Infectious Disease Epidemiology

IV.1

Susanne Straif-Bourgeois, Raoult Ratard

Introduction 1.1

The following chapter intends to give the reader an overview of the current field of applied infectious disease epidemiology. Prevention of disease by breaking the chain of transmission has traditionally been the main purpose of infectious disease epidemiology. While this goal remains the same, the picture of infectious diseases is changing. New pathogens are identified and already known disease agents are changing their behavior. The world population is aging; more people develop underlying disease conditions and are therefore more susceptible to certain infectious diseases or have long term sequelae after being infected.

Infectious diseases are not restricted to certain geographic areas anymore because of the increasing numbers of world travelers and a worldwide food distribution. The fear of a bioterrorist attack adds a new dimension in infectious disease epidemiology, and health departments enhance their surveillance systems for early detection of suspicious disease clusters and for agents used as weapons of mass destruction.

Improvements in laboratory techniques and mapping tools help to expand the knowledge of transmission of disease agents and enhanced surveillance techniques are feasible as a result of software progress and reporting of diseases via secure internet sites.

Surveillance and outbreak investigations remain the major responsibilities in public health departments. Epidemiologic methods and principles are still the basis for these tasks but surveillance techniques and outbreak investigation are changing and adapting to improvements and the expanded knowledge.

Conducting surveys is a useful way to gather information on diseases where surveillance data or other data sources are not available, especially when dealing with emerging or re-emerging pathogens. Program evaluation is an important tool to systematically evaluate the effectiveness of intervention or prevention programs for infectious diseases.

The Global Burden of Infectious Diseases 1.1.1

Infectious diseases are a major cause of human suffering in terms of both morbidity and mortality. In 1995, out of an estimated total of 52 million deaths, 17 million were due to infectious diseases (WHO 2000a,b). The most common cause of infectious disease deaths were pneumonia (5 million), diarrhea (3 million) followed by tuberculosis, malaria, AIDS and hepatitis B. Not surprisingly, there is a large imbalance in diseases between developing and industrialized countries (see Table 1.1).

Morbidity due to infectious diseases is very common in spite of the progress accomplished in recent decades. Even in industrialized countries, the prevalence of infection is very high for some infectious agents. Serologic surveys found that by young adulthood the prevalence of antibodies was 80% against herpes simplex virus type 1, 15–20% against type 2, 95% against human herpes virus, 33%

Table 1.1. Proportion of principal causes of deaths

	Developing countries	Industrialized countries
Infectious diseases	43%	1%
Cardiovascular diseases	24%	46%
Cancer	10%	21%
Respiratory diseases	10%	8%

Source: WHO, World Health Report 2000

against *Hepatitis A*, 2% against *Hepatitis C*, 5–8% against *Hepatitis B*, and 50% against *Chlamydia pneumoniae* (American Academy of Pediatrics 2003; Mandell et al. 2000). Annually, approximately 267,000,000 episodes of diarrhea leading to 612,000 hospitalizations and 3000 deaths occur among adults in the United States (Mounts et al. 1999). The Center for Disease Control and Prevention (CDC) estimates that each year 76 million people in the US get sick, more than 300,000 are hospitalized and 5000 die as a result of foodborne illnesses (CDC 2004). Every year influenza circulates widely, infecting from 10% to 40% of the world population.

1.1.2 The Importance of Infectious Disease Epidemiology for Prevention

It is often said that epidemiology is the basic science of preventive medicine. To prevent diseases it is important to understand the causative agents, risk factors and circumstances that lead to a specific disease. This is even more important for infectious disease prevention, since simple interventions may break the chain of transmission. Preventing cardiovascular diseases or cancer is much more difficult because it usually requires multiple long term interventions requiring lifestyle changes and behavior modification, which are difficult to achieve.

In 1900, the American Commission of Yellow Fever, headed by Walter Reed, was sent to Cuba. The commission showed that the infective agent was transmitted by the mosquito *Aedes aegypti*. This information was used by the then Surgeon General of the US Army William Gorgas, to clean up the 200 year old focus of yellow fever in Havana by using mosquito proofing or oiling of the larval habitat, dusting houses with pyrethrum powder and isolating suspects under a mosquito net. This rapidly reduced the number of cases in Havana from 310 in 1900 to 18 in 1902 (Goodwin 1996).

A complete understanding of the causative agent and transmission is always useful but not absolutely necessary. The most famous example is that of John Snow who was able to link cholera transmission to water contamination during the London cholera epidemic of 1854 by comparing the deaths from those households served by the Southwark & Vauxhall Company versus those served by another water company. John Snow further confirmed his hypothesis by the experiment of removing the Broad street pump handle (Wills 1996a).

The Changing Picture of Infectious Disease Epidemiology

1.1.3

Over the past three decades, more than 40 new pathogens have been identified, some of them with global importance: *Bartonella henselae*, *Borrelia burgdorferi*, *Campylobacter*, *Cryptosporidium*, *Cyclospora*, *Ebola virus*, *Escherichia coli* 0157:H7, *Ehrlichia*, *Hantaan virus*, *Helicobacter*, *Hendra virus*, *Hepatitis C* and *E*, *HIV*, *HTLV-I & II*, *Human herpesvirus 6* and *8*, *Human metapneumovirus*, *Legionella*, new variant Creutzfeldt-Jakob disease agent, *Nipah virus*, *Parvovirus B19*, *Rotavirus*, severe acute respiratory syndrome (SARS) etc..

While there are specific causative agents for infectious diseases, these agents may undergo some changes over time. The last major outbreak of pneumonic plague in the world occurred in Manchuria in 1921. This scourge, which had decimated humans for centuries, is no longer a major threat. The plague bacillus cannot survive long outside its animal host (humans, rodents, fleas) because it lost the ability to complete the Krebs cycle on its own. While it can only survive in its hosts, the plague bacillus also destroys its hosts rapidly. As long as susceptible hosts were abundant, plague did prosper. When environmental conditions became less favorable (lesser opportunities to sustain the host to host cycles), less virulent strains had a selective advantage (Wills 1996b).

Changes in Etiologic Agent

The influenza virus is the best example of an agent able to undergo changes leading to renewed ability to infect populations that had been already infected and immune. The influenza virus is a single stranded RNA virus with a lipophilic envelope. Two important glycoproteins from the envelope are the hemagglutinin (HA) and neuraminidase (NA). The HA protein is able to agglutinate red blood cells (hence its name). This protein is important as it is a major antigen for eliciting neutralizing antibodies. *Antigenic drift* is a minor change in surface antigens that result from point mutations in a gene segment. Antigenic drift may result in epidemics, since incomplete protection remains from past exposures to similar viruses. *Antigenic shift* is a major change in one or both surface antigens (H and/or N) that occurs at varying intervals. Antigenic shifts are probably due to genetic recombination (an exchange of a gene segment) between influenza A viruses, usually those that affect humans and birds. An antigenic shift may result in a worldwide pandemic if the virus can be efficiently transmitted from person to person.

Changes in Populations at Risk

In the past three decades throughout the world, there has been a shift towards an increase in the population of individuals at high risk for infectious diseases.

In industrialized nations, the increase in longevity leads to higher proportion of the elderly population who are more prone to acquiring infectious diseases and developing life threatening complications. For example, a West Nile Virus (WNV) infection is usually asymptomatic or causes a mild illness (West Nile fever); rarely

does it cause a severe neuro-invasive disease. In the 2002 epidemic of West Nile in Louisiana, the incidence of neuro-invasive disease increased progressively from 0.3 per 100,000 in the 0 to 14 age group to 9 per 100,000 in the 60 to 75 year old age group and jumped to 32 per 100,000 in the age group 75 and older. Mortality rates showed the same pattern, a gradual increase to 0.7 per 100,000 in the 60 to 75 age group with a sudden jump to 11 per 100,000 for the oldest age group of 75 and older.

Improvement in health care in industrialized nations has caused an increase in the number of immune-deficient individuals, be it cancer survivors, transplant patients or people on immuno-suppressive drugs for long term auto-immune diseases. Some of the conditions that may increase susceptibility to infectious diseases are: cancers, particularly patients on chemo or radiotherapy, leukemia, lymphoma, Hodgkin's disease, immune suppression (HIV infection), long term steroid use, liver disease, hemochromatosis, diabetes, alcoholism, chronic kidney disease and dialysis patients. For example persons with liver disease are 80 times more likely to develop *Vibrio vulnificus* infections than are persons without liver disease. Some of these infections may be severe, leading to death.

In developing countries a major shift in population susceptibility is associated with the high prevalence of immune deficiencies due to HIV infections and AIDS. In Botswana which has a high prevalence of HIV (sentinel surveillance revealed HIV seroprevalence rates of 36% among women presenting for routine antenatal care), tuberculosis rates increased from 202 per 100,000 in 1989 to 537 per 100,000 in 1999 (Lockman et al. 2001) while before the HIV/AIDS epidemics, rates above 100 were very rare.

Changes in lifestyles have increased opportunities for the transmission of infectious disease agents in populations previously at low risk. Intravascular drug injections have increased the transmission of agents present in blood and body fluids (e.g. HIV, hepatitis B and C). Consumption of raw fish, shell fish and ethnic food expanded the area of distribution of some parasitic diseases. Air travel allows people to be infected in a country and be half-way around the globe before becoming contagious.

By the same token, insects and other vectors have become opportunistic global travelers. *Aedes albopictus*, the Asian Tiger mosquito, was thus imported in 1985 to Houston, Texas inside Japanese tires. Subsequently, it has invaded 22 US states.

Changes in Knowledge About Transmission of Disease Agents

With the advent of nucleic acid tests, it has become possible to detect the presence of infectious disease agents in the air and environmental surfaces. For example, the use of air samplers and polymerase chain reaction analysis has shown that *Bordetella pertussis* DNA can be found in the air surrounding patients with *B. pertussis* infection, providing further evidence of airborne spread (Aintablian et al. 1998) and thus leading to re-evaluate the precautions to be taken. However the presence of nucleic acids in an environmental medium does not automatically mean that transmission will occur. Further studies are necessary to determine the significance of such findings.

Bioterrorism Adds a New Dimension

Infectious disease agents, when used in bioterrorism events, have often been re-engineered to have different physical properties and are used in quantities not usually experienced in natural events. There is little experience and knowledge about the human body's response to large doses of an infectious agent inhaled in aerosol particles that are able to be inhaled deep into lung alveolae. During the 2001 anthrax letter event, there was considerable discussion about incubation period, recommended duration of prophylaxis, and minimum infectious dose. This lack of knowledge base has led to confusion in recommendations being made.

New Approaches in Infectious Disease Epidemiology 1.2

Although the basics of infectious disease epidemiology have not changed and the discipline remains strongly anchored on some basic principles, technological developments such as improved laboratory methods and enhanced use of informatics (such as advanced mapping tools, web based reporting systems and statistical analytical software) have greatly expanded the field of infectious disease epidemiology.

Improved Laboratory Methods 1.2.1

Molecular techniques are being used more and more as a means to analyze epidemiological relationships between microorganisms. Hence the term molecular epidemiology refers to epidemiologic research studies made at the molecular level.

The main microbial techniques used, target plasmids and chromosomes. More specifically, plasmid fingerprinting and plasmid restriction endonuclease (REA) digestion, chromosomal analysis including pulse field gel electrophoresis (PFGE), restriction fragment length polymorphism (RFLP), multi-locus sequence type (MLST) and spa typing to name a few of these techniques. Polymerase chain reaction (PCR) is used to amplify the quantity of genomic material present in the specimen. Real-time PCR detection of infectious agents is now possible in a few hours. These techniques are becoming more widely used, even in public health laboratories for routine investigations.

It is beyond the scope of this text to describe these methods in more detail.

Applications of molecular epidemiology methods have completely changed the knowledge about infectious disease transmission for many microorganisms.

The main application is within outbreak investigations. Being able to characterize the nucleic acid of the microorganisms permits an understanding of how the different cases relate to each other.

Molecular epidemiology methods have clarified the controversy about the origin of tuberculosis cases: is it an endogenous (reactivation) or exogenous (re-infection) origin? Endogenous origin postulates that *Mycobacterium tuberculosis*

can remain alive in the human host for a lifetime and can start multiplying and producing lesions. On the other hand exogenous origin theory postulates that reinfection plays a role in the development of tuberculosis. The immunity provided by the initial infection is not strong enough to prevent another exposure to *Mycobacterium tuberculosis* and a new infection leads to disease. In countries with low tuberculosis transmission, for example the Netherlands, most strains have unique RFLP fingerprints. Each infection is unique and there are hardly any clusters of infections resulting from a common source. Most cases are the result of reactivation. This is in contrast with areas of high endemicity where long chains of transmission can be identified with few RFLP fingerprinting patterns (Alland et al. 1994). In some areas, up to 50% of tuberculosis cases are the result of reinfection.

Numerous new immunoassays have been developed. They depend on an antigen-antibody reaction, either using a test antibody to detect an antigen in the patient's specimen or using a test antigen to detect an antibody in the patient's specimen.

An indicator system is used to show that the reaction has taken place and to quantify the amount of patient antigen or antibody. The indicator can be a radioactive molecule (radioimmunoassay [RIA]), a fluorescent molecule (fluorescent immunoassay [FIA]), a molecule with an attached enzyme that catalyzes a color reaction (enzyme-linked immunoassay [ELISA or EIA]), or a particle coated with antigen or antibody that produces an agglutination (latex particle agglutination [LA]).

The reaction can be a simple antigen/antibody reaction or a "sandwich" immunoassay where the antigen is "captured" and a second "read out" antibody attaches to the captured antigen. The antibody used may be polyclonal (i.e. a mixture of immunoglobulin molecules secreted against a specific antigen, each recognizing a different epitope) or monoclonal (i.e. immunoglobulin molecules of single-epitope specificity that are secreted by a clone of B cells). It may be directed against an antigen on an epitope (i.e. a particular site within a macromolecule to which a specific antibody binds).

1.2.2 Mapping as an Epidemiological Tool

Plotting diseases on a map is one of the very basic methods epidemiologists do routinely. As early as 1854 John Snow, suspecting water as a cause of cholera, plotted the cases of cholera in the districts of Golden Square, St. James and Berwick, in London. The cases seemed to be centered around the Broad Street pump and less dense around other pumps. The map supplemented by other observations led to the experiment of removing the handle on the Broad Street pump and subsequent confirmation of his hypothesis (Snow 1936).

Geographical information systems (GIS) have been a very useful tool in infectious disease research. GIS are software programs allowing for integration of a data bank with spatial information. The mapping component includes physical layout of the land, towns, buildings, roads, administrative boundaries, zip codes etc. Data

may be linked to specific locations in the physical maps or to specific aggregates. A GIS system includes tools for spatial analysis. Climate, vegetation and other data may be obtained through remote sensing and combined with epidemiologic data to predict vector occurrence.

However, these tools should be used with caution. They can be useful to generate hypotheses and identify possible associations between risk of disease and environmental exposures. Because of potential bias, mapping should never be considered as more than an initial step in the investigation of an association. "The bright color palettes tend to silence a statistical conscience about fortuitous differences in the raw data" (Boelaert et al. 1998). For statistical methods in geographical epidemiology see Chap. II.8 of this handbook.

Computer Reporting and Software Progress 1.2.3

Web based reporting, use of computer programs and developments of sophisticated reporting and analytical software have revolutionized epidemiologic data collection and analysis. These tools have provided the ability to collect large amounts of data and handle large databases. However this has not been without risks. It remains crucial to understand the intricacies of data collected to avoid misinterpretation. For example, one should be aware that diseases and syndromes are initially coded by a person who may not be very software proficient, using shortcuts and otherwise could enter data of poor quality.

What Are the Questions to Be Answered? 1.3

Too often one sees epidemiologists and statisticians preparing questionnaires, carrying out surveys, gathering surveillance information, processing data and producing reports, tables, charts and graphs in a routine fashion. Epidemiology describes the distribution of health outcomes and determinants for a purpose. It is important to question the goals and objectives of all epidemiologic activities and tailor these activities to meet these objectives.

The description of disease patterns includes analysis of demographic, geographical, social, seasonal and other risk factors.

Age groups to be used differ depending on the disease e.g. diseases affecting young children should have numerous age groups among children; sexually transmitted diseases require detailed age groups in late adolescence and early adulthood. Younger age groups may be lumped together for diseases affecting mainly the elderly. Gender categorization, while important for sexually transmitted diseases and other diseases with a large gender gap (such as tuberculosis), may not be important for numerous other diseases.

Geographical distribution is important to describe diseases linked to environmental conditions but may not be so useful for other diseases.

1.4 Surveillance Issues

Surveillance, both active and passive, is the systematic collection of data pertaining to the occurrence of specific diseases, the analysis and interpretation of these data, and the dissemination of consolidated and processed information to contributors to the program and other interested persons (CDC 2001b).

1.4.1 Passive Surveillance

In a *passive surveillance system* the surveillance agency has devised and put a system in place. After the placement, the recipient waits for the provider of care to report.

Passive case detection has been used for mortality and morbidity data for decades. It is almost universal. Most countries have an epidemiology section in the health department that is charged with centralizing the data in a national disease surveillance system collecting mortality and morbidity data.

In theory, a passive surveillance system provides a thorough coverage through space and time and gives a thorough representation of the situation. Practically, compliance with reporting is often irregular and incomplete. In fact, the main flaws in passive case detection are incomplete reporting and inconsistencies in case definitions.

The main advantages are the low cost of such a program and the sustained collection of data over decades. The purpose is to produce routine descriptive data on communicable diseases, generate hypotheses and prompt more elaborate epidemiologic studies designed to evaluate prevention activities.

Some conditions must be met to maximize compliance with reporting:

1. Make reporting easy: Provide easy to consult lists of reportable diseases, provide pre-stamped cards for reporting, provide telephone or fax reporting facilities.
2. Do not require extensive information: Name, age, sex, residence, diagnosis. Some diseases may include data on exposure, symptoms, method of diagnosis etc.
3. Maintain confidentiality and assure reporters that confidentiality will be respected.
4. Convince reporters that reporting is essential: provide feedback; show how the data are used for better prevention.

Confidentiality of data is essential, particularly for those reporting health care providers who are subject to very strict confidentiality laws. Any suspicion of failure of maintaining secure data would rapidly ruin a passive surveillance program.

1.4.2 Active Surveillance

In an *active surveillance system*, the recipient will actually take some action to identify the cases.

In an active surveillance program, the public health agency organizes a system by searching for cases or maintaining a periodic contact with providers. Regular contacting boosts the compliance of the providers. Providers are health agencies but also as in passive case detection, there may be day care centers, schools, long term care facilities, summer camps, resorts, and even public involvement. The agency takes the step to contact the health providers (all of them or a carefully selected sample) and requests reports from them at regular intervals. Thus no reports are missing.

Active surveillance has several advantages:

- It allows the collection of more information. A provider sees that the recipient agency is more committed to surveillance and is therefore more willing to invest more time her/himself.
- It allows direct communication and opportunities to clarify definitions or any other problems that may have arisen.

Active surveillance provides much better, more uniform data than passive case detection but active case detection is much more expensive (see Tables 1.2 and 1.3).

Table 1.2. Reports per 100 physicians of cases of hepatitis, salmonellosis, measles, and rubella by active and passive reporting, Monroe County (NY), 1980–1981

Disease	Active	Passive	Ratio
Hepatitis	78	27	2.9
Measles	11	8	1.4
Rubella	7	3	2.3
Salmonellosis	44	9	4.9
Total	140	48	2.9

Table 1.3. Comparison of Health Department estimated costs for active and passive surveillance systems, Vermont 1981

	Type of surveillance system	
	Active ($)	Passive ($)
Paper	114	80
Mailing	185	48
Telephone	1947	175
Personnel		
Secretary	3000	2000
Public health nurses	14,025	0
State laboratory	700	500
Post exposure prophylaxis	10,890	8250
Total	30,861	11,053

Active Surveillance Through Active Case Detection

Active surveillance systems are usually designed when a passive system is deemed insufficient to accomplish the goals of disease monitoring. This type of surveillance is reserved for special programs, usually when it is important to identify every single case of a disease. Active surveillance is implemented in the final phases of an eradication program: smallpox eradication, poliomyelitis eradication, Guinea worm eradication and malaria eradication in some countries. Active surveillance is also the best approach in epidemic or outbreak investigations to elicit all cases.

In the smallpox eradication program, survey agents visited providers, asking about suspected cases and actually investigating each suspected case. In polio eradication programs, all cases of acute flaccid paralysis are investigated.

In malaria eradication programs and some malaria control programs, malaria control agents go from house to house asking who has fever or had fever recently (in the past week or month for example). A blood smear is collected from those with fever.

1.4.3 Case Register

A case register is a complete list of all the cases of a particular disease in a definite area over a certain time period. Registers are used to collect data on infections over long periods of time. Registers should be population based, detailed and complete. A register will show an unduplicated count of cases. They are especially useful for long term diseases, diseases that may relapse or recur and diseases for which the same cases will consult several providers and therefore would be reported on more than one occasion.

Case registers contain identifiers, locating information, disease, treatment, outcome and follow-up information as well as contact management information. They are an excellent source of information for epidemiologic studies. In disease control, case registers are indispensable tools for follow up of chronic infections disease such as tuberculosis and leprosy.

The contents and quality of a case register determine its usefulness. It should contain

- Patient identifiers with names (all names), age, sex, place and date of birth, complete address with directions on how to reach the patient,
- Name and address of a "stable" relative that knows the patient's whereabouts,
- Diagnosis information with disease classification, brief clinical description (short categories are better than detailed descriptions),
- Degree of infectiousness (bacteriological, serological results),
- Circumstances of detection,
- Initial treatment and response with specific dose, notes on compliance, side effects, clinical response,
- Follow-up information with clinical response, treatment regimen, compliance, side effects,
- Locating information; for some diseases contact information is also useful.

Updating a register is a difficult task. It requires cooperation from numerous persons. Care must be taken to maintain the quality of data. It is important to only request pertinent information for program evaluation or information that would remind users to collect data or to perform an exam. For example, if compliance is often a neglected issue, include a question on compliance. Further details concerning the use of registries in general are given in Chap. I.4 of this handbook.

Sentinel Disease Surveillance 1.4.4

For sentinel disease surveillance, only a sample of health providers is used. The sample is selected according to the objectives of the surveillance program. Providers most likely to serve the population affected by the infection are selected, for example child health clinics and pediatricians should be selected for surveillance of childhood diseases. A sentinel system allows cost reduction and is combined with active surveillance.

A typical surveillance program for influenza infections includes a selected numbers of general practitioners who are called every week to obtain the number of cases presented to them. This program may include the collection of samples for viral cultures or other diagnostic techniques. Such a level of surveillance would be impossible to maintain on the national level.

Evaluation of a Surveillance System 1.4.5

Surveillance systems are evaluated on the following considerations (CDC 2001b):

- Usefulness: Some surveillance systems are routine programs that collect data and publish results; however it appears that they have no useful purpose – no conclusions are reached, no recommendations are made. A successful surveillance system would provide information used for preventive purposes.
- Sensitivity or the ability to identify every single case of disease is particularly important for outbreak investigations and eradication programs.
- Predictive value positive (PVP) is the proportion of reported cases that actually have the health-related event under surveillance. Low PVP values mean that non-cases might be investigated, outbreaks may be exaggerated or pseudo outbreaks may even be investigated. Misclassification of cases may corrupt the etiologic investigations and lead to erroneous conclusions. Unnecessary interventions and undue concern in the population under surveillance may result.
- Representativeness ensures that the occurrence and distribution of cases accurately represent the real situation in the population.
- Simplicity is essential to gain acceptance, particularly when relying on outside sources for reporting.
- Flexibility is necessary to adapt to changes in epidemiologic patterns, laboratory methodology, operating conditions, funding or reporting sources.
- Data quality is evaluated by the data completeness (blank or unknown variable values) and validity of data recorded (cf. Chap. I.13 of this handbook).

- Acceptability is shown in the participation of providers in the system.
- Timeliness is more important in surveillance of epidemics.
- Stability refers to the reliability (i.e., the ability to collect, manage and provide data properly without failure) and availability (the ability to be operational when it is needed) of the public health surveillance system.

1.4.6 Elements of a Surveillance System

The major elements of a surveillance system as summarized by WHO are: Mortality registration, morbidity reporting, epidemic reporting, laboratory investigations, individual case investigations, epidemic field investigations, surveys, animal reservoir and vector distribution studies, biologics and drug utilization, knowledge of the population and the environment. Traditional surveillance methods rely on counting deaths and cases of diseases. However, these data represent only a small part of the global picture of infectious disease problems.

Mortality Registration

Mortality registration was one of the first elements of surveillance implemented. The earliest quantitative data available on infectious disease is about mortality. The evolution of tuberculosis in the US for example, can only be traced through its mortality. Mortality data are influenced by the occurrence of disease but also by the availability and efficacy of treatment. Thus mortality cannot always be used to evaluate the trend of disease occurrence.

Morbidity Reporting

Reporting of infectious diseases is one of the most common requirements around the world. A list of notifiable diseases is established on a national or regional level. The numbers of conditions vary; it ranges usually from 40 to 60 conditions. In general, a law requires that health facility staff, particularly physicians and laboratories, report these conditions with guaranteed confidentiality. It is also useful to have other non-health related entities report suspected communicable diseases such as day care centers, schools, restaurants, long term care facilities, summer camps and resorts. Regulations on mandatory reporting are often difficult to enforce. Voluntary compliance by the institution's personnel is necessary. Reporting may be done in writing, by phone or electronically in the most advanced system. Since most infectious diseases are confirmed by a laboratory test, reporting by the laboratory may be more reliable. The advantage of laboratory reporting is the ability to computerize the reporting system. Computer programs may be set up to automatically report a defined set of tests and results.

For some infectious diseases, only clinical diagnoses are made. These syndromes may be the consequences of a large number of different microorganisms for which laboratory confirmation is impractical.

When public or physician attention is directed at a specific disease, reporting may be biased. When there is an epidemic or when the press focuses on a particular disease, patients are more prone to look for medical care and physicians are more

likely to report. Reporting rates were evaluated in several studies. In the US, studies show report rates of 10% for viral hepatitis, *Hemophilus influenzae* 32%, meningococcal meningitis 50% and shigellosis 62%.

Morbidity Case Definition

It is important to have a standardized set of definitions available to providers. Without standardized definitions, a surveillance system may be counting different entities from one provider to another. The variability may be such that the epidemiologic information obtained is meaningless.

Most case definitions in infectious disease epidemiology are based on *laboratory tests*, however some clinical syndromes such as toxic shock syndrome do not have confirmatory laboratory tests. Most case definitions include a brief *clinical description* useful to differentiate active disease from colonization or asymptomatic infection. Some diseases are diagnosed based on epidemiologic data. As a result many case definitions for childhood vaccine preventable diseases and foodborne diseases include epidemiologic criteria (e.g., exposure to probable or confirmed cases of disease or to a point source of infection). In some instances, the anatomic site of infection may be important; for example, respiratory diphtheria is notifiable, whereas cutaneous diphtheria is not (CDC 1997).

Cases are classified as a confirmed case, a probable or a suspected case. An epidemiologically linked case is a case in which 1) the patient has had contact with one or more persons who either have/had the disease or have been exposed to a point source of infection (including confirmed cases) and 2) transmission of the agent by the usual modes is plausible. A case may be considered epidemiologically linked to a laboratory-confirmed case if at least one case in the chain of transmission is laboratory confirmed. Probable cases have specified laboratory results that are consistent with the diagnosis yet do not meet the criteria for laboratory confirmation. Suspected cases are usually cases missing some important information in order to be classified as a probable or confirmed case.

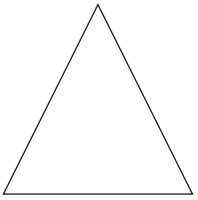
Case definitions are not diagnoses. The usefulness of public health surveillance data depends on its uniformity, simplicity and timeliness. Case definitions establish uniform criteria for disease reporting and should not be used as the sole criteria for establishing clinical diagnoses, determining the standard of care necessary for a particular patient, setting guidelines for quality assurance, or providing standards for reimbursement. Use of additional clinical, epidemiological and laboratory data may enable a physician to diagnose a disease even though the formal surveillance case definition may not be met.

Which Stage of Disease Should Be Collected? The Morbidity Iceberg

Surveillance programs collect data on the overt cases diagnosed by the health care system. However these cases may not be the most important links in the chain of transmission. Cases reported are only the tip of the iceberg. They may not at all be representative of the true endemicity of an infectious disease.

There is a continuous process leading to an infectious disease: exposed, colonized, incubating, sick, clinical form, convalescing, cured. Even among those who have overt disease there are several disease stages that may not be included in a surveillance system:

- some have symptoms but do not seek medical attention
- some do get medical attention but do not get diagnosed or get misdiagnosed
- some get diagnosed but do not get reported



Cases reported
Cases diagnosed but not reported
Cases who seek medical attention but were not diagnosed
Cases who were symptomatic but did not seek medical attention
Cases who were not symptomatic

Infectious disease cases play different roles in the epidemiology of an infectious disease; some individuals are the indicators (most symptomatic), some are the reservoir of microorganisms (usually asymptomatic, not very sick), some are amplifiers (responsible for most of the transmission), some are the victims (those who develop severe long term complications). Depending on the specific disease and the purpose of the surveillance program, different disease stages should be reported. For example

- In a program to prevent rabies in humans exposure to a suspect rabid animal (usually a bite) needs to be reported. At the stage where the case is a suspect, prevention will no longer be effective.
- For bioterrorism events, reporting of suspects is of paramount importance to minimize consequences. Waiting for confirmation causes too long of a delay. In the time necessary to confirm cases, opportunities to prevent co-infections may be lost and secondary cases may already be incubating, depending on the transmissibility of the disease.
- Surveillance for West Nile viral infections best rests on the reporting of neuro-invasive disease. Case reports of neuro invasive diseases are a better indicator than West Nile infection or West Nile fever cases that are often benign, go undiagnosed and are reported haphazardly.
- For Gonorrhea, young males are the indicators because of the intensity of symptoms. Young females are the main reservoir because of the high proportion of asymptomatic infections. Females of reproductive age are the victims because of pelvic invasive disease (PID) and sterility.
- A surveillance program for hepatitis B that only would include symptomatic cases of hepatitis B could be misleading. A country with high transmission of hepatitis B from mother to children would have a large proportion of infected newborn becoming asymptomatic carriers and a major source of infection during their lifetime. Typically in countries with poor reporting of symptomatic hepatitis, the reporting of acute cases of hepatitis B would be extremely low in

spite of high endemicity which would result in high rates of chronic hepatitis and hepatic carcinoma.

Individual Cases or Aggregate Data?

Most morbidity reporting collects data about individual cases. Reporting of individual cases includes demographic and risk factor data which are analyzed for descriptive epidemiology and for implementation of preventive actions. For example, any investigation leading to contact identification and prophylaxis requires a start from individual cases.

However, identification of individuals may be unnecessary and aggregate data sufficient for some specific epidemiologic purposes. Monitoring an influenza epidemic for example, can be done with aggregate data. Obtaining individual case information would be impractical since it would be too time consuming to collect detailed demographics on such a large number of cases. Aggregate data from sentinel sites consists of a number of influenza-like illnesses by age group and the total number of consultants or the total number of 'participants' to be used as denominators. Such data is useful to identify trends and determine the extent of the epidemic and geographic distribution.

Collection of aggregate data of the proportion of school children by age group and sex is a useful predictive tool to identify urinary schistosomiasis endemic areas (Lengeler et al. 2002) without having to collect data on individual school children.

Investigations of Cases, Outbreaks, Epidemics and Surveys

Epidemics of severe diseases are almost always reported. This is not the case for epidemics of milder diseases such as rashes or diarrheal diseases. Many countries do not want to report an outbreak of disease that would cast a negative light on the countries. For example, many countries that are tourism dependent do not report cholera or plague cases. Some countries did not report AIDS cases for a long time.

Case investigations are usually not undertaken for individual cases unless the disease is of major importance such as hemorrhagic fever, polio, rabies, yellow fever, any disease that has been eradicated and any disease that is usually not endemic in the area.

Outbreaks or changes in the distribution pattern of infectious diseases should be investigated and these investigations should be compiled in a comprehensive system to detect trends. While the total number of infectious diseases may remain the same, changes may occur in the distribution of cases from sporadic to focal outbreaks. For example the distribution of WNV cases in Louisiana shifted from mostly focal outbreaks the first year the West Nile Virus arrived in the state in 2002, to mostly sporadic cases the following year in 2003 (see Fig. 1.1).

Surveys are a very commonly used tool in public health, particularly in developing countries where routine surveillance is often inadequate (cf. Chap. IV.6 of this handbook). Survey data needs to be part of a comprehensive surveillance database. One will acquire a better picture from one or a series of well constructed surveys than from poorly collected surveillance data. Surveys are used in control

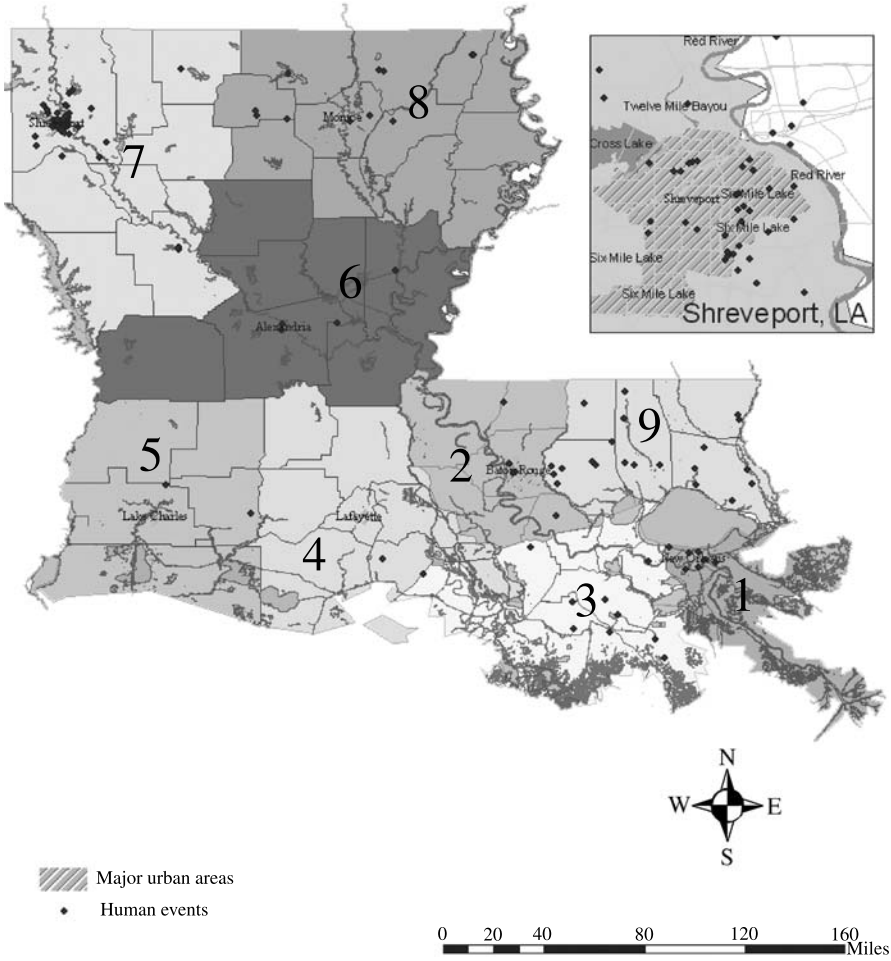


Figure 1.1. Human West Nile Virus Cases Louisiana 2003

programs designed to control major endemic diseases: spleen and parasite surveys for malaria, parasite in urine and stools for schistosomiasis, clinical surveys for leprosy or guinea-worm disease and skin test surveys for tuberculosis.

Surveillance of Microbial Strains

Surveillance of microbial strains is designed to monitor, through active laboratory based surveillance, the bacterial and viral strains isolated. Examples of these systems are:

- In the US, the *PulseNet program* is a network of public health laboratories that performs DNA fingerprinting of bacteria causing foodborne illnesses (Swaminathan et al. 2001). Molecular sub-typing methods must be standardized to allow comparisons of strains and the building of a meaningful data bank. The

method used in PulseNet is pulse field gel electrophoresis (PFGE). The use of standardized subtyping methods has allowed isolates to be compared from different parts of the country, enabling recognition of nationwide outbreaks attributable to a common source of infection, particularly those in which cases are geographically separated.

- The US National Antimicrobial Resistance Monitoring System (NARMS) for enteric bacteria is a collaboration between CDC, participating state and local health departments and the US Food and Drug Administration (FDA) to monitor antimicrobial resistance among foodborne enteric bacteria isolated from humans. NARMS data are also used to provide platforms for additional studies including field investigations and molecular characterization of resistance determinants and to guide efforts to mitigate antimicrobial resistance (CDC 2003).
- Monitoring of antimicrobial resistance is routinely done by requiring laboratories to either submit all, or a sample of their bacterial isolates.

Surveillance of Animal Diseases

Surveillance for zoonotic diseases should start at the animal level, thus providing early warning for impending increases of diseases in the animal population.

- Rabies surveillance aims at identifying the main species of animals infected in an area, the incidence of disease in the wild animals and the prevalence of infection in the asymptomatic reservoir (bats). This information will guide preventive decisions made when human exposures do occur.
- Malaria control entomologic activities must be guided by surveillance of *Anopheles* population, biting activities, *Plasmodium* infection to biting acivities and *Plasmodium* infection rates in the *Anopheles* population.
- Surveillance for dead birds, infection rates in wild birds, infection in sentinel chickens and horse encephalitis are all part of West Nile encephalitis surveillance. These methods provide an early warning system for human infections.
- The worldwide surveillance for influenza is the best example of the usefulness of monitoring animals prior to spread of infection in the human population. Influenza surveillance programs aim to rapidly obtain new circulating strains to make timely recommendations about the composition of the next vaccine. The worldwide surveillance priority is given to the establishment of regular surveillance and investigation of outbreaks of influenza in the most densely populated cities in key locations, particularly in tropical or other regions where urban markets provide opportunities for contacts between humans and live animals (Snacken et al. 1999).

Rationale of Selecting Diseases for Surveillance Purposes

The rationale for selecting infectious diseases and an appropriate surveillance method is based on the goal of the preventive program. Table 1.4 shows a few examples of different surveillance methods based on the disease and the objectives of the surveillance.

Table 1.4. Examples of different surveillance methods based on the disease and the objectives of the surveillance system

Disease	Objectives	Surveillance method
Anthrax	Limit bioterrorism event	Active or passive syndromic surveillance
Antibiotic resistance	Description	Active laboratory reporting of antibiograms
Aseptic meningitis	Sentinel event for West Nile Identification of outbreak	Passive surveillance by health care providers
Gonorrhea	Description of epidemic Treatment of cases	Passive case detection by health care practitioners Systematic screening of young females (Family planning, prenatal, student health services etc.)
Hepatitis B	Description of endemicity	Survey of representative groups
Hepatitis B	Prevention of perinatal transmission	Screening of pregnant women
Influenza	Quantify epidemic	Sentinel surveillance with aggregate data from physicians' offices, emergency departments, nursing homes and schools
Poliomyelitis	Identification of residual cases before complete eradication	Active surveillance of acute flaccid paralysis
Rabies	Prevent human cases	Passive reporting of exposure to potentially rabid animals
Staphylococcus aureus Methicillin resistant (MRSA)	Provide information for management of suspected staphylococcal infections	Active laboratory surveillance of aggregate data on proportion of staphylococci resistant to Methicillin
Staphylococcus aureus Vancomycin resistant (VRSA)	Identification of an emerging infection	Laboratory submission of specimens
Tuberculosis	Description of endemicity Case management	Case register
West Nile	Early warning for public and mosquito control	Passive reporting of dead birds by the public, passive reporting of encephalitic horses, sentinel chicken serology, survey of wildlife by serologic methods
West Nile	Description of epidemic	Passive and active case finding of neuro-invasive disease

1.5 Outbreak Investigations

Outbreaks of acute infectious diseases are common and investigations of these outbreaks are an important task for public health professionals, especially epidemiol-

ogists. In 2001, a total of 1238 foodborne outbreaks with 25,035 cases involved were reported in the US (CDC 2004) with Norovirus being the most common confirmed etiologic agent associated with these outbreaks (see Table 1.5).

Table 1.5. Confirmed etiologic agents of foodborne outbreaks in the US in 2001

Etiology	Number of Outbreaks
Bacillus cereus	5
Brucella spp.	1
Campylobacter spp.	16
Clostridium botulinum	3
Clostridium perfringens	30
Enterohemorrhagic Escherichia coli	4
Enterohemorrhagic Escherichia coli O157:H7	16
Enterotoxigenic Escherichia coli	2
Listeria monocytogenes	1
Salmonella spp.	112
Shigella spp.	15
Staphylococcus aureus	23
Vibrio spp.	4
Yersinia enterocolitica	3
Total Bacteria	235
Ciguatera	23
Histamin	10
Other Chemical	1
Scrombroid	18
Total Chemical	52
Cyclospora cayetanensis	2
Giardia lamblia	1
Trichinella spp	2
Total Parasitic	5
Hepatitis A	6
Norovirus	150
Total Viral	156

Source: CDC Foodborne Outbreak Response and Surveillance Unit, 2004

Outbreaks or epidemics are defined as the number of disease cases above what is normally expected in the area for a given time period. Depending on the disease, it is not always known if the case numbers are really higher than expected and some outbreak investigations can reveal that the reported case numbers did not actually increase.

The nature of a disease outbreak depends on a variety of circumstances, most importantly the suspected etiologic agent involved, the disease severity or case fatality rate, population groups affected, media pressure, political inference and investigative progress. There are certain common steps for outbreak investiga-

tions as shown in Table 1.6. However, the chronology and priorities assigned to each phase of the investigation have to be decided individually, based on the circumstances of the suspected outbreak and information available at the time.

Table 1.6. Common steps in outbreak investigations

1.	"Outbreak" detected based on initial report or analysis of surveillance data
2.	Collect basic numbers and biologic specimens
3.	Investigate or not?
4.	Think prevention first
5.	Get information on the disease or condition
6.	Sometimes numbers do not count
7.	Is the increase real or artificial?
8.	Verify the diagnosis
9.	Prepare a case definition
10.	Put the information in a database
11.	Find additional cases
12.	Basic descriptive epidemiology (time, place and person)
13.	Hypothesis testing and measures of association
14.	Final report and communications

For example, in 2002, 21 outbreaks of acute gastroenteritis on cruise ships with travel destinations outside the US were reported to the CDC (CDC 2002). In only five of these outbreaks about 1400 persons, with an average 280 cases per cruise, had symptoms of viral acute gastroenteritis. Norovirus outbreaks begin usually as a food or water borne disease but often continue because of the easy person to person transmission in a closed environment and low infectious dose (100 viral particles can be infectious) (CDC 2001a).

1.5.1 Basic Steps in Outbreak Investigations

1. The initial report can originate from very different sources. Examples are:
   - A physician is calling the local or state health department about an increase of number of patients seen and diagnosed with a specific disease,
   - A high number of patients with similar signs and symptoms are showing up in the emergency room,
   - A school principal or daycare owner is reporting a high number of absent students,
   - A nursing home health care professional is seeing a lot of residents with gastrointestinal illnesses,
   - A person is complaining to the health department that she/he got sick after eating at a certain restaurant.

   Another way to detect an increase of cases is if the surveillance system of reportable infectious diseases reveals an unusually high number of people

with the same diagnosis over a certain time period at different health care facilities.

Outbreaks of benign diseases like self-limited diarrhea are often not detected because people are not seeking medical attention and therefore medical services are not aware of them. Furthermore, early stages of a disease outbreak are often undetected because single cases are diagnosed sporadically. It is not until a certain threshold is passed, that it becomes clear that these cases are related to each other through a common exposure or secondary transmission. Depending on the infectious disease agent, there can be a sharp or a gradual increase of number of cases. It is sometimes difficult to differentiate between sporadic cases and the early phase of an outbreak. In the 2001 St. Louis Encephalitis (SLE) outbreak in Louisiana, the number of SLE cases increased from 9 to 18 between week one and two and then the numbers gradually decreased over the next 9 weeks to a total of 63 cases (Jones et al. 2002).

2. After the initial report is received, it is important to collect and document basic information: Contact information of persons affected, a good and thorough event description, names and diagnosis of hospitalized persons (and depending on the presumptive diagnosis their underlying conditions and travel history), laboratory test results and other useful information to get a complete picture and to confirm the initial story of the suspected outbreak. It also might be necessary to collect more biological specimens such as food items and stool samples for further laboratory testing.
3. Based on the collected information the decision to investigate must be made. It may not be worthwhile to start an investigation if there are only a few people who fully recovered after a couple of episodes of a self-limited, benign diarrhea. Other reasons not to investigate might be that this type of outbreak occurs regularly every summer or that it is only an increase in number of reported cases which are not related to each other.

   On the other hand, however, there should be no time delay in starting an investigation if there is an opportunity to prevent more cases or the potential to identify a system failure which can be caused, for example, by poor food preparation in a restaurant or poor infection control practices in a hospital or to prevent future outbreaks by acquiring more knowledge of the epidemiology of the agent involved. Additional reasons to investigate include the interest of the media, politicians and the public in the disease cluster and the pressure to provide media updates on a regularly basis. Another fact to consider is that outbreak investigations are good training opportunities for newly hired epidemiologists.

   Sometimes lack of data and lack of sufficient background information make it difficult to decide early on if there is an outbreak or not. The best approach then is to assume that it is an outbreak until proven otherwise.
4. Prevention of more cases is the most important goal in outbreak investigations and therefore a rapid evaluation of the situation is necessary. If there are precautionary measures to be recommended to minimize the impact of the outbreak and the spread to more persons, they should be implemented

before a thorough investigation is completed. Most likely control measures implemented by public health professionals in foodborne outbreaks are:
- Recall or destruction of contaminated food items,
- Restriction of infected food handlers from food preparation,
- Correction of any deficiency in food preparation or conservation.

5. After taking immediate control measures, the next step is to know more about the epidemiology of the suspected agent. The most popular books for public health professionals include the "Red Book" (American Academy of Pediatrics 2003), the "Control of Communicable Diseases Manual" from the American Public Health Association (APHA 2000) or other infectious disease epidemiology books as well as the CDC website (www.cdc.gov). If the disease of interest is a reportable disease or a disease where surveillance data are available, baseline incidence rates can be calculated. Then a comparison is made to determine if the reported numbers constitute a real increase or not. Furthermore, the seasonal and geographical distribution of the disease is important as well as the knowledge of risk factors. Many infectious diseases show a seasonal pattern such as Rotavirus or *Neisseria meningitides*. For example in suspected outbreaks where cases are associated with raw oyster consumption, the investigator should know that in the US Gulf states *Vibrio* cases increase in the summer months because the water conditions are optimal for the growth of the bacteria in water and in seafood. This kind of information will help to determine if the case numbers show a true increase and if it seems likely to be a real outbreak.
6. For certain diseases, numbers are not important. Depending on the severity of the disease, its transmissibility and its natural occurrence, certain diseases should raise a red flag for every health care professional and even a single case should warrant a thorough public health investigation. For example a single confirmed case of a rabid dog in a city (potential dog to dog transmission within a highly populated area), a case of dengue hemorrhagic fever or a presumptive case of smallpox would immediately trigger an outbreak investigation.
7. Sometimes an increase of case numbers is artificial and not due to a real outbreak. In order to differentiate between an artificial and a natural increase in numbers, the following changes have to be taken into consideration:
   - Alterations in the surveillance system,
   - A new physician who is interested in the disease and therefore more likely to diagnose or report the disease,
   - A new health officer strengthening the importance of reporting,
   - New procedures in reporting (from paper to web based reporting),
   - Enhanced awareness or publicity of a certain disease that might lead to increased laboratory testing,
   - New diagnostic tests,
   - A new laboratory,
   - An increase in susceptible population such as a new summer camp.

8. It is important to be sure that reported cases of a disease actually have the correct diagnosis and are not misdiagnosed. Is there assurance that all the cases have the same diagnosis? Is the diagnosis verified and were other differential diagnoses excluded? In order to be correct, epidemiologists have to know the basis for the diagnosis. Are laboratory samples sufficient? If not, what kind of specimens should be collected to ascertain the diagnosis? What are the clinical signs and symptoms of the patient?
   In an outbreak of restaurant associated botulism in Canada only the 26th case was correctly diagnosed. The slow progression of symptoms and misdiagnosis of the dispersed cases made it very difficult to link these cases and identify the source of the outbreak (CDC 1985, 1987).
9. The purpose of a case definition is to standardize the identification and counting of the number of cases. The case definition is a standard set of criteria and is not a clinical diagnosis. In most outbreaks the case definition has components of person, place and time, such as the following: Persons with symptoms of X and Y after eating at the restaurant Z between Date1 and Date2. The case definition should be broad enough to get most of the true cases but not too narrow so that true cases will not be misclassified as controls. A good method is to analyze the data, identify the frequency of symptoms and include symptoms that are more reliable than others. For example, diarrhea and vomiting are more specific than nausea and headache in the case definition of a food related illness.
10. What kind of information is necessary to be collected? It is sufficient to have a simple database with basic demographic information such as name, age, sex and information for contacting the patient. More often, date of reporting and date of onset of symptoms are also important. Depending on the outbreak and the potential exposure or transmission of the agent involved further variables such as school, grade of student or occupation in adults might be interesting and valuable.
11. During an outbreak investigation it is important to identify additional cases that may not have been known or were not reported. There are several approaches:
    - Interview known cases and ask them if they know of any other friends or family members with the same signs or symptoms,
    - Obtain a mailing list of frequent customers in an event where a restaurant is involved,
    - Set up an active surveillance with physicians or emergency departments,
    - Call laboratories and ask for reports of suspected and confirmed cases.

    Another possibility is to review surveillance databases or to establish enhanced surveillance for prospective cases. Occasionally it might be worthwhile to include the media for finding additional cases through press releases. However the utility of that technique depends on the outbreak and the etiologic agent; the investigator should always do a benefit risk analysis before involving the media.

12. After finding additional cases, entering them in the database and organizing them, the investigator should try to get a better understanding of the situation by performing some basic descriptive epidemiology techniques such as sorting the data by time, place and person. For a better visualization of the data, an epidemic or "epi" curve should be graphed. The curve shows the number of cases by date or time of onset of symptoms. This helps to understand the nature and dynamic of the outbreak as well as to get a better understanding of the incubation period if the time of exposure is known. It also helps to determine whether the outbreak had a single exposure and no secondary transmission (single peak) or if there is a continuous source and ongoing transmission. Figures 1.2 and 1.3 show "epi" curves of two different outbreaks: a foodborne outbreak in a school in Louisiana, and the number of WNV human cases in Louisiana in the 2002 outbreak, respectively.
Sometimes it is useful to plot the cases on a map to get a better idea of the nature and the source of an outbreak. Mapping may be useful to track the spread by water (see John Snow's cholera map) or by air or even a person to person transmission. If a contaminated food item was the culprit, food distribution routes with new cases identified may be helpful. Maps, however, should be taken with caution and carefully interpreted. For example, WNV cases are normally mapped by residency but do not take into account that people might have been exposed or bitten by an infective mosquito far away from where they live. For outbreak investigations, spot maps are usually more useful than rate maps or maps of aggregate data.
Depending on the outbreak it might be useful to characterize the outbreak by persons' demographics such as age, sex, address and occupation or health status. Are the cases at increased susceptibility or at high risk of infection?

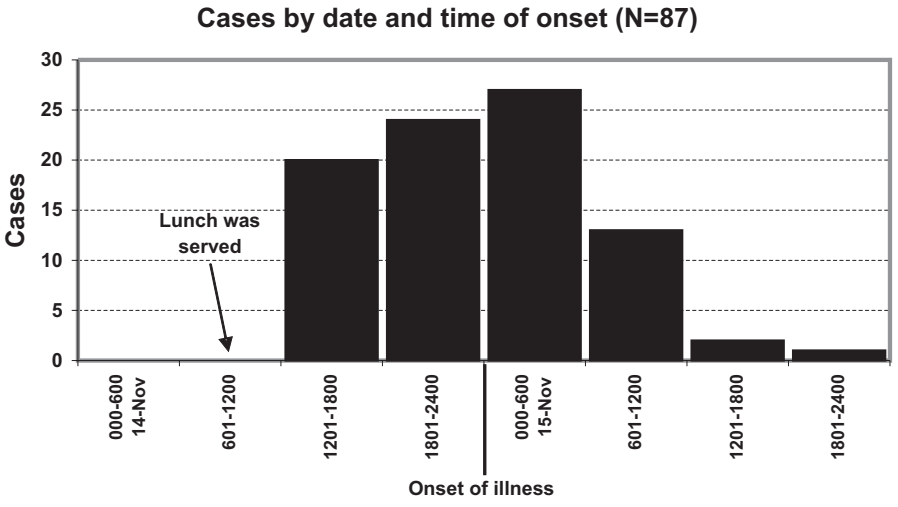


Figure 1.2. Gastroenteritis outbreak in a school in Louisiana, 2001

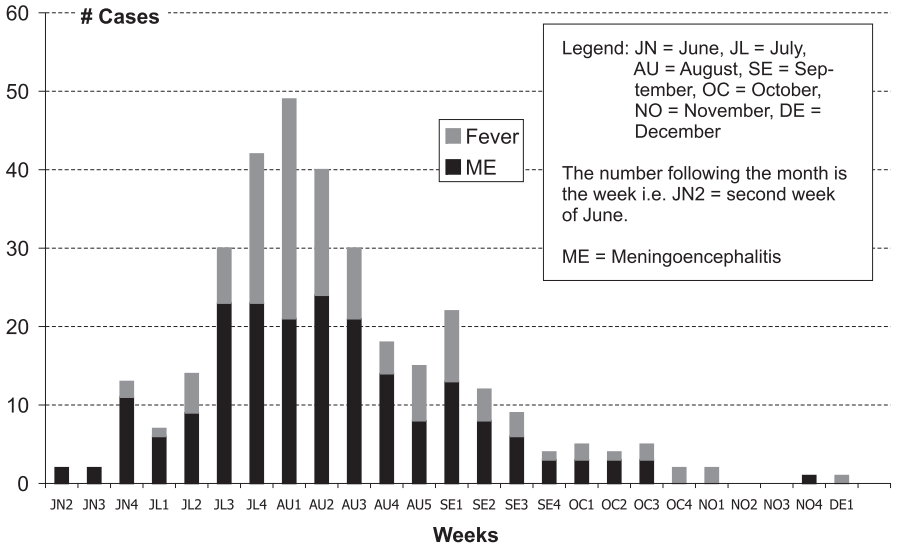


Figure 1.3. Human West Nile Virus cases, Louisiana 2002

These kinds of variables might give the investigator a good idea if the exposure is not yet known. For typical foodborne outbreaks however, demographic information is not very useful because the attack rates will be independent of age and sex. More details on methods used in descriptive epidemiology are given in Chap. I.3 of this handbook.

13. Based on the results of basic descriptive epidemiology and the preliminary investigation, some hypotheses should be formulated in order to identify the cause of the outbreak. A hypothesis will be most likely formulated such as "those who attended the luncheon and ate the chicken salad are at greater risk than those who attended and did not eat the chicken salad". It is always easier to find something after knowing what to look for and therefore a hypothesis should be used as a tool. However, the epidemiologist should be flexible enough to change the hypothesis if the data do not support it. If data clues are leading in another direction, the hypothesis should be reformulated such as "those who attended the luncheon and ate the baked chicken are at greater risk than those who attended and did not eat the baked chicken".

    To verify or deny hypotheses, measures of risk association such as the relative risk (RR) or the odds ratio (OR) have to be calculated (as described in Chaps. I.2, I.5, and I.6 of this handbook). The CDC has developed the software program 'EpiInfo' which is easy to use in outbreak investigations, and, even more importantly, free of charge. It can be downloaded from the CDC website (http://www.cdc.gov/epiinfo/). Measures of association, however, should be carefully interpreted; even a highly significant measure of association can not give enough evidence of the real culprit or the contaminated food item. The measure of association is only as good and valid as the data. Most people

have recall problems when asked what they ate, when they ate and when their symptoms started. Even more biases or misclassifications of cases and controls can hide an association. A more confident answer comes usually from the laboratory samples from both human samples and food items served at time of exposure. Agents isolated from both food and human samples that are identified as the same subtype, in addition to data results supporting the laboratory findings, are the best evidence beyond reasonable doubt.

14. As the last step in an outbreak investigation, the epidemiologist writes a final report on the outbreak and communicates the results and recommendations to the public health agency and facilities involved. In the US, public health departments also report foodborne outbreaks electronically to CDC via a secure web based reporting system, the Electronic Foodborne Outbreak Reporting System (EFORS).

1.5.2 Types of Outbreaks

The "Traditional" Foodborne Outbreak

The "traditional" foodborne outbreak is usually a small local event such as family picnic, wedding reception, or other social event and occurs often in a local restaurant or school cafeteria. This type of outbreak is highly local with a high attack rate in the group exposed to the source. Because it is immediately apparent to those in the local group such as the group of friends who ate at the restaurant or the students' parents, public health authorities are normally notified early in the outbreak while most of the cases are still symptomatic. Epidemiologists can start early on with their investigation and therefore have a much better chance to collect food eaten and stool samples of cases with gastroenteritis for testing and also to detect the etiologic agent in both of them.

In a 2001 school outbreak in Louisiana, eighty-seven persons (sixty-seven students and twenty faculty members) experienced abdominal cramps after eating at the school's annual "Turkey Day" the day before. Stool specimens and the turkey with the gravy were both positive for *Clostridium perfringens* with the same pulse field gel electrophoresis (PFGE) pattern (Merlos 2002). The inspection of the school cafeteria revealed several food handling violations such as storing, cooling and reheating of the food items served. Other than illnesses among food handlers, these types of improper food handling or storage are the most common causes of foodborne outbreaks.

New Types of Outbreaks

A different type of outbreak is emerging as the world is getting smaller. In other words persons and food can travel more easily and faster from continent to continent and so do infectious diseases with them. Foodborne outbreaks related to imported contaminated food items are normally widespread, involving many states and countries and therefore are frequently identified. In 1996, a large outbreak of *Cyclospora cayetanensis* occurred in 10 US states and Ontario, Canada and was

linked to contaminated raspberries imported from South America. Several hundred laboratory confirmed cases were reported, most of them in immunocompetent persons (CDC 1996).

A very useful molecular tool to identify same isolates from different geographic areas is sub-typing enteric bacteria with PFGE. In the US, the PulseNet database allows state health department to compare their isolates with other states and therefore increase the recognition of nationwide outbreaks linked to the same food item (Swaminathan et al. 2001).

In a different scenario, a widely distributed food item with low-level contamination might result in an increase of cases within a large geographic area and therefore might be not get detected on a local level. This kind of outbreak might only be detected by chance if the number of cases increased in one location and the local health department alerts other states to be on the lookout for a certain isolate.

Another type of outbreak is the introduction of a new pathogen into a new geographic area as it happened in 1991 when *Vibrio cholerae* was inadvertently introduced in the waters off the Gulf Coast of the United States. In the U.S., however, most cases are usually traced back to people who traveled to areas with a high cholera risk or to people who ate food imported from cholera-risk countries and only sporadic *Vibrio cholerae* cases are associated with the consumption of raw or undercooked shellfish from the Gulf of Mexico (CDC 1999b).

Food can not only be contaminated by the end of the food handling process i.e. by infected food handlers but also can be contaminated by any event earlier in the chain of food production. In 1996, an ice cream outbreak of *Salmonella enteritidis* in a national brand of ice cream resulted in 250,000 illnesses. The outbreak was detected by routine surveillance because of a dramatic increase of *Salmonella enteritidis* in South Minnesota. The cause of the outbreak was a basic failure on an industrial scale to separate raw products from cooked products. The ice cream premix was pasteurized and then transported to the ice cream factory in tanker trucks which had been used to haul raw eggs. This resulted in the contamination of the ice cream and subsequent salmonella cases (Hennessey et al. 1996).

Surveys 1.6

Surveys are useful to provide information for which there is no data source or no reliable data source. Surveys are time consuming and are often seen as a last choice to obtain information. However, too often unreliable information is used because it is easily available. For example, any assessment of the *Legionella* problem using passive case detection will be unreliable due to under-diagnosis and under-reporting. Most cases of legionellosis are treated empirically as community acquired pneumonias and are never formally diagnosed.

In developing countries, surveys are often necessary to evaluate health problems since data collected routinely (disease surveillance, hospital records, case registers)

are often incomplete and of poor quality. In industrialized nations, although many sources of data are available, there are some circumstances where surveys may be necessary.

Prior to carrying out surveys involving human subjects, special procedures need to be followed. In industrialized countries, a human subject investigation review board has to evaluate the project's value and ethics. In developing countries, however, such boards may not be formalized but it is important to obtain permission from medical, national and local political authorities before proceeding.

1.6.1 Survey Methods

Surveys of human subjects are carried out by mail, telephone, personal interviews, and behavioral observations. In infectious diseases, the collection of biological specimens in humans (i.e. blood for serologic surveys) or the collection of environmental samples (food, water, environmental surfaces) is very common. Personal interviews and specimen collection require face to face interaction with the individual surveyed. These are carried out in offices or by house to house surveys.

Non-respondents are an important problem for infectious disease surveys. Those with an infection may be absent from school, may not answer the door or may be unwilling to donate blood for a serologic survey, thus introducing a systematic bias into the survey results.

Since surveys are expensive, they cannot be easily repeated. All field procedures, questionnaires, biological sample collection methods and laboratory tests should be tested prior to launching the survey itself. Feasibility, acceptability and reliability can be tested in a small scale pilot study. More details on survey methods are to be found in Chap. I.10 of this handbook.

1.6.2 Sampling

Since surveys are labor intensive, they are rarely carried out on an entire population but rather on a sample. To do a correct sampling, it is necessary to have a sampling base (data elements for the entire population) from which to draw the sample. Examples of sampling bases are population census, telephone directory (for the phone subscriber population), school roster or a school list. In developing countries such lists are not often available and may have to be prepared before sampling can start. More information on sampling designs can be found in Chap. IV.5 of this handbook.

1.6.3 Community Surveys (House to House Surveys)

Most community surveys are carried out in developing countries because reliable data sources are rare. The sampling base often ends up to the physical layout of the population. A trip and geographical reconnaissance of the area are necessary.

The most common types of surveys undertaken in developing countries are done at the village level; they are based on maps and a census of the village.

In small communities, it is important to obtain the participation of the population. Villagers are often wary of government officials counting people and going from door to door. To avoid misinterpretations and rumors, influential people in the community should be told about the survey. Their agreement is indispensable and their help is needed to explain the objectives of the survey and particularly its potential benefits. Increasing the knowledge about disease, disease prevention and advancing science are abstract notions that are usually poorly understood or valued by villagers who are, in general, very practical people. If a more immediate benefit can be built into the survey, there will be an increase in cooperation of the population. Incentives such as offering to diagnose and treat an infection or drugs for the treatment of common ailments such as headaches or malaria enhance the acceptance of the survey.

In practically all societies the household is a primary economic and social unit. It can be defined as the smallest social unit of people who have the same residency and maintain a collective organization. The usual method for collecting data is to visit each household and collect samples or administer a questionnaire.

Medical staff may feel left out or even threatened whenever a medical intervention (such as a survey) is done in their area. A common concern is that people will go to their medical care provider and ask questions about the survey or about specimen collection and results. It is therefore important to involve and inform local medical providers as much as practical.

A rare example of a house to house survey in an industrialized nation was carried out in Slidell, Louisiana for the primary purpose of determining the prevalence of West Nile infection in a southern US focus. Since the goal was to obtain a random sample of serum from humans living in the focus, the only method was a survey of this type. A cluster sampling design was used to obtain a representative number of households. The area was not stratified because of its homogeneity. Census blocks were grouped so that each cluster contained a minimum of 50 households. The probability of including an individual cluster was determined by the proportion of houses selected in that cluster and the number of persons participating given the number of adults in the household. A quota sampling technique was used, with a goal of enlisting 10 participating households in each cluster.

Inclusion criteria included age (at least 12 years of age) and length of residence (at least 2 years). The household would be included only if an adult household resident was present. A standardized questionnaire was used to interview each participant. Information was collected on demographics, any recent febrile illness, knowledge, attitudes, and behaviors to prevent WNV infection and potential exposures to mosquitoes. A serum sample for WNV antibody testing was drawn. In addition, a second questionnaire regarding selected household characteristics and peridomestic mosquito reduction measures was completed. Informed consent was

obtained from each participant, and all participants were advised that they could receive notification of their blood test results if they wished. Institutional Review Board approvals were obtained.

Logistics for specimen collection, preservation and transportation to the laboratory were arranged. Interpretation of serologic tests and necessary follow up were determined prior to the survey and incorporated in the methods submitted to the ethics committee.

Sampling weights, consisting of components for block selection, household-within-block selection, and individual-within-household participation, were used to estimate population parameters and 95% confidence intervals (CI). Statistical tests were performed incorporating these weights and the stratified cluster sampling design.

In this survey, 578 households were surveyed (a 54% response rate), including 1226 participants. There were 23 IgM seropositive persons, for a weighted seroprevalence of 1.8% (with a 95% confidence interval of 0.9%–2.7%) (Vicari et al. 2003).

1.7 Program Evaluation

Program evaluation is a systematic way to determine if prevention or intervention programs for the infectious disease of interest are effective and to see how they can be improved. It is beyond the scope of this chapter to explain program evaluation in detail however there is abundant information available i.e. the CDC's Framework for Program Evaluation in Public Health (CDC 1999a) as well as text books on program evaluation (Fink 1993).

Most importantly, evaluators have to understand the program such as the epidemiology of the disease of interest, the program's target population and their risk factors, program activities and resources. They have to identify the main objectives of the control actions and determine the most important steps. Indicators define the program attributes and translate general concepts into measurable variables. Data are then collected and analyzed so that conclusions and recommendations for the program are evidence based.

Evaluating an infectious disease control program requires a clear understanding of the microorganism, its mode of transmission, the susceptible population and the risk factors. The following example of evaluation of tuberculosis control shows the need to clearly understand the priorities.

Most of tuberculosis transmission comes from active pulmonary tuberculosis cases who have positive sputum smear (confirmed as tuberculosis *Mycobacteria* on culture). To a lesser extent, smear negative culture positive pulmonary cases are also transmitting the infection. Therefore priority must be given to find sputum positive pulmonary cases. The incidence of smear positive tuberculosis cases is the most important incidence indicator. Incidences of active pulmonary cases and of all active cases (pulmonary and extra-pulmonary) are also calculated but are

of lesser interest. The following proportions are used to detect anomalies in case finding or case ascertainment:

- all tuberculosis cases who are pulmonary versus extra-pulmonary,
- smear positive, culture positive, pulmonary cases versus smear negative, culture positive, pulmonary cases,
- culture positive, pulmonary cases versus culture negative, pulmonary cases.

Poor laboratory techniques or low interest in obtaining sputa for smears or cultures may result in underestimating bacteriological confirmed cases. Excessive diagnosis of tuberculosis with reliance on chest X-rays on the other hand may overestimate unconfirmed tuberculosis cases.

Once identified, tuberculosis cases are placed under treatment. Treatment of infectious cases is an important preventive measure. Treatment efficacy is evaluated by sputum conversion (both on smear and culture) of the active pulmonary cases. After 2 months of an effective regimen, 85% of active pulmonary cases should have converted their sputum from positive to negative. Therefore the rate of sputum conversion at 2 months becomes an important indicator of program effectiveness. This indicator must be calculated for those who are smear positive and with a lesser importance for the other active pulmonary cases.

To ensure adequate treatment and prevent the development of acquired resistance, tuberculosis cases are placed under directly observed therapy (DOT). This measure is quite labor intensive. Priority must therefore be given to those at highest risk of relapse. These are the smear positive culture proven active pulmonary cases. DOT on extra-pulmonary cases is much less important from a public health standpoint.

The same considerations apply to contact investigation and preventive treatment in countries that can afford a tuberculosis contact program. A recently infected contact is at the highest risk of developing tuberculosis the first year after infection; hence the best preventive return is to identify contacts of infectious cases. Those contacts are likely to have been recently infected. Systematic screening of large population groups would also identify infected individuals but most would be 'old' infections at lower risk of developing disease. Individuals infected with tuberculosis and HIV are at extremely high risk of developing active tuberculosis. Therefore the tuberculosis control program should focus on the population at high risk of HIV infection.

Often, program evaluation is performed by epidemiologists who have not taken the time to understand the dynamics of a disease in the community. Rates or proportions are calculated, no priorities are established and precious resources are wasted on activities with little preventive value. For example, attempting to treat all tuberculosis cases, whether pulmonary or not with DOT, investigating all contacts regardless of the bacteriologic status of the index case, would be wasteful.

1.8 Conclusions

Today the world is smaller than ever before, international travel and a worldwide food market make us all potentially vulnerable to infectious diseases no matter where we live.

New pathogens are emerging such as the SARS or spreading through new territories such as WNV. WNV introduced in the US in 1999, became endemic in the US over the next years. Hospital-associated and community-associated Methicillin Resistant Staphylococcus Aureus (MRSA) and resistant tuberculosis cases and outbreaks are on the rise. Public health professionals are concerned that a novel recombinant strain of influenza will cause a new pandemic.

But not only the world and the etiologic agents are changing, the world population is changing as well. In industrialized countries, the life expectancy is increasing and the elderly are more likely to acquire a chronic disease, cancer or diabetes in their lifetime. Because of underlying conditions or the treatment of these diseases, older populations also have an increased susceptibility for infectious diseases and are more likely to develop life-threatening complications.

Knowledge in the field of infectious disease epidemiology is expanding. While basic epidemiological methods and principles still apply today, improved laboratory diagnoses and techniques help to confirm cases faster, see how cases are related to each other and therefore can support the prevention of spread of the specific disease. Better computers can improve the data analysis and internet allows access to in depth disease specific information. Computer connectivity improves disease reporting for surveillance purposes and the epidemiologist can implement faster preventive measures if necessary and is also able to identify disease clusters and outbreaks on a timelier basis.

The global threat of bioterrorism adds a new dimension. The intentional release of anthrax spores, and the infection and death of persons who contracted the disease created a scare of contaminated letters in the US population.

With all these changes, there is renewed emphasis on infectious disease epidemiology and makes it a challenging field to work in.

References

Aintablian N, Walpita P, Sawyer MH (1998) Detection of Bordetella pertussis and respiratory syncytial virus in air samples from hospital rooms. Infect Control Hosp Epidemiol 19:918-923

Alland D, Kalkut GE, Moss AR, McAdam RA, Hahn JA, Bosworth W, Drucker E, Bloom BR (1994) Transmission of tuberculosis in New York City. An analysis by DNA fingerprinting and conventional epidemiologic methods. NEJM 330(24):1710–1716

American Academy of Pediatrics (AAP) (2003) In: Pickering LK (ed) Red Book: 2003. Report of the Committee on Infectious Diseases, 26th edn. AAP, Elk Grove Village, IL

American Public Health Association (2000) In: Chin J (ed) Control of communicable diseases manual, 17th edn. United Book Press Inc, Baltimore, MD

Boelaert M, Arbyn M, Van der Stuyft P (1998) Geographical Information System (GIS), gimmick or tool for health district management? Trop Med Int Health 3:163–165

CDC (1985) Update: International outbreak of restaurant-associated botulism – Vancouver, British Columbia, Canada. MMWR 34(41):643

CDC (1987) Epidemiologic notes and reports restaurant associated botulism from mushrooms bottled in-house – Vancouver, British Columbia, Canada. MMWR 36(7):103

CDC (1996) Outbreaks of Cyclospora cayetanensis infection – United States, 1996. MMWR 45(25):549–551

CDC (1997) Case definitions for infectious conditions under public health surveillance. MMWR 46(10):1–55

CDC (1999a) Framework for program evaluation in Public Health. MMWR 48(11):1–40

CDC (1999b) Summary of infections reported to *Vibrio* surveillance system (http://www.cdc.gov/ncidod/dbmd/diseaseinfo/files/VibCSTE99web.pdf) Accessed May 25, 2004

CDC (2001a) "Norowalk-like viruses": Public health consequences and outbreak management. MMWR 50 (9):1–17

CDC (2001b) Updated guidelines for evaluating public health surveillance systems: Recommendations from the Guidelines Working Group. MMWR 50(13): 1–35

CDC (2002) Outbreaks of gastroenteritis associated with noroviruses on cruise ships – United States, 2002. MMWR 51(49):1112–1115

CDC (2003) National antimicrobial resistance monitoring system for enteric bacteria (NARMS): 2001 Annual Report. U.S. Department of Health and Human Services, Atlanta, Georgia

CDC (2004) Diagnosis and management of foodborne illnesses: A primer for Physicians and other health care professionals. MMWR 53(4):1–33

EpiInfo (http://www.cdc.gov/epiinfo/) Accessed May 25, 2004

Fink A (1993) Evaluation fundamentals: Guiding health programs, research and policy. Sage, Newbury Park, CA

Goodwin LG, Gordon Smith CE (1996) Yellow fever. In: Cox CR (ed) The Wellcome Trust illustrated history of tropical diseases. Wellcome Trust, London, p 147

Hennessy TW, Hedberg CW, Slutsker L, White KE, Besser-Wiek JM, Moen ME, Feldman J, Coleman WW, Edmonson LM, MacDonald KL, Osterholm MT (1996) A national outbreak of salmonella enteritidis Infections from ice cream. NEJM 334(20): 1281–1286

Jones SC, Morris J, Hill G, Alderman M, Ratard RC (2002) St. Louis encephalitis outbreak in Louisiana in 2001. J La State Med Soc 154:303–306

Lengeler C, Makwala J, Ngimbi D, Utzinger J (2000) Simple school questionnaires can map both Schistosoma mansoni and Schistosoma haematobium in the Democratic Republic of Congo. Acta Trop 74(1):77–87

Lockman S, Sheppard JD, Braden CR (2001) Molecular and conventional epidemiology of Mycobacterium tuberculosis in Botswana: A population-based prospective study of 301 pulmonary tuberculosis patients. J Clin Microbiol 39 (3):1042–1047

Mandell GL, Bennett JE, Dolin R (eds) (2000) Mandell, Douglas, and Bennett's principles and practice of infectious diseases, 5th edn. Churchill Livingstone, Philadelphia

McMahon B, Pugh TF (1970) Epidemiology, principles and methods. Little, Brown and Company, Boston, p 149

Merlos II (2002) An uninvited guest at "Turkey Day" Louisiana morbidity report 13(1): 1–2 (http://www.oph.dhh.state.la.us/infectiousdisease/docs/Lmr/janfeb02.pdf) Accessed May 25, 2004

Mounts AW, Holman RC, Clarke MJ, Bresee JS, Glass RI (1999) Trends in hospitalizations associated with gastroenteritis among adults in the United States, 1979–1995. Epidemiol Infect 123:1–8

Snacken R, Kendal AP, Haaheim LR, Wood JM (1999) The next influenza pandemic: Lessons from Hong Kong, 1997. Emer Inf Dis 5 (2):195–203

Snow J (1936) On the mode of communication of cholera. In: Snow on cholera. A reprint of two papers by John Snow. The Commonwealth Fund, New York, pp 1–175

Swaminathan B, Barrett TJ, Hunter SB, Tauxe RV, and the CDC PulseNet Task Force (2001) PulseNet: The Molecular Subtyping Network for Foodborne Bacterial Disease Surveillance, United States. Emer Inf Dis 7 (3):382–389

Vicari SA, Zielinski-Gutierrez E, Montgomery S, Chow C, O'Leary D, Grayson K, Biggerstaff B, Martin D, Ratard R, Campbell C, Bunning M (2003) Household-based seroepidemiologic survey of West Nile Virus infection-Slidell, Louisiana, 2002. Late-breaker report presented at the CDC "Annual Epidemic Intelligence Service Conference" held in Atlanta, GA, on April 4, 2003

WHO (2000a) World health report 2000 (http://www.who.int/whr2001/2001/archives/index.htm) Accessed May 25, 2004

WHO (2000b) World health report 2000. Statistical annex. (http://www.who.int/whr2001/2001/archives/2000/en/pdf/StatisticalAnnex.pdf) Accessed May 25, 2004

Wills C (1996a) Cholera, the Black one. In: Yellow fever black goddess, the coevolution of people and plagues. Addison-Wesley Publishers, Reading, MA, p 115

Wills C (1996b) Four tales from the new decameron. In: Yellow fever black goddess, the coevolution of people and plagues. Addison-Wesley Publishers, Reading, MA, p 84