1 Disease Surveillance, a Public Health Priority

Joseph S. Lombardo, David Ross

Pandemic influenza, West Nile virus, severe acute respiratory syndrome (SARS), and bioterrorism are a few of the current challenges facing public health officials. The need for early notification of, and response to, an emerging health threat is gaining increasing visibility as public opinion increases the pressure to reduce the mortality and morbidity of health threats. With the greater emphasis on the early recognition and management of health threats, federal, state, and local health departments are turning to modern technology to support their disease surveillance activities. Several modern disease surveillance systems are in operational use today. This book presents the components of an effective automated disease surveillance system and is intended for use by public health informatics students, masters of public health students interested in modern disease surveillance techniques, and health departments seeking to improve their disease surveillance capabilities.

This introductory chapter provides an overview of the changing requirements for disease surveillance from the perspective of past, present, and future concerns. It includes a brief history of how technology has evolved to enhance disease surveillance, as well as a cursory look at modern disease surveillance technology and activities.

1.1 INTRODUCTION

Control of infectious diseases is a cornerstone of public health. Various surveillance methods have been used over the centuries to inform health officials of the presence and spread of disease. The practice of disease surveillance began in the Middle Ages and evolved into the mandatory reporting of infectious disease cases to authorities responsible for the health of populations.

A common definition of surveillance is "the ongoing systematic collection, analysis, and interpretation of outcome-specific data for use in planning, implementation, and evaluation of public health practice" [1]. One of the more challenging aspects of public health surveillance is the early identification of infectious disease outbreaks..."
that have the potential to cause high morbidity and mortality. In recent years, concern over potential uncontrolled outbreaks due to bioterrorism or the appearance of highly virulent viruses such as avian influenza has placed increased pressure on public health officials to monitor for abnormal diseases. Public concern was heightened when at the beginning of the twenty-first century, the dissemination of a biological warfare agent through the U.S. mail system revealed weaknesses in the ability of existing public health surveillance systems to provide early detection of a biological attack.

Containment of potential outbreaks is also confounded by advances in transportation technology. Modern transportation systems permit communicable diseases to be carried around the world in hours over many public health jurisdictions. Health authorities can no longer simply be concerned only with the health status of the populations they serve; they must also cooperate and collaborate in surveillance and containment activities at regional, national, and international levels.

The Internet is an enabling technology for collaboration across wide geographic areas. Information technology in general is also playing a vital role in the timely capture and dissemination of information needed for identification and control of outbreaks. The subject of this book is the use of modern information technology to support the public health mission for early disease recognition and containment.

1.2 THE EMERGING ROLE OF INFORMATICS IN PUBLIC HEALTH PRACTICE

For more than 50 years, public health has been undergoing a change in identity that strongly affects how the public health sector envisions the use of information technologies. Public health is best viewed as an emergent industry. It has grown from a collection of single-purpose disease prevention and intervention programs to a national network of professionals linked through professional and organizational bonds. The 1988 Institute of Medicine report titled "The Future of Public Health" recognized that public health was established around three core functions and 10 essential services.

The core functions are:

- Assessment
- Assurance
- Policy development

The 10 essential services are:

1. Monitor health status to identify and solve community health problems.
2. Diagnose and investigate health problems and health hazards in the community.
3. Inform, educate, and empower people about health issues.
4. Mobilize community partnerships to identify and solve health problems.
5. Develop policies and plans that support individual and community health efforts.
6. Enforce laws and regulations that protect health and ensure safety.
7. Link people to needed personal health services and assure the provision of health care when otherwise unavailable.
8. Assure a competent public health and personal health care workforce.
9. Evaluate effectiveness, accessibility, and quality of personal and population-based health services.
10. Research for new insights and innovative solutions to health problems.

Information is one of the central products produced by public health. Protecting community health, promoting health, and preventing disease, injury and disability require vigorous monitoring and surveillance of health threats and aggressive application of information and knowledge by those able to prevent and protect the public’s health. Thus, public health informatics supports the activities, programs, and needs of those entrusted with assessing and ensuring that the health status of entire populations is protected and improves over time.

Public health informatics has been defined as the systematic application of information and computer science and technology to public health practice [2]. The topic supports the programmatic needs of agencies, improves the quality of population-based information upon which public health policy is based, and expands the range of disease prevention, health promotion, and health threat assessment capability extant in every locale throughout the world [3]. In the future, public health informatics may change to be defined as informatics supporting the public’s health, a discipline that may be practiced beyond the walls of the health department.

In 1854, John Snow conducted the first comprehensive epidemiological study by linking the locations of cholera patients’ homes to a single water pump. In doing so, he established that cholera was a waterborne disease. Using visual data, Snow quickly convinced the authorities to remove the pump handle. Following that simple intervention, the number of infections and deaths fell rapidly [4].

Over the past 30–50 years, public health programs have emerged around specific diseases, behaviors, or intervention technologies (e.g., immunization for vaccine-preventable diseases), each having specific data and information needs. Not surprisingly, information systems were developed to meet the specific needs of each categorical program, and a culture of program-specific information system design permeated public health thinking. By the mid-1990s, leaders in public health acknowledged the need to rethink public health information systems, conceive of systems as support tools for enterprise goals, and do so through nationally adopted standards. As noted in [3] "Public health has lagged behind health care delivery and other sectors of industry in adopting new information technologies, in part because public health is a public enterprise depending on funding action by legislative bodies (local, state, and federal).
Additionally, adoption of new technologies requires significant effort to work through government procurement processes.” A 1995 Centers for Disease Control and Prevention (CDC) study reported that integrated information and surveillance systems “can join fragments of information by combining or linking the data systems that hold such information. What holds these systems together are uniform data standards, communications networks, and policy-level agreements regarding confidentiality, data access, sharing, and reduction of the burden of collecting data” [5].

In the late 1990s, it became apparent that public health should be more comprehensive in understanding disease and injury threats. Reassessing its information mission has led federal programs such as CDC and Health Resources Service Administration (HRSA), to view information system integration as the driver for future information system funding. Integration across programs and organizations requires interoperability: data from various sources being brought together, collated in a common format, analyzed, and interpreted without manual intervention. Interoperability also requires an underlying architecture for data coding, vocabularies, message formats, message transmission packets, and system security. Interoperability implies connectedness among systems, which requires agreements that cover data standards, communications protocols, and sharing or use agreements. Interconnected, interoperable information systems will allow public health to address larger aspects of the public’s health. The twenty-first century will probably be seen as the enterprise era of public health informatics. Once the domain of humans alone, the process of gathering and interpreting data should now be mediated by computers. Major advances in the quality, timeliness, and use of public health data will require a degree of machine intelligence not presently embedded in public health information systems [6].

The context in which informatics can contribute to public health progress is changing. New initiatives within public health and throughout the health care industry portend changes in how data are captured, the breadth of data recorded, the speed with which data are exchanged, the number of parties involved in the exchange of data, and how results of analyses are shared. Increasing use of electronic health record systems provides an opportunity to gather more granular, discrete data from a variety of sources, including nursing, pharmacy, laboratory, radiology, and physician notes, thereby changing the specificity and timeliness of knowledge about the distribution of risk factors, preventive measures, disease, and injury within subpopulations.

As agreements are reached on the major information architectural standards (data, transmission, and security) and appropriate approaches to governance and viable business models can be demonstrated, health information exchanges will emerge to assist and transform how health care is delivered. Public health considerations must be central to this transformation, and public health informatics will be central to how public health agencies participate in this rapidly evolving environment.
1.3 EARLY USE OF TECHNOLOGY FOR PUBLIC HEALTH PRACTICE

There are historical accounts in the bible of social distancing as a control measure to stop the spread of leprosy. During the spread of plague in Europe in the fourteenth century, public health authorities searched vessels looking for signs of disease in passengers waiting to disembark. In the United States, the practice of disease surveillance by public health inspection at immigration has been highly publicized as a result of the renovation of Ellis Island. The immigration law of 1891 required a health inspection of all immigrants coming into the United States by Public Health Service physicians. Between 1892 and 1924, over 22 million immigrants seeking to become American citizens were subject to health inspections (Fig. 1.1). The law stipulated the exclusion of "all idiots, insane persons, paupers or persons likely to become public charges, persons suffering from a loathsome or dangerous contagious diseases" [7]. Technology was limited to paper-and-pencil recordkeeping for these surveillance and control activities.

![Public health inspectors at Ellis Island looking at the eyes of immigrants for signs of trachoma. (Photo courtesy of the National Library of Medicine)](image)

Fig. 1.1 Public health inspectors at Ellis Island looking at the eyes of immigrants for signs of trachoma. (Photo courtesy of the National Library of Medicine)

1.3.1 Early Use of Analytics, Visualization, and Communications

One of the earliest technologies used in disease surveillance was the statistical interpretation of mortality data. In 1850, William Farr analyzed the 1849 cholera outbreak in London by deriving a mathematical solution using multiple causation [8]. Florence Nightingale used statistical methods to fight for reform in the British military. She developed the polar-area diagram to demonstrate the needless deaths
caused by unsanitary conditions during the Crimean War (1854–1856). Nightingale was an innovator in the collection, tabulation, interpretation, and graphical display of descriptive statistics. Figure 1.2 is Florence Nightingale’s famous diagram depicting the causes of mortality for British troops during the Crimean War. The circle in the figure is divided into wedges, each representing a month of the war. The radius of each wedge is equal to the square root of the number of deaths for the month. The area of each wedge, measured from the center, is proportional to the statistic being represented. Dark gray wedges represent deaths from “preventable or mitigable zymotic” diseases (contagious diseases such as cholera and typhus), medium gray wedges represent deaths from wounds, and light gray wedges are deaths from all other causes [9].

Another example of the use of graphics to support epidemiological investigations is the 1869 chart by C.J. Minard describing Napoleon’s ill-fated 1812–1813 march to Moscow and back [10]. Figure 1.3 is Minard’s chart. The upper portion of the chart provides the strength of the French forces as a function of time superimposed on a map of Russia. The gray band is a measure of the size and location of the force as it advanced to Moscow; the black band represents the size and location of the retreating forces. On the lower portion of the chart is a record of the temperatures that the army encountered upon their retreat. Napoleon’s army numbered 422,000 when it crossed the Polish border on the way to Russia. Only 100,000 survived to participate in the battle at Moscow. The returning army facing the Russians at the Battle of Berezina numbered only 19,000. The returning forces suffered massive casualties due to disease and hypothermia associated with the declining temperatures. Temperatures in Russia dropped to −45 degrees Celsius during the campaign.

The invention of the telegraph and Morse code in the mid-nineteenth century provided a means for rapid dissemination of information over a wide geographic area. This technology had important implications for public health surveillance. During the Spanish Flu outbreak in 1918, the telegraph and the weekly Public Health Reports became essential tools to provide the Public Health Service with surveillance data on the progression of the pandemic.

1.3.2 Early Informatics Applications in Medicine & Public Health

Medical computing applications evolved with the development of computing technology. The very earliest applications were patient records to support diagnosis and clinical laboratory work. Bruce Blum describes the objects that are processed by computers as data, information, or knowledge [11]. A data point is a single measurement, element of demographics, or physical condition made available to the computer application or analyst. Information is a set of data with some interpretation or processing to add value. Knowledge is a set of rules, formulas, or heuristics applied to the information and data to create greater understanding.
**DIAGRAM OF THE CAUSES OF MORTALITY**

**IN THE ARMY IN THE EAST.**

**AUGUST 1855 TO MARCH 1856.**

---

The area wedges are each measured from the centre as the common vertex:

- **Deaths from preventible or Mitigable Zymotic Diseases**
- **Deaths from wounds**
- **Deaths from all other causes**

**Original Diagram by Florence Nightingale, corrected by Hugh Small**

---

Fig. 1.2 Florence Nightingale’s visualization of the causes of mortality in British troops during the Crimean War. (From Cohen [9])
Fig. 1.3  C. J. Minard's 1869 chart showing the condition of Napoleon's forces during the 1812-1813 campaign (From Dursteler [10]).
Applications using data were introduced in the 1960s when the IBM 1401 mainframe computer found use in university and research settings. In the 1970s, with the advent of low-cost minicomputers, such as the DEC PDP series or Data General Nova series, computer processing applications were developed to create information to support diagnosis in various branches of medicine. Medical imaging made great advances because images could now be acquired, stored, and processed as individual pixels, permitting multidimensional slices with high resolution. In 1970, a prototype computerized tomography system, developed by Grant [12], enabled multi-axis images to be acquired of a region under investigation. By 1973, Ledley had begun development of a whole-body CT scanner, called the automatic computerized transverse scanner (ACTA), which began clinical service early in 1974 [13].

One of the initial languages developed specifically for the organization of files in the healthcare industry was the Massachusetts General Hospital Utility Multi-Programming System (MUMPS). The language was developed by Neil Pappalardo, an MIT student working in the animal laboratory at Massachusetts General Hospital in Boston during 1966 and 1967. The original MUMPS system was built on a spare DEC minicomputer. MUMPS was designed for building database applications that help programmers develop applications that use as few computing resources as possible. The core feature of MUMPS is that database interaction is built transparently into the language [14].

The Veterans’ Health Administration (VHA) adopted MUMPS as the programming language for an integrated laboratory/pharmacy/patient admission, tracking, and discharge system in the early 1980s. This system, known originally as the Decentralized Hospital Computer Program (DHCP), has been extended continuously in the years since. In March 1988, the Department of Defense launched the Composite Health Care System (CHCS), based on the VHA’s DHCP software, for all of its military hospitals [15]. DHCP and CHCS form the largest medical records archiving systems in the United States. These archives are sources of indicators of emerging diseases and outbreaks.

1.3.3 Public Health Records Archiving

In the United States, state and local health departments have taken on the role of collecting and archiving vital statistics for the populations they serve. Health departments issue certified copies of birth, death, fetal death, and marriage certificates for events that occur in their population. Many departments also provide divorce verifications and registries on adoption and act as adjudicators of paternity.

The National Center for Health Statistics (NCHS) is the lead U.S. federal government agency for collecting, sharing, and developing procedures and standards for vital statistics. The NCHS is the oldest and one of the first examples of intergovernmental data sharing in public health. The data are provided through contracts between NCHS and individual record systems operated in the various jurisdictions legally responsible for the registration of vital events: births, deaths, marriages, divorces, and fetal deaths. In the United States, legal authority for maintaining registries of vital events
and for issuing copies of birth, marriage, divorce, and death certificates resides with the states, some individual cities (Washington, DC, and New York City), and the territories [16, 17].

In 1916, the Illinois Department of Public Health (IDPH) assumed responsibility for collecting data on vital events such as live births, still births, and deaths. In 1938, the department acquired IBM tabulation equipment for the generation of vital statistics and other health data. A computer was first used in population monitoring to support the Census Bureau in tabulating data from the 1950 census. In 1962, the IDPH became the first state health department to convert its applications on tabulation equipment to the newly acquired IBM 1401 computer. Many applications were developed for the IDPH computers, one of the most famous for a large salmonellosis outbreak in 1985. The IDPH identified communications with local health departments as a major weakness to the response. As a result, a minicomputer network was established that used modems and phone lines to pass information among state and local health departments. This system was known as the Public Health Information Network [18].

1.4 GUIDING PRINCIPLES FOR DEVELOPMENT OF PUBLIC HEALTH APPLICATIONS

The Public Health Informatics Institute (PHI) was formed in 1993 with a grant from the Robert Wood Johnson Foundation. The institute helps to foster applications that provide value to public health rather than just using the latest technology for technology's sake [19]. The Institute has outlined a set of principles to assist in guiding the development and use of computer applications for public health [20]:

1. Engage all stakeholders throughout the life cycle of the project.

2. Consider the business processes and operational constraints and develop the requirements prior to system design. In other words, think logically before physically.

3. Plan for the system to be interoperable with emerging standards such as the Public Health Informatics Network.

4. Manage the project and maintain accountability through the use of detailed plans, status reports, and meetings to help focus the project on obtaining its goals.

Figure 1.4 provides a graphical representation of the PHI principles and the four major steps in the development of a public health informatics application. The first step is to determine how the new system can improve health outcomes by quantifying the health problem, developing a business case for the system, and defining the indicators for measuring success. The second step is to determine how the work will be accomplished through a series of analyses to define the workflow and business processes that will support the application. The third step is to determine the requirements for the
application through performance requirements analysis and system design. Once the system is implemented, the final step is to determine how success will be measured through an evaluation and a series of metrics that measure the performance of the system. For advanced disease surveillance systems, the Centers for Disease Control and Prevention (CDC) has developed a framework for evaluating syndromic surveillance systems that contains a series of metrics [21, 22]. The framework assumes that the system has been fully developed and operational for several years; thus, a comprehensive evaluation in the early implementation stages of the system using the framework is not possible. It is one of the most comprehensive sets of metrics developed for disease surveillance systems. See Chapter 10 for a discussion of this and other frameworks.

Fig. 1.4 Principles and approach for planning and design of an enterprise informatics system. (From Public Health Informatics Institute [20]. ©PHII)
1.5 INFORMATION REQUIREMENTS FOR AUTOMATED DISEASE SURVEILLANCE

James Jekel describes surveillance as the entire process of collecting, analyzing, interpreting, and reporting data concerning the incidence of death, diseases, and injuries and the prevalence of certain conditions whose knowledge is considered important for promoting the health of the public [23]. Most surveillance systems are developed and implemented with a clear objective of the specific outcome being sought. Examples are the linkage of specific environmental risk factors to chronic diseases such as cancer or monitoring of behavioral factors associated with the transfer of sexually transmitted diseases (STDs). As mentioned earlier, a main focus of this book is surveillance systems for the early recognition of outbreaks due to highly infectious diseases that have a potential for high morbidity and mortality, such as virulent forms of influenza or disease agents of bioterrorism. A main objective of a system developed around this focus is to reduce the number of cases by enabling the administration of prophylaxis rapidly or by allowing for social distancing to reduce the spread of disease. To achieve this objective, a disease outbreak must be recognized in the very early stages, such as influenza or during the initial symptoms of a disease like anthrax so that treatment and control efforts still have a high chance of a successful outcome. Traditional disease surveillance and response can be represented by the steps shown in Fig. 1.5. Health departments have traditionally relied on reporting from health care providers or laboratories before initiating epidemiological investigations. This surveillance approach is highly specific, but neither sensitive or timely. In the case of anthrax, preventing the mortality of those infected relies on the rapid identification and treatment of the disease.

One potential approach for early identification of abnormal disease in a community is to collect and analyze data that are not used traditionally for surveillance and may contain early indicators of the outbreak. This approach relies on capturing health-seeking information when a person becomes ill. The concept of how such a system may operate is illustrated in Fig. 1.6. The concept is based on the assumption that a pathogen is released into the environment either in the air or in the water supply. If some type of sensor is present that can detect the presence of the pathogen and determine its identity, the detection phase is complete, but it is not possible for sensors to be located everywhere. Also, environmental sensors may be of little value if the health threat is due to highly contagious persons rather than pathogens released into the environment. If biological or chemical material has been released into the environment, the effect may be seen in animals, birds, and plant life, as well as in humans. Zoonotic diseases such as West Nile virus may first present with animal illness and death before presenting in humans.

Several types of data are collected routinely for purposes other than disease surveillance could contain indicators and warnings of an abnormal health event. When continual feeds are established for these data, analytical techniques can be applied to identify abnormal behavior. Signals identified through this process can fall into several different classes, where the most important is an outbreak with the potential
Fig. 1.5 A traditional method of public health surveillance and response for infectious diseases.

Fig. 1.6 Concept for a disease surveillance system using data sources that may contain early indicators and warnings of a health event.
for high morbidity or mortality in the population. Once it has been established that
the signal is of importance, additional data are needed to understand what is occurring
before a public health response can be executed.

Following the detection of a statistical aberration in surveillance data, several
questions must be answered: What disease is present, and what agent is causing it?
What are the characteristics of the disease and what methods are used to treat the
disease? Where and when did people get infected? Was the exposure at a single
point over a short duration, or was exposure over an extended time period and a large
geographic area? Knowledge of the population at risk is also necessary to assess the
potential public health implications of a surveillance alarm. If the disease is highly
contagious, is it contagious before symptoms develop, and which persons are at risk
of being infected by contact with those initially infected? Where are those who have
been infected, and how can they be contacted? These are just a few of the questions
for which answers would be urgently needed.

Health departments need the answers to these questions to develop and execute
a response to contain an outbreak. However, surveillance systems that use non-
specific data as early indicators of disease cannot provide many answers; traditional
epidemiological investigations are still needed. The best modern disease surveillance
systems recognize this burden and attempt to collect as much data as possible to assist
investigators in pulling together as much information as possible in a timely manner.

### 1.6 HISTORICAL IMPACT OF INFECTIOUS DISEASE OUTBREAKS

Modern medicine has had a significant impact on the control of infectious disease
outbreaks. During the majority of the past century, Western countries have had
abundant supplies of vaccines and antibiotics to control emerging outbreaks. A large
outbreak of an unknown strain of an infectious disease agent or a large bioterrorist
event could overburden the ability of the medical communities to give high-quality
care to all those infected. A review of the history of significant outbreaks provides
insight into the challenges facing the public health community.

#### 1.6.1 Smallpox

One of the most significant diseases in the history of humankind is smallpox. Early
accounts of smallpox date back to 10,000 B.C., when it appeared in the agricultural
settlements of northeastern Africa [24]. Egyptian merchants helped to spread the
disease to India in the last millennium B.C. Lesions resembling smallpox were found
on the faces of mummies, including the well-preserved mummy of Ramses V, who
died in 1157 B.C.

Western civilization has been affected greatly by smallpox. The plague of Antio-
hine, around A.D. 180, killed between 3.5 and 7 million persons and coincided with
the beginning of the decline of the Roman Empire [25, 26]. Arab expansionism,
the Crusades, and the discovery of the West Indies all contributed to the spread of
smallpox. The disease was introduced into the new world by Spanish and Portuguese conquistadors and contributed to the fall of the Aztec and Inca empires. During the decade following the Spanish arrival in Mexico, the population decreased from 25 million to 1.6 million, with disease contributing significantly to the decline [27].

The diseases that ravaged Europe and Asia for centuries were for some time unknown to Native North Americans. Ultimately, infectious diseases introduced by expansionism devastated the American Indian, with the greatest number of deaths caused by smallpox — sometimes intentionally. During the Indian siege of Fort Pitt in the summer of 1763, the British sent smallpox-infected blankets and handkerchiefs to the Indians in a deliberate attempt to start an epidemic [28]. The plan to infect the Indians and quell the siege was documented in a letter written by Colonel Henry Bouquet to Sir Jeffery Amherst, the commander-in-chief of British forces in North America.

In 1796, Edward Jenner, an English physician, observed that dairymaids who contracted cowpox, a much milder disease, were immune to smallpox. With serum taken from a dairymaid, Jenner began vaccination. When it was available, vaccination became an effective way of controlling the spread of smallpox.

In 1947, the Soviet Union established its first smallpox weapons factory in Zagorsk just northwest of Moscow. Animal tests showed that fewer than five viral particles were needed to cause infection in 50 percent of subjects. In comparison, 1500 plague cells and 10,000 anthrax spores were needed to achieve the same results. By 1970, smallpox was considered so important to the biological weapons arsenal that over 20 tons were stored annually at Zagorsk for immediate use [29].

In 1967, the World Health Organization (WHO) initiated a mass vaccination program that resulted in the eradication of smallpox by 1978 [30, 31, 32]. On May 8, 1980, WHO announced that smallpox had been eradicated from the planet. Smallpox immunization programs were discontinued, and only limited quantities of the virus were retained for research purposes at the Centers for Disease Control in Atlanta and the Ivanovsky Institute of Virology in Moscow. Coincidently, the Soviet weapons program, Biopreparat, included smallpox in the weapons improvement list in its five-year 1981–1985 plan [29].

1.6.2 Plague

Bubonic plague, or Black Death, left an indelible mark on history. In 1346, there were fearful rumors of plague in the East at major European seaports. India was depopulated; Tartary, Mesopotamia, Syria, and Armenia were covered with dead bodies. The disease traveled from the Black Sea to the Mediterranean in galleys following the trade routes to Constantinople, Messina, Sicily, Sardinia, Genoa, Venice, and Marseilles. By 1348, the Black Death had taken a firm grip on Italy. Between the years 1347 and 1352, plague accounted for the destruction of one third to one half the population of Europe, approximately 25 million victims. The disease terrified the populations of European cities because it struck so swiftly and consumed a town
or city within weeks. Victims died within days in agony from fevers and infected swellings [33].

Plague had been around London since it first appeared in Britain in 1348, but in 1665, a major outbreak occurred. Two years earlier, plague ravaged Holland. Trade was restricted with the Dutch, but despite the precautions, plague broke out in London, starting in the poorer sections of the city. Initially, the authorities ignored it, but as spring turned into one of the hottest summers in recent years, the number of deaths increased dramatically. In July, over 1000 deaths per week were reported, and by August, the rate peaked at over 6000 deaths per week. A rumor that dogs and cats caused the spread resulted in a drastic reduction in their numbers, leaving the plague-carrying rats without predators.

Control measures consisted of quarantining families in their homes. When a person in a household became infected, the house was sealed until 40 days after the victim either recovered or died. Guards were posted at the door to see that no one left. The guard had to be bribed to allow any food to pass to the homes. Accounting for victims was difficult because the quarantine measures were so harsh that families were not willing to report the death of family members. Nurses went from door to door in an attempt to quantify the number dead. Estimates are that over 100,000 people (about a quarter of the population of London) perished in the outbreak. In 1666, the Great Fire of London burned down the city slums and brought the plague under control.

1.6.3 Spanish Influenza, 1918

In colonial times, laws were passed mandating the reporting of smallpox, yellow fever, and cholera [34]. By the nineteenth century, mandatory reporting at the state and federal levels became common. During the twentieth century, increasing use of vaccines and antibiotics, improvements in communication, and the dedication of individuals and organizations led to a significant decline in morbidity and mortality due to highly contagious diseases. The twentieth century also saw the pandemic or world-wide epidemic of the Spanish influenza of 1918 and the belief by government leadership that modern medicine had conquered the risk of infectious disease outbreaks by the end of the century. These beliefs led to complacency in allocating funding to improve disease surveillance activities.

There were three major pandemic influenza outbreaks in the twentieth century [34]. In 1918–1919, Spanish influenza, caused by the H1N1 subtype of the influenza A virus, infected up to one-third of the world’s populations. The pandemic erupted during the final stages of World War I and ultimately killed more people than the war. The number of dead is estimated at between 20 and 40 million, with the exact

---

1Influenza A virus subtypes are labeled by an H number and an N number. The H number represents HA antigens or hemagglutinin protein and varies from H1 to H16. The HA antigen is responsible for binding the virus to the cell. The N number represents the NA antigen, or neuraminidase enzyme, and varies from N1 to N9. The NA antigen is responsible for releasing the virus from infected cells. H1N1 is a subtype of the swine influenza virus species.
numbers unknown due to inadequate reporting. In the United States, the outbreak claimed 675,000 lives. It has been cited as the most devastating epidemic in recorded world history. More people died of influenza in a single year than in the four years of the Black Death from 1347 to 1351.

From analysis to determine the virulence of the H1N1 virus strain, a U.S. Armed Forces Institute of Pathology study determined that the Spanish influenza could first have appeared in a young British soldier during the Battle of the Somme in 1916 [35]. In 1916, supply lines stretching through the French town of Etaples comprised not only hundreds of thousands of troops but also pigsties and chicken coops to supply food for the forces. Etaples could have been the incubation site for the transfer of the virus from chickens and pigs to humans. The Institute of Pathology study also included the collection of virus samples from victims buried in the Alaska permafrost. Using documentary evidence and new genetic clues, researchers have been able to trace the flu’s spread in three waves around the world. These studies are being used to speculate about the impact of a potential H5N1 Avian Influenza pandemic [36].

Camp Funston provides a graphic example of how the 1918 pandemic ravaged communities. The 29th Field Artillery Battalion was constituted on July 5, 1918, as part of the Army’s 10th Division at Camp Funston, Kansas. There, they underwent equipment issue and tactical training and began preparations to deploy to Europe. However, during this period, Camp Funston suffered an influenza outbreak that devastated the installation. Figure 1.7 shows an emergency hospital set up at Camp Funston to care for the influenza patients. By the end of October 1918, there were 14,000 reported cases and 861 deaths in Camp Funston alone. The State of Kansas reported a total of 12,000 deaths by the time the flu had run its course and the units were healthy, the war had ended. Camp Funston was originally considered the initial site of the Spanish Influenza outbreak.

There are still several questions regarding the characteristics of the 1918–1919 pandemic. Figure 1.8 gives the mortality rate in the United Kingdom for the Spanish Flu. Three distinct waves occurred: in the spring of 1918, the fall of 1918, and the late winter of 1919. The first two waves of the pandemic occurred at a time of the year unfavorable to normal influenza virus strains. Could the virus have mutated around the world so quickly and simultaneously?

Another major difference between the pandemic strain and normal flu related to the groups affected. Mortality for influenza typically occurs among the very young or aged populations. In the 1918–1919 pandemic, disproportionate numbers of healthy young adults became victims. One theory is that earlier circulating influenza strains provided partial immunity for those exposed to a similar strain of the virus. The elderly would have been exposed to many more strains. Because most elderly could be expected to have weaker immune systems, the rates remained high. Figure 1.9 provides a comparison of the number of deaths per 100,000 persons in the United States by age group during 1911–1917 with those that occurred during 1918.
1.6.4 Influenza Pandemics after 1918

Two influenza pandemics have swept the world since 1919: the Asian influenza pandemic of 1957 (H2N2) and the Hong Kong influenza pandemic of 1968 (H3N2),
both of which were avian influenza viruses. The Asian flu pandemic probably made more people sick than the pandemic of 1918, but the availability of antibiotics to treat the secondary infections resulted in a much lower death rate. Asian flu was first identified in China in February 1957. The virus was quickly identified due to advances in scientific technology, and vaccine production began in May 1957, before the disease spread to the United States in June 1957. By August 1957, vaccine was available in limited supply in the United States. The virus claimed 1 million victims worldwide.

The Hong Kong flu pandemic strain of H3N2 evolved from H2N2 by antigenic shift. Antigenic shift is the process by which two different strains of influenza combine to form a new subtype with a mixture of the surface antigens of the two original strains. Annual flu virus mutation occurs through a process called antigenic drift, where the surface proteins change slowly over time. The body’s immune system can react to slow changes but cannot readily adapt to a rapid antigenic shift. Because of its similarity to the 1957 Asian flu and, possibly, the subsequent accumulation of related antibodies in the affected population, the Hong Kong flu resulted in far fewer casualties than in most pandemics. Casualty estimates vary; between 750,000 and 2 million people died of the virus worldwide during the two years (1968–1969) that it was active [37].

A highly virulent form of the avian virus H5N1 is currently being spread across the world by migrating waterfowl. Domestic poultry catch the virus from contact with migratory birds. Humans have caught H5N1 from close contact with infected chickens. Originally endemic only in birds in Southeast Asia, migratory patterns threaten to infect birds everywhere. Tens of millions of birds have died of the H5N1 virus, with hundreds of millions slaughtered in an attempt to control the disease.
Figure 1.10 shows an example of the flyways currently being used by migratory birds. The flyway patterns cover most populated areas of the globe.

![Map of H5N1 Outbreaks in 2005 and Major Flyways of Migratory Birds](image)

Fig. 1.10 Flyway patterns of migratory birds. (Adapted from United Nations Food and Agriculture Organization Figure [38])

The present form of the H5N1 virus does not pass efficiently between humans. However, as the virus continues to evolve, another pandemic on the order of the Spanish flu is feared. Table 1.1 presents the number of human cases of H5N1 and related deaths from 2003 until March 16, 2006. Of the 176 confirmed cases, there have been 97 fatalities, yielding a case fatality rate of 55.4%. The rate far exceeds that of previous pandemics [40].

Table 1.2 provides a list of major outbreaks considered pandemics from answers.com. There were undoubtedly many more episodes that did not make this list due to the lack of documented historical evidence prior to the eighteenth century. For the first entry, severe acute respiratory syndrome (SARS), there were fewer than 10,000 cases of the disease, but air travel spread the previously unknown contagious disease quickly.

### 1.7 DISEASE AS A WEAPON

Before the twentieth century, biological weapons were relatively simple. Infected materials were used to induce illness in an opponent's forces, or food or water supplies were poisoned. In the sixth century B.C., the Assyrians poisoned the drinking water of
Table 1.1  Cumulative Number of Confirmed Human Cases of Avian Influenza A(H5N1) Reported to WHO as of March 10, 2006. Source: World Health Organization [39].

<table>
<thead>
<tr>
<th>Year</th>
<th>Cambodia</th>
<th>China</th>
<th>Indonesia</th>
<th>Iraq</th>
<th>Thailand</th>
<th>Turkey</th>
<th>Viet Nam</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Cases</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>Deaths</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>2004</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>17</td>
<td>0</td>
<td>29</td>
<td>46</td>
</tr>
<tr>
<td>Cases</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>17</td>
<td>0</td>
<td>29</td>
<td>46</td>
</tr>
<tr>
<td>Deaths</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>20</td>
<td>32</td>
</tr>
<tr>
<td>2005</td>
<td></td>
<td>4</td>
<td>8</td>
<td>17</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>61</td>
</tr>
<tr>
<td>Cases</td>
<td>4</td>
<td>8</td>
<td>17</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>61</td>
<td>95</td>
</tr>
<tr>
<td>Deaths</td>
<td>4</td>
<td>5</td>
<td>11</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>19</td>
<td>41</td>
</tr>
<tr>
<td>2006</td>
<td></td>
<td>0</td>
<td>7</td>
<td>11</td>
<td>2</td>
<td>0</td>
<td>12</td>
<td>32</td>
</tr>
<tr>
<td>Cases</td>
<td>0</td>
<td>7</td>
<td>11</td>
<td>2</td>
<td>0</td>
<td>12</td>
<td>0</td>
<td>32</td>
</tr>
<tr>
<td>Deaths</td>
<td>0</td>
<td>5</td>
<td>10</td>
<td>2</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>24</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>4</td>
<td>15</td>
<td>28</td>
<td>2</td>
<td>22</td>
<td>12</td>
<td>93</td>
</tr>
<tr>
<td>Cases</td>
<td>4</td>
<td>15</td>
<td>28</td>
<td>2</td>
<td>22</td>
<td>12</td>
<td>93</td>
<td>176</td>
</tr>
<tr>
<td>Deaths</td>
<td>4</td>
<td>10</td>
<td>21</td>
<td>3</td>
<td>14</td>
<td>4</td>
<td>42</td>
<td>97</td>
</tr>
</tbody>
</table>

their enemies; in medieval times Mongol and Turkish armies catapulted the diseased corpses of animals or humans into fortified castles; and as late as 1710, Russian armies used plague corpses as weapons. During World War I, German agents in the United States inoculated horses and cattle with glanders before they were shipped to France for use by the Allied powers.

In 1925, the first international agreement, known as the Geneva Protocol, to limit the use of chemical and biological weapons was signed. The Protocol prohibited the use in war of asphyxiating gases and of bacteriological methods of warfare. The agreement did not address production, storage, or verification mechanisms and could not be used to support disarmament. As a result, significant research was performed in the twentieth century to increase the performance of biowarfare agents and delivery methods. Biological weapons could be developed very cheaply and cause large numbers of casualties compared with conventional weapons [41].

The Soviet Union established its biological weapons program in the late 1920s after a typhus epidemic in Russia from 1918 to 1922 killed between 2 and 10 million, illustrating graphically the destructive and disruptive power of biological weapons. From the occupation of Manchuria in 1931 to the end of World War II in 1945, the Imperial Japanese Army experimented with biological weapons on thousands of Chi-
<table>
<thead>
<tr>
<th>Year(s)</th>
<th>Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>165-180</td>
<td>Antonine plague (smallpox)</td>
</tr>
<tr>
<td>541</td>
<td>Plague of Justinian (bubonic plague)</td>
</tr>
<tr>
<td>1300s</td>
<td>The Black Death (plague)</td>
</tr>
<tr>
<td>1732-1733</td>
<td>Influenza</td>
</tr>
<tr>
<td>1775-1776</td>
<td>Influenza</td>
</tr>
<tr>
<td>1816-1826</td>
<td>Cholera</td>
</tr>
<tr>
<td>1829-1851</td>
<td>Cholera</td>
</tr>
<tr>
<td>1847-1848</td>
<td>Influenza</td>
</tr>
<tr>
<td>1852-1860</td>
<td>Cholera</td>
</tr>
<tr>
<td>1857-1859</td>
<td>Influenza</td>
</tr>
<tr>
<td>1863-1875</td>
<td>Cholera</td>
</tr>
<tr>
<td>1899-1923</td>
<td>Cholera</td>
</tr>
<tr>
<td>1918-1919</td>
<td>Spanish flu (influenza)</td>
</tr>
<tr>
<td>1957-1958</td>
<td>Asian flu (influenza)</td>
</tr>
<tr>
<td>1959-present</td>
<td>AIDS</td>
</tr>
<tr>
<td>1960s</td>
<td>El Tor (cholera)</td>
</tr>
<tr>
<td>1968-1969</td>
<td>Hong Kong flu (influenza)</td>
</tr>
<tr>
<td>1993-1994</td>
<td>Plague, Gujarat, India</td>
</tr>
<tr>
<td>2002-2003</td>
<td>SARS</td>
</tr>
</tbody>
</table>

These experiments were conducted in a disguised water purification plant known as Unit 731 at Pingfan, near the city of Harbin in northeastern China [42]. Japanese scientists tested plague, cholera, smallpox, botulism, and other diseases on prisoners. Their research led to the development of a defoliation bacillus bomb to destroy crops and a flea bomb to spread bubonic plague. Initial successes with this technology stimulated other developments, which enabled Japanese soldiers to launch biological attacks with anthrax, plague-carrying fleas, typhoid, dysentery, cholera, and other deadly pathogens. At least 14 Chinese cities were attacked with biological weapons, resulting in an estimated 10,000 to 200,000 deaths. In addition, there are firsthand accounts of the Japanese infecting civilians through the distribution of infected food and contaminated water supplies, with estimated casualties of over 580,000 from plague and cholera. Following the war, the United States granted amnesty to the Japanese
scientists in exchange for their experimentation data. Figure 1.11 shows a human vivisection experiment conducted by Unit 731 during World War II, in which a team of Japanese surgeons is removing organs while another is taking measurements on the organs.

![Figure 1.11](image)

In 1941, a biological weapons development program initiated by the United States, the United Kingdom, and Canada in response to German and Japanese weapons development activities resulted in the weaponization of anthrax, brucellosis, and botulinum toxin. During World War II, the United Kingdom developed the Allies' first anthrax bomb by experimenting with sheep on Graudn Island in Scotland. Sheep were used because they were similar in weight to humans, are highly susceptible to anthrax, and are plentiful in the area. The research left the island contaminated with anthrax spores (Fig. 1.12).

In another World War II program, termed Operation Vegetarian, the UK manufactured and planned to drop 5 million anthrax cattle cakes on German beef and dairy herds. The plan was to wipe out the German herds and simultaneously infect the German human population. Because antibiotics were not available to the general population, the operation could have caused thousands, if not millions, of human deaths. The operation was abandoned due to the success of the Normandy invasion. At the end of 1945, the British incinerated 5 million anthrax cattle cakes.

Stockpiles of biological weapons were destroyed after President Nixon unilaterally ended the United States' offensive biological warfare program. This initiative ultimately resulted in the Biological Weapons Convention in 1972. Signers of the Convention pledged to never develop, produce, stockpile, acquire, or retain biological warfare agents or the means to deliver them.
Following World War II, the Soviet Union formulated a doctrine on the production and use of biological weapons. Two types of biological weapons were developed: strategic weapons, consisting of such highly lethal agents as anthrax, smallpox, and plague, for use on deep targets inside the United States and other countries, and operational weapons, to be used to incapacitate vital civilian and military activities well behind the battlefront. The latter weapons contained agents causing diseases such as tularemia, glanders, and Venezuelan equine encephalomyelitis. Biological weapons were not considered for tactical targets because they were not immediately effective in stopping advancing forces [43].

Concern over the use of biological weapons against civilian populations resulted in a research effort within the United States. In June 1965, the U.S. Central Intelligence Agency released a harmless simulat into the New York City subway system during peak traffic periods to demonstrate the vulnerability of U.S. cities to a covert biological warfare attack. These experiments were performed in secret; commuters had no knowledge that they had been exposed to the simulat.

Despite signing the Biological Weapons Convention, the Soviet Union continued research and production of biological weapons in a program called Biopreparat [29]. The United States was unaware of the program until the first deputy director of Biopreparat, Dr. Karatjan Alibekov, defected in 1992. The program employed 30,000 in the research and development of biological weapons and antidotes. Pathogens weaponized or under development included smallpox, bubonic plague,
antirax, Venezuelan equine encephalitis, tularemia, influenza, brucellosis, Marburg virus, Ebola virus, and Machupo virus.

Documented testimony indicated that the Soviets conducted aerosol attacks on Laos, Kampuchea, and, eventually, Afghanistan using "yellow rain" (trichothecene mycotoxins) and causing thousands of deaths between 1974 and 1981. In 1979, an accidental release of Bacillus anthracis spores from the Compound 19 production facility in the town of Sverdlovsk resulted in at least 66 fatalities. The Soviets initiated mass prophylaxis of the population, burying victims using special procedures without the attendance of family members. A massive cover-up of the incident has made it difficult to reconstruct the event to determine the actual death toll, but estimates have ranged from 200 to 1000. In 1992, President Boris Yeltsin acknowledged that the Sverdlovsk incident was an accident involving aerosol release of anthrax spores.

In 1991, the United Nations' Bioweapons Inspection Team found evidence that the Iraqis were in the early stages of developing an offensive biological warfare capability. Inspectors found several state-of-the-art facilities that could have been used for agent production, as well as evidence of the weaponization of anthrax, botulinum toxin, and aflatoxin [44]. Fortunately, these weapons were not used during Desert Shield or Desert Storm. Pressure from the United Nations resulted in the destruction of the Iraqi offensive program by 1996. Several other countries have biological warfare programs in place or under development, including Russia, Israel, China, Iran, Libya, Syria, and North Korea.

1.7.1 Bioterrorism

In 1995, the religious cult Aum Shinrikyo released sarin nerve gas in a Japanese subway system. The group was subsequently found to have been developing biological weapons, including anthrax, botulism, and Q fever. Following an Ebola outbreak in 1993, the group sent cult doctors and nurses to Zaire to bring back samples of the virus for a possible biological weapon. The group staged several unsuccessful attacks using their biological weapons before resorting to sarin for the subway attack.

In September and October 2001, letters containing anthrax spores were mailed to addresses in Florida, New York City, and Washington, DC (Fig. 1.13). The incident resulted in five fatalities, with more than a dozen victims developing full-blown infections. Tens of thousands at risk of exposure were prescribed antibiotics prophylactically. The perpetrator of the attacks has not yet been identified, but it is known that the strain of anthrax was obtained from the U.S. Army Medical Research Institute for Infectious Diseases at Fort Detrick, Maryland. The letters contained 2 to 3 grams of weaponized spores of remarkable purity, indicating use of the latest technology and a well-funded and sizable research program with possible government support. The anthrax letters revealed how unprepared the public health infrastructure in the United States was to respond to acts of bioterrorism or biowarfare.
1.8 MODERN DISEASE SURVEILLANCE APPLICATIONS

1.8.1 Components of an Early Recognition Disease Surveillance System

In response to the need for earlier recognition of significant health events, health departments, academia, and information technologists have developed surveillance systems that use data which may provide early indications of disease, but are not specific enough to confirm the presence of any particular disease. These routinely collected data include records of over-the-counter (OTC) medication sales; school absenteeism; school nurse visits; 911 calls; calls to poison control centers; reports of illness from nursing homes; animal health data; health maintenance organization encounter data; and reports of chief complaints from emergency medical services and hospital emergency departments. These data sources have some features in common. For example, although they may provide an early indication of a health event, they do not typically provide a specific signal. OTC medication sales can increase due to sales promotions, consumers stocking-up, or just to movement of product displays in the store. Data generated by interactions with health care providers are typically more specific but arise only when symptoms become uncomfortable enough for a person to seek professional help. Figure 1.14 shows data sources that may contain indicators of health status. They are arranged from left to right, with the sources on the left more likely to provide an earlier but less specific indicator and the sources on the right likely to be more specific but less timely. Chapter 2 addresses the value of various data sources as indicators of events of interest for public health surveillance.

Since the attacks of September 11, 2001, in the United States, organizations acquiring data containing health indicators have been willing to provide data feeds to health departments for disease surveillance. Data can be acquired in a variety of
modes, including real-time feeds of the data via a secure connection to the facility or batched transmission where data are aggregated over time and sent periodically to the surveillance system. Chapter 3 addresses the most common data feeds as well as data privacy issues and standards used in the formatting and transmission of data.

Once data are acquired, a variety of different analytical processes can be applied to convert them into information that can be used in surveillance. Statistical algorithms are used to find anomalies in individual data streams or in many data streams where the data elements are the same, but are coming from different facilities. Examples are sales of OTC medications from stores distributed across a region or chief complaint data from hospitals distributed across the same region. Analytic techniques may also be used to fuse data or information from several data sources to look for abnormal patterns that may not be obvious in a single data stream but become evident when data sources are used together. There are also analytic techniques for identifying clusters in time and space from single or multiple sources of data. Chapter 4 provides an introduction to some of the more popular analytical techniques used in modern disease surveillance systems.

Continued operation of a disease surveillance system is an important issue for health departments. IT resources must be allocated to operate and maintain the application, and an epidemiologist must take time to review the system’s outputs. One system’s architecture may fit more readily into a health department’s business processes than others. Visualizing data in a specific format may fit more easily into a health department’s review protocol than others. Chapter 5 presents different
architectures, data processing, and visualization options available to developers of disease surveillance systems.

Because surveillance data may be nonspecific, and because algorithms detect spurious statistical anomalies as well as events of epidemiological interest, algorithms often give rise to false triggers, alarms, or alerts. The greater the number of data sources and the larger the number of algorithms applied to the data, the greater the potential for false alarms. The astute epidemiologist who is experienced in looking at local surveillance data and the alerts coming from a system can dismiss many alarms quickly. An experienced epidemiologist also can use the data and information within the system to make decisions efficiently about the health status of a population. When an epidemiologist cannot dismiss an alerts quickly, additional information may be needed to determine its importance to public health. Chapter 6 describes the business processes used by health departments to perform surveillance with nonspecific data sources.

The first place to look to for additional data to resolve a suspicious alert is the organization that provided the data causing the alert. For example, chief complaint data provided by a hospital emergency department may not contain a diagnosis or the personal identifiers needed to contact the person or persons causing the alerts. A health department can, however, request that the hospital perform a chart review to capture the information needed to resolve the alert.

Presenting large amounts of disease surveillance data or information in a manner that is comprehensible to the users of a surveillance system is a challenge. Data can be represented as the aggregate count of patients with the same syndrome, the number of OTC medication products sold, or the number of students absent. Information can be the outputs of various detector algorithms applied to one or more data streams. Information can be presented in graphical terms, such as time-series graphs of counts over time, geographic representation of counts by zip code, or census tracks overlaid on maps, along with other information needed by the user of the system.

Figure 1.15 is an example of outbreaks indicated by a time series of counts of the number of patients presenting to military clinics in San Diego County with respiratory illness. The data in the example are simulated, but they contain many of the characteristics of previous large respiratory events in the region. The example is taken from an exercise performed to evaluate the ability of the ESENCE surveillance system to identify the health status of the population during a simulated bioterrorism event (see Section 1.8.2). Several types of data can be shown on the same graph. As seen, the graph displays both the total count of patients seen at clinics and the number of patients who return after being seen some time during the previous 14 days. Activity decreases for two days (Saturday and Sunday), followed by an increase early in the week. Counts increase near the end of the time series, which is one indication of the beginning of a synthetic outbreak. The detector output is noted by the change in shade of the small dot representing the daily patient count. Two levels of threshold levels are provided as outputs from the algorithm. The grey shade represents a warning level and the black an alert level. Because time-series plots provide an easily interpreted
overview of the data, they have become an important visualization tool in modern disease surveillance systems.

![Daily Visit Counts](chart.png)

**Fig. 1.15** Example of a time-series representation of respiratory syndrome counts. The solid lines represent daily counts and the dashed lines represent counts for new patients that have not been seen for at least the previous 14 days.

An example of a geographic representation of data is provided in Fig. 1.16. A map of zip codes in San Diego County is overlaid with small squares representing the sites of medical treatment facilities. The shade of the square represents the level of activity at the facility for the syndrome of interest. The intensity of the shading represents the number of patients residing in that zip code who were seen at the treatment facilities. This technique allows spatial clusters of disease to be readily identified. Another informative representation would be the number of patients seen by zip code where people spend most of their time during the day. An example would be work zip codes. The representation may identify exposure at the worksite.

The work zip codes of persons seeking treatment are also an important demographic. These data are rarely available for analysis because most disease surveillance systems do not capture them. Working adults tend to travel large distances to work, so their working zip code is probably different from their zip code of residence. School-aged and elderly persons spend more time closer to their zip code of residence.

For regions of the country where there is a large transient population due to tourism, sporting events, or other activities, the local geographic representation of data may be of limited value. Other representations for counts and detection results would be needed for patients living outside the region under surveillance.
Fig. 1.16 Geographical representation of a simulated outbreak in San Diego County from an ESSENCE simulation.

Most modern disease surveillance systems provide some map graphing feature. The example shown in Fig. 1.16 is a geographic presentation of the data provided in Fig. 1.15. Different visualizations may be required for different users: epidemiologists reviewing the data would require detail, whereas higher level decision makers would require a summary view. Chapter 5 discusses approaches to the visualization of data used in modern disease surveillance applications.

The appropriate definition of regions for the aggregation and analysis of data in surveillance at a national or multinational level poses a problem. Algorithms that form clusters using all the zip codes or census tracts in the country could be a processing bottleneck if innovative analytical techniques are not employed. These concepts are explored in more detail in Chapter 4.

1.8.2 Modern Surveillance Applications for Use by State and Local Health Departments

In the mid to late 1990s, the fear of the reemergence of highly virulent forms of naturally occurring infectious diseases such as influenza and tuberculosis (TB), combined with the ever-increasing threat of bioterrorism, spurred increased development of disease surveillance systems. These systems focused on early detection rather than specificity of disease identification to reduce the risk of high mortality and morbidity.

One of the first systems in use was the Electronic System for the Early Notification of Community-based Epidemics (ESSENCE), which grew out of a pilot project for the Maryland Department of Health and Mental Hygiene and a preventive medicine
project at the Walter Reed Army Institute of Research [45, 46]. The initial pilot of ESSENCE was developed for surveillance during year 2000 celebrations. Development for ESSENCE included the acquisition and evaluation of several data sources that could contain early indicators of infectious diseases. An important characteristic of ESSENCE is that it was developed in close coordination with the stakeholders in health departments, taking into consideration their business processes and operational requirements. It was the first system to integrate health indicators from both the military and civilian populations. ESSENCE became operational across the Department of Defense and was implemented by the District of Columbia, Maryland, and Virginia Health departments in a network that performs surveillance across the National Capital Region. The ESSENCE software is provided free to any health department that wants to set up its own surveillance system. ESSENCE is designed to be hosted locally by health departments so that they can keep the health indicator records within their jurisdictions [46, 47].

The Real-Time Outbreak Detection System (RODS) was developed by the University of Pittsburgh in conjunction with Carnegie Mellon University. RODS was originally developed for use by large medical centers receiving real-time data feeds from emergency departments. It was converted for use by health departments in two modes. It was the first system to provide a version of its software in open-source form on the Internet for download and installation by local users. RODS is also provided as an application service provider, connecting local hospitals to archives in the RODS Laboratory at the University of Pittsburgh and providing web access to health departments using the service [48].

The New York City Health Department has responsibility for one of the largest and most concentrated populations in the United States. The city is therefore thought to be an attractive target for terrorist activities and a favorable environment for the spread of naturally occurring diseases. Following the attack on the World Trade Center in New York in September 2001, the New York City Department of Health and Mental Hygiene initiated a fully operational syndromic surveillance project to collect data from emergency departments, pharmacy chains, and other data sources [49]. During the first year of operation, the system was able to capture data from 35 hospitals covering 2.5 million patient visits, or approximately 78% of the total visits. The system was able to provide early recognition of seasonal influenza and gastrointestinal illness shortly after it became operational.

The Early Aberration Reporting System (EARS) began as a CDC initiative to provide health departments with a set of easy-to-implement analytical tools for advanced disease surveillance applications, including bioterrorism monitoring during large-scale events. Following the terrorist attacks of September 11, 2001, the EARS tool evolved into a complete standalone application for download and use by health departments. Because it is easy to download, install, and use and is available at no cost, various city, county, and state public health officials in the United States and abroad have used or are currently using the EARS application [50].

These surveillance systems use data captured for routine business purposes in the health care industry so that little additional burden is placed on the facilities providing
the data. Another model exists where data are obtained specifically for surveillance purposes. Data collected with one of these systems can be much more specific in recognizing abnormal disease occurrences. One of the first systems to exploit this feature is the Rapid Syndrome Validation Project (RSVP), developed by Los Alamos National Laboratory [31]. Physicians enter records of patient visits to a secure website. The physician is made aware of abnormal cases of disease in his or her area. The system also works with personal digital assistants (PDAs) to facilitate data entry in mobile environments. This form of data capture permits easy entry of animal health data by veterinarians and handlers on farms and ranches. This system is available commercially under the name SYRIS.

1.8.3 National Disease Surveillance Initiatives

Historically, advances in disease surveillance have been made first at the national level by federal agencies with resources and requirements sufficient to respond to political pressures regarding health matters. In the United States, the National Centers for Disease Control and Prevention (CDC) has the clearest mandate at the federal level for disease surveillance and control. The Centers have several programs for conducting advanced surveillance at the national level and supporting state and local health departments in performing their responsibilities within their jurisdictions. Support comes in the form of personnel assigned to health departments through the Epidemic Intelligence Service (EIS), which is a two-year postgraduate program of service and on-the-job training for health professionals interested in the practice of epidemiology. At least 25% of all EIS trainees are assigned to local health departments. Funding has also been provided to health departments of states and large cities through CDC cooperative agreements on public health preparedness and response for bioterrorism. These funds are intended to upgrade the preparedness of state and local public health jurisdictions for responding to bioterrorism, outbreaks of infectious disease, and other public health threats and emergencies. Many of the states have used these funds to upgrade their surveillance systems.

1.8.3.1 National Electronic Telecommunications Surveillance System

The National Electronic Telecommunications System for Surveillance (NETSS) is a computerized public health surveillance information system that provides the CDC with weekly data regarding cases of nationally notifiable diseases. Through NETSS, the CDC receives reports of notifiable diseases from the 50 state health departments, New York City, the District of Columbia, and five U.S. Territories. These reports are initiated when health care providers or laboratory directors suspect or diagnose a case of disease that is notifiable in their state. When a case of disease is reported at the local level, staff members in the local or county health department conduct further investigation, implement control measures as needed, and forward the report to the state health department.

Only designated staff in state and territorial health departments or in the New York City or District of Columbia health departments may transmit data to the CDC.
through NETSS. In some states, city and county staff enter data that will ultimately be transmitted to the CDC, but the weekly transmission of all reported data is overseen by the appropriate state or territorial health department staff. NETSS does not require the use of a specific computer software program. However, data are transmitted in common ASCII format, which allows the NETSS system to integrate data from surveillance systems throughout the United States.

Provisional weekly reports of notifiable diseases are published in the CDC's Morbidity and Mortality Weekly Report (MMWR). Final, corrected data are published in the annual MMWR Summary of Notifiable Diseases, United States [52]. The NETSS program began in 1984 as the Epidemiologic Surveillance Project. By 1989, all 50 states were reporting to the CDC.

1.8.3.2 National Electronic Disease Surveillance System In 1995, the CDC initiated the National Electronic Disease Surveillance System (NEDSS). The goal of NEDSS is the automated capture and analysis of data of public health significance from public and private health entities. The vision for NEDSS is a network of complementary electronic information systems that automatically gather health data from a variety of sources on a real-time basis to facilitate the monitoring of community health and to assist in the ongoing analysis of trends and detection of emerging public health problems. The foundation of NEDSS is a series of standards for the collection, archiving, and reporting of significant health events through the use of low-cost commercial off-the-shelf (COTS) products to support state and local systems for data collection and analysis. The NEDSS system architecture is intended to integrate and eventually replace several current CDC surveillance systems, including NETSS and systems for reporting HIV/AIDS, vaccine-preventable diseases, and tuberculosis and infectious diseases [52].

The NEDSS Base System is a platform to support state-notifiable disease surveillance and analysis activities in a secure environment. The Base System is a modular platform that provides a seamless view and management of cross-program data, supports the storage and maintenance of data in an integrated database, and supports data analysis and visualization activities through the use of specific COTS products [53]. States are not required to use the Base System, but funds provided under the CDC cooperative agreements require the use of NEDSS standards in the communication among NEDSS systems.

1.8.3.3 Public Health Information Network and BioSense The CDC's Public Health Information Network (PHIN) initiative began in 2004 with the objective of implementing a multorganizational business and technical architecture for interoperable public health information systems. PHIN includes a portfolio of standards and software solutions to build and maintain the connectivity among information systems throughout the public health sector at the local, state, and federal levels. Applications using PHIN standards include systems for disease surveillance, national health status indicators, data analysis, public health decision support, information resources and
knowledge management, alerting and communications, and the management of public health responses [54].

BioSense is a CDC initiative to perform advanced disease surveillance at the national level [55, 56]. BioSense collects and analyzes data from hospitals at the national level, Department of Defense and Veterans Health Affairs ambulatory visits, and laboratory test orders from the Laboratory Corporation of America. The application summarizes and presents data trends and data visualizations by source, day, and syndrome for each zip code, state, and metropolitan area through maps, graphs, and tables. BioSense data are analyzed at CDC’s Biointelligence Center and made available to local health departments via a secure website. Substantial investments in standards and common infrastructure are also being made through BioSense to collect real-time hospital data. A goal of BioSense is to permit hospital data feeds to be sent to local health department surveillance systems in parallel with the data feed to CDC for BioSense. Figure 1.17 provides an example of how the early event detection capabilities of BioSense fit into the framework of PHIN.

---

**Fig. 1.17  Applications using the network standards proposed by PHIN. (From CDC [55, 56])**

---

1.8.3.4 U.S. Department of Defense Disease Surveillance  The U.S. Department of Defense (DoD) operates its own version of ESSENCE for surveillance of all U.S. military treatment facilities worldwide. Data for the DoD instance of ESSENCE come from the TriCare system, which acquires data under the Composite Health Care System (CHCS) program. The system is operated by the Office of the Secretary of Defense for Health Affairs; users are provided with web access across the globe.
The DoD ESSENCE system currently provides service to approximately 800 users worldwide, making it the largest modern informatics program for disease surveillance.

1.9 SUMMARY

Public health organizations are facing increased challenges in rapidly identifying outbreaks in their communities. Health indicator surveillance data and modern information technology has helped to automatically collect, archive, process, and present summaries of a community's health status. Most implementations of automated surveillance systems lack the desired specificity or timeliness, but provide valuable information to monitors of disease surveillance. It is hoped that the information contained in the following chapters will provide insights to the readers to advance the technology and to better meet the challenges of the future.

REFERENCES


38. United Nations Food and Agriculture Organization. H5N1 outbreaks in 2005 and major flyways of migratory birds. Compiled by FAO AGAH, EMPRESS


