be highly available, and the accuracy of data is dependent on type but tends also to be high.

Barriers to integration of clinical systems into public health include legal, administrative, and technical barriers. At present, health systems typically provide detailed information to public health only about patients with notifiable diseases. Although public health has a legal basis to access any data needed for public health purposes, routine access of data needed for early detection is a new area. There are few models for how such use would be supervised or supported financially. The technical barriers include differences in how data are coded and represented that would involve expensive interface development. A bright note is that the clinical computing industry has been working on the problem of interfacing and data integration for several decades, so there is a large body of work already completed towards solutions that can be applied directly to the problem of integrating clinical data into public health practice.

Questions for Further Study

1. What proportion of different types of clinical data needed for public health surveillance is available at present electronically from secondary sources? Which types of clinical information systems contain the data?
2. How many vendors offer each type of clinical information system? What are their market shares?

3. If we were to focus exclusively on application service providers, what proportion of the data needed would be available?

Chapter 5. Emergency Call Centers

Emergency call centers are facilities that receive telephone calls from persons requiring immediate assistance to prevent the imminent loss of life or property, to prevent injury, or to address a potentially hazardous situation requiring immediate attention. Government agencies or private contractors may operate these facilities. Emergency call centers include 911 call and dispatch centers; military call and dispatch centers; commercial call centers such as General Motors' "On-Star"®; and transport company (truck or rail) control centers.

911 Call Center/Dispatch Computer Systems

How 911 Systems Work

911 systems were developed to allow a person to contact emergency services by dialing an easily recalled, short string of digits. Incoming calls are routed automatically by the telephone company's central office, via dedicated lines, to a Primary Service Answering Point (PSAP). In the vast majority of jurisdictions, law enforcement personnel operate the PSAP because most of the call volume over 911 systems is either law-enforcement related or may require the involvement of a peace officer. In some systems, especially in smaller communities, the PSAP operator will handle all callers regardless of the nature of their emergency needs. Other PSAPs, such as those in cities, will route the caller to a Secondary Service Answering Point, operated by the appropriate fire or emergency medical service. The PSAP operator talks to the caller; simultaneously, he/she enters information via computer keyboard to emergency service workers, and/or assist the SSAP in doing so. The operator(s) can also speak to emergency personnel by radio.

911 systems are widely available in the United States**.** A 1997 Bureau of Justice Statistics survey found that Enhanced 911 (E911, defined below) was operational at 83% of all American law enforcement agencies (90% of municipal police agencies, 80% of county police agencies, 79% of sheriff's departments) and the National Emergency Number Association estimates that 85% of jurisdictions are equipped with some form of 911. *This means that 90% of people living in the United States can call emergency services by dialing 911.* There are still jurisdictions, however, which do not offer this service. In those areas, callers must dial specific, separate 7-digit (or 10-digit) telephone numbers to reach law enforcement, fire or emergency medical services.

Basic 911 systems, deployed decades ago, cannot provide an address to the PSAP operator. Basic 911 may be equipped to provide the caller's telephone number to the operator and the majority of 911 systems have this capability. E911 adds Automatic Number Identification (ANI) and Automatic Location Identification (ALI) to Basic 911. In enhanced 911, ANI is used to retrieve a street address or precise location from an ALI database (ALI is actually a stand-alone database in its own right, and could be accessed by any Basic 911 system if the operator merely had an additional computer terminal in front of him/her).

Emergency call centers in most medium-sized, and virtually all large cities utilize Computer-Aided Dispatch (CAD) systems. CAD stores and retrieves for the operator call histories for a given location or caller (e.g., previous altercations with police officers, previous fire, previous call for cardiac arrest, previous "HazMat" dispatch); the location of desired response resources, such as patrol officers, fire trucks, ambulances and special response teams; and additional notes received from responders in the field. CAD systems help the 911 operators to efficiently manage the call and resources deployed in response to the call. The operator can locate and dispatch the most appropriate responders to a call. CAD systems may also be linked to Geographic Information Systems (GIS).

Figure 5.1 illustrates the architecture of a Basic 911 system:

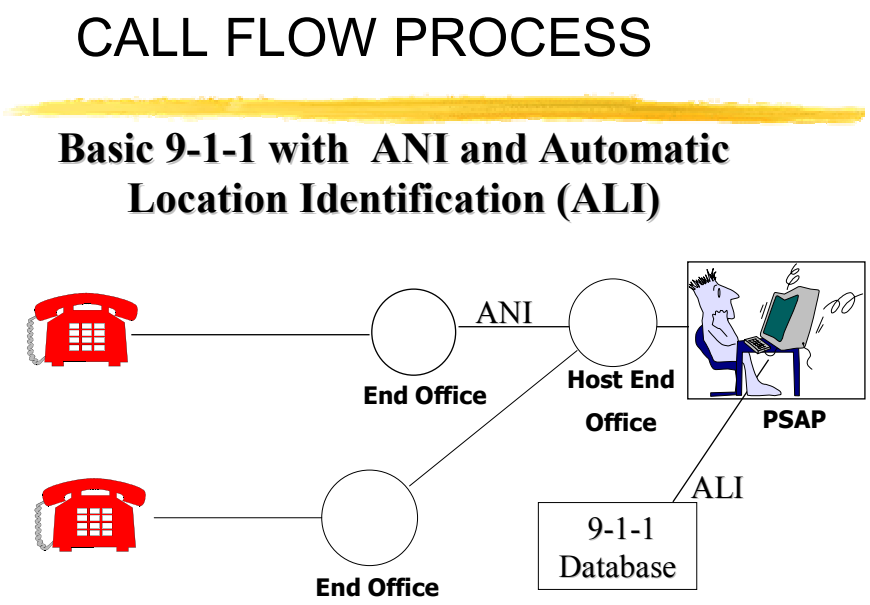


FIGURE 5.1 – COURTESY OF NATIONAL EMERGENCY NUMBER ASSOCIATION
End Office = Telephone company switching office local to caller
Host End Office = switching office operating 911 PSAP

Industry professionals estimate that roughly half of 911 systems in the US are equipped with an integrated CAD system. Figure 5.2 shows a typical E911 setup.

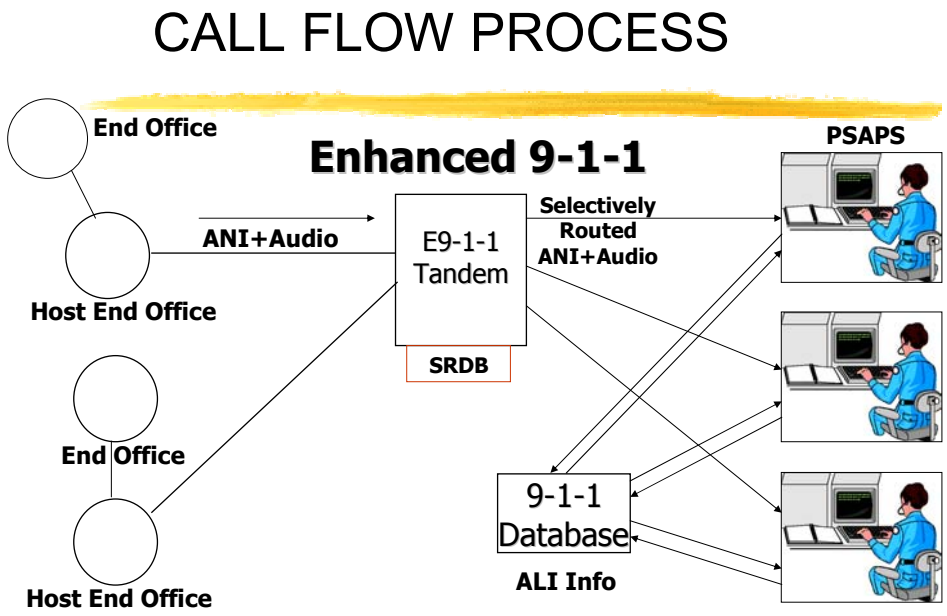


FIGURE 5.2 – COURTESY OF NATIONAL EMERGENCY NUMBER ASSOCIATION
End Office = Telephone company switching office local to caller
Host End Office = Switching Office operating 911 PSAP

E911 systems currently do not work with mobile phones (cellular phones, satellite phones) because the technical means to automatically locate a caller using a mobile phone have not yet been implemented in most jurisdictions. If the caller specifies the address where emergency response is needed and a particular jurisdiction has installed CAD, the operator will be able to retrieve and display address-specific information stored in CAD.

The rapid penetration of cellular phones into the marketplace is already limiting the utility of automatic caller location feature of E911 for emergency services. This growing problem has sparked efforts to implement technical improvements that would allow the automatic location determination of cellular phone callers. This determination may be accomplished in two general ways: by triangulation of a signal using multiple cell towers, or by the use of Global Positioning System (GPS) satellites. The latter technique is already used by many commercial transport companies, which equip their trucks with GPS receivers.

The difficulties in automatic location detection of cell phone users are illustrated by Globalstar, a provider of handheld satellite phones. If a caller using a Globalstar® phone, with satellite service enabled, calls 911, the PSAP operator can contact Globalstar on a company-supported hotline, where a technician can locate the caller by triangulation. The phone must be within line of sight (LOS) of at least two of Globalstar's 48 satellites, so locating the caller may require a delay of up to several seconds if initially only one satellite was within LOS of the caller. Unfortunately, the location precision achievable currently is limited to 10 square miles. Because Globalstar® phones are also compatible with cellular systems, and can be used as cellular phones, locating a Globalstar caller by triangulation of cellular towers would also be feasible.

Other potentially useful systems include Lo-Jack®, which, when activated, allows police to track the location of a stolen vehicle).

Implications for Bioterrorism Detection and Public Health Surveillance

A 911 system is actually a set of components that work together to provide communications and data storage functions. The component with the most immediate potential for public health surveillance is the Computer-Aided Dispatch (CAD) database, which stores, in real-time, the complaint, previous call history from the location, and information supplied by emergency responders at the location. This information includes pre hospital care information regarding patients, as well as information about sick animals, contaminants (and human/animal exposure to them) and other directly relevant facts. The information overlaps with information collected by clinical information systems to some degree, but not entirely. Some 911 callers may not seek medical attention, or in large-scale outbreak situations, may be referred or routed to locations that do not have clinical information systems.

Other 911 components play supporting roles: The ALI function provides instant location identification. A Geographic Information System (GIS) can help place all of the above in context with other significant sites in the area, however it makes more sense architecturally for that level of analysis to reside on the public health analytic side, not at the data collection side. Therefore, geographic information systems will be described in another chapter.

Side Bar: Field Use of On-Line Databases

One very important aspect of police dispatch, directly related to data accessibility, is the recent introduction of CDPD (Cellular Digital Packet Data). CDPD is a wireless encrypted data network employing 128-bit (DES) encryption. CDPD coverage is very widespread, and although operates at a low rate of data transfer (19.2 k) it is anticipated to increase to 28.8 within the coming months.

Following the Los Angeles riots of 1992, supporters have touted CDPD as a strong supplemental system assisting police agencies in the dissemination and accessing of data through police units employing MDT's (Mobile Data Terminals).

Linking MDT's with CDPD enables law enforcement officers in the field to readily access their Record Management Systems (RMS) and other databases. While many of these databases, such as the National Crime Information Center (NCIC) and/or State Crime Information Center (SCIC), are clearly outside the scope of this report, CDPD technology also makes feasible access to databases of direct interest to public health authorities.

MDTs are not limited to laptops; recent advances are also allowing local law enforcement to consider wireless palm pilot services or BlackBerry® handheld computers for similar applications. The World-Wide Web can be used for this purpose; virtually any Web-enabled device can participate.

CDPD / wireless web applications must be kept simple, with little image transfers or large data files kept to a minimum. Data distributive systems using MDT/CDPD could also be invaluable during major catastrophes (such as the bombing of the World Trade Center) because they are supported by the inherent robustness and flexibility of the Internet; however, there still exists the potential for failure if the telephone system itself fails.

CDPD is available in virtually all telephone company service areas in the United States, except a few remote, rural areas. Not every 911 call center in a CDPD-capable area harnesses it however.

Military Call Center/Dispatch

Although military systems are beyond the scope of this report, we provide a brief summary of military call center functionality. The National Emergency Number Association considers military bases to be underserved by 911 when compared to civilian jurisdictions. The majority of emergency services on military bases in the United States are equipped with at least Basic 911, but there are still many locations, which require callers to dial seven or ten digit numbers to reach emergency services. The 911-equipped centers are, however, not uniformly equipped with CAD technology. A particular 911-equipped base may have E911 with CAD, or it may have Basic 911 with ANI and ALI, or ANI only.

When callers must dial all seven or ten digits, the dispatcher may be able to display the originating telephone number, and may also be able to manually retrieve an address (but this information would not be available on-line). Often, base services are all dispatched centrally, meaning an incoming call is handled by one PSAP, and the same dispatcher will initiate and coordinate the response of military or DoD police, fire and medical services as needed. Some military bases are tied into civilian 911 systems. For example, Fort Belvoir, Virginia is served by Fairfax County's 911 system. A caller who is located on the base dials 911 and reaches a Fairfax County PSAP operator. That operator acts as the base' dispatcher, alerting military police, fire or EMS workers as needed.

As is the case in the civilian world, E911 and CAD functions are not yet available for callers using cellular phones.

Military and DoD personnel can also reach 911 via satellite phones. The Department of Defense awarded Iridium a contract to provide services to the department. Iridium is a service provider operating a network of 66 satellites in low-orbit, and is a direct competitor to Globalstar. Iridium's phones, built by Motorola, can offer automatic 911 localization in a manner similar to Globalstar's service; however, Iridium® phones are not compatible with cellular systems, and must rely on a terrestrial wireless switch to connect a satellite phone user with landlines or cellular callers. The location of an Iridium user cannot be determined by cellular tower triangulation. This means locating, in real time, military callers (and placing this information in ALI) using Iridium is more difficult than locating civilian callers using satellite phones when they can access the cellular phone network.

Commercial Assistance Call Centers

Emergency dispatch centers for commercial transport companies do not participate in 911; however, they employ data systems that contain information of direct interest to public health. Railroads and trucking companies set up control centers that stay in contact with their trucks and trains. Larger enterprises track the location of their trains and trucks with systems that rely on global positioning system satellites. Their data systems contain information from train engineers and truck drivers who are in direct contact with control centers that dispatch assistance in emergencies. Railroads usually post this number at railroad grade crossings, so as to allow passers-by to call the railroad in the event of an accident or derailment. Passers-by witnessing an

emergency involving a truck or train must either use 911 or, in the case of railroads, a toll-free number to reach a dispatcher.

The systems currently employed by trucking companies can automatically track not only where a truck is located, but also whether the trailer is empty or full, or disconnected from the tractor. Some systems can even remotely monitor engine performance to detect breakdowns and automatically summon aid to the driver. Work logs can be submitted automatically as well. Drivers can reach a control center operator at any time, as needed, and so the company's call center will often become aware of a problem before a public 911 center will.

These companies, with fleets of thousands of tractors and trailers, also employ advanced, computerized logistical systems allowing a shipped item to be tracked continuously from shipping dock to receiving dock. Since some are licensed to carry hazardous materials, pharmaceuticals, biological samples and other items of interest to a bioterrorist, the companies' tracking systems and call centers are important sources of data to consider. Since these modern systems use relational databases, they will generally be technically amenable to modifications for purposes of public health authorities. Access to these databases is as valuable to public health authorities as access to CAD databases, described earlier.

[Sources: JB Hunt Transport, Schneider National Carriers, Werner Enterprises, CRST, CSX, and Norfolk Southern.]

Poison Information Centers

Poison information dissemination is the primary role of poison information centers. The information is provided by Specialists in Poison Information (nurses and pharmacists) and clinical toxicologists. Each patient interaction is documented electronically and constitutes the poison center medical record. The record contains the standard personal and demographic information about the patient as well as specific information that are categorized by substance (the poison), treatment, patient symptoms, route of exposure and laboratory values. Each of these areas constitutes a fully searchable section of the medical record. The values within each section are standardized and coded so that data from all poison centers in the United States can be incorporated into a single database. There is also a free-text documentation section. The free-text section is not searchable as part of the standard poison center medical record and is therefore not included in the national database on poisoning exposures.

Currently, data are submitted to the American Association of Poison Control Centers Toxic Exposure Surveillance System (AAPCC TESS) on a monthly, quarterly or semi-annual basis by participating poison centers. Sixty-three poison centers submitted data in 2000 and only Alaska, Hawaii, Mississippi, northwest Ohio, South Carolina and Vermont are not represented by the AAPCC TESS data [30]. Participation in AAPCC TESS and submission of all data is mandatory for Certified Regional Poison Information Centers and voluntary for non-certified centers. Data are reviewed and published on an annual basis approximately nine months after the end of the calendar year. Since data submission is intermittent and partially voluntary, there is no real-time surveillance of national poison center data for toxidromes that are consistent with exposure to biological and chemical terrorism agents. Furthermore, most individual centers do not conduct real-time toxico-surveillance. Figure 5.3 shows the type of surveillance that is feasible using poison center data.

However, poison information centers collect a large volume of information on exposures involving both humans and animals. For example, the Pittsburgh Poison Center responded to approximately 78,000 inquiries in 2000. Nearly 5,000 of the inquiries involved animal exposures. This reflects the public awareness of poison information services. This call volume, the active data reservoir, and the reliance upon the poison center by the public and medical professionals make the poison center an ideal active surveillance center to identify sentinel events and to profile medically and demographically the nature of a biological or chemical terrorist event.

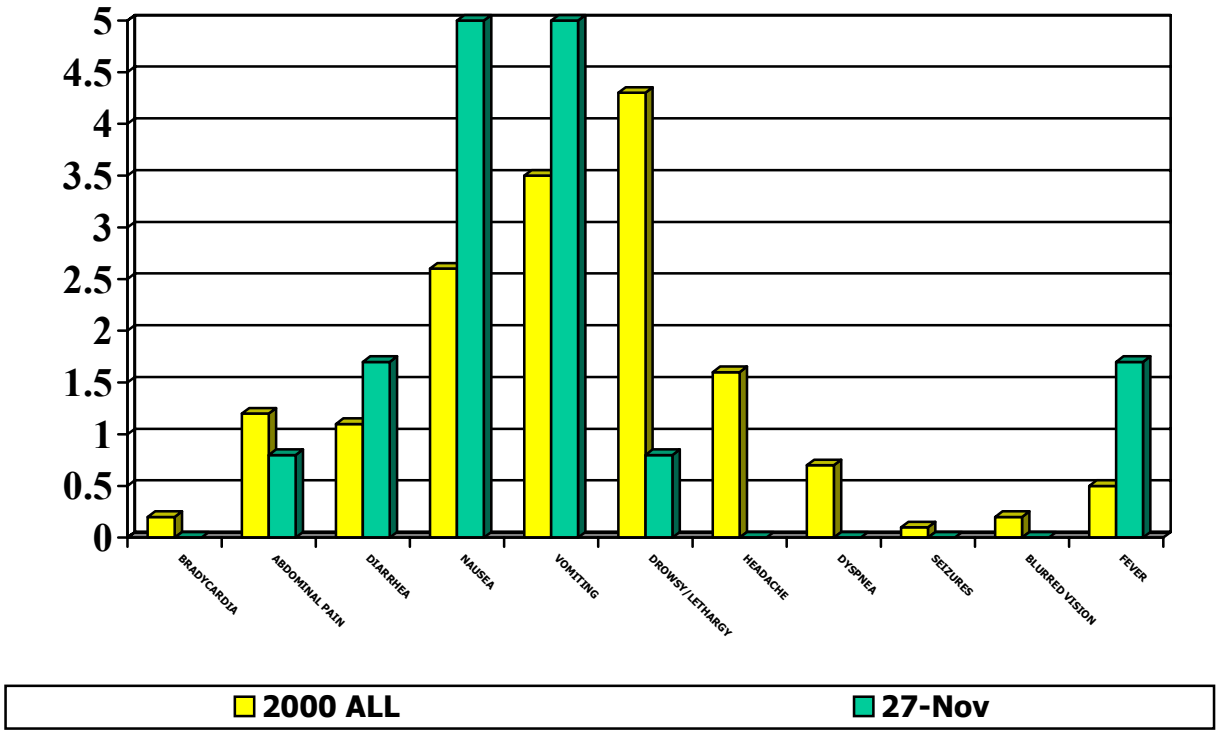


Figure 5.3 Daily Surveillance report from Pittsburgh Poison Center for November 27, 2001. Abscissa is the percent of calls normalized by total calls. Note the increased proportion of calls for this date relative to the average for diarrhea, nausea and vomiting.

Software to facilitate the identification of sentinel events or toxidromes through the surveillance of real-time poison center data has not been developed. Currently, the only way to conduct analysis of data is through the identification of toxidromes based on expert opinion and then the tedious analysis of data to identify the presence of those toxidromes. The data are available readily and are accessible easily through the use of any database management system. Most poison centers document cases in real-time, which make the data fields available for immediate analysis. It would not be a challenging task to conduct real-time background surveillance of a single poison center's medical record database. However, poison centers within most states and throughout the nation are not linked in real-time and all national data analysis is conducted retrospectively. The challenge is to network US poison centers so that data are contributed real-

time into a central repository. Thresholds and sentinel effects need to be identified. The use of computerized artificial intelligence may help to identify sentinel events that are occurring, below the established threshold of recognition. The collaboration of data that includes poison center data, non-prescription and prescription pharmaceutical sales, emergency department diagnoses, veterinary clinic data on companion and large animal problems, etc. could help to identify problems in the early stages of their evolution.

Summary

Emergency call centers are equipped with data systems consisting of a number of elements. Among the databases used in these centers, Computer-Aided Dispatch (CAD) databases are the most useful, and immediately relevant, to public health authorities, because they can reveal, in real-time, information from the scene of an incident involving a sick person or animal, including patients' pre-hospital data, past history at the scene, and, if appropriately linked to GIS, geographic information. The data include many items entered by professional responders and caregivers, and so are generally reliable. These systems cover a wide area of the United States, since Enhanced 911, with CAD, is widely deployed in this country. Systems currently being deployed utilize modern database technology, so access does not present unduly difficult technical challenges.

Public health users of call center data must observe proscriptions on the use of confidential, personal information about callers. Cost is also a factor. Although obtaining access is not very difficult, municipal and county budgets are often constrained, so that public health authorities must be prepared to support the necessary software and hardware components

Other systems that may contribute information of use to bioterrorism detection are commercial emergency call centers. Poison information centers collect large amounts of information from callers about poisoning incidents, and this information can be readily extracted from their databases. However, truly useful, nationwide public health surveillance, utilizing poison center data, awaits the linking of these centers in a network, as well as new developments in automating toxidromes surveillance and recognition.

Questions for Further Study

1. What factors prevent 100% coverage of the US with E911 services?
2. How is E911 with CAD being implemented for cellular phones and when might full functionality be expected?
3. What additional kinds of data would be desirable to record in CAD? E911?
4. What kind of data exchange exists between 911 and commercial call data centers?
5. What percentage of the US commercial truck fleet is monitored by call centers? What percentage of Canadian and Mexican trucks using US highways is monitored by such systems?
6. What potential is there for integration of poison control and hazardous material data, 911 data systems into public health surveillance?

Chapter 6. Public Health Systems

Epidemiologists have been developing systems to collect data ever since John Snow systematically collected information about cases of cholera in London, England.[31] In a previous report to AHRQ entitled *The Nation's Current Capacity for the Early Detection of Public Health Threats including Bioterrorism,* we discussed current US systems. In this chapter, we provide a brief summary of current systems and focus on issues of data content, availability, timeliness, and other data characteristics.

Current Public Health Systems

Data are generated by local and state health departments directly through their own investigations and practices, and are also received from medical laboratories, health practitioners, and private citizens. The data become available electronically relatively late in the collection process in most settings. The data eventually are transmitted via the national reportable disease system to the Centers for Disease Control though the National Electronic Telecommunications Surveillance System (NETSS)[32] Fifty states, the District of Columbia, New York City, and five US territories report a core set of data (date, county, age, sex, and race/ethnicity) on the occurrence of each reportable diseases using this system.

Other infectious diseases surveillance systems include PHLIS (Public Health Laboratory Information System) [33], which collects culture results. There are also special purpose, "single-disease" systems used to collect data not available through NETSS but important to federal goals. These single-disease systems function either through the states (e.g., the National Malaria Surveillance Systems[34]), cities (e.g., 121-cities mortality reporting systems) [35], or clinical settings (e.g., United States Influenza Sentinel Physicians Surveillance Network) [36]

The data collected by Pulse Net—the national molecular subtyping network for food borne disease—are interesting. The Pulse Net network of laboratories *fingerprints* (determine the genotype or other very specific information about organisms) and then places the results in

databases where they can be matched against other submitted organisms. This system, however, is not yet universally implemented. These data are available in a central location.

Non infectious-disease surveillance systems were not specifically covered in our previous report. Most do not collect data directly relevant to bioterrorism detection, nor do they satisfy the timeliness requirement. Nevertheless, these systems are a resource of interest because their available personnel, telephones, and computers could be incorporated into an emergency system for the conduct of rapid case control or cohort studies via telephone. Such studies could be helpful to determine the extent of illness in a community as well as provided data necessary to calculate attack rates for different demographic and epidemiological conditions. These non-contagious disease systems include:

- Special Disease Reporting Systems (e.g. lead, child abuse and neglect, cancer, new born screening)
- National Health Care Survey
- National Health Interview Survey (NHIS)
- National Survey of Family Growth (NSFG)
- National Health and Examination Survey
- State and Local Area Integrated Telephone Survey
- Pregnancy Risk Factor Surveillance System (not PA)
- Behavioral Risk Factor Surveillance Survey
- Youth Risk Factor Surveillance Survey (not PA)
- NIOSH mortality data set
- Cancer: Surveillance, Epidemiology and End Results (SEER) Data Request (Not PA)
- National Electronic Disease Surveillance Survey
- NHSDA National Household Survey of Drug Abuse: Prevalence and correlates of substance use
- DASIS Drug and Alcohol Services Information System: Substance abuse treatment facilities data
- DAWN Drug Abuse Warning Network: Emergency department and medical examiner data
- Fatal Accident Reporting System - National Highway Traffic Safety Administration
- Sentinel Event Notification Systems for Occupational Risks (SENSOR) model

Administrative data systems provide information for monitoring compliance with laws, program management, and service reimbursement. The most useful systems in this category are the U.S. Census and hospital discharge data. Census data (mandated by the constitution for voting purposes) is useful for the calculation of incidence rates needed to determine if one subpopulation is being infected more than another subpopulation. Hospital discharge data can provide baseline data for the past prevalence of disease in the population being studied. Other types of data, not collected in a consistent manner but possibly useful in specific situations, include:

- Program Process Data (Inspections, clients served, inquiries responded to, outbreaks investigated, number of facility beds)
- Census data (including the Current Population Survey
- Hospital uniform bill data
- Medicaid billing data

- Census data
- Hospital Discharge Uniform Bills
- Immunization School Records
- WIC Records

Timeliness and Availability

Although electronic availability of data in public health varies dramatically by state and funding source, only rarely is data available electronically sufficiently early for bioterrorism detection. The National Electronic Data Surveillance System [37] is attempting to improve the timeliness of reporting of public health data. Forty-nine states received a total of over nine million dollars from the federal government related to preparation for NEDSS in FY 2000 and FY 2001 funding more than doubled.[38]

Data Quality

Existing public health systems vary in data completeness, but there is under-detection of outbreaks and under-reporting of infectious diseases. A 1998 CSTE report found that state programs without CDC funding tended to be inadequately staffed and equipped given their mission and the volume and nature of the work to be done.[39] The Centers for Disease Control recently published "Updated Guidelines for Evaluating Public Health Surveillance Systems Recommendations for the Guidelines Working Group" to help states improve the data quality of their surveillance systems.[40]. If followed these guidelines would improve data quality. However with no mandate and no funding, data quality will remain questionable in areas with few resources.

Legal or Administrative Barriers

Since this analysis is being done from the perspective of public health, there are by definition no barriers to public health access to public health data. Inter-jurisdictional reporting of data is more problematic, especially when the data are personally identifiable.

Technical Barriers

The technical barriers to data integration of public health data are formidable. Data are kept in a variety of systems that have existed over the years including Epi Info (sometimes in customized databases provided by the CDC), commercial PC and mainframe (Dbase, Access, Oracle, etc.), in-house programs and programs customized by outside vendors. This diversity of platforms and languages is explained by the categorical funding at the program level (programs funded by separate entities did not always communicate or cooperate).

Summary

Public health surveillance systems contain a wide array of data potentially of value for early detection. The majority of this information is simply not collected, transmitted or stored in a manner that would permit the timely analysis of the data. NEDSS offers the promise of improvement; however with insufficient funding in comparison to the number of states, health departments and systems, fulfillment of that promise remains elusive.

Questions for Further Study

1. What funding, technical resources, and strategies are available for implementing NEDSS-compliant systems more rapidly?

2. How can public health data collection and processing in states with few resources and unsupportive political leadership be improved?
3. What data need to be collected by Public Health? What data can be better collected by other methods?

Chapter 7. Food and Water Supply

Food and water are key potential transmission routes for a bioterrorist's pathogens. There are two strategies for monitoring for such contaminations: First, we can monitor food and water directly and prospectively to detect contamination and prevent or mitigate a potential outbreak very early. Second, once an outbreak begins, we can trace back from affected individuals to the source as quickly as possible to prevent additional exposures, and to direct prophylactic treatment to individuals likely to have been exposed to the pathogen. This chapter examines data needed to trace back from affected individuals to the original source of contamination and also discusses for water supply routine monitoring for contamination.[2]

Food Supply

Systems likely to contain information relevant to food contamination include systems in processing and manufacturing plants, including Pasteurization and Irradiation Records, and systems that track the distribution of food from the source to the grocery store.

Trace back investigations[41]involve tracing a product by using shipping and purchase records at each place in the distribution chain back to common points that can explain the occurrence of illness among all or most affected individuals. When consumers falls ill after eating food purchased from a retailer or restaurant, investigators will ask the affected individuals to identify the restaurants or retailers they patronized, who, in turn, identify the wholesalers. The wholesalers identify their suppliers, who identify the farm or farms that were the ultimate food source. The supply chains' information systems are not vertically integrated. Each entity knows only to whom they shipped, or from whom they received food. Public health police powers can provide the legal basis for trace backs within state lines; however, in many cases the FDA may become involved. Trace back by the FDA involves many challenges including the absence of records, a lack of authority to require records be kept or provided, multiple sources of product,

[2] Governmental prospective monitoring of food is limited. Spot checks of food and produce are made when they enter the country with high risk items getting more attention, however only a small fraction of food gets inspected prospectively. State authority does not extend outside its boundaries and FDA does not have authority to require that comprehensive records be kept.

complex distribution systems and the resource-intensive nature of the process which may or may not confirm a contamination.[41]

At each point in the trace back, the procedures for storing or processing the food are observed for abuses that would result in contamination of the food. Specimens may be taken to determine if other similarly handled food or the environment that the food was stored or processed in is contaminated. If there is evidence to suggest that contaminated food is in the distribution chain than a recall of that food is initiated.

Water Supply

The public in the United States obtains drinking water from but a few sources. Municipal water departments deliver water from watershed sources (springs, streams, rivers, reservoirs), through mains, to taps and to water fountains. Bottled water from privately owned watersheds, which may be filtered and/or ozonated to kill microorganisms, is sold to the public in retail outlets. Some landowners obtain drinking water from privately held watersheds or from wells. Although law governs the drilling of such wells, their operation and maintenance is left to their owners, with government authorities providing education and water testing services as required.

In cases where water is supplied by private or publicly held companies or individuals, as opposed to municipalities, these entities may or may not filter the water, as required by government authorities such as the EPA, and/or chlorinate the water prior to its delivery to the customer. .

Municipal water utilities regularly test water supplies for contaminants and microorganisms, employing either their own work forces and facilities, or contract laboratories. In addition, their employees field calls regarding alleged illnesses resulting from consumption of drinking water. The water utility may be the first to hear of an alleged incident; others who may be the first to be alerted include the victim's personal healthcare provider, or a public health agency, such as a municipal or county health department. It is important to note that, often, problems with drinking water are discovered through water testing, and corrected, before anyone has the opportunity to become ill.

Timeliness of Incident Reporting

When victims report alleged illnesses due to contaminated water to physicians or health departments, investigations will result in data collection by the health department or physician, and these data sources are discussed in other sections of this report, but may result in significant latencies due to current public health surveillance practices. The water utility will record these incidents itself, of course, but the timing of their appearance in its own records, whether electronic or otherwise, will depend on a number of factors: where the original report is made; efficiency of verification and laboratory testing; and efficiency of its information systems.

If the incident is reported first to the utility, the appearance of an incident may or may not be recognized as a public-health threat and so may or may not appear promptly in the utility's database.

One suburban water supplier in southeastern Pennsylvania employs technicians who are trained to receive, assess and immediately plot such reports to facilitate rapid dispatch of personnel to the incident. However, as stated before, if this utility is first notified by a county health department, the time lag between the incident and its recording in a database can be as long as

several days. This occurs despite the company's efforts to maintain a close working relationship with health department staff. When interviewed, the company stated it was unusual for a utility to operate a telephone special response center. Complaints directed at other utilities are directed to plant managers for evaluation and response.

Once a report is received, the veracity of the report must be confirmed. Water utilities perform water chemistry and microbiological testing. Some utilities operate their own laboratories, while others out-source this work to commercial laboratories. The resulting logistical issues influence the time latency for relevant data; if a contract laboratory processes samples for several utilities the "turnaround" times for sample reporting may be longer than for utilities with their own laboratories. Some utilities have their sophisticated, highly automated laboratories sited adjacent to water intake facilities. Water samples collected in the morning can be tested, and results reported, by afternoon. Since the modern laboratory's instruments can, following certification by a technician, send all results to a relational database with just a keystroke, these data can be available immediately to public health authorities.

In summary, a detection strategy based on continuous monitoring for contaminants should work well for those contaminants that are routinely being monitored in municipal facilities that have good monitoring. Since this monitoring occurs at the same locus as control measures, this process is very effective and cannot be improved upon much other than to extend the number of water sources being monitored and the number of organisms and threats monitored for.

For those contaminants not being routinely monitored, or for non-monitored water, the second strategy based on working backwards from affected individuals comes into play. Here the existing strategy involves substantial delays, especially if cases are reported through the public health system.

Utility Laboratory Systems

Nearly all large water utilities employ laboratories with highly automated equipment capable of populating a database with data with minimal human intervention. Human input is needed for microbiology culture interpretation, as well as other categories of data; thus, the operation of these laboratories is not dissimilar to the operation of clinical laboratories. The laboratory systems receive data soon after samples are collected and submitted; however, because the latency from incident to utility notification can vary, depending on how the notification occurred (see above), the additive effect of the laboratory time lag, itself quite short, and the notification lag, can significantly reduce the laboratory system's early detection potential from the point of view of incident reporting.

Utilities perform proactive sampling, with the most relevant contaminants being sought daily. If funding is provided to water utilities to perform daily sampling of potential public health threats, the short "turn-around" time and highly efficient laboratory systems combine to provide a potentially very effective early-warning system for bioterrorism and other threats (see "Water Supply Sampling and Surveillance" below).

The means by which utilities recognize a "cluster" of cases can vary. One suburban utility does not utilize any form of geographic information system (GIS), but relies on a small group of response technicians who are intimately familiar with their service territory. Another uses a wall-mounted pin map.

It is important to note that the current state of water utility information systems does not reflect utility response capability per se. The "standard of care" at all large utilities is to respond immediately to complaints; the source of a reported problem is usually determined within a day by utility technicians, and most certainly within 48 hours, when a culture is used to evaluate contamination incidents. But this information may or may not be available in electronic form and if it is, it may not be easily accessible to government health authorities.

Water Supply Sampling and Surveillance

Water utilities regularly, proactively, test samples of drinking water for contaminants, whether or not customers call to report a problem. A given utility will decide what to test for, and how often to test for it, based on what threats are prevalent in its service area. For example, a utility in a northern farm state may be concerned, and extra vigilant, about nitrates from farm runoff but not spend as much time on looking for Cryptosporidium, due to that organism's rarity in the utility's water sources. An urban water utility in the Northeast may have somewhat different priorities. Regardless of specific local priorities, the purpose of proactive testing is to detect contamination of drinking water before anyone becomes ill from it.

The Safe Drinking Water Act, as amended in 1996, requires the U.S. Environmental Protection Agency (EPA) to establish "criteria for a program to monitor unregulated contaminants and, by August 6, 1999, to publish a list of contaminants to be monitored" (from the Act Summary). EPA's response to this has been to write the Unregulated Contaminant Monitoring regulation. This final rule includes a list of contaminants to be monitored, procedures for selecting a representative nationwide sample of small PWSs that will be required to monitor, the frequency and schedule for monitoring, the sampling points, the approved analytical methods to be used, and procedures for entering the monitoring data in the National Drinking Water Contaminant Occurrence Database (NCOD), as required under section 1445 of SDWA, as amended" (from the Act). According to EPA's website, the NCOD is intended to contain data without analysis; EPA intends for interested parties to download datasets, and perform analyses off-line. This makes it less of an early-warning tool, and more of a quality assurance or analytical tool.

All utilities serving at least 10,000 people each are required to comply, as is a representative group of water utilities serving fewer than 10,000 people. Utilities were to submit water sampling data on a quarterly basis to the EPA, intended for the NCOD database. The EPA required these data to be submitted in a particular format; the utilities interviewed complied with varying degrees of difficulty and cost, in part related to delays by the government in supplying necessary software. While NCOD does not appear to be intended as an early warning system, the challenges created by its deployment point to similar challenges which might be expected in deploying an early-warning or sentinel water contamination system. A more promising approach involves the querying of utility laboratory systems.

As mentioned above, most large utilities operate, or contract with, highly automated laboratories. Many of the test results achieved are reported automatically by the instruments to a database. In the case of contaminants or chemical levels which are tested for daily, this provides a consistent and reliable source of data which is no more than 24 hours old, or 48 hours old in the case of microbiological cultures. Since most databases in use today by water utilities are relational, SQL-compliant systems, they are amenable to query by public health authorities. These systems can be powerful detection tools; moreover, their use would provide opportunities to detect a problem before anyone becomes ill.

In order to fully utilize these systems, public health authorities would have to decide which types of tests the utilities would be required to perform more frequently than they do currently. Water utilities would expect the government to pay for the cost of both increased testing frequency (manpower, lab usage, reagents etc.), as well as the necessary software or hardware modifications, which would be required, including such measures as providing "mirror" databases. These are copies of the master databases maintained at the utility; their use ensures that public health operations do not cause problems to the master data files.

A Model System for Water Monitoring

One utility interviewed had set up a system that could be considered a model. Contaminants of special importance to the local region are sampled for daily at water intake facilities and treatment plants; laboratories are located directly adjacent to water sampling points; laboratory instruments perform tests and transmit data directly to either an Oracle® or SQLServer® relational database; the system is Web-enabled and data are available on the Web to view. Other tests not performed daily can be changed to a daily routine as required. Testing performed in response to a complaint produces results within hours that are posted to the database and available for viewing on the Web. Extensive verification procedures are in place. This utility's data is readily available to the government, either from the Web or through customized views through the relational database. While not all utilities maintain this kind of database and Web access, the majority of laboratory systems currently in use can populate a Web site with data if appropriate modifications are made. The reason for this is that most laboratory systems in use today are SQL-compliant, relational systems. Successful adaptation and use in this way would still require significant investments of technical resources. Yet, this model system demonstrates what can be achieved.

In summary, although incident reporting through water utilities may not offer the kind of timely reporting needed for early detection of bioterrorism and other threats, proactive water sampling surveillance and reporting through laboratory computer systems does hold much promise.

Summary

The logistical systems supporting food supply in the United States are not vertically integrated. Each level of the industry is concerned only with its immediate suppliers and customers, and not with the entire supply chain. Therefore, a trace back investigation is virtually the only method available to public health authorities to determine the source of a problem. Challenges facing public health authorities include a complex web of relationships between producers and consumers, limitations to governmental authority and to existing information systems.

For water, the detection strategy based on continuous monitoring for contaminants should work well for those contaminants that are routinely being monitored in municipal facilities that have good monitoring. Since this monitoring occurs at the same locus as control measures, this process is very effective and cannot be improved upon much other than to extend the number of water sources being monitored and the number of organisms and threats monitored for. For those contaminants not being routinely monitored, or for unmonitored water, the second strategy based on working backwards from affected individuals comes into play. Here the existing strategy involves substantial delays, especially if cases are reported through the public health system. Improvements in the speed and completeness of case detection and epidemiological investigation for all threats will enhance the ability to detect water borne illnesses. Additional

data required to detect that cases are water borne include the ability to collect information from patients about their exposures to water and collate that information with similar information form other individuals with such symptoms and with detailed representations of the water supply system, including municipal, private, temporary water, and bottled water distributions.

Questions for Further Study

1. To what degree can modern procurement technologies be deployed to create vertically integrated information systems and audit trails, from farm to meal?
2. Can information technology use the pattern of cases to help focus the trace back investigation onto a subset of the points in the distribution chain shared by affected individuals, especially the labor intensive part of inspection and observation of food processing practices?
3. For which organisms, chemicals or other substances is daily drinking water sampling and reporting cost-effective? What methodology is available to determine the cost-effectiveness of such surveillance?
4. Should the federal government require periodic sampling of private water sources?

Chapter 8. Pharmaceutical Sales

The pharmaceutical industry invests heavily in information technology to support research and development of new medicines; clinical trials; manufacturing; marketing and distribution; post-marketing surveillance; and adverse events reporting. Electronic systems that collect data that could be useful in the detection of public health threats, such as bioterrorism, include on-line prescription verification and assignment systems, and post-marketing surveillance systems; the latter encompasses adverse event reporting (AER) systems. Over-the-counter sales data are also of great potential. We consider each group, mindful that the combined use of two different kinds of systems may be more effective than one type used alone.

Prescription Benefits Management /Prescription Verification

Pharmacies utilize on-line systems to verify health plan prescription coverage and to assign automatically the prescription to the appropriate supplier. The information recorded includes geographic and age information, the name and dosage of the drug, cost and, in a few cases, an ICD-9 code signifying the underlying diagnosis. Prescription data enter a transactional system in real-time; periodically--often weekly-- these data are uploaded to a data warehouse designed to provide analytical "data mining" services. The data warehouse, built around modern relational technology, can be made readily accessible to public health authorities; its value is limited by the inherent lag time. Coupling warehouse access with direct access to the transactional system feeding it, can potentially produce and up-to-the minute ability to detect and interpret trends in prescription drug activity that can indicate a threat to public health.

The transactional systems receive data about prescriptions as soon as they are presented to pharmacists for filling. Several companies who run such systems have been using legacy mainframe systems to store data, usually utilizing VSAM schemes, and these systems are somewhat difficult to interface with. To modernize these systems, and facilitate compliance with HIPAA, relational databases are currently being introduced, and will be on-line this year and make interfaces easier. External entities such as those charged with public health surveillance can obtain data from this transactional system in two ways: by receiving transactions in parallel to the transactional system, or by launching SQL-based queries, or using decision-support tools, against the transactional relational database. The external entities charged with public health surveillance may need several years of historical data for temporal analysis. They can obtain

these data one time from data warehouses, or can query the warehouses when the data are needed.

Although this scheme is technically feasible, it is important to note that the companies collecting these data will require payment for access to the data, as well as to offset the costs of any software modifications required to provide this new service.

Post-Marketing Surveillance Systems

Post-marketing surveillance of medicines falls under the jurisdiction of the US Food and Drug Administration (FDA). Intentional contamination of drugs has occurred and therefore it makes sense to review how pharmaceutical adverse-event reporting occurs in the United States, to identify and examine related electronic systems, and to ask whether additional systems are required to counter this threat. If the contamination is with a rapidly acting toxin, the relationship with the medication may be obvious from the co-location of the victim and the medication and there is a chance that the problem may be detected through post marketing surveillance. However, in most cases it is likely that the problem will be detected and characterized through public health of criminal investigation. In any event, both pharmaceutical manufacturers and the FDA maintain adverse-event reporting, or safety, databases. The FDA's system--described below--contains reports that have been vetted by pharmaceutical makers, but that are weeks old. Manufacturers' databases are populated with reports whose actual latencies for serious or life-threatening incidents may be as little as 48 hours, or as much as several days to weeks for minor or trivial drug reactions. In the latter cases, the patient may not have even bothered presenting to a physician for care. Because monitoring for adverse events is more extensive in clinical trials involving investigational drugs relevant information may be available earlier than 48 hours; public health authorities may use this to improve detection of events, such as drug tampering, which can occur during such trials, but this represents a small part of the overall problem or drug tampering.

An adverse event report may vary in content, depending on the source of the report. A report from a physician will generally be the most complete, including a description of the complaint itself, the affected patient's relevant medical history, results of physical examination (if performed) and results of laboratory tests (if any). A call to an adverse drug reporting hotline by a consumer may yield a complaint that something is wrong, as well as some medical history, but will not include physical exam results or other data obtainable from a physician. A competent hotline interviewer has the responsibility; in that case, of ensuring that information relevant to most, if not all, categories of public health threats may be elicited.

FDA-operated Adverse Event Reporting System

The FDA operates AERS, the Adverse Event Reporting System. This system resides on an Oracle database maintained by the agency. Adverse event reporting consists of a mix of voluntary and mandatory reporting. Reporting by manufacturers, health care providers, and patients may be accomplished in a number of ways: by telephone to an FDA hotline; by fax; or by submission of a form through the FDA's website. Manufacturers are required by law to report adverse events. When they do so, names and addresses of patients suffering an adverse event are not reported to the FDA in this initial report, but the reporting manufacturer must maintain raw information used to file the report for at least ten years.

Mandatory reporting of adverse events is governed by sections of the Code of Federal Regulations, including Part 21 sections 310.305, 312.32, 314.80 and 600.80. These regulations cover use of prescription drugs and biological products, investigational new drug (IND) safety reporting and post-marketing reporting of adverse experiences. These regulations require a pharmaceutical maker or distributor to report unexpected adverse events brought to its attention within fifteen days. Although the regulations state such reporting should be accomplished as soon as possible, the absolute deadline is fifteen days. Expected adverse events, meaning events noted during pre-licensing trials, must be reported quarterly; in cases of serious or life-threatening events, the FDA requests that companies file within fifteen days, a guideline which is generally accepted and adhered to. At least one company has stated it files adverse events reports with the FDA, for investigational drugs, within seven business days.

The fifteen-day reporting requirement reduces the timeliness, and thus the potential value of AERS as a contributor to early detection of public health threats. According to FDA staff, the agency has not investigated or documented the average lag time for reporting; one staffer remarked that pharmaceutical makers, while adhering faithfully to the regulations, frequently filed reports very close to the deadline, implying that the average lag between the time of awareness of an event by a company and its report to FDA AERS, may be close to the 15-day limit. Officials at three pharmaceutical manufacturers confirmed this, though interviews with them revealed the average may be 8-12 days. They explained that the reason lay in a need to confirm the veracity of information collected. Moreover, this lag does not include the time between a patient's initial call or presentation for examination, and the mandatory reporter's (i.e., pharmaceutical maker) becoming aware of it. This lag may be minimal, in the case of serious or life-threatening events, or it may be several days or even weeks for less serious events.

Another limitation is the misattribution of a report of illness or ill effect to another cause. Unless a patient's malady is identified as a drug-related adverse event, it will not be reported to AERS.

Physicians, and others may voluntarily report an adverse event to the FDA's AERS via the Medwatch program, but there is no standard here for timeliness of reporting, and inclusion of data on the AERS system.

Pharmaceutical Maker Safety-related Databases

Pharmaceutical makers maintain their own safety-related databases of adverse events, reported by health care providers and consumers. These database systems generate adverse event reports, as appropriate, to the FDA. The databases may be purchased from vendors, or may be adaptations of software developed in house by the manufacturer to record untoward events during clinical trials of investigational drugs. Recent installations all utilize relational database technology. Trained personnel, usually registered nurses, who are familiar with the drug manufacturer's reporting requirements and protocols, accomplish recording of reporting data.

The time lag between receipt of report by a company, and its appearance in a safety-related database, is typically at least 48 hours, because the unit receiving the report must pass it on to other units which evaluate and code it prior to entering it in the database. In the case of serious or life-threatening events, the pharmaceutical maker receives the information very quickly, but query results from the safety database will still be at least 48 hours old. They may be even older, depending on the workload of the reporting unit, and the employees' ability to "turn around" reports quickly. The following hypothetical example, based on actual events in the recent past,

will illustrate how this process may be temporarily slowed: Company A's drug, intended to treat hypertension, is found to have caused a number of fatalities, and a highly publicized recall occurs. Company B markets a competing drug in the same class. Company B's consumer hotline is inundated with calls from people concerned that they are being adversely affected by Company B's drug. Each report, regardless of how improbable the actual connection to the drug may be, must be documented and followed up, resulting in an increased lag time from initial report to inclusion in Company B's database.

However, there is still some potential for safety databases' use in early warning. The company receives reports from its affiliates/programs, which, in turn, have stored reports only 24 hours old. In these cases, it may be possible for public health authorities to tap these databases. In addition, when drug manufacturers run clinical trials involving investigational drugs, tighter monitoring for adverse events occurs, which often enables the manufacturer's database to receive an adverse event report within one day. This makes the database useful for threats involving investigational drugs.
Pharmaceutical makers are very reluctant to allow outside parties, except the FDA under Part 21 of the CFR, to have access to company databases, for fear of revealing proprietary information to competitors; inadvertently releasing information, out of context, which the public may misinterpret; and, of course, violating patients' right to privacy under HIPAA. Thus, the government would have to present a very compelling reason before any manufacturer or distributor would be willing to grant access to this kind of information. The manufacturers would also require the government to pay for any software or hardware development performed.

Information from the safety database may be useful when coupled to the use of prescription-tracking systems. One provides early, inferential warning, the other more direct confirmation. Their complementary use may enhance detection of public health threats.

Prescription Pharmaceutical Sales/Marketing Systems

Pharmaceutical companies routinely outsource the tracking of the sales of their prescription medicines to firms that specialize in health-care or pharmacy-related marketing and sales analysis. These firms, such as IMS Health and National Data Corp., provide extensive sales tracking, market analysis, and customized reports to their clients. Each pharmaceutical maker is shown specific sales results in its markets, data that are not shown to competitors. These market surveillance firms will also provide data to researchers, but not at the same levels of granularity as afforded their primary clients, to ensure compliance with confidentiality agreements.

Pharmaceutical makers are provided with a range of customized reports. They detail many things, including where each drug was prescribed and sold, comparing one time period's sales with another, and comparing one sales region with another. The primary purpose of these services is to facilitate analysis of market performance for each product, and aid the pharmaceutical maker in making needed adjustments to its manufacturing, shipping and marketing efforts. These data do not include diagnostic information or other data elements found in medical records. Their value is likely to be inferential, in that an unusual or unexpected increase in dispensing or shipments of a given drug may raise suspicion of an imminent public health threat.

The data collected by the market surveillance firms arrives both electronically and on paper, with collection frequency usually ranging from weekly to monthly. Thus, raw data regarding a

sales event is received at least one week after the drug has been sold; the resulting report package is sent to the client (pharmaceutical maker) with a lag of as much as four to six weeks. The lag is primarily related to cost, not technological obstacles.

Market-surveillance firms are likely to be very leery of offering public health authorities access to detailed, vendor-specific data. They decline to provide such details to researchers. But, even if a public health agency were to address such reservations, and pay for software modifications with "early warning" functions, the results gained would probably be marginal. The time lag in both data collection and reporting significantly reduces the potential of pharmaceutical sales-tracking systems to aid in the early detection of public health threats. They were not designed for this purpose.

Over-The-Counter (OTC) Pharmaceutical Sales/Marketing Systems

Just as prescription sales tracking is outsourced, so too is Over-The-Counter (OTC) drug sales tracking. Two vendors are dominant is this business. There is one important difference is that, unlike the situation in prescription sales, OTC sales are also tracked by the retailers' supply chains. Retail supply chains provide highly detailed sales tracking in real time. The OTC sales tracking firms obtain data directly from the supply chain systems and integrate it. However, at least one of the firms integrates data weekly, not daily, presenting the same type of latency problem cited for prescription drug sales. It must be stressed that this latency is strictly price-related, and that daily or even hourly data feeds are, from a technical standpoint, very feasible. ***The tracking firms cover a substantial proportion of the market, but not to the same degree as firms involved in prescription drug sales. At least one large retailer, Wal-Mart, does not outsource any sales data tracking. The firm's management jealously guards its data, because they enable Wal-Mart to maintain its negotiating advantage over suppliers. This implies that obtaining data from Wal-Mart for public health surveillance purposes might be difficult.***

Summary

Electronic systems that collect data that could be useful in the detection of public health threats, such as bioterrorism, include on-line prescription verification and assignment systems, over-the-counter sales, and post-marketing surveillance systems.

On-line prescription assignment systems offer the potential of a very rapid "trigger," allowing public health authorities to infer the presence of a threat by examining statistical "spikes" in the sale of drugs. We note that prescription drugs are a relatively late indicator relative to over-the-counter drugs.

On the one hand, the high level of integration in prescription drug sales tracking, and the presence of vendors with 70-80% market coverage, makes data collection technically very feasible; on the other hand, the sale of such data provides nearly all of the vendors' revenue, which poses two challenges to public health authorities. First, public health authorities would be required to negotiate contracts with vendors. Second, the granularity of these data may not be optimal for public health purposes, because vendors would be concerned about breaking lucrative agreements with commercial customers. Even if proprietary constraints were not present, public health officials would have to safeguard the privacy of patients, especially those whose prescriptions are intended to treat mental illness and AIDS make them especially vulnerable to societal discrimination.

Over-The-Counter (OTC) sales data are available from specialty firms; their market penetration is significant, and their data could be highly useful. Because data are obtained directly from retailers' electronic supply chain sources (beginning with the cash register's barcode wand), the information collected is highly reliable. Weekly collection introduces a latency problem, but there are essentially no technical barriers to daily or even real-time collection. Limited public health budgets pose a problem.

Post-marketing surveillance of medicines falls under the jurisdiction of the US Food and Drug Administration (FDA). Intentional contamination of drugs has occurred and therefore it made sense to review how pharmaceutical adverse-event reporting occurs in the United States, to identify and examine related electronic systems, and to ask whether additional systems are required to counter this threat. Upon examination, however, the government's AERS system has worst-case latencies that will impair early detection. The drug manufacturers' safety databases, however, do have some potential in this regard. To use the latter effectively, however, some standardization of internal reporting may be required.

The successful use of all these systems will require cooperation with the pharmaceutical industry, addressing legal objections, as well as some public investment to provide incentives for the industry to provide services.

Questions for Further Study

1. What data specifications would an effective pharmaceutical benefits management system-based public health data feed be required to meet?
2. Who are the key players in the pharmaceutical benefits management industry? What percentage of the population is covered by integrated data systems? Where are the gaps in coverage? Describe the state of the art in military, VA and HMOs.
3. How can public health authorities best persuade commercial data suppliers to cooperate in establishing early-warning networks for public health threats?

Chapter 9. Absenteeism

Absenteeism can be determined directly from absenteeism reporting systems, or indirectly through other measures of a person's physical presence at a location. Both offer, in theory, the potential of inferring the occurrence of a public health threat very early, because people tend to self-treat and stay home from work before they see a physician (if they see a physician at all).

The value of an absenteeism reporting system to a public health authority seeking to achieve early detection and response would depend on how well it samples the population, its availability to the authority, the frequency of its reporting, and the specificity of information contained within it. The threat would have to result in illness severe enough to result in absence from school or work.

Populations for Whom Attendance is Taken

In the United States, employment and education occur in a wide variety of settings; in some, physical attendance is meaningless, and the only measure recorded is the quantity and quality of work submitted. For example, a publishing house editor does not know, or care, of the specific daily whereabouts of an author writing a novel for which he or she has received an advance. Similarly, college class rosters often include students who will appear in class solely at exam time; the college maintains a record of required exams completed or missed, not classroom sessions attended.

One could hypothesize that systems which closely track personnel, and store data electronically, would be more valuable to public health authorities than others. Such systems would include absenteeism reporting systems in the compulsory attendance school grades, employer-run reporting systems, especially in restricted facilities such as defense contractors and nuclear generating facilities, and military bases.

Many organizations do, in fact keep track of attendance electronically, and store records in a manner which makes them available on-line to authorized users. However, the information gathered with regard to the reasons for absenteeism is generally not very specific, and thus the use of these systems as an aid to public-health threat detection must focus on their geographic

and statistical power – their ability to alert users to the presence of statistical trends and anomalies, and the ability to map these events. The presence of restricted or high-security facilities does not appear to influence attendance keeping, overall, but the presence of such tools as "electronic time-cards" may ensure that more up-to-date information is present in an employer's database.

Access to such sources of data by public health authorities would require measures to protect the privacy of individuals whose personal information is stored in personnel or school databases; establishing legal agreements with school districts and employers; and compensating the owners of these data for any costs incurred while creating such access.

Attendance-recording Systems

Public/Private School Attendance

Large school districts use computer technology to track student attendance. Most large districts are computerized, and a large proportion use relational technology, while some, especially small districts with smaller budgets, rely on hierarchical or file technology. The most up-to-date systems provide on-line query capability. These systems could be configured to accept on-line queries from public health authorities. Such actions would require much legal groundwork first, as well as a source of funding. School districts that we interviewed, protecting their students' privacy, will not routinely offer any access to state or federal health authorities without a court order, statutory authority, and a specific contractual relationship. Moreover, they will not provide access to entire databases; rather, they will provide access only to relevant portions. One large school district's CIO indicated that the preference would be for "off-line" access, that is, information provided via a school district staff person on a request basis; on-line periodic "polling" access, if granted, would be considered a full-time commitment requiring full compensation by public authorities for use of resources. . In older systems, using network or hierarchical database technology, or file-based technology, such access becomes more expensive to arrange because the public health authority would be required to install more up to date software, and perhaps hardware as well.

School districts which include primary and/or high schools, typically record daily attendance at a "homeroom" or other pre-set period during the day, a procedure which generates a daily expected roster. Individual teachers will also report absences from specific classes if their head counts do not match the distributed roster for the day. This will generate a record or view containing the student's name and the date absent, the school attended and home address, as well as a given reason. Reasons for absence can include anything from a class or family trip, to an illness, or an unexcused, unauthorized absence. Sometimes, the latter is recorded when the student's guardian has neglected to supply the real reason for the absence. Data entry in such systems is usually straightforward, with teachers or administrators supplying the data on a daily basis.

A parent or guardian typically supplies the reason for absence, and if an illness is involved, can be quite vague. For example, one may discern that the student is ill by examining the attendance record but be entirely unaware of the circumstances around, or cause or nature of, the illness. In schools staffed with nurses, physicians, or athletic trainers, a more specific reason may be known or recorded, if the school's health provider became involved in a given case. This will be the

exception, not the rule. Thus, only part of the data (the individual, school and geography, and the fact that he/she is absent) is reliable; the purported reason is not.

Employer-Run Attendance Reporting

Employer-run attendance reporting at most large employers is computerized, with a wide variety of systems in place. Newer systems use relational databases, facilitating the creation of different views of these data. Moreover, vendors such as Peoplesoft and SAP have created Web-based systems, which can take attendance daily and make lists of absentees available in real-time to authorized users. Whether these systems are actually used this way is up to the employer, not the software vendor. It is estimated that at least three quarters or more of all customers for this type of software are utilizing relatively up-to-date systems, defined as systems deployed within the last three years. Other clients are using older systems, often deployed on mainframes, which may rely on older network or hierarchical database technologies; queries not anticipated on initial deployment are more difficult and costly to add later. Government systems are more likely to be of an older generation than those deployed in the private sector. Some agencies, such as the Department of Defense, are currently upgrading to newer systems.

Even if attendance is recorded daily, there can be substantial latencies before data are available in electronic form. Payroll-based attendance recording typically involves a time card. If this is a paper time card, attendance or absence information will not be entered into the electronic database until the end of the pay period, which could be one week, two weeks, a half-month, or a full month, depending on the employer. Some employers must follow specific recording guidelines promulgated by appropriate government regulations, when these employers hold government contracts. These regulations focus mostly on program-specific attendance, that is, the accurate charging of precisely demarcated blocks of work-time to a given government program. They are not focused on determining when or why an employee is absent from work.

Attendance information stored in employer databases contains the employee's name, position and work location. If the employee is absent, a category of absence is noted, but this may not be specific to illness. Many employers use a category known as "PT," or personal time, which represents an absence on request of the employee for any reason, including a short-term illness. If an employee falls ill, the record will not reflect this. Once the employee presents to either a personal or occupational physician, a medical note can be made.

If an employee is absent for more than one week, he or she may apply for short-term disability. A disability record is created in that case, and this is often stored electronically. Thus, the database containing short-term disability claims can be queried, but all data in this file will be older than one week, limiting its usefulness for early threat detection.

These or similar reporting conditions prevail at most large employers, including banks, defense contractors, power utilities, large movie studios, and medical centers.

Military Attendance/Compliance

The US armed forces monitor absenteeism in a manner similar to civilian employers, except that punishment for unauthorized absence is more severe. The services do not maintain centralized attendance systems; however, personnel rosters are maintained and an unauthorized absence noted and recorded for appropriate response. Until relatively recently this was a "paper and pencil" type of process but has begun to change. However, authorized absences for illness or

other personal reasons are often noted only by department heads; a record of absence for reason of illness is recorded when a member of service presents to a military health facility; those records may be electronic or manual. Personnel records will reflect whether a service member is on duty or on leave.

The reason stems from the kind of information unit commanders require. The following example illuminates this very well: The commander of a US Navy warship needs to know, and be able to report to superiors, the state of readiness of that warship; for this he/she needs to know the state of readiness of all ship's departments. The identities or even (to a point) number of service members absent is a secondary matter, so long as all departments are functioning properly. Each department head, dealing with a relatively small number of people, will no doubt recognize when a full complement is not present and know exactly who is missing and why; the vessel's captain will not, and, in any case, has matters of higher priority to deal with. This is particularly true of large vessels, such as an aircraft carrier with a crew complement of over five thousand. Hence, there is little incentive, currently, to record and keep track of, attendance information in this manner.

Indirect Measures of Absenteeism

We should bear in mind that attendance-keeping systems are not the only ways to determine absenteeism. We can imagine existing data sources that could be used to derive whether an individual is at work, school or home. Such signals might include utilization of various resources at the place of work (or home) especially those that require self-identification. Examples include logging into a computer network, utilization of computer and network resources, and entry into facilities (buildings, parking garages, elevators) that require pass cards. Even phone activity (not content, just time of placement and minutes) may provide a useful indicator. To our knowledge, these approaches have not been tried in public health surveillance.

Summary

Although absenteeism is a relatively early behavior of sick individuals, a relatively small proportion of the population of any region is subject to attendance recording. Moreover, for those individuals for whom absenteeism is recorded, existing practice is such that data are often transferred into electronic form on a periodic basis introducing substantial latencies. Additionally, analysis of absenteeism data must take into account work and school schedules that include weekends, nights, holidays, and vacations.

Attendance-keeping systems are not the only ways to measure absenteeism. We can imagine existing data sources that could be used to derive whether an individual is at work, school or home. Such signals might include utilization of various resources at the place of work (or home) especially those that require self-identification. Logging into a computer network, utilization of computer and network resources, entry into facilities or elevators that require pass cards, parking garages. Even phone activity (not content, just minutes) may provide a useful indicator.

Although absenteeism from work or school is not a good indicator during periods when businesses or schools are closed, we can generalize away from absenteeism to the broader concept of "change in normal activity patterns." Then we can start to consider how to measure indirectly changes in activity patterns in general that may be reflective of a change in health status.

Because of the inherent earliness of absenteeism as an indicator, may be especially useful in selected situations that have the above characteristics. For example, a large corporation that controls access to its computer systems or physical spaces could implement a potentially extremely sensitive system for detecting a building contamination.

Questions for Further Study

1. What percent of the population and of the school age population in a given city are subject to attendance record keeping? To what degree is state-of-the-art database technology utilized?
2. What is the actual time latency of attendance keeping? Can it be reduced?
3. What indirect measures of absenteeism exist? How available are they? How well do they work in detection?
4. What can employers due to provide detection of building contaminations using their own data systems?

Chapter 10. Satellite and Weather

We discuss weather and satellite imaging data together because they represent existing, nationally scoped data sources.

Satellite Imaging

The United States and Russia, the first space-faring nations in the world, have long-recognized the value of observing territory using spacecraft. The original motives were military in nature. In 1961, the US government established the National Reconnaissance Office (NRO) to control reconnaissance satellites. In addition to military "birds," the US also deploys weather-observing satellites, operated the National Atmospheric and Oceanic Administration (NOAA). Still other satellites gather data for use in environmental studies and natural resources, and the data they gather can be quite relevant to a public health mission.

Of the many satellites in earth orbit, a portion is devoted to communication or extra planetary observation. The remainder is engaged in earth observation for military or civilian purposes and their output is relevant to public health surveillance.

In the past, military reconnaissance spacecraft periodically dropped film capsules by parachute into the atmosphere; they were snatched in mid-air by specially equipped C-130 Hercules aircraft. Today, all satellites transmit data in real-time to ground stations. Communications satellites, such as the US military's Milstar series, are used to relay these transmissions to ground antennas.

Occasionally, there may be need for more directed, real-time observation of an event than satellites alone can deliver; or observation with instruments not typically carried by satellites may be desired. In these cases, aircraft can supplement satellite-based imaging.

Satellites and Living Beings

Space-based observation can be of great help in detecting the presence of deceased humans and animals. Satellites that photograph the earth using various portions of the electromagnetic spectrum may accomplish this.

Space-based observation has depended on a system that includes hardware, software, and trained human observers. A reconnaissance satellite photographs a scene of interest, and transmits its images in digital form to an interpretation center on the ground, either directly, or relayed by a communications satellite. The value of the images depends on the type of image; its resolution and the time elapsed between the occurrence of an event and receipt of imagery by a trained interpreter. These factors are, obviously, still important were the human interpreter to be replaced by a computer program.

The data are available in real-time, and reliable. Degree of coverage depends on how many satellites of a particular type are available to cover the US landmass at any given time. Temporary gaps in coverage have been known to occur when malfunctioning satellites, or those reaching the end of their useful lives, have not been replaced in time.

Image Technology

Photographs of use in detecting stricken people or animals are most likely taken using optical cameras, infrared sensors, and perhaps radar antennas using synthetic-aperture techniques, assuming a high enough resolution is achieved.

The resolution of photography varies among platforms. The best non-military satellites offer imagery to customers with a resolution of one meter per pixel (Ikonos; four meters per pixel for multispectral images). Older spacecraft offer 10 meters per pixel (Spot) to 30 meters per pixel (Landsat). The resolution of military reconnaissance satellites is classified, but civilian experts, such as those at the Federation of American Scientists, have stated that 10 centimeters per pixel would not be unrealistic. In addition, these satellites' sensors can also take pictures in a night sky.

Satellites are equipped with maneuvering thrusters allowing their controllers to adjust their orbits and quickly bring their cameras to bear on scenes of interest. Should an image taken provoke curiosity or concern, the satellite can be maneuvered to pay special attention to the location where the image was taken.

Images retrieved from satellites are usually processed by sophisticated software that enhances the images.

Image Recognition Technology

Image recognition technology is relatively young. Automated recognition is being pursued with a number of techniques, such as pattern matching, use of "primitives," and others. The military has demonstrated its utility; in 1991, cruise missiles fired from aircraft and naval vessels employed image recognition technology during the final homing phase of their flight to strike their targets. One key to their success was the relative simplicity of the images their seeker heads sought, for example, a building window. Current weapons utilize improved software; the exact capability of that software is classified.

This simplicity can, potentially, be found in public health-related tasks as well. For example, a dead cow assumes a profile quite distinct from a sleeping cow, or one that is contentedly munching on alfalfa. If a computer program can be designed to recognize the first case, the latter

two cases should provide more than enough variance to convince the program to reject them, given sufficient image resolution. A large group of dead cows would be spotted in the same way.

Humans could present a more difficult problem. Our bodies are smaller than cattle, and discrimination between death and sleep, or even death and any stationary pose, more challenging. However, significant changes in the condition or activities of groups of humans are readily documented by satellite data; fully automated detection of desired signals, however, may require technologies that, if they exist, are classified.

It is clear, however, that if satellites offering very high-resolution photography, coupled to appropriate enhancement and interpretive or detection software, were to be available to public health authorities, they could add to the US ability to detect bioterrorism.

Limitations

Satellites are affected by weather. Heavy cloud cover hides scenes of interest from cameras. If bioterrorists understand that satellites are being used for surveillance, they are likely to attempt to strike when weather conditions are most favor them, or to try as much as possible to conceal their attacks for as long as possible.

In addition to the technical limitations related to human bodies discussed above, other issues that would confront public health authorities are ethical (especially those related to privacy), and financial (spacecraft observation time must be paid for). Indeed, it is likely that certain advocacy groups may protest the utilization of high-resolution satellite imagery for observation of people.

Aircraft Imaging

Aircraft are currently available that are equipped with very high-resolution cameras, thermal and infrared sensors. Thermal and infrared sensors are, in particular, very useful for distinguishing between live and dead animals, because dead bodies do not generate heat. As time passes, the contrast between a live body and a dead one increases.

Unlike satellites, which have mission durations measured in years, aircraft must be launched and recovered, and the cost of continuous over flights for public health missions would be prohibitive. The successful deployment of pilot less, remote-controlled reconnaissance platforms by the US military, such as the Predator aircraft used most recently in Afghanistan, promises to considerably reduce the cost of aerial surveillance. Other aerial tools with some potential include tethered aerostats; whether any of these options will prove to be affordable to public health authorities is an open question.

Even if they are expensive to operate, aircraft may be useful, and cost-effective, in ancillary or follow-up roles, gathering data in specific situations where an observational orbit around a site first identified by satellite is desired. The images gathered could then be processed by the same tools that processed the satellite's data.

Weather

Weather and climate data currently used in epidemiological analysis include temperature, wind direction and speed (for bioaerosol related analyses), and precipitation. Of possible value, but not used according to our experts, might be barometric pressure and ultraviolet exposure.

In the United States, weather data are already highly available. The specific data available include temperature, wind speed, wind direction, and precipitation. Up-to-the-minute information for the entire nation is available because data are collected in real-time, in standard formats, and integrated in a central location that is publicly available without any technical or administrative barriers (http://weather.noaa.gov/). We note that the weather system--in addition to a mature source of data for public health surveillance and early warning of bioterrorism--represents a good case study in how to integrate data from many sources and independent monitoring stations using communication networks and standards to achieve a national surveillance capability.

Satellite Weather Imaging

The satellite weather systems most interesting and relevant to public health authorities are those that measure water temperature and ecological conditions such as floods, fires, and conditions of vegetation have been used for epidemic prediction. These ecological factors are often predictive of changes in intermediate host and vector activity that are, in turn, predictive of outbreaks of human disease.

In the United States, satellite data are already highly available. As with reconnaissance satellites, data for the entire US are usually constantly available, so long as enough satellites are in orbit. There have been occasions when the US has not launched a satellite in time to replace a failing "bird," resulting in temporary gaps in coverage. The raw data is of different types ranging from numeric to image. The ability to extract useful information for public health purposes, from images, has not been highly developed. An exception is the GEIS Rift Valley Information System [42]. We have yet to determine other factors that determine availability and utility such as the extent to which they are available in standard formats, and integrated in a central location that is publicly available without any barriers.

Aircraft and Weather Observation

Just as aircraft can be useful for reconnaissance purposes, so can they serve a supporting role in monitoring the weather. The US Air Force, for example, has for years dispatched C-130 Hercules aircraft to fly into the eyes of hurricanes to gather data. These flights have made invaluable contributions to our understanding of how storms develop and evolve. The flights are expensive to operate, however, and high cost is always a concern for public health authorities.

Summary

Up-to-the-minute weather information for the entire nation is available because data are collected in real-time, in standard formats, and integrated in a central location that is publicly available without any technical or administrative barriers.

Space-based platforms can provide a great deal of valuable data in real-time to public health authorities concerned with detecting bioterrorism. These data are collected by very reliable and proven platforms with which the US has decades of operating experience. Images and other data collected concerning animals, people and weather are available for the entire United States. These data may be supplemented by data collected by aircraft.

The barriers facing access for public health authorities include limited budgets, concerns about privacy, and access to technology, either classified or open, enabling automatic detection of a bioterrorism incident. In addition, public health officials will need agreements with the agencies

or corporations operating spacecraft or aircraft in order to be assigned observation time on these systems.

Questions for Further Study

1. What are the advantages and disadvantages of optical, infrared, and radar-based sensors in public-health settings?
2. What weather conditions are ideal for a bioterrorist attack? How do these conditions vary by pathogen type?
3. What administrative and organizational arrangements are needed to provide public health authorities with adequate access to satellites and aircraft?

Chapter 11. Claims Data and Billing

Claims, which are submissions used to pay for health services provided to patients, are potentially one of the most available sources of data available for detecting selected public health threats. Claims include a large amount of information, including the patient's age and gender, geographic location, ethnicity, preliminary and revised diagnoses, and orders for laboratory, radiology and special procedures, as well as results. Claims submission has become a highly refined process over the years; insurers have started investing heavily in computerized database applications because they have come to recognize the potential benefits of statistical analysis in reducing costs. This investment has resulted in the creation of large data warehouses, equipped with relational database technology to facilitate "data mining, which includes the creation of different views of the data, large scale "number crunching," trend discovery, analysis and prediction. Because the data warehouse can readily accommodate users with different query requests, it can serve public health authorities; it offers highly specific data about patients and illnesses. Its value, as part of an early warning system, is reduced, however, by the age of the data it contains, and that will place limitations on its inclusion in an early-warning scheme.

Characteristics of Claims Data

Availability

There is a great deal of heterogeneity in the American health insurance industry, but claims warehouses to a large extent hide that complexity. Employers offering health-insurance benefits to employees may do so either by arranging for group or individual insurance policies from insurance companies, or by "self-insuring." Employers pursuing the latter strategy, typically very large organizations with thousands of employees, typically hire an insurance carrier to administer the plan. Senior citizens are covered by Medicare, and may purchase additional insurance to complete benefit coverage. There are many different models of insurance, such as fee-for-service, HMO, PPO, etc.

The common threat is that patients generate claims whenever they present for care. The claims are submitted for payment either manually (on paper) or electronically. The claims ultimately enter a transactional system and accumulate there, awaiting transfer to a central data warehouse.

These transfers may be accomplished entirely electronically, or via reels of magnetic tape, or other media, which are shipped to a data center. Claims data are comprehensive. They include demographic information, (name, gender, age, ethnicity, address of residence and dependents), medical information (the presenting complaint, physician's notes and diagnostic codes), laboratory, pathology and radiology tests and results, and other data. Claims data warehouses also represent the most inexpensive, though not the only, portal into claims data on a national scale.

Timeliness

The usefulness of claims data to detect public-health threats is greatly limited, however, by the age of these data. Health-care providers file claims through electronic and manual means, and one cannot assume that the day a claim is filed is the day a patient visit occurred. In addition, the time lag acknowledged by industry representatives to entry of data in the warehouse can be as long as 30 days from the time of primary filing of a claim into the transactional system. The delay can be even longer, in actual practice, because the efficiency of claim filing depends on the diligence of a health care provider's (physician, hospital etc.) billing procedures and staff. Delays of several months, due to inattention by a hospital's billing department, have been known to occur.

Insurers' main incentive to streamlining claims processing is reduction in costs. This incentive does not translate directly to a shorter lag time between a patient's visit and a record's availability in the warehouse. Most of the data warehouse' clients are interested in outcomes research, trend analysis, cost-effectiveness studies, and the like, so timeliness is not a requirement for them.

The data warehouses deployed in the US typically receive a huge volume of data monthly, from transactional systems. In 1998, the Health Care Financing Administration reported that over 850 million Medicare claims were processed at a cost of over $650 million dollars. One large insurer's data warehouse holds several terabytes (trillions of bytes) of information; its associated transactional systems process over 450,000 claims daily. These transactional systems are often tiered; claims are submitted daily to a transactional system, and then passed along to a weekly transactional system. This system then provides the data to the warehouse. Each warehouse may receive data from dozens of transactional systems; these systems are owned and operated by different entities, which may or may not be related to the owner of the claims data. Transactional systems may or may not utilize modern database technology; some are relational, and some still rely on network or hierarchical systems. The relational systems may not be fully compatible with each other; the mere use of Structured Query Language (SQL) does not guarantee easy interfacing between systems.

In theory, a query against a data warehouse could retrieve up-to-date information, if polling functions were available, and applied to all the feeding systems, and if the volume of transactions was tractable. Banks use such functions to allow customers to withdraw money from automatic teller machines (ATMs). When a customer requests a withdrawal, the ATM sends a transaction request to the bank's central database. The database verifies the available balance in the customer's account; it then "polls"(a process whereby a central computer queries its associated peripheral systems to produce a more updated view of data than it holds itself) other computer systems which record current transactions from ATMs or point-of-purchase locations, such as retail stores, in real time to determine if any recent transactions affecting the true account balance

have been memo-posted but not yet uploaded to the main database. Data warehouses holding medical claim information do not utilize "polling" systems, because their users are not interested in such functions. Their priorities derive from a need to reduce the costs of providing care, and are focused on such tasks as chronic disease management, physician profiling, and "best practices" development. Timeliness of reporting at the warehouse level is not important to these efforts.

Technical Barriers

A number of technical options are available to create access for public health authorities to more recent, claims data, each of which offers advantages and disadvantages. In order to insure that these data are truly not obsolete, however, each option's implementation would have to be accompanied by improvements in claims filing. As organizations progressively adopt electronic filing, the average lag between a health care service and its claim filing can be expected to shrink, but much depends on the providers.

Before agreeing to offer public health authorities access to the data warehouse, the operator would demand an agreement protecting privacy and confidentiality, and relieving the warehouse operator of any liability resulting from inappropriate release or dissemination of information from the warehouse. Further, permitting the implementation of any option, or combination of options, the data warehouse's operator would require the government to pay all costs associated with such development work, including additional costs incurred while serving existing customers. Even accounting for the availability of new technology which aids in the movement of information across systems representing different storage technologies and data layouts, the cost of providing access to one data warehouse, with its associated transactional systems, could easily reach millions of dollars; this does not include the day-to-day cost of operating the resulting electronic public health surveillance system, also to be borne entirely by the government. A data warehouse's operator earns revenues by selling information to clients; a public health authority operating a public health threat detection system would represent another paying client.

One option would involve arranging for the transactional systems to upload data to the warehouse daily, instead of monthly. The advantages of this approach include the utilization of existing software and procedures, and exploiting the robustness of the warehouse itself. Data integrity is assured; with up to date data present, the warehouse, in concert with its associated transactional systems, may now function within an early warning system. Its sensitivity has been increased with no loss of specificity. The disadvantages of this approach include a tremendous increase in transaction volume arriving at the warehouse, an increase in demand on personnel, and possibly an increase in data storage requirements. Moreover, this increase in data collection includes much information, such as billing rates, which are not relevant to the public health mission. The warehouse' operators will be reluctant to undertake without significant compensation. Hence, this option is expensive.

A second option involves the public health authority creating its own data warehouse to receive uploads from the transactional systems themselves. The advantages of this option include the creation of a warehouse completely under the control of public health personnel and optimized for public health uses; as well as the use of relatively straightforward upload procedures that are easy to implement. Another advantage is that only the data portion of interest to public health is uploaded. Disadvantages include the need to compensate each transactional system vendor or

operator for the cost of extracting and formatting just the desired data in a form appropriate for uploading, as well as the cost of performing daily uploads to the public health warehouse in addition to the normal uploads already performed; and the need to keep track of individual transactional systems as they are replaced or upgraded. Sensitivity for detection by public health authorities is improved, again at a significant cost.

A third option involves the creation of a public health warehouse and data collection system aimed directly at providers. The advantage here includes the creation and use of standard, replicable, public-health optimized interfaces; elimination of the need to burden either transactional systems or the commercial data warehouse with public health-related functions, and thus a lower cost of operation. An important disadvantage is the imposition of additional reporting requirements on health care workers. Doctors, nurses and others would be required to enter data twice for each patient they see. When a situation-specific "syndromic surveillance" system requiring duplicate entry of this type was set up in Philadelphia at a large convention, many complaints were noted from emergency room providers who felt they were already overburdened. This could lead to incomplete reporting, which would adversely affect the integrity of data being submitted. This would reduce the intended sensitivity of such a system. This is the least expensive, financially, of the options listed, but clearly it incurs costs of a different type.

There are other options as well. Regardless of the option chosen, a public health authority seeking to gain access to such information would have to conclude legal agreements with the owners of claims data, the vendors of all software systems involved in the claims process, and their clients. One pilot project, begun as a cooperative effort between the state of Minnesota and a large HMO, has involved a few percent of the population of Minneapolis. It utilizes claims submitted nightly by physicians, and uploaded to a public health database. Thus, the data arrives with a very short latency. However, tools with which public health authorities can conduct automated signal detection on these data are still few in number and experimental.

Certain claims data originate from other information systems, such as laboratory data systems, digital or teleradiology systems, and, where deployed, point-of-care systems. Optimizing early detection of a threat intuitively points to acquisition of data contained in these "point of care" systems, where less time has elapsed between the onset of illness and data generation, compared to claims records.

Claims data warehouses were designed and deployed to serve purposes much different than early detection of public health threats. While modifications are possible, the required investment may still not produce the necessary improvements in timeliness. Thus, it may be that public health authorities should look to other data systems for elements of an early warning system. The deployment of electronic medical record systems, including clinical notes, laboratory tests and radiology reports offer public health authorities a source of current information, which can be used in an early-warning system. However, as the 1990's ended only a quarter of all US hospitals used such systems, and many did not do so comprehensively, i.e., throughout all services; even fewer of all physician's offices did so. POC systems were reviewed in Chapter 4.

Summary

A claims data warehouse is a source of comprehensive health information about people, readily retrievable for analysis. The huge numbers of claims processed and stored by data warehouses,

and their logical organization, make possible very sophisticated trend detection and analysis. Unfortunately, its optimization for look-back and analysis gives its data latency unacceptable for use in public health early-warning systems. Transactional systems that feed the data warehouses are characterized by a relative absence of commonality in hardware and software, and by latency inherent in systems not requiring providers to submit claims the same day patients are seen. It is possible to develop an effective public health surveillance system using claims data. A number of different strategies have been outlined, each balancing sensitivity and specificity of threat detection differently. Regardless of the approach used, developers would have to address significant financial, legal and ethical challenges.

Questions for Further Study

1. What is the most cost-effective way to generate and extract claims data in a manner useful to early warning for bioterrorism and other public health threats?

Chapter 12. Physiological Monitoring

For purposes of this chapter, we shall consider physiological monitoring systems to be those which help assess aspects of human performance and either respond with a corrective action, or alert a human being to do so. The kinds of data physiologic monitors can measure and record includes heart rate, temperature, respiration, and alertness and activity. We will consider the potential of these systems to assist in the early detection of bioterrorism by their equipping a "sentinel population." A sentinel population is any group of people who are willing to participate in a monitoring scheme. Security professionals, for example, could act as the proverbial canary in the mineshaft in a variety of settings. Examples of this approach could include advance members of the President's Secret Service detail, or other executive protection forces, or security agents patrolling special events (e.g., the Republican National Convention). In the latter case, we imagine that devices worn unobtrusively by the sentinels would provide early warning of spiking fevers, tachycardia or other signs arising in a modest fraction of the sentinel population. The information transmitted automatically from these sentinels would be timely.

To offer this capability, the equipment worn must be able to carry out accurate measurements while withstanding, and compensating for, the activities of an active subject. For example, a thermometer worn on the body must be able to compensate for clothing, weather, and exercise. Drawbacks to this approach include having to deal with the same limitations accompanying other early signs of disease – a lack of specificity.

Although most people would, no doubt, associate physiological monitoring with hospitals, physiology laboratories and other specialized settings, some elementary forms of physiological monitoring are practiced, routinely, in everyday life; Control cabs in railroad locomotives and commuter/subway railcars are equipped with the so-called "Dead Man's Handle." If the engineer does not touch this device periodically, it automatically deploys the train's emergency brake, on the presumption that the engineer is disabled for some reason.

Aerospace and Remote Terrestrial Applications

Aerospace

Remote physiological monitoring, in one form or another, has a long, established history in aviation and space applications.

Gear intended to fit under, or on, clothing, to measure various parameters of bodily function and environmental conditions has existed since the late 1940's. Beginning soon after World War II, the US Air Force, the National Advisory Committee on Aeronautics (NACA), and its successor, the National Aeronautics and Space Administration (NASA), built a series of rocket-powered experimental aircraft to explore flight at very high speeds and altitudes. In 1947, Charles Yeager flew the X-1 past Mach 1; in 1956, the X-2 flew at nearly Mach 3. Two years later an X-2 test pilot set a world altitude record of nearly 126,000 feet. By 1967 a pilot flying North American Aviation's X-15 had reached Mach 6.7, soaring 60 miles above the Earth's surface. Test pilots flying in the late 1950's and after needed special suits, helmets and oxygen supply systems to stay alive, and scientists needed to collect data on human performance in their cockpits. Physiologic data was not collected consistently throughout the test flight programs, due to resistance by test pilots to becoming "guinea pigs." When collected, this information was generally recorded on the aircraft, and collected for study after landing.

Collecting data on human bodily function was also a vital part of research in the space program. In 1958, the Mercury program ushered in the first group of physicians specifically assigned to care for astronauts. NASA decided that monitoring astronaut physiology in real-time was of critical importance, and so physiologic data collected by sensors was included in spacecraft telemetry from the beginning of the space program. These data included heart rate and rhythm, respiration and core and skin temperature. In the 1960's, Hamilton-Standard, now part of United Technologies Corp, developed a series of space suits designed to accept NASA-developed biometric-measuring gear. These suits were used in the Apollo, Skylab, and Space Shuttle programs. Data from these systems was recorded during flight and added to other telemetry sent to the ground. To these systems the Russians added (on the space station, not the suits), a blood analyzer, built by Daimler-Benz in Germany, which automatically sent blood chemistry results to ground control from the Mir space station.

Currently, the only biosensor worn by astronauts is an EKG probe. It should be noted, however, that, given preflight respiratory function test data, respiratory rate might be inferred from the amplitude of the R wave transmitted by the EKG probe. (Sources: NASA; United Technologies Corp/Hamilton Sundstrand.)

Remote Terrestrial

Medical researchers have successfully demonstrated, on a small scale, the potential of physiological monitoring technology to aid in the remote monitoring of personnel functioning in extremely inhospitable environments. Examples of this include trials of biosensors designed and built by FitSense Technologies.

Working at the NASA-funded Commercial Space Center of Yale University's Department of Surgery, researchers led by Dr. Richard Satava organized climbing expeditions on Mount Everest, during the climbing seasons of 1998-99, in order to demonstrate the reliability of FitSense biometrics gear and its value to the health maintenance of people in remote

environments. Three climbers wore sensors measuring heart rate, skin temperature, core body temperature, and activity level, as well as a GPS receiver to determine location. Data from these devices was transmitted to the Everest Base Camp, and relayed to Yale University. The outcome measures were correlated via time-stamp identification. Sensor availability ranged from 78% to 100%. The researchers concluded that this application of biosensor technology was feasible, but that improvements in reliability and robustness were needed.

FitSense also equipped 16 Marines at Quantico, VA, and nine US Army Rangers at Fort Benning, GA with biosensors during 10-day war fighting exercises in 1998, 1999 and 2000. The sensors were taped on or otherwise worn by soldiers wearing standard-issue Battle Dress Uniforms (BDU). The Rangers' biosensors transmitted data in real-time to a command post; the Marine's sensor data was recorded by a device worn by each soldier. The company reported 100% availability of sensors, with reporting accuracy of 97%. [SOURCE: INTERVIEW W/T. BLACKADAR, CEO FITSENSE] When interviewed, the company's CEO noted that pulse oximeters proved problematic at first, and the company expended a good deal of effort to improve them.

As shown by these trials, these types of sensors are not intended exclusively for the military. Law enforcement, medical surveillance, search and rescue, and public health applications offer possible uses. The idea common to all these applications is to enable a control center to determine if people under its control are injured or otherwise disabled, as well as how severely. These include a military command post watching over soldiers embarking on a reconnaissance mission, a police unit commander watching over officers who are dispersed over a wide area, or a fire department battalion chief keeping a close watch on fire fighters advancing into a burning building. A photo showing the transceivers appears in Figure 12.1.

One drawback of this equipment is its expense. It costs nearly $3,500.00 to outfit a soldier with FitSense Technologies' equipment.

Figure 12.1 Courtesy FitSense Technologies

Cardiology and Pulmonology

There are several devices currently available in the market, which will monitor heart function and transmit data in real time to a monitoring station. These include both externally worn and implantable monitors. One-to-three lead EKG monitors no larger than a music tape cassette can monitor heart rhythm and transmit information by landline, cellular or satellite phone. Small, light, battery-powered Holter monitors are also available. Implantable cardiac pacemakers employ low-power sensors to measure heart rhythm and deliver a counter shock when a dysthrhythmia is noted. These devices have been limited to employing 8-bit microprocessors, in order to minimize heat generation during operation within the body. Conceivably these devices could be adapted to store and forward physiological data for public health purposes [Source: Schiller website].

Pulse oximeters measure the degree of blood oxygenation non-invasively. Models currently available weigh a few ounces, are battery-powered, and feature digital signal processing. Adding a transmitter to such devices is technically quite feasible. [Sources include Nellcor's website]

Mobile Clinical Trial/Health and Exercise Monitoring

Commercial vendors sell a variety of wearable devices that are intended for monitoring human performance in clinical trial settings or exercise settings. Wrist and elbow devices manufactured by Pittsburgh-based BodyMedia, for example, measures movement, heat flow, skin temperature, ambient temperature, and galvanic skin response. A transceiver mounted on each unit allows the instrument to both transmit data in real-time to a workstation, and to receive data from another, third-party device.

Limitations of Physiological Monitoring

Biometric gear offers great potential, but is expensive and evaluated only with small groups of volunteers thus far. Government agencies have yet to adopt these systems as standard issue.

Technology

The development of electronics has followed Moore's Law (first described by Gordon Moore, cofounder of Intel Corp.) in that the number of transistors placed on a computer chip has roughly doubled every one to two years, causing a corresponding increase in computer processor speed and capability. As electronics become smaller, more sophisticated, and cheaper, proposals to create and outfit sentinel populations will become more realistic. As more advanced low-voltage power sources are developed, devices intended for implantation will do more, and perform more accurately. Applications include local and national security, war fighting, automated treatment of diseases such as diabetes, and early detection of bioterrorism through analysis of the bodily functions in a sentinel population.

Bioethics

The widespread use of sentinel populations, wearing physiological monitors, will require the development of ethical principles governing their use. If a sentinel's monitor alerts a control center to changes in physiological functioning, under what conditions should the sentinel be withdrawn, and under what conditions should he/she remain in place? How will such policies balance, on the one hand, maximizing the sensitivity and specificity of detection, and on the other, the sentinel's safety?

Summary

Professionals who work to protect others from harm, and, by the nature of their jobs, expose themselves to higher risk than members of the public would accept, are ideal candidates for sentinel populations. The effect of biological agents on these individuals includes changes in physiological functions. These changes can be measured by monitoring equipment. The space program has long used physiologic monitors. Advancements in functional sophistication and miniaturization have culminated in their experimental deployment, on a very small scale, during combat training.

The use of monitoring gear in these settings has resulted in many technological improvements and demonstrations of increasing reliability. While the results have been encouraging, it is also clear that physiologic monitors are still very expensive systems which cannot be widely deployed in the military, much less a chronically cash-starved public health system. Therefore, the single most important development, which would allow the use of physiologic monitoring gear, is affordable technology.

The widespread use of sentinel populations, wearing physiological monitors, will require the development of ethical principles governing their use. The sentinel's safety must be considered as we seek to maximize the sensitivity and specificity of the detection policy we impose on him or her.

Questions for Further Study

1. There are good physiological measurements available from physician's offices and hospitalized patients. Can the annual influenza outbreak be inferred from those individuals?
2. If a population of ambulatory patients were equipped with such devices, would it be possible to detect outbreaks early?
3. What ethical issues must be addressed before the federal government can request or require individuals to wear such gear? How will these issues affect policies and procedures? What did NASA learn in the course of the space program, and are those lessons transferable here?

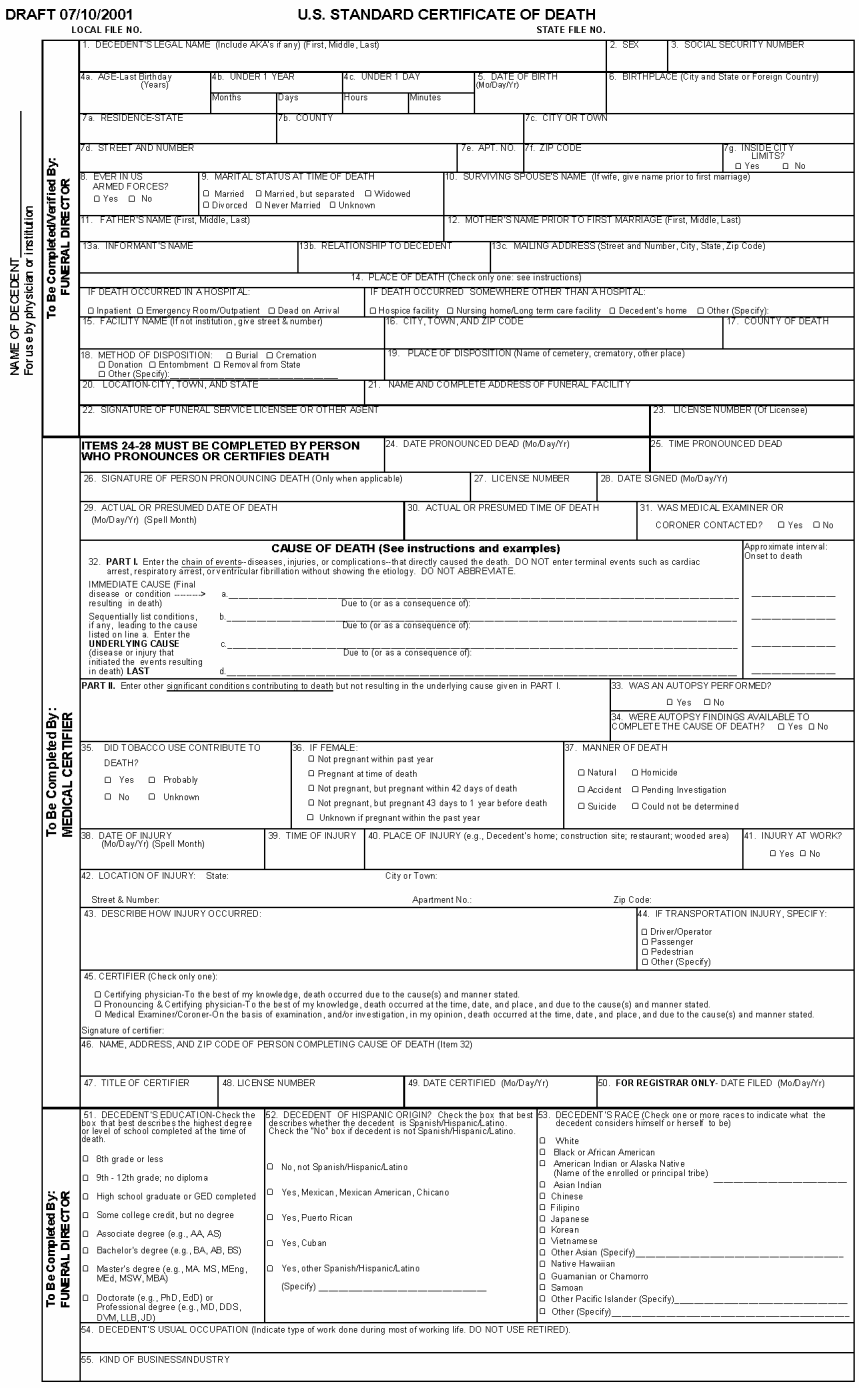
Chapter 13. Vital Statistics

Vital statistics are generated from records of births, adoptions, marriages, divorces, and deaths (collectively referred to as *vital records*). This section will deal only with birth and death certificates—the vital records of relevance to public health and the only two certificates required for every American. Although vital records are issued by states, uniformity across states is encouraged through specifications developed by workgroups sponsored by the Division of Vital Statistics, Department of Health and Human Services. Although the use of these specifications is voluntary, and implementations vary, there exists a fair degree of uniformity among the states.

State governments issue death certificates and their purpose is to certify death. They are needed to bury an individual, to probate a will, to sell property of the deceased, and for a variety of other civil functions. Figure 13.1is the draft 2003 revisions for death certificates.

Increases in disease-specific or overall death rates are used by public health to detect the existence of an outbreak. For example, selected cities report total deaths and deaths from pneumonia and influenza to the CDC weekly as a means of tracking outbreaks of influenza.[35] In death certificates, the cause of death, age and place of death may be helpful in characterizing an outbreak and may give clues as to the identity of the causative agent.

Birth certificates have potential utility for infectious agents that may cause premature delivery (*Neisseria gonorrhoeae, Treponema pallidum, Herpes simplex*, group b streptococcus etc.) or birth defects (rubella, cytomegalovirus, *Toxoplasma gondii*). Birth certificates, however, are only issued for live births; thus, birth certificates do not have potential to detect increases in fetal deaths or miscarriages. Figure 13.2 is a copy of the draft recommendations for birth certificates fields beginning in 2003. The majority of the fields in the proposed certificate are already present in current birth certificates.



DRAFT 07/10/2001

U.S. STANDARD CERTIFICATE OF DEATH

LOCAL FILE NO. STATE FILE NO.

NAME OF DECEDENT
For use by physician or institution

To Be Completed/Verified By: FUNERAL DIRECTOR

1. DECEDENT'S LEGAL NAME (Include AKA's if any) (First, Middle, Last) | 2. SEX | 3. SOCIAL SECURITY NUMBER

4a. AGE-Last Birthday (Years) | 4b. UNDER 1 YEAR: Months, Days | 4c. UNDER 1 DAY: Hours, Minutes | 5. DATE OF BIRTH (Mo/Day/Yr) | 6. BIRTHPLACE (City and State or Foreign Country)

7a. RESIDENCE-STATE | 7b. COUNTY | 7c. CITY OR TOWN

7d. STREET AND NUMBER | 7e. APT. NO. | 7f. ZIP CODE | 7g. INSIDE CITY LIMITS? ☐ Yes ☐ No

8. EVER IN US ARMED FORCES? ☐ Yes ☐ No | 9. MARITAL STATUS AT TIME OF DEATH: ☐ Married ☐ Married, but separated ☐ Widowed ☐ Divorced ☐ Never Married ☐ Unknown | 10. SURVIVING SPOUSE'S NAME (If wife, give name prior to first marriage)

11. FATHER'S NAME (First, Middle, Last) | 12. MOTHER'S NAME PRIOR TO FIRST MARRIAGE (First, Middle, Last)

13a. INFORMANT'S NAME | 13b. RELATIONSHIP TO DECEDENT | 13c. MAILING ADDRESS (Street and Number, City, State, Zip Code)

14. PLACE OF DEATH (Check only one: see instructions)

IF DEATH OCCURRED IN A HOSPITAL: ☐ Inpatient ☐ Emergency Room/Outpatient ☐ Dead on Arrival

IF DEATH OCCURRED SOMEWHERE OTHER THAN A HOSPITAL: ☐ Hospice facility ☐ Nursing home/Long term care facility ☐ Decedent's home ☐ Other (Specify)

15. FACILITY NAME (If not institution, give street & number) | 16. CITY, TOWN, AND ZIP CODE | 17. COUNTY OF DEATH

18. METHOD OF DISPOSITION: ☐ Burial ☐ Cremation ☐ Donation ☐ Entombment ☐ Removal from State ☐ Other (Specify): _____ | 19. PLACE OF DISPOSITION (Name of cemetery, crematory, other place)

20. LOCATION-CITY, TOWN, AND STATE | 21. NAME AND COMPLETE ADDRESS OF FUNERAL FACILITY

22. SIGNATURE OF FUNERAL SERVICE LICENSEE OR OTHER AGENT | 23. LICENSE NUMBER (Of Licensee)

To Be Completed By: MEDICAL CERTIFIER

ITEMS 24-28 MUST BE COMPLETED BY PERSON WHO PRONOUNCES OR CERTIFIES DEATH

24. DATE PRONOUNCED DEAD (Mo/Day/Yr) | 25. TIME PRONOUNCED DEAD

26. SIGNATURE OF PERSON PRONOUNCING DEATH (Only when applicable) | 27. LICENSE NUMBER | 28. DATE SIGNED (Mo/Day/Yr)

29. ACTUAL OR PRESUMED DATE OF DEATH (Mo/Day/Yr) (Spell Month) | 30. ACTUAL OR PRESUMED TIME OF DEATH | 31. WAS MEDICAL EXAMINER OR CORONER CONTACTED? ☐ Yes ☐ No

CAUSE OF DEATH (See instructions and examples)

Approximate interval: Onset to death

32. **PART I.** Enter the chain of events--diseases, injuries, or complications--that directly caused the death. DO NOT enter terminal events such as cardiac arrest, respiratory arrest, or ventricular fibrillation without showing the etiology. DO NOT ABBREVIATE.

IMMEDIATE CAUSE (Final disease or condition --------> resulting in death) a. _____ Due to (or as a consequence of):

Sequentially list conditions, if any, leading to the cause listed on line a. Enter the **UNDERLYING CAUSE** (disease or injury that initiated the events resulting in death) **LAST** b. _____ Due to (or as a consequence of):

c. _____ Due to (or as a consequence of):

d. _____

PART II. Enter other significant conditions contributing to death but not resulting in the underlying cause given in PART I.

33. WAS AN AUTOPSY PERFORMED? ☐ Yes ☐ No

34. WERE AUTOPSY FINDINGS AVAILABLE TO COMPLETE THE CAUSE OF DEATH? ☐ Yes ☐ No

35. DID TOBACCO USE CONTRIBUTE TO DEATH? ☐ Yes ☐ Probably ☐ No ☐ Unknown

36. IF FEMALE: ☐ Not pregnant within past year ☐ Pregnant at time of death ☐ Not pregnant, but pregnant within 42 days of death ☐ Not pregnant, but pregnant 43 days to 1 year before death ☐ Unknown if pregnant within the past year

37. MANNER OF DEATH ☐ Natural ☐ Homicide ☐ Accident ☐ Pending Investigation ☐ Suicide ☐ Could not be determined

38. DATE OF INJURY (Mo/Day/Yr) (Spell Month) | 39. TIME OF INJURY | 40. PLACE OF INJURY (e.g., Decedent's home; construction site; restaurant; wooded area) | 41. INJURY AT WORK? ☐ Yes ☐ No

42. LOCATION OF INJURY: State: City or Town: Street & Number: Apartment No.: Zip Code:

43. DESCRIBE HOW INJURY OCCURRED: | 44. IF TRANSPORTATION INJURY, SPECIFY: ☐ Driver/Operator ☐ Passenger ☐ Pedestrian ☐ Other (Specify)

45. CERTIFIER (Check only one):

☐ Certifying physician-To the best of my knowledge, death occurred due to the cause(s) and manner stated.
☐ Pronouncing & Certifying physician-To the best of my knowledge, death occurred at the time, date, and place, and due to the cause(s) and manner stated.
☐ Medical Examiner/Coroner-On the basis of examination, and/or investigation, in my opinion, death occurred at the time, date, and place, and due to the cause(s) and manner stated.

Signature of certifier:

46. NAME, ADDRESS, AND ZIP CODE OF PERSON COMPLETING CAUSE OF DEATH (Item 32)

47. TITLE OF CERTIFIER | 48. LICENSE NUMBER | 49. DATE CERTIFIED (Mo/Day/Yr) | 50. **FOR REGISTRAR ONLY**- DATE FILED (Mo/Day/Yr)

To Be Completed By: FUNERAL DIRECTOR

51. DECEDENT'S EDUCATION-Check the box that best describes the highest degree or level of school completed at the time of death.
☐ 8th grade or less
☐ 9th - 12th grade; no diploma
☐ High school graduate or GED completed
☐ Some college credit, but no degree
☐ Associate degree (e.g., AA, AS)
☐ Bachelor's degree (e.g., BA, AB, BS)
☐ Master's degree (e.g., MA, MS, MEng, MEd, MSW, MBA)
☐ Doctorate (e.g., PhD, EdD) or Professional degree (e.g., MD, DDS, DVM, LLB, JD)

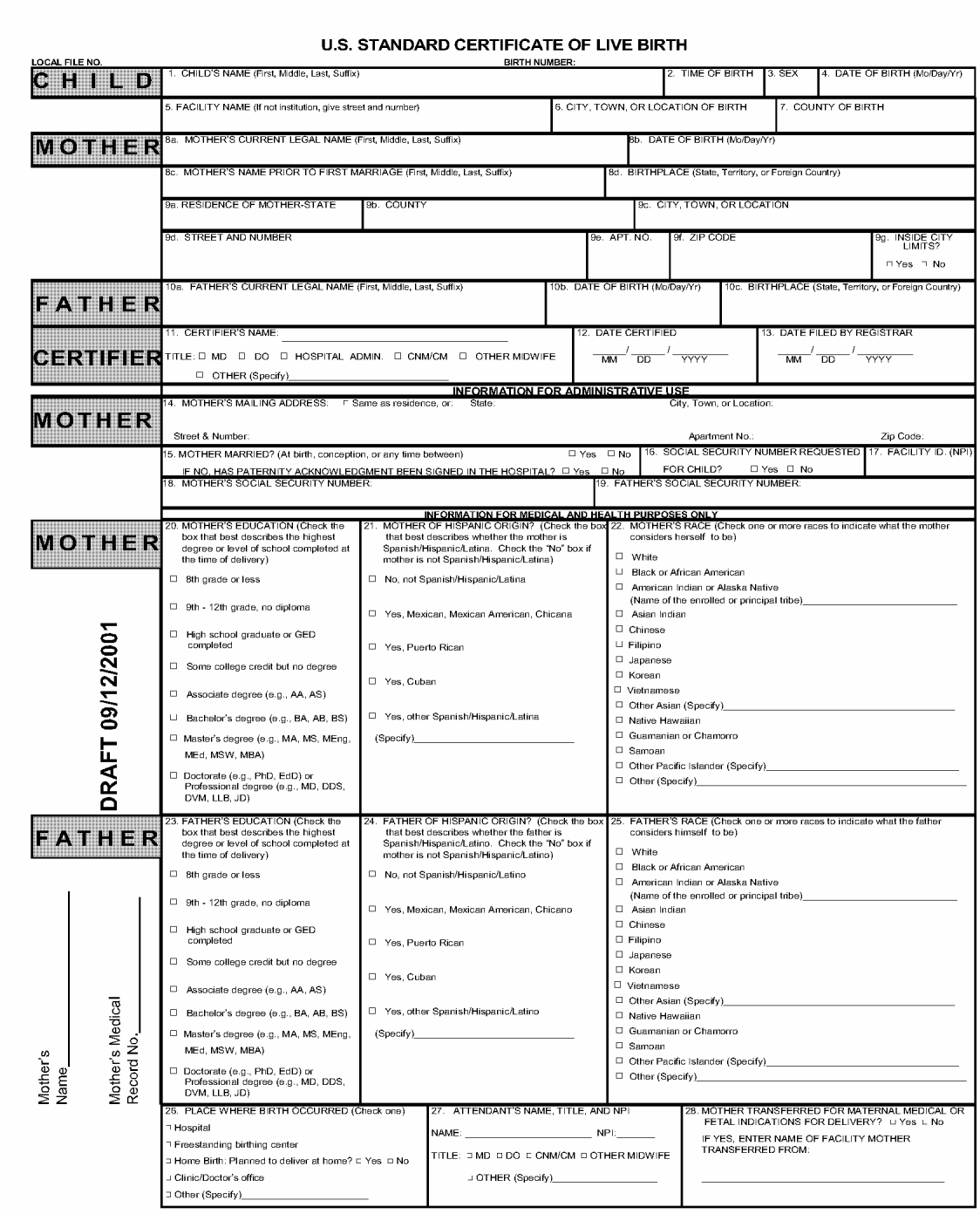
52. DECEDENT OF HISPANIC ORIGIN? Check the box that best describes whether the decedent is Spanish/Hispanic/Latino. Check the "No" box if decedent is not Spanish/Hispanic/Latino.
☐ No, not Spanish/Hispanic/Latino
☐ Yes, Mexican, Mexican American, Chicano
☐ Yes, Puerto Rican
☐ Yes, Cuban
☐ Yes, other Spanish/Hispanic/Latino
(Specify) _____

53. DECEDENT'S RACE (Check one or more races to indicate what the decedent considers himself or herself to be)
☐ White
☐ Black or African American
☐ American Indian or Alaska Native (Name of the enrolled or principal tribe) _____
☐ Asian Indian
☐ Chinese
☐ Filipino
☐ Japanese
☐ Korean
☐ Vietnamese
☐ Other Asian (Specify)_____
☐ Native Hawaiian
☐ Guamanian or Chamorro
☐ Samoan
☐ Other Pacific Islander (Specify)_____
☐ Other (Specify)_____

54. DECEDENT'S USUAL OCCUPATION (Indicate type of work done during most of working life. DO NOT USE RETIRED).

55. KIND OF BUSINESS/INDUSTRY

Figure 13.1 **Draft** 2003 revision of the national death certificates[43]



U.S. STANDARD CERTIFICATE OF LIVE BIRTH

LOCAL FILE NO. BIRTH NUMBER:

CHILD

1. CHILD'S NAME (First, Middle, Last, Suffix) | 2. TIME OF BIRTH | 3. SEX | 4. DATE OF BIRTH (Mo/Day/Yr)

5. FACILITY NAME (If not institution, give street and number) | 6. CITY, TOWN, OR LOCATION OF BIRTH | 7. COUNTY OF BIRTH

MOTHER

8a. MOTHER'S CURRENT LEGAL NAME (First, Middle, Last, Suffix) | 8b. DATE OF BIRTH (Mo/Day/Yr)

8c. MOTHER'S NAME PRIOR TO FIRST MARRIAGE (First, Middle, Last, Suffix) | 8d. BIRTHPLACE (State, Territory, or Foreign Country)

9a. RESIDENCE OF MOTHER-STATE | 9b. COUNTY | 9c. CITY, TOWN, OR LOCATION

9d. STREET AND NUMBER | 9e. APT. NO. | 9f. ZIP CODE | 9g. INSIDE CITY LIMITS? ☐ Yes ☐ No

FATHER

10a. FATHER'S CURRENT LEGAL NAME (First, Middle, Last, Suffix) | 10b. DATE OF BIRTH (Mo/Day/Yr) | 10c. BIRTHPLACE (State, Territory, or Foreign Country)

CERTIFIER

11. CERTIFIER'S NAME: _____
TITLE: ☐ MD ☐ DO ☐ HOSPITAL ADMIN. ☐ CNM/CM ☐ OTHER MIDWIFE
☐ OTHER (Specify)_____

12. DATE CERTIFIED ____/____/____ (MM DD YYYY)

13. DATE FILED BY REGISTRAR ____/____/____ (MM DD YYYY)

INFORMATION FOR ADMINISTRATIVE USE

MOTHER

14. MOTHER'S MAILING ADDRESS: ☐ Same as residence, or: State: City, Town, or Location:
Street & Number: Apartment No.: Zip Code:

15. MOTHER MARRIED? (At birth, conception, or any time between) ☐ Yes ☐ No
IF NO, HAS PATERNITY ACKNOWLEDGMENT BEEN SIGNED IN THE HOSPITAL? ☐ Yes ☐ No

16. SOCIAL SECURITY NUMBER REQUESTED FOR CHILD? ☐ Yes ☐ No

17. FACILITY ID. (NPI)

18. MOTHER'S SOCIAL SECURITY NUMBER:

19. FATHER'S SOCIAL SECURITY NUMBER:

INFORMATION FOR MEDICAL AND HEALTH PURPOSES ONLY

MOTHER

20. MOTHER'S EDUCATION (Check the box that best describes the highest degree or level of school completed at the time of delivery)
☐ 8th grade or less
☐ 9th - 12th grade, no diploma
☐ High school graduate or GED completed
☐ Some college credit but no degree
☐ Associate degree (e.g., AA, AS)
☐ Bachelor's degree (e.g., BA, AB, BS)
☐ Master's degree (e.g., MA, MS, MEng, MEd, MSW, MBA)
☐ Doctorate (e.g., PhD, EdD) or Professional degree (e.g., MD, DDS, DVM, LLB, JD)

21. MOTHER OF HISPANIC ORIGIN? (Check the box that best describes whether the mother is Spanish/Hispanic/Latina. Check the "No" box if mother is not Spanish/Hispanic/Latina)
☐ No, not Spanish/Hispanic/Latina
☐ Yes, Mexican, Mexican American, Chicana
☐ Yes, Puerto Rican
☐ Yes, Cuban
☐ Yes, other Spanish/Hispanic/Latina
(Specify)_____

22. MOTHER'S RACE (Check one or more races to indicate what the mother considers herself to be)
☐ White
☐ Black or African American
☐ American Indian or Alaska Native
(Name of the enrolled or principal tribe)_____
☐ Asian Indian
☐ Chinese
☐ Filipino
☐ Japanese
☐ Korean
☐ Vietnamese
☐ Other Asian (Specify)_____
☐ Native Hawaiian
☐ Guamanian or Chamorro
☐ Samoan
☐ Other Pacific Islander (Specify)_____
☐ Other (Specify)_____

FATHER

23. FATHER'S EDUCATION (Check the box that best describes the highest degree or level of school completed at the time of delivery)
☐ 8th grade or less
☐ 9th - 12th grade, no diploma
☐ High school graduate or GED completed
☐ Some college credit but no degree
☐ Associate degree (e.g., AA, AS)
☐ Bachelor's degree (e.g., BA, AB, BS)
☐ Master's degree (e.g., MA, MS, MEng, MEd, MSW, MBA)
☐ Doctorate (e.g., PhD, EdD) or Professional degree (e.g., MD, DDS, DVM, LLB, JD)

24. FATHER OF HISPANIC ORIGIN? (Check the box that best describes whether the father is Spanish/Hispanic/Latino. Check the "No" box if mother is not Spanish/Hispanic/Latino)
☐ No, not Spanish/Hispanic/Latino
☐ Yes, Mexican, Mexican American, Chicano
☐ Yes, Puerto Rican
☐ Yes, Cuban
☐ Yes, other Spanish/Hispanic/Latino
(Specify)_____

25. FATHER'S RACE (Check one or more races to indicate what the father considers himself to be)
☐ White
☐ Black or African American
☐ American Indian or Alaska Native
(Name of the enrolled or principal tribe)_____
☐ Asian Indian
☐ Chinese
☐ Filipino
☐ Japanese
☐ Korean
☐ Vietnamese
☐ Other Asian (Specify)_____
☐ Native Hawaiian
☐ Guamanian or Chamorro
☐ Samoan
☐ Other Pacific Islander (Specify)_____
☐ Other (Specify)_____

26. PLACE WHERE BIRTH OCCURRED (Check one)
☐ Hospital
☐ Freestanding birthing center
☐ Home Birth: Planned to deliver at home? ☐ Yes ☐ No
☐ Clinic/Doctor's office
☐ Other (Specify)_____

27. ATTENDANT'S NAME, TITLE, AND NPI
NAME: _____ NPI:_____
TITLE: ☐ MD ☐ DO ☐ CNM/CM ☐ OTHER MIDWIFE
☐ OTHER (Specify)_____

28. MOTHER TRANSFERRED FOR MATERNAL MEDICAL OR FETAL INDICATIONS FOR DELIVERY? ☐ Yes ☐ No
IF YES, ENTER NAME OF FACILITY MOTHER TRANSFERRED FROM:
_____

DRAFT 09/12/2001

Mother's Name_____ Mother's Medical Record No._____

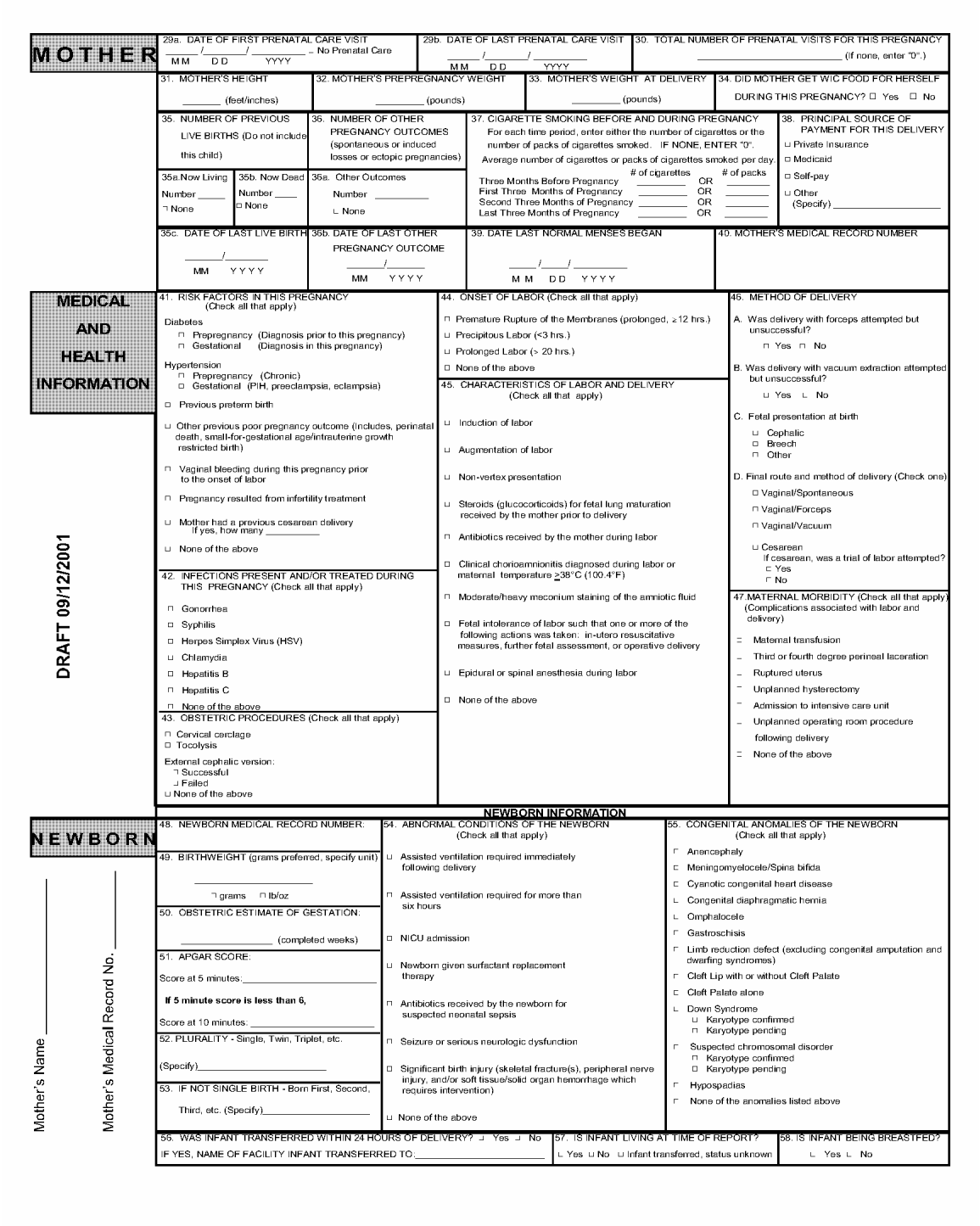


Figure 13.2 Copy of the 2003 draft U.S. Standard Certificate of Live Birth [44]

Inherent Timeliness and Sensitivity

Births can in theory be monitored in real-time through the use of birth certificates. In states such as New Jersey that utilize electronic transmission of birth certificates on daily basis, the certification of birth can be entered by the obstetrician as soon as mother and child are out of immediate danger and hands are washed. However such timely entry does not always occur, and is not legally mandated. Perhaps of more importance, completion of the birth certificate usually

waits for the parent to select a name. Some fields are not completed in a timely manner such as the recording of birth defects on the confidential portion of the birth certificate--precise diagnosis of birth defects often requires extensive evaluation so that a diagnosis is often not available by the time a certificate must be submitted to the state.

For fatal diseases, death certificates in theory can be the basis for a sensitive surveillance system capable of detecting outbreaks of fatal diseases of size as small as a single case (because every death must be certified). This potential assumes, however, (1) the accurate determination of the cause of death by the clinicians, and (2) the adequacy of the International Classification of Diseases - 10th Revision (ICD-10) (http://www.who.int/whosis/icd10/) to represent the cause. We note that ICD-10 is adequate to express causes of death due to most organisms of bioterroristic concern (e.g., anthrax pneumonia is a nosologic entity in ICD-10). However, the accuracy of clinical ascertainment and recording of cause of death is a significant limiting factor.

Timeliness is also a problem with death certificates because of the complicated nature of death-certificate issuance in most states. Unlike births, where the majority occurs at a limited number of health-care facilities, deaths occur in more locations and death certification depends on information from more individuals and institutions including funeral homes. This difficulty could be addressed by a system where a first responder could immediately record the fact of death and relevant individuals could record the other information as they become available. Such systems are being developed and promulgated. Issuance of a death certificate would still require completion of all information; however, deaths and diagnoses could be monitored as they become available.

Data and System Integrity

The accuracy of birth-certificate data varies with the individuals entering the data, and the uses of the data. Work in the early 1990's in New Jersey (prior to an electronic birth certificate) showed differences in coding for C-section rates between the birth certificate and the hospital bill Because the hospital bill was used for payment and subject to audit, more institutional efforts were put into accurate completion.

The work group of The Division of Vital Statistics at the Department of Health and Human Services that develops specifications for the 2003 revision[45] has included specifications of the types of edits (e.g. the editing program will not accept a mothers age as 6 years old, and the program would question but ultimately allow a mothers age to be coded as 11 years old. Efforts such as this will improve data and system integrity.

Dual usage of data helps to improve the accuracy of data. For example, even after the implementation of the electronic birth certificate in New Jersey (including immediate data entry, audits, and real time edits) occasional mistakes still occurred. One clerk was found to be coding 700 grams instead of 7 lbs. This error was picked up when several mothers of 7-pound infants complained about receiving letters requesting that their infants be provided hearing screening because of low birth weight.

The recording of the *fact-of-death* in death certificates should be very reliable since the next of kin need these documents to bury the deceased and for other legal purposes as described above. Fact-of-death would have most utility for detecting pathogens that are fatal so quickly that the individual dies before interaction with the health-care system. Rare errors in *fact-of-death* occur

when the body is missing and death must be presumed. In these cases, death certificates are often not issued for a number of years, however exceptions have been made in cases such as the September 11th tragedy where a well documented accident or fire precluded the recovery of bodies. Errors of misidentification have also occurred, however modern technology now allows identification if even a small amount of DNA remains.

Death certificate information other than *fact-of-death* are less reliable. Age is often incorrectly recorded especially for the very old and those who have few social supports. Cause of death can be inaccurate especially if an individual has not been autopsied. Nosologists are employed by state health departments in order to improve the coding of deaths. However all of the after coding improvements won't work if the initial clinical diagnosis is wrong and no autopsy is performed.[46, 47]

Legal or Administrative Barriers

States vary in their policies governing the release of death information. Many states require approval by a specific board for research. At the national level, there is a model state law developed by the Centers for Disease Control /National Center for Health Statistics that addresses these issues. These model regulations keep birth certificate information confidential for 100 years and death certificate information confidential for 50 years. Information is permitted to be released to the next of kin, those with specific legal need, for governmental administrative or research need approved by the registrar. [48]

Technical Barriers

All states transmit birth and death data electronically to the federal government. Most if not all states accept filing of birth data electronically. The mechanism for initial filing of a report is typically a stand-alone computer system into which the filer enters data and can print a certificate. There is little if any information provided from hospital medical records or other data systems.

The technical barriers to achieving national electronic surveillance of deaths are greater because of the complexity of the process and the number of signature needed. The process of filing a certificate requires three components: pronouncement of death signed by the pronouncer, certification of cause of death (and medical facts) signed by the certifier (or medical examiner), and decedent demographics completed and signed by the funeral director. Each part requires a separate signature, which means that the physical form be hand carried or delivered to several locations before it is finally filed at the local registrar and then submitted to the state. The report "Toward an Electronic Death Registration System in the United States: Report of the Steering Committee to Reengineer the Death Registration Process" concludes that *...registration processes remain labor intensive, employ disparate and limited automated procedures, and require several professionals at different locations to complete each of the more than 2.3 million death certificates registered each year.* [49] However, the national impetus towards electronic death certificates is high because of the desire of the Social Security Administration to immediately stop checks to individuals who have died. Because of this economic motivator, barriers are documented and are being addressed by NAPHSIS and Social Security.[49]

Current Availability of Data in Electronic Form

One vendor (Genesis Inc of PA) has electronic birth-certificate contracts with more than 30 states and covers more than half of US births. The vendor provides DOS and windows based

products customized to each state. Other states are covered by their own or other vendor systems. Typically two processes are involved. First the hospital enters data into a computer and prints out a certificate of birth to be signed by the certifier and mother. The hospital then sends the paper certificates via mail or currier and the electronic data via diskette, batch mode modem transmission, or, less frequently, via the Internet. At the state a separate software system allows the state to accept the certificates, edit the certificate (e.g. nosologist cause of death or adoptions where changes in parents are made.), and issue certificates. Data from this system is then sent to the federal government usually in batch mode via tape.

Electronic filing of death certificates is much less common, however this may change rapidly in the next couple of years. The Social Security Administration funded five states and one city (California, Minnesota, New Hampshire, New Jersey, New York City, and New York State) to demonstrate the feasibility of online verification of social security numbers by funeral directors. These demonstrations are part of the Electronic Death Registration Project. The project's objectives are to develop guidelines and standards for electronic death registration. [49] New Hampshire and District of Columbia are expected to obtain additional funds that will allow them to support Web-based electronic incremental filing of death certificates entry, in which information can be entered as it is received. These systems are designed to provide more immediate information about changes in death rates.

Summary

Birth certificate data are of potential use for only a few organisms that cause early births or birth defects. They are: *Neisseria gonorrhoeae, Treponema pallidum, Herpes simplex*, group b streptococcus, syphilis, rubella, cytomegalovirus, *and Toxoplasma gondii*. None of these agents are likely to be used for bioterrorism. From death certificates, the fact of death is ascertained with accuracy, as are demographics. The cause of death is less accurate. Death certificate data are already used in public health to discover trends in death rates in different demographic subpopulations (e.g., oral cancer women in the southwest) over time. Whether these data can be useful in bioterrorism detection is an open question. The value of death records for bioterrorism detection is limited by their inherent lateness and accuracy. Their utility is further reduced by delays introduced by the (largely) manual process of ascertainment of cause of death and of recording of demographic information by funeral directors. It is too early to comment about the reliability and availability of death certificate data in real-time systems as such systems are still under construction. We would expect they to be accurate and reliable. Legal and administrative issues related to public-health use of such data in real-time have not been addressed, nor have issues of cost of access. Technical barriers to integration of such information also would have to be addressed, especially agreement about how to encode the different data elements captured.

Therefore, for many public health threats, we would expect that other types of data would provide earlier warning. Exceptions are small outbreaks due to lethal pathogens that result in single or a few deaths. Here, with complementary information from hospital or medical examiner records discussed in Chapter 16, the detection of an unusual cluster of deaths over the background might be possible.

Questions for Further Study

1. What are the latencies between the initial symptoms of a fatal illness and death for all diseases of interest?

2. What are the smallest outbreaks that could be detected using death certificate data (fact of death, demographics, and the cause of death)--assuming 100% reporting? Assuming that the reporting occurs in real-time, what would be the timeliness of detection?
3. How could death certificate data be used with other types of data for detecting an outbreak automatically?
4. Questions 2 and 3 for birth certificate data?
5. What data coding schemes are available for death and birth certificates? What additional data coding schemes are needed? What additional data should be collected?

Chapter 14. Monitoring of Animals

There are three reasons to monitor animals for pathogens and disease. First, diseases can be transmitted from animals to humans so monitoring of disease prevalence and distribution in animal populations may contribute to the detection of outbreaks in human populations. Second, diseases can cause sickness in both animals and humans and in some cases animals are more susceptible than humans. The classic example of this effect is the sensitivity of canaries to coal gases. Therefore, monitoring of animals may provide an early warning of a co-occurent outbreak in humans. Third, animals have economic value; hence, disease in animals can threaten the economic health—and in the extreme case physical health through starvation—of human populations.

The crossover threat posed by microbes in animals is partly determined by the degree to which we share DNA with them. For example, the chimpanzee genome contains 99% of the human genome, and, consequently, chimpanzees can fall ill with many illnesses that also strike humans. The reverse is also true. In contrast, a human is unlikely to be affected by many pathogens, which make a lion ill, with the notable exception of enteric pathogens such as salmonella. Examples of diseases that can be transmitted from animals to humans include cat-scratch disease, rat bite fever, rabies, tularemia and plague. Other mammals and birds can fall ill with the same viruses that affect humans, examples of which include encephalitides and West Nile virus. Reptiles have even less in common with humans, but can carry salmonella, and infect humans who handle the animals. Young children are especially vulnerable when manifesting an all-too-common habit of not washing hands after handling reptiles.

There are several fields and practices that are relevant to monitoring of disease in animals including veterinary medicine, agribusiness, and public health. Each of these fields uses information systems to track animal health and disease in populations of animals. We note that veterinarians and veterinary medicine play a prominent role in each field. Although veterinarians have special interests, similar to physicians, the practice of veterinary medicine is divided, generally, into these three spheres. One sphere of practice is the care of domesticated animals, which includes animal hospitals, private clinics and offices of such groups as the Society for Prevention of Cruelty to Animals (SPCA). Another is institutional veterinary medicine, which includes zoos, wild animal parks, aquariums and aviaries. Yet a third is public

practice, essentially comprised of government veterinarians who watch over agribusiness (cattle, swine), and organisms of relevance to public health (e.g., birds and mosquitoes). These distinctions are of importance, because the three categories of practice utilize what we will call electronic veterinary records (EVR) very differently.

Veterinary Medicine

The practice of veterinary medicine includes the examination, diagnosis, and treatment of animals-- whether mammals, reptiles or fish. Veterinary medicine has an important role to play in the detection of public health threats.

An EVR is defined as software running on a computer system that allows the recording of veterinary history, physical examination notes, lab results, surgical notes, diagnoses, and treatment. The function of an EVR parallels that of a point-of-care system in human patients. Veterinary practice is primarily a cash business; hence, third-party-payer-oriented coding schemes such as ICD-9 are not utilized. Nevertheless, EVRs are often designed to be compatible with, and accept, a controlled vocabulary, most often SNOVED, a cousin of medicine's SNOMED. But SNOVED is optimized for pathology, which limits its acceptance among many veterinarians.

EVR Use: Domestic Veterinary Practice

The degree of acceptance of EVRs in veterinary hospitals focused on domesticated animals is about the same as POC acceptance in medical hospitals. No more than 10% of veterinary hospitals have them installed. These systems range from complete electronic veterinary systems with entries for history, physical examination, laboratory data, radiology data, diagnoses and procedures, to systems, which merely list pedigree, vaccinations, current, diagnoses and what tests or procedures have been ordered (but not their results or follow-up). A prime example of the former is the EVR developed at the veterinary teaching hospital of the University of California at Davis; the latter, functioning more as a billing tool than a medical record, is typified by the University of Pennsylvania's veterinary hospital. Many veterinary hospitals, even those affiliated with schools of veterinary medicine, have no electronic medical record systems. In veterinary practices that have them, one may find that some veterinarians utilize the system, while others refuse to do so. Some systems are extensions to practice management software systems, most commonly used for billing, with the EVR capability unused. (Sources: American Animal Hospital Association, vet. Interviews)

The advantages of EVRs are similar to POC systems in human medical settings: they receive data from veterinarians in real-time or near real time, and, as with relational databases, allow multiple views through the data. Limitations to EVRs usefulness to public health authorities in this practice setting are also similar to those in human medical settings. The number of systems installed is relatively small. Confidentiality must be respected; the diagnosis and treatment of an animal by a veterinarian can be disclosed only with permission of the pet's owner. Additionally, laboratory systems can be a little less efficient than in human settings. Many veterinary laboratory instruments are not connected to databases; hence, a technician must manually enter all lab data into an EVR, which can introduce a delay of one to two days in entering that data, if the veterinarian is very busy and behind schedule.

EVR Use: Institutional Veterinary Practice

Institutional veterinarians, who sometimes also maintain private practices, work in zoos, wild animal parks, aviaries and aquariums. These practitioners, and the medical facilities they run, specialize in undomesticated animals.

Institutional veterinary centers most often deal with publicly owned, or publicly supported, animals, birds and fish. The care of these patients is not subject to the same privacy and confidentiality regulations as privately owned pets; moreover, public zoos and parks are subject to the Freedom of Information Act, wherein interested parties may demand to inspect documents related to the care of animals at a given facility. (That fact notwithstanding, private, for-profit or not-for-profit facilities can still be cooperative regarding the sharing of veterinary information especially for valid research purposes.)

However, this relative openness is not limitless. Zoos are understandably sensitive to disclosures, which may discourage the public from visiting. Such disclosures may include the death of an animal, especially when a report may imply, intentionally, or otherwise, that visitors are not safe. Zoos also bear the brunt of frequent legal assaults and harassment by animal rights advocates, and as a result, zoo administrators have tried to restrict the flow of information in certain ways.

The vast majority of large and medium-sized zoos and animal parks belong to ISIS; ISIS stands for the International Species Identification System, which is both a standardized way of classifying and inventorying animal collections in zoos, parks and aquariums, and an organization which maintains those standards. ISIS membership is comprised of over 450 such facilities, half in the United States, with the other half located throughout the world. Years ago, ISIS developed a DOS-based program called Arks, which functioned as an animal cataloguing and inventory system. All ISIS members received this software as a benefit of ISIS membership. Subsequently, an EVR extension was added to Arks, called MedArks. Currently about 90% of ISIS member facilities use MedArks as their EVR, and, in contrast to the situation often encountered, in POC-equipped human medical facilities, the zoo or park will utilize MedArks to record the care of its entire animal population MedArks is designed around a DOS-based, SQL-compliant Foxpro® database, which, unfortunately, precludes downloading or uploading data files directly through the Web. However, this problem has been addressed: MedArks is designed to allow one institutional veterinary center to transmit EVRs to any other MedArks-equipped facility, as well as ISIS' central database, by email, or by mailing a diskette containing the desired files. Currently, ISIS members send data (sets of EVRs) to the ISIS veterinary database on a monthly basis; this schedule is subject to manpower availability, and is not limited by technical means. Of course, these same transmissions could be also be received by a public health agency, if appropriate arrangements were made.

A few zoos, such as the San Diego Zoo, and the Denver Zoo, do not use MedArks, and have installed a modern's Windows based EVR equipped with a relational database. It is important to note, however, that nearly 100% of medium-to-large sized institutions use EVRs of one type or another. The distribution of institutions in most metropolitan areas of the United States (as well as many areas of the world), their lower privacy and confidentiality barriers, and their commonality of EVRs, combine to offer a unique opportunity to develop an early warning sentinel system for those public health threats that can affect both humans and *sentinel species* in a manner not currently possible in human medicine.

The continued commonality of this system is threatened by the age of its software. There is currently no modern, Web-based or client-server version of MedArks available to replace existing implementations, and some institutions have threatened to purchase or develop customized systems, causing some divergence in function and potentially complicating compatibility across ISIS members. The National Aquarium in Baltimore, the New England Aquarium and the (institution in Chicago) are attempting to reengineer MedArks for the Web, but their version's widespread implementation would have to overcome cost and other hurdles.

The detection spectrum of such a sentinel network, if erected, is likely to be selective, and possibly quite narrow. In zoos and wild animal parks, wild animals are not allowed to circulate with humans; thus opportunities for disease transmission are reduced. Zoos exchange animals with each other, not with the general public (exceptions exist, of course). Institutional veterinarians do not deal with animals owned as pets. However, some animals, such as birds, and stray dogs or cats, may cross zoo boundaries and come in contact with park visitors.

On the other hand, institutional veterinarians, and their ubiquitous EVRs, will have opportunities to intervene in cases that private veterinarians may turn away: a veterinarian in private practice, for example, cannot be counted on to locate and treat stray animals. Moreover, they frequently turn away an animal if the owner does not have enough money to pay for care. In contrast, institutions employ zookeepers who inspect the grounds of their institutions and collect dead animals for examination, whether or not they originate in the institution's collection. In such cases, the institutional veterinarian will often perform a necropsy, especially when there are signs of an unusual or unexpected illness. In settings where veterinarians must practice population medicine rather than treat individual patients, (for example, a school of fish), a suspicion of illness may be addressed by sacrificing one animal to determine the nature of the illness.

This has important implications to public health. To cite one example, it was a zoo veterinarian, not a physician or public health service officer, who uncovered an epidemic of West Nile virus in 1999 in New York; subsequently, the public health authorities confirmed her finding.

DARPA has recognized the importance of veterinary medicine in the early detection of threats, such as bioterrorism, and is conducting studies in this area.

Public Veterinary Practice-Agribusiness

Both state and federal government agencies employ veterinarians to inspect and help care for cattle and swine.

In addition, ranches and feedlot operators, described below, hire veterinarians and nutritionists to help them optimize their cattle yield. Information about food animals, and their veterinary care, is collected at several points in the journey from the farm to the packing plant.

It should be noted here that veterinary medicine, as practiced with regard to food animals (cattle, swine), emphasizes population medicine and cost/benefit analysis to a much greater degree than human medical practice or private veterinary practice. This is because the veterinarians are supporting food safety and economic goals: growers want to maximize the percentage of the animals they grow which are accepted by beef processors without spending inordinate amounts of money on veterinary care. Each cow is worth a few hundred dollars to the rancher;

administration of medicines costing as little as $20 to that animal may significantly reduce the rancher's profit margin. Beyond that, the growers are also interested in optimizing the quality of beef these animals represent; the grade of beef assigned to each animal upon slaughter by the USDA's Food Safety and Inspection Service (FSIS) will help determine the price paid to the animal's growers by the beef processor.

The Ranch

A cattle ranch purchases steers (castrated bulls) and heifers (cows) from a variety of farms, and raises the animals on pastures until they are sufficiently mature, and have gained enough weight, to be sent to a feedlot. The primary goals are to use cost-effective means to keep the herd healthy and encourage optimal weight gain. Ranches do not generally track animals on an individual basis, but, rather, by pen assignment. Veterinary care is delivered through periodic inspection of herds by experienced ranch hands, who receive training from veterinarians. These ranch hands, or cowboys, take note of the presence of animals exhibiting the symptoms of infectious illnesses. As examples, a cow may have a nasal discharge, may wheeze or cough audibly, may refuse to hold its head level, may have a "glassy" stare, may vomit or may refuse to follow the herd around the pasture. The ranch hands "pull" the sick animal to a "sick pen" in order to minimize the opportunity for the spread of disease through the herd. They then summon a veterinarian to evaluate the animals in the pen and authorize appropriate treatment.

Evaluation of a food animal emphasizes symptomatology because, in contrast to veterinary practice in animal hospitals and private offices, veterinarians dealing with food animals do not perform a complete physical examination on each sick animal individually, using a stethoscope for the heart and examining the pharynx. There are far too many animals to survey and it would not be cost-effective. Veterinarians will perform an abbreviated necropsy on each dead animal they find. If a carcass has an unusual appearance, or the veterinarian suspects the presence of a particularly virulent pathogen, he/she may arrange for the dead animal's shipment to an animal hospital for complete necropsy, which would include laboratory testing of bodily fluids, weighing of organs and more extensive pathological examination than is possible in the field.

Data collection at the ranch emphasizes aggregate and statistical reporting. Animals are not assigned individual health records; although there has been some use of individual records on a trial basis, these have not been proven, in the eyes of ranchers, to be cost-effective.

Data systems in use can vary. Larger feedlots utilize modern relational databases; this makes their data accessible and their databases amenable to modifications for public healthy purposes.

The Feed Lot

Feedlots accept cattle from farmers and ranchers feed them and provide them with veterinary care, as at the ranch, the primary goals are to use cost-effective means to keep the herd healthy and encourage optimal weight gain. The feedlot is also concerned with "finishing," the process by which the animal's feed is adjusted to optimize beef yield at the processing plant. Whereas, at the ranch, a cow feeds on pasture grass, or, during the winter, on irrigated wheat patches grown especially to compensate for the shortage of appropriate grazing areas, a cow in a feedlot will be switched to other feeds, such as alfalfa and corn. The latter feeds increase the fat content of the muscle tissue ("marbling") and, hence, the percentage of beef likely to be graded "prime" or "choice" by the FSIS. This, in turn, increases the price paid by the processor for the animal in question.

When an animal reaches maturity, the feedlot ships it to a beef processor's slaughterhouse. Beef processors own some feedlots; many others are independent companies that have contracted with beef processors to supply them with animals.

The information collected by feedlots about the animals they raise can vary. Large feedlots are in the minority; these will track each animal individually, collecting data on weight, temperature, feeding, history of illnesses and treatment, including any visits to a veterinary hospital. Essentially, each animal receives a "passport" reflecting and recording its journey through life. The larger feedlots maintain modern databases to assist them with quality assurance and operating efficiency. Identification of individual animals is accomplished by means of an ear tag. Currently, the majority of ear tags are simply plastic cards with numbers printed on them. Feedlot managers, recognizing that there are advantages to systems that can help overcome the challenges brought on by mud, snow and rain, are introducing new technologies. Some new ear tags are equipped with passive transistors; the ranch hand points a transmitter "gun" at the tag and the tag returns a code to the gun's display, much like an airliner's transponder returns an identification code to a radar screen. Yet another technology, still in its infancy in this business, is the retinal scanner.

The majority of feedlots will note the pen the animal is kept in, but not actually assign an individual identification number to the animal. Veterinary care is reflected not by records kept about a sick animal, but in the form of general ledger entries indicating how much money is spent on medications and veterinary visits during a given fiscal period. Thus, one could not necessarily differentiate between the ten animals in a given pen. This approach is used for two reasons. First, the beef processor's priority is obtaining a shipment of healthy animals for slaughter, and a feedlot's management believes that this can be accomplished without the bother of tracking individual animals. Second, managers at smaller feedlots may perceive a desire by larger operations to share the risk of lawsuits with them through a "traceback" capability.

When an animal arrives at a slaughterhouse, it is accompanied by paperwork certifying that its physical condition satisfies federal and state law, as well as the contract between feedlot and beef processor. This is not equivalent to a complete record for the animal, and, in many situations today, is not transmitted electronically. Rather the feedlot provides a subset of its data required by law and contract to the beef processor, which then enters the data into its own information systems. A number of different systems currently serve the feedlot market.

The Slaughterhouse

Slaughterhouses, or beef processing plants**,** accept animals from feedlots for slaughtering and manufacturing of meat and meat-derived products. In the United States, three large beef processors dominate this part of the meat industry: IBP, a subsidiary of Tyson Foods; ConAgra, and Excel Corp., a subsidiary of Cargill. Of these, only ConAgra and Excel own feedlots; company-owned feedlots account for a minority of the food animals shipped to slaughterhouses. Primary responsibility for inspection of food animals at the slaughterhouse rests with the US Department of Agriculture.

The US Department of Agriculture's Food Safety and Inspection Service (FSIS) employs well over 1,000 veterinarians for this purpose. These veterinarians perform both antemortem and postmortem examinations of animals at commercial slaughterhouses, such as those operated by

the large meat processors Excell, ConAngra and IBP. At they arrive at the slaughterhouse, animals are screened by USDA inspectors; animals who raise suspicions are diverted to a holding pen where an FSIS veterinarian examines them for evidence of disease. The FSIS veterinarians record observations on standardized paper forms, which are batched and forwarded to a data entry facility in De Moine, Iowa at the end of each workweek. These observations include number of tumors found, condition of skin and muscle, presence of pneumonia or other infection, and presence of injuries. The forms sent to De Moine are not complete history-and-physical examination records; rather, they are lists of anomalies found in each animal. The forms themselves are pathology-oriented. All forms are submitted over the weekend, and data from them appear in USDA's database the following Monday. Thus, the newest data in the database are 72 hours old; the oldest (from that week) is 10 days old. The database itself is housed on a minicomputer running an aging, proprietary, hierarchical database, posing potential difficulties for adapting the system to public health surveillance purposes. The system's design, its age, its range of time lags, and the current absence of EVRs in the slaughterhouses, tend to limit the usefulness of the USDA's legacy database in the detection of bioterrorism.

In the spring of 2002, the USDA's FSIS plans to roll out a new relational database, coupled to an on-line submission system replacing the paper forms. If this is accomplished successfully, FSIS veterinarians will be able to enter antemortem and postmortem reports into the database in real time; the new database will be very amenable to modification. Public health agencies will be able to set up their own views into the database and perform surveillance and signal detection.

Company policy regarding recording of veterinary information may vary; IBP, for example, relies on FSIS for its in-plant documentation and does not generate any paperwork of its own. IBP has a staff veterinarian who is concerned with quality control, specifically, responding to FSIS findings and devising strategies to reduce the percentage of sick animals which IBP has paid for but cannot slaughter and process. IBP does not have an in-house EVR.

ConAgra, and the feedlots it owns or are under contract to it, are promoting an initiative to create information systems that will closely track the progress of each food animal from the date of birth until the date of slaughter and distribution. The efforts are not necessarily smoothly coordinated, in that each level of the industry must still protect its own interests, who can conflict with the purported goals of the initiative; different companies create different visions of the future. This initiative is another part of the effort mentioned in the feedlot section to give each animal a "passport," or a completely integrated life, health, and sale record.

Public Veterinary Practice: Public Health

For completeness, we mention information about animal health that is collected, stored, and analyzed by health departments. These data include dead bird necropsies, mosquito pool studies, and other vector surveillance information. These data are described further in our first report.

Summary

The care of animals in the United States involves population medicine to a greater extent than the care of people. While private veterinarians do not store diagnosis and treatment data electronically to any greater degree than physicians do, and a pet's veterinary record is subject to the same confidentiality rules as its owner's, veterinary care in institutional (zoos, animal parks, aquariums) and cattle industry/public practice settings is well supported by computers, databases and software. Institutions are universally equipped with electronic veterinary record (EVR)

systems linked by email into a loose but effective network. The typical cattle industry veterinarian is a sophisticated user of statistical tools, more so than his/her physician counterpart, though the emphasis on population medicine has, historically, meant that EVR use for individual animals has not been a priority. That is changing, however, as beef processors and feedlot operators recognize the economic value of an animal health record. Individual animal illness and treatment records have been successfully deployed on laptop computers for field use. These data are recorded by experienced ranch hands and veterinarians, and are intended, ultimately to be a part of the system that produces a "passport" for each animal, from birth to slaughterhouse.

These data are highly accessible, and already very well suited for public health purposes, because of the emphasis on population medicine and a sharing of goals. Early disease detection in animals translates into lower herd losses and higher profits for the rancher and feedlot operator. While providing this information to public health authorities is likely to require considerable financial compensation, the industry as a whole is likely to recognize that it is in its best interest, after all, to cooperate.

The US Department of Agriculture's Food Safety and Inspection Service (FSIS) collects pathology-oriented data from slaughterhouses, and is in the process of deploying a new data system which will allow entry in real time. As this data includes information regarding carcasses, it is accurate and reliable, though not oriented toward infectious diseases *per se.* The marginal cost to the government for using FSIS data for public health purposes will be low.

Animal monitoring is a potentially rich source of data for detection of bioterrorism and other public health threats. As in many other fields of endeavor, money, competition and politics will greatly influence the outcomes of recent initiatives.

Questions for Further Study

1. Institutional use of EVRs is nearly universal, yet perpetually threatened by budget limitations. To what degree should the federal government subsidize EVR deployment at zoos and animal parks for public health purposes?
2. Efforts by beef producers to create "passports" for cattle may result in a set of competing standards. How wills this affect threat detection by public health authorities?
3. How should the federal government respond to the beef industry? What is the nature of the relationship between the Department of Health and Human Services (DHHS) and the US Department of Agriculture (USDA)?
4. In view of the threat bioterrorism poses, should DHHS play an increased role in promulgating standards in the cattle industry?

Chapter 15. Information Retrieval

When people search for information, the questions they ask, if they can be accessed by public health, can provide early indications of illness.

Searching for Information on the World-Wide Web

A person searching for information on the World-Wide Web can conduct that search using several approaches and tools. First, the inquirer can point his/her browser to a specific URL (Universal Record Locator), or address, of a website he/she wishes to visit, or even a specific page within the website. This is the most efficient approach when the inquirer already knows where to look for the desired information. A variation of this approach involves visiting many sites, whose contents are generally known by the inquirer.

When the inquirer does not know precisely where to look, he/she can utilize Web directories or search engines to attempt to locate desired information. A Web directory, such as Yahoo! (R), is a compilation of websites and their contents, created by the directory's staff. A search engine is a sophisticated software suite designed to search websites for key words and phrases, and present them to the inquirer based on some criteria. Some search engines order the results based, in part, on fees paid by website owners who want the search engines to include their sites in searches.

Tracking Web-based Inquiry Behavior

The measurement of the public's use of the web is a highly controversial topic. Advertisers, marketers and website owners often cannot agree on basic issues, such as, "How many people visited my site today," and "What did they look at?" Firms specializing in measuring the market share of entertainment content, such as A.C. Nielsen, have entered the fracas, and the various competitors cannot agree on what is the "right" way to measure the popularity of websites. This controversy is relevant to the question at hand, because, in order to tie web-searching behavior to the presence of an illness, one should know what a given inquirer is looking at and why.

Use of Directories and Search Engines

Directly observing the use of a directory or search engine is, conceptually, quite simply. Logging each "hit" on the site, and the word or phrase requested in the search will produce a list

of search terms submitted, along with the number of times each term was requested. This is true for both general-purpose search engines and directories, and special-purpose tools, such as those found on health-related websites.

Deciphering the reason(s) for a search, or concluding that a large group of searches are related, however, is more difficult. First, the term being submitted may reflect a personal circumstance, academic curiosity or other reason. For example, a person submitting the term, "cough" may have a cough and want to do something about it, may be a medical student assigned to research the physiology of the cough reflex, or may be someone looking for "cough" as part of the title of a novel, which has nothing to do with health. We cannot know with certainty why someone is submitting a search term unless the inquirer offers a reason during the encounter. Few search engines and directories collect this kind of information.

Second, the search engine or directory selected, and the way it presents information in response to a query, directly affect what the inquirer views. Did the inquirer select the search engine for a particular reason, or was the choice a random one?

We can attempt to simplify our consideration of the inferential significance of web searches by restricting ourselves to only health-related directories and search engines. Presumably, an inquirer utilizing the search functions of WebMD or a naturalopathic website to look for "cough" is not doing so for non-health related reasons.

Health-related Websites

American are heavy users of web sites for health information. Health-related websites may cater to health-care providers, or to the lay public, or both. Material from these sites may be retrieved as searches, or following a menu may retrieve them. A "search-term" driven inquiry is keyed to the term submitted, and is the more obvious of the two to keep track of. A menu-driven search can also reveal something useful about the inquirer's purpose; how much is revealed depends on how specific the menu choices become as the inquirer proceeds through them.

In either case, tracking can be accomplished through passive surveillance software, now widely employed commercially, both by media analysis organizations, and by employers who monitor their employees' on-line activities. Employing this kind of software for public health use will incur costs for acquiring and deploying these tools, and will require authorities to address objections raised by those who view such monitoring as an invasion of privacy. While such monitoring has been judged acceptable for employee monitoring, its use for public health surveillance may face constitutional scrutiny.

Searching by Healthcare Providers

The number of on-line sources of peer-reviewed professional information has increased dramatically in the last several years. Web sites of interest to healthcare providers include those run by professional societies and charities, medical center websites, pharmaceutical company websites, websites run by journal publishers, and for-profit services such as WebMD and Medscape. However, this development has not been accompanied by increased utilization of electronic sources by healthcare providers. For example, both physicians and nurse practitioners tend to consult each other, textbooks and reference manuals rather than Web sites when they have questions they want to answer. Moreover, physicians do not actively seek answers to every

question that comes up during the work day; according to one study, as many as seven questions out of 10 remain unanswered.[50-56]

Of course, we are more interested here in detecting significant changes in query patterns, than the raw number of queries, but the existence of a pattern depends on the submission of queries, and the reluctance by providers to fully embrace the Web and to pursue answers to their questions may degrade the quality of the data available for analysis.

Searching by Lay Public

The public's use of the Web has increased dramatically. [52-56] Queries about health information have increased. People use the web in a variety of ways. Many inquirers are seeking a second opinion after having spoken to their physicians. Others, exhibiting very active information-seeking behavior, are educating themselves rather than choosing to receive information only from their physicians. On-line information-seeking behavior varies among people, depending on degree of access to computers and the Internet, age and socioeconomic background, and education.

Nonetheless, use of the Internet by the public has increased so much, that it could be safely assumed that there are more than enough queries occurring to provide a good basis for analyzing them.

Other Information Retrieval Systems

There are many other types of information systems besides web sites that potentially can be helpful. Web portals for health systems, discussed in Chapter 4, are one such example. Physician's Database Searches as a tool for early detection of epidemics is another with proven potential[57].

Summary

People use information retrieval systems to find information when they are ill, and physicians use these systems when they are seeing ill people. Because of the inherent earliness of such data, the potential of the query and exploration patterns of sick individuals is an important area for further study.

Chapter 16. Coroners and Medical Examiners

The public agencies charged with investigation of causes of death, and assisting with identification of deceased persons are either the Medical Examiner's office, headed by a physician, or a coroner's office. The coroner may be an elected official who entrusts the performance of this work to a physician. For our convenience, we will use the term *medical examiner* to refer to all such agencies.

The medical examiner must prepare a record, for a deceased person, not unlike the medical record other physicians prepare for their living patients. The medical examiner must carefully review and record a history and pertinent past medical history, supplied by witnesses, family members and medical records obtained from the deceased person's health care providers; he/she must then perform a thorough physical examination, which includes inspection of the body, examination, weighing and dissection of organs. The medical examiner may order radiological and laboratory tests as appropriate. Pathology specimens may be prepared as well. The purpose of this work is to reach a conclusion regarding the cause of death. The cause of death may be considered the deceased person's *diagnosis*. The types of information collected and recorded by the medical examiner make this type of record helpful in considering every category of lethal public-health threat.

Electronic Medical Records for Medical Examiners: Accessibility

Although medical records and medical examiner records share common elements, there are, obviously, some aspects of medical examiner records, such as organ weights and dissection observations that are unique. Hence, electronic medical record systems were developed especially for medical examiner use. The vast majority of medical examiners covering large US cities and suburban jurisdictions has installed, or are installing now, electronic record systems. Early generation systems--developed in-house or sold by commercial vendors--used relatively inflexible data storage schemes. Current systems rely on relational databases and are as adaptable to new uses as modern electronic medical record systems. Moreover, their prevalence among most large medical examiner jurisdictions in the US means they can play an important role in detecting a public-health threat. This market penetration is in contrast to point-of-care

systems for live patients, whose less frequent deployment reduces their potential in the public health early-warning role (although other limitations of medical examiner data such as the limited number of people that die and come to autopsy offsets this deployment advantage).

Of note is that medical examiner electronic record systems utilize laboratory information systems in the same way that point-of-care systems do for live patients. The medical examiner's office will receive data from laboratory systems regarding a deceased person. In theory, public health authorities would be able to tap these data either from the laboratory system, or the medical examiner's system.

Technically, access to medical examiner databases by public health authorities is achievable in ways very similar to the access achievable with standard electronic medical records. The required software modifications may be extensive, depending on the system involved.

Timeliness

As would be the case with any electronic medical record, data regarding a deceased person can be available for public health as soon as it is entered into a database (assuming that technical integration and administrative barriers have been addressed). Some data, such as laboratory values or pathologic examination, might be entered at different times than the results of the physical examination. Additionally, while live patients can usually supply a past medical history immediately, the medical examiner may have to wait longer to collect such information.

As a result, the time latency between time of death and a medical examiner's reporting of data regarding the deceased can be protracted. The post-mortem examination takes time. If only body parts are found, the true cause of death may take longer to establish. A heavy caseload can delay completion of autopsy reports. A mass-casualty incident, which is a rare event, can cause delays measured in days. Unless special assistance is requested for a mass casualty incident (and such help may itself be delayed by days), the medical examiner's office must deal with whatever the daily demand is with a fixed staff. This means that, on a given day, some bodies, without obvious law-enforcement or immediately discernible public-health priority, may be placed in a freezer compartment until attention can be paid to them. As an example, a year ago, an efficient and experienced mortuary team conducting autopsies of 19 air-crash victims required three days to report all results; this reporting did not include laboratory tests. A medical examiner reporting a jumbo jet crash did not complete the required work for weeks. A caveat here, however, is that such data can usually be recorded and released sooner to law enforcement or public health authorities than to families.

Given an average workload, significant delays are the exception rather than the rule. Whether efficiency can be maintained in the face of a bioterrorist attack depends on the effect on the medical examiner's workload and the medical examiner's ability to quickly deploy additional staff.

Summary

Medical examiners utilize information systems that are similar, in many respects, to point-of-care systems used by physicians. Thus, reliability and data accuracy would be expected to be high. Virtually all large jurisdictions use these systems, so availability should also be high and they constitute a potentially useful source of information for public health. Post-mortem medical examinations, however, are conducted rarely and produce data about an event that is inherently

late (death) using methods that are time consuming (establishing with certainty the cause of death). Mass-casualty incidents can cause significant delays in data entry. The utility of medical examiner data during, or after a bioterrorist attack also depends on the degree to which the information systems are integrated and linked to public health, and the degree to which they can continue functioning efficiently. Thus, medical examiner data seems most helpful in outbreaks that have a long-window of opportunity for intervention, are fatal, and are distinctive enough or large enough to come to post mortem examination. There are many pathogens and outbreaks, for example, which have this characteristic including Hantavirus (where there is a continuous source of exposure that will continue to produce death until it is identified), or person-to-person contagious diseases, or vector borne illness such as West Nile. At the other extreme, an extremely rapidly progressing outbreak such as one caused by a point bioaerosol release would benefit from data that are inherently earlier.

Questions for Further Study

1. How did the recent terrorist incidents involving anthrax affect the daily operations of medical examiner offices? Had the number of cases been larger, and more personnel deployed in a short period of time, how would the timeliness of information been affected?
2. What kinds of public investments are required to provide adequate information system capability to a medical examiner faced with a bioterrorist attack? A large-scale attack?
3. What are the legal or administrative barriers to real-time use of medical examiner data in public health?
4. What are the actual time delays for different types of data collected by medical examiners? What is the accuracy of medical examiner data useful for public health?
5. How do we integrate medical examiner observations about an individual with observations from clinical and other sources about the same individual?

PART III: Case Study and Conclusions

Chapter 17. Pittsburgh Case Study

Public health surveillance depends ultimately on availability of data. This chapter is a study of the availability of data for one large metropolitan area. Our understanding of data availability in this region is not theoretical. It is a result of a 2-year-old project, described below, developing an advanced public health surveillance capability for this region.

Pittsburgh MSA and MMRS

The Pittsburgh Metropolitan Statistical Area (as defined by Office of Management and Budget, 6/30/99) consists of the six counties in South Western Pennsylvania (Table 17.1). The total population of the Pittsburgh MSA on April 2000 was 2,358,695. Pittsburgh, the central city of the MSA has a population of 334,563. The Pennsylvania Region 13 Working Group Metropolitan Response System (MMRS) includes an additional seven counties. The total population of the thirteen counties was 3,008,921 in 2000. The population of each county is given in Table 17.1

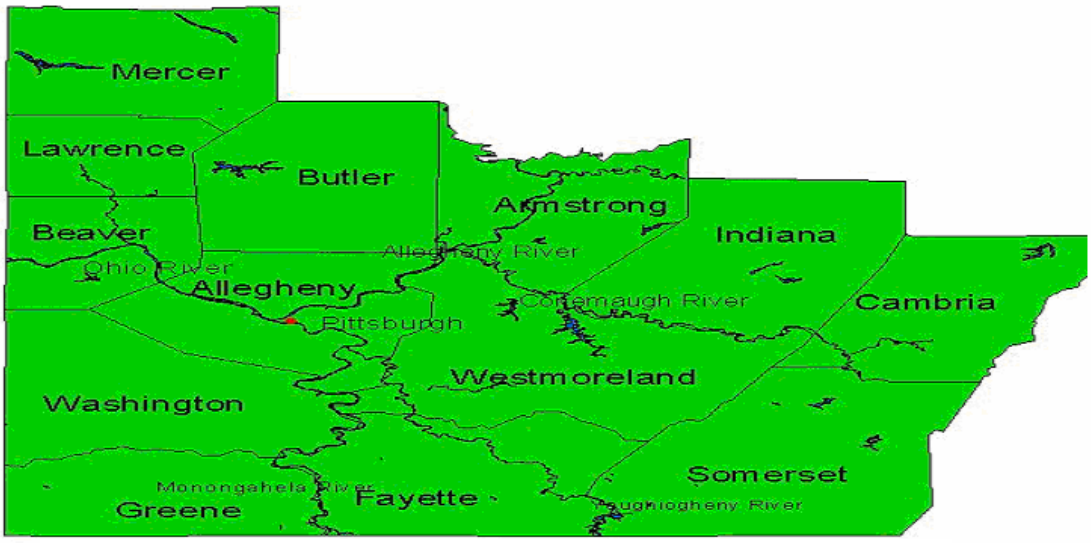


Figure 17.1 Counties (13) in the Metropolitan Medical Response System (MMRS) in southwestern Pennsylvania.

RODS

The Real-time Outbreak and Disease Surveillance (RODS) system is a public health surveillance system deployed in Western Pennsylvania. RODS collects and analyzes data automatically and in real-time. RODS was deployed with the assistance of the local health systems.

RODS receives data from emergency departments and hospitals in the 13-county metropolitan medical response system, centered on Pittsburgh. Figure 17.1 shows how the data provide a picture of the overall volume of patients, as well as of those presenting with chief complaints of diarrhea, rash, respiratory illness and other key symptoms (the term *syndromic surveillance* refers to the practice of monitoring a population for increased incidence of non specific, early presentations of illnesses). RODS—described below—utilizes data that are already being collected routinely during emergency department visits, thus no special (labor-intensive) data collection is necessary. RODS has been in operation for two years and it is developed by a research group at the Center for Biomedical Informatics

At present, RODS, operating under a trusted broker arrangement with the health systems and the health department, receives data from 13 emergency departments. RODS collects microbiology data and other data from 19 hospitals. The percent coverage of regional ED visits is as follows: 52% of the central urban region (population 1.3M), 26% of urban and suburban (population of 2.3 M), and 25% of the region encompassing a total of 13 counties that participate in the Metropolitan Medical Response System (population 3 M).

Although RODS does not at present receive non clinical data, a reason for the detailed analysis in this report is to identify additional useful data sources as a preliminary to addressing the administrative, legal, and technical barriers needed to obtain and integrate such data in real-time.

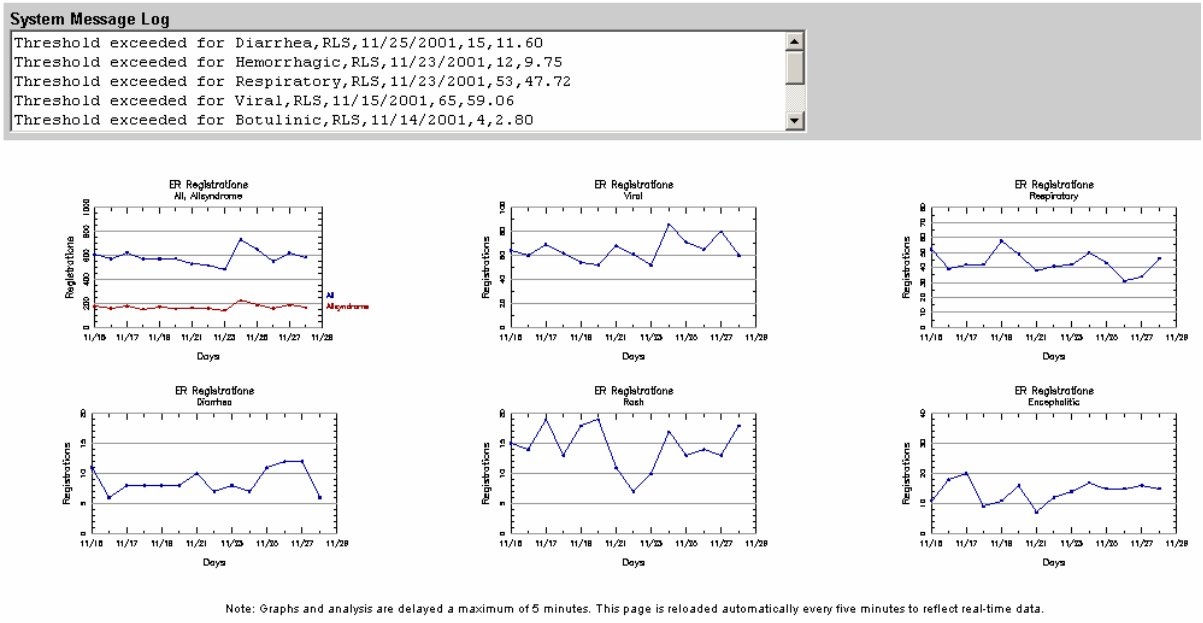


Figure 17.1 RODS Prodrome Monitoring. Shown are daily plots of ED registrations for 10 hospitals. Legend: *All,* all causes; *Viral, Respiratory, Diarrhea, Rash, Encephalitic:* daily registrations with chief complaints of these types.

Providers of Clinical Data

Like most large metropolitan regions, the MMRS region is served by multiple health systems with approximately 65 hospitals, many private offices, long-term care facilities, and insurers. The market has a relatively low penetration of HMOs. In this section, we summarize available clinical information relevant to the detection of bioterrorism.

Emergency Department Registration Systems

Table 17.1 shows the numbers of emergency departments by county for both the metropolitan statistical area (population 2.3 M) and the MMRS area (population approximately 3M). In the MMRS region, there are 54 emergency departments. Most emergency departments use commercial registration systems such as Siemens Invision. Registration data containing chief complaint information is widely available, although the majority of the hospitals encode the chief complaint as free text and others use ICD-9 codes. Most hospitals use HL7 messages to transmit registration information to other systems. Those HL7 messages are processed by HL7 message routers that represent an excellent point of potential connection to public health computers, provided that patient confidentiality and health system proprietary information concerns can be satisfied. We have integrated clinical data in this region and have real-time data feeds of this type from 13/58 emergency departments. The process has been facilitated by the existence of large health systems with multiple emergency departments. The table shows the current level of coverage as percent of emergency department visit records being monitored. Integration of virtually all emergency departments in real-time using HL7 interfaces is likely to be possible, based on information we have obtained in our research. We have also are working with a large application services provider (Siemens Medical Systems) for clinical information systems that has approximately a 30-50% market share in Pennsylvania and will provide additional emergency department registration data feeds.

Table 17.1. Emergency Departments in Pittsburgh MSA and MMRS.

Counties in Pittsburgh MSA and MMRS	Population	Number of EDs	Annual total (7/1/99-6/30/00) ED visits	RODS ED surveillance by county (%)
MSA COUNTIES				
ALLEGHENY COUNTY	1,281,666	22	596,313	52.0%
BEAVER COUNTY	181,412	2	53,105	
BUTLER COUNTY	174,083	2	47,396	
FAYETTE COUNTY	148,644	3	65,851	
WASHINGTON COUNTY	202,897	3	83,636	
WESTMORELAND COUNTY	369,993	6	124,345	
MSA TOTALS (population 2.3M)	2,358,695	**38**	**970,646**	**26.2%**
ADDITIONAL COUNTIES IN MMRS				
ARMSTRONG COUNTY	72,392	1	21,520	
CAMBRIA COUNTY	152,598	3	65,718	44.8%
GREENE COUNTY	40,672	1	13,054	
INDIANA COUNTY	89,605	1	28,708	
LAWRENCE COUNTY	94,643	3	54,024	
MERCER COUNTY	120,293	4	78,424	38.5%
SOMERSET COUNTY	80,023	3	29,969	
MMRS TOTALS (population 3M)		**54**	**1,262,063**	**24.9%**

Source: State of Pennsylvania

Laboratory Information Systems

In the region, laboratory information is available from laboratory information systems in each health system and from commercial laboratory vendors such as Labcorp and Quest. One health system, comprising 19 hospitals, uses a single laboratory information system (Sunquest) and a single data dictionary for microbiology cultures. Additionally, the health system routes all results through a single HL7 message router. This setup provides enormous leverage and is to be exploited wherever it is found. The microbiology HL7 messages are, however, not entirely suitable for usage for public health surveillance and we had to write parsing routines to understand whether a given message was an update, a correction, or a first report. This problem is typical of microbiology outbound interfaces because they are designed to provide a snapshot of the current state of the work-up of a specimen, and they do not identify explicitly for what reason a new snapshot is being sent. It is an open question as to how much work it will be to integrate a new source of microbiology results into the RODS system.

The State of Pennsylvania Bureau of Epidemiology receives daily batches of microbiology and other laboratory results from Quest and LabCorp, two commercial laboratory vendors with large market shares in private doctors offices and specialized testing. The State Department of Health also operates specialized laboratories. These data and systems are not available at present in real-time for public health practice in Western PA, to our knowledge. They are discussed further in the section below on Public Health Systems.

Other laboratory test results of value to public health, such as white blood counts, influenza titers, immunoassays, and cerebrospinal fluid analyses, are available from these laboratory information systems; however, usability is decreased by lack of standard terminologies and data structures in different information systems.

Radiology Reports and other Dictations

In the region, radiology departments generate imaging reports in each hospital system and freestanding radiology practices. The results are dictated typically, and therefore in free-text and are time lagged by the process of transcription. We and others have found that it is feasible to extract information from these reports, especially chest radiographs that contain information relevant to the detection of different types of pneumonias, or of the detection of the radiographic findings of the disease inhalational anthrax. RODS obtains radiology information from 19 hospitals from the HL7 interface. We have still to characterize the numbers of radiology information systems in the region or the issues in their integration.

In clinical medicine, physician notes in the office, emergency department, hospital floor, and intensive care unit can be handwritten in charts or dictated. These notes contain important information about symptoms, signs, and diagnoses. RODS obtains such information from 19 hospitals from the HL7 interface.

Orders

One health system captures orders electronically for microbiology tests (e.g., blood cultures) and other orderable items and routes them through an HL7 message router and RODS captures these

data and uses them for public health surveillance. From discussions with other health systems, we believe such information systems are common in larger settings.

Scheduling, Billing, Pathology, Pharmacy

These types of systems seem to be relatively common in clinical medicine but we have not tried to integrate them into our public health surveillance yet. Scheduling data are of most interest to us because of the potential to have information possibly indicative of early illness.

Data Warehouse

In clinical medicine, data warehousing is not uncommon. In one health system, data warehousing is available on a large-scale real-time basis and real-time interfacing is possible. However, the data are stored in a proprietary format with a proprietary query language that is difficult to use.

Point-of-Care Systems

In one health system, there are multiple point-of-care systems in ICU, ED, and hospital. They contain useful data but are not yet widely used by physicians for purposes other than results review. We do not have detailed information about the other hospitals but we expect they are similar.

Patient Web Portals and Call Centers

In the region, a search of the health systems for web portals revealed no web portals that supported collection of symptom-level data or information useful for public health surveillance. Call centers of different types exist. An example is MedCall, which supports physician-to-physician communication for referrals.

In summary, the most useful clinical data that we have identified, taking into account the requirement for inherent value for early detection as well as availability are registration chief complaints accompanied by demographic information such as age, gender, home zip code. Laboratory data are equally available and useful for public health surveillance.

Sources of Public Health Data

Public health services for Allegheny County (including Pittsburgh) are provided by The Allegheny County Health Department. The Allegheny County Health Department has broad powers delegated through Act 315 of state law. That act specifically says that Local Boards of Health "shall exercise the rule-making power for the prevention of disease, for the prevention and removal of conditions which constitute a menace to health and for the promotion and preservation of the public health generally."

The Allegheny County Health Department is active in public health surveillance and data collection in the area. Health-care providers send data to the Allegheny County Health Department and they are then forwarded to the State Health Department. Data include information about Influenza mortality (on a weekly basis as part of the 121-Cities Influenza Reporting Program) and reportable diseases. The data are entered manually into proprietary systems such as Epi Info 6.0 or into Microsoft Access databases.[3] Health departments also

[3] N.B. The State of Pennsylvania Bureau of Epidemiology has chosen Allegheny County Health Department as a pilot site for its NEDSS implementation. This program will eventually provide an electronic laboratory results

record treatment and contact information, and the results of any epidemic investigations they conduct.

Public health services and information systems for the Pittsburgh MSA and MMRS region other than Allegheny County are provided by the Pennsylvania Department of Health through its Bureau of Community Health Systems, which operates six medical districts (each with a district office), 57 health centers, and three community health pilot sites. The health centers operate under the direction of the district offices, whose staff provides coordination, consultative and administrative support to these delivery units. South Western Pennsylvania falls within the Southwestern District office located in Pittsburgh PA. The Bureau does not have the same degree of funding, staffing or authority as the Allegheny County Health.

Data collected by public health are inherently useful, but difficult to access and relatively late.

Water Supply Data

The Pittsburgh Water and Sewer Authority (PWSA) supplies water to the residents of City of Pittsburgh as well as some of the surrounding municipalities. PWSA draw its water from the Allegheny River upstream of the Highland Park lock and dam. The intakes are situated on the north shore of the river at a point eight miles upstream of the junction of the Allegheny and Monongahela. No ground or well water is used. The Pittsburgh Water and Sewer Authority monitors for constituents in drinking water (on a continuous basis - 365 days per year) according to the Federal and State laws. The PWSA lab is fully certified for water analysis by the Pennsylvania Department of Environmental Protection. The Pittsburgh Water and Sewer Authority tests for contaminants that may be present in the source water prior to treatment. These include: microbial contaminants, inorganic chemical contaminants, pesticides and herbicides, Organic chemical contaminants and radioactive contaminants. Water quality reports are available on the web at **www.pgh2o.com.**

By contrast to Pittsburgh, Fayette County is served by the following water companies: Municipal Authority of the Borough of Belle Vernon; Municipal Authority of Washington Township; Municipal Authority of Perryopolis Borough, Municipal Authority of Newell; Washington Run Water Company; Masontown Borough Municipal Water System; Smithfield Borough Municipal Water System; Springhill Water Company; Fairchance Borough Municipal Water System; Albert Gallatin Municipal Authority; Mountain Water Association; The Southwestern Pennsylvania Water Authority; The Tri-County Joint Municipal Authority; Brownsville Water Company; National Mines-Isabella Braznell Water Company; Redstone Water Company-Redstone Division; Redstone Water Company - Royal-Allison Division; Western Pennsylvania Water Company-Uniontown Division; North Fayette County Municipal Authority; Municipal Authority of Westmoreland County; Western Pennsylvania Water Company-Connellsville District; North Fayette County Municipal Authority; Municipal Authority of Ohiopyle; and the Indian Creek Valley Municipal Authority. The sources of raw water for these water companies are the Monongahela and Youghiogheny Rivers, along with numerous streams and reservoirs**.**

In addition, there are many private sources of water that are certified by the State of PA. This analysis simply identifies the set of entities that one would have to interact with in the region to obtain detailed water supply anatomy for public health purposes.

---

reporting system for reportable diseases, a Web-based console for physicians to report diseases disease, and a communication network for urgent public health notices.

Absenteeism

Table 17.2 shows the top 10 employers in two counties in the Pittsburgh MSA. Note that the largest employers in each county represent but a fraction of the workforce, and an even smaller fraction of the population. We did not analyze in detail which of these employers keep attendance records and for what proportion of the workforce. This table convinced us that absenteeism records only can play a niche role in public health surveillance—limited to building or work place centered outbreaks in those organizations that keep electronic attendance records in real-time.

School attendance (not shown) similarly involves a large numbers of schools and school districts and a relatively small proportion of population covered. The fact that attendance is kept, and that epidemics occur in schools, however, is an argument for further analysis of the problem of collecting school absenteeism information.

Table 17.2 Ten Largest Employers for Allegheny and Beaver Counties

ALLEGHENY COUNTY EMPLOYERS	Employees (est.)	Normalized by county pop, (1,281,666)	BEAVER COUNTY EMPLOYERS	Employees (est.)	Normalized by county pop, (181,412)
UPMC Health System	28,000	2.2%	Zinc Corporation of America	775	0.4%
U.S. Government	20,300	1.6%	Koppel Steel Corporation	715	0.4%
Commonwealth of Pennsylvania	15,800	1.2%	Anchor Hocking Corporation	650	0.4%
West Penn Allegheny Health System	12,615	1.0%	Cutler-Hammer / Eaton Corp.	650	0.4%
U.S. Airways Group Inc.	11,440	0.9%	NOVA Chemicals, Inc.	604	0.3%
Mellon Financial	9,037	0.7%	J&L Speciality Steel, Inc.	567	0.3%
University of Pittsburgh	8,005	0.6%	LTV Steel Company	508	0.3%
PNC Bank Corp.	6,914	0.5%	Veka, Inc.	425	0.2%
Allegheny County	6,344	0.5%	TeleSpectrum	373	0.2%
USX Corp.	5,960	0.5%	McCarls Inc.	350	0.2%
TOTAL	**124,415**	**9.7%**		**5,617**	**3.1%**

Claims and Prescription Data

There are three large payers in the region: UPMC Health System, Blue Cross/Blue Shield and Welfare. The claims information systems in these systems represent a leverage point for obtaining prescription and diagnosis information. The pharmacy benefits managers for these systems represent a real-time point of leverage for obtaining prescription information.

Animal Monitoring

In this section, we describe the organization of animal monitoring in the region. We are still in the process of collecting information about data and data systems that contain information about vectors, and animal health.

In Pennsylvania, there are a number of state-level agencies that are responsible for regulating, monitoring and protecting animal species and their health. These include the Bureau of Animal Health in the Pennsylvania Department of Agriculture, The Pennsylvania Game Commission, Department of Conservation and Natural Resources, Pennsylvania Fish and Boat Commission and Department of Environmental Protection.

The **Bureau of Animal Health** is responsible for the control and eradication of diseases in livestock and poultry that affect human health. It also administers regulatory programs for animal health certification, containment of diseased animals and elimination of disease agents. The Bureau of Animal Health has seven regional veterinarians. The Pittsburgh MSA falls in Region 4 that comprises the following 10 counties: Allegheny, Armstrong, Beaver, Butler, Fayette, Greene, Indiana, Lawrence, Washington, and Westmoreland. These counties are all part of the 13-county MRSA. In addition, the Allegheny County through its Health Department has a vector control program and also has retains the services of a veterinarian. We are in the process of collecting information about data systems used in this bureau.

The second state agency---the **Pennsylvania Game Commission**--maintains a regional office in southwestern Pennsylvania that serves the following 10 counties that do not overlap with the Region 4 counties fully: Allegheny, Armstrong, Beaver, Cambria, Fayette, Greene, Indiana, Somerset, Washington, Westmoreland. We are in the process of collecting information about data systems used in this bureau.

The domestic animal veterinarians in Allegheny County utilize approximately four emergency veterinarian clinics for almost all after-hours care. One of these clinics is jointly owned by 18 area veterinarians. The clinic is only open 8PM to 8AM, weekdays and 24 hours on weekends and holidays. That clinic uses a proprietary veterinarian system Veterinary Specialist System for billing and for record keeping. The system currently only contains billing data (Name of owner, address, phone number and charges for services including all pharmacy products, euthanasia, Class III hospitalization (isolation). The codes were created by this practice and are not standardized across practices. They will be upgrading their current software to a new product Animal Intelligence System offered by the same company Animal Intelligence Software, Inc. found at http://www.animalintelligence.com. The clinic is planning to upgrade and should be able to provide data feeds about after-hour and weekend care. During the day, the information is widely distributed among private practices. Information that should be available includes animals that die (from billing codes for equipment needed), and animals admitted with suspected infections (billing for extra equipment needed for isolation).

Table 17.3. Animals in Pittsburgh MSA and MMRS

Bats	Beaver	Black Bear
Blackbirds, Orioles, Cowbird & Starling	Bobcat	Bobwhite Quail
Canada Goose	Chickadees	Chipmunk
Cottontail Rabbit	Crows & Ravens	Diving Ducks
Eagles and Osprey	Eastern Coyote	Elk
Fisher	Flycatchers	Foxes
Hawks	Heron Family	Mallard

Mice & Voles	Mink/muskrats	Mourning Dove
Opossum	Owls	Porcupine
Puddle Ducks	Raccoon	Ring-Necked Pheasant
River Otter	Ruffed Grouse	Shrews
Squirrels	Striped Skunk	Turkey
Vultures	Varying Hare	Weasels
White-tailed Deer	Woodchuck	Woodcock
Wood Duck	Woodpeckers	

Table 17.4. Counts of Animals in MSA and MMRS (***being completed)

	LIVESTOCK				WILDLIFE		DOMESTIC	
Counties in Pittsburgh MSA and MMRS	**Cattle**	**Sheep**	**Chickens**	**Turkey**	**Deer**	**Bear**	**Dogs**	**Cats**
MSA COUNTIES								
ALLEGHENY COUNTY	2,700	600						
BEAVER COUNTY	9,400	900						
BUTLER COUNTY	21,200	2,200						
FAYETTE COUNTY	20,400	900						
WASHINGTON COUNTY	33,500	7,400						
WESTMORELAND COUNTY	23,300	2,500						
MSA TOTALS	110,500	14,500						
ADDITIONAL COUNTIES IN MMRS								
ARMSTRONG COUNTY	16,200	900						
CAMBRIA COUNTY	12,500	2,600						
GREENE COUNTY	18,000	5,000						
INDIANA COUNTY	21,000	1,300						
LAWRENCE COUNTY	20,500	1,200						
MERCER COUNTY	33,800	2,100						
SOMERSET COUNTY	46,100	2,600						
MMRS TOTALS	278,600	30,200						

Other Regional Data Providers

Table 17.5 summarizes other potential data providers in the region. It includes sources for prescription drug information, over-the-counter drug sales information, claims data/billing, vital statistics, and county coroner's/medical examiners.

Prescription Information

Information about prescriptions can be purchased on a daily basis at present from National Data Corp, but the data is aggregated at the regional level. Sub regional (e.g., neighborhood) data is only available on a weekly basis. The industry is working towards providing new products for public health, but they are not yet available. A single or small number of interfaces that will be able to provide substantial prescription information would involve National Data Corp.'s NDC

Health unit, and IMS Health. For the present, interfaces with many pharmacies and pharmacy chains would be required. As discussed under claims above, pharmacy benefits managers and insurers have real-time data collection systems for pre-authorization of prescriptions and co-payment determinations that can potentially be exploited.

Over-the-counter Drug Sales

Over-the-counter drug information can be obtained from sales tracking firms such as IRI and A. C. Nielsen; they gather information directly from electronic supply chain systems on a weekly basis, but could do so on a daily or even hourly basis, were they paid to do so; some information may also be gleaned from retail tracking and supermarket sales tracking, involving cash register level real-time transaction systems owned by the grocery chains and pharmacy chains.

Poison Center

Poison center calls for the region are centralized in a single poison information center. This model is representative of the situation nationally, and the poison information centers of the country are rapidly moving towards developing real-time services for public health surveillance. There is administrative and technical effort required, but in our experience in this region the effort was not large and the fact that a single poison centered covers Western Pennsylvania is a highly leveraged situation.

Enhanced 911 Electronic Records

The City of Pittsburgh's 911-call center is operated by the city Bureau of Communications, which is an organization separate from the services (police, fire, EMS) its operators dispatch. The center is equipped with Enhanced 911, which includes Automatic Number Identifications (ANI) and Automatic Location Identification (ALI). Computer-Aided Dispatch (CAD) is used as well.

Emergency services in the counties surrounding Pittsburgh are controlled and coordinated by 911 centers in those counties, creating integration problems both for local response as well as for public health secondary uses of the data collected by these systems.

Vital Statistics

The Allegheny County Health Department has been negotiating with the state to receive electronic feeds of birth certificates. We do not know the current status of that request, however many hospitals in the county use the Genesis birth certificate software so in theory electronic feeds of birth certificates are possible. Death certificates are a different issue due to a law that permits delayed submission of certificates. A result is that ACHD often submits late to the 121-Cities Mortality Reporting System.

Table 17.5 Other Potential Regional Data Providers

Regional Data Provider	Number of Systems (if appl.)	Coverage (est.)
OVER-THE-COUNTER DRUGS		
Grocery stores and convenience stores	"pages"	100%
Grocery chains	2	>50%
Pharmacies	"pages"	100%
Pharmacy chains	5	>50%
IRI inc, Chicago IL AC Nielsen inc., Schaumburg IL	2	~70%
PRESCRIPTION DRUGS		
Pharmacies	"pages"	100%
Pharmacy chains	5	>50%
NDC Health, Atlanta GA IMS Health, Blue Bell PA	2	?
Pharmacy benefits managers	Few	?
CALL CENTERS/EMS		
Pittsburgh Poison Center	1	100%
911 Call Centers	14	100%
City Of Pittsburgh Bureau Of Communications E911 ANI ALI CAD Allegheny County Department Of Emergency Services E911 ANI ALI CAD (6 Regional Call Centers) Beaver County Department Of Emergency Services E911 ANI ALI CAD Butler County Department Of Emergency Services E911 ANI ALI CAD Fayette County Emergency Management Agency E911 ANI ALI Greene County Department Of Emergency Services (911 Center Under Renovation) Washington County Department Of Emergency Services E911 ANI ALI CAD Westmoreland County Department Of Public Safety E911 ANI ALI CAD Cambria County Department Of Emergency Services E911 ANI ALI CAD Indiana County Emergency Management Agency E911 ANI ALI CAD Lawrence County Department Of Public Safety E911 ANI ALI CAD Mercer County Department Of Public Safety E911 ANI ALI CAD Somerset County Emergency Management Agency E911 ANI ALI Armstrong County Emergency Management Agency E911 ANI ALI		
VITAL STATISTICS		
Birth (Genesis)	1	?
Death (not electronic)	N/a	0
CORONER/MED EXAMINERS		
?		
CLAIMS DATA/BILLING		
UPMC Health Plan		
BC/BS		
Aetna/US Healthcare Welfare		

"pages" refers to multiple pages of entries in the regional yellow pages

National Data Providers

Satellite Systems/Weather and information retrieval systems that contain data relevant to the Pittsburgh MSA and MMRS include the national weather service, WebMD, and Pub Med.

Summary

This case study provides a sobering look at the numbers of systems that would have to be integrated to obtain all data of potential value for public health. Types of systems that would involve a relatively high amount of system integration relative to data yield are employer attendance systems. At the other extreme, systems with national scope and pre-existing data integration such as weather information represent easy choices for inclusion. Examples of data and systems that involve more decision tension, but nevertheless appear to represent favorable integration to data value ratios include drug sales—both over-the-counter and prescription, and clinical data. Some types of systems or data identified by prior chapters such as sentinel population/physiological monitoring were not represented in the region, and some types of data such as symptom and sign data are available only on limited bases, would represent enormous investment to obtain, but are probably worthwhile pursuing nevertheless.

The most obvious overall conclusion to be drawn is the need for a strategy to integrate these data sources—a strategy that triages and prioritizes the effort to develop interfaces at both the regional and national levels. In particular, some effort to optimize the balance between the value of the data and the numbers of systems to be integrated is required. The next chapter attempts to summarize the results of this report in this manner.

Chapter 18. Summary and Conclusions

The complexity of the analysis undertaken in this report derives in part from the numbers of diseases that represent public health threats, and, for each disease, the range of outbreaks that can occur. The spectrum ranges from single cases to large-scale outbreaks, from sudden outbreaks with exponential growth rates to outbreaks with time latencies measured in years. The spectrum also includes outbreaks involving transmission via air, food, water, person, and intermediate hosts. Also complicating the analysis is lack of real experience with many of the threats because they have not occurred in recent times, hence their effects on routinely collected data cannot be assessed directly. The analysis is also complicated by its attempt to include types of data for which detection analytic methods are still being developed. Thus, the conclusions that we draw in this chapter will be as general as possible but no doubt some reader's favorite disease will represent an exception to anything we say. Thus, we ask that anytime such an exception is identified, that the reader add a qualifying clause to the wording of our conclusion that excepts that disease, and then examines whether the more restricted conclusion is still valid, and more to the point, whether it contributes helpful insights or direction to the field.

The importance of data availability adds to the complexity of an analysis that would be complex enough if the goal were simply to identify which data would be required to achieve complete and early detection of public health threats.

Thus the following set of conclusions attempt to be a bottom-line analysis, integrating over issues of availability, the range of diseases, and assuming that what is desired ultimately by readers of this report is the highest distillation of available information into a set of recommendations about which data the country should pursue for public health surveillance.

Table 18.1 summarizes our findings for different types of data, organized in order of timeliness.

For Detection of Outbreaks that Present with Many Cases

Initial detection of an outbreak that presents with many cases (as in the case of some food poisonings or bioaerosol releases of pathogens) ideally should occur when individuals are still

very early in their illness, before the illness is sufficiently well-defined that the diagnosis can be made from information collected about a single person. For outbreaks that present with many cases, strategies should exploit either the sheer numbers or rate of increase in incidence of disease, and or epidemiological correlations among the cases.

Physiological data are among the earliest data covered in this report (biosensors and intelligence data are earlier, but they are outside our scope). The value of these data for early detection of outbreaks has not been well studied and these data are not widely available. Therefore for the next few years they are likely to play at best niche roles such as protection of key resources such as the president (by monitoring the health of bodyguards and other staff). Additional research is needed to explore this approach and such research should be a priority.

Information seeking by sick individuals using telephone or the Internet is a behavior that can, in theory, be inferred from examination of existing data sources. This approach has not been applied to public health surveillance, to our knowledge. We have examined WebMD data and found suggestion that query words can potentially provide indirect evidence of sickness (see Figure 2.1. This approach has obvious potential because of real-time availability, good coverage of both the population and the set of threats, and the relatively small number of systems that would have to be integrated to achieve national surveillance. Additional research is needed to explore this approach should be a priority.

Self-treatment by sick individuals in the form of over-the-counter-drug purchases are highly available in real time. The numbers of systems that would have to be integrated to achieve regional or national surveillance is relatively small because the numbers of pharmacies and supermarkets that would have to be integrated is comparatively small, as suggested by our Pittsburgh analysis. The evidence that such data are of value is better than for many types of data. We know, for example, that these data show effects of Influenza epidemic and that the signal is early. Some limitations include blindness during holidays and when stores are closed; and diseases presenting with weakness do not have OTC remedies. Development of real-time access to this type of data should be a priority.

Animal health data is a broad heterogeneous category. Such data can provide information about pathogens that can be transmitted from animals to people as well as illnesses that are not transmissible, but to which both animals and humans are susceptible. In the latter case, the value of animals is as canaries, so to speak. Depending on the disease, animal data can be early or late; our decision here to discuss its value as an early indicator is motivated by threats such as West Nile and anthrax for which it provides an early signal. For the majority of threats, needed data are not being collected routinely at present, or are being collected late by a large number of systems whose integration poses significant barriers. Many open research questions were identified by our consideration of the question *How can animal health data be used optimally to build early warning systems for public health surveillance?* We conclude that, for a few selected diseases, the level of maturity of information systems provides a solid base for development of early detection systems.

Work or school attendance information, although an early indicator of disease, are only recorded for a small fraction of the population. Moreover, such data are affected by weekends, holidays, work vacations and school vacations. The data is significantly time delayed and the number of systems to be integrated is large. Indirect measures of attendance indicated by records of access to parking facilities, buildings, or computing resources have potential either to provide

absenteeism information about a sample of the overall population, or--in facilities with card-restricted access--complete monitoring. Research on how to measure absenteeism indirectly is needed. Indirect (or even direct) measures of absenteeism can and should play a key, niche role in building-level surveillance.

Call centers such as 911 or poison centers collect data about sick individuals in real-time. They are available 24 hours per day. Call centers cover metropolitan or county-size regions and their information systems should be amenable to integration. In the case of poison centers, significant national aggregation for the purposes of public-health surveillance is already underway. The value of these data for detection of outbreaks is largely anecdotal and requires further study to delineate both the outbreaks that can be identified as well as the analytic methods to apply to such data. Limitations are that these data will not be sensitive to diseases that do not trigger calls. In most cases, sick individuals are advised to seek medical attention, thus there are questions of overlap with signals coming from clinical sources. Relevant additional questions include what percent of individuals who have a consultation with a call center are advised to seek treatment, and of those, what percent seek treatment.

Clinical data, like animal health data, is a heterogeneous category with some high availability of some types of data and low availability of others. Registration data, including reason for visit, are among the most widely available data being collected in real-time 24 hours per day by the majority of healthcare systems in the U.S. The value of these data has been partly established. Health system registration systems, with the exception of the VA, large HMOs, and possibly the military, cover sub-regional areas; thus, the number of systems to integrate is large. Leverage point may be the relatively small number of vendors of electronic medical record software products that could be induced to build into their product the technical means to achieve data integration. Development of real-time access to this type of data should be a priority.

Laboratory test results including microbiology cultures are similarly available in real-time. A few commercial laboratories with national scope service smaller hospitals and practices. These companies represent a resource that can reduce the number of interfaces that need to be created. The value of these data in public health is well accepted. A barrier to such integration, identified by the Pittsburgh case study, is the complexity of microbiology data and the lack of standards. Limitations are that laboratory data are relatively late. Development of real-time access to this type of data should be a priority.

There are many key data discussed in the Gap Analysis below that are not being collected by clinical systems, or are being collected in forms that are not amenable to computer processing. Examples include symptoms and vital signs. The value of these data for public health surveillance is well accepted and additional research to identify ways to obtain these types of information pre-clinically should be a priority.

Prescription drug sales have high real-time availability due to existing inventory tracking systems, pharmacy benefits managers etc. These resources will make national integration of prescriptions relatively easy to accomplish. Such data are known to show antidiarrheal and antibiotic sales peaks prior to public health detection of diarrheal outbreaks. Population coverage is moderate because not every person sees a physician. Threat coverage requires additional research although it is safe to say that many threats would be such that they would result in treatment by physicians with prescriptions. Their full potential for the range of threats

is a subject of future research. Development of real-time access to this type of data should be a priority.

In general, earliness of detection can be boosted using **epidemiogical information** including location of patients over recent time, food and water consumption, physical contacts, and even something as crude as international travel history. These data are unfortunately not fully available from secondary sources. Home address, work address, and derived from them, family contacts and work contacts may be either recorded directly or can be reconstructed from widely available data, however, such information is considered sensitive by many and therefore there are barriers to its use or reconstruction from existing sources. International travel history can be partly derived from passport control. Water consumption can partly be inferred from home and work address. Food consumption, however, is very difficult to ascertain from secondary sources of data, although credit card receipts and grocery purchases using advantage cards can provide partial information. Many of these data sources do not involve high numbers of interfaces, so there is a potential for tractability if privacy issues can be addressed.

Relevant **information about food and water supplies** needed for early detection also include anatomy of supply chains--needed to trace back from affected individuals to points of common exposure. Additionally data about any known or possible contaminations are needed to trace forward to affected individuals. Unfortunately, the required data are collected by multiple government agencies, private and semi private firms. The numbers of systems that would have to be integrated are large. The utility of food and water data is well established by their use in public health surveillance. Methods to automatically interpret those data are an area for additional research.

Weather and data available from satellites are almost ideal from the perspective of availability and national coverage--in most cases being integrated to the point of being available from a single, open source via the Internet. They have proven utility in predicting selected outbreaks. This is a type of data that represents a resource for public health and does not seem to require additional work to develop as a data feed (although considerable research on analytic approaches to combining weather and satellite data with other data will be required).

For Detection of Outbreaks in Which Cases Accumulate More Slowly

The previously discussed data can also be used to detect slowly growing outbreaks once the numbers of cases have become large, however early detection depends on noticing an early case, which requires relatively accurate information. The information can be either a diagnosis by a clinician, specific culture results, pathognomic findings, or less specific findings that when combined become distinctive. At the present time, most of the necessary data—with the exception of cultures--are not widely available in electronic form in real-time, although they are often available in a delayed but still relevant time frame.

The current **public health system** collects laboratory data, submitted case reports, special purpose surveillance data, and the results of active investigations. These data are highly relevant to public health surveillance. However the types of data are inherently late and they are also subject to delays in reporting. To improve this situation, a large number of public health systems that would have to be integrated that use diverse technical approaches and data representation standard. Population coverage is not good because of underreporting. We discussed public

health systems in detail in the previous report and provide suggestions for their improvement. The National Electronic Disease Surveillance System (NEDSS) project is focusing on improving the timeliness and completeness of reporting of such data. Deployment of NEDSS should be a priority.

Claims data/billing contain information about health services delivered. The information content has overlap with clinical systems and prescription drug systems. The value of claims data for outbreak detection has not been well demonstrated. For medical claims, data warehouses provide a fair amount of data integration to generate regional or national data sets, however claims data is inherently extremely delayed in time, thus the utility of claims data is for the detection of public health threats that have very long time latency and a very wide window of opportunity for mitigation.

Death certificate data have proven surprisingly difficult to obtain electronically with reasonable time latencies. Additional key limitation of vital statistics and County Coroner's Office/Medical Examiner results are the inherent lateness of the health event of death. At present, both types of data involve considerable additional time latencies that could be reduced by administrative and information technology measures. Although the fact of death is established with accuracy, the cause of death may not be. Population coverage for the fact of death is nearly 100%; by contrast, autopsies that accurately establish the cause of death are rarely performed, although they are more likely to be performed in cases of unexpected death in previously healthy people. These types of data obviously are useful for fatal diseases, and of those diseases, the types of outbreaks are limited to those that are relatively protracted such that they are still mitigable at the time they are detectable by death information. Research to identify how to use these data fully is a priority. Development of real-time access to this type of data should be a priority.

For Characterizing an Outbreak

Characterization of an outbreak means going beyond the recognition of an anomalous pattern to elucidate the route of transmission, the causative organism, and the specifics of the outbreak including other potentially affected individuals. The data needed to do this include the more specific types of clinical information just discussed, epidemiological data just discussed, and results of environmental testing. The latter two types of data are typically collected by public health, so are subject to delays inherent in public health information processing.

Table 18.1 Comparative Value of Required Data Elements

Type of Data	Availability	Number of Interfaces	Coverage of Population	Limitations	Recommendation
PHYSIOLOGICAL MONITORING	Poor	?	Poor		Research, Deploy for special cases
ANIMAL MONITORING		Many			
Agribusiness	Good	Many	Good*		Develop real-time access
Veterinarian	Poor	Many			Deploy EVR
Vector monitoring	?				?
INFORMATION SEEKING					
Phone to MD	Good	Few	Good	Not clear how to analyze	Research
Web query	Good	Few	Good		Research
ATTENDANCE					
Attendance records	Poor	Many	Poor	Nights and holidays	Deploy for building monitoring
Indirect evidence of attendance	?	Many	Poor	Nights and holidays	Deploy for building monitoring
CALL CENTERS					
911	Fair	Few	Good		Develop real-time access
Poison Centers	Good	Few	Good		Develop real-time access
CLINICAL DATA					
Reason for visit	Good	Fair**	Fair**		Develop real-time access
Symptoms	Poor	Fair**	Fair**		Deploy EMRs
Signs	Poor	Fair**	Fair**		Deploy EMRs
Diagnoses	Poor	Fair**	Fair**		Deploy EMRs
Tests including micro	Good	Fair**	Fair**		Develop real-time access
Orders	Fair	Fair**	Fair**		Develop real-time access
CLAIMS DATA	Good	Good	Fair**	Very late	Develop real-time access
PUBLIC HEALTH					
Mandatory reporting	Good	Good	Underreporting	Very late	Deploy NEDSS
Laboratory reporting	Good	Good	Underreporting	Very late	Deploy NEDSS
PRESCRIPTIONS	Good	Good	Moderate	Nights and holidays	Develop real-time access
OTC DRUG PURCHASES	Good	Fair	Good	Not all diseases have OTC treatments. Nights and holidays. People may have in stock	Develop real-time access
DEATH					
Vital statistics	Poor	?	Good	Very late	Develop real-time access
Autopsy	Fair	?	Few autopsies	Late	Develop real-time access
FOOD WATER SUPPLY					
Food	Poor	?	?		
Water	Good	Many	Good	Water is low risk	

**not everyone sees doctors

Gaps

The biggest gaps in existing data are early information about the symptoms and signs that patients are experiencing and epidemiological data. Sick people do not seek medical attention promptly, if at all. There is undoubtedly indirect evidence of their illness available in records of purchases, utility usage at home or work, web queries and other secondary data sources, but all of those sources are marked by noise and it is an open question how accurately the health status of an individual can be determined even with good integration of all such sources. Research to develop methods whereby ill individuals communicate their illness to health care or public health early should be a top priority.

Epidemiological data needed to enhance earliness of outbreak recognition as well as its characterization are also very difficult to obtain from secondary data sources. Research to develop methods whereby epidemiological information can be obtained quickly about the sick should also be a top research priority.

Conclusions

The conclusions from this study are that near-term improvement in detection of outbreaks that presents with many cases (e.g., bioaerosol releases) should be pursued by developing real-time access to clinical and veterinary information systems, call center data, poison center data, and over-the-counter drug sales data. Research and development on ways to detect health-information seeking behavior, or to encourage ill individuals to self-report could provide results that could quickly translate into improved capabilities. Physiological monitoring and absenteeism data (direct or indirect) may provide improved detection in limited settings such as facility protection.

Longer-term improvement in the ability for early detection of outbreaks that presents with many cases will be facilitated by deployment of point-of-care clinical information systems, veterinary electronic medical records, and physiological monitoring gear. Additionally, research on methods to obtain epidemiological information more quickly can be expected to result in additional improvements in detection. Deployment of interoperable systems adherent to NEDSS standards will also improve detection of widely distributed outbreaks that cross jurisdictions.

Near-term improvement in the detection of small, or slowly growing outbreaks can be attained by developing better real-time access to pathology systems, claims data, death records, and autopsy results.

Longer-term improvement in the detection of small, or slowly growing outbreaks can be facilitated by research in natural language processing of imaging reports, and obituaries.

Glossary

AAPCC TESS: American Association of Poison Control Centers Toxic Exposure Surveillance System
ABCs: Active Bacterial Core Surveillance
AER: Adverse Event Reporting
AHRQ: Agency for Healthcare Research and Quality
ALI: Automatic Location Identification
ANI: Automatic Number Identification
APHIS: Animal and Plant Health Inspection Services
ASD: Adult/Adolescent Spectrum of HIV Disease
ASTHO: Association of State and Territorial Health Officials
ATM: Automatic Teller Machines

BIDS: Border Infectious Disease Surveillance Project
BDU: Battle Dress Uniforms

CAD: Computer-Aided Dispatch
CDPD: Cellular Digital Packet Data
CDC: Centers for Disease Control and Prevention
CSTE: Council of State and Territorial Epidemiologists

DARPA: Defense Advanced Research Projects Agency
DAWN: Drug Abuse Warning Network
DES: Data Encryption Standard
DoD: Department of Defense
DTRA: Defense Threat Reduction Agency

ED Systems: Emergency Department Systems
EIN: Emerging Infections Network

EMRs: Electronic Medical Record System
EMERGEncy ID NET: Emergency Department-Based Emerging Infections Sentinel Network
EPA: Environmental Protection Agency
EVR: Electronic Veterinary Records
E911: Enhanced 911

FBDO: Foodborne- Disease Outbreak Surveillance System
FBI: Federal Bureau of Investigation
FDA: Food and Drug Administration
FoodNet: Foodborne Diseases Active Surveillance Network
FSIS: Food Safety and Inspection Service

GIS: Geographic Information Systems
GISP: Gonococcal Isolate Surveillance Project
GRITS: Georgia Registry of Immunization Transactions & Services
GULC: Georgetown University Law Center

HL7: Health Level Seven
HMO: Health Maintenance Organization
HSS: Hemophilia Surveillance System

ICD: International Classification of Diseases
IDSA: Infectious Diseases Society of America
IND: Investigational New Drug
IRB: Institutional Review Board
ISAM: Indexed Sequential Access Method
ISIS: International Species Identification System

LOS: Line of Sight

ME: Medical Examiner
MDT: Mobile Data Terminals
MMRS: Metropolitan Medical Response System
MUMPS:?

NACA: National Advisory Committee on Aeronautics
NAPHSIS: National Association for Public Health Statistics & Information Systems
NASA: National Aeronautics and Space Administration
NATO: North Atlantic Treaty Organization
NCIC: National Crime Information Center
NCOD: National Drinking Water Contaminant Occurrence Database
NCSL: National Conference of State Legislatures
NEDSS: National Electronic Disease Surveillance System
NETSS: National Notifiable Diseases Surveillance System
NHIS: National Health Interview Survey
NHSDA National Household Survey of Drug Abuse

NIOSH: National Institute for Occupation Safety & Health
NIS: National Immunization Survey
NJDHSS: New Jersey's Department of Health and Senior Services
NMSS: National Malaria Surveillance System
NREVSS: National Respiratory and Enteric Virus Surveillance System
NSFG: National Survey of Family Growth Surveillance
NSSS: National Salmonella Surveillance System
NVSL: National Veterinary Services Laboratories

PHLIS: Public Health Laboratory Information System
PPO: Preferred Provider Organization
PPQ: Plant Protection and Quarantine

PSAP: Primary Service Answering Point
PWSA: Pittsburgh Water and Sewer Authority
.
RMS: Record Management Systems
RODS: Real-time Outbreak and Disease Surveillance
RSVP: Rapid Syndrome Validation Project

SCIC: State Crime Information Centers
SEER: Epidemiology and End Results
SENDSS: Statewide Electronic Notifiable Disease Surveillance System
SENSOR: Sentinel Event Notification Systems for Occupational Risks
SODA: Salmonella Outbreak Detection Algorithm

SPCA: Society for Prevention of Cruelty to Animals
SQL: Structured Query Language
SWTR: ?

TSN: The Surveillance Network

UPMC: University of Pittsburgh Medical Center
USDA: United State Department of Agriculture

VIN: Vehicle Identification Number
VHSP: Viral Hepatitis Surveillance Program
VS: Veterinary Services
VSAM: Virtual Sequential Access Method

WIC: Women Infant & Children

References

1. **Quenel P, Dab W, Hannoun C, Cohen JM. Sensitivity, specificity and predictive values of health service based indicators for the surveillance of influenza A epidemics. *International Journal of Epidemiology.* 1994;23(4):849-855.**
2. **Tsui F-C, Wagner MM, Dato V, Chang C-CH. Value of ICD-9-coded chief complaints for detection of epidemics. March 8, 2001 2001:711-715.**
3. **Kaufmann A, Meltzer M, Schmid G. The economic impact of a bioterrorist attack: Are prevention and postattack intervention programs justifiable? *Emerging Infectious Diseases.* 1997;3(2):83-94.**
4. **Wagner M, Tsui F-C, Espino J, et al. The emerging science of very early detection of disease outbreaks. *Journal of Public Health Management and Practice.* 2001;6(6):50-58.**
5. **Meltzer ML, Damon I, LeDuc JW, Millar JD. Modeling potential responses to Smallpox as a bioterrorist weapon. *EID.* 2001;7(6):959-969.**
6. **Update: Investigation of Bioterrorism-Related Anthrax and Interim Guidelines for Clinical Evaluation of Persons with Possible Anthrax. *MMWR - Morbidity & Mortality Weekly Report.* 2001(50):582-585.**
7. **Lober W, Bryant, T., Wagner, M., Overhage, J.M., Espino, J., and Tsui, F. Roundtable on Bioterrorism Detection: Information System-based Surveillance. *JAMIA (under review).* 2001.**
8. **Dato V, Wagner M, Allswede M, Aryel R, Fapohunda A. The nation's current capacity for the early detection of public health threats including bioterrorism. June 8, 2001.**
9. **Zeng X, Wagner MM. Modeling the effects of epidemics on routinely collected data. *JAMIA-Symposium Supplement.* 2001:781-785.**
10. **Quality FbtAfHRa. The nation's current capacity for the early detection of public health threats including bioterrorism. June 8, 2001.**
11. **Espino J, Wagner M. The accuracy of ICD-9 coded chief complaints for detection of acute respiratory illness. *JAMIA Symposium Issue.* 2001:164-168.**
12. **Lober W, Bryant T, Wagner M, et al. Roundtable on Bioterrorism Detection: Information System-based Surveillance. *JAMIA.* 2001;in press.**

13. Effler P, Ching-Lee M, Bogard A, Man-Cheng L, Nekomoto T, Jernigan D. Statewide system of electronic notifiable disease reporting from clinical laboratories: comparing automated reporting with conventional methods. *JAMA.* 1999;282(19):1845-1850.

14. Panackal AA, M'ikanatha NM, Tsui F-C, et al. *Evaluation of automatic electronic laboratory reporting of notifiable infectious diseases at a large health system.* Pittsburgh: University of Pittsburgh; 2001.

15. Jain NL, Knirsch CA, Friedman C, Hripcsak G. Identification of suspected tuberculosis patients based on natural language processing of chest radiograph reports. *Proceedings / AMIA Annual Fall Symposium.* 1996:542-546.

16. Fiszman M, Chapman W, Aronsky D, Evans R, Haug P. Automatic detection of acute bacterial pneumonia from chest X-ray reports. *Journal of the American Medical Informatics Association.* 2000;7(6):593-604.

17. Chapman W, Haug P. Comparing expert systems for identifying chest x-ray reports that support pneumonia. *Proceedings / AMIA Annual Symposium.* 1999:216-220.

18. Hripcsak G, Knirsch C, Jain N, Stazesky RJ, Pablos-Mendez A, Fulmer T. A health information network for managing innercity tuberculosis: bridging clinical care, public health, and home care. *Computers & Biomedical Research.* 1999;32(1):67-76.

19. Knirsch C, Jain N, Pablos-Mendez A, Friedman C, Hripcsak G. Respiratory isolation of tuberculosis patients using clinical guidelines and an automated clinical decision support system. *Infection Control & Hospital Epidemiology.* 1998;19(2):94-100.

20. Tierney WM, Miller ME, Overhage JM, McDonald CJ. Physician inpatient order writing on microcomputer workstations: Effects on resource utilization. *JAMA.* 1993;269(3):379-373.

21. Sittig DF, Stead WW. Computer-based physician order entry: the state of the art. *Journal of the American Medical Informatics Association.* 1994;1(2):108-123.

22. Bates DW, Boyle DL, Teich JM. Impact of computerized physician order entry on physician time. *Proceedings - the Annual Symposium on Computer Applications in Medical Care.* 1994:996.

23. McDonald CJ, Tierney WM, Overhage JM, et al. The Regenstrief Medical Record System--experience with MD order entry and community-wide extensions. *Proceedings - the Annual Symposium on Computer Applications in Medical Care.* 1994:1059.

24. Hripcsak G, Clayton P, Jenders R, Cimino J, Johnson S. Design of a clinical event monitor. *Comp. Biomed. Res.* 1996;29:194-221.

25. Wagner M, Pankaskie M, Hogan W, Tsui F, Vries J. Clinical event monitoring at the University of Pittsburgh. *Journal of the American Medical Informatics Association-Symposium supplement.* 1997:188-192.

26. Hripcsak G, Knirsch C, Jain N, Pablos-Mendez A. Automated tuberculosis detection. *Journal of the American Medical Informatics Association.* 1997;4(5):376-381.

27. Wagner M, Espino J, Tsui F, et al. *The role of clinical event monitors in public health surveillance.* Pittsburgh: Center for Biomedical Informatics; 2001.

28. Wagner MM, Hogan WR. Capturing medication data in an outpatient electronic medical record. Paper presented at: AMIA Spring Congress, 1995; Boston, Mass.

29. Hogan WR, Wagner MM. The accuracy of data in clinical information systems: A structured review of the literature. February, 1996 1996.

30. Litovitz TL, Klein-Schwartz W, White S, et al. 2000 Annual report of the American Association of Poison Control Centers Toxic Exposure Surveillance System. *Am J Emerg Med.* Sep 2001;19(5):337-395.

31. **Snow J. *Snow on cholera, being a reprint of two papers by John Snow, M.D., together with a biographical memoir by B.W. Richardson, M.D., and an introduction by Wade Hampton Frost, M.D.* New York: The Commonwealth Fund; 1936.**
32. **NETSS National Electronic Telecommunications System for Surveillance. Available at: http://www.cdc.gov/epo/dphsi/netss.htm.**
33. **Salmonella Outbreak Detection Algorithm (SODA). Available at: http://www.cdc.gov/ncidod/dbmd/phlisdata/default.htm.**
34. **National Malaria Surveillance System (NMSS). Available at: "MMWR February 26, 1999 / 48(SS-1);1-21;http://www.cdc.gov/mmwr/preview/mmwrhtml/00056518.htm".**
35. **121 Cities Mortality Reporting System. Available at: http://www.cdc.gov/drugresistance/technical/surveillance.htm http://wonder.cdc.gov/mmwr/about.htm.**
36. **United States Influenza Sentinel Physicians Surveillance Network. Available at: "http://www.cdc.gov/ncidod/diseases/flu/flusurv.htm,http://www.cdc.gov/drugresistance/technical/surveillance.htm htmhttp://www.cdc.gov/ncidod/diseases/flu/weekly.htm".**
37. **National Electronic Disease Surveillance System (NEDSS). Available at: http://www.cdc.gov/od/hissb/docs.htm#nedss.**
38. **CDC. Surveillance Information.**
39. **Brown W. *Epidemiologic capacity of state health agencies: A guide for assessment.* Atlanta: Council of State and Territorial Epidemiologists; January 23, 1998 1998.**
40. **CDC. Updated Guidelines for Evaluating Public Health Surveillance Systems, Recommendations from the Guidelines Working Group,. *MMWR Recommendations and Reports.* July 27, 2001 2001;50(RR-13):1-36.**
41. **Traceback Investigation.**
42. **GEIS. GEIS Rift Valley Fever Monitor. Available at: http://www.geis.ha.osd.mil/riftvalleyfever/index.htm. Accessed May 10, 2001.**
43. **National Death Certificate. Available at: http://www.cdc.gov/nchs/about/major/dvs/vital_certs_rev/death-w-ins.pdf, 2001.**
44. **Birth Certificate. Available at: http://www.cdc.gov/nchs/about/major/dvs/vital_certs_rev/birth1ac.pdf, 2001.**
45. **Birth/Death Certificate. Available at: http://www.cdc.gov/nchs/about/major/dvs/vital_certs_rev/specs-birth-death-a.pdf, 2001.**
46. **Lee PN. Comparison of autopsy, clinical and death certificate diagnosis with particular reference to lung cancer. A review of the published data. *APMIS Suppl.* 1994;45:1-42.**
47. **Kircher T, Nelson J, Burdo H. The autopsy as a measure of accuracy of the death certificate. *N Engl J Med.* Nov 14 1985;313(20):1263-1269.**
48. **U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES Public Health Service Centers for Disease Control and Prevention National Center for Health Statistics Hyattsville M. Model State Vital Statistics Act and Regulations From the 1992 Revision. *DHHS Publication.* February 1994:No. (PHS) 94-1115.**
49. **Background on Electronic Death Registration.**
50. **Lomax EC, Lowe, H.J. Information needs research in the era of the digital medical library. *Proc AMIA Annu Fall Symp.* 1998:658 - 662.**
51. **McKnight M, Peet, M.,. Health care providers' information seeking: recent research. *Med Ref Ser Q.* Summer 2000;19(2):27 - 50.**
52. **Navarro FH, Wilkins S.T. A new perspective on consumer health Web use: "valuegraphic" profiles of health information seekers. *Managed Care Q.* 2001;9(2):35 - 43.**

53. **Eysenbach G, Diepgen, T.L. Patients looking for information on the Internet and seeking teleadvice: motivation, exceptions, and misconceptions as expressed in emails sent to physicians. *Arch Dermatol.* Feb 1999;135(2):151 - 156.**
54. **Pennbridge J, Moya, R., Rodrigues, L. Questionnaire survey of California consumers' use and rating of sources of health care information including the Internet. *West J Med.* Nov - Dec 1999;171((5 - 6)):302 - 305.**
55. **Gollop CJ. Health information-seeking behavior and older African American women. *Bull Med Libr Assoc.* 1997;85(2):141 - 146.**
56. **Gore P, Madhavan S, McClung G, Riley D. Consumer involvement in nonprescription medicine purchase decisions. *J Health Care Mark.* Summer 1994;14(2):16-23.**
57. **Saxonov S, Daizadeh I, Fedorov A, Gilbert W. EID: the Exon-Intron Database-an exhaustive database of protein-coding intron-containing genes. *Nucleic Acids Res.* Jan 1 2000;28(1):185-190.**
58. **GEIS Rift Valley Fever Monitor. Available at: http://www.geis.ha.osd.mil/riftvalleyfever/index.htm.**
59. **Georgia Registry of Immunization Transactions & Services (GRITS). Available at: http://www.ph.dhr.state.ga.us/inpho/statewide/index.shtml.**
60. **Global Outbreak Alert and Response Network. Available at: http://www.who.int/emc.**
61. **ProMed. Available at: http://www.cdc.gov/ncidod/eid/vol4no3/kay.htm, http://www.promedmail.org/pls///promed/promed.homepromed.**
62. **Global Emerging Infections Sentinel Network (GeoSentinel). Available at: http://www.istm.org/.**
63. **GPHIN Global Public Health Information Network. Available at: http://www.carleton.ca/Capital_News/12031999/f4.htm;http://gphin-rmisp.hc-sc.gc.ca/index.html.**
64. **WHONET: an information system for monitoring antimicrobial resistance. Available at: " Emerging Infectious Diseases. 1(2):66, 1995 Apr-Jun;http://www.who.int/emc/WHONET/WHONET.html".**
65. **The Surveillance Network (TSN)TMDatabase. Available at: "http://www.mrlpharmaceutical.com/,http://www8.techmall.com/techdocs/TS990913-8.html".**
66. **Surveillance for Emerging Antimicrobial Resistance Connected to Healthcare (SEARCH). Available at: http://www.cdc.gov/ncidod/hip/ARESIST/SEARCH.htm.**
67. **National Antimicrobial Resistance Monitoring System: Enteric Bacteria (NARMS:EB). Available at: http://www.cdc.gov/ncidod/dbmd/narms/.**
68. **INSPEAR- International Network for the Study and Prevention of Emerging Antimicrobial Resistance. Available at: http://www.cdc.gov/ncidod/hip/SURVEILL/inspear.HTM.**
69. **"National Respiratory and Enteric Virus Surveillance System (NREVSS),". Available at: http://www.cdc.gov/ncidod/dvrd/nrevss/.**
70. **National West Nile Virus (WNV) Surveillance System. Available at: http://www.cdc.gov/ncidod/dvbid/westnile/surv&control.htm.**
71. **U.S. Department of Agriculture National Veterinary Services Laboratories (NVSL). Available at: http://www.aphis.usda.gov/vs/nvsl/.**
72. **Tillett HE, de Louvois J, Wall PG. Surveillance of outbreaks of waterborne infectious disease: categorizing levels of evidence. *Epidemiology & Infection.* 1998;120(1):37-42.**

73. **EMERGEncy ID NET:An Emergency Depart.-Based Emerging Infections Sentinel Network. Available at: " Talan: Ann Emerg Med, Volume 32(6).December 1998.703-711".**
74. **National Surveillance System for Health Care Workers. Available at: http://www.cdc.gov/ncidod/hip/SURVEILL/nash.htm.**
75. **Anonymous. Carbon monoxide poisoning associated with use of LPG-powered (propane) forklifts in industrial settings--Iowa, 1998. *MMWR Morb Mortal Wkly Rep.* Dec 17 1999;48(49):1121-1124.**
76. **Anonymous. From the Centers for Disease Control and Prevention. Carbon monoxide poisoning deaths associated with camping--Georgia, March 1999. *JAMA.* 1999;282(14):1326.**
77. **Anonymous. Houseboat-associated Carbon Monoxide Poisonings on lake Powell --- Arizona and Utah, 2000,. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49:1105-1108.**
78. **Anonymous. From the Centers for Disease Control and Prevention. Houseboat-associated carbon monoxide poisonings on Lake Powell--Arizona and Utah, 2000. *JAMA.* 2001;285(5):530-532.**
79. **Anonymous. Cluster of HIV-positive young women--New York, 1997-1998. *MMWR - Morbidity & Mortality Weekly Report.* 1999;48(20):413-416.**
80. **Anonymous. Drug-susceptible tuberculosis outbreak in a state correctional facility housing HIV-infected inmates--South Carolina, 1999-2000. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(46):1041-1044.**
81. **Anonymous. HIV-related tuberculosis in a transgender network--Baltimore, Maryland, and New York City area, 1998-2000. *MMWR Morb Mortal Wkly Rep.* Apr 21 2000;49(15):317-320.**
82. **Anonymous. Fluoroquinolone-resistance in Neisseria gonorrhoeae, Hawaii, 1999, and decreased susceptibility to azithromycin in N. gonorrhoeae, Missouri, 1999. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(37):833-837.**
83. **New Mexico Department of Health. RSVP. Available at: http://epi.health.state.nm.us/rsvpdesc/default.shtml. Accessed May 10, 2001.**
84. **Rapid Syndrome Validation Project (RSVP). Available at: http://epi.health.state.nm.us/rsvpdesc/default.shtml.**
85. **CDC Unexplained Deaths and Critical Illness Surveillance Project. Available at: " June 16, 2000 / 49(23);509-512 http://www.cdc.gov/ncidod/osr/EIP.htm http://www.cdc.gov/ncidod/eid/vol2no1/perkins.htmhttp://www.slackinc.com/child/idc/200012/unknown.asp".**
86. **Electronic Surveillance System for Early Notification of Community-based Epidemics (ESSENCE). Available at: http://www.geis.ha.osd.mil/ http://www.geis.ha.osd.mil/getpage.asp?page=SyndromicSurveillance.htm&action=7&click=KeyPrograms.**
87. **New Jersey's Active Influenza-like Illness Surveillance System. Available at: http://www.state.nj.us/health/fluinfo/.**
88. **New York City Emergency Management System. Available at: http://www.ci.nyc.ny.us/html/oem/html/oem_creation.html.**
89. **Anonymous. Epidemiologic Notes and Reports Pneumocystis carinii Pneumonia among Persons with Hemophilia A. *MMWR.* July 16, 1982 1982;31(27):365-367.**
90. **Anonymous. Pneumocystis Pneumonia --- Los Angeles. June 5, 1981 1981;Vol. 30(No. 21).**

91. **Anonymous. Outbreak of Hendra-like virus--Malaysia and Singapore, 1998-1999. *MMWR - Morbidity & Mortality Weekly Report.* 1999;48(13):265-269.**

92. **Antibodies to a Retrovirus Etiologically Associated with Acquired Immunodeficiency Syndrome (AIDS) in Populations with Increased Incidences of the Syndrome. *MMWR.* July 13, 1984; 33(27);377-9(27):337-339.**

93. **Anonymous. Outbreak of Escherichia coli O157:H7 and Campylobacter among attendees of the Washington County Fair-New York, 1999. *MMWR - Morbidity & Mortality Weekly Report.* 1999;48(36):803-805.**

94. **Anonymous. Outbreak of Salmonella serotype Muenchen infections associated with unpasteurized orange juice--United States and Canada, June 1999. *MMWR - Morbidity & Mortality Weekly Report.* 1999;48(27):582-585.**

95. **Anonymous. Outbreaks of gastrointestinal illness of unknown etiology associated with eating burritos--United States, October 1997-October 1998. *MMWR - Morbidity & Mortality Weekly Report.* 1999;48(10):210-213.**

96. **Anonymous. Outbreak of Rift Valley fever--Saudi Arabia, August-October, 2000. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(40):905-908.**

97. **Anonymous. Escherichia coli O111:H8 outbreak among teenage campers--Texas, 1999. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(15):321-324.**

98. **Anonymous. Outbreak of Shigella sonnei infections associated with eating a nationally distributed dip--California, Oregon, and Washington, January 2000. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(3):60-61.**

99. **Anonymous. Fatal yellow fever in a traveler returning from Venezuela, 1999. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(14):303-305.**

100. **Dialysis Surveillance Network. Available at: http://www.cdc.gov/ncidod/hip/DIALYSIS/dsn.htm.**

101. **"Adult/Adolescent Spectrum of HIV Disease (ASD) sentinel surveillance project,". *MMWR.* April 16, 1999 (48(SS-2)):1-22.**

102. **CDC's National Botulism Surveillance and Reference Laboratory. Available at: "September 01, 2000 / 49(34);778-780.**

103. **Influenza Activity as Assessed by State and Territorial Epidemiologists. Available at: http://www.cdc.gov/ncidod/diseases/flu/flusurv.htm.**

104. **IDSA.EIN or IDSA.NEThe Infectious Diseases Society of America (IDSA) Emerging Infections Network (EIN). Available at: http://www.idsociety.org/EIN/TOC.htm.**

105. **Anonymous. Pseudomonas dermatitis/folliculitis associated with pools and hot tubs--Colorado and Maine, 1999-2000. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(48):1087-1091.**

106. **Active Bacterial Core Surveillance (ABCs). Available at: "http://www.cdc.gov/drugresistance/technical/surveillance.htm, http://aspe.os.dhhs.gov/datacncl/datadir/cdc4.htm".**

107. **Cholera. Available at: http://aspe.os.dhhs.gov/datacncl/datadir/cdc4.htm.**

108. **Community-acquired MRSA infection in patients without established risk factors. Available at: "MMWR August 20, 1999 / 48(32);707-710".**

109. **"Gonococcal Isolate Surveillance Project (GISP),". Available at: http://www.cdc.gov/mmwr/preview/mmwrhtml/mm4937a1.htm.**

110. **Foodborne Diseases Active Surveillance Network (FoodNet). Available at: http://www.cdc.gov/ncidod/dbmd/foodnet/.**

111. **HIV/AIDS Reporting System. Available at: http://www.os.dhhs.gov/progorg/aspe/minority/mincdc4.htm.**

112. **"PulseNet, the national molecular subtyping network for foodborne disease". Available at: http://www.cdc.gov/ncidod/dbmd/pulsenet/pulsenet.htm#What%20is.**

113. **Anonymous. Serogroup W-135 meningococcal disease among travelers returning from Saudi Arabia--United States, 2000. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(16):345-346.**

114. **Anonymous. Outbreak of aseptic meningitis associated with multiple enterovirus serotypes--Romania, 1999. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(29):669-671.**

115. **Anonymous. Trichinellosis outbreaks--Northrhine-Westfalia, Germany, 1998-1999. *MMWR - Morbidity & Mortality Weekly Report.* 1999;48(23):488-492.**

116. **Anonymous. Tuberculosis outbreaks in prison housing units for HIV-infected inmates--California, 1995-1996. *MMWR - Morbidity & Mortality Weekly Report.* 1999;48(4):79-82.**

117. **Anonymous. Outbreaks of Shigella sonnei infection associated with eating fresh parsley--United States and Canada, July-August 1998. *MMWR - Morbidity & Mortality Weekly Report.* 1999;48(14):285-289.**

118. **Anonymous. Use of pulsed-field gel electrophoresis for investigation of a cluster of invasive group A streptococcal illness--Spokane, Washington, 1999. *MMWR - Morbidity & Mortality Weekly Report.* 1999;48(31):681-683.**

119. **Anonymous. Multistate outbreak of listeriosis--United States, 2000. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(50):1129-1130.**

120. **Anonymous. From the Centers for Disease Control and Prevention. Fatal illnesses associated with a New World arenavirus--California, 1999-2000. *Jama.* 2000;284(10):1237-1238.**

121. **Arness MK, Feighner BH, Canham ML, et al. Norwalk-like viral gastroenteritis outbreak in U.S. Army trainees. *Emerg Infect Dis.* Mar-Apr 2000;6(2):204-207.**

122. **Anonymous. Outbreak of poliomyelitis--Iraq, 1999. *MMWR - Morbidity & Mortality Weekly Report.* 1999;48(38):858-859.**

123. **Anonymous. Outbreak of poliomyelitis--Kunduz, Afghanistan, 1999. *MMWR - Morbidity & Mortality Weekly Report.* 1999;48(34):761-762.**

124. **Anonymous. Outbreak of influenza A infection among travelers--Alaska and the Yukon Territory, May-June 1999. *MMWR - Morbidity & Mortality Weekly Report.* 1999;48(25):545-546, 555.**

125. **Anonymous. Outbreak of West Nile-like viral encephalitis--New York, 1999. *MMWR - Morbidity & Mortality Weekly Report.* 1999;48(38):845-849.**

126. **Anonymous. Public health response to a potentially rabid bear cub--Iowa, 1999. *MMWR - Morbidity & Mortality Weekly Report.* 1999;48(42):971-973.**

127. **Ma JZ, Peterson DR, Ackerman E. Parameter sensitivity of a model of viral epidemics simulated with Monte Carlo techniques. IV. Parametric ranges and optimization. *International Journal of Bio-Medical Computing.* 1993;33(3-4):297-311.**

128. **Water-borne Disease Outbreak Surveillance System. Available at: http://www.cdc.gov/mmwr/preview/mmwrhtml/00001596.htm.**

129. **Foodborne- Disease Outbreak Surveillance System (FBDO). Available at: http://www.cdc.gov/mmwr/preview/mmwrhtml/ss4901a1.htm.**

130. **Hemophilia Surveillance System (HSS). Available at: http://www.cdc.gov/ncidod/osr/data-reports/dr-bysurvsys.htm.**

131. **National TB Genotyping and Surveillance Network. Available at: "http://www.cdc.gov/ncidod/dastlr/tb/tb_tgsn.htm - MMWR April 21, 2000 / 49(15);317-320".**

132. Anonymous. Outbreak of poliomyelitis--Angola, 1999. *MMWR - Morbidity & Mortality Weekly Report.* 1999;48(16):327-329.
133. Anonymous. From the Centers for Disease Control and Prevention. Four pediatric deaths from community-acquired methicillin-resistant Staphylococcus aureus--Minnesota and North Dakota, 1997-1999. *JAMA.* 1999;282(12):1123-1125.
134. Anonymous. Update: multistate outbreak of listeriosis--United States, 1998-1999. *MMWR - Morbidity & Mortality Weekly Report.* 1999;47(51-52):1117-1118.
135. Anonymous. Rubella outbreak--Westchester County, New York, 1997-1998. *MMWR - Morbidity & Mortality Weekly Report.* 1999;48(26):560-563.
136. Anonymous. Scombroid Fish Poisoning --- Pennsylvania, 1998. *MMWR.* May 12, 2000 2000;49(18):398-400.
137. Anonymous. Human Ingestion of Bacillus Anthracis-Contaminated Meat --- Minnesota, August 2000. *MMWR.* September 15, 2000 2000;49(36):813-816.
138. Anonymous. Foodborne botulism from eating home-pickled eggs--Illinois, 1997. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(34):778-780.
139. Anonymous. Outbreak of poliomyelitis--Dominican Republic and Haiti, 2000. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(48):1094, 1103.
140. Anonymous. Outbreak of poliomyelitis--Cape Verde, 2000. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(47):1070.
141. Anonymous. Measles outbreak--Netherlands, April 1999-January 2000. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(14):299-303.
142. Anonymous. Rubella among Hispanic adults--Kansas, 1998, and Nebraska, 1999. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(11):225-228.
143. Anonymous. From the Centers for Disease Control and Prevention. Serogroup W-135 meningococcal disease among travelers returning from Saudi Arabia--United States, 2000. *Jama.* May 24-31 2000;283(20):2647.
144. Anonymous. Probable locally acquired mosquito-transmitted Plasmodium vivax infection--Suffolk County, New York, 1999. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(22):495-498.
145. Anonymous. Coccidioidomycosis in travelers returning from Mexico--Pennsylvania, 2000. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(44):1004-1006.
146. Anonymous. Cluster of HIV-infected adolescents and young adults--Mississippi, 1999. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(38):861-864.
147. Anonymous. Outbreaks of Escherichia coli O157:H7 infections among children associated with farm visits--Pennsylvania and Washington, 2000. *MMWR - Morbidity & Mortality Weekly Report.* 2001;50(15):293-297.
148. Anonymous. Salmonellosis associated with chicks and ducklings--Michigan and Missouri, Spring 1999. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(14):297-299.
149. Anonymous. Reptile-associated salmonellosis--selected states, 1996-1998. [erratum appears in MMWR Morb Mortal Wkly Rep 1999 Nov 19;48(45):1051]. *MMWR - Morbidity & Mortality Weekly Report.* 1999;48(44):1009-1013.
150. Anonymous. Human rabies--Virginia, 1998. *MMWR - Morbidity & Mortality Weekly Report.* 1999;48(5):95-97.
151. Anonymous. Blastomycosis acquired occupationally during prairie dog relocation--Colorado, 1998. *MMWR - Morbidity & Mortality Weekly Report.* 1999;48(5):98-100.
152. Anonymous. Outbreak of Rift Valley fever--Yemen, August-October 2000. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(47):1065-1066.

153. Anonymous. Update: outbreak of Rift Valley Fever--Saudi Arabia, August-November 2000. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(43):982-985.
154. Anonymous. Varicella outbreaks among Mexican adults--Alabama, 2000. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(32):735-736.
155. Anonymous. Reptile-associated salmonellosis--selected states, 1996-1998. *Canada Communicable Disease Report.* 2000;26(3):19-23.
156. Anonymous. From the Centers for Disease Control and Prevention. Human ingestion of Bacillus anthracis-contaminated meat--Minnesota, August 2000. *Jama.* Oct 4 2000;284(13):1644-1646.
157. Anonymous. Outbreaks of Salmonella serotype enteritidis infection associated with eating raw or undercooked shell eggs--United States, 1996-1998. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(4):73-79.
158. Anonymous. Outbreak of Escherichia coli O157:H7 infection associated with eating fresh cheese curds--Wisconsin, June 1998. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(40):911-913.
159. Anonymous. Foodborne outbreak of Group A rotavirus gastroenteritis among college students--District of Columbia, March-April 2000. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(50):1131-1133.
160. Anonymous. Outbreaks of Norwalk-like viral gastroenteritis--Alaska and Wisconsin, 1999. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(10):207-211.
161. Anonymous. Outbreak of gastroenteritis associated with an interactive water fountain at a beachside park--Florida, 1999. *MMWR - Morbidity & Mortality Weekly Report.* 2000;49(25):565-568.
162. Anonymous. Outbreak of Vibrio parahaemolyticus infection associated with eating raw oysters and clams harvested from Long Island Sound--Connecticut, New Jersey, and New York, 1998. *MMWR - Morbidity & Mortality Weekly Report.* 1999;48(3):48-51.
163. Anonymous. From the Centers for Disease Control and Prevention. Scombroid fish poisoning--Pennsylvania, 1998. *JAMA.* 2000;283(22):2927-2928.
164. Anonymous. Update: outbreak of Nipah virus--Malaysia and Singapore, 1999. *MMWR - Morbidity & Mortality Weekly Report.* 1999;48(16):335-337.
165. Anonymous. Probable locally acquired mosquito-transmitted Plamodium vivax infection--Suffolk County, New York, 1999. *Canada Communicable Disease Report.* 2000;26(16):137-139.
166. Anonymous. From the Centers for Disease Control and Prevention. Probable locally acquired mosquito-transmitted Plasmodium vivax infection--Suffolk County, New York, 1999. *Jama.* 2000;284(4):431-432.

Appendices

Appendix A. Required Data

Table A.1 Types of data/data sources identified by different methods (column two combines first principles and care seeking literature)

	Theoretical	Collected by Public Health	Used in Outbreak Investigations	Used for CDC definitions
Pre-outbreak	911-call system (suspicious activities, hazardous conditions) Police activity Veterinary data systems especially public veterinary practices (diseases that animals transmit to humans including animals which will eventually be distrbuted to consumers in the form of pets or food) Contact tracing to find infectious individiduals before additional outbreaks occur.	Environmental condition as detected by satellite. (e.g., vegetation, climate, sea surface temperature, cloudiness, rainfall)[58] Immunization registries information [59](e.g., child's name, date and place of birth, identifiers, contact information, dates, types and lots of vaccinations) Outbreaks occurring in other regions that have the potential for spread[60, 61] Emergence of new infections in other regions [61, 62] Keywords in the Internet and electronic reports related to or indicative of outbreaks in other areas[63] Antimicrobial resistance patterns [64-68] Routine testing of animal food and water supplies Monitoring temporal and geographic patterns associated with the detection diseases (respiratory syncytial virus (RSV), human parainfluenza viruses (HPIV), respiratory and enteric adenoviruses, and rotavirus) [69] Avian morbidity and mortality, captive or free ranging sentinel		

	Theoretical	Collected by Public Health	Used in Outbreak Investigations	Used for CDC definitions
		animals [70] Port inspections for foreign pests and disease[71]		
Attack/ release/ exposure	911-call data (e.g., report envelop with white power) Poison center data (e.g., report of exposure to infectious or toxic agent) Vector surveillance data Environmental monitoring: special event, permanent	Monitoring of mosquito pools or other vectors for agents [70] Routine testing of water supplies[72] Animal Exposures (rabies post exposure prophylaxis practices) [73] Injury related exposures to infectious disease (reports of percutaneous injuries in hospitals) [74] 3 Horse or other terminal host surveillance to observe for the presence of organisms in the environment which also infect humans[70]	CO2 monitor[75-78]	
Pre-symptomatic	Biometric data (temperature, increase heart rate) If animals are more susceptible that humans, veterinary data systems Routine testing	Tuberculin skin testing programs[74] .	Contact tracing[79-82]	
Non-specific syndrome	Pharmaceutical sales/marketing systems (over-the-counter drug sales) Clinical (medical records, physical exam, non-specific laboratory tests, non-specific diagnosis, HMO usage, outpatient clinical volumes,	Chief complaints Symptoms (hepatitis[83] and febrile rash[83], fever, water diarrhea, altered mental status[83], bloody diarrhea[73, 83], dehydration, vomiting, jaundice, seizure[73], paralysis, flu like illness[84] cramps) Adult respiratory distress syndrome[83]	Chest abnormalities, Chest X-ray[89], deaths[90, 91] Signs and symptoms of illness (fatigue, fever[90], arthralgia, myalgia nausea, headache, abdominal cramps diarrhea, vomiting, fever, pneumonia, ataxia,	Diarrhea, abdominal cramps, loss of appetite, low-grade fever, nausea, and vomiting.(Cryptosporidiosis) Diarrhea and/or vomiting; (Cholera) Acute onset of fever, headache, myalgia, and/or malaise. Nausea, vomiting, or rash may be present in some cases.(Ehrlichiosis) Fever, headache, stiff neck, and pleocytosis. altered mental status, confusion, coma , paresis or paralysis, cranial nerve palsies, sensory

	Theoretical	**Collected by Public Health**	**Used in Outbreak Investigations**	**Used for CDC definitions**
	emergency room volume and chief complaints, unusual numbers of unexplained deaths) Absenteeism reporting Systems (non-specific increase in absence at work or school) Vital statistics (fact of death, nonspecific diagnosis on death certificate, acute in crease in low birth rate or birth defects. Web-related health sites (queries related to words like cough) Newspaper reports (obituaries, articles of unusual deaths of animals and humans, clustering of victims) Sentinel population monitoring Social behavior patterns and changes Poison control system usage (reports of illness)	Unexplained deaths and critical illness [85] Deaths from pneumonia and influenza [35] Number of patients reporting, influenza like illness[36] fevers, headaches, diarrhea, vomiting, stuffy noses, coughs or rashes [86] Information on sudden deaths [86] School absentee rates and predominant reason for school absentee[87] Nursing home respiratory illness rate[87]s Emergency medical services dispatching[88] Widened mediastinum Gram positive rods Number of medical visits with gastrointestinal illness	lymphadenopathy[92], conjunctivitis, hepatic abnormalities, prolonged shock, unexplained fever, gastroenteritis)[89, 93-99] Clinical laboratory results (abnormal blood test such as white blood count, pathological tests) Emergency room medications Absentee records	deficits, abnormal reflexes, generalized convulsions, and abnormal movements.("Encephalitis or meningitis, arboviral") Diarrhea (often bloody) and abdominal cramps. (Enterohemorrhagic Escherichia coli) Meningitis, bacteremia, epiglottitis, or pneumonia. ("Haemophilus influenzae, invasive disease") Watery diarrhea, loss of appetite, weight loss, abdominal bloating and cramping, increased flatus, nausea, fatigue, and low-grade fever. vomiting (Cyclosporiasis) Meningitis or bacteremia;, fetal loss through miscarriage or stillbirth, or neonatal meningitis or bacteremia.(Listeriosis)**:** Fatigue, fever, headache, mildly stiff neck, arthralgia, or myalgia. (Listeriosis) Fever. headache, back pain, chills, sweats, myalgia, nausea, vomiting, diarrhea, and cough. coma, renal failure, pulmonary edema, and death.(Malaria) Fever, chills, headache, photophobia, cough, and myalgia (Psittacosis) Encephalomyelitis, coma, death("Rabies, human") Diarrhea, abdominal pain, nausea, ,vomiting. (Salmonellosis) Diarrhea, fever, nausea, cramps, and tenesmus.(Shigellosis) Cutaneous infection (e.g., cellulitis, erysipelas, or infection of a surgical or nonsurgical wound), deep soft-tissue infection (e.g., myositis or necrotizing fasciitis), meningitis, peritonitis, osteomyelitis, septic arthritis, postpartum sepsis (i.e., puerperal fever), neonatal sepsis, and nonfocal bacteremia. ("Streptococcal disease, invasive, Group A") Acute otitis media, pneumonia, bacteremia, or meningitis("Streptococcus pneumoniae, drug-

	Theoretical	Collected by Public Health	Used in Outbreak Investigations	Used for CDC definitions
				resistant invasive"). Acute onset maculopapular rash Temperature greater than 99.0 F (greater than 37.2 C) Arthralgia/arthritis, lymphadenopathy, or conjunctivitis (Rubella) Sustained fever, headache, malaise, anorexia, relative bradycardia, constipation or diarrhea, and nonproductive cough(Typhoid fever)
Specific syndrome	Vital statistics (e.g., specific diagnosis related to bioterrorist agents) Physician data base searches (web-based medical queries) County coroner's office/medical examiner information Electronic medical records (specific diagnosis in outpatient, hospital or emergency room) Public health reporting (reportable diseases) Medication usage (usage of medications reserved for unusual diseases)	Cause of death on death certificate (if syndromic) Medical record (e.g., dialysis related infection, [100]opportunistic infection[101]) Disease reports which match a case definition (botulism)[102] Iinfluenza-like activity as assessed by state epidemiologists[103] Discharge diagnoses (AIDS indicator diseases,[101] necrotizing fasciitis, streptococcal toxic shock syndrome (STSS)[104], neurocysteicercosis,[73] CD4 Counts [101] Frequency of prescriptions for antiretroviral therapy and prophylaxis for *Pneumocystis carinii* pneumonia and *Mycobacterium avium* complex disease.[101]	Medical diagnosis (PCP diagnosis, toxoplasmosis, CO poisoning[75-78]) Reportable disease e.g. folliculitis/dermatitis[105]	Prodrome resembling a viral respiratory illness, hypoxia and dyspnea, radiographic evidence of mediastinal widening , severe abdominal distress, fever and signs of septicemia (Anthrax) Diplopia, blurred vision, and bulbar weakness. Symmetric paralysis.("Botulism, foodborne") Acute or insidious onset of fever, night sweats, undue fatigue, anorexia, weight loss, headache, and arthralgia (Brucellosis) painful genital ulceration and inflammatory inguinal adenopathy ulcer(s) not typical of disease caused by (Chancroid) urethritis, epididymitis, cervicitis, acute salpingitis, lymphogranuloma venereum and trachoma.("Chlamydia trachomatis, genital infections") fever, chest pain, cough, myalgia, arthralgia, and headache) Pneumonia or other pulmonary lesion, diagnosed by chest radiograph, Erythema nodosum or erythema multiforme rash Involvement of bones, joints, or skin by dissemination,meningitis involvement of viscera and lymph nodes (Coccidioidomycosis) Acute onset of fever, headache, myalgia, and/or malaise,nausea, vomiting, or rash.(Ehrlichiosis) Thrombocytopenia, leukopenia, and/or elevated liver enzymes. Intracytoplasmic bacterial aggregates (morulae) may be visible in the leukocytes of some patients. .(Ehrlichiosis) Hemolytic uremic syndrome (HUS) or thrombotic

	Theoretical	Collected by Public Health	Used in Outbreak Investigations	Used for CDC definitions
				thrombocytopenic purpura (TTP)(Enterohemorrhagic Escherichia coli) Tuberculoid: one or a few well-demarcated, hypopigmented, and anesthetic skin lesions, frequently with active, spreading edges and a clearing center; peripheral nerve swelling or thickening also may occur (Hansen disease (Leprosy)) Lepromatous: a number of erythematous papules and nodules or an infiltration of the face, hands, and feet with lesions in a bilateral and symmetrical distribution that progress to thickening of the skin (Hansen disease (Leprosy)) Borderline (dimorphous): skin lesions characteristic of both the tuberculoid and lepromatous forms Indeterminate: early lesions, usually hypopigmented macules, without developed tuberculoid or lepromatous features (Hansen disease (Leprosy)) Febrile illness (i.e., temperature greater than 101.0 F [greater than 38.3 C]) characterized by bilateral diffuse interstitial edema that may radiographically resemble ARDS, with respiratory compromise requiring supplemental oxygen, developing within 72 hours of hospitalization, and occurring in a previously healthy person An unexplained respiratory illness resulting in death, with an autopsy examination demonstrating noncardiogenic pulmonary edema without an identifiable cause(Hantavirus Pulmonary Syndrome) Acute onset of microangiopathic hemolytic anemia, renal injury, and low platelet count.(Hemolytic uremic syndrome post-diarrheal) Anemia (acute onset) with microangiopathic changes (i.e., schistocytes, burr cells, or helmet

	Theoretical	Collected by Public Health	Used in Outbreak Investigations	Used for CDC definitions
				cells) on peripheral blood smear and Renal injury (acute onset) evidenced by either hematuria, proteinuria, or elevated creatinine level (Hemolytic uremic syndrome post-diarrheal) Discrete onset of symptoms and jaundice or elevated serum aminotransferase levels("Hepatitis, viral, acute") Fever, myalgia, cough, pneumonia (Legionellosis) Erythema migrans, recurrent, brief attacks (weeks or months) of objective joint swelling in one or a few joints, sometimes followed by chronic arthritis in one or a few joints, lymphocytic meningitis; cranial neuritis, particularly facial palsy (may be bilateral); radiculoneuropathy; or, rarely, encephalomyelitis, (2nd-degree or 3rd-degree) atrioventricular conduction defects that resolve in days to weeks and are sometimes associated with myocarditis.(Listeriosis) Generalized rash, length of time of rash, height of temperature,cough, coryza, or conjunctivitis(Measles) Meningitis, meningococcemia, purpura fulminans, shock, and death. (Meningococcal disease) Unilateral or bilateral tender, self-limited swelling of the paraotid or other salivary gland, lasting greater than or equal to 2 days, and without other apparent cause (Mumps) Cough illness, length of illness, paroxysms of coughing, inspiratory "whoop," or post-tussive vomiting, without other apparent cause (Pertussis) Flaccid paralysis of one or more limbs, Decreased or absent tendon reflexes, affected limbs, other apparent cause, sensory or cognitive loss, neurologic deficit, time after initial symptoms, death, follow-up status("Polio, paralytic") Fever, rigors, myalgia, malaise, and retrobulbar

	Theoretical	Collected by Public Health	Used in Outbreak Investigations	Used for CDC definitions
				headache, acute hepatitis, pneumonia, meningoencephalitis, liver enzyme levels, chest film findings. endocarditis, chronic fatigue-like syndrome (Q fever) Acute onset, myalgia, headache, and petechial rash (on the palms and soles in two thirds of the cases) time from hospitalization (Rocky Mountain spotted fever) Hypotension Renal impairment: Creatinine, history of renal disease, platelet count, disseminated intravascular coagulation.: alanine aminotransferase, aspartate aminotransferase,total bilirubin, preexisting liver disease, acute onset of diffuse pulmonary infiltrates and hypoxemia, cardiac failure, diffuse capillary leak, acute onset of generalized edema, or pleural or peritoneal effusions, hypoalbuminemia. generalized erythematous macular rash, desquamation, soft-tissue necrosis, necrotizing fasciitis or myositis, or gangrene.(Streptococcal toxic-shock syndrome) One or more ulcers (chancres) localized or diffuse mucocutaneous lesions, often with generalized lymphadenopathy. central nervous system infection inflammatory lesions of the cardiovascular system, skin, and bone.(Syphilis) Diffuse (generalized) maculo-papulovesicular rash(Varicella (Chickenpox) - 1999 Non-Notifiable) Acute onset of hypertonia and/or painful muscular contractions (usually of the muscles of the jaw and neck) and generalized muscle spasms without other apparent medical cause(Tetanus) Fever: temperature greater than or equal to 102.0°F (greater than or equal to 38.9°C) rash: diffuse macular erythroderma Desquamation: 1-2 weeks after onset of illness, particularly on the palms and soles hypotension: orthostatic

	Theoretical	Collected by Public Health	Used in Outbreak Investigations	Used for CDC definitions
				syncope, or orthostatic dizziness gastrointestinal: vomiting, diarrhea ,-muscular: severe myalgia or creatine phosphokinase,-mucous membrane: vaginal, oropharyngeal, or conjunctival hyperemia, blood urea nitrogen or creatinine, urinary sediment with pyuria, total bilirubin, alanine aminotransferase enzyme, or asparate aminotransferase enzyme platelets, -central nervous system: disorientation or alterations in consciousness, focal neurologic signs,fever, hypotension (Toxic-shock syndrome) Eosinophilia, fever, myalgia, and periorbital edema.(Trichinosis) A positive tuberculin skin test, chest radiographs, or clinical evidence of current disease-(Tuberculosis) Cutaneous ulcer, regional lymphadenopathy ,: regional lymphadenopathy with no ulcer,: conjunctivitis with preauricular lymphadenopathy, stomatitis or pharyngitis or tonsillitis and cervical lymphadenopathy, intestinal pain, vomiting, diarrhea, primary pleuropulmonary disease,(Tularemia) Acute onset, constitutional symptoms, brief remission, recurrence of fever, hepatitis, albuminuria, renal failure, shock, and generalized hemorrhages (Yellow fever)
Definitive	Radiology reports (mediastinal widening for Inhalational anthrax) Microbial culture results Microbial resistance patterns CoHB levels	Bacterial culture and sensitivity (group A Streptococcus,[106] group B Streptococcus,[106] *Haemophilius influenzae,*[106] *Neisseria meningitidis,*[106] *Streptococcus pneumoniae*,[106] Cholera,[107] MSRA,[108]. N. gonorrhoeae,[109] *Staphylococcus aureus* with reduced susceptibility to Vancomycin,	Serology[92, 113, 114] Pathology (trans bronchial biopsy[89], microscopic examination, lung biopsy specimens[90, 91]) Post mortem examination,[90] (postmortem COHB levels, toxicologic and Cultures (sputum,stool,	Screening test for HIV antibody (enzyme immunoassay), followed by a positive result on a confirmatory (sensitive and more specific) test for HIV antibody (e.g., Western blot or immunofluorescence antibody test) HIV virologic (nonantibody) tests: HIV nucleic acid (DNA or RNA) detection (e.g., DNA polymerase chain reaction [PCR] or plasma HIV-1 RNA), HIV p24 antigen test, including neutralization assay, HIV isolation (viral

	Theoretical	**Collected by Public Health**	**Used in Outbreak Investigations**	**Used for CDC definitions**
		Salmonella (serogroup/serotype),[67, 110] Vibrio (species),[110] Shigella (serogroup/species),[67, 110] *Listeria monocytogenes*,[110] *Campylobacter* [110](species),[67] *Yersinia enterocolitical*[110] *E. coli O157*,[67, 110] other H antigen positive nonmotile, shiga-like toxin producing)[73] *Cryptosporidum, Cyclospora* Serologic titers (HIV)[111] Detection of botulinum toxin in serum, stool, or patient's food[102] Rapid Diagnosis Tests (Directigen, EIA, fluorescent antibodies) Number of isolates of Salmonella (current compared to historical)[33] DNA fingerprints[112],or RNA[70], (West Nile Virus)	blood, joint for bacterial or viral pathogens)[115] Assays (Cell culture, biologic toxicity, polymerase chain reaction[96], immunologic) Chemical analyses for toxins DNA fingerprint[80, 81, 116], Pulsed-field gel electrophoresis[117-119], reverse transcriptase[120, 121], enzyme records, plamid analysis, Viral cultures[122-124], Serologic testing[92] and assays[125] , rubella specific IGM, Skin biopsy, Direct fluorescent antibody test[126], rapid influenza A antigen-detection tests[124] Pathological examination (intercellular parasites), PCR testing[96], DNA sequencing, phylogenetic analysis, antimicrobial susceptibility pattern comparisons Autopsy[90],	culture)("HIV Infection, adult") Botulinum toxin in serum, stool, or patient's food, or isolation of Clostridium botulinum from stool ("Botulism, foodborne") Isolation of Brucella sp. from a clinical specimen, or Fourfold or greater rise in Brucella agglutination titer between acute- and convalescent-phase serum specimens obtained greater than or equal to 2 weeks apart and studied at the same laboratory, or demonstration by immunofluorescence of Brucella sp. in a clinical specimen(Brucellosis) Isolation of H. ducreyi from a clinical specimen (Chancroid) Treponema pallidum infection by darkfield microscopic examination of ulcer exudates, by a serologic test for syphilis performed greater than or equal to 7 days after onset of ulcers herpes simplex virus (HSV) culture negative for HSV.(Chancroid) Isolation of C. trachomatis by culture, or Demonstration of C. trachomatis in a clinical specimen by detection of antigen or nucleic acid ("Chlamydia trachomatis, genital infections") Isolation of toxigenic (i.e., cholera toxin-producing) Vibrio cholerae O1 or O139 from stool or vomitus, or Serologic evidence of recent infection (Cholera) Cryptosporidiosis oocysts in stool by microscopic examination, or in intestinal fluid or small-bowel biopsy specimens, or oocyst or sporozoite antigens by immunodiagnostic methods, e.g., ELISA, or by PCR techniques when routinely available, or demonstration of reproductive stages in tissue preparations.(Cryptosporidiosis) Demonstration of a four-fold change in antibody titer to E. chaffeensis antigen by indirect immunofluorescence assay (IFA) in paired serum samples, or Positive polymerase chain

	Theoretical	Collected by Public Health	Used in Outbreak Investigations	Used for CDC definitions
				reaction (PCR) assay and confirmation of E. chaffeensis DNA, or Identification of morulae in leukocytes, and a positive IFA titer to E. chaffeensis antigen (based on cutoff titers established by the laboratory performing the assay), or Immunostaining of E. chaffeensis antigen in a biopsy or autopsy sample, or Culture of E. chaffeensis from a clinical specimen. HGE: Demonstration of a four-fold change in antibody titer to E. phagocytophila antigen by IFA in paired serum samples, or Positive PCR assay and confirmation of E. phagocytophila DNA, or Identification of morulae in leukocytes, and a positive IFA titer to E. phagocytophila antigen (based on cutoff titers established by the laboratory performing the assay), or Immunostaining of E. phagocytophila antigen in a biopsy or autopsy sample, or Culture of E. phagocytophila from a clinical specimen. Ehrlichiosis (other or unspecified agent): Demonstration of a four-fold change in antibody titer to more than one Ehrlichia species by IFA in paired serum samples, in which a dominant reactivity cannot be established, or Identification of an Ehrlichia species other than E. chaffeensis or E. phagocytophila by PCR, immunostaining, or culture. (Ehrlichiosis) Fourfold or greater change in virus-specific serum antibody titer, or Isolation of virus from or demonstration of specific viral antigen or genomic sequences in tissue, blood, cerebrospinal fluid (CSF), or other body fluid, or Virus-specific immunoglobulin M (IgM) antibodies demonstrated in CSF by antibody-capture enzyme immunoassay (EIA), or Virus-specific IgM antibodies demonstrated in serum by antibody-capture EIA and confirmed by demonstration of virus-specific serum

	Theoretical	Collected by Public Health	Used in Outbreak Investigations	Used for CDC definitions
				immunoglobulin G (IgG) antibodies in the same or a later specimen by another serologic assay (e.g., neutralization or hemagglutination inhibition).("Encephalitis or meningitis, arboviral") Demonstration of G. lamblia cysts in stool, or - Demonstration of G. lamblia trophozoites in stool, duodenal fluid, or small bowel biopsy, or - Demonstration of G. lamblia antigen in stool by a specific immunodiagnostic such as enzyme-linked immunosorbent assay (ELISA) (Giardiasis) Isolation of typical gram-negative, oxidase-positive diplococci (presumptive Neisseria gonorrhoeae) Demonstration of N. gonorrhoeae in a clinical specimen by detection of antigen or nucleic acid, or Observation of gram-negative intracellular diplococci in a urethral smear obtained from a male(Gonorrhea) Isolation of H. influenzae from a normally sterile site (e.g., blood or cerebrospinal fluid [CSF] or, less commonly, joint, pleural, or pericardial fluid) ("Haemophilus influenzae, invasive disease") Cyclospora oocysts in stool by microscopic examination, or in intestinal fluid or small bowel biopsy specimens, or demonstration of sporulation, or DNA (by polymerase chain reaction) in stool, duodenal/jejunal aspirates or small bowel biopsy specimens.(Cyclosporiasis) Demonstration of acid-fast bacilli in skin or dermal nerve, obtained from the full-thickness skin biopsy of a lepromatous lesion (Hansen disease (Leprosy)) Detection of hantavirus-specific immunoglobulin M or rising titers of hantavirus-specific immunoglobulin G, or Detection of hantavirus-specific ribonucleic acid sequence by polymerase chain reaction in clinical specimens,

	Theoretical	Collected by Public Health	Used in Outbreak Investigations	Used for CDC definitions
				or Detection of hantavirus antigen by immunohistochemistry (Hantavirus Pulmonary Syndrome) Hepatitis A: Immunoglobulin M (IgM) antibody to hepatitis A virus (anti-HAV) positive Hepatitis B: IgM antibody to hepatitis B core antigen (anti-HBc) positive or hepatitis B surface antigen (HBsAg) positive IgM anti-HAV negative (if done) Hepatitis C: Revised 2000 Serum aminotransferase levels greater than 7 times the upper limit of normal, and IgM anti-HAV negative, and IgM anti-HBc negative (if done) or HBsAg negative, and Antibody to hepatitis C virus (anti-HCV) positive, verified by a n additional more specific assay Non-A, Non-B hepatitis: Serum aminotransferase levels greater than 2.5 times the upper limit of normal, and IgM anti-HAV negative, and IgM anti-HBc negative (if done) or HBsAg negative, and Anti-HCV negative (if done) Delta hepatitis*: HBsAg or IgM anti-HBc positive and antibody to hepatitis delta virus positive ("Hepatitis, viral, acute") Isolation of Legionella from respiratory secretions, lung tissue, pleural fluid, or other normally sterile fluids, or Demonstration of a fourfold or greater rise in the reciprocal immunofluorescence antibody (IFA) titer to greater than or equal to 128 against Legionella pneumophila serogroup 1 between paired acute- and convalescent-phase serum specimens, or Detection of L. pneumophila serogroup 1 in respiratory secretions, lung tissue, or pleural fluid by direct fluorescent antibody testing, or Demonstration of L. pneumophila serogroup 1 antigens in urine by radioimmunoassay or enzyme-linked immunosorbent assay(Legionellosis)

	Theoretical	Collected by Public Health	Used in Outbreak Investigations	Used for CDC definitions
				Isolation of L. monocytogenes from a normally sterile site (e.g., blood or cerebrospinal fluid [CSF] or, less commonly, joint, pleural, or pericardial fluid) isolation of L. monocytogenes from placental or fetal tissue (Listeriosis) Isolation of Borrelia burgdorferi from a clinical specimen or Demonstration of diagnostic immunoglobulin M or immunoglobulin G antibodies to B. burgdorferi in serum or cerebrospinal fluid (CSF). A two-test approach using a sensitive enzyme immunoassay or immunofluorescence antibody followed by Western blot is recommended (Listeriosis) Demonstration of malaria parasites in blood films (Malaria) Positive serologic test for measles mmunoglobulin M antibody, or Significant rise in measles antibody level by any standard serologic assay, or Isolation of measles virus from a clinical specimen(Measles) Isolation of Neisseria meningitidis from a normally sterile site (e.g., blood or cerebrospinal fluid [CSF] or, less commonly, joint, pleural, or pericardial fluid) (Meningococcal disease) Isolation of mumps virus from clinical specimen, or Significant rise between acute- and convalescent-phase titers in serum mumps immunoglobulin G (IgG) antibody level by any standard serologic assay, or Positive serologic test for mumps immunoglobulin M (IgM) antibody (Mumps) Isolation of Bordetella pertussis from clinical specimen Positive polymerase chain reaction (PCR) for B. pertussis(Pertussis) Isolation of Chlamydia psittaci from respiratory secretions, or Fourfold or greater increase in antibody against C. psittaci by complement fixation or microimmunofluorescence (MIF) to a

	Theoretical	Collected by Public Health	Used in Outbreak Investigations	Used for CDC definitions
				reciprocal titer of greater than or equal to 32 between paired acute- and convalescent-phase serum specimens, or Presence of immunoglobulin M antibody against C. psittaci by MIF to a reciprocal titer of greater than or equal to 16 (Psittacosis) Fourfold or greater change in antibody titer to C. burnetii phase II or phase I antigen in paired serum specimens ideally taken 3-6 weeks apart, or, Isolation of C. burnetii from a clinical specimen by culture, or Demonstration of C. burnetti in a clinical specimen by detection of antigen or nucleic acid.(Q fever) Rabies, human encephalomyelitis, coma, death("Rabies, human") Detection by direct fluorescent antibody of viral antigens in a clinical specimen (preferably the brain or the nerves surrounding hair follicles in the nape of the neck), or Isolation (in cell culture or in a laboratory animal) of rabies virus from saliva, cerebrospinal fluid (CSF), or central nervous system tissue, or Identification of a rabies-neutralizing antibody titer greater than or equal to 5 (complete neutralization) in the serum or CSF of an unvaccinated person. ("Rabies, human") Fourfold or greater rise in antibody titer to Rickettsia rickettsii antigen by immunofluorescence antibody (IFA), complement fixation (CF), latex agglutination (LA), microagglutination [127], or indirect hemagglutination antibody (IHA) test in acute- and convalescent-phase specimens ideally taken greater than or equal to 3 weeks apart, or Positive polymerase chain reaction assay to R. rickettsii, or Demonstration of positive immunofluorescence of skin lesion (biopsy) or organ tissue (autopsy), or Isolation of R.

	Theoretical	Collected by Public Health	Used in Outbreak Investigations	Used for CDC definitions
				rickettsii from clinical specimen (Rocky Mountain spotted fever) Isolation of rubella virus, or Significant rise between acute- and convalescent-phase titers in serum rubella immunoglobulin G antibody level by any standard serologic assay, or Positive serologic test for rubella immunoglobulin M (IgM) antibody (Rubella) Isolation of Salmonella from a clinical specimen (Salmonellosis) Isolation of Shigella from a clinical specimen (Shigellosis) Streptococcus (Streptococcus pyogenes) by culture from a normally sterile site (e.g., blood or cerebrospinal fluid, or, less commonly, joint, pleural, or pericardial fluid) ("Streptococcal disease, invasive, Group A") Isolation of group A Streptococcus. .(Streptococcal toxic-shock syndrome) Isolation of S. pneumoniae from a normally sterile site (e.g., blood, cerebrospinal fluid, or, less commonly, joint, pleural, or pericardial fluid) and "Nonsusceptible" isolate (i.e., intermediate- or high-level resistance of the S. pneumoniae isolate to at least one antimicrobial agent currently approved for use in treating pneumococcal infection ("Streptococcal disease, invasive, Group A")* Isolation of S. pneumoniae from a normally sterile site (e.g., blood, cerebrospinal fluid, or, less commonly, joint, pleural, or pericardial fluid) ("Streptococcus pneumoniae, Invasive, (Children < 5 years)") Demonstration of T. pallidum in clinical specimens by darkfield microscopy, direct fluorescent antibody (DFA-TP), or equivalent methods.("Syphilis, (1) primary") Trichinella larvae in tissue obtained by muscle biopsy, or Positive serologic test for Trichinella (Trichinosis)

	Theoretical	Collected by Public Health	Used in Outbreak Investigations	Used for CDC definitions
				Isolation of M. tuberculosis from a clinical specimen or Demonstration of M. tuberculosis from a clinical specimen by nucleic acid amplification test, or Demonstration of acid-fast bacilli in a clinical specimen (Tuberculosis) Presumptive: -Elevated serum antibody titer(s) to F. tularensis antigen (without documented fourfold or greater change) in a patient with no history of tularemia vaccination or -Detection of F. tularensis in a clinical specimen by fluorescent assay Confirmatory: -Isolation of F. tularensis in a clinical specimen or -Fourfold or greater change in serum antibody titer to F. tularensis antigen(Tularemia) Isolation of S. typhi from blood, stool, or other clinical specimen (Typhoid fever) Fourfold or greater rise in yellow fever antibody titer in a patient who has no history of recent yellow fever vaccination and cross-reactions to other flaviviruses have been excluded or Demonstration of yellow fever virus, antigen, or genome in tissue, blood, or other body fluid (Yellow fever) Isolation of varicella virus from a clinical specimen, or Direct fluorescent antibody (DFA), or Polymerase chain reaction (PCR), or Significant rise in serum varicella immunoglobulin G (IgG) antibody level by any standard serologic assay(Varicella (Chickenpox) - 1999 Non-Notifiable) Reactive serologic test (nontreponemal: Venereal Disease Research Laboratory [VDRL] or rapid plasma reagin [RPR]; treponemal: fluorescent treponemal antibody absorbed [FTA-ABS] or microhemagglutination assay for antibody to T. pallidum [MHA-TP]) ("Syphilis, (1) primary") Elevated CSF protein or leukocyte count("Syphilis, (7) neuro")

	Theoretical	Collected by Public Health	Used in Outbreak Investigations	Used for CDC definitions
				Isolation of Escherichia coli O157:H7 from a specimen, or Isolation of Shiga toxin-producing E. coli from a clinical specimen,(Enterohemorrhagic Escherichia coli)
Epidemiologic Data	911 GIS information system (geographic dispersion of victims) Medical records systems (age, sex, race, employer, religion, past history) Worksite logs, guest lists, itinerary Food history, food handler interviews, food preference questionnaire, Social aggregations: families, co-workers, recreational groups Location of recent travel or gatherings Food and water (specific intake of ill individuals, population distribution patterns.)	Mortality Information[110] (Date of death, location of death Descriptive information (age[83], gender[83], race, zipcode[83] ethnicity, occupation[83]) Outbreak information[110, 128](number ill, exposed, hospitalized: Type of outbreak date of outbreak, location of outbreak, food specific attack rates[129], location of food consumption) Illness related information (incubation period, duration of illness, symptoms of illness, submitting physician, patient status, hospital date of admission and discharge, outcome (died, alive), hospital transfer status, case-control study, treatment history, county, travel history, onset of illness, use of chemoprophylaxis) Identifying and contact Information (phone number, etc.) Disease specific information (hemophila:inhibitor/immune tolerance usage,[130] severity, functional health status, service expenditure, service utilization (hospital bed usage)[73], infection control policies, volume	Duration of hospitalization, age, sex, illness in family members, recent illness, standardized, medical history questionnaire Deaths[91, 132, 133], travel/ work history, age, sex Hospital admissions[93, 99], potential risk factors, pregnancy[134], age, sex, deaths, miscarriage[134, 135] Discharge record, histories of exposures[91], health center visits, food handler interviews[117], food preference questionnaire, food history[136-138] Hospitalization[97], birth place, age race, vaccination status[139-142], travel history[99, 113, 124, 143-145], guest list, worksite, hospital & provider referrals Age, hospitalization diagnoses, death certificate, location of residence, travel history, Sexual partners[79, 82, 146], sexual orientation[81], onset of illness, hospitalization	Epidemiologic link (e.g., ingestion of a home-canned food within the previous 48 hours) ,persons who ate the same food as persons who have laboratory-confirmed botulism (Botulism) Epidemiologically linked to a confirmed case (Brucellosis) Exposure is defined as having been (less than or equal to 30 days before onset of EM) in wooded, brushy, or grassy areas (i.e., potential tick habitats) in a county in which Lyme disease is endemic. A history of tick bite is not required. Disease endemic to county. A county in which Lyme disease is endemic is one in which at least two confirmed cases have been previously acquired or in which established populations of a known tick vector are infected with B. burgdorferi.(Listeriosis) Location infection acquired, number of episodes, acquired through artificial means (e.g., blood transfusion, common syringes, or malariotherapy), relapses, presence of other cases(Malaria) Rash onset occurs within 18 days after entering the jurisdiction, and illness cannot be linked to local transmission. Imported cases should be classified as: -International. A case that is imported from another country -Out-of-State. A case that is imported from another state in the United States. The possibility that a patient was exposed within his or her state of residence should be excluded; therefore, the patient either must have been out of state continuously for the entire period of possible exposure (at least 7-18 days before onset of rash) or have had one of the following

	Theoretical	Collected by Public Health	Used in Outbreak Investigations	Used for CDC definitions
		depletion, Device associated infection rates,[100] surgical site infection rates, incidence rates, frequency of specific TB strains geographically,[131] spread of related strains in communities,[131]geographic mobility of related strains,[131]relatedness of strains in persons at high risk for tuberculosis,[131], Laboratory related information (type and method of laboratory test, specimen collection date, submitting laboratory, source of specimen, specimen id number)	Pet history[126], location of residence, travel history[124] Visitors sign-in-log, at zoo, direct & indirect animal contact,[126, 134, 147-155] Food source[94, 95, 98, 115, 117, 119, 121, 134, 136-138, 156-160] Party attendance, Illnesses in other health care workers, Onset of illness,	types of exposure while out of state: a) face-to-face contact with a person who had either a probable or confirmed case or b) attendance in the same institution as a person who had a case of measles (e.g., in a school, classroom, or day care center). An indigenous case is defined as a case of measles that is not imported. Cases that are linked to imported cases should be classified as indigenous if the exposure to the imported case occurred in the reporting state. Any case that cannot be proved to be imported should be classified as indigenous.(Measles) Age ("Streptococcus pneumoniae, Invasive, (Children < 5 years)"; "Syphilis, (3) latent") date of initial infection("Syphilis, (3) latent")death(Toxic-shock syndrome; Varicella (deaths only)) Treatment with two or more antituberculosis medications., lost to supervision or discharged from treatment greater than 12 months (Tuberculosis) evidence or history of a tick or deerfly bite, exposure to tissues of a mammalian host of Francisella tularensis, or exposure to potentially contaminated water. (Tularemia)
Zoonotic Data	(See syndromic, incubation and pre-outbreak sections)	Reportable zoonotic diseases, see also data in the pre-outbreak, exposure and incubation data sections,	Specimens from reptiles[155]	Rabies, animal A positive direct fluorescent antibody test (preferably performed on central nervous system tissue) Isolation of rabies virus (in cell culture or in a laboratory animal) Confirmed: a case that is laboratory confirmed("Rabies, animal")
Environmental Data	Civil engineering data (Heating, ventilation and air conditioning system records) Food cultures, food handling methods, sanitary inspections, food delivery schedule	Food (source, marketing, processing, preparation, sale, packaging, refrigeration, storage, cooking, culture)[129] Water (type , chlorination, contaminated, circumstances)[128] Work stations	Sources of environmental contamination (in door gasoline-powered forklifts[75], propane stove, charcoal grill, furnace with poor venting, contaminated water[93, 161])	

	Theoretical	**Collected by Public Health**	**Used in Outbreak Investigations**	**Used for CDC definitions**
	Water supply facilities		Conditions conducive to development of infections (use of gasoline-powered auger, hand trowels, dog tunnels and burrows, lack of personal protective equipment, poor plant ventilation[75], air-conditioner system and exhaust fan, contact with manure, inadequate pool or hot tub filtration system[105, 161], poor cooking practices[95, 115, 159]) Trace back investigation (including check of water temperature at time of seafood harvest[162, 163], farm inspection[117, 164], shell fish tags, Contamination via transportation & worksites[91], Environmental sampling (mosquito trapping)[125, 144, 165, 166]	

Table A.2 Types of Information Actually Contributing to Detection of Previous Outbreaks by Type of Outbreak Type or by CDC Case Definition

Outbreak Type	Type of Information Used in Outbreaks	Referenced in CDC Case Definitions
Large-scale bioaerosol		Inhalational Anthrax: onset of symptoms, brief, prodrome resembling a viral respiratory illness, hypoxia, dyspnea, mediastinal widening, Isolation of Bacillus anthracis, Anthrax electrophoretic immunotransblot (EITB) reaction to the protective antigen and/or lethal factor bands, Demonstration of B. anthracis by immunofluorescence Brucellosis: onset of fever, night sweats, undue fatigue, anorexia, weight loss, headache, and arthralgia, isolation of Brucella sp,fourfold or greater rise in Brucella agglutination titer, date of titers, immunofluorescence of Brucella sp, type of specimen epidemiologically linked to a confirmed case Q fever: fever, rigors, myalgia, malaise, retrobulbar headache, acute hepatitis, pneumonia,meningoencephalitis, liver enzyme levels, chest film findings. endocarditis, chronic fatigue-like syndrome, antibody titer to C. burnetii phase II or phase I antigen, date of specimens, isolation of C. burnetii,,demonstration of C. burnetti by detection of antigen or nucleic acid. Tularemia: ulceroglandular: cutaneous ulcer,regional lymphadenopathy conjunctivitis, preauricular lymphadenopathy, stomatitis, pharyngitis, tonsillitis, cervical lymphadenopathy,intestinal pain, vomiting, and diarrhea, primary pleuropulmonary disease, history of a tick or deerfly bite, exposure to tissues of a mammalian host of Francisella tularensis, exposure to potentially contaminated water. Elevated serum antibody titer(s) to F. tularensis antigen, history of tularemia vaccination or -detection of F. tularensis by fluorescent assay
Aerosolized Continuous Release	Clinical symptoms e.g. fatigue, fever, arthralgia & myalgia Medical diagnosis Serology, trans bronchial biopsy, microscopic	Coccidioidomycosis fever, chest pain, cough, myalgia, arthralgia, headache) Pneumonia, other pulmonary lesion, chest radiograph result, erythema nodosum, erythema multiforme rash, bones, joints, or skin involvement, meningitis, viscera and lymph nodes involvement Legionellosis fever, myalgia, cough, pneumonia, Isolation of Legionella, type of fluid, reciprocal immunofluorescence antibody (IFA) titer against Legionella

Outbreak Type	Type of Information Used in Outbreaks	Referenced in CDC Case Definitions
	examination, lung biopsy specimens, CD4 count, chest xray abnormalities,	pneumophila serogroup 1, date of titer, Detection of L. pneumophila serogroup 1 by direct fluorescent antibody testing, type of specimen, or L. pneumophila serogroup 1 antigen by radioimmunoassay or enzyme-linked immunosorbent assay
	Duration of hospitalization, age, sex, illness in family members, recent illness, standardized, medical history questionnaire	Psittacosis fever, chills, headache, photophobia, cough, myalgia Isolation of Chlamydia psittaci from respiratory secretions, antibody against C. psittaci by complement fixation or microimmunofluorescence (MIF), presence of immunoglobulin M antibody against C. psittaci by MIF
	Gasoline-powered auger, hand trowels, dog tunnels and burrows, personal protective equipment	
Building Contamination	CO monitors Medical symptoms & diagnosis, Chest X-ray, Post mortem examination, postmortem COHB levels, sputum specimen, cultures, DNA fingerprint, toxicologic and microbiologic culture Deaths, travel/ work history, age, sex Presence of propanestove, charcoal grill, location of furnace and vent, location of forklifts, plant ventilation, air-conditioner system and exhaust fan	Anthrax: clinically compatible illness, evidence of B. anthracis DNA by polymerase chain reaction (PCR), type of specimen, demonstration of B. anthracis in a clinical specimen by immunohistochemical staining, or other laboratory tests (e.g., serology).
Food borne	Symptoms of illness e.g.	Botulism, foodborne: diplopia, blurred vision, and bulbar weakness. Symmetric

Outbreak Type	Type of Information Used in Outbreaks	Referenced in CDC Case Definitions
	nausea, headache, abdominal cramps diarrhea, vomiting, fever,	paralysis botulinum toxin, type of specimen, isolation of Clostridium botulinum , ingestion of a home-canned food, time of food ingestion, other individuals who ate same food.
	Clinical laboratory results (stool culture, cell culture assays, biologic toxicity assays, chemical analyses for toxins, pulsed-field gel electrophoresis, reverse transcriptase poly-merase chain reaction assay, enzyme immunoassay, serotyping records, plamid analysis)	Intestinal Anthrax: severe abdominal distress, fever, signs of septicemia Isolation of Bacillus anthracis , Anthrax electrophoretic immunotransblot (EITB) reaction to the protective antigen and/or lethal factor bands, onset of symptoms, number of specimens,demonstration of B. anthracis by immunofluorescence
		Cholera: diarrhea and/or vomiting; isolation of toxigenic (i.e., cholera toxin-producing) Vibrio cholerae O1 or O139 from stool or vomitus, or serology
		Enterohemorrhagic Escherichia coli: diarrhea (often bloody),a bdominal cramps. hemolytic uremic syndrome (HUS) or thrombotic thrombocytopenic purpura (TTP) isolation of Escherichia coli O157:H7 from a specimen, or Isolation of Shiga toxin-producing E. coli from a clinical specimen
	Emergency room medications	Hemolytic Uremic Syndrome: acute onset of microangiopathic hemolytic anemia, renal injury, low platelet count, central nervous system (CNS) involvement, fever , type of onset of symptoms, anemia (acute onset) with microangiopathic changes (i.e., schistocytes, burr cells, or helmet cells) on peripheral blood smear , hematuria, proteinuria, or elevated creatinine level , age.
	Hospital admissions, potential risk factors, pregnancy, age, sex, deaths, miscarriage discharge record, histories of exposures, health center visits, food handler interviews, food preference questionnaire, food history	Cyclosporiasis: watery diarrhea, loss of appetite, weight loss, abdominal bloating, cramping, increased flatus, nausea, fatigue, low-grade fever. vomiting Cyclospora oocysts by microscopic examination, type of specimen, demonstration of sporulation, or DNA (by polymerase chain reaction)
		Hepatitis, viral, acute, type of onset of illness, jaundice, elevated serum aminotransferase levels, Hepatitis A: Immunoglobulin M (IgM) antibody to hepatitis A virus (anti-HAV) positive
	Trace back investigation, water temperature, farm inspection, shell fish tags, hospital alert, cooking practices, absentee records,	Listeriosis meningitis or bacteremia: fetal loss through miscarriage or stillbirth, or neonatal meningitis or bacteremia. Isolation of L. monocytogenes, type of specimen,
		Salmonellosis:diarrhea, abdominal pain, nausea, ,vomiting. isolation of Salmonella.
		Shigellosis diarrhea: fever, nausea, cramps, and tenesmus, Isolation of Shigella
		Trichinosis: eosinophilia, fever, myalgia, periorbital edema,Trichinella larvae in tissue, type of biopsy, serologic test for Trichinella

Outbreak Type	Type of Information Used in Outbreaks	Referenced in CDC Case Definitions
Contagious Person-to person aerosol	Clinical diagnoses e.g. pneumonia ataxia, lymphadenopathy, conjunctivitis, symptoms of illness, Reports of illness from public health authorities Blood culture, autopsy, viral cultures, serologic testing, joint culture, rubella specific IGM, rapid influenza A antigen-detection tests Hospitalization, birth place, age race, vaccination status, travel history, guest list, worksite, hospital & provide referrals Contamination via transportation & worksites,	Haemophilus influenzae, invasive disease meningitis: bacteremia, epiglottitis, or pneumonia. Isolation of H. influenzae, type of site Hansen disease (Leprosy): number and type of skin lesions, type of spreading edges, type of clearing, peripheral nerve swelling or thickening also may occur, number and type of papules and nodules, an infiltration of the face, hands, and feet with lesions, type of distribution, thickening of the skin, Demonstration of acid-fast bacilli in skin or dermal nerve, type of biopsy Measles: type of rash, days of rash, days of fever, height of fever, cough, coryza, or conjunctivitis, serologic test for measles mmunoglobulin M antibody, date of titers, type of titer, isolation of measles virus, type of specimen, travel history, exposure to others with measles. Meningococcal disease meningitis and/or meningococcemia: purpura fulminans, shock, death. Isolation of Neisseria meningitides, site of isolation. Mumps: type of onset, type of swelling, location of swelling, duration of swelling, Isolation of mumps virus, or titers for mumps immunoglobulin G (IgG) antibody, mumps immunoglobulin M (IgM) antibody titers, date of titers. Pertussis duratin of cough, paroxysms of coughing, inspiratory "whoop," or post-tussive vomiting, presence of other cause., isolation of Bordetella pertussis, type of specimen, polymerase chain reaction (PCR) for B. pertussis. Rubella: Type of onset and type of rash, Temperature. arthralgia/arthritis, lymphadenopathy, or conjunctivitis, isolation of rubella virus, serum rubella immunoglobulin G antibody titers, dates of titers, serologic test for rubella immunoglobulin M (IgM) antibody Streptococcal disease, invasive, Group A pneumonia: bacteremia, type of cutaneous infection (e.g., cellulitis, erysipelas, or infection of a surgical or nonsurgical wound), deep soft-tissue infection (e.g., myositis or necrotizing fasciitis), meningitis, peritonitis, osteomyelitis, septic arthritis, postpartum sepsis (i.e., puerperal fever), neonatal sepsis, and nonfocal bacteremia. Isolation of group A Streptococcus (Streptococcus pyogenes) by culture, specimen site. Streptococcal toxic-shock syndrome, onset of illness, onset of syndrome, blood pressure, age, renal impairment: creatinine, elevation in creatinine, platelet count,

Outbreak Type	Type of Information Used in Outbreaks	Referenced in CDC Case Definitions
		clotting times, fibrinogen level, the presence of fibrin degradation products, alanine aminotransferase, aspartate aminotransferase, or total bilirubin levels, date of levels, onset of symptoms, diffuse pulmonary infiltrates, hypoxemia, cardiac failure, acute onset of generalized edema, type of effusions, hypoalbuminemia, type of rash, desquamation of rash. Soft-tissue necrosis, necrotizing fasciitis, myositis, gangrene. Isolation of group A Streptococcus. Streptococcus pneumoniae: acute otitis media, pneumonia, bacteremia, or meningitis. Isolation of S. pneumoniae, site , sensitivity Tuberculosis: tuberculin skin test, chest radiographs, type of treatment -Isolation of M. tuberculosis, type of specimen or demonstration of M. tuberculosis by nucleic acid amplification test, or demonstration of acid-fast bacilli. Varicella fatal diffuse (generalized) maculo-papulovesicular rash, Death, Isolation of varicella virus, type of specimen. direct fluorescent antibody (DFA), polymerase chain reaction (PCR), serum varicella immunoglobulin G (IgG) antibody levels, date of specimens.
Premonitory Release of Agents	Illnesses in other health care workers, abnormal blood test, chest X-ray, hepatic abnormalities, prolonged shock, pathological test Skin biopsy, white blood count, direct fluorescent antibody test, pathological examination, serological assays, PCR test, CT Scan Age, hospitalization diagnoses, death certificate, location of residence, travel	Tetanus: type of onset, hypertonia, painful muscular contractions, generalized muscle spasms, other apparent medical cause Toxic-shock syndrome: Temperature type of rash(diffuse macular erythroderma) desquamation, onset of illness, onset of desquamation, location of descquamation, hypotension, age,: orthostatic syncope, or orthostatic dizziness gastrointestinal: vomiting or diarrhea, onset of vomiting or diarrhea -Muscular: severe myalgia or creatine phosphokinase level, vaginal, oropharyngeal, or conjunctival hyperemia -Renal: blood urea nitrogen, urinary tract infection, total bilirubin, alanine aminotransferase enzyme, or asparate aminotransferase enzyme levels: platelets, disorientation or alterations in consciousness, focal neurologic signs.

Outbreak Type	Type of Information Used in Outbreaks	Referenced in CDC Case Definitions
	history, environmental sampling, laboratory mishap	
Sexually Transmitted Diseases	Contact tracing Onset of illness, symptoms of illness, unexplained fever PCP diagnosis, toxoplasmosis in gay men Medications prescribed , PCR testing, DNA sequencing, phylogenetic analysis, antimicrobial susceptibility testing Sexual partners, sexual orientation, onset of illness, hospitalization	AIDS/HIV: Screening test for HIV antibody (enzyme immunoassay), confirmatory (sensitive and more specific) test for HIV antibody (e.g., Western blot or immunofluorescence antibody test) HIV virologic (nonantibody) tests: HIV nucleic acid (DNA or RNA) detection (e.g., DNA polymerase chain reaction [PCR] or plasma HIV-1 RNA), HIV p24 antigen test, including neutralization assay, HIV isolation (viral culture) Chlamydia trachomatis:, genital infections, urethritis, epididymitis, cervicitis, acute salpingitis, lymphogranuloma venereum and trachoma.: Isolation of C. trachomatis by culture, or demonstration of C. trachomatis by detection of antigen or nucleic acid, type of specimen. Gonorrhea: urethritis, cervicitis, or salpingitis. isolation of typical gram-negative, oxidase-positive diplococci (presumptive Neisseria gonorrhoeae) Demonstration of N. gonorrhoeae by detection of antigen or nucleic acid, Observation of gram-negative intracellular diplococci, type of specimen, sex. Hepatitis, viral, acute: type of onset of symptoms ,jaundice, serum aminotransferase levels, Hepatitis B: IgM antibody to hepatitis B core antigen (anti-HBc) positive or hepatitis B surface antigen (HBsAg) positive IgM anti-HAV negative (if done) Hepatitis C: Serum aminotransferase levels, IgM anti-HAV negative, and IgM anti-HBc negative or HBsAg negative, and antibody to hepatitis C virus (anti-HCV) positive, more specific assay Non-A, Non-B hepatitis: Serum aminotransferase, IgM anti-HAV, IgM anti-HBc HBsAg negative, Anti-HCV negative, Delta hepatitis: HBsAg or IgM anti-HBc ,antibody to hepatitis delta virus Syphilis, number and type of ulcers, mucocutaneous lesions, lymphadenopathy. central nervous system infection inflammatory lesions of the cardiovascular system, skin, and bone, nontreponemal: Venereal Disease Research Laboratory [VDRL] or rapid plasma reagin [RPR]; treponemal: fluorescent treponemal antibody absorbed

Outbreak Type	Type of Information Used in Outbreaks	Referenced in CDC Case Definitions
		[FTA-ABS] or microhemagglutination assay for antibody to T. pallidum [MHA-TP]) Demonstration of T. pallidum by darkfield microscopy, direct fluorescent antibody (DFA-TP), or equivalent methods, type of specimen
Vector/Host borne	Autopsy, direct florescent antibody, DNA sequencing, blood culture, pulsed-field gel electrophoresis; Intercellular parasites Pet history, location of residence, travel history Visitors sign-in-log, at zoo, direct & indirect animal contact, Specimens from reptiles, mosquito trapping	Ehrlichiosis, type of onset of fever, headache, myalgia, malaise. nausea, vomiting, or rash, thrombocytopenia, leukopenia, liver enzymes. Intracytoplasmic bacterial aggregates (morulae) in the leukocytes of some patients. antibody titer to E. chaffeensis antigen by indirect immunofluorescence assay (IFA) ,positive polymerase chain reaction (PCR) assay, confirmation of E. chaffeensis DNA, identification of morulae in leukocytes, IFA titer to E. chaffeensis antigen, or immunostaining of E. chaffeensis antigen in a biopsy or autopsy sample, or culture of E. chaffeensis, specimen,. antibody titer to E. phagocytophila antigen by IFA , type of specimens, date of specimens, PCR assay and confirmation of E. phagocytophila DNA, or Identification of morulae in leukocytes, IFA titer to E. phagocytophila antigen oriImmunostaining of E. phagocytophila antigen in a biopsy or autopsy sample, or Culture of E. phagocytophila from a clinical specimen. Ehrlichiosis (other or unspecified agent): antibody titer to more than one Ehrlichia species by IFA in paired serum samples, or identification of an Ehrlichia species other than E. chaffeensis or E. phagocytophila by PCR, immunostaining, or culture. Encephalitis or meningitis, arboviral: Fever, headache, stiff neck, pleocytosis, altered mental status, confusion, coma, paresis or paralysis, cranial nerve palsies, sensory deficits, abnormal reflexes, generalized convulsions, and abnormal movements. virus-specific serum antibody titer, date of titers,isolation of virus from or demonstration of specific viral antigen or genomic sequences in tissue, blood, cerebrospinal fluid (CSF), or other body fluid, or Virus-specific immunoglobulin M (IgM) antibodies demonstrated in CSF by antibody-capture enzyme immunoassay (EIA), or Virus-specific IgM antibodies demonstrated in serum by antibody-capture EIA and confirmed by demonstration of virus-specific serum immunoglobulin G (IgG) antibodies in the same or a later specimen by another serologic assay (e.g., neutralization or hemagglutination inhibition). Hantavirus Pulmonary Syndrome: temperature, bilateral diffuse interstitial edema, chest XRAY, supplemental oxygen, time until oxygen required, previous health

Outbreak Type	Type of Information Used in Outbreaks	Referenced in CDC Case Definitions
		status, unexplained respiratory illness resulting in death, autopsy examination, noncardiogenic pulmonary edema, other identifiable cause, hantavirus-specific immunoglobulin M, hantavirus-specific immunoglobulin G, date of titers, detection of hantavirus-specific ribonucleic acid sequence by polymerase chain reaction in clinical specimens, or Detection of hantavirus antigen by immunohistochemistry
		Lyme disease:Isolation of Borrelia burgdorferi from a clinical specimen or Demonstration of diagnostic immunoglobulin M or immunoglobulin G antibodies to B. burgdorferi in serum or cerebrospinal fluid (CSF). Sensitive enzyme immunoassay or immunofluorescence antibody,Western blot Erythema migrans. fatigue, fever, headache, mildly stiff neck, arthralgia, or myalgia. Objective joint swelling, number of joints, duration of swelling. Lymphocytic meningitis; cranial neuritis, particularly facial palsy, radiculoneuropathy; encephalomyelitis, high-grade (2nd-degree or 3rd-degree) atrioventricular conduction defects, duration and onset myocarditis. Type of exposure to wooded, brushy or grassy area, , days of exposure, location of exposure.
		Malaria fever. headache, back pain, chills, sweats, myalgia, nausea, vomiting, diarrhea, and cough. coma, renal failure, pulmonary edema, and death. Demonstration of malaria parasites in blood filmsLocation infection acquired, number of episodes, acquired through artificial means (e.g., blood transfusion, common syringes, or malariotherapy), relapses, presence of other cases
		Rabies, animal direct fluorescent antibody test,isolation of rabies virus (in cell culture or in a laboratory animal)
		Rabies, human encephalomyelitis, coma, death,direct fluorescent antibody of viral antigens, type of specimen, isolation of rabies virus, identification of a rabies-neutralizing antibody titer., vaccination status.
		Rocky Mountain spotted Fever: acute onset, myalgia, headache, and petechial rash, location of rash, antibody titer to Rickettsia rickettsii antigen by immunofluorescence antibody (IFA), complement fixation (CF), latex agglutination (LA), microagglutination [127], or indirect hemagglutination antibody (IHA), dates of titers.positive polymerase chain reaction assay to R. rickettsii, or demonstration of positive immunofluorescence of skin lesion (biopsy) or organ tissue (autopsy), or

Outbreak Type	Type of Information Used in Outbreaks	Referenced in CDC Case Definitions
		Isolation of R. rickettsii from clinical specimen Yellow, temperature, graph of temperature, type of onset, symptoms hepatitis, albuminuria, , renal failure, shock, hemorrhages. yellow fever antibody titer, dates of titers, history of recent yellow fever vaccination, cross-reactions to other flaviviruses or demonstration of yellow fever virus, antigen, or genome, type of specimen.
Water Supply	Diagnosis of illness e.g. gastroenteritis PCR testing, serological assays for ITG & ITM Food source, Contact with manure, water system documentation	Cryptosporidiosis: diarrhea, abdominal cramps, loss of appetite, low-grade fever, nausea, vomiting. Cryptosporidiosis oocysts by microscopic examination,type of specimens, oocyst or sporozoite antigens by immunodiagnostic methods, e.g., ELISA, or by PCR techniques, or demonstration of reproductive stages in tissue preparations. Giardiasis: diarrhea, abdominal cramps, bloating, weight loss, or malabsorption. demonstration of G. lamblia cysts or -demonstration of G. lamblia trophozoites, type of specimen, demonstration of G. lamblia antigen by a specific immunodiagnostic such as enzyme-linked immunosorbent assay (ELISA) Typhoid fever: sustained fever, headache, malaise, anorexia, relative bradycardia, constipation or diarrhea, nonproductive cough isolation of S. typhi, type of specimen
Recreational Water	Reportable disease e.g. folliculitis/dermatitis Stool culture, pulsed-field gel electrophoresis Party attendance Pool or hot tub, filtration system	Polio, paralytic: flaccid paralysis, number of limbs, type of onset, type of reflexes, other causes, sensory or cognitive loss. neurologic deficit, duration, death,

Appendix B: Interviews Conducted

CHAPTER 4 – Clinical Data Systems

L.Ursini,
CIO
US Quality Algorithms
Middletown, CT

P. Gaffney,
Information Technology,
Blue Cross/Blue Shield,
Philadelphia PA

H. Bowers,
Informatics Division
Quest Diagnostics Inc.
Baltimore MD

(Various sales staff)
Clinsoft Inc.
Chandler AZ

A. Ferretti,
VP, Marketing
Sunquest Information Systems
Tucson AZ

CHAPTER 5 – Emergency Call Centers

R. Hicxon,
Technical Issues Director
National Emergency Number Association
Columbus, OH

W.Glover,
President
Association of Public Safety Officials
Communication International Inc.
Daytona Beach, FL

R. Lehr,
TRW Public Safety Communications Systems
Reston VA

W. Lutz,
Senior Analyst
Camden Police Department.
Camden NJ

Lt. F.Church,
Commanding Officer,
Communications Division
New York City Police Department
New York NY

Major J. Dove,
US Army
Chief of Uniformed Police
Fort Dix, NJ

J. Trego,
Sales Representative
Globalstar USA
Walnut Creek, CA

R. Allenbach,
Chief of Police
Camden Police Dept.
Camden NJ

Sgt. First Class C. Zila
US Army Military Police
Fort Belvoir, VA

G. Lama,
Manager
General Motors – OnStar
Pontiac, MI

G. Aupperlee, EMT-P
EMS Program Specialist
Emergency Medical Services Office
Pennsylvania Dept. of Health
Harrisburg PA

J. Bimblis, President
Software Development
CIA Co.
Coroapolis PA

M. Godin
Director, Product Management & System Engineering
CML Emergency Systems
Quebec, Canada

D. Plante
Director of Engineering
Kinzler Ritter and Plante LLC
Philadelphia PA

CHAPTER 7 – Food, Water

L. Kosty,*
Director of Regulatory Affairs,
American Meat Institute,
Arlington VA

K. Payne, *
Food Safety and Inspection Service,
USDA
Washington DC

R. Cook, PhD, *
National Manager, Microbiology,
MAF Food Assurance Authority,
Wellington, New Zealand

P. Lutweiler, PhD.,
Senior Manager of Water Resources,
Philadelphia Suburban Corp.,
Bryn Mawr PA

G. Brock,
Director of Laboratories,
Philadelphia Water Department
Philadelphia PA

L.D. McMullen,
CEO, DeMoine Water Works,
DeMoine IA

K. Schmeck,
Director of Water Quality Management
Montgomery County Health Dept.
Norristown PA

**also interviewed for Chapter 14*

CHAPTER 8 - Pharmaceutical Sales

S. Om,
Account Manager,
IMS Health, Blue Bell, PA

M. Weintraub,
Account/Sales Manager
NDC Health
Atlanta GA

F. Pignataro, MD,
Senior Therapeutic Team Manager
Medical/Professional Information Dept
Pfizer Inc, Morris, Plains, NJ

B. Cobert, PhD
Schering-Plough Corp,
Kenilworth NJ

L. Hostelley,
Executive Director,
Adverse Events Reporting
Merck Inc., Blue Bell PA

J. Dubberley,
DUR Manager
ACS-Consultec, Atlanta GA

K. Jensen,
Account Executive
A.C. Nielsen Co.
Schaumberg IL

CHAPTER 9 - Absenteeism Reporting

J. Fetig,
Director of Media Relations
Lockheed Martin Corp., Bethesda, MD

P. Barrett,
Sales Lead Management
Peoplesoft Inc, Pleasanton CA

Lt. M. Hinesley,
US Navy
Bethesda MD

Brig. Gen. K. Schafer,
US Air Force
Washington DC

K. Kumar,
CIO,
New York City Board of Education
Brooklyn, NY

T. Brown,
Bureau of Attendance,
New York City Board of Education,
Brooklyn, NY

CHAPTER 12 – Physiologic Monitoring

A. Teller, CEO
BodyMedia
Pittsburgh PA

T. Blackdar, CEO
Fit Sense Technologies
Southboro, MA

S. Joffe, President,
Fit Sense Technologies
Southboro, MA

J. Genovese,
Chief of Advanced Technology
United Technologies Corp./Hamilton
Sundstrand Space Systems Intl
Windsor Locks, CT

E. Hodgson,
Project Engineering Manager
Advanced Extra-Vehicular Activity
Technology
United Technologies Corp./Hamilton
Sundstrand Space Systems Intl
Windsor Locks, CT

CHAPTER 14 – Monitoring of Animals

K. C. Hinshaw, DVM, MPVM
Vice President,
Animal Health
Philadelphia Zoo, Philadelphia PA

D. Travis, DVM
Veterinarian, Lincoln Park Zoo
Chicago IL

T. McNamara, DVM
Veterinarian
New York City Wildlife Conservation Society, NY

D. Stremme, VMD
Veterinarian, NJ State Aquarium at Camden (also in private practice)
Camden NJ

R.R. Bunke,
Controller
Impromed Inc.
Oshkosh WI

P. Hughes,
Consultant
Animal Intelligence Inc.
Silverdale, WA

J.A. Teare, DVM,
Designer, MedArks
Associate Veterinarian
Jacksonville Zoo
Jacksonville, FL

K. Mullin, RHIA, CTR,
Medical Records Consultant
Veterinary Hospital of the University of Pennsylvania
Philadelphia PA

M. Khan, DVM
Veterinary Medical Officer
Food Safety and Inspection Service/Technical Service Center
US Dept. of Agriculture
Omaha, NE

William O'Brien
Ow ner
Texas Beef Producers Group
Amarillo TX

David Huseman,
CIO
Texas Beef Producers Group
Amarillo TX

Jeff Tharpe,
CIO
ConAgra Beef Co.
Greeley CO

CHAPTER 15 – Coroners and Medical Examiners

H. Mirchandani, MD,
Chief Medical Examiner
Philadelphia PA

I. Hood, MD.
Deputy Medical Examiner
Philadelphia, PA

Appendix C. Research Team

Michael M. Wagner, M.D., Ph.D.
Principal Investigator
Assistant Professor of Medicine and Intelligent Systems
Center for Biomedical Informatics
Department of Medicine
University of Pittsburgh

Michael P. Allswede, D.O.
Visiting Associate Professor
Section Chief, Special Emergency Medical Response
Department of Emergency Medicine
University of Pittsburgh

Kathleen M. Carley, Ph.D.
Professor of Sociology
Department of Social and Decision Sciences
Carnegie Mellon University

Wendy W. Chapman, Ph.D.
Postdoctoral Fellow
Center for Biomedical Informatics
Department of Medicine
University of Pittsburgh

Virginia M. Dato, M.D., MPH
Assistant Professor of Health Services Administration
Senior Public Health Physician
Center for Public Health Practice
Graduate School of Public Health
University of Pittsburgh

Jeremy U. Espino, M.D.
Postdoctoral Fellow
Center for Biomedical Informatics
Department of Medicine
University of Pittsburgh

Douglas B. Fridsma, M.D.
Assistant Professor of Medicine
Center for Biomedical Informatics
Department of Medicine
University of Pittsburgh

Anna Goldenberg
Graduate Student
Center for Automated Learning and Discovery
Computer Science Department
Carnegie Mellon University

Edward Krenzelok, Pharm.D.
Director, Pittsburgh Poison Center
Childrens' Hospital of Pittsburgh
Professor, School of Pharmacy and Medicine
University of Pittsburgh

Zhongwei Lu, M.S., M.Ed.
Programmer Analyst
Center for Biomedical Informatics
Department of Medicine
University of Pittsburgh

Galit Shmueli, Ph.D., M.Sc.
Visiting Assistant Professor
Department of Statistics
Carnegie Mellon University

Fu-Chiang Tsui, Ph.D.
Research Assistant Professor of Medicine
Center for Biomedical Informatics
Department of Medicine
University of Pittsburgh

Xiaoming Zeng, M.D.
Graduate Student Researcher
Center for Biomedical Informatics
Department of Medicine
University of Pittsburgh

Abi Fapohunda Dr.P.H.

Principal
FOB Group, LLC
Monroeville, PA

Ron Aryel, M.D., M.B.A.
Principal Consultant, Healthcare
Del Rey Technologies LLC.
Philadelphia PA

Appendix D. External Scientific Advisory Board

We wish to thank the following individuals who have served as an external advisory board during the conduct of this research.

Bruce W. Dixon, M.D.	Associate Professor of Medicine and Director, Allegheny County Health Department (of Pennsylvania) Room 301A Juvenile Court Building 3333 Forbes Avenue Pittsburgh, Pennsylvania 15213
Charles Friedman, Ph.D.	Associate Vice Chancellor and Director, Center for Biomedical Informatics Professor of Medicine and Information Science Suite 8084 Forbes Tower University of Pittsburgh 200 Lothrop Street Pittsburgh, Pennsylvania 15213
John W. Loonsk, M.D	Associate Director for Informatics and Director, Information Resource Management Office Centers for Disease Control and Prevention 1600 Clifton Road, N.E. Mail Stop D45, Bld. 16, Room 2114 Atlanta, Georgia 30333

Major Julie Pavlin, M.D.	Chief, Department of Field Studies Walter Reed Army Institute of Research Division of Preventive Medicine Building 503, Room 3A34 503 Robert Grant Avenue Silver Spring, Maryland 20910
Dennis M. Perrotta, Ph.D	Chief, Bureau of Epidemiology Texas Department of Health 1100 West 49th Street Austin, Texas 78756-3199
Loren H. Roth, M.D., MPH	Professor, Department of Psychiatry Associate Sr. Vice Chancellor, Health Sciences University of Pittsburgh Senior Vice President, Medical Services UPMC Health Systems Suite 11082 Forbes Tower 200 Lothrop Street Pittsburgh, PA 15213
Tracee Treadwell, DVM	Epidemiologist National Center for Infectious Diseases Centers for Disease Control and Prevention 1600 Clifton Road, N.E. Mail Stop C18 Atlanta, Georgia 30333