1
CHAPTER

Introduction

Michael M. Wagner

RODS Laboratory, Center for Biomedical Informatics, University of Pittsburgh, Pittsburgh, Pennsylvania

1. INTRODUCTION

Biosurveillance is a process that detects disease in people, plants, or animals. It detects and characterizes outbreaks of disease. It monitors the environment for bacteria, viruses, and other biological agents that cause disease. The biosurveillance process *systematically* collects and analyzes data for the purpose of detecting cases of disease, outbreaks of disease, and environmental conditions that predispose to disease.

Detection of disease *outbreaks* is of particular importance. Unlike explosions, disease outbreaks are silent. Disease outbreaks sicken or kill individuals before they are detected. Disease outbreaks can inflict this damage quickly, and they can also spread quickly. The window of opportunity to limit this damage can be as brief as a few days in the worst case (Wagner et al., 2001).

The United States spends billions of dollars per year on various forms of biosurveillance. The major expenditures are for hospital infection control, public health surveillance, surveillance of the air and water, training, improvement of the information technology infrastructure for public health, and research.[1]

Biosurveillance is also a rapidly growing scientific field at the intersection of epidemiology, artificial intelligence, microbiology, computer science, statistics, system engineering, medicine, and veterinary medicine.

2. THE BIOSURVEILLANCE PROCESS

The biosurveillance process is a continuous one (Figure 1.1). An organization conducting biosurveillance collects and analyzes *surveillance data* continuously. The organization also faces decisions continuously about whether to act based on the results of these analyses.

The biosurveillance process involves a positive feedback loop: when the continuous collection and analysis of surveillance data identifies an anomalous number of sick individuals (or a single case of a dangerous disease), investigators collect additional information that feeds back into the analytic process, resulting in better characterization of the event. The improved understanding of the event may lead to more questions, which drive further collection of data and additional analyses. Concurrent with these cycles of data collection and analysis, the organization may initiate response actions such as vaccinations and quarantine to control the outbreak. The net effect of this process, when viewed over time, is a series of actions that lead to characterization and control of the outbreak.

3. THE SCOPE OF BIOSURVEILLANCE

The word *biosurveillance* is of recent origin.[2] *Biosurveillance* overlaps with two existing terms: *disease surveillance* and *public health surveillance*. These terms are defined as systematic methods for the collection and analysis of data for the purpose of detecting disease (Thacker and Berkelman, 1988; Halperin and Baker, 1992; Teutsch and Churchill, 2000).

As with any new word, we could speculate whether its invention and growing usage signals the appearance of a new field or simply reflects an inadequacy of existing terminology.

---

1 Although it is difficult to identify components in federal and other government budgets that correspond to biosurveillance, as these organizations have broader missions, the overall 2005 budget of the Centers for Disease Control and Prevention (CDC) was $8 billion and the median budget for a state health department in 2004 was $2.9 billion (Hearne et al., 2004). The total U.S. government civilian biodefense funding for 2005 was estimated at $7.6 billion (Schuler, 2004), with $452 million allocated for agricultural laboratories, monitoring, and research and $129 million allocated for air monitoring (*http://www.whitehouse.gov/omb/pdf/Homeland-06.pdf*). A 1975 estimate of the cost of hospital infection control programs in the United States (updated to 2005 dollars) was $261 million (Haley, 1977).

2 In fact, the current edition of the *Oxford English Dictionary* (OED) does not define the word *biosurveillance*. although it is in widespread usage, as evinced by Google search results (13,000 hits on May 8, 2005) as well as its routine use by government agencies, politicians, journalists, and academics. There is no doubt that *biosurveillance* has been inducted into the common vernacular. Even those without technical expertise or training in the field understand the term intuitively, just as they understand the meaning of *bioterrorism*, another word currently left undefined in the OED. The absence of a standard definition reflects the need to synthesize the multidisciplinary work being done in the field. Indeed, this book is our effort to present a unified approach to and understanding of biosurveillance.
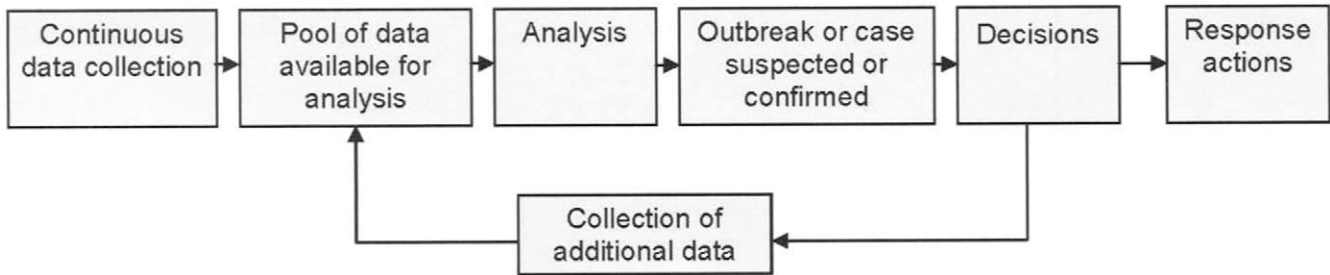
FIGURE 1.1 The biosurveillance process. When the continuous collection and analysis of surveillance data raises suspicion of an outbreak or a single case of a dangerous disease, biosurveillance personnel must decide whether to react to the information. They may decide to collect additional data that feed back into the analytic process, resulting in better characterization of the event. They may decide to take actions such as the issuance of a boil-water advisory (in the case of suspected water contamination), closure of a restaurant, or treatment of individuals with antibiotics or vaccines. Whenever the staff (or an automatic system) decides to collect additional data, the biosurveillance process exhibits a feedback loop. If the event is confirmed, staff will make many decisions over time about additional data to collect, directed by the analysis of data accumulated to that point (a positive feedback loop).

Instead, we simply state why, after considerable deliberation, we selected the term *biosurveillance* for this book. The terms *disease surveillance* and *public health surveillance* connote disease surveillance practiced by governmental public health. *Biosurveillance* allows us to broaden the scope of our discussion to include many other organizations that monitor for disease, such as hospitals, agribusinesses, and zoos. These organizations share the same basic goals of identifying individuals (people or animals) with disease, understanding disease transmission patterns in a population, elucidating the root causes of disease outbreaks, and monitoring the microbiological status of the environment. They collect similar types of data (clinical, microbiological, and environmental), use similar techniques to analyze the data, and they all face difficult decisions regarding how to react to the data. They often interact with each other to achieve the goals of disease detection and characterization. The similarities in goals and techniques suggested that we should unify them conceptually.

We also decided against *disease surveillance* and *public health surveillance* because these terms, to an epidemiologist, connote surveillance for noninfectious disease, child mortality, injury, cigarette smoking, and dental diseases such as enamel fluorosis (CDC, 2005). To keep what was already a very large topic manageable, we decided against discussing surveillance for these conditions. The principles and techniques that we discuss, nevertheless, apply to surveillance for any disease or condition.

Importantly, we decided against *disease surveillance* and *public health surveillance* because we consider outbreak characterization (i.e., determining the organism, source, route of transmission, spatial distribution, and number of affected individuals) a key process in biosurveillance. Epidemiologists may not consider outbreak characterization as falling under disease surveillance or public health surveillance (e.g., Buehler, 1998: Chapter 22; Teutsch and Churchill, 2000). To an epidemiologist, the process of public health surveillance detects an outbreak, and then an investigation characterizes it. *Biosurveillance,* as we use the term, encompasses both detection and characterization.

As we will discuss later, processes that detect outbreaks also partially characterize them. Future advances in biosurveillance techniques will facilitate even better characterization of an outbreak at the time it is first detected. The blurring of the boundary between detection and characterization suggested that we should unify these processes conceptually.

4. FUNDAMENTAL PROPERTIES OF THE BIOSURVEILLANCE PROCESS

In addition to its continuous, cyclic nature, the biosurveillance process has several other properties, which are frequent themes in this book. These properties have implications for the design of a biosurveillance system, the types of computer systems that are required to support biosurveillance, and the training of individuals working in this field:

- Multidisciplinary
- Multiorganizational
- Time critical
- Probabilistic
- Decision oriented
- Data intensive
- Dependent on information technology
- Knowledge intensive
- Complex

4.1. Multidisciplinary

To conduct biosurveillance (i.e., to collect and analyze data to detect cases and outbreaks and characterize outbreaks), an organization must draw on the expertise of individuals with diverse professional backgrounds: epidemiologists, physicians, nurses, veterinarians, computer scientists, statisticians, water quality specialists, biologists, and microbiologists. Table 1.1 summarizes the expertise and training in different biosurveillance tasks (e.g., diagnosis of individuals and "diagnosis" of outbreaks) of many of the professionals that participate in biosurveillance. These individuals have different backgrounds

TABLE 1.1 Professional Training Related to Biosurveillance Tasks

Professional	Diagnosis* (Human)	Diagnosis* (Animal)	Epidemiological Analysis	Interpretation of Laboratory Tests	Environmental Monitoring	Information Technology†	Other Training	Typical Degrees
Epidemiologist	Varies†	Varies	Extensive	Moderate	Moderate	Limited	Limited forensics	MPH, PhD, DPH
Physician§	Extensive	—	Limited	Extensive	—	—	Disease reporting	MD, DO
Nurse practitioner	Extensive	—	—	Extensive	—	—	Disease reporting	RN, BSN
Registered nurse	Moderate	—	—	Limited	—	—	Disease reporting	MSN
Radiologist	Extensive	—	—	Extensive		—	Disease reporting	MD, DO
Pathologist	Extensive	—	—	Extensive		—	Forensics, disease reporting	MD, DO
Medical technologist	Limited	—	—	Conducts of tests		Limited	Disease reporting	BS
Hospital ICP¶	Limited	—	Of hospitals	Extensive	Extensive	—	Disease reporting	BS
Veterinarian	Cross-over diseases	Extensive	Limited	Extensive		Minimal	Disease reporting	DVM,
Farm worker	—	Informal	—	—	—	—	—	Ag science
Microbiologist	—	—	—	Extensive	Varies	—	—	PhD
Medical examiner	Extensive	—	—	Extensive		None	Forensics	MD
Water inspector	—	—	—	—	Extensive	–	Water systems	BS
Food inspector	—	—	—	—	Extensive	–	Food systems	BS
Systems analyst	—	—	—	—	—	Extensive	Functional requirement definition	BS or MS
Database administrator	—	—	—	—	—	Extensive	—	BS or MS
Computer scientist	—	—	—	—	—	Extensive	—	PhD
Computer programmer	—	—	—	—	—	Extensive	—	BS or MS
Decision analyst	—	—	—	—	—	Extensive	Decision analysis	MS/PhD
Economist	—	—	—	—	—	—	Economics	PhD
Medical informaticist	Varies	Varies	Varies	Extensive	—	Extensive	Clinical information systems	PhD/PhD
Lawyer	—	—	—	—	—	—	Law	JD
Ethicist	—	—	—	—	—	—	Ethics	Varies

Extensive means that the skill is a primary objective of the professional training for this individual typically involving many courses and apprenticeship experience.

*Diagnosis means that the individual is trained to establish a diagnosis from symptom, sign, and laboratory data.

†*Information technology* refers to training in programming, system architecture, and algorithms.

‡*Varies* Epidemiologists and medical informaticians often have dual degrees, including MD or DVM (Doctor of Veterinary Medicine) so the ability related to diagnosis may vary widely.

§*Physicians* refers to physicians that are practicing medicine. There are thousands of MDs who have expertise in other areas such as environmental testing (e.g., physicians who supervise environmental testing by industrial hygienists, Association of Occupational and Environmental Medicine [ACOEM]).

¶*Hospital ICP*: hospital infection control practitioner.

and may have difficulty communicating with each other unless their training includes exposure to the roles, methods, and concepts used by other individuals.

Many of these individuals have different "day jobs." Most physicians, for example, work in medical care, not biosurveillance. They, nevertheless, must be competent in the skills related to biosurveillance, such as diagnosing rare diseases and reporting the existence of a person with a communicable disease to the appropriate authorities. The professional and continuing education of physicians, veterinarians, and medical technologists must ensure that they have the necessary training in the diseases and procedures related to their biosurveillance roles—skills that will allow them to operate as components in a larger system whose goal is disease and outbreak detection and characterization. Conversely, a biosurveillance system must ensure that these individuals have the information that they need when they need it.

4.2. Multiorganizational

Biosurveillance of just a single city requires the cooperation of many organizations, including hospitals, infection control units within hospitals, laboratories within hospitals, medical practices, commercial laboratories, water suppliers, and health departments (Table 1.2).

In a region the size of the United States, there are more than 7,500 hospitals (U.S. Census Bureau, 2005), 40,000 long-term

care facilities (National Center for Health Statistics, 1986), 160,000 water departments, and 185,000 clinical laboratories (CLMA, 2005). There are 3,000 local public health agencies and 60 state tribal, or territorial health departments (Hearne et al., 2004). In addition, 1.3 million farms carry livestock (Kellog, 2002), and there exist large numbers of manufacturers and distributors of food and drugs.

These organizations work collaboratively in the service of biosurveillance by communicating, exchanging data, and acting in concert during outbreaks. As we will see in Chapter 2, disease does not respect national boundaries. Outbreaks of diseases, such as severe acute respiratory syndrome (SARS) and influenza, spread quickly around the world. The set of organizations that may have to communicate and exchange information about an international outbreak may be 100-fold larger than the set that would have to collaborate for a nationwide outbreak, and the difficulty of coordination among these organizations is exacerbated by differences in languages, customs, and laws.

An engineer or a computer scientist reading the previous paragraphs would immediately characterize biosurveillance as a highly distributed process. The engineer would realize that each person and organization must perform specific functions for such an arrangement to work, and that were she to attempt to improve the process, significant attention would need to be paid to communication and coordination among the components.

TABLE 1.2 Responsibilities of Organizations for Monitoring of Environment, Detection of Cases, Detection of Outbreaks, and Outbreak Characterization

Organization	Monitoring Water, Food, Drug, Air	Case Detection	Outbreak Detection	Outbreak Characterization
Hospital	Primary in hospital	Primary	Primary in hospital, otherwise supportive	Primary in hospital, otherwise supportive
Laboratory	Supportive	Primary	Supportive	Supportive
Coroner	—	Primary	Supportive	Supportive
State/local health department	Primary	Primary (public health clinics/contact tracing)	Primary	Primary
CDC	Supportive	Supportive	Primary for multistate rare isolates	Primary for bioterrorism
WHO	Supportive	Supportive	Supportive	Supportive
FDA	Primary for drugs	Primary	Supportive	Primary for food or drug trace-back
Water Supplier	Primary (water)	—	Supportive	Supportive
EPA	Supportive, water	—	—	Supportive
Drug Manufacturer	Primary for drugs	—	—	Supportive
Food Manufacturer	Primary for food	—	—	Supportive
USDA	Primary for food, meat	—	—	Primary for agriculture
State department of agriculture	—	Through state laboratories	Supportive	Primary for agriculture
Large farm	Primary for farm	Primary for farm	Primary for farm	Primary for farm
Animal hospital	Primary in hospital	Primary	Primary in hospital, otherwise supportive	Primary in hospital, otherwise supportive
Zoo	Primary for zoo	Primary for zoo	Primary for zoo	Primary for zoo
DHS	Primary for air (bioterrorism)	—	—	Primary for bioterrorism
U.S. Postal Service	Primary for facilities	—	—	Supportive
DOD	Primary for DOD	Primary for DOD	Primary for DOD	Primary for DOD
Transit system	Primary for system	—	—	Supportive

CDC indicates Centers for Disease Control and Prevention; WHO, World Health Organization; FDA, Food and Drug Administration; EPA, Environmental Protection Agency; USDA, U.S. Department of Agriculture; DHS, Department of Homeland Security; DOD, Department of Defense; and Transit, airlines and mass transit systems.

4.3. Time Critical

Early detection of outbreaks is perhaps the most important requirement for a biosurveillance system. Morbidity and economic loss accumulate rapidly, beginning with the first sick individual. In a worst-case scenario, such as a surreptitious aerosol release of the organism *Bacillus anthracis* by a terrorist on a large city, hundreds of thousands of individuals would be exposed nearly simultaneously to a biological organism that is lethal and fast acting (Figure 1.2).

The implications of time criticality are profound for the design of biosurveillance systems. Reducing the time delay between the start of an outbreak and its detection is a key goal of research and development in biosurveillance. This requirement for early detection pervades the design of new biosurveillance systems, which are designed to collect and analyze new types of surveillance data in real time.

4.4. Probabilistic

Because early detection is important, biosurveillance increasingly involves analysis of novel types of data, such as sales of diarrhea remedies and numbers of visits to emergency departments for respiratory complaints. These data are more difficult to interpret than are definitive diagnoses (e.g., a patient has anthrax) because the former are not diagnostically precise. Detection of an outbreak depends on noticing an increase in the numbers of sales or visits relative to usual levels.

The challenge of early detection is that most outbreaks present weaker signals (increases) in the data streams earlier in the outbreak than they do in the middle of the event or after the event. This means that earlier detection requires the detection of smaller signals.

Early (and reliable) detection is necessarily probabilistic because early detection requires detection when signals are small and when few signal sources may yet be active. The goal is to detect a case or an outbreak before the signals are large enough and present in enough data streams for detection to be 100% certain.

The assessment of the probability that an anomalous event is occurring places demands on a biosurveillance system and its algorithms. Analytic techniques must handle multiple, independent, yet correlated data streams. The need for probabilistic detection from multiple data streams strongly suggests the need for a detection system based on Bayesian inference. A well-organized Bayesian approach allows for rational combination of many small indicators into a big picture. We discuss Bayesian methods in detail in this book.

4.5. Decision Oriented

Biosurveillance does not exist in a vacuum. Its purpose is to collect and analyze information that people use to guide decision making and action. Biosurveillance personnel make decisions under time pressure. They make decisions based on incomplete and uncertain information. Early in the course of an outbreak,
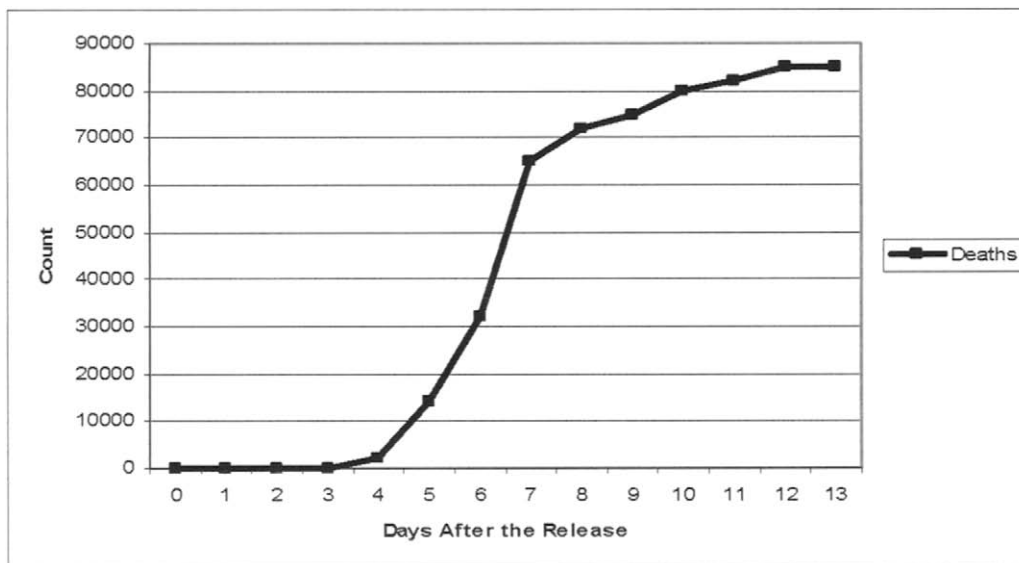

FIGURE 1.2 Hypothetical cumulative mortality from a surreptitious aerosol release of *Bacillus anthracis* by a terrorist on a major city. Such a release could expose hundreds of thousands of individuals nearly simultaneously to a biological organism that is lethal and fast acting. The window of opportunity to detect this event and administer antibiotics to those exposed is brief. (We estimated the shape of this curve from published data on the incubation period and mortality observed in the 1979 release of *B. anthracis* [*Kirov* strain] from Soviet Biological Weapons Compound 19 described in Chapter 2 and in the 2001 U.S. postal attacks).

they may not know the cause of the illness in patients, the number of affected individuals in the community, or the source of the infections. Nevertheless, they must form conjectures and hypotheses based on the available information and make decisions about how to direct resources to investigate, treat, and even quarantine individuals.

Psychological research has shown that human decision makers perform most poorly under conditions of uncertainty and time pressure. The effect of uncertainty on decision making can be profound as demonstrated by tabletop exercises (see Inglesby et al., 2001; O'Toole et al., 2002). Fortunately, the sciences of decision making and of economics provide methods to improve decision making under uncertainty. These methods elucidate the tradeoff between the risk of waiting and the cost of taking the wrong action. Biosurveillance organizations can use these methods to develop guidelines for such decision situations, or they can build these methods into computer systems that provide decision support to frontline personnel facing specific decisions. We discuss the science of decision making and economic studies in detail in Part V of this book.

4.6. Data Intensive

It is perhaps obvious, but worth stating, that biosurveillance is not a vaccine or drug that can save lives directly. Biosurveillance is a process that collects and analyzes data to guide the application of vaccines, drugs, quarantine, and other disease control strategies that can save lives.

The role of biosurveillance in disease control is to gather and process data—to collect, communicate, and analyze data. A chain of data-processing steps links raw surveillance data to "actionable information," as illustrated by Figure 1.3. The link between biosurveillance and response occurs at the point that biosurveillance personnel make decisions to act.

Each step in the chain may involve information systems and people–all of which must function effectively if an outbreak is to be quickly detected, characterized, and controlled. Any breakdown or delay in the chain can reduce the efficacy of the biosurveillance system and its ability to contribute, ultimately, to the prevention of mortality and morbidity.

4.7. Dependent on Information Technology

Societies, especially cities, have conducted biosurveillance in some form for centuries, so it is self-evident that organizations can conduct biosurveillance without the assistance of information technology. However, information technology is of increasing importance in biosurveillance because it can address the problem of time criticality. Information technology has the potential to speed up and improve the accuracy of almost every aspect of the biosurveillance process. Information technology can assist or fully automate data collection, transmission, storage, and communication. It can assist or partially automated even the most cognitively challenging steps–patient diagnosis, outbreak detection, outbreak characterization, and decision making.

4.8. Knowledge Intensive

Biosurveillance is a knowledge-intensive process. To diagnose a patient with an infectious disease, a physician must be familiar with the symptoms, signs, radiological characteristics, and laboratory tests for hundreds of diseases. This information can fill several large textbooks and requires years of study to master. A veterinarian must master an even larger body of knowledge as veterinary medicine concerns large numbers of animal species. An epidemiologist must similarly master a large body of knowledge, including a subset of human and animal diseases, as well as the subject of epidemiology, which concerns patterns of disease transmission. This knowledge also fills large textbooks, as does the knowledge required for the conduct of infection control in hospitals.

The human ability to master and apply large bodies of knowledge varies but, in general, is imperfect. Fortunately, there are technologies such as *diagnostic expert systems* and *knowledge-based* systems that professionals in many fields use to extend the range of their competencies. We discuss these technologies in Chapter 13.

4.9. Complex

The biosurveillance process is complex. There is complexity inherent in a system that distributes its functions over a large number of individuals and organizations. There is cognitive
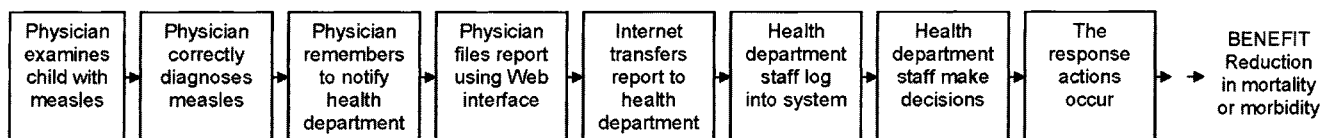
FIGURE 1.3 The indirect connection between information and benefit. A physician evaluates a child with measles. The physician must correctly diagnose the patient and remember to notify the local department of health. If the physician uses a Web-based disease reporting system, the local computer, the Internet, and the health department information systems must be functioning. Staff must review the report and make correct decisions about collection of additional information and appropriate control measures to institute. At some time later, the benefit of the information is realized by control of the outbreak and reduction in the level of morbidity and possibly mortality.

complexity inherent in reasoning and taking decisions from partial and uncertain information. There is complexity due to the number of biological agents that can cause disease and the myriad ways that they can present as outbreaks (e.g., airborne pattern, food-contamination pattern, subway system—contamination pattern, mail system–contamination pattern, and building-contamination pattern).

This complexity makes it difficult to design a biosurveillance system. In the past, organizations and people have managed the complexity of biosurveillance by specialization and prioritization. *Specialization* is a divide-and-conquer technique in which people or organizations manage complexity by, for example, creating separate biosurveillance capabilities for communicable diseases and for water-borne diseases. Specialization is not without its drawbacks, as demonstrated by the existence of many specialized information systems that cannot interoperate. *Prioritization* refers to paying more attention to certain diseases, which is a polite way of saying that, to some extent, people manage complexity by sometimes ignoring it.

One of the key benefits of information technology in professional domains, such as engineering, medicine, and biosurveillance, is that it can help to manage complexity for the professional working in that field. By managing both data and knowledge, information technology can make previously impossible or Herculean tasks possible. Information technology is a way of managing the ever-increasing complexity of biosurveillance without relying as heavily on specialization and prioritization. We discuss information systems that manage data throughout this book, and we examine systems that assist biosurveillance personnel with analytic and cognitive tasks in Parts III and V.

5. BIOSURVEILLANCE SYSTEMS

The definitions of *biosurveillance, disease surveillance,* and *public health surveillance* all include the word *systematic.* A system is any organized way of doing something. Because of the numbers of individuals, organizations, and steps in the biosurveillance process, a basic property of biosurveillance is that it is systematic. A biosurveillance system may be manual, automated, or, more commonly, a mixture of manual and automated processes. Biosurveillance systems of all types exist. The systematic, process-oriented nature of biosurveillance can be represented diagrammatically, as illustrated in Figure 1.4, which represents a highly automated system. The developers of manual biosurveillance systems often represent the organization and flow of information in a system diagrammatically as well. The diagrammatic representation of biosurveillance systems finds its fullest expression in the concept of an *architecture* for a biosurveillance system. We discuss architecture in detail in Chapter 33.

6. SCIENTIFIC FOUNDATIONS OF BIOSURVEILLANCE

The scientific foundations of biosurveillance have evolved over the centuries, parallel to advances in medicine, microbiology, veterinary science, laboratory science, epidemiology, mathematics, and many other fields.

Over the past 5 years, the scientific foundations of biosurveillance have changed rapidly. Bioterrorism and the threat posed by emerging infectious diseases triggered this change by creating a new requirement—very early detection of disease outbreaks (Wagner et al., 2001). New techniques are being introduced rapidly from diverse scientific fields and
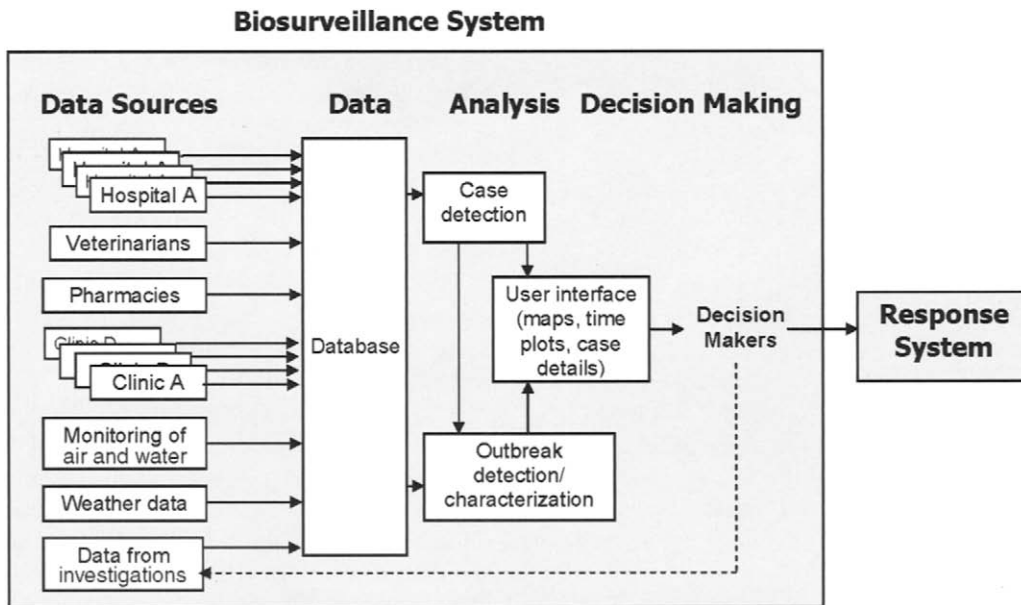
FIGURE 1.4 A generic biosurveillance system. The key elements of a biosurveillance system are data sources, a database, analysis, and decision making.

include mathematical models of the process of medical diagnosis and of decision making, as well as mathematical models of the process of "epidemic" diagnosis.

Perhaps the most important new techniques can be traced to Ledley and Lusted (1959), who first introduced the idea that medical diagnosis and decision making could be modeled mathematically. This idea spawned a large body of research about how physicians use diagnostic information, how a computer could represent medical knowledge, and how to construct computer programs that perform medical (and veterinary) diagnosis. Approximately 5 years ago, it became apparent that these same techniques could be applied to *epidemiological* diagnosis and decision making (Wagner et al., 2001). These techniques represent new core subject matter for the professional training of researchers and practitioners. We discuss these new approaches in Parts III and V.

This recent expansion of the scientific foundations of biosurveillance has been abrupt and large. The philosopher of science Kuhn termed such changes *paradigm shifts* (Kuhn, 1962). A paradigm shift is associated with changes in the curriculums of professional schools, the structure and functions of organizations, the appearance of new journals, the workforce, and the tables of contents in standard textbooks. There is evidence of such changes in biosurveillance (Logan-Henfrey, 2000; Yasnoff et al., 2001; Wagner, 2002).

7. OPEN RESEARCH PROBLEMS IN BIOSURVEILLANCE

Biosurveillance has been an active area of research since the seminal work of John Snow (1855). However, the recent requirement for very early detection caused a change in the direction (and intensity) of research. Researchers now focus on improving the timeliness and accuracy of case detection, outbreak detection, and outbreak characterization. Researchers are developing more rapid and accurate diagnostic tests, methods for sensing microbes or their effects in the environment, new detection algorithms that can extract maximum signal from early but noisy data, and research to identify types of surveillance data that provide an earlier indication of an outbreak.

Examples of the questions that research attempts to answer related to detection of individual cases of disease include the following: *What are the optimal data to collect to detect a case of disease X? What is the optimal analytic method to detect a case of disease X? What are the sensitivity, specificity, and timeliness of the current best methods for detecting a case of disease X?*

Research pursues the same set of questions for *outbreaks* of disease but also pursues additional questions, such as the following: *When can we expect to detect an outbreak of disease X that affects 1% of the population by analysis of some class of surveillance data? What is the smallest outbreak of X that we can detect?*

8. THE ROLE OF BIOSURVEILLANCE IN BIODEFENSE

Biodefense is a set of activities that together function to provide security against disease due to biological agents. Biosurveillance

is one of these activities—along with sanitation, vaccination, quarantine, intelligence, interdiction (of terrorists and materiel), forensic science, and control of technologies used to create biological weapons.

Many organizations, in addition to biosurveillance organizations, play a role in biodefense, including governmental public health (in its response role), intelligence agencies, the police, the military, and pan-national organizations, such as the World Health Organization.

9. ORGANIZATION OF THE BOOK

We have organized this book into six parts.

Part I: The Problem of Biosurveillance comprises this introductory chapter and Chapters 2 through 4. Chapter 2 ("Outbreaks and Investigations") provides examples of outbreaks that have been investigated by governmental public health, hospital infection control, and the animal healthcare system. Chapter 3 ("Case Detection, Outbreak Detection, and Outbreak Characterization") provides an overview of the basic tasks of biosurveillance, explaining in detail the methods used to detect and characterize the outbreaks described in Chapter 2. Chapter 4 ("Functional Requirements for Biosurveillance") discusses biosurveillance from the perspective of a system analyst or engineer.

Part II: Organizations that Conduct Biosurveillance and the Data They Collect (Chapters 5–12) discusses governmental public health, the human healthcare system, the animal healthcare system, laboratories, water departments, the food and drug industries, and other organizations that conduct biosurveillance. The chapters discuss the types of professionals that work in these organizations, the organizations themselves, and the information systems used by these organizations.

Part III: Data Analysis (Chapters 13–20) discusses methods for detection of individual cases, methods for detecting anomalous numbers of cases in a population, and methods for elucidating characteristics of outbreaks. The first two chapters discuss algorithms for detection of individual cases ("Case-Detection Algorithms") and the simplest algorithms for detecting outbreaks ("Classical Time-Series Methods for Biosurveillance"). The last chapter (Chapter 20) discusses methods for evaluating both case-detection and outbreak-detection algorithms. The remaining chapters cover more advanced topics, including spatial scanning, multivariate analysis, atmospheric dispersion modeling, natural language processing, and Bayesian biosurveillance.

Part IV: Newer Types of Biosurveillance Data (Chapters 21–28) discusses what research has found about the value of newer types of biosurveillance data, such as school absenteeism, sales of over-the-counter medications, and data from sensors (including physiological sensors and remote sensing from space-based satellites). Because many of these types of data are still the subject of active research, we devote the first chapter in Part IV to research methods for evaluating surveillance data.

Part V: Decision Making (Chapters 29–31) discusses the types of decisions faced by biosurveillance personnel, the types of errors in judgment to which decision makers are prone, and formal methods for modeling decisions. Part V uses an extended example of a common decision problem that is currently at the forefront in the biosurveillance community: Whether and how to react to anomalies in newer types of surveillance data.

Part VI: Building and Field Testing Biosurveillance Systems (Chapters 32–37) covers implementation issues. To avoid hearing "Professor, you left out a whole bunch of stuff,"[3] the final part of this book covers pragmatic issues related to building biosurveillance systems. Although data and analysis are the foundations of biosurveillance, organizations that wish to build biosurveillance systems must attend to proper architectural design, use of standards, legal issues, and project management. Chapter 37 discusses methods for field testing of operational biosurveillance systems.

10. SUMMARY

We can perhaps best summarize this introductory chapter about the biosurveillance process and its role in biodefense with an analogy.

An individual or an organization that is operating a biosurveillance system is like a military commander who cannot see directly every threat he faces. The commander relies on "surveillance systems" that include reports from frontline units, reconnaissance, and sensing systems located in aircraft or in orbit to make tactical decisions. Both the military commander and a biosurveillance organization are decision makers who make high-stakes decisions under time pressure by using incomplete and uncertain information. Both the military commander and the biosurveillance organization have the ability to increase the information available (by sending out additional reconnaissance or by initiating an outbreak investigation), but that new information comes at a price, which involves both the cost of the investigation as well as the "cost" of waiting for additional information before acting. This latter cost can be quite high should either disease or a military opponent gain the upper hand during the delay. Interestingly, the technologies that the military and biosurveillance organizations use–signal processing, risk-benefit analyses, methods for decision support—are either already similar or converging.

Like most analogies, this one breaks down the further one goes into detail. The level of training required for the surveillance task in the military is less than that required for the biosurveillance task, and the data required by the commander and the biosurveillance organization are completely different.

Nevertheless, like a military commander, an organization that conducts biosurveillance is primarily action oriented and conducts biosurveillance as a means, not as an end.

Biosurveillance is a means to the end of protecting health. It is an indispensable means to that end, which is absolutely dependent on the quality and timeliness of biosurveillance.

ADDITIONAL RESOURCES

Teutsch, S. and Churchill, R. (2000). *Principles and Practice of Public Health Surveillance*. Oxford: Oxford University Press. Provides an interesting account of the history of public health surveillance, as well as a description of the practice of this surveillance as of the time of writing.

Bennett, J. and Brachman, P., eds. (1998). *Hospital Infections*. Philadelphia: Lippincott-Raven. A textbook for hospital infection control practitioners.

O'Carroll, P., et al., eds. (2003). *Public Health Informatics and Information Systems*. New York, Springer. Discusses informatics principles as they relate to public health practice, as well as the current and future role of information technology in public health practice

REFERENCES

Buehler, J. (1998). Surveillance. In *Modern Epidemiology*, 2nd ed. (K. Rothman and S. Greenland, eds.). Philadelphia: Lippencott-Raven.

Centers for Disease Control and Prevention [CDC]. (2005). Surveillance for Dental Carries, Dental Sealants, Tooth Retention, Edentulism, and Enamel Fluorosis–United States, 1988–1994 and 1999–2002. *The Morbidity and Morality Weekly Report* vol. 54:1–48.

Halperin, W. and Baker, E. L. J., eds. (1992). *Public Health Surveillance*. Van New York: Nostrand Reinhold.

Hearne, S., Segal, L., Earls, M., and Unruh, P. (2004). *Ready or Not? Protecting the Public's Health in the Age of Bioterrorism*. Washington, DC: Trust for America's Health.

Inglesby, T. V., Grossman, R., and O'Toole, T. (2001). A Plague on Your City: Observations from TOPOFF. *Clinical Infectious Diseases* vol. 32: 436–445.

Kellog, R. (2002). Profile of Farms with Livestock in the United States: A Statistical Summary. Washington, DC: Natural Resources Conservation Service, U.S. Department of Agriculture. *http://www.nrcs.usda.gov/technical/land/pubs/livestockfarm.html.*

Kuhn, T. (1962). *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press.

Ledley, R. S. and Lusted, L. B. (1959). Reasoning foundations of medical diagnosis. *Science* vol. 130:9–21.

Logan-Henfrey, L. (2000). Mitigation of bioterrorist threats in the 21st century. *Annals of the New York Academy of Sciences* vol. 916:121-33.

National Center for Health Statistics. (1986). Inventory of Long-Term Care Places/National Master Facility Inventory (NMFI) Public-Use Data Files. Hyattsville, MD: National Center for Health Statistics, Centers for Disease Control and Prevention.

---

3   Rodney Dangerfield playing Thornton Mellon in the film *Back to School*.

O'Toole, T., Mair, M., and Inglesby, T. V. (2002). Shining Light on "Dark Winter." *Clinical Infectious Diseases* vol. 34:972–983.

Snow, J. (1855). *On the Mode of Communication of Cholera.* London: John Churchill.

Teutsch, S. and Churchill, R. (2000). *Principles and Practice of Public Health Surveillance.* Oxford: Oxford University Press.

Thacker, S. and Berkelman, R. (1988). Public health surveillance in the United States. *Epidemiology Review* vol. 10:164–190.

U.S. Census Bureau. (2005). Facts for Features. CB05-FFSE.02-2. Bethesda, MD: U.S. Census Bureau. *http://www.census.gov/ Press-Release/www/releases/archives/facts_for_features_special_ editions/004491.html.*

Wagner, M. M. (2002). The space race and biodefense: Lessons from NASA about big science and the role of medical informatics.

Journal of the American Medical Information Association vol. 9:120–122. *http://www.jamia.org/cgi/content/full/9/2/120.*

Wagner, M. M., et al. (2001). The emerging science of very early detection of disease outbreaks. *Journal of Public Health Management Practice* vol. 7:51–59. *http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db =PubMed&dopt=Citation&list_uids=11710168.*

Yasnoff, W. A., et al. (2001). A national agenda for public health informatics. *Journal of Public Health Management Practice* vol. 7:1–21. *http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd= Retrieve&db=PubMed&dopt=Citation&list_uids=11713752.*