AFRICA POPULATION INSTITUTE (API)

PUBLIC HEALTH SHORT COURSE HANDBOOK

Website: www.africapopulation.net
Email: info@africapopulation.net

TABLE OF CONTENTS

TABLE OF CONTENTS---2
Forward:--3
How we work:---3
We aim to help you: ---3
Our approach --3
Courses offered in our training workshops that are client tailored ---3
Objectives --5
History of public health --5
Public health programs---7
Economics of public health---8
Research--9
What is Environmental Health?--- 10
What is Tracking?-- 10
Biomonitoring -- 10
Environmental Burden of Disease-- 10
Environmental Epidemiology-- 10
Environmental Exposure--- 10
Environmental Health -- 11
Environmental Justice-- 11
Environmental Standard --- 11
Epidemiology-- 11
Exposure Assessment -- 11
Global Environmental Change-- 11
Health Impact Assessment (HIA) -- 12
Interdisciplinary Research --- 12
National Environmental Health Action Plans (NEHAPs)--- 12
NEHAPs are government documents that address environmental health problems in a comprehensive, holistic and intersectoral manner. Most NEHAPs address environmental health management (including information systems, environmental health services, public information and participation, education, research and economic tools), environmental health hazards (including ambient air, drinking and bathing water, noise, radiation, food safety, housing and workplace), and economic sectors (including industry, agriculture, transportation and energy).-- 12
Public Health -- 12
Risk Assessment (in Environmental Health)-- 12
Risk Communication -- 12
Objectives of health psychology-- 53
Applications of Health Psychology -- 54
Human resources for health -- 56
Planning for human resources--- 57
Human resources contribute to keep a healthy and productive population------------------------------ 57
References --- 72

Forward:
API is a registered institution with 4 years' experience of supporting voluntary organizations, agencies and individuals in developing quality systems. A major part of our work is providing external evaluations and trainings to organizations or specific projects and also building capacities of the members to have relevant skills applicable to their working environment.

How we work:
We aim to understand the precise needs of your organization and to offer you good value, an integrated service, and work which is based on clear principles. Our style is inclusive, participatory and flexible.

We aim to help you:
- Develop your skills, reflect and gain focus
- Make your organization more confident, effective and efficient, and able to demonstrate this to others
- Help you plan more effectively and strategically for the future
- Demonstrate the benefits (or outcomes) for your service users.

Our approach
- Starts by listening carefully to what you need and tailoring our services accordingly
- Includes clear and practical advice, plans and reports
- Is based on extensive knowledge and experience of the voluntary sector
- Is supportive and friendly.

Courses offered in our training workshops that are client tailored
Monitoring and Evaluation Training
Project Planning and Management
Public Health and HIV/AIDS management
Guidance and Counseling Techniques
Family Planning and RH issues
Research Methods and Data management
Specialized Statistical Packages for data analysis (SPSS, STATA, EVIEWs, ATLAS TI, SUDAN, EPINFO and Epi Data etc)
Training of Trainers Course
Management and Leadership Skills Development
Procurement and Contract Management
Peace and conflict Management/Resolution
Disaster Preparedness and Management Course
Communication Skills and Technique
Heath Care Administration (HCAD)
Interdisciplinary Environmental Health Studies (ENVHs)
Substance Abuse and Addictions Management (SAAM)
Advocacy and Lobbying Techniques
Strategic Planning and Management
Business Sales and Marketing Strategies
Health Marketing and Health Promotion
Logistics, Transport and Supply Chain Management

Background to Public Health Course

A certificate in public health will provide professionals working in health-related fields an opportunity to receive graduate academic training in critical areas in public health. Participants will receive training in the five core modules of Africa Population Institute Public health program with this certificate. This Certificate in Public Health is a module designed program which is designed to:

- offer a dynamic program that prepares students/participants to become leaders within their own professional area
- prepare practitioners who will be able to exercise higher levels of judgment and problem solving in relation to health promotion and encourage development in others
- promote opportunities for inter-professional education; this will represent the collaborative nature of public health practice
- develop and promote the skills of synthesis, together with critical, analytical and reflective thinking, and promote an ethos of lifelong learning
- promote learning opportunities which allow practitioners to enhance professional knowledge and expertise in their sphere of practice, and to utilise best evidence to promote research-based change
- provide theoretical underpinning for the key areas of competency in public health practice.

Program Structure and Features

The philosophy of this program is underpinned by evidence-based practice. It puts greater emphasis on practical application of public health evidence into practice problems.

The program consists of the following modules, all of which have internal integrity and the capacity to be linked to other modules to form a coherent program of learning:

- Health Promotion in a Public Health Practice
- Principles of Epidemiology
- Inequalities in Health

The content of the program takes into account the need for Public Health practitioners to develop high levels of expertise on three domains of Public Health practice:

- Health Protection
- Health Improvement
- Health Service Delivery and Quality

The content is underpinned by and reflects the framework for International Standards for Specialist Practice in Public Health , and is linked to a higher education accreditation system, with existing mechanisms for the award of credit for appropriate learning and prior experiential learning.

All five modules must be taken to complete a general plan of this study in the following manner.

PHC MODULE 1— Introduction to Public Health Administrative Systems

Overview of the public health system, including public health concepts and practice and health care delivery and financing. Focus on understanding of organization and administration of health services, structure and functions of public health system, and health insurance programs.

PHC MODULE 2— Environmental Health Concepts in Public Health

Survey of major topics of environmental health. Sources, routes, media, and health outcomes associated with biological, chemical, and physical agents in environment. Effects of agents on disease, water quality, air quality, food safety, and land resources. Current legal framework, policies, and practices associated with environmental health and intended to improve public health.

PHC MODULE 3— Principles of Epidemiology in Public Health

Overview of epidemiology methods used in research studies that address disease patterns in community and clinic-based populations. Includes distribution and determinants of health-related states or events in specific populations and application to control of health problems.

PHC MODULE 4— Statistical Methods for Health Science

Appropriate use of data summarization and presentation of basic statistical methods, including ANOVA, nonparametric methods, inference on discrete data, inference on survival data, and regression methods for continuous, binary, and survival data.

PHC MODULE 5—Psychological, Behavioral, and Social Issues in Public Health

Health and behavior from social and community perspectives, including comparison of various theories regarding social and behavioral approaches to public health.

MODULE 1— Introduction to Public Health Administrative Systems

Public health is "the science and art of preventing disease, prolonging life and promoting health through the organized efforts and informed choices of society, organizations, public and private, communities and individuals." (1920, C.E.A. Winslow). It is concerned with threats to the overall health of a community based on population health analysis. The population in question can be as small as a handful of people or as large as all the inhabitants of several continents (for instance, in the case of a pandemic). Public health is typically divided into epidemiology, biostatistics and health services. Environmental, social, behavioral, and occupational health are also important subfields.
There are 2 distinct characteristics of public health:
1. It deals with preventive rather than curative aspects of health
2. It deals with population-level, rather than individual-level health issues
The focus of public health intervention is to prevent rather than treat a disease through surveillance of cases and the promotion of healthy behaviors. In addition to these activities, in many cases treating a disease may be vital to preventing it in others, such as during an outbreak of an infectious disease. Hand washing, vaccination programs and distribution of condoms are examples of public health measures. The goal of public health is to improve lives through the prevention and treatment of disease. The United Nations' World Health Organization defines health as "a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity.

Objectives of Module 1
The focus of a public health intervention is to prevent rather than treat a disease through surveillance of cases and the promotion of healthy behaviors. In addition to these activities, in many cases treating a disease can be vital to preventing its spread to others, such as during an outbreak of infectious disease or contamination of food or water supplies. Vaccination programs and distribution of condoms are examples of public health measures.

Most countries have their own government public health agencies, sometimes known as ministries of health, to respond to domestic health issues. In the United States, the front line of public health initiatives are state and local health departments. The United States Public Health Service (PHS), led by the Surgeon General of the United States, and the Centers for Disease Control and Prevention, headquartered in Atlanta and a part of the PHS, are involved with several international health activities, in addition to their national duties.

There is a vast discrepancy in access to health care and public health initiatives between developed nations and developing nations. In the developing world, public health infrastructures are still forming. There may not be enough trained health workers or monetary resources to provide even a basic level of medical care and disease prevention. As a result, a large majority of disease and mortality in the developing world results from and contributes to extreme poverty. For example, many African governments spend less than USD$10 per person per year on health care, while, in the United States, the federal government spent approximately USD$4,500 per capita in 2000.

Many diseases are preventable through simple, non-medical methods. For example, research has shown that the simple act of hand washing can prevent many contagious diseases.
Public health plays an important role in disease prevention efforts in both the developing world and in developed countries, through local health systems and through international non-governmental organizations.

History of public health

In some ways, public health is a modern concept, although it has roots in antiquity. From the beginnings of human civilization, it was recognized that polluted water and lack of proper waste disposal spread communicable diseases (theory of miasma). Early religions attempted to regulate behavior that specifically related to health, from types of food eaten, to regulating certain indulgent behaviors, such as drinking alcohol or sexual relations. The establishment of governments placed responsibility on leaders to develop public health policies and programs in order to gain some understanding of the causes of disease and thus ensure social stability prosperity, and maintain order.

Early public health interventions

By Roman times, it was well understood that proper diversion of human waste was a necessary tenet of public health in urban areas. The Chinese developed the practice of variolation following a smallpox epidemic around 1000 BC. An individual without the disease could gain some measure of immunity against it by inhaling the dried crusts that formed around lesions of infected individuals. Also, children were protected by inoculating a scratch on their forearms with the pus from a lesion. This practice was not documented in the West until the early-1700s, and was used on a very limited basis. The practice of vaccination did not become prevalent until the 1820s, following the work of Edward Jenner to treat smallpox.

During the 14th century Black Death in Europe, it was believed that removing bodies of the dead would further prevent the spread of the bacterial infection. This did little to stem the plague, however, which was most likely spread by rodent-borne fleas. Burning parts of cities resulted in much greater benefit, since it destroyed the rodent infestations. The development of quarantine in the medieval period helped mitigate the effects of other infectious diseases. However, according to Michel Foucault, the plague model of governmentality was later controverted by the cholera model. A Cholera pandemic devastated Europe between 1829 and 1851, and was first fought by the use of what Foucault called "social medicine", which focused on flux, circulation of air, location of cemeteries, etc. All those concerns, born of the miasma theory of disease, were mixed with urbanistic concerns for the management of populations, which Foucault designated as the concept of "biopower". The German conceptualized this in the *Polizeiwissenschaft* ("Science of police").

The science of epidemiology was founded by John Snow's identification of a polluted public water well as the source of an 1854 cholera outbreak in London. Dr. Snow believed in the germ theory of disease as opposed to the prevailing miasma theory. Although miasma theory correctly teaches that disease is a result of poor sanitation, it was based upon the prevailing theory of spontaneous generation. Germ theory developed slowly: despite Anton van Leeuwenhoek's observations of Microorganisms, (which are now known to cause many of the most common infectious diseases) in the year 1680, the modern era of public health did not begin until the 1880s, with Louis Pasteurs germ theory and production of artificial vaccines. Public health nursing made available through child welfare services. Other public health interventions include latrinization, the building of sewers, the regular collection of garbage followed by incineration or disposal in a landfill, providing clean water and draining standing water to prevent the breeding of mosquitos.

Modern public health

As the prevalence of infectious diseases in the developed world decreased through the 20th century, public health began to put more focus on chronic diseases such as cancer and heart disease. An emphasis on physical exercise was reintroduced. In America, public health worker Dr. Sara Josephine Baker lowered the infant mortality rate using preventative methods. She established many programs to help the poor in New York City keep their infants healthy. Dr. Baker led teams of nurses into the crowded neighborhoods of Hell's Kitchen and taught mothers how to dress, feed, and bathe their babies. After World War I many states and countries followed her example in order to lower infant mortality rates.

During the 20th century, the dramatic increase in average life span is widely credited to public health achievements, such as vaccination programs and control of infectious diseases, effective safety policies such as motor-vehicle and occupational safety, improved family planning, fluoridation of drinking water, anti-smoking measures, and programs designed to decrease chronic disease. Meanwhile, the developing world remained plagued by largely preventable infectious diseases, exacerbated by malnutrition and poverty.

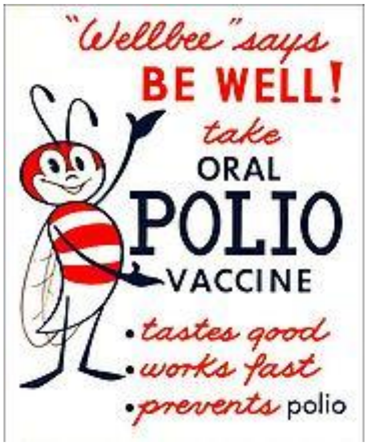
Front-page headlines continue to present society with public health issues on a daily basis: emerging infectious diseases such as SARS, making its way from China (see Public health in China) to Canada and the United States; prescription drug benefits under public programs such as Medicare; the increase of HIV-AIDS among young heterosexual women and its spread in South Africa; the increase of childhood obesity and the concomitant increase in type II diabetes among children; the impact of adolescent pregnancy; and the ongoing social, economic and health disasters related to the 2004 Tsunami and Hurricane Katrina in 2005. These are all ongoing public health challenges.
Since the 1980s, the growing field of population health has broadened the focus of public health from individual behaviors and risk factors to population-level issues such as inequality, poverty, and education.

Modern public health is often concerned with addressing determinants of health across a population, rather than advocating for individual behaviour change. There is a recognition that our health is affected by many factors including where we live, genetics, our income, our educational status and our social relationships - these are known as "social determinants of health." A social gradient in health runs through society, with those that are poorest generally suffering the worst health. However even those in the *middle classes* will generally have worse health outcomes than those of a higher social stratum. The *new* public health seeks to address these health inequalities by advocating for population-based policies that improve the health of the whole population in an equitable fashion.

The burden of treating conditions caused by unemployment, poverty, unfit housing and environmental pollution have been calculated to account for between 16-22% of the clinical budget of the British National Health Service.
UK Public health functions include:

- Health surveillance, monitoring and analysis
- Investigation of disease outbreaks, epidemics and risk to health
- Establishing, designing and managing health promotion and disease prevention programmes
- Enabling and empowering communities to promote health and reduce inequalities
- Creating and sustaining cross-Government and intersectoral partnerships to improve health and reduce inequalities
- Ensuring compliance with regulations and laws to protect and promote health
- Developing and maintaining a well-educated and trained, multi-disciplinary public health workforce
- Ensuring the effective performance of NHS services to meet goals in improving health, preventing disease and reducing inequalities
- Research, development, evaluation and innovation
- Quality assuring the public health function

Public health programs



This 1963 poster featured CDC's national symbol of public health, the "Wellbee", encouraging the public to receive an oral polio vaccine. Today, most governments recognize the importance of public health programs in reducing the incidence of disease, disability, and the effects of aging, although public health generally receives significantly less government funding compared with medicine. In recent years, public health programs providing vaccinations have made incredible strides in promoting health, including the eradication of smallpox, a disease that plagued humanity for thousands of years.

An important public health issue facing the world currently is HIV/AIDS. Antibiotic resistance is another major concern, leading to the reemergence of diseases such as Tuberculosis.
Another major public health concern is diabetes. In 2006, according to the World Health Organization, at least 171 million people worldwide suffered from diabetes. Its incidence is increasing rapidly, and it is estimated that by the year 2030, this number will double.

A controversial aspect of public health is the control of smoking. Many nations have implemented major initiatives to cut smoking, such as increased taxation and bans on smoking in some or all public places. Proponents argue by presenting evidence that smoking is one of the major killers in all developed countries, and that therefore governments have a duty to reduce the death rate, both through limiting passive (second-hand) smoking and by providing fewer opportunities for smokers to smoke. Opponents say that this undermines individual freedom and personal responsibility (often using the phrase nanny state in the UK), and worry that the state may be emboldened to remove more and more choice in the name of better population health overall. However, proponents counter that inflicting disease on other people via passive smoking is not a human right, and in fact smokers are still free to smoke in their own homes. There is also a link between public health and veterinary public health which deals with zoonotic diseases, diseases that can be transmitted from animals to humans. (See also Vector control).

Public hygiene

Public hygiene includes public behaviors individuals can take to improve their personal health and wellness. Topics include public transportation, food preparation and public washroom use. These are steps individuals can take themselves. Examples would include avoiding crowded subways during the flu season, using gloves when touching the handrails and opening doors in public malls as well as going to clean restaurants.

Economics of public health

The application of economics to the realm of public health has been rising in importance since the 1980s. Economic studies can show, for example, where limited public resources might best be spent to save lives or cause the greatest increase in quality of life.

Research
Public health investigates sources of disease and descriptors of health through scientific methodology. This can lead to a public health solution to an epidemic, or a community based intervention for chronic diseases. Either way, research can provide the link between cause and effect for public health issues.

Community based participatory research
In contrast to clinical, patient oriented, or literature review research, community based participatory research (CBPR) investigates community-based etiology, involves community leaders, and overall respects the forces under which the community and its participants preside toward promoting and sustaining public health matters. CBPR describes the approach embraced by such institutions as the Gamelan Council and, as described by the WK Kellogg Foundation Community Health Scholars Program, is a *"collaborative approach to research that equitably involves all partners in the research process and recognizes the unique strengths that each brings. CBPR begins with a research topic of importance to the community, has the aim of combining knowledge with action and achieving social change to improve health outcomes and eliminate health disparities.* CBPR methods have been necessary for implementation of certain public health actions. This have been difficult to accomplish because communities in poorer, less well developed areas often distrust researchers and scientists from "outside."

MODULE 2— Environmental Health Concepts in Public Health
Sources, routes, media, and health outcomes associated with biological, chemical, and physical agents in environment. Effects of agents on disease, water quality, air quality, food safety, and land resources. Current legal framework, policies, and practices associated with environmental health and intended to improve public health

Environmental factors have a major impact on human health. We depend on clean air, clean water, and a safe food supply to sustain ourselves. Activities in agriculture, transportation, construction, energy production and use, and manufacturing have a significant impact on the environment. Issues such as air pollution and contamination of food or drinking water appear frequently in the news and require the expertise of trained professionals to determine the source of the pollution and implement preventive measures against occurrences.

Are environmental laws adequately enforced? Should ground beef be irradiated to protect against E Coli contamination? Do the benefits from eating conventionally grown produce outweigh the risks from pesticide residues? These are a few questions that confront environmental health professionals.

Environmental health science is a multi-faceted field incorporating a variety of professionals, including epidemiologists, toxicologists, clinicians, industrial hygienists, policy analysts, attorneys, and molecular biologists. Environmental health professionals typically hold positions in industry either in compliance programs or as health and safety officers; in consulting firms performing environmental monitoring; in not-for-profit environmental organizations as researchers; or in government in the areas of inspections and permits, water quality, food, sanitation, or air quality.

The Program in Environmental Health Science at the School of Public Health provides the conceptual framework and practical tools to recognize and address environmental hazards. Required concentration courses familiarize the student with issues of air and water quality as well as pollution, environmental law, and risk assessment. Students learn analytical and problem-solving skills as well as how to communicate with technical and non-technical groups.

What is Environmental Health?
While there are many ways to define environmental health, we think about it as how the environment might affect a person's health. The environment is the air we breathe, our water, our food and our surroundings. It's the chemicals, radiation and microbes as well as the physical world that we have contact with everyday. Understanding how we interact with the environment is complicated; so is understanding how the environment may affect our health.
This web site focuses mainly on how chemicals in the environment might affect health. It provides links to the following kinds of data:

- **Environmental hazard data** about chemicals or other substances in the environment.
- **Exposure data** about the level of a chemical in people's bodies.
- **Data about health conditions and diseases,** such as asthma or birth defects.
- **Other data** helpful to explore relationships between exposures and health effects, such as data that describes people (age, race and sex), or behavior or lifestyle choices that could affect the likelihood that people are diagnosed with particular health condition.

What is Tracking?
One of the tools used to examine trends in environmental health is tracking or surveillance. Tracking means continuously collecting, examining and sharing data. In environmental health tracking, we are interested in examining trends of environmental hazards, exposures, health conditions and population. We do this for individual data sets or combinations of data sets.
Tracking information can be shared in different ways, such as maps or tables. You may be familiar with some environmental health tracking information such as EPA's Air Quality Index that reports daily air quality, or New York State's Cancer Registry that collects and reports cancer incidence data across the state.

Biomonitoring
An effective tool to assess human exposure to environmental pollutants. Biomonitoring is the direct measurement of a substance, or its metabolite, in biological specimens such as blood or urine, using biomarkers. Measurement by biomonitoring allows for the most health-relevant assessment of exposure, since it indicates the amount of the chemical that actually gets into the body from environmental sources (such as air, soil, water, dust or food), rather than the amount that may have come from these sources.

Environmental Burden of Disease
The environmental burden of disease quantifies the amount of disease caused by environmental risks. Disease burden can be expressed in deaths, incidence or in Disability-Adjusted Life Years (DALY). The DALY combines the burden due to death and disability in a single index.

Environmental Epidemiology
Environmental epidemiology deals with environmental conditions and hazards that may pose a risk to human health. Environmental exposures include air pollution, hazardous waste, metals, pesticides, and radiation, and health effects include cancer, cardiovascular disease, neurological effects, and reproductive effects.

Environmental Exposure
The contact of a chemical or other agent with the skin or the openings into the body such as the mouth, the nostrils, and punctures and lesions in the skin. Exposure may be either short term (acute) or long term (chronic).

Environmental Health
Environmental health comprises those aspects of human health and disease that are determined by factors in the environment. It also refers to the theory and practice of assessing and controlling factors in the environment that can potentially affect health. Important sources of environmental health risks include industrial waste, air emissions and water discharges, human waste, consumer products, living conditions, transport, and ionizing and non-ionizing radiation. Health effects with known or suspected environmental etiologies include cancer, cardio-pulmonary diseases, asthma and other respiratory diseases, allergies, neuro-toxicity and neurological impairment, gastro-intestinal diseases, developmental and congenital abnormalities, and acute poisonings.

Environmental Justice
Equal protection from environmental health hazards for individuals, groups or communities, regardless of race, ethnicity, or economic status. Environmental justice applies to the setting, implementation and enforcement of laws, regulations and environmental policy. Environmental justice aims to prevent situations where a particular social group is forced to shoulder a disproportionate share of damage from pollution, environmental health risks and the like, due to lack of economic or political power.

Environmental Standard
An environmental standard is a criterion (typically quantifiable) against which environmental quality can be assessed. The standard may apply to, among other things, industrial processes, pollutants, various facilities, products and vehicles. The adoption of the standard by a legally authorized body grants it an official and binding status. However, a standard can also be part of recommended guidelines, or a voluntary professional code, whose legitimacy derives from another source, such as the scientific authority or the market power of the standard-setting body. An environmental standard is typically based on risk assessment, while taking into account, among other things, economic considerations, the availability of technology and safety factors.

Epidemiology
The study of the distribution of diseases in populations and of factors that influence the occurrence of disease. Epidemiology examines epidemic (excess) and endemic (always present) diseases; it is based on the observation that most diseases do not occur randomly, but are related to environmental and personal characteristics that vary by place, time, and subgroup of the population. The epidemiologist attempts to determine who is prone to a particular disease; where risk of the disease is highest; when the disease is most likely to occur and its trends over time; what exposure its victims have in common; how much the risk is increased through exposure; and how many cases of the disease could be avoided by eliminating the exposure.

Exposure Assessment
A quantitative or qualitative evaluation of environmental exposure, which considers the intensity, frequency and duration of contact with a chemical (or other) agent, as well as the number of people exposed, the route of exposure (e.g. dermal, oral or respiratory), chemical intake or uptake rate, the amount of the chemical a person or population were exposed to and the amount absorbed.

Global Environmental Change
Global environmental change includes changes in the physical and biogeochemical environment such as climate change, stratospheric ozone depletion, and stresses on ecosystems, either caused naturally or influenced by human activities such as deforestation, fossil fuel, water and land consumption, urbanization, intensive agriculture, over-exploitation of fisheries and waste production. Global environmental change causes direct and indirect hazards to human health and requires the recognition

that the long-term safeguarding of human health depends on the functioning of the biosphere's life-supporting systems.

Health Impact Assessment (HIA)
A combination of procedures, methods, and tools by which a policy, program, or project may be judged as to its potential effects on the health of a population, and the distribution of those effects within the population. An HIA can provide recommendations to increase positive health outcomes and to minimize adverse outcomes. A major benefit of the HIA is that it brings public health issues to the attention of decision-makers in areas, such as transportation or land use, which do not traditionally deal with health.

Interdisciplinary Research
Research that integrates information, concepts, methodology, data, tools, training, perspectives and/or theories from two or more disciplines or bodies of specialized knowledge, so as to advance knowledge or to solve problems whose solutions are beyond the scope of a single discipline or research field. Interdisciplinary research is not just the combination of two disciplines but, rather, an integration and synthesis of ideas and methods to create a new field. As such it differs from multidisciplinary research, which involves various disciplines, but in which researchers typically focus on their particular fields of research.

National Environmental Health Action Plans (NEHAPs)
NEHAPs are government documents that address environmental health problems in a comprehensive, holistic and intersectoral manner. Most NEHAPs address environmental health management (including information systems, environmental health services, public information and participation, education, research and economic tools), environmental health hazards (including ambient air, drinking and bathing water, noise, radiation, food safety, housing and workplace), and economic sectors (including industry, agriculture, transportation and energy).

Public Health
A field which combines basic and applied research as well as education, social sciences, economics and management, in order to promote the health and welfare of the individual and the community. The field includes activities focused on reducing factors that contribute to the burden of disease as well as activities that promote the health of the community, which include, among others, the assessment and monitoring of public health, policy formulation, preventive medicine and health education.

Risk Assessment (in Environmental Health)
Risk assessment is conducted to estimate damage or injury, which could be caused from exposure to a given environmental hazard. It is a methodology that combines information on the risk from various sources in order to provide a basis for decision-making and regulation. A health risk assessment typically includes hazard identification, exposure assessment, a dose-response assessment and risk characterization.

Risk Communication
The National Research Council (NRC) defines risk communication as "an interactive process of exchange of information and opinion among individuals, groups, and institutions." The concept relates to a range of activities directed at increasing the public's knowledge of risk issues and participation in risk management. Such activities include, for example, warning labels for consumers about potential hazards in the use of a product, development of publicly accessible databases, or a public hearing on the risks of a new technology. Risk communication is a relatively new field, which developed in the early 1980s following research that showed a difference in risk perception between experts and the public, a

fact that made justifying decisions on risks post facto problematic. Nowadays, risk communication is viewed as a dialogue among interested parties: risk experts, policy makers, and affected segments of the public.

EPA's Seven Cardinal Rules of Risk Communication: Risk Management
The process of weighing policy alternatives and selecting appropriate regulation based on the results of risk assessment and incorporating social, economic, political and moral concerns. Risk assessment provides input for the decision-maker, but cannot answer the question: 'is the risk acceptable?' Risk management provides an answer to this question as well as other questions such as the tradeoff between the cost of preventing or reducing the risk and the expected benefits, how uncertainties should be considered, what risks may arise as a consequence of regulation and which risks should be prioritized. Increasingly, risk perception and risk communication with the public are recognized as integral elements in the process of risk management.

MODULE 3— Principles of Epidemiology in Public Health

Recently, a news story described an inner-city neighborhood's concern about the rise in the number of children with asthma. Another story reported the revised recommendations for who should receive influenza vaccine this year. A third story discussed the extensive disease-monitoring strategies being implemented in a city recently affected by a massive hurricane. A fourth story described a finding published in a leading medical journal of an association in workers exposed to a particular chemical and an increased risk of cancer. Each of these news stories included interviews with public health officials or researchers who called themselves epidemiologists. Well, who are these epidemiologists, and what do they do? What is epidemiology? This lesson is intended to answer those questions by describing what epidemiology is, how it has evolved and how it is used today, and what some of the key methods and concepts are. The focus is on epidemiology in public health practice, that is, the kind of epidemiology that is done at health departments.

Objectives
By the end of this Module, you will be able to:
- *Define epidemiology*
- *Summarize the historical evolution of epidemiology*
- *Name some of the key uses of epidemiology*
- *Identify the core epidemiology functions*
- *Describe primary applications of epidemiology in public health practice*
- *Specify the elements of a case definition and state the effect of changing the value of any of the elements*
- *List the key features and uses of descriptive epidemiology*
- *List the key features and uses of analytic epidemiology*
- *List the three components of the epidemiologic triad*
- *Describe the different modes of transmission of communicable disease in a population*

Definition of Epidemiology
The word epidemiology comes from the Greek words *epi*, meaning on or upon, *demos*, meaning people, and *logos*, meaning the study of. In other words, the word epidemiology has its roots in the study of what befalls a population. Many definitions have been proposed, but the following definition captures the underlying principles and public health spirit of epidemiology:

Epidemiology is the ***study*** *of the* ***distribution*** *and* ***determinants*** *of* ***health-related states or events*** *in* ***specified populations****, and the* ***application*** *of this study to the control of health problems.*
Key terms in this definition reflect some of the important principles of epidemiology.

Study:
Epidemiology is a scientific discipline with sound methods of scientific inquiry at its foundation. Epidemiology is data-driven and relies on a systematic and unbiased approach to the collection, analysis, and interpretation of data. Basic epidemiologic methods tend to rely on careful observation and use of valid comparison groups to assess whether what was observed, such as the number of cases of disease in a particular area during a particular time period or the frequency of an exposure among persons with disease, differs from what might be expected. However, epidemiology also draws on methods from other scientific fields, including biostatistics and informatics, with biologic, economic, social, and behavioral sciences.

In fact, epidemiology is often described as the basic science of public health, and for good reason. First, epidemiology is a quantitative discipline that relies on a working knowledge of probability, statistics, and sound research methods. Second, epidemiology is a method of causal reasoning based on developing and testing hypotheses grounded in such scientific fields as biology, behavioral sciences, physics, and ergonomics to explain health-related behaviors, states, and events. However, epidemiology is not just a research activity but an integral component of public health, providing the foundation for directing practical and appropriate public health action based on this science and causal reasoning.

Distribution:
Epidemiology is concerned with the **frequency** and **pattern** of health events in a population:
Frequency refers not only to the number of health events such as the number of cases of meningitis or diabetes in a population, but also to the relationship of that number to the size of the population. The resulting rate allows epidemiologists to compare disease occurrence across different populations.
Pattern refers to the occurrence of health-related events by time, place, and person. Time patterns may be annual, seasonal, weekly, daily, hourly, weekday versus weekend, or any other breakdown of time that may influence disease or injury occurrence. Place patterns include geographic variation, urban/rural differences, and location of work sites or schools. Personal characteristics include demographic factors which may be related to risk of illness, injury, or disability such as age, sex, marital status, and socioeconomic status, as well as behaviors and environmental exposures. Characterizing health events by time, place, and person are activities of **descriptive epidemiology**, discussed in more detail later in this lesson.

Determinants:
Epidemiology is also used to search for **determinants**, which are the causes and other factors that influence the occurrence of disease and other health-related events. Epidemiologists assume that illness does not occur randomly in a population, but happens only when the right accumulation of risk factors or determinants exists in an individual. To search for these determinants, epidemiologists use analytic epidemiology or epidemiologic studies to provide the "Why" and "How" of such events. They assess whether groups with different rates of disease differ in their demographic characteristics, genetic or immunologic make-up, behaviors, environmental exposures, or other so-called potential risk factors. Ideally, the findings provide sufficient evidence to direct prompt and effective public health control and prevention measures.

Health-related states or events:
Epidemiology was originally focused exclusively on epidemics of communicable diseases but was subsequently expanded to address endemic communicable diseases and non-communicable infectious

diseases. By the middle of the 20th Century, additional epidemiologic methods had been developed and applied to chronic diseases, injuries, birth defects, maternal-child health, occupational health, and environmental health. Then epidemiologists began to look at behaviors related to health and well-being, such as amount of exercise and seat belt use. Now, with the recent explosion in molecular methods, epidemiologists can make important strides in examining genetic markers of disease risk. Indeed, the term healthrelated states or events may be seen as anything that affects the well-being of a population. Nonetheless, many epidemiologists still use the term “disease” as shorthand for the wide range of health-related states and events that are studied.

Specified populations:
Although epidemiologists and direct health-care providers (clinicians) are both concerned with occurrence and control of disease, they differ greatly in how they view “the patient.” The clinician is concerned about the health of an individual; the epidemiologist is concerned about the collective health of the people in a community or population. In other words, the clinician’s “patient” is the individual; the epidemiologist’s “patient” is the community. Therefore, the clinician and the epidemiologist have different responsibilities when faced with a person with illness. For example, when a patient with diarrheal disease presents, both are interested in establishing the correct diagnosis. However, while the clinician usually focuses on treating and caring for the individual, the epidemiologist focuses on identifying the exposure or source that caused the illness; the number of other persons who may have been similarly exposed; the potential for further spread in the community; and interventions to prevent additional cases or recurrences.

Application:
Epidemiology is not just “the study of” health in a population; it also involves applying the knowledge gained by the studies to community-based practice. Like the practice of medicine, the practice of epidemiology is both a science and an art. To make the proper diagnosis and prescribe appropriate treatment for a patient, the clinician combines medical (scientific) knowledge with experience, clinical judgment, and understanding of the patient.Similarly, the epidemiologist uses the scientific methods of descriptive and analytic epidemiology as well as experience, epidemiologic judgment, and understanding of local conditions in “diagnosing” the health of a community and proposing appropriate, practical, and acceptable public health interventions to control and prevent disease in the community.

Summary:
Epidemiology is the study (scientific, systematic, data-driven) of the distribution (frequency, pattern) and determinants (causes, risk factors) of health-related states and events (not just diseases) in specified populations (patient is community, individuals viewed collectively), and the application of (since epidemiology is a discipline within public health) this study to the control of health problems.

Core Epidemiologic Functions
In the mid-1980s, five major tasks of epidemiology in public health practice were identified: **public health surveillance, field investigation, analytic studies, evaluation, and linkages.** A sixth task, **policy development**, was recently added. These tasks are described below.

Public health surveillance:
Public health surveillance is the ongoing, systematic collection, analysis, interpretation, and dissemination of health data to help guide public health decision making and action. Surveillance is equivalent to monitoring the pulse of the community. The purpose of public health surveillance, which is sometimes called “information for action,”18 is to portray the ongoing patterns of disease occurrence and disease potential so that investigation, control, and prevention measures can be applied efficiently

and effectively. This is accomplished through the systematic collection and evaluation of morbidity and mortality reports and other relevant health information, and the dissemination of these data and their interpretation to those involved in disease control and public health decision making.

Morbidity and mortality reports are common sources of surveillance data for local and state health departments. These reports generally are submitted by health-care providers, infection control practitioners, or laboratories that are required to notify the health department of any patient with a reportable disease such as pertussis, meningococcal meningitis, or AIDS. Other sources of health-related data that are used for surveillance include reports from investigations of individual cases and disease clusters, public health program data such as immunization coverage in a community, disease registries, and health surveys.

Most often, surveillance relies on simple systems to collect a limited amount of information about each case. Although not every case of disease is reported, health officials regularly review the case reports they do receive and look for patterns among them. These practices have proven invaluable in detecting problems, evaluating programs, and guiding public health action.

While public health surveillance traditionally has focused on communicable diseases, surveillance systems now exist that target injuries, chronic diseases, genetic and birth defects, occupational and potentially environmentally-related diseases, and health behaviors. Since September 11, 2001, a variety of systems that rely on electronic reporting have been developed, including those that report daily emergency department visits, sales of over-the-counter medicines, and worker absenteeism.19,20 Because epidemiologists are likely to be called upon to design and use these and other new surveillance systems, an epidemiologist's core competencies must include design of data collection instruments, data management, descriptive methods and graphing, interpretation of data, andscientific writing and presentation.

Field investigation:
As noted above, surveillance provides information for action. One of the first actions that results from a surveillance case report or report of a cluster is investigation by the public health department.
The investigation may be as limited as a phone call to the healthcare provider to confirm or clarify the circumstances of the reported case, or it may involve a field investigation requiring the coordinated efforts of dozens of people to characterize the extent of an epidemic and to identify its cause.
The objectives of such investigations also vary. Investigations often lead to the identification of additional unreported or unrecognized ill persons who might otherwise continue to spread infection to others. For example, one of the hallmarks of investigations of persons with sexually transmitted disease is the identification of sexual partners or contacts of patients. When interviewed, many of these contacts are found to be infected without knowing it, and are given treatment they did not realize they needed. Identification and treatment of these contacts prevents further spread. For some diseases, investigations may identify a source or vehicle of infection that can be controlled or eliminated.

For example, the investigation of a case of *Escherichia coli* O157:H7 infection usually focuses on trying to identify the vehicle, often ground beef but sometimes something more unusual such as fruit juice. By identifying the vehicle, investigators may be able to determine how many other persons might have already been exposed and how many continue to be at risk. When a commercial product turns out to be the culprit, public announcements and recalling the product may prevent many additional cases.

Occasionally, the objective of an investigation may simply be to learn more about the natural history, clinical spectrum, descriptive epidemiology, and risk factors of the disease before determining what disease intervention methods might be appropriate. Early investigations of the epidemic of SARS in

2003 were needed to establish a case definition based on the clinical presentation, and to characterize the populations at risk by time, place, and person. As more was learned about the epidemiology of the disease and communicability of the virus, appropriate recommendations regarding isolation and quarantine were issued.

Field investigations of the type described above are sometimes referred to as "shoe leather epidemiology," conjuring up images of dedicated, if haggard, epidemiologists beating the pavement in search of additional cases and clues regarding source and mode of transmission. This approach is commemorated in the symbol of the Epidemic Intelligence Service (EIS), CDC's training program for disease detectives — a shoe with a hole in the sole.

Analytic studies:
Surveillance and field investigations are usually sufficient to identify causes, modes of transmission, and appropriate control and prevention measures. But sometimes analytic studies employing more rigorous methods are needed. Often the methods are used in combination — with surveillance and field investigations providing clues or hypotheses about causes and modes of transmission, and analytic studies evaluating the credibility of those hypotheses.

Clusters or outbreaks of disease frequently are investigated initially with descriptive epidemiology. The descriptive approach involves the study of disease incidence and distribution by time, place, and person. It includes the calculation of rates and identification of parts of the population at higher risk than others. Occasionally, when the association between exposure and disease is quite strong, the investigation may stop when descriptive epidemiology is complete and control measures may be implemented immediately.

John Snow's 1854 investigation of cholera is an example. More frequently, descriptive studies, like case investigations, generate hypotheses that can be tested with analytic studies. While some field investigations are conducted in response to acute health problems such as outbreaks, many others are planned studies.

The hallmark of an analytic epidemiologic study is the use of a valid comparison group. Epidemiologists must be skilled in all aspects of such studies, including design, conduct, analysis, interpretation, and communication of findings.
• **Design** includes determining the appropriate research strategy and study design, writing justifications and protocols, calculating sample sizes, deciding on criteria for subject selection (e.g., developing case definitions), choosing an appropriate comparison group, and designing questionnaires.
• **Conduct** involves securing appropriate clearances and approvals, adhering to appropriate ethical principles, abstracting records, tracking down and interviewing subjects, collecting and handling specimens, and managing the data.
• **Analysis** begins with describing the characteristics of the subjects. It progresses to calculation of rates, creation of comparative tables (e.g., two-by-two tables), and computation of measures of association (e.g., risk ratios or odds ratios), tests of significance (e.g., chi-square test), confidence intervals, and the like. Many epidemiologic studies require more advanced analytic techniques such as stratified analysis, regression, and modeling.
• Finally, **interpretation** involves putting the study findings into perspective, identifying the key take-home messages, and making sound recommendations. Doing so requires that the epidemiologist be knowledgeable about the subject matter and the strengths and weaknesses of the study.

Evaluation

Epidemiologists, who are accustomed to using systematic and quantitative approaches, have come to play an important role in evaluation of public health services and other activities. Evaluation is the process of determining, as systematically and objectively as possible, the relevance, effectiveness, efficiency, and impact of activities with respect to established goals.22

- **Effectiveness** refers to the ability of a program to produce the intended or expected results in the field; effectiveness differs from **efficacy**, which is the ability to produce results under ideal conditions.
- **Efficiency** refers to the ability of the program to produce the intended results with a minimum expenditure of time and resources.

The evaluation itself may focus on plans (formative evaluation), operations (process evaluation), impact (summative evaluation), or outcomes — or any combination of these. Evaluation of an immunization program, for example, might assess the efficiency of the operations, the proportion of the target population immunized, and the apparent impact of the program on the incidence of vaccine-preventable diseases. Similarly, evaluation of a surveillance system might address operations and attributes of the system, its ability to detect cases or outbreaks, and its usefulness.

Linkages:
Epidemiologists working in public health settings rarely act in isolation. In fact, field epidemiology is often said to be a "team sport." During an investigation an epidemiologist usually participates as either a member or the leader of a multidisciplinary team. Other team members may be laboratorians, sanitarians, infection control personnel, nurses or other clinical staff, and, increasingly, computer information specialists.

Many outbreaks cross geographical and jurisdictional lines, so co-investigators may be from local, state, or federal levels of government, academic institutions, clinical facilities, or the private sector. To promote current and future collaboration, the epidemiologists need to maintain relationships with staff of other agencies and institutions. Mechanisms for sustaining such linkages include official memoranda of understanding, sharing of published or on-line information for public health audiences and outside partners, and informal networking that takes place at professional meetings.

Policy development:
The definition of epidemiology ends with the following phrase: "...and the application of this study to the control of health problems." While some academically minded epidemiologists have stated that epidemiologists should stick to research and not get involved in policy development or even make recommendations, public health epidemiologists do not have this luxury. Indeed, epidemiologists who understand a problem and the population in which it occurs are often in a uniquely qualified position to recommend appropriate interventions. As a result, epidemiologists working in public health regularly provide input, testimony, and recommendations regarding disease control strategies, reportable disease regulations, and health-care policy.

The Epidemiologic Approach

As with all scientific endeavors, the practice of epidemiology relies on a systematic approach. In very simple terms, the epidemiologist:

- **Counts** cases or health events, and describes them in terms of time, place, and person;
- **Divides** the number of cases by an appropriate denominator to calculate rates; and
- **Compares** these rates over time or for different groups of people.

Before counting cases, however, the epidemiologist must decide what a case is. This is done by developing a case definition. Then, using this case definition, the epidemiologist finds and collects information about the case-patients. The epidemiologist then performs descriptive epidemiology by characterizing the cases collectively according to time, place, and person. To calculate the disease rate, the epidemiologist divides the number of cases by the size of the population. Finally, to determine whether this rate is greater than what one would normally expect, and if so to identify factors contributing to this increase, the epidemiologist compares the rate from this population to the rate in an appropriate comparison group, using analytic epidemiology techniques. These epidemiologic actions are described in more detail below. Subsequent tasks, such as reporting the results and recommending how they can be used for public health action, are just as important, but are beyond the scope of this lesson.

Defining a case:
Before counting cases, the epidemiologist must decide what to count, that is, what to call a case. For that, the epidemiologist uses a **case definition**. A case definition is a set of standard criteria for classifying whether a person has a particular disease, syndrome, or other health condition. Some case definitions, particularly those used for national surveillance, have been developed and adopted as national standards that ensure comparability. Use of an agreedupon standard case definition ensures that every case is equivalent, regardless of when or where it occurred, or who identified it.

Furthermore, the number of cases or rate of disease identified in one time or place can be compared with the number or rate from another time or place. For example, with a standard case definition, health officials could compare the number of cases of listeriosis that occurred in Forsyth County, North Carolina in 2000 with the number that occurred there in 1999. Or they could compare the rate of listeriosis in Forsyth County in 2000 with the national rate in that same year. When everyone uses the same standard case definition and a difference is observed, the difference is likely to be real rather than the result of variation in how cases are classified.

Components of a case definition for outbreak investigations:
A case definition consists of clinical criteria and, sometimes, limitations on time, place, and person. The clinical criteria usually include confirmatory laboratory tests, if available, or combinations of symptoms (subjective complaints), signs (objective physical findings), and other findings. Case definitions used during outbreak investigations are more likely to specify limits on time, place, and/or person than those used for surveillance. Contrast the case definition used for surveillance of listeriosis with the case definition used during an investigation of a listeriosis outbreak in North Carolina in 2000.
Both the national surveillance case definition and the outbreak case definition require a clinically compatible illness and laboratory confirmation of *Listeria monocytogenes* from a normally sterile site, but the outbreak case definition adds restrictions on time and place, reflecting the scope of the outbreak.

Many case definitions, such as that shown for listeriosis, require laboratory confirmation. This is not always necessary, however; in fact, some diseases have no distinctive laboratory findings.
Kawasaki syndrome, for example, is a childhood illness with fever and rash that has no known cause and no specifically distinctive laboratory findings. Notice that its case definition (see box below) is based on the presence of fever, at least four of five specified clinical findings, and the lack of a more reasonable explanation.

Criteria in case definitions:
A case definition may have several sets of criteria, depending on how certain the diagnosis is. For example, during an investigation of a possible case or outbreak of measles, a person with a fever and rash might be classified as having a suspected, probable, or confirmed case of measles, depending on what evidence of measles is present. A case might be classified as suspected or probable while waiting for the laboratory results to become available. Once the laboratory provides the report, the case can be reclassified as either confirmed or “not a case,” depending on the laboratory results. In the midst of a large outbreak of a disease caused by a known agent, some cases may be permanently classified as

suspected or probable because officials may feel that running laboratory tests on every patient with a consistent clinical picture and a history of exposure (e.g., chickenpox) is unnecessary and even wasteful. Case definitions should not rely on laboratory culture results alone, since organisms are sometimes present without causing disease.

Modifying case definitions:
Case definitions can also change over time as more information is obtained. The first case definition for SARS, based on clinical symptoms and either contact with a case or travel to an area with SARS transmission, was published in CDC's Morbidity and Mortality Weekly Report (MMWR) on March 21, 2003 (see box below). Two weeks later it was modified slightly. On March 29, after a novel coronavirus was determined to be the causative agent, an interim surveillance case definition was published that included laboratory criteria for evidence of infection with the SARSassociated coronavirus. By June, the case definition had changed several more times. In anticipation of a new wave of cases in 2004, a revised and much more complex case definition was published in December 2003.

Variation in case definitions:
Case definitions may also vary according to the purpose for classifying the occurrences of a disease. For example, health officials need to know as soon as possible if anyone has symptoms of plague or anthrax so that they can begin planning what actions to take. For such rare but potentially severe communicable diseases, for which it is important to identify every possible case, health officials use a sensitive case definition. A sensitive case definition is one that is broad or "loose," in the hope of capturing most or all of the true cases. For example, the case definition for a suspected case of rubella (German measles) is "any generalized rash illness of acute onset." This definition is quite broad, and would include not only all cases of rubella, but also measles, chickenpox, and rashes due to other causes such as drug allergies. So while the advantage of a sensitive case definition is that it includes most or all of the true cases, the disadvantage is that it sometimes includes other illnesses as well.

On the other hand, an investigator studying the causes of a disease outbreak usually wants to be certain that any person included in a study really had the disease. That investigator will prefer a specific or "strict" case definition. For instance, in an outbreak of *Salmonella* Agona infection, the investigators would be more likely to identify the source of the infection if they included only persons who were confirmed to have been infected with that organism, rather than including anyone with acute diarrhea, because some persons may have had diarrhea from a different cause. In this setting, the only disadvantages of a strict case definition are the requirement that everyone with symptoms be tested and an underestimation of the total number of cases if some people with salmonellosis are not tested.

Using counts and rates:
As noted, one of the basic tasks in public health is identifying and counting cases. These counts, usually derived from case reports submitted by health-care workers and laboratories to the health department, allow public health officials to determine the extent and patterns of disease occurrence by time, place, and person. They may also indicate clusters or outbreaks of disease in the community. Counts are also valuable for health planning. For example, a health official might use counts (i.e., numbers) to plan how many infection control isolation units or doses of vaccine may be needed. However, simple counts do not provide all the information a health department needs. For some purposes, the counts must be put into context, based on the population in which they arose. Rates are measures that relate the numbers of cases during a certain period of time (usually per year) to the size of the population in which they occurred.

Descriptive Epidemiology

As noted earlier, every novice newspaper reporter is taught that a story is incomplete if it does not describe the what, who, where, when, and why/how of a situation, whether it be a space shuttle launch or a house fire. Epidemiologists strive for similar comprehensiveness in characterizing an epidemiologic event, whether it be a pandemic of influenza or a local increase in allterrain vehicle crashes. However, epidemiologists tend to use synonyms for the five W's listed above: case definition, person, place, time, and causes/risk factors/modes of transmission.

Descriptive epidemiology covers **time**, **place**, and **person**. Compiling and analyzing data by time, place, and person is desirable for several reasons.

• First, by looking at the data carefully, the epidemiologist becomes very familiar with the data. He or she can see what the data can or cannot reveal based on the variables available, its limitations (for example, the number of records with missing information for each important variable), and its eccentricities (for example, all cases range in age from 2 months to 6 years, plus one 17-yearold.).

• Second, the epidemiologist learns the extent and pattern of the public health problem being investigated — which months, which neighborhoods, and which groups of people have the most and least cases.

• Third, the epidemiologist creates a detailed description of the health of a population that can be easily communicated with tables, graphs, and maps.

• Fourth, the epidemiologist can identify areas or groups within the population that have high rates of disease. This information in turn provides important clues to the causes of the disease, and these clues can be turned into testable hypotheses.

Time:

The occurrence of disease changes over time. Some of these changes occur regularly, while others are unpredictable. Two diseases that occur during the same season each year include influenza (winter) and West Nile virus infection (August– September). In contrast, diseases such as hepatitis B and salmonellosis can occur at any time. For diseases that occur seasonally, health officials can anticipate their occurrence and implement control and prevention measures, such as an influenza vaccination campaign or mosquito spraying. For diseases that occur sporadically, investigators can conduct studies to identify the causes and modes of spread, and then develop appropriately targeted actions to control or prevent further occurrence of the disease. In either situation, displaying the patterns of disease occurrence by time is critical for monitoring disease occurrence in the community and for assessing whether the public health interventions made a difference.

Day of week and time of day. For some conditions, displaying data by day of the week or time of day may be informative. Analysis at these shorter time periods is particularly appropriate for conditions related to occupational or environmental exposures that tend to occur at regularly scheduled intervals. Farm tractor fatalities are displayed by days of the week.32 Note that the number of farm tractor fatalities on Sundays was about half the number on the other days. The pattern of farm tractor injuries by hour, as displayed in Figure 1.8 peaked at 11:00 a.m., dipped at noon, and peaked again at 4:00 p.m. These patterns may suggest hypotheses and possible explanations that could be evaluated with further study. The hourly number of survivors and rescuers presenting to local hospitals in New York following the attack on the World Trade Center on September 11, 2001.

Epidemic period. To show the time course of a disease outbreak or epidemic, epidemiologists use a graph called an epidemic curve. As with the other graphs presented so far, an epidemic curve's y-axis shows the number of cases, while the x-axis shows time as either date of symptom onset or date of diagnosis. Depending on the incubation period (the length of time between exposure and onset of symptoms) and routes of transmission, the scale on the x-axis can be as broad as weeks (for a very prolonged epidemic) or as narrow as minutes (e.g., for food poisoning by chemicals that cause symptoms within minutes). Conventionally, the data are displayed as a histogram (which is similar to a bar chart but has no gaps between adjacent columns). Sometimes each case is displayed as a square. The shape and other features of an epidemic curve can suggest hypotheses about the time and source of exposure, the mode of transmission, and the causative agent.

Place:

Describing the occurrence of disease by place provides insight into the geographic extent of the problem and its geographic variation. Characterization by place refers not only to place of residence but to any geographic location relevant to disease occurrence. Such locations include place of diagnosis or report, birthplace, site of employment, school district, hospital unit, or recent travel destinations. The unit may be as large as a continent or country or as small as a street address, hospital wing, or operating room. Sometimes place refers not to a specific location at all but to a place category such as urban or rural, domestic or foreign, and institutional or noninstitutional.

Analyzing data by place can identify communities at increased risk of disease. Even if the data cannot reveal why these people have an increased risk, it can help generate hypotheses to test with additional studies. For example, is a community at increased risk because of characteristics of the people in the community such as genetic susceptibility, lack of immunity, risky behaviors, or exposure to local toxins or contaminated food? Can the increased risk, particularly of a communicable disease, be attributed to characteristics of the causative agent such as a particularly virulent strain, hospitable breeding sites, or availability of the vector that transmits the organism to humans? Or can the increased risk be attributed to the environment that brings the agent and the host together, such as crowding in urban areas that increases the risk of disease transmission from person to person, or more homes being built in wooded areas close to deer that carry ticks infected with the organism that causes Lyme disease

Person:

Because personal characteristics may affect illness, organization and analysis of data by "person" may use inherent characteristics of people (for example, age, sex, race), biologic characteristics (immune status), acquired characteristics (marital status), activities (occupation, leisure activities, use of medications/tobacco/drugs), or the conditions under which they live (socioeconomic status, access to medical care). Age and sex are included in almost all data sets and are the two most commonly analyzed "person" characteristics. However, depending on the disease and the data available, analyses of other person variables are usually necessary. Usually epidemiologists begin the analysis of person data by looking at each variable separately. Sometimes, two variables such as age and sex can be examined simultaneously. Person data are usually displayed in tables or graphs.

Age. Age is probably the single most important "person" attribute, because almost every health-related event varies with age. A number of factors that also vary with age include: susceptibility, opportunity for exposure, latency or incubation period of the disease, and physiologic response (which affects, among other things, disease development). When analyzing data by age, epidemiologists try to use age groups that are narrow enough to detect any age-related patterns that may be present in the data. For some diseases, particularly chronic diseases, 10-year age groups may be adequate. For other diseases, 10-year and even 5-year age groups conceal important variations in disease occurrence by age. Consider the graph of pertussis occurrence by standard 5-year age groups.

Ethnic and racial groups. Sometimes epidemiologists are interested in analyzing person data by biologic, cultural or social groupings such as race, nationality, religion, or social groups such as tribes and other geographically or socially isolated groups. Differences in racial, ethnic, or other group variables may reflect differences in susceptibility or exposure, or differences in other factors that influence the risk of disease, such as socioeconomic status and access to health care. Infant mortality rates for 2002 are shown by race and Hispanic origin of the mother.

Socioeconomic status. Socioeconomic status is difficult to quantify. It is made up of many variables such as occupation, family income, educational achievement or census track, living conditions, and social standing. The variables that are easiest to measure may not accurately reflect the overall concept. Nevertheless, epidemiologists commonly use occupation, family income, and educational achievement, while recognizing that these variables do not measure socioeconomic status precisely. The frequency of many adverse health conditions increases with decreasing socioeconomic status. For example, tuberculosis is more common among persons in lower socioeconomic strata. Infant mortality and time lost from work due to disability are both associated with lower income.

These patterns may reflect more harmful exposures, lower resistance, and less access to health care. Or they may in part reflect an interdependent relationship that is impossible to untangle: Does low socioeconomic status contribute to disability, or does disability contribute to lower socioeconomic status, or both? What accounts for the disproportionate prevalence of diabetes and asthma in lower socioeconomic areas? A few adverse health conditions occur more frequently among persons of higher socioeconomic status. Gout was known as the “disease of kings” because of its association with consumption of rich foods. Other conditions associated with higher socioeconomic status include breast cancer, Kawasaki syndrome, chronic fatigue syndrome, and tennis elbow. Differences in exposure account for at least some if not most of the differences in the frequency of these conditions.

Analytic Epidemiology

As noted earlier, descriptive epidemiology can identify patterns among cases and in populations by time, place and person. From these observations, epidemiologists develop hypotheses about the causes of these patterns and about the factors that increase risk of disease. In other words, epidemiologists can use descriptive epidemiology to generate hypotheses, but only rarely to test those hypotheses. For that, epidemiologists must turn to analytic epidemiology.

The key feature of analytic epidemiology is a comparison group. Consider a large outbreak of hepatitis A that occurred in Pennsylvania in 2003.38 Investigators found almost all of the casepatients had eaten at a particular restaurant during the 2–6 weeks (i.e., the typical incubation period for hepatitis A) before onset of illness. While the investigators were able to narrow down their hypotheses to the restaurant and were able to exclude the food preparers and servers as the source, they did not know which particular food may have been contaminated. The investigators asked the case-patients which restaurant foods they had eaten, but that only indicated which foods were popular.

The investigators, therefore, also enrolled and interviewed a comparison or control group — a group of persons who had eaten at the restaurant during the same period but who did not get sick. Of 133 items on the restaurant’s menu, the most striking difference between the case and control groups was in the proportion that ate salsa (94% of case-patients ate, compared with 39% of controls). Further investigation of the ingredients in the salsa implicated green onions as the source of infection. Shortly thereafter, the Food and Drug Administration issued an advisory to the public about green onions and risk of hepatitis A. This action was in direct response to the convincing results of the analytic epidemiology, which compared the exposure history of case-patients with that of an appropriate comparison group. When investigators find that persons with a particular characteristic are more likely than those without the characteristic to contract a disease, the characteristic is said to be associated with the disease. The characteristic may be a:

- Demographic factor such as age, race, or sex;
- Constitutional factor such as blood group or immune status;

• Behavior or act such as smoking or having eaten salsa; or
• Circumstance such as living near a toxic waste site.

Identifying factors associated with disease help health officials appropriately target public health prevention and control activities. It also guides additional research into the causes of disease.
Thus, analytic epidemiology is concerned with the search for causes and effects, or the why and the how. Epidemiologists use analytic epidemiology to quantify the association between exposures and outcomes and to test hypotheses about causal relationships. It has been said that epidemiology by itself can never prove that a particular exposure caused a particular outcome. Often, however, epidemiology provides sufficient evidence to take appropriate control and prevention measures.
Epidemiologic studies fall into two categories: **experimental** and **observational**.

Experimental studies:
In an experimental study, the investigator determines through a controlled process the exposure for each individual (clinical trial) or community (community trial), and then tracks the individuals or communities over time to detect the effects of the exposure. For example, in a clinical trial of a new vaccine, the investigator may randomly assign some of the participants to receive the new vaccine, while others receive a placebo shot. The investigator then tracks all participants, observes who gets the disease that the new vaccine is intended to prevent, and compares the two groups (new vaccine vs. placebo) to see whether the vaccine group has a lower rate of disease. Similarly, in a trial to prevent onset of diabete among high-risk individuals, investigators randomly assigned enrollees to one of three groups — placebo, an anti-diabetes drug, or lifestyle intervention. At the end of the follow-up period, investigators found the lowest incidence of diabetes in the lifestyle intervention group, the next lowest in the anti-diabetic drug group, and the highest in the placebo group.

Observational studies:
In an observational study, the epidemiologist simply observes the exposure and disease status of each study participant. John Snow's studies of cholera in London were observational studies. The two most common types of observational studies are cohort studies and case-control studies; a third type is cross-sectional studies.

Cohort study. A cohort study is similar in concept to the experimental study. In a cohort study the epidemiologist records whether each study participant is exposed or not, and then tracks the participants to see if they develop the disease of interest. Note that this differs from an experimental study because, in a cohort study, the investigator observes rather than determines the participants' exposure status. After a period of time, the investigator compares the disease rate in the exposed group with the disease rate in the unexposed group. The unexposed group serves as the comparison group, providing an estimate of the baseline or expected amount of disease occurrence in the community. If the disease rate is substantively different in the exposed group compared to the unexposed group, the exposure is said to be associated with illness.

The length of follow-up varies considerably. In an attempt to respond quickly to a public health concern such as an outbreak, public health departments tend to conduct relatively brief studies.
On the other hand, research and academic organizations are more likely to conduct studies of cancer, cardiovascular disease, and other chronic diseases which may last for years and even decades.
The Framingham study is a well-known cohort study that has followed over 5,000 residents of Framingham, Massachusetts, since the early 1950s to establish the rates and risk factors for heart disease. The Nurses Health Study and the Nurses Health Study II are cohort studies established in 1976 and 1989, respectively, that have followed over 100,000 nurses each and have provided useful information on oral contraceptives, diet, and lifestyle risk factors. These studies are sometimes called

follow-up or **prospective** cohort studies, because participants are enrolled as the study begins and are then followed prospectively over time to identify occurrence of the outcomes of interest.

An alternative type of cohort study is a **retrospective** cohort study. In this type of study both the exposure and the outcomes have already occurred. Just as in a prospective cohort study, the investigator calculates and compares rates of disease in the exposed and unexposed groups. Retrospective cohort studies are commonly used in investigations of disease in groups of easily identified people such as workers at a particular factory or attendees at a wedding. For example, a retrospective cohort study was used to determine the source of infection of cyclosporiasis, a parasitic disease that caused an outbreak among members of a residential facility in Pennsylvania in 2004. The investigation indicated that consumption of snow peas was implicated as the vehicle of the cyclosporiasis outbreak.

Case-control study. In a case-control study, investigators start by enrolling a group of people with disease (at CDC such persons are called case-patients rather than cases, because case refers to occurrence of disease, not a person). As a comparison group, the investigator then enrolls a group of people without disease (controls). Investigators then compare previous exposures between the two groups. The control group provides an estimate of the baseline or expected amount of exposure in that population. If the amount of exposure among the case group is substantially higher than the amount you would expect based on the control group, then illness is said to be associated with that exposure.

The study of hepatitis A traced to green onions, described above, is an example of a case-control study. The key in a case-control study is to identify an appropriate control group, comparable to the case group in most respects, in order to provide a reasonable estimate of the baseline or expected exposure.

Cross-sectional study. In this third type of observational study, a sample of persons from a population is enrolled and their exposures and health outcomes are measured simultaneously. The cross-sectional study tends to assess the presence (prevalence) of the health outcome at that point of time without regard to duration. For example, in a cross-sectional study of diabetes, some of the enrollees with diabetes may have lived with their diabetes for many years, while others may have been recently diagnosed. From an analytic viewpoint the cross-sectional study is weaker than either a cohort or a case-control study because a crosssectional study usually cannot disentangle risk factors for occurrence of disease (incidence) from risk factors for survival with the disease. On the other hand, a cross-sectional study is a perfectly fine tool for descriptive epidemiology purposes. Crosssectional studies are used routinely to document the prevalence in a community of health behaviors (prevalence of smoking), health states (prevalence of vaccination against measles), and health outcomes, particularly chronic conditions (hypertension, diabetes).

In summary, the purpose of an analytic study in epidemiology is to identify and quantify the relationship between an exposure and a health outcome. The hallmark of such a study is the presence of at least two groups, one of which serves as a comparison group. In an experimental study, the investigator determines the exposure for the study subjects; in an observational study, the subjects are exposed under more natural conditions. In an observational cohort study, subjects are enrolled or grouped on the basis of their exposure, then are followed to document occurrence of disease. Differences in disease rates between the exposed and unexposed groups lead investigators to conclude that exposure is associated with disease. In an observational case-control study, subjects are enrolled according to whether they have the disease or not, then are questioned or tested to determine their prior exposure. Differences in exposure prevalence between the case and control groups allow investigators to conclude that the exposure is associated with the disease. Cross-sectional studies measure exposure and disease status at the same time, and are better suited to descriptive epidemiology than causation.

Concepts of Disease Occurrence

A critical premise of epidemiology is that disease and other health events do not occur randomly in a population, but are more likely to occur in some members of the population than others because of risk factors that may not be distributed randomly in the population. As noted earlier, one important use of epidemiology is to identify the factors that place some members at greater risk than others.

Causation:
A number of models of disease causation have been proposed. Among the simplest of these is the epidemiologic triad or triangle, the traditional model for infectious disease. The triad consists of an external **agent**, a susceptible **host**, and an **environment** that brings the host and agent together. In this model, disease results from the interaction between the agent and the susceptible host in an environment that supports transmission of the agent from a source to that host. Two ways of depicting this model are shown in Agent, host, and environmental factors interrelate in a variety of complex ways to produce disease. Different diseases require different balances and interactions of these three components. Development of appropriate, practical, and effective public health measures to control or prevent disease usually requires assessment of all three components and their interactions.

Agent originally referred to an infectious microorganism or pathogen: a virus, bacterium, parasite, or other microbe. Generally, the agent must be present for disease to occur; however, presence of that agent alone is not always sufficient to cause disease. A variety of factors influence whether exposure to an organism will result in disease, including the organism's pathogenicity (ability to cause disease) and dose. Over time, the concept of agent has been broadened to include chemical and physical causes of disease or injury. These include chemical contaminants (such as the L-tryptophan contaminant responsible for eosinophilia-myalgia syndrome), as well as physical forces (such as repetitive mechanical forces associated with carpal tunnel syndrome). While the epidemiologic triad serves as a useful model for many diseases, it has proven inadequate for cardiovascular disease, cancer, and other diseases that appear to have multiple contributing causes without a single necessary one.

Host refers to the human who can get the disease. A variety of factors intrinsic to the host, sometimes called risk factors, can influence an individual's exposure, susceptibility, or response to a causative agent. Opportunities for exposure are often influenced by behaviors such as sexual practices, hygiene, and other personal choices as well as by age and sex. Susceptibility and response to an agent are influenced by factors such as genetic composition, nutritional and immunologic status, anatomic structure, presence of disease or medications, and psychological makeup.

Environment refers to extrinsic factors that affect the agent and the opportunity for exposure. Environmental factors include physical factors such as geology and climate, biologic factors such as insects that transmit the agent, and socioeconomic factors such as crowding, sanitation, and the availability of health services.

Component causes and causal pies:
Because the agent-host-environment model did not work well for many non-infectious diseases, several other models that attempt to account for the multifactorial nature of causation have been proposed. One such model was proposed by Rothman in 1976, and has come to be known as the Causal Pies. An individual factor that contributes to cause disease is shown as a piece of a pie. After all the pieces of a pie fall into place, the pie is complete — and disease occurs. The individual factors are called **component causes**. The complete pie, which might be considered a causal pathway, is called a **sufficient cause**. A disease may have more than one sufficient cause, with each sufficient cause being composed of several component causes that may or may not overlap. A component that appears in every pie or pathway is called a **necessary cause**, because without it, disease does not occur.

The component causes may include intrinsic host factors as well as the agent and the environmental factors of the agent-hostenvironment triad. A single component cause is rarely a sufficient cause by itself. For example, even exposure to a highly infectious agent such as measles virus does not invariably result in measles disease. Host susceptibility and other host factors also may play a role. At the other extreme, an agent that is usually harmless in healthy persons may cause devastating disease under different conditions.

Pneumocystis carinii is an organism that harmlessly colonizes the respiratory tract of some healthy persons, but can cause potentially lethal pneumonia in persons whose immune systems have been weakened by human immunodeficiency virus (HIV). Presence of *Pneumocystis carinii* organisms is therefore a necessary but not sufficient cause of pneumocystis pneumonia. As the model indicates, a particular disease may result from a variety of different sufficient causes or pathways. For example, lung cancer may result from a sufficient cause that includes smoking as a component cause. Smoking is not a sufficient cause by itself, however, because not all smokers develop lung cancer.

Neither is smoking a necessary cause, because a small fraction of lung cancer victims have never smoked. Suppose Component Cause B is smoking and Component Cause C is asbestos. Sufficient Cause I includes both smoking (B) and asbestos (C). Sufficient Cause II includes asbestos without smoking, and Sufficient Cause C includes smoking without asbestos. But because lung cancer can develop in persons who have never been exposed to either smoking or asbestos, a proper model for lung cancer would have to show at least one more Sufficient Cause Pie that does not include either component B or component C.

Note that public health action does not depend on the identification of every component cause. Disease prevention can be accomplished by blocking any single component of a sufficient cause, at least through that pathway. For example, elimination of smoking (component B) would prevent lung cancer from sufficient causes I and II, although some lung cancer would still occur through sufficient cause III.

Natural History and Spectrum of Disease

Natural history of disease refers to the progression of a disease process in an individual over time, in the absence of treatment. For example, untreated infection with HIV causes a spectrum of clinical problems beginning at the time of seroconversion (primary HIV) and terminating with AIDS and usually death. It is now recognized that it may take 10 years or more for AIDS to develop after seroconversion.43 Many, if not most, diseases have a characteristic natural history, although the time frame and specific manifestations of disease may vary from individual to individual and are influenced by preventive and therapeutic measures.

The process begins with the appropriate exposure to or accumulation of factors sufficient for the disease process to begin in a susceptible host. For an infectious disease, the exposure is a microorganism. For cancer, the exposure may be a factor that initiates the process, such as asbestos fibers or components in tobacco smoke (for lung cancer), or one that promotes the process, such as estrogen (for endometrial cancer). After the disease process has been triggered, pathological changes then occur without the individual being aware of them. This stage of subclinical disease, extending from the time of exposure to onset of disease symptoms, is usually called the **incubation period** for infectious diseases, and the **latency period** for chronic diseases. During this stage, disease is said to be asymptomatic (no symptoms) or inapparent. This period may be as brief as seconds for hypersensitivity and toxic reactions to as long as decades for certain chronic diseases. Even for a single disease, the characteristic incubation period has a range. For example, the typical incubation period for hepatitis A is as long as 7 weeks. The latency period for leukemia to become evident among survivors of the atomic bomb blast in Hiroshima ranged from 2 to 12 years, peaking at 6-7 years.44 Incubation periods of selected exposures and diseases varying from minutes to decades are displayed in

Although disease is not apparent during the incubation period, some pathologic changes may be detectable with laboratory, radiographic, or other screening methods. Most screening programs attempt to identify the disease process during this phase of its natural history, since intervention at this early stage is likely to be more effective than treatment given after the disease has progressed and become symptomatic. The onset of symptoms marks the transition from subclinical to clinical disease. Most diagnoses are made during the stage of clinical disease. In some people, however, the disease process may never progress to clinically apparent illness. In others, the disease process may result in illness that ranges from mild to severe or fatal. This range is called the **spectrum of disease**. Ultimately, the disease process ends either in recovery, disability or death.

For an infectious agent, **infectivity** refers to the proportion of exposed persons who become infected. **Pathogenicity** refers to the proportion of infected individuals who develop clinically apparent disease. **Virulence** refers to the proportion of clinically apparent cases that are severe or fatal. Because the spectrum of disease can include asymptomatic and mild cases, the cases of illness diagnosed by clinicians in the community often represent only the tip of the iceberg. Many additional cases may be too early to diagnose or may never progress to the clinical stage. Unfortunately, persons with inapparent or undiagnosed infections may nonetheless be able to transmit infection to others. Such persons who are infectious but have subclinical disease are called **carriers**. Frequently, carriers are persons with incubating disease or inapparent infection. Persons with measles, hepatitis A, and several other diseases become infectious a few days before the onset of symptoms. However carriers may also be persons who appear to have recovered from their clinical illness but remain infectious, such as chronic carriers of hepatitis B virus, or persons who never exhibited symptoms. The challenge to public health workers is that these carriers, unaware that they are infected and infectious to others, are sometimes more likely to unwittingly spread infection than are people with obvious illness.

Chain of Infection

As described above, the traditional epidemiologic triad model holds that infectious diseases result from the interaction of agent, host, and environment. More specifically, transmission occurs when the agent leaves its **reservoir** or host through a **portal of exit**, is conveyed by some **mode of transmission**, and enters through an appropriate **portal of entry** to infect a **susceptible host**. This sequence is sometimes called the chain of infection.

Portal of exit:

Portal of exit is the path by which a pathogen leaves its host. The portal of exit usually corresponds to the site where the pathogen is localized. For example, influenza viruses and *Mycobacterium tuberculosis* exit the respiratory tract, schistosomes through urine, cholera vibrios in feces, *Sarcoptes scabiei* in scabies skin lesions, and enterovirus 70, a cause of hemorrhagic conjunctivitis, in conjunctival secretions. Some bloodborne agents can exit by crossing the placenta from mother to fetus (rubella, syphilis, toxoplasmosis), while others exit through cuts or needles in the skin (hepatitis B) or blood-sucking arthropods (malaria).

Modes of transmission:

An infectious agent may be transmitted from its natural reservoir to a susceptible host in different ways. There are different classifications for modes of transmission. Here is one classification:

- Direct (Direct contact and Droplet spread)
- Indirect (Airborne and Vehicleborne Vectorborne (mechanical or biologic))

In **direct transmission**, an infectious agent is transferred from a reservoir to a susceptible host by direct contact or droplet spread.

Direct contact occurs through skin-to-skin contact, kissing, and sexual intercourse. Direct contact also refers to contact with soil or vegetation harboring infectious organisms. Thus, infectious mononucleosis ("kissing disease") and gonorrhea are spread from person to person by direct contact. Hookworm is spread by direct contact with contaminated soil.

Droplet spread *r*efers to spray with relatively large, short-range aerosols produced by sneezing, coughing, or even talking. Droplet spread is classified as direct because transmission is by direct spray over a few feet, before the droplets fall to the ground. Pertussis and meningococcal infection are examples of diseases transmitted from an infectious patient to a susceptible host by droplet spread.

Indirect transmission refers to the transfer of an infectious agent from a reservoir to a host by suspended air particles, inanimate objects (vehicles), or animate intermediaries (vectors).

Airborne transmission occurs when infectious agents are carried by dust or droplet nuclei suspended in air. Airborne dust includes material that has settled on surfaces and become resuspended by air currents as well as infectious particles blown from the soil by the wind. Droplet nuclei are dried residue of less than 5 microns in size. In contrast to droplets that fall to the ground within a few feet, droplet nuclei may remain suspended in the air for long periods of time and may be blown over great distances. Measles, for example, has occurred in children who came into a physician's office after a child with measles had left, because the measles virus remained suspended in the air.

Vehicles that may indirectly transmit an infectious agent include food, water, biologic products (blood), and fomites (inanimate objects such as handkerchiefs, bedding, or surgical scalpels). A vehicle may passively carry a pathogen — as food or water may carry hepatitis A virus. Alternatively, the vehicle may provide an environment in which the agent grows, multiplies, or produces toxin — as improperly canned foods provide an environment that supports production of botulinum toxin by *Clostridium botulinum*.

Vectors such as mosquitoes, fleas, and ticks may carry an infectious agent through purely mechanical means or may support growth or changes in the agent. Examples of mechanical transmission are flies carrying *Shigella* on their appendages and fleas carrying *Yersinia pestis*, the causative agent of plague, in their gut. In contrast, in biologic transmission, the causative agent of malaria or guinea worm disease undergoes maturation in an intermediate host before it can be transmitted to humans.

Portal of entry:

The portal of entry refers to the manner in which a pathogen enters a susceptible host. The portal of entry must provide access to tissues in which the pathogen can multiply or a toxin can act. Often, infectious agents use the same portal to enter a new host that they used to exit the source host. For example, influenza virus exits the respiratory tract of the source host and enters the respiratory tract of the new host. In contrast, many pathogens that cause gastroenteritis follow a so-called "fecal-oral" route because they exit the source host in feces, are carried on inadequately washed hands to a vehicle such as food, water, or utensil, and enter a new host through the mouth. Other portals of entry include the skin (hookworm), mucous membranes (syphilis), and blood (hepatitis B, human immunodeficiency virus).

Host:

The final link in the chain of infection is a susceptible host. Susceptibility of a host depends on genetic or constitutional factors, specific immunity, and nonspecific factors that affect an individual's ability to resist infection or to limit pathogenicity. An individual's genetic makeup may either increase or decrease susceptibility. For example, persons with sickle cell trait seem to be at least partially protected from a particular type of malaria. Specific immunity refers to protective antibodies that are directed against a specific agent. Such antibodies may develop in responsemto infection, vaccine, or toxoid (toxin that has been deactivated but retains its capacity to stimulate production of toxin antibodies) or may be acquired by transplacental transfer from mother to fetus or by injection of antitoxin or immune globulin. Nonspecific factors that defend against infection include the skin, mucous membranes, gastric acidity, cilia in the respiratory tract, the cough reflex, and nonspecific immune response. Factors that may increase susceptibility to infection by disrupting host defenses include malnutrition, alcoholism, and disease or therapy that impairs the nonspecific immune response.

Implications for public health:
Knowledge of the portals of exit and entry and modes of transmission provides a basis for determining appropriate control measures. In general, control measures are usually directed against the segment in the infection chain that is most susceptible to intervention, unless practical issues dictate otherwise. For some diseases, the most appropriate intervention may be directed at controlling or eliminating the agent at its source. A patient sick with a communicable disease may be treated with antibiotics to eliminate the infection. An asymptomatic but infected person may be treated both to clear the infection and to reduce the risk of transmission to others. In the community, soil may be decontaminated or covered to prevent escape of the agent.

Some interventions are directed at the mode of transmission:
Interruption of direct transmission may be accomplished by isolation of someone with infection, or counseling persons to avoid the specific type of contact associated with transmission. Vehicleborne transmission may be interrupted by elimination or decontamination of the vehicle. To prevent fecal-oral transmission, efforts often focus on rearranging the environment to reduce the risk of contamination in the future and on changing behaviors, such as promoting handwashing. For airborne diseases, strategies may be directed at modifying ventilation or air pressure, and filtering or treating the air. To interrupt vectorborne transmission, measures may be directed toward controlling the vector population, such as spraying to reduce the mosquito population.

Some strategies that protect portals of entry are simple and effective. For example, bed nets are used to protect sleeping persons from being bitten by mosquitoes that may transmit malaria. A dentist’s mask and gloves are intended to protect the dentist from a patient’s blood, secretions, and droplets, as well to protect the patient from the dentist. Wearing of long pants and sleeves and use of insect repellent are recommended to reduce the risk of Lyme disease and West Nile virus infection, which are transmitted by the bite of ticks and mosquitoes, respectively.

Some interventions aim to increase a host’s defenses. Vaccinations promote development of specific antibodies that protect against infection. On the other hand, prophylactic use of antimalarial drugs, recommended for visitors to malaria-endemic areas, does not prevent exposure through mosquito bites, but does prevent infection from taking root.

Finally, some interventions attempt to prevent a pathogen from encountering a susceptible host. The concept of **herd immunity** suggests that if a high enough proportion of individuals in a population are resistant to an agent, then those few who are susceptible will be protected by the resistant majority, since the pathogen will be unlikely to “find” those few susceptible individuals. The degree of herd immunity necessary to prevent or interrupt an outbreak varies by disease. In theory, herd immunity means that not everyone in a community needs to be resistant (immune) to prevent disease spread and occurrence of an outbreak. In practice, herd immunity has not prevented outbreaks of measles and rubella in populations with immunization levels as high as 85% to 90%. One problem is that, in highly immunized populations, the relatively few susceptible persons are often clustered in subgroups defined by socioeconomic or cultural factors. If the pathogen is introduced into one of these subgroups, an outbreak may occur.

Epidemic Disease Occurrence

Level of disease

The amount of a particular disease that is usually present in a community is referred to as the baseline or **endemic** level of the disease. This level is not necessarily the desired level, which may in fact be zero, but rather is the observed level. In the absence of intervention and assuming that the level is not high enough to deplete the pool of susceptible persons, the disease may continue to occur at this level indefinitely. Thus, the baseline level is often regarded as the expected level of the disease. While some diseases are so rare in a given population that a single case warrants an epidemiologic investigation (e.g., rabies, plague, polio), other diseases occur more commonly so that only deviations from the norm warrant investigation.

Sporadic refers to a disease that occurs infrequently and irregularly.

Endemic refers to the constant presence and/or usual prevalence of a disease or infectious agent in a population within a geographic area.

Hyperendemic refers to persistent, high levels of disease occurrence. Occasionally, the amount of disease in a community rises above the expected level.

Epidemic refers to an increase, often sudden, in the number of cases of a disease above what is normally expected in that population in that area.

Outbreak carries the same definition of epidemic, but is often used for a more limited geographic area.

Cluster refers to an aggregation of cases grouped in place and time that are suspected to be greater than the number expected, even though the expected number may not be known.

Pandemic refers to an epidemic that has spread over several countries or continents, usually affecting a large number of people.

Epidemics occur when an agent and susceptible hosts are present in adequate numbers, and the agent can be effectively conveyed from a source to the susceptible hosts. More specifically, an epidemic may result from:

- A recent increase in amount or virulence of the agent,
- The recent introduction of the agent into a setting where it has not been before,
- An enhanced mode of transmission so that more susceptible persons are exposed,
- A change in the susceptibility of the host response to the agent, and/or
- Factors that increase host exposure or involve introduction through new portals of entry.

The previous description of epidemics presumes only infectious agents, but non-infectious diseases such as diabetes and obesity exist in epidemic proportion.

Epidemic Patterns:

Epidemics can be classified according to their manner of spread through a population:

- Common-source
- Point
- Continuous
- Intermittent
- Propagated
- Mixed
- Other

A **common-source outbreak** is one in which a group of persons are all exposed to an infectious agent or a toxin from the same source. If the group is exposed over a relatively brief period, so that everyone who becomes ill does so within one incubation period, then the common-source outbreak is further classified as a **pointsource outbreak**. The epidemic of leukemia cases in Hiroshima following the atomic bomb blast and the epidemic of hepatitis A among patrons of the Pennsylvania restaurant who ate green onions each had a point source of exposure.38, 44 If the number of cases during an epidemic were plotted over time, the resulting graph, called an epidemic curve, would typically have a steep upslope and a more gradual downslope (a so-called “log-normal distribution”).

MODULE 4— Statistical Methods for Health Science

Biostatistics involves the theory and application of statistical science to analyze public health problems and to further biomedical research. The faculty includes leaders in the development of statistical methods for clinical trials and observational studies, studies on the environment, and genomics/genetics. The department's research in statistical methods and interdisciplinary collaborations provide many opportunities for student participation.

Current departmental research on statistical and computing methods for observational studies and clinical trials includes survival analysis, missing-data problems, and causal inference. Other areas of investigation are environmental research (methods for longitudinal studies, analyses with incomplete data, and meta-analysis); statistical aspects of the study of AIDS and cancer; quantitative problems in health-risk analysis, technology assessment, and clinical decision making; statistical methodology in psychiatric research and in genetic studies; Bayesian statistics; statistical computing; statistical genetics and computational biology; and collaborative research activities with biomedical scientists in other Harvard-affiliated institutions.

Measures of Risk

This Lesson describes the measures of central location and spread, which are useful for summarizing continuous variables. However, many variables used by field epidemiologists are categorical variables, some of which have only two categories — exposed yes/no, test positive/negative, case/control, and so on. These variables have to be summarized with frequency measures such as ratios, proportions, and rates. Incidence, prevalence, and mortality rates are three frequency measures that are used to characterize the occurrence of health events in a population.

Objectives

At the end of this module, you will be able to:
• Calculate and interpret the following epidemiologic measures:
– Ratio
– Proportion
– Incidence proportion (attack rate)
– Incidence rate
– Prevalence
– Mortality rate
• Choose and apply the appropriate measures of association and measures of public health impact

Frequency Measures

A measure of central location provides a single value that summarizes an entire distribution of data. In contrast, a frequency measure characterizes only part of the distribution. Frequency measures compare one part of the distribution to another part of the distribution, or to the entire distribution. Common frequency measures are **ratios**, **proportions**, and **rates**. All three frequency measures have the same basic form: *numerator denominator x 10n*
Recall that:
100 = 1 (anything raised to the 0 power equals 1)
101 = 10 (anything raised to the 1st power is the value itself)
102 = 10 x 10 = 100
103 = 10 x 10 x 10 = 1,000
So the fraction of (numerator/denominator) can be multiplied by 1, 10, 100, 1000, and so on. This multiplier varies by measure and will be addressed in each section.

Ratio:

A ratio is the relative magnitude of two quantities or a comparison of any two values. It is calculated by dividing one interval- or ratio-scale variable by the other. The numerator and denominator need not be related. Therefore, one could compare apples with oranges or apples with number of physician visits.

Method for calculating a ratio

Number or rate of events, items, persons, etc. in one group

Number or rate of events, items, persons, etc. in another group

After the numerator is divided by the denominator, the result is often expressed as the result "to one" or written as the result ":1."

Note that in certain ratios, the numerator and denominator are different categories of the same variable, such as males and females, or persons 20–29 years and 30–39 years of age. In other ratios, the numerator and denominator are completely different variables, such as the number of hospitals in a city and the size of the population living in that city.

EXAMPLE: Calculating a Ratio — Different Categories of Same Variable

Between 1971 and 1975, as part of the National Health and Nutrition Examination Survey (NHANES), 7,381 persons ages 40–77 years were enrolled in a follow-up study.1 At the time of enrollment, each study participant was classified as having or not having diabetes. During 1982–1984, enrollees were documented either to have died or were still alive. The results are summarized as follows.

Original Enrollment Dead at Follow-Up

(1971–1975) (1982–1984)

Diabetic men 189 100

Nondiabetic men 3,151 811

Diabetic women 218 72

Nondiabetic women 3,823 511

Of the men enrolled in the NHANES follow-up study, 3,151 were nondiabetic and 189 were diabetic. Calculate the ratio of non-diabetic to diabetic men.

Ratio = 3,151 / 189 x 1 = 16.7:1

Properties and uses of ratios:

• Ratios are common descriptive measures, used in all fields. In epidemiology, ratios are used as both descriptive measures and as analytic tools. As a descriptive measure, ratios can describe the male-to-female ratio of participants in a study, or the ratio of controls to cases (e.g., two controls per case). As an analytic tool, ratios can be calculated for occurrence of illness, injury, or death between two groups. These ratio measures, including risk ratio (relative risk), rate ratio, and odds ratio, are described later in this lesson.

• As noted previously, the numerators and denominators of a ratio can be related or unrelated. In other words, you are free to use a ratio to compare the number of males in a population with the number of females, or to compare the number of residents in a population with the number of hospitals or dollars spent on over-the-counter medicines.

• Usually, the values of both the numerator and denominator of a ratio are divided by the value of one or the other so that either the numerator or the denominator equals 1.0. So the ratio of non-diabetics to diabetics cited in the previous example is more likely to be reported as 16.7:1 than 3,151:189.

EXAMPLES: Calculating Ratios for Different Variables

Example A: A city of 4,000,000 persons has 500 clinics. Calculate the ratio of clinics per person.

500 / 4,000,000 x 10n = 0.000125 clinics per person

To get a more easily understood result, you could set 10n = 104 = 10,000. Then the ratio becomes:

0.000125 x 10,000 = 1.25 clinics per 10,000 persons

You could also divide each value by 1.25, and express this ratio as 1 clinic for every 8,000 persons.
Example B: Delaware's infant mortality rate in 2001 was 10.7 per 1,000 live births.2 New Hampshire's infant mortality rate in 2001 was 3.8 per 1,000 live births. Calculate the ratio of the infant mortality rate in Delaware to that in New Hampshire.
10.7 / 3.8 x 1 = 2.8:1
Thus, Delaware's infant mortality rate was 2.8 times as high as New Hampshire's infant mortality rate in 2001.

A commonly used epidemiologic ratio: death-to-case ratio:
Death-to-case ratio is the number of deaths attributed to a particular disease during a specified period divided by the number of new cases of that disease identified during the same period. It is used as a measure of the severity of illness: the death-to-case ratio for rabies is close to 1 (that is, almost everyone who develops rabies dies from it), whereas the death-to-case ratio for the common cold is close to 0.
For example, in the United States in 2002, a total of 15,075 new cases of tuberculosis were reported.3 During the same year, 802 deaths were attributed to tuberculosis. The tuberculosis death-tocase ratio for 2002 can be calculated as 802 / 15,075. Dividing both numerator and denominator by the numerator yields 1 death per 18.8 new cases. Dividing both numerator and denominator by the denominator (and multiplying by 10n = 100) yields 5.3 deaths per 100 new cases. Both expressions are correct.
Note that, presumably, many of those who died had initially contracted tuberculosis years earlier. Thus many of the 802 in the numerator are not among the 15,075 in the denominator. Therefore, the death-to-case ratio is a ratio, but not a proportion.

Proportion:
A proportion is the comparison of a part to the whole. It is a type of ratio in which the numerator is included in the denominator. You might use a proportion to describe what fraction of clinic patients tested positive for HIV, or what percentage of the population is younger than 25 years of age. A proportion may be expressed as a decimal, a fraction, or a percentage.
Method for calculating a proportion
Number of persons or events with a particular characteristic
Total number of persons or events, of which the numerator is a subset x 10n
For a proportion, 10n is usually 100 (or n=2) and is often expressed as a percentage.

EXAMPLE: Calculating a Proportion
Example A: Calculate the proportion of men in the NHANES follow-up study who were diabetics.
Numerator = 189 diabetic men
Denominator = Total number of men = 189 + 3,151 = 3,340
Proportion = (189 / 3,340) x 100 = 5.66%
Example B: Calculate the proportion of deaths among men.
Numerator = deaths in men
= 100 deaths in diabetic men + 811 deaths in nondiabetic men
= 911 deaths in men
Notice that the numerator (911 deaths in men) is a subset of the denominator.
Denominator = all deaths
= 911 deaths in men + 72 deaths in diabetic women + 511 deaths in nondiabetic women
= 1,494 deaths
Proportion = 911 / 1,494 = 60.98% = 61%
Your Turn: What proportion of all study participants were men? (Answer = 45.25%)

Properties and uses of proportions:

• Proportions are common descriptive measures used in all fields. In epidemiology, proportions are used most often as descriptive measures. For example, one could calculate the proportion of persons enrolled in a study among all those eligible (“participation rate”), the proportion of children in a village vaccinated against measles, or the proportion of persons who developed illness among all passengers of a cruise ship.
• Proportions are also used to describe the amount of disease that can be attributed to a particular exposure. For example, on the basis of studies of smoking and lung cancer, public health officials have estimated that greater than 90% of the lung cancer cases that occur are attributable to cigarette smoking.
• In a proportion, the numerator must be included in the denominator. Thus, the number of apples divided by the number of oranges is not a proportion, but the number of apples divided by the total number of fruits of all kinds is a proportion. Remember, the numerator is always a subset of the denominator.
• A proportion can be expressed as a fraction, a decimal, or a percentage. The statements “one fifth of the residents became ill” and “twenty percent of the residents became ill” are equivalent.
• Proportions can easily be converted to ratios. If the numerator is the number of women (179) who attended a clinic and the denominator is all the clinic attendees (341), the proportion of clinic attendees who are women is 179 / 341, or 52% (a little more than half). To convert to a ratio, subtract the numerator from the denominator to get the number of clinic patients who are not women, i.e., the number of men (341 – 179 = 162 men.)Thus, ratio of women to men could be calculated from the proportion as:

Ratio = 179 / (341 – 179) x 1
= 179 / 162
= 1.1 to 1 female-to-male ratio

Conversely, if a ratio’s numerator and denominator together make up a whole population, the ratio can be converted to a proportion. You would add the ratio’s numerator and denominator to form the denominator of the proportion, as illustrated in the NHANES follow-up study examples (provided earlier in this lesson).

A specific type of epidemiologic proportion: proportionate mortality:

Proportionate mortality is the proportion of deaths in a specified population during a period of time that are attributable to different causes. Each cause is expressed as a percentage of all deaths, and the sum of the causes adds up to 100%. These proportions are not rates because the denominator is all deaths, not the size of the population in which the deaths occurred. Table 3.1 lists the primary causes of death in the United States in 2003 for persons of all ages and for persons aged 25–44 years, by number of deaths, proportionate mortality, and rank.

Rate:

In epidemiology, a rate is a measure of the frequency with which an event occurs in a defined population over a specified period of time. Because rates put disease frequency in the perspective of the size of the population, rates are particularly useful for comparing disease frequency in different locations, at different times, or among different groups of persons with potentially different sized populations; that is, a rate is a measure of risk.

To a non-epidemiologist, rate means how fast something is happening or going. The speedometer of a car indicates the car’s speed or rate of travel in miles or kilometers per hour. This rate is always reported per some unit of time. Some epidemiologists restrict use of the term rate to similar measures that are expressed per unit of time. For these epidemiologists, a rate describes how quickly disease occurs in a population, for example, 70 new cases of breast cancer per 1,000 women per year. This measure conveys a sense of the speed with which disease occurs in a population, and seems to imply that this pattern has occurred and will continue to occur for the foreseeable future. This rate is an ***incidence rate***, described in the next section.

Other epidemiologists use the term rate more loosely, referring to proportions with case counts in the numerator and size of population in the denominator as rates. Thus, an **attack rate** is the proportion of the population that develops illness during an outbreak. For example, 20 of 130 persons developed diarrhea after attending a picnic. (An alternative and more accurate phrase for attack rate is **incidence proportion**.) A **prevalence rate** is the proportion of the population that has a health condition at a point in time. For example, 70 influenza case-patients in March 2005 reported in County A. A **case-fatality rate** is the proportion of persons with the disease who die from it. Forexample, one death due to meningitis among County A's population. All of these measures are proportions, and none is expressed per units of time. Therefore, these measures are not considered "true" rates by some, although use of the terminology is widespread.

Incidence refers to the occurrence of new cases of disease or injury in a population over a specified period of time. Although some epidemiologists use incidence to mean the number of new cases in a community, others use incidence to mean the number of new cases per unit of population. Two types of incidence are commonly used — **incidence proportion** and **incidence rate**.

Incidence proportion or risk:
Incidence proportion is the proportion of an initially disease-free population that develops disease, becomes injured, or dies during a specified (usually limited) period of time. Synonyms include attack rate, risk, probability of getting disease, and cumulative incidence. Incidence proportion is a proportion because the persons in the numerator, those who develop disease, are all included in the denominator (the entire population).
Method for calculating incidence proportion (risk)
Number of new cases of disease or injury during specified period
Size of population at start of period

EXAMPLES: Calculating Incidence Proportion (Risk)
Example A: In the study of diabetics, 100 of the 189 diabetic men died during the 13-year follow-up period. Calculate the risk of death for these men.
Numerator = 100 deaths among the diabetic men
Denominator = 189 diabetic men
10n = 102 = 100
Risk = (100 / 189) x 100 = 52.9%
Example B: In an outbreak of gastroenteritis among attendees of a corporate picnic, 99 persons ate potato salad, 30 of whom developed gastroenteritis. Calculate the risk of illness among persons who ate potato salad.
Numerator = 30 persons who ate potato salad and developed gastroenteritis
Denominator = 99 persons who ate potato salad
10n = 102 = 100
Risk = "Food-specific attack rate" = (30 / 99) x 100 = 0.303 x 100 = 30.3%

Properties and uses of incidence proportions:
• Incidence proportion is a measure of the risk of disease or the probability of developing the disease during the specified period. As a measure of incidence, it includes only new cases of disease in the numerator. The denominator is the number of persons in the population at the start of the observation period. Because all of the persons with new cases of disease (numerator) are also represented in the denominator, a risk is also a proportion.

• In the outbreak setting, the term **attack rate** is often used as a synonym for risk. It is the risk of getting the disease during a specified period, such as the duration of an outbreak. A variety of attack rates can be calculated.

Overall attack rate is the total number of new cases divided by the total population.
A **food-specific attack rate** is the number of persons who ate a specified food and became ill divided by the total number of persons who ate that food, as illustrated in the previous potato salad example.
A **secondary attack rate** is sometimes calculated to document the difference between community transmission of illness versus transmission of illness in a household, barracks, or other closed population. It is calculated as:
Number of cases among contacts of primary cases
Total number of contacts x 10n
Often, the total number of contacts in the denominator is calculated as the total population in the households of the primary cases, minus the number of primary cases. For a secondary attack rate, 10n usually is 100%.

EXAMPLE: Calculating Secondary Attack Rates
Consider an outbreak of shigellosis in which 18 persons in 18 different households all became ill. If the population of the community was 1,000, then the overall attack rate was 18 / 1,000 x 100% = 1.8%. One incubation period later, 17 persons in the same households as these "primary" cases developed shigellosis. If the 18 households included 86 persons, calculate the secondary attack rate.
Secondary attack rate = (17 / (86 - 18)) x 100% = (17 / 68) x 100% = 25.0%

Incidence rate or person-time rate:
Incidence rate or person-time rate is a measure of incidence that incorporates time directly into the denominator. A person-time rate is generally calculated from a long-term cohort follow-up study, wherein enrollees are followed over time and the occurrence of new cases of disease is documented. Typically, each person is observed from an established starting time until one of four "end points" is reached: onset of disease, death, migration out of the study ("lost to follow-up"), or the end of the study. Similar to the incidence proportion, the numerator of the incidence rate is the number of new cases identified during the period of observation. However, the denominator differs. The denominator is the sum of the time each person was observed, totaled for all persons. This denominator represents the total time the population was at risk of and being watched for disease. Thus, the incidence rate is the ratio of the number of cases to the total time the population is at risk of disease.
Method for calculating incidence rate
Number of new cases of disease or injury during specified period
Time each person was observed, totaled for all persons
In a long-term follow-up study of morbidity, each study participant may be followed or observed for several years. One person followed for 5 years without developing disease is said to contribute 5 person-years of follow-up.

What about a person followed for one year before being lost to follow-up at year 2? Many researchers assume that persons lost to follow-up were, on average, disease-free for half the year, and thus contribute ½ year to the denominator. Therefore, the person followed for one year before being lost to follow-up contributes 1.5 person-years. The same assumption is made for participants diagnosed with the disease at the year 2 examination — some may have developed illness in month 1, and others in months 2 through 12. So, on average, they developed illness halfway through the year. As a result, persons diagnosed with the disease contribute ½ year of follow-up during the year of diagnosis.

The denominator of the person-time rate is the sum of all of the person-years for each study participant. So, someone lost to follow-up in year 3, and someone diagnosed with the disease in year 3, each contributes 2.5 years of disease-free follow-up to the denominator.

Properties and uses of incidence rates:

• An incidence rate describes how quickly disease occurs in a population. It is based on person-time, so it has some advantages over an incidence proportion. Because person-time is calculated for each subject, it can accommodate persons coming into and leaving the study. As noted in the previous example, the denominator accounts for study participants who are lost to follow-up or who die during the study period. In addition, it allows enrollees to enter the study at different times. In the NHANES follow-up study, some participants were enrolled in 1971, others in 1972, 1973, 1974, and 1975.

• Person-time has one important drawback. Person-time assumes that the probability of disease during the study period is constant, so that 10 persons followed for one year equals one person followed for 10 years. Because the risk of many chronic diseases increases with age, this assumption is often not valid.

• Long-term cohort studies of the type described here are not very common. However, epidemiologists far more commonly calculate incidence rates based on a numerator of cases observed or reported, and a denominator based on the mid-year population. This type of incident rate turns out to be comparable to a person-time rate.

• Finally, if you report the incidence rate of, say, the heart disease study as 2.5 per 1,000 person-years, epidemiologists might understand, but most others will not. Person-time is epidemiologic jargon. To convert this jargon to something understandable, simply replace “person-years” with “persons per year.” Reporting the results as 2.5 new cases of heart disease per 1,000 persons per year sounds like English rather than jargon. It also conveys the sense of the incidence rate as a dynamic process, the speed at which new cases of disease occur in the population.

EXAMPLES: Calculating Incidence Rates

Example A: Investigators enrolled 2,100 women in a study and followed them annually for four years to determine the incidence rate of heart disease. After one year, none had a new diagnosis of heart disease, but 100 had been lost to follow-up. After two years, one had a new diagnosis of heart disease, and another 99 had been lost to follow-up. After three years, another seven had new diagnoses of heart disease, and 793 had been lost to follow-up. After four years, another 8 had new diagnoses with heart disease, and 392 more had been lost to follow-up. The study results could also be described as follows: No heart disease was diagnosed at the first year. Heart disease was diagnosed in one woman at the second year, in seven women at the third year, and in eight women at the fourth year of follow-up. One hundred women were lost to follow-up by the first year, another 99 were lost to followup after two years, another 793 were lost to follow-up after three years, and another 392 women were lost to followup after 4 years, leaving 700 women who were followed for four years and remained disease free. Calculate the incidence rate of heart disease among this cohort. Assume that persons with new diagnoses of heart disease and those lost to follow-up were disease-free for half the year, and thus contribute ½ year to the denominator.

Numerator = number of new cases of heart disease

= 0 + 1 + 7 + 8 = 16

Denominator = person-years of observation

= (2,000 + ½ x 100) + (1,900 + ½ x 1 + ½ x 99) + (1,100 + ½ x 7 + ½ x 793) +
(700 + ½ x 8 + ½ x 392)

= 6,400 person-years of follow-up

or

Denominator = person-years of observation

= (1 x 1.5) + (7 x 2.5) + (8 x 3.5) + (100 x 0.5) + (99 x 1.5) + (793 x 2.5) +
(392 x 3.5) + (700 x 4)

= 6,400 person-years of follow-up
Person-time rate = Number of new cases of disease or injury during specified period
Time each person was observed, totaled for all persons
= 16 / 6,400
= .0025 cases per person-year
= 2.5 cases per 1,000 person-years
In contrast, the incidence proportion can be calculated as 16 / 2,100 = 7.6 cases per 1,000 population during the four-year period, or an average of 1.9 cases per 1,000 per year (7.6 divided by 4 years). The incidence proportion underestimates the true rate because it ignores persons lost to follow-up, and assumes that they remained diseasefree for all four years.

Example B: The diabetes follow-up study included 218 diabetic women and 3,823 nondiabetic women. By the end of the study, 72 of the diabetic women and 511 of the nondiabetic women had died. The diabetic women were observed for a total of 1,862 person-years; the nondiabetic women were observed for a total of 36,653 person-years.
Calculate the incidence rates of death for the diabetic and non-diabetic women.
For diabetic women, numerator = 72 and denominator = 1,862
Person-time rate = 72 / 1,862
= 0.0386 deaths per person-year
= 38.6 deaths per 1,000 person-years
For nondiabetic women, numerator = 511 and denominator = 36,653
Person-time rate = 511 / 36,653 = 0.0139 deaths per person-year
= 13.9 deaths per 1,000 person-years

Example C: In 2003, 44,232 new cases of acquired immunodeficiency syndrome (AIDS) were reported in the United
States.5 The estimated mid-year population of the U.S. in 2003 was approximately 290,809,777.6
Calculate the incidence rate of AIDS in 2003.
Numerator = 44,232 new cases of AIDS
Denominator = 290,809,777 estimated mid-year population
10n = 100,000
Incidence rate = (44,232 / 290,809,777) x 100,000
= 15.21 new cases of AIDS per 100,000 population

Prevalence:
Prevalence, sometimes referred to as **prevalence rate**, is the proportion of persons in a population who have a particular disease or attribute at a specified point in time or over a specified period of time. Prevalence differs from incidence in that prevalence includes all cases, both new and preexisting, in the population at the specified time, whereas incidence is limited to new cases only.
Point prevalence refers to the prevalence measured at a particular point in time. It is the proportion of persons with a particular disease or attribute on a particular date.
Period prevalence refers to prevalence measured over an interval of time. It is the proportion of persons with a particular disease or attribute at any time during the interval.
Method for calculating prevalence of disease
All new and pre-existing cases during a given time period
Population during the same time period x 10n
Method for calculating prevalence of an attribute
Persons having a particular attribute during a given time period
Population during the same time period x 10n

The value of 10n is usually 1 or 100 for common attributes. The value of 10n might be 1,000, 100,000, or even 1,000,000 for rare attributes and for most diseases.

EXAMPLE: Calculating Prevalence
In a survey of 1,150 women who gave birth in Maine in 2000, a total of 468 reported taking a multivitamin at least 4 times a week during the month before becoming pregnant.7 Calculate the prevalence of frequent multivitamin use in this group.
Numerator = 468 multivitamin users
Denominator = 1,150 women
Prevalence = (468 / 1,150) x 100 = 0.407 x 100 = 40.7%

Properties and uses of prevalence:
• Prevalence and incidence are frequently confused. Prevalence refers to proportion of persons who *have* a condition at or during a particular time period, whereas incidence refers to the proportion or rate of persons who *develop* a condition during a particular time period. So prevalence and incidence are similar, but prevalence includes new and pre-existing cases whereas incidence includes new cases only. The key difference is in their numerators.
Numerator of incidence = new cases that occurred during a given time period
Numerator of prevalence = all cases present during a given time period
• The numerator of an incidence proportion or rate consists only of persons whose illness began during the specified interval.
The numerator for prevalence includes all persons ill from a specified cause during the specified interval **regardless of when the illness began**. It includes not only new cases, but also preexisting cases representing persons who remained ill during some portion of the specified interval.
• Prevalence is based on both incidence and duration of illness. High prevalence of a disease within a population might reflect high incidence or prolonged survival without cure or both. Conversely, low prevalence might indicate low incidence, a rapidly fatal process, or rapid recovery.
• Prevalence rather than incidence is often measured for chronic diseases such as diabetes or osteoarthritis which have long duration and dates of onset that are difficult to pinpoint.

EXAMPLES: Incidence versus Prevalence
10 new cases of illness over about 15 months in a population of 20 persons whereby each horizontal line represents one person, The down arrow indicates the date of onset of illness. The solid line represents the duration of illness. The up arrow and the cross represent the date of recovery and date of death, respectively.

Example A: Calculate the incidence rate from October 1, 2004, to September 30, 2005, using the midpoint population (population alive on April 1, 2005) as the denominator. Express the rate per 100 population. Incidence rate numerator = number of new cases between October 1 and September 30
= 4 (the other 6 all had onsets before October 1, and are not included)
Incidence rate denominator = April 1 population
= 18 (persons 2 and 8 died before April 1)
Incidence rate = (4 / 18) x 100
= 22 new cases per 100 population
Example B: Calculate the point prevalence on April 1, 2005. Point prevalence is the number of persons ill on the date divided by the population on that date. On April 1, seven persons (persons 1, 4, 5, 7, 9, and 10) were ill.
Point prevalence = (7 / 18) x 100
= 38.89%

Example C: Calculate the period prevalence from October 1, 2004, to September 30, 2005. The numerator of period prevalence includes anyone who was ill any time during the period. In Figure 3.1, the first 10 persons were all ill at some time during the period.
Period prevalence = (10 / 20) x 100
= 50.0%

Mortality Frequency Measures
Mortality rate:
A mortality rate is a measure of the frequency of occurrence of death in a defined population during a specified interval. Morbidity and mortality measures are often the same mathematically; it's just a matter of what you choose to measure, illness or death. The formula for the mortality of a defined population, over a specified period of time, is:
Deaths occurring during a given time period
Size of the population among which the deaths occurred x 10n
When mortality rates are based on vital statistics (e.g., counts of death certificates), the denominator most commonly used is the size of the population at the middle of the time period. In the United States, values of 1,000 and 100,000 are both used for 10n for most types of mortality rates. Table 3.4 summarizes the formulas of frequently used mortality measures.

Crude mortality rate (crude death rate):
The crude mortality rate is the mortality rate from all causes of death for a population. In the United States in 2003, a total of 2,419,921 deaths occurred. The estimated population was 290,809,777. The crude mortality rate in 2003 was, therefore,
(2,419,921 / 290,809,777) x 100,000, or 832.1 deaths per 100,000 population.

Cause-specific mortality rate:
The cause-specific mortality rate is the mortality rate from a specified cause for a population. The numerator is the number of deaths attributed to a specific cause. The denominator remains the size of the population at the midpoint of the time period. The fraction is usually expressed per 100,000 population. In the United States in 2003, a total of 108,256 deaths were attributed to accidents (unintentional injuries), yielding a cause-specific mortality rate of 37.2 per 100,000 population.8

Age-specific mortality rate:
An age-specific mortality rate is a mortality rate limited to a particular age group. The numerator is the number of deaths in that age group; the denominator is the number of persons in that age group in the population. In the United States in 2003, a total of
130,761 deaths occurred among persons aged 25-44 years, or an age-specific mortality rate of 153.0 per 100,000 25–44 year olds. Some specific types of age-specific mortality rates are neonatal, postneonatal, and infant mortality rates, as described in the following sections.

Infant mortality rate:
The infant mortality rate is perhaps the most commonly used measure for comparing health status among nations. It is calculated as follows:
Number of deaths among children < 1 year of age reported during a given time period
Number of live births reported during the same time period x 1,000
The infant mortality rate is generally calculated on an annual basis.
It is a widely used measure of health status because it reflects the health of the mother and infant during pregnancy and the year thereafter. The health of the mother and infant, in turn, reflects a wide variety of factors, including access to prenatal care, prevalence of prenatal maternal health behaviors (such as

alcohol or tobacco use and proper nutrition during pregnancy, etc.), postnatal care and behaviors (including childhood immunizations and proper nutrition), sanitation, and infection control.

Is the infant mortality rate a ratio? Yes. Is it a proportion? No, because some of the deaths in the numerator were among children born the previous year. Consider the infant mortality rate in 2003. That year, 28,025 infants died and 4,089,950 children were born, for an infant mortality rate of 6.951 per 1,000.8 Undoubtedly, some of the deaths in 2003 occurred among children born in 2002, butthe denominator includes only children born in 2003.

Is the infant mortality rate truly a rate? No, because the denominator is not the size of the mid-year population of children < 1 year of age in 2003. In fact, the age-specific death rate for children < 1 year of age for 2003 was 694.7 per 100,000.8

Obviously the infant mortality rate and the age-specific death rate for infants are very similar (695.1 versus 694.7 per 100,000) and close enough for most purposes. They are not exactly the same, however, because the estimated number of infants residing in the

United States on July 1, 2003 was slightly larger than the number of children born in the United States in 2002, presumably because of immigration.

Neonatal mortality rate:

The neonatal period covers birth up to but not including 28 days. The numerator of the neonatal mortality rate therefore is the number of deaths among children under 28 days of age during a given time period. The denominator of the neonatal mortality rate, like that of the infant mortality rate, is the number of live births reported during the same time period. The neonatal mortality rate is usually expressed per 1,000 live births. In 2003, the neonatal mortality rate in the United States was 4.7 per 1,000 live births.8

Postneonatal mortality rate:

The postneonatal period is defined as the period from 28 days of age up to but not including 1 year of age. The numerator of the postneonatal mortality rate therefore is the number of deaths among children from 28 days up to but not including 1 year of age during a given time period. The denominator is the number of live births reported during the same time period. The postneonatal mortality rate is usually expressed per 1,000 live births. In 2003, the postneonatal mortality rate in the United States was 2.3 per 1,000 live births

Maternal mortality rate:

The maternal mortality rate is really a ratio used to measure mortality associated with pregnancy. The numerator is the number of deaths during a given time period among women while pregnant or within 42 days of termination of pregnancy, irrespective of the duration and the site of the pregnancy, from any cause related to or aggravated by the pregnancy or its management, but not from accidental or incidental causes. The denominator is the number of live births reported during the same time period. Maternal mortality rate is usually expressed per 100,000 live births. In 2003, the U.S. maternal mortality rate was 8.9 per 100,000 live births.

Sex-specific mortality rate:

A sex-specific mortality rate is a mortality rate among either males or females. Both numerator and denominator are limited to the one sex.

Race-specific mortality rate:

A race-specific mortality rate is a mortality rate related to a specified racial group. Both numerator and denominator are limited to the specified race.

Combinations of specific mortality rates:

Mortality rates can be further stratified by combinations of cause, age, sex, and/or race. For example, in 2002, the death rate from diseases of the heart among women ages 45–54 years was 50.6 per 100,000.9 The death rate from diseases of the heart among men in the same age group was 138.4 per 100,000, or more than 2.5 times as high as the comparable rate for women. These rates are a cause-, age-, and sex-specific rates, because they refer to one cause (diseases of the heart), one age group (45–54 years), and one sex (female or male).

EXAMPLE: Calculating Mortality Rates
The number of deaths from all causes and from accidents (unintentional injuries) by age group in the United States in 2002. Review the following rates. Determine what to call each one, then calculate it using the data provided. a. Unintentional-injury-specific mortality rate for the entire population
This is a cause-specific mortality rate.
Rate = number of unintentional injury deaths in the entire population x 100,000 estimated midyear population
= (106,742 / 288,357,000) x 100,000
= 37.0 unintentional-injury-related deaths per 100,000 population
b. All-cause mortality rate for 25–34 year olds
This is an age-specific mortality rate.
Rate = number of deaths from all causes among 25–34 year olds x 100,000 estimated midyear population of 25–34 year olds
= (41,355 / 39,928,000) x 100,000
= 103.6 deaths per 100,000 25–34 year olds
c. All-cause mortality among males
This is a sex-specific mortality rate.
Rate = number of deaths from all causes among males x 100,000 estimated midyear population of males
= (1,199,264 / 141,656,000) x 100,000
= 846.6 deaths per 100,000 males
d. Unintentional-injury-specific mortality among 25- to 34-year-old males
This is a cause-specific, age-specific, and sex-specific mortality rate
Rate = number of unintentional injury deaths among 25–34 year old males x 100,000 estimated midyear population of 25–34 year old males
= (9,635 / 20,203,000) x 100,000
= 47.7 unintentional-injury-related deaths per 100,000 25–34 year olds

Age-adjusted mortality rate: a mortality rate statistically modified to eliminate the effect of different age distributions in the different populations. Mortality rates can be used to compare the rates in one area with the rates in another area, or to compare rates over time. However, because mortality rates obviously increase with age, a higher mortality rate among one population than among another might simply reflect the fact that the first population is older than the second.

Consider that the mortality rates in 2002 for the states of Alaska and Florida were 472.2 and 1,005.7 per 100,000, respectively. Should everyone from Florida move to Alaska to reduce their risk of death? No, the reason that Alaska's mortality rate is so much lower than Florida's is that Alaska's population is considerably younger. Indeed, for seven age groups, the age-specific mortality rates in Alaska are actually higher than Florida's.

To eliminate the distortion caused by different underlying age distributions in different populations, statistical techniques are used to adjust or standardize the rates among the populations to be compared. These techniques take a weighted average of the agespecific mortality rates, and eliminate the effect of different age distributions among the different populations. Mortality rates computed with these techniques are **age-adjusted** or **age-standardized mortality rates**. Alaska's 2002 age-adjusted mortality rate (794.1 per 100,000) was higher than Florida's (787.8 per 100,000), which is not surprising given that 7 of 13 agespecific mortality rates were higher in Alaska than Florida.

Death-to-case ratio:

The death-to-case ratio is the number of deaths attributed to a particular disease during a specified time period divided by the number of new cases of that disease identified during the same time period. The death-to-case ratio is a ratio but not necessarily a proportion, because some of the deaths that are counted in the numerator might have occurred among persons who developed disease in an earlier period, and are therefore not counted in the denominator.

EXAMPLE: Calculating Death-to-Case Ratios

Between 1940 and 1949, a total of 143,497 incident cases of diphtheria were reported. During the same decade, 11,228 deaths were attributed to diphtheria. Calculate the death-to-case ratio.
Death-to-case ratio = 11,228 / 143,497 x 1 = 0.0783 or
= 11,228 / 143,497 x 100 = 7.83 per 100

Case-fatality rate:

The case-fatality rate is the proportion of persons with a particular condition (cases) who die from that condition. It is a measure of the severity of the condition. The formula is:
Number of cause-specific deaths among the incident cases
Number of incident cases x 10n
The case-fatality rate is a proportion, so the numerator is restricted to deaths among people included in the denominator. The time periods for the numerator and the denominator do not need to be the same; the denominator could be cases of HIV/AIDS diagnosed during the calendar year 1990, and the numerator, deaths among those diagnosed with HIV in 1990, could be from 1990 to the present.

EXAMPLE: Calculating Case-Fatality Rates

In an epidemic of hepatitis A traced to green onions from a restaurant, 555 cases were identified. Three of the casepatients died as a result of their infections. Calculate the case-fatality rate.
Case-fatality rate = (3 / 555) x 100 = 0.5%
The case-fatality rate is a proportion, not a true rate. As a result, some epidemiologists prefer the term **case-fatality ratio**. The concept behind the case-fatality rate and the death-to-case ratio is similar, but the formulations are different. The death-tocase ratio is simply the number of cause-specific deaths that occurred during a specified time divided by the number of new cases of that disease that occurred during the same time. The deaths included in the numerator of the death-to-case ratio are not restricted to the new cases in the denominator; in fact, for many diseases, the deaths are among persons whose onset of disease was years earlier. In contrast, in the case-fatality rate, the deaths included in the numerator are restricted to the cases in the denominator.

Proportionate mortality:

Proportionate mortality describes the proportion of deaths in a specified population over a period of time attributable to different causes. Each cause is expressed as a percentage of all deaths, and the sum of the causes must add to 100%. These proportions are not mortality rates, because the denominator is all deaths rather than the population in which the deaths occurred.
Method for calculating proportionate mortality
For a specified population over a specified period,
Deaths caused by a particular cause
Deaths from all causes x 100

The distribution of primary causes of death in the United States in 2003 for the entire population (all ages) and for persons ages 25– 44 years are provided in Table 3.1. As illustrated in that table, accidents (unintentional injuries) accounted for 4.3% of all deaths, but 21.6% of deaths among 25–44 year olds.8 Sometimes, particularly in occupational epidemiology, proportionate mortality is used to compare deaths in a population of interest (say, a workplace) with the proportionate mortality in the broader population. This comparison of two proportionate mortalities is called a **proportionate mortality ratio**, or PMR for short. A PMR greater than 1.0 indicates that a particular cause accounts for a greater proportion of deaths in the population of interest than you might expect. For example, construction workers may be more likely to die of injuries than the general population. However, PMRs can be misleading, because they are not based on mortality rates. A low cause-specific mortality rate in the population of interest can elevate the proportionate mortalities for all of the other causes, because they must add up to 100%. Those workers with a high injury-related proportionate mortality very likely have lower proportionate mortalities for chronic or disabling conditions that keep people out of the workforce. In other words, people who work are more likely to be healthier than the population as a whole — this is known as the healthy worker effect.

Years of potential life lost:
Years of potential life lost (YPLL) is one measure of the impact of premature mortality on a population. Additional measures incorporate disability and other measures of quality of life. YPLL is calculated as the sum of the differences between a predetermined end point and the ages of death for those who died before that end point. The two most commonly used end points are age 65 years and average life expectancy.

The use of YPLL is affected by this calculation, which implies a value system in which more weight is given to a death when it occurs at an earlier age. Thus, deaths at older ages are "devalued." However, the YPLL before age 65 (YPLL65) places much more emphasis on deaths at early ages than does YPLL based on remaining life expectancy (YPLLLE). In 2000, the remaining life expectancy was 21.6 years for a 60-year-old, 11.3 years for a 70- year-old, and 8.6 for an 80-year-old. YPLL65 is based on the fewer than 30% of deaths that occur among persons younger than 65. In contrast, YPLL for life expectancy (YPLLLE) is based on deaths among persons of all ages, so it more closely resembles crude mortality rates. YPLL rates can be used to compare YPLL among populations of different sizes. Because different populations may also have different age distributions, YPLL rates are usually age-adjusted to eliminate the effect of differing age distributions.

Method for calculating YPLL from a line listing:
Step 1. Decide on end point (65 years, average life expectancy, or other).
Step 2. Exclude records of all persons who died at or after the end point.
Step 3. For each person who died before the end point, calculate that person's YPLL by subtracting the age at death from the end point. YPLLindividual = end point – age at death
Step 4. Sum the individual YPLLs. YPLL = . YPLLindividual

Method for calculating YPLL from a frequency:
Step 1. Ensure that age groups break at the identified end point
(e.g., 65 years). Eliminate all age groups older than the endpoint.
Step 2. For each age group younger than the end point, identify the midpoint of the age group, where midpoint = age group's youngest age in years + oldest age + 12
Step 3. For each age group younger than the end point, identify that age group's YPLL by subtracting the midpoint from the end point.
Step 4. Calculate age-specific YPLL by multiplying the age group's YPLL times the number of persons in that age group.

Step 5. Sum the age-specific YPLL's.
The **YPLL rate** represents years of potential life lost per 1,000 population below the end-point age, such as 65 years. YPLL rates should be used to compare premature mortality in different populations, because YPLL does not take into account differences in population sizes.

The formula for a YPLL rate is as follows:
Years of potential life lost Population under age 65 years x 10n

EXAMPLE: Calculating YPLL and YPLL Rates
Use the data in Tables 3.9 and 3.10 to calculate the leukemia-related mortality rate for all ages, mortality rate for persons under age 65 years, YPLL, and YPLL rate.
1. Leukemia-related mortality rate, all ages
= (21,498 / 288,357,000) x 100,000 = 7.5 leukemia deaths per 100,000 population
2. Leukemia-related mortality rate for persons under age 65 years
= 125 + 316 + 472 + 471 + 767 + 1,459 + 2,611 = x 100,000
(19,597 + 41,037 + 40,590 +39,928 + 44,917 + 40,084 + 26,602)
= 6,221 / 252,755,000 = x 100,000
= 2.5 leukemia deaths per 100,000 persons under age 65 years
3. Leukemia-related YPLL
a. Calculate the midpoint of each age interval. Using the previously shown formula, the midpoint of the age group 0–4 years is (0 + 4 + 1) / 2, or 5 / 2, or 2.5 years. Using the same formula, midpoints must be determined for each age group up to and including the age group 55 to 64 years.
b. Subtract the midpoint from the end point to determine the years of potential life lost for a particular age group. For the age group 0–4 years, each death represents 65 minus 2.5, or 62.5 years of potential life lost (see column 4 of Table 3.10).
c. Calculate age-specific years of potential life lost by multiplying the number of deaths in a given age group by its years of potential life lost. For the age group 0–4 years, 125 deaths x 62.5 = 7,812.5 YPLL
d. Total the age-specific YPLL. The total YPLL attributed to leukemia in the United States in 2002 was 117,033 years (see Total of column 5, Table 3.10).
4. Leukemia-related YPLL rate
= YPLL65 rate
= YPLL divided by population to age 65
= (117,033 / 252,755,000) x 1,000
= 0.5 YPLL per 1,000 population under age 65

Natality (Birth) Measures
Natality measures are population-based measures of birth. These measures are used primarily by persons working in the field of maternal and child health. Table 3.11 includes some of the commonly used measures of natality.

Measures of Association
The key to epidemiologic analysis is comparison. Occasionally you might observe an incidence rate among a population that seems high and wonder whether it is actually higher than what should be expected based on, say, the incidence rates in other communities. Or, you might observe that, among a group of casepatients in an outbreak, several report having eaten at a particular restaurant. Is the restaurant just a popular one, or have more casepatients eaten there than would be expected? The way to address that concern is by comparing the observed group with another group that represents the expected level.

A measure of association quantifies the relationship between exposure and disease among the two groups. Exposure is used loosely to mean not only exposure to foods, mosquitoes, a partner with a sexually transmissible disease, or a toxic waste dump, but also inherent characteristics of persons (for example, age, race, sex), biologic characteristics (immune status), acquired characteristics (marital status), activities (occupation, leisure activities), or conditions under which they live (socioeconomic status or access to medical care). The measures of association described in the following section compare disease occurrence among one group with disease occurrence in another group. Examples of measures of association include risk ratio (relative risk), rate ratio, odds ratio, and proportionate mortality ratio.

Risk ratio:

A risk ratio (RR), also called relative risk, compares the risk of a health event (disease, injury, risk factor, or death) among one group with the risk among another group. It does so by dividing the risk (incidence proportion, attack rate) in group 1 by the risk (incidence proportion, attack rate) in group 2 . The two groups are typically differentiated by such demographic factors as sex (e.g., males versus females) or by exposure to a suspected risk factor (e.g., did or did not eat potato salad). Often, the group of primary interest is labeled the exposed group, and the comparison group is labeled the unexposed group.

Method for Calculating risk ratio:

The formula for risk ratio (RR) is:

Risk of disease (incidence proportion, attack rate) in group of primary interest

Risk of disease (incidence proportion, attack rate) in comparison group

A risk ratio of 1.0 indicates identical risk among the two groups. A risk ratio greater than 1.0 indicates an increased risk for the group in the numerator, usually the exposed group. A risk ratio less than 1.0 indicates a decreased risk for the exposed group, indicating that perhaps exposure actually protects against disease occurrence.

EXAMPLES: Calculating Risk Ratios

Example A: In an outbreak of tuberculosis among prison inmates in South Carolina in 1999, 28 of 157 inmates residing on the East wing of the dormitory developed tuberculosis, compared with 4 of 137 inmates residing on the West wing. These data are summarized in the two-by-two table so called because it has two rows for the exposure and two columns for the outcome. Here is the general format and notation.

Example B: In an outbreak of varicella (chickenpox) in Oregon in 2002, varicella was diagnosed in 18 of 152 vaccinated children compared with 3 of 7 unvaccinated children. Calculate the risk ratio.

Rate ratio:

A rate ratio compares the incidence rates, person-time rates, or mortality rates of two groups. As with the risk ratio, the two groups are typically differentiated by demographic factors or by exposure to a suspected causative agent. The rate for the group of primary interest is divided by the rate for the comparison group.

Rate for group of primary interest Rate ratio = *Rate for comparison group*

The interpretation of the value of a rate ratio is similar to that of the risk ratio. That is, a rate ratio of 1.0 indicates equal rates in the two groups, a rate ratio greater than 1.0 indicates an increased risk for the group in the numerator, and a rate ratio less than 1.0 indicates a decreased risk for the group in the numerator.

EXAMPLE: Calculating Rate Ratios

Public health officials were called to investigate a perceived increase in visits to ships' infirmaries for acute respiratory illness (ARI) by passengers of cruise ships in Alaska in 1998.13 The officials compared passenger visits to ship infirmaries for ARI during May–August 1998 with the same period in 1997. They recorded 11.6 visits for ARI per
1,000 tourists per week in 1998, compared with 5.3 visits per 1,000 tourists per week in 1997. Calculate the rate ratio.
Rate ratio = 11.6 / 5.3 = 2.2
Passengers on cruise ships in Alaska during May–August 1998 were more than twice as likely to visit their ships' infirmaries for ARI than were passengers in 1997. (Note: Of 58 viral isolates identified from nasal cultures from passengers, most were influenza A, making this the largest summertime influenza outbreak in North America.)

Odds ratio:
An odds ratio (OR) is another measure of association that quantifies the relationship between an exposure with two categories and health outcome. The odds ratio is calculated as *a c Odds ratio* = (*b*)(*d*) = *ad / bc* where
a = number of persons exposed and with disease
b = number of persons exposed but without disease
c = number of persons unexposed but with disease
d = number of persons unexposed: and without disease
a+c = total number of persons with disease (case-patients)
b+d = total number of persons without disease (controls)
The odds ratio is sometimes called the **cross-product ratio** because the numerator is based on multiplying the value in cell "a" times the value in cell "d," whereas the denominator is the product of cell "b" and cell "c." A line from cell "a" to cell "d" (for the numerator) and another from cell "b" to cell "c" (for the denominator) creates an x or cross on the two-by-two table.

EXAMPLE: Calculating Odds Ratios
Use the data provided to calculate the risk and odds ratios.
1. Risk ratio
5.0 / 1.0 = 5.0
2. Odds ratio
(100 x 7,920) / (1,900 x 80) = 5.2
Notice that the odds ratio of 5.2 is close to the risk ratio of 5.0. That is one of the attractive features of the odds ratio — when the health outcome is uncommon, the odds ratio provides a reasonable approximation of the risk ratio.
Another attractive feature is that the odds ratio can be calculated with data from a case-control study, whereas neither a risk ratio nor a rate ratio can be calculated.
In a case-control study, investigators enroll a group of case-patients (distributed in cells a and c of the two-by-two table), and a group of non-cases or controls (distributed in cells b and d).
The odds ratio is the measure of choice in a case-control study.

A case-control study is based on enrolling a group of persons with disease ("case-patients") and a comparable group without disease ("controls"). The number of persons in the control group is usually decided by the investigator. Often, the size of the population from which the case-patients came is not known. As a result, risks, rates, risk ratios or rate ratios cannot be calculated from the typical case-control study. However, you can calculate an odds ratio and interpret it as an approximation of the risk ratio, particularly when the disease is uncommon in the population.

Measures of Public Health Impact

A measure of public health impact is used to place the association between an exposure and an outcome into a meaningful public health context. Whereas a measure of association quantifies the relationship between exposure and disease, and thus begins to provide insight into causal relationships, measures of public health impact reflect the burden that an exposure contributes to the frequency of disease in the population. Two measures of public health impact often used are the attributable proportion and efficacy or effectiveness.

Attributable proportion:

The attributable proportion, also known as the attributable risk percent, is a measure of the public health impact of a causative factor. The calculation of this measure assumes that the occurrence of disease in the unexposed group represents the baseline or expected risk for that disease. It further assumes that if the risk of disease in the exposed group is higher than the risk in the unexposed group, the difference can be attributed to the exposure. Thus, the attributable proportion is the amount of disease in the exposed group attributable to the exposure. It represents the expected reduction in disease if the exposure could be removed (or never existed). Appropriate use of attributable proportion depends on a single risk factor being responsible for a condition. When multiple risk factors may interact (e.g., physical activity and age or health status), this measure may not be appropriate.

Method for calculating attributable proportion

Attributable proportion is calculated as follows:
Risk for exposed group – risk for unexposed group
Risk for exposed group x 100%
Attributable proportion can be calculated for rates in the same way.

EXAMPLE: Calculating Attributable Proportion

In another study of smoking and lung cancer, the lung cancer mortality rate among nonsmokers was 0.07 per 1,000 persons per year. The lung cancer mortality rate among persons who smoked 1–14 cigarettes per day was 0.57 lung cancer deaths per 1,000 persons per year. Calculate the attributable proportion.
Attributable proportion = (0.57 – 0.07) / 0.57 x 100% = 87.7%
Given the proven causal relationship between cigarette smoking and lung cancer, and assuming that the groups are comparable in all other ways, one could say that about 88% of the lung cancer among smokers of 1-14 cigarettes per day might be attributable to their smoking. The remaining 12% of the lung cancer cases in this group would have occurred anyway.

Vaccine efficacy or vaccine effectiveness:

Vaccine efficacy and vaccine effectiveness measure the proportionate reduction in cases among vaccinated persons. Vaccine efficacy is used when a study is carried out under ideal conditions, for example, during a clinical trial. Vaccine effectiveness is used when a study is carried out under typical field (that is, less than perfectly controlled) conditions. Vaccine efficacy/effectiveness (VE) is measured by calculating the risk of disease among vaccinated and unvaccinated persons and determining the percentage reduction in risk of disease among vaccinated persons relative to unvaccinated persons. The greater the percentage reduction of illness in the vaccinated group, the greater the vaccine efficacy/effectiveness. The basic formula is written as:
Risk among unvaccinated group – risk among vaccinated group
Risk among unvaccinated group
OR: *1 – risk ratio*
In the first formula, the numerator (risk among unvaccinated – risk among vaccinated) is sometimes called the risk difference or excess risk. Vaccine efficacy/effectiveness is interpreted as the proportionate reduction in disease among the vaccinated group. So a VE of 90% indicates a 90%

reduction in disease occurrence among the vaccinated group, or a 90% reduction from the number of cases you would expect if they have not been vaccinated.

EXAMPLE: Calculating Vaccine Effectiveness
Calculate the vaccine effectiveness from the varicella data in Table 3.13.
VE = (42.9 – 11.8) / 42.9 = 31.1 / 42.9 = 72%
Alternatively, VE = 1 – RR = 1 – 0.28 = 72%
So, the vaccinated group experienced 72% fewer varicella cases than they would have if they had not been vaccinated.

Summary
Because many of the variables encountered in field epidemiology are nominal-scale variables, frequency measures are used quite commonly in epidemiology. Frequency measures include ratios, proportions, and rates. Ratios and proportions are useful for describing the characteristics of populations. Proportions and rates are used for quantifying morbidity and mortality. These measures allow epidemiologists to infer risk among different groups, detect groups at high risk, and develop hypotheses about causes — that is, why these groups might be at increased risk.

The two primary measures of morbidity are incidence and prevalence.
• **Incidence** rates reflect the occurrence of new disease in a population.
• **Prevalence** reflects the presence of disease in a population.
A variety of **mortality** rates describe deaths among specific groups, particularly by age or sex or by cause. The hallmark of epidemiologic analysis is comparison, such as comparison of observed amount of disease in a population with the expected amount of disease. The comparisons can be quantified by using such measures of association as risk ratios, rate ratios, and odds ratios. These measures provide evidence regarding causal relationships between exposures and disease. Measures of public health impact place the association between an exposure and a disease in a public health context. Two such measures are the attributable proportion and vaccine efficacy.

Analyzing and Interpreting Data
After morbidity, mortality, and other relevant data about a health problem have been gathered and compiled, the data should be analyzed by time, place, and person. Different types of data are used for surveillance, and different types of analyses might be needed for each. For example, data on individual cases of disease are analyzed differently than data aggregated from multiple records; data received as text must be sorted, categorized, and coded for statistical analysis; and data from surveys might need to be weighted to produce valid estimates for sampled populations.

For analysis of the majority of surveillance data, descriptive methods are usually appropriate. The display of frequencies (counts) or rates of the health problem in simple tables and graphs, as discussed in Lesson 4, is the most common method of analyzing data for surveillance. Rates are useful — and frequently preferred — for comparing occurrence of disease for different geographic areas or periods because they take into account the size of the opulation from which the cases arose. One critical step before calculating a rate is constructing a denominator from appropriate population data. For state- or countywide rates, general population data are used. These data are available from the U.S. Census Bureau or from a state planning agency.

For other calculations, the population at risk can dictate an alternative denominator. For example, an infant mortality rate uses the number of live-born infants; rates of surgical wound infections in a hospital nrequires the number of such procedures performed. In addition to calculating\ frequencies and rates, more sophisticated methods (e.g., space-time cluster analysis, time series analysis, or computer mapping) can be applied. To determine whether the incidence or prevalence of a health problem has increased, data must be compared either over time or across areas. The selection of data for comparison depends on the health problem under surveillance and what is known about its typical temporal and geographic patterns of occurrence.

For example, data for diseases that indicate a seasonal pattern (e.g., influenza and mosquito-borne diseases) are usually compared with data for the corresponding season from past years. Data for diseases without a seasonal pattern are commonly compared with data for previous weeks, months, or years, depending on the nature of the disease. Surveillance for chronic diseases typically requiresdata covering multiple years. Data for acute infectious diseases might only require data covering weeks or months, although data extending over multiple years can also be helpful in the analysis of the natural history of disease. Data from one geographic area are sometimes compared with data from another area. For example, data from a county might be compared with data from adjacent counties or with data from the state. We now describe common methods for, and provide examples of, the analysis of data by time, place, and person.

MODULE 5—Psychological, Behavioral, and Social Issues in Public Health

Illness and poor health bring a unique set of psychological issues. Daily stress, work, family and financial problems can cause illness and poor health. Is there an answer? That is what a public health psychologist strives to answer as they work with medical personnel or in private practice. Hospitals, medical centers, nursing homes and public health centers employ experts in the field of public health psychology. Government programs can use public psychologists to help design and implement programs that are aimed at prevention and intervention in areas of health that affect the general population. Not only is the area of public health psychology personally rewarding but the job outlook is excellent. The Bureau of Labor Statistics, 'www.bls.gov *reports that 50% of psychologists earn between $45,000 and $77,000 annually*

Public Health Psychology - Making a Difference:
Health psychology is concerned with understanding how biology, behavior, and social context influence health and illness. Health psychologists work alongside other medical professionals in clinical settings, work on behaviour change in public health promotion, teach at universities, and conduct research. Although its early beginnings can be traced to the kindred field of clinical psychology, four different divisions within health psychology and one allied field have developed over time: clinical health psychology, occupational health psychology (an allied field), public health psychology, community health psychology, and critical health psychology. Organizations closely associated with the field of health psychology include Division 38 of the American Psychological Association and the Division of Health Psychology of the British Psychological Association.

Recent advances in psychological, medical, and physiological research have led to a new way of thinking about health and illness. This conceptualization, which has been labeled the biopsychosocial model, views health and illness as the product of a combination of factors including biological characteristics (e.g., genetic predisposition), behavioral factors (e.g., lifestyle, stress, health beliefs), and social conditions (e.g., cultural influences, family relationships, social support).

Psychologists who strive to understand how biological, behavioral, and social factors influence health and illness are called health psychologists. The term "health psychology" is often used synonymously with the terms "behavioral medicine" and "medical psychology". Health psychologists work with many different health care professionals (e.g., physicians, dentists, nurses, physician's assistants, dietitians, social workers, pharmacists, physical and occupational therapists, and chaplains) to conduct research and provide clinical assessments and treatment services. Many health psychologists focus on prevention research and interventions designed to promote health and reduce the risk of disease.

While more than half of health psychologists provide clinical services as part of their duties, many health psychologists function in non-clinical roles, primarily involving teaching and research.

- **Clinical health psychology** (**ClHP**) is a term for a division of health psychology that reflects the fact that the field was originally a branch of clinical psychology. ClHP is also a major contributor to the field of behavioral medicine within psychiatry. Clinical practice includes education, the techniques of behavior change, and psychotherapy. In some countries, a clinical health psychologist, with additional training, can become a medical psychologist and, thereby, obtain prescription privileges.
- **Occupational health psychology** (**OHP**) is a relatively new discipline allied with health psychology. The ancestry of OHP includes health psychology, industrial/organizational psychology, and occupational health. OHP has own doctoral programs, journals, and professional organizations. The field is concerned with identifying psychosocial characteristics of workplaces that give rise to health-related problems in people who work.

These problems can involve physical health (e.g., cardiovascular disease) or mental health (e.g., depression). Examples of psychosocial characteristics of workplaces that OHP has investigated include amount of decision latitude a worker can exercise and the supportiveness of supervisors. OHP is also concerned with the development and implementation of interventions that can prevent or ameliorate work-related health problems. In addition, OHP research has important implications for the economic success of organizations.

Other research areas of concern to OHP include workplace incivility and violence, work-home carryover, unemployment and downsizing, and workplace safety and accident prevention.

- **Public health psychology** (**PHP**) is population oriented. A major aim of PHP is to investigate potential causal links between psychosocial factors and health at the population level. PH psychologists present research results to educators, policy makers, and health care providers in order to promote better public health. PHP is allied to other public health disciplines including epidemiology, nutrition, genetics and biostatistics. Some PHP interventions are targeted toward at-risk population groups (e.g., undereducated, single pregnant women who smoke) and not the population as a whole (e.g., all pregnant women).
- **Community health psychology** (**CoHP**) investigates community factors that contribute to the health and well-being of individuals who live in communities. CoHP also develops community-level interventions that are designed to combat disease and promote physical and mental health. The community often serves as the level of analysis, and is frequently sought as a partner in health-related interventions.
- **Critical health psychology** (**CrHP**) is concerned with the distribution of power and the impact of power differentials on health experience and behavior, health care systems, and health policy. CrHP prioritizes social justice and the universal right to health for people of all races, genders, ages, and socioeconomic positions. A major concern is health inequalities. The CrH psychologist is an agent of change, not simply an analyst or cataloger.

Health psychology is both a theoretical and applied field. Health psychologists employ diverse research methods. These methods include controlled randomized experiments, quasi-experiments, longitudinal studies, time-series designs, cross-sectional studies, and case-control studies as well as action research. Health psychologists study a broad range of variables including genotype, cardiovascular disease, smoking habits, religious beliefs, alcohol use, social support, living conditions, emotional state, social class, and much more. Some health psychologists treat individuals with sleep problems, headaches, alcohol problems, etc. Other health psychologists work to empower community members by helping community members gain control over their health and improve quality of life of entire communities.

Objectives of health psychology

Understanding behavioral and contextual factors, Health psychologists conduct research to identify behaviors and experiences that promote health, give rise to illness, and influence the effectiveness of health care. They also recommend ways to improve health care and health-care policy. Health psychologists have worked on developing ways to reduce smoking and improve daily nutrition in order to promote health and prevent illness. They have also studied the association between illness and individual characteristics. For example, health psychology has found a relation between the personality characteristics thrill seeking, impulsiveness, hostility/anger, emotional instability, and depression, on one hand, and high-risk driving, on the other.

Health psychology is also concerned with contextual factors, including economic, cultural, community, social, and lifestyle factors that influence health. The biopsychosocial model can help in understanding the relation between contextual factors and biology in affecting health. Physical addiction plays an important role in smoking cessation. However, seductive advertising also contributes to psychological dependency on tobacco. Research in occupational health psychology indicates that people in jobs that combine little decision latitude with a high psychological workload are at increased risk for cardiovascular disease. Other OHP research reveals a relation between unemployment and elevations in blood pressure. OHP research also documents a relation between social class and cardiovascular disease. Health psychologists also aim to change health behaviors for the dual purpose of helping people stay healthy and helping patients adhere to disease treatment regimens. Health psychologists employ cognitive behavior therapy and applied behavior analysis for that purpose.

Preventing illness

Health psychologists work towards promoting health through behavioral change, as mentioned above; however, they attempt to prevent illness in other ways as well. Campaigns informed by health psychology have targeted tobacco use. Those least able to afford tobacco products consume them most. Tobacco provides individuals with a way controlling aversive emotional states accompanying daily experiences of stress that characterize the lives of deprived and vulnerable individuals. Practitioners emphasize education and effective communication as a part of illness prevention because many people do not recognize, or minimize, the risk of illness present in their lives. Moreover, many individuals are often unable to apply their knowledge of health practices owing to everyday pressures and stresses. A common example of population-based attempts to motivate the smoking public to reduce its dependence on cigarettes is anti-smoking campaigns.

Health psychologists also aim at educating health professionals, including physicians and nurses, in communicating effectively with patients in ways that overcome barriers to understanding, remembering, and implementing effective strategies for reducing exposures to risk factors and making health-enhancing behavior changes. There is also evidence from occupational health psychology that stress-reduction interventions at the workplace can be effective. For example, Kompier and his colleagues have shown that a number of interventions aimed at reducing stress in bus drivers has had beneficial effects for employees and bus companies.

The effects of disease

Health psychologists investigate how disease affects individuals' psychological well-being. An individual who becomes seriously ill or injured faces many different practical stressors. The stresssors include problems meeting medical and other bills; problems obtaining proper care when home from the hospital; obstacles to caring for dependents; having one's sense of self-reliance compromised; gaining a new, unwanted identity as a sick person; and so on. These stressors can lead to depression, reduced self-esteem, etc Health psychology also concerns itself with bettering the lives of individuals with terminal illness. When there is little hope of recovery, health psychologist therapists can improve the quality of life of the patient by helping the patient recover at least some his or her psychological well-being. Health psychologists are also concerned with identifying the best ways for providing therapeutic services for the bereaved.

Critical analysis of health policy
Critical health psychologists explore how health policy can influence inequities, inequalities, and social injustice. These avenues of research expand the scope of health psychology beyond the level of individual health to an examination of the social and economic determinants of health both within and between regions and nations. The individualism of mainstream health psychology has been critiqued and deconstructed by critical health psychologists using newer qualitative methods and frameworks for investigating health experience and behavior.

Applications of Health Psychology
Improving doctor-patient communication, Health psychologists attempt to aid the process of communication between physicians and patients during medical consultations. There are many problems in this process, with patients showing a considerable lack of understanding of many medical terms, particularly anatomical terms (e.g., intestines). One main area of research on this topic involves 'doctor-centered' or 'patient-centered' consultations. Doctor-centered consultations are generally directive, with the patient answering questions and playing less of a role in decision-making. Although this style is preferred by elderly people and others, many people dislike the sense of hierarchy or ignorance that it inspires. They prefer patient-centered consultations, which focus on the patient's needs, involve the doctor listening to the patient completely before making a decision, and involving the patient in the process of choosing treatment and finding a diagnosis.

Improving adherence to medical advice
Getting people to follow medical advice and adhere to their treatment regimens is a difficult task for health psychologists. People often forget to take their pills or are inhibited by the side effects of their medicines. Failing to take prescribed medication is costly and wastes millions of usable medicines that could otherwise help other people. Estimated adherence rates are difficult to measure (see below); there is, however, evidence that adherence could be improved by tailoring treatment programs to individuals' daily lives.

Ways of measuring adherence
Health psychologists have identified a number of ways of measuring patients' adherence to medical regimens.

- Counting the number of pills in the medicine bottle - although this has problems with privacy and/or could be deemed patronizing or showing lack of trust in patients
- Using self-reports - although patients may fail to return the self-report or lie about their adherence
- Asking a doctor or health worker - although this presents problems on doctor-patient confidentiality
- Using 'Trackcap' bottles, which track the number of times the bottle is opened; however, this either raises problems of informed consent or, if informed consent is obtained, influence through demand characteristics.

Managing pain

Health psychology attempts to find treatments to reduce and eliminate pain, as well as understand pain anomalies such as episodic analgesia, causalgia, neuralgia, and phantom limb pain. Although the task of measuring and describing pain has been problematic, the development of the McGill Pain Questionnaire[36] has helped make progress in this area. Treatments for pain involve patient-administered analgesia, acupuncture (found by Berman to be effective in reducing pain for osteoarthritis of the knee), biofeedback, and cognitive behavior therapy.

Mind/Body Health: Stress

We've probably all felt stress. Sometimes it's brief and highly situational, like being in heavy traffic. Other times, it's more persistent and complex—relationship problems, an ailing family member, a spouse's death. And sometimes, stress can motivate us to accomplish certain tasks.

Dangerous Stress

Stress becomes dangerous when it interferes with your ability to live a normal life for an extended period of time. You may feel "out of control" and have no idea of what to do, even if the cause is relatively minor. This in turn, may cause you to feel continually fatigued, unable to concentrate, or irritable in otherwise relaxed situations. Prolonged stress may also compound any emotional problems stemming from sudden events such traumatic experiences in your past, and increase thoughts of suicide.

Natural reactions

Stress can also affect your physical health because of the human body's built-in response mechanisms. You may have found yourself sweating at the thought of an important date, or felt your heartbeat pick up while watching a scary movie. These reactions are caused by hormones that scientists believe helped our ancestors cope with the threats and uncertainties of their world. If the cause of your stress is temporary, the physical effects are usually short-term as well. In one study, the pressure of taking exams led to increased severity of acne among college students, regardless of how they ate or slept. The condition diminished after exams were over. Abdominal pain and irregularity have also been linked to situational stress. The longer your mind feels stressed, however, the longer your physical reaction systems remain activated. This can lead to more serious health issues.

Physical wear and tear

The old saying that stress "ages" a person faster than normal was recently verified in a study of women who had spent many years caring for severely ill and disabled children. Because their bodies were no longer able to fully regenerate blood cells, these women were found to be physically a decade older than their chronological age. Extended reactions to stress can alter the body's immune system in ways that are associated with other "aging" conditions such as frailty, functional decline, cardiovascular disease, osteoporosis, inflammatory arthritis, type 2 diabetes, and certain cancers. Research also suggests that stress impairs the brain's ability to block certain toxins and other large, potentially harmful molecules. This condition is also common to patients suffering from Alzheimer's Disease.

Pressure points

Although sudden emotional stress has been linked to severe heart dysfunction in otherwise healthy people, scientists are uncertain whether chronic stress alone causes cardiovascular disease. What is clear is that excessive stress can worsen existing risk factors such as hypertension and high cholesterol levels. Studies also show that people who are quick to anger or who display frequent hostility—a behavior common to those under stress—have an increased risk of heart disease and crying fits. Feelings of despair that accompany stress can easily worsen into chronic depression, a condition that can lead you to neglect good diet and activity habits. This, in turn, can put you at a greater risk for heart disease, obesity, and kidney dysfunction. Stress can also complicate your ability to recover from a serious illness. A Swedish study found that women who have suffered heart attacks tend to have poorer chances of recovery if they are also experiencing marital stressors such as infidelity, alcohol abuse, and a spouse's physical or psychiatric illness. On the other hand, stress management training is a proven method for helping speed recovery follow a heart attack.

What you can do

Learning to deal with stress effectively is a worthwhile effort, even if you already consider yourself capable of handling anything life sends your way. Many of the most common long-term stressors—family illness, recovery after injury, career pressures—often arise without warning and simultaneously. Stress management is particularly valuable if your family has a history of hypertension and other forms of heart disease.

Identify the cause. You may find that your stress arises from something that's easy to correct. A psychologist can help you define and analyze these stressors, and develop action plans for dealing with them.

Monitor your moods. If you feel stressed during the day, write down what caused it along your thoughts and moods. Again, you may find the cause to be less serious than you first thought.

Make time for yourself at least two or three times a week. Even ten minutes a day of "personal time" can help refresh your mental outlook and slow down your body's stress response systems. Turn off the phone, spend time alone in your room, exercise, or meditate to your favorite music.

Walk away when you're angry. Before you react, take time to mentally regroup by counting to 10. Then look at the situation again. Walking or other physical activities will also help you work off steam.

Analyze your schedule. Assess your priorities and delegate whatever tasks you can (e.g., order out dinner after a busy day, share household responsibilities). Eliminate tasks that are "shoulds" but not "musts."

Health Service Delivery and Quality

Poor people - in both developed and developing countries - experience more ill health and shorter life spans than their richer fellow citizens. Although people's health is influenced by a wide range of socio-economic and lifestyle factors, access to high quality and affordable health care and public health services makes a critical contribution to health status. Health services are failing poor people - with lower rates of child immunisation, skilled attendance at child birth, and TB and malaria treatment. It's also true that richer groups tend to benefit more from public sector subsidies to health care - hospitals in urban centres often receive disproportionate funds compared with primary care in poor rural areas. And in most poor countries which lack formal taxation and insurance systems, out-of-pocket payments are paid to both private and public providers, consume household income and assets, and contribute to impoverishment.

Improving service delivery to the poor involves all the major stakeholders in the health system - the policymakers in ministries of health, finance, and public administration, health service managers and workers, public and private providers and clients and communities themselves. Better access depends on a wide range of factors - on health policies, strategy and plans that prioritise health needs and set out revenue sources and resource requirements (including mechanisms to address inequalities), on motivated and properly trained and remunerated health workers, on infrastructure, drugs and equipment, on good referral links and communication, and – last but not least - on well-informed clients and their representative bodies.

Human resources for health

This **human resources for health dossier** offers practical up to date information about how to address human resource problems and issues, drawing upon evidence about what works, and identifying innovations in approaches, policy and practice. Developing countries have committed to achieving the Millennium Development Goals (MDGs). They will need to make the most effective use of all available resources to achieve the MDGs - this includes human resources. Many countries are improving their short and medium term financial planning and budgetary processes but in the past, few, if any, have given human resource management a similar focus. Now however human resources are being seen to be as crucial as money in improving services for poor people.

Planning for human resources

Planning the supply of and demand for human resources in healthcare is a neglected topic with little consensus on how it should be done. This section covers institutional and system considerations, gender, labour markets, civil service reform and more.

Human resources contribute to keep a healthy and productive population

Poverty Reduction Strategy Papers (PRSPs) are the major aid instrument in tackling poverty. They provide a mechanism for countries to have a clear focus on this central objective. The link to debt relief for heavily indebted poor countries (HIPC countries) provides real incentives for those very poor countries to make progress. In a World Bank publication, **Review of the Poverty Reduction Strategy Paper (PRSP) approach: main findings**, the authors recognise that there has been widespread acceptance of the PRSP approach and that PRSPs are leading to better informed decision-making. But they also recognise that strategies alone are not sufficient—they must be followed up by actions. The implementation of PRSPs cannot realistically be achieved without clear human resource plans.

An analysis of PRSP and HIPC documentation illustrates the problem. Most of the country based documentation, interim PRSPs, PRSPs and progress reports refer to human resource problems but few indicate how these problems are to be addressed. HIPC documentation is even weaker in this respect. Of course it can be argued that PRSP documents and HIPC documents are not designed to address these issues. But the fact remains that without complementary human resources strategies and plans; few PRSPs will achieve the improvements they envisage.

A Bank review of the HIPC/PRSP process in 18 African countries, **Enhancing human development in the HIPC/PRSP context: progress in the Africa region during 2000**, observes that it is important to take advantage of the specific opportunities presented by the HIPC initiative to strengthen the link between debt relief and improvements in social services and eventually, progress in poverty reduction. As well as the human resource issues of increasing coverage of services to poor people, there is an emphasis on improving efficiency of human resource utilisation. Often civil service reform has focused on reducing staffing numbers but the sustained improvements have been elusive. Civil Service Reform: a review of World Bank assistance provides recommendations on how the implementation of World Bank supported civil service reforms can be improved.

GLOSSARY

Active immunity see immunity, active.
active surveillance see surveillance, active.
age-adjusted mortality rate see mortality rate, age-adjusted.
agent a factor (e.g., a microorganism or chemical substance) or form of energy whose presence, excessive presence, or in the case of deficiency diseases, relative absence is essential for the occurrence of a disease or other adverse health outcome.
age-specific mortality rate see mortality rate, age-specific.
alternative hypothesis see hypothesis, alternative.
analytic epidemiology see epidemiology, analytic.
analytic study see study, analytic.
antibody any of a variety of proteins in the blood that are produced in response to an antigen as an immune response.

antigen any substance (e.g., a toxin or the surface of a microorganism or transplanted organ) recognized as foreign by the human body and that stimulates the production of antibodies.
applied epidemiology see epidemiology, applied.
arbovirus any of a group of viruses that are transmitted between hosts by mosquitoes, ticks, and other arthropods.
arithmetic mean see mean, arithmetic.
arithmetic-scale line graph see line graph, arithmetic-scale.
arthropod an organism that has jointed appendages and segmented external skeleton (e.g., flies, mosquitoes, ticks, or mites).
association the statistical relation between two or more events, characteristics, or other variables.
asymmetrical a type of distribution where the shape to the right and left of the central location is not the same. Often referred to as a skewed distribution; the mean, median, and mode of an asymmetrical distribution are not the same.
asymptomatic without symptoms.
attack rate a form of incidence that measures the proportion of persons in a population who experience an acute health event during a limited period (e.g., during an outbreak), calculated as the number of new cases of a health problem during an outbreak divided by the size of the population at the beginning of the period, usually expressed as a percentage or per 1,000 or 100,000 population (see also **incidence proportion**).
attack rate, secondary a measure of the frequency of new cases of a disease among the contacts of known patients.
attributable proportion see **proportion, attributable**.
attributable risk percent see **proportion, attributable**.
attribute a risk factor that is an intrinsic characteristic of the individual person, animal, plant, or other type of organism under study (e.g., genetic susceptibility, age, sex, breed, weight).
axis one of the dimensions of a graph in a rectangular graph, the x-axis is the horizontal axis, and the y-axis is the vertical axis.

Bar chart a visual display in which each category of a variable is represented by a bar or column bar charts are used to illustrate variations in size among categories.
bar chart, 100% component a stacked bar chart in which all bars or columns are the same length, and the measured axis represents 0%–100%.
bar chart, deviation a bar chart displaying either positive or negative differences from a baseline.
bar chart, grouped a bar chart displaying quantities of two variables, represented by adjoining bars or columns (i.e., a group) of categories of one variable, separated by space between groups.
bar chart, stacked a bar chart displaying quantities of two variables, represented by subdivided bars or columns (the subdivisions representing the categories of one variable) separated by space between bars or columns.
bias a systematic deviation of results or inferences from the truth or processes leading to such systematic deviation; any systematic tendency in the collection, analysis, interpretation,
publication, or review of data that can lead to conclusions that are systematically different from the truth. In epidemiology, does not imply intentional deviation.
bias, information systematic difference in the collection of data regarding the participants in a study (e.g., about exposures in a case-control study, or about health outcomes in a cohort study) that leads to an incorrect result (e.g., risk ratio or odds ratio) or inference.
bias, selection systematic difference in the enrollment of participants in a study that leads to an incorrect result (e.g., risk ratio or odds ratio) or inference.
bimodal having two data peaks.
biologic transmission see **transmission, biologic**.
birth cohort see **cohort, birth**.

birth rate, crude the number of live births during a specified period divided by the mid-period population, usually expressed per 1,000 population.
box plot a visual display that summarizes data by using a "box and whiskers" format to indicate the minimum and maximum values (ends of the whiskers), interquartile range (length of the box), and median (line through the box).

Carrier a person or animal that harbors the infectious agent for a disease and can transmit it to others, but does not demonstrate signs of the disease. A carrier can be asymptomatic (never indicate signs of the disease) or can display signs of the disease only during the incubation period, convalescence, or postconvalescence. The period of being a carrier can be short (a transient carrier) or long (a chronic carrier).
case an instance of a particular disease, injury, or other health conditions that meets selected criteria (see also **case definition**). Using the term to describe the person rather than the health condition is discouraged (see also **case-patient**).
case-control study see **study, case-control**.
case definition a set of uniformly applied criteria for determining whether a person should be identified as having a particular disease, injury, or other health condition. In epidemiology, particularly for an outbreak investigation, a case definition specifies clinical criteria and details of time, place, and person.
case-fatality rate (also called **case-fatality ratio)** the proportion of persons with a particular condition (e.g., patients) who die from that condition. The denominator is the number of persons with the condition; the numerator is the number of cause-specific deaths among those persons.
case, index the first case or instance of a patient coming to the attention of health authorities.
case-patient in a case-control study, a person who has the disease, injury, or other health condition that meets the case definition (see also **case**).
case, source the case or instance of a patient responsible for transmitting infection to others; the instance of a patient who gives rise to an outbreak or epidemic.
cause, component a factor that contributes to a sufficient cause (see **cause, sufficient**).
cause of disease a factor (e.g., characteristic, behavior, or event) that directly influences the occurrence of a disease. Reducing such a factor among a population should reduce occurrence of the disease.
cause, necessary a factor that must be present for a disease or other health problem to occur.
cause-specific mortality rate see **mortality rate, cause-specific**.
cause, sufficient a factor or collection of factors whose presence is always followed by the occurrence of a particular health problem.
census the enumeration of an entire population, usually including details on residence, age, sex, occupation, racial/ethnic group, marital status, birth history, and relationship to the head of household.
central location (also called **central tendency)** a statistical measurement to quantify the middle or the center of a distribution. Of the multiple ways to define central tendency, the most common are the mean, median, and mode.
chain of infection the progression of an infectious agent that leaves its reservoir or host through a portal of exit, is conveyed by a mode of transmission, and then enters through an appropriate portal of entry to infect a susceptible host.
"chartjunk" unnecessary or confusing visual elements in charts, illustrations, or graphs. The term was first used by Edward Tufte in his book, *The Visual Display of Quantitative Information* (1983).
class interval the span of values of a continuous variable that are grouped into a single category (see **class**), usually to create a frequency distribution for that variable.
class limits the values at the upper and lower ends of a class interval.
clinical criteria the medical features (e.g., symptoms, medical examination findings, and laboratory results) that are used in a case definition.
clinical disease a disease that has been manifested by its symptoms and features.
clinical trial see **trial, clinical**.

cluster an aggregation of cases of a disease, injury, or other health condition (particularly cancer and birth defects) in a circumscribed area during a particular period without regard to whether the number of cases is more than expected (often the expected number is not known).
cohort a well-defined group of persons who have had a common experience or exposure and are then followed up, as in a cohort study or prospective study, to determine the incidence of new diseases or health events.
cohort, birth a group of persons born during a particular period or year.
cohort study see **study, cohort**.
common-source outbreak see **outbreak, common-source**.
community immunity see **immunity, herd**.
community trial see **trial, community**.
comparison group a group in an analytic study (e.g., a cohort or case-control study) with whom the primary group of interest (exposed group in a cohort study or case-patients in a case-control study) is compared. The comparison group provides an estimate of the background or expected incidence of disease (in a cohort study) or exposure (in a case-control study).
confidence interval a range of values for a measure (e.g., rate or odds ratio) constructed so that the range has a specified probability (often, but not necessarily, 95%) of including the true value of the measure.
confidence limits the end points (i.e., the minimum and maximum values) of a confidence.
confounding the distortion of the association between an exposure and a health outcome by a third variable that is related to both.
contact exposure to a source of an infection; a person who has been exposed.
contact, direct exposure or transmission of an agent from a source to a susceptible host through touching (e.g., from a human host by kissing, sexual intercourse, or skin-to-skin contact) or from touching an infected animal or contaminated soil or vegetation.
contagious capable of being transmitted from one person to another by contact or close proximity.
contingency table a two-variable table of cross-tabulated data.
continuous variable see **variable, continuous**.
control in a case-control study, a member of the group of persons without the health problem under study (see also **comparison group** and **study, case-control**).
crude when referring to a rate, an overall or summary rate for a population, without adjustment.
crude birth rate see **birth rate, crude**.
crude death rate see **mortality rate, crude**.
crude mortality rate see **mortality rate, crude**.
cumulative frequency in a frequency distribution, the number or proportion of observations with a particular value and any smaller value.
cumulative frequency curve a plot of the cumulative frequency rather than the actual frequency for each class interval of a variable. This type of graph is useful for identifying medians and quartiles and other percentiles.

Death-to-case ratio the number of deaths attributed to a particular disease, injury, or other health condition during a specified period, divided by the number of new cases of that disease, injury, or condition identified during the same period.
decision analysis application of quantitative methods to decision-making.
decision tree a branching chart that represents the logical sequence or pathway of a clinical or public health decision.
demographic information personal characteristics of a person or group (e.g., age, sex, race/ethnicity, residence, and occupation) demographic information is used in descriptive epidemiology to characterize patients or populations.

dendrogram see **phylogenetic tree**.
denominator the lower portion of a fraction; used in calculating a ratio, proportion, or rate. For a rate, the denominator is usually the midinterval population.
dependent variable see **variable, dependent**.
descriptive epidemiology see **epidemiology, descriptive**.
determinant any factor that brings about change in a health condition or in other defined characteristics (see also **cause** and **risk factor**).
direct transmission see **transmission, direct**.
discrete variable (or **data**) see **variable (**or **data), discrete**.
distribution in epidemiology, the frequency and pattern of health-related characteristics and events in a population. In statistics, the frequency and pattern of the values or categories of a variable.
dose-response association between an exposure and health outcome that varies in a consistently increasing or decreasing fashion as the amount of exposure (dose) increases.
dot plot a visual display of the specific data points of a variable.
droplet nuclei the residue of dried droplets of infectious agents that is easily inhaled and exhaled and can remain suspended in air for relatively long periods or be blown over great distances.
droplet spread the direct transmission of an infectious agent by means of the aerosols produced in sneezing, coughing, or talking that travel only a short distance before falling to the ground.

Effect the result of a cause.
effectiveness the ability of an intervention or program to produce the intended or expected results in the field.
efficacy the ability of an intervention or program to produce the intended or expected results under ideal conditions.
efficiency the ability of an intervention or program to produce the intended or expected results with a minimum expenditure of time and resources.
EIS Epidemic Intelligence Service; CDC's 2-year training program in applied epidemiology for public health professionals (http://www.cdc.gov/eis).
endemic the constant presence of an agent or health condition within a given geographic area or population; can also refer to the usual prevalence of an agent or condition.
environmental factor an extrinsic factor (e.g., geology, climate, insects, sanitation, or health services) that affects an agent and the opportunity for exposure.
epidemic the occurrence of more cases of disease, injury, or other health condition than expected in a given area or among a specific group of persons during a particular period. Usually, the cases are presumed to have a common cause or to be related to one another in some way (see also **outbreak**).
epidemic curve a histogram that displays the course of an outbreak or epidemic by plotting the number of cases according to time of onset.
epidemic period the time span of an outbreak or epidemic.
epidemiologic triad the traditional model of infectious disease causation having three components: an external agent, a susceptible host, and an environment that brings the host and agent together so that disease occurs.
epidemiology the study of the distribution and determinants of health conditions or events among populations and the application of that study to control health problems.
epidemiology, analytic the aspect of epidemiology concerned with why and how a health problem occurs. Analytic epidemiology uses comparison groups to provide baseline or expected values so that associations between exposures and outcomes can be quantified and hypotheses about the cause of the problem can be tested (see also **study, analytic)**.
epidemiology, applied the application or practice of epidemiology to control and prevent health problems.

epidemiology, descriptive the aspect of epidemiology concerned with organizing and summarizing data regarding the persons affected (e.g., the characteristics of those who became ill), time (e.g., when they become ill), and place (e.g., where they might have been exposed to the cause of illness).
epidemiology, field applied epidemiology (i.e., the application or practice of epidemiology to control and prevent health problems), particularly when the epidemiologist(s) must travel to and work in the community in which the health problem is occurring or has occurred.
evaluation systematic and objective examination of activities to determine their relevance, effectiveness, and impact.
excess risk risk difference, calculated as the risk among the exposed group minus the risk among the unexposed group.
experimental study see **study, experimental**.
exposed group a group whose members have had contact with a suspected cause of, or possess a characteristic that is a suspected determinant of, a particular health problem.
exposure having come into contact with a cause of, or possessing a characteristic that is a determinant of, a particular health problem.

False-negative a negative test result for a person who actually has the condition similarly, a person who has the disease (perhaps mild or variant) but who does not fit the case definition, or a patient or outbreak not detected by a surveillance system.
false-positive a positive test result for a person who actually does not have the condition. Similarly, a person who does not have the disease but who nonetheless fits the case definition, or a patient or outbreak erroneously identified by a surveillance system.
field epidemiology see **epidemiology, field**.
follow-up study see **study, cohort**.
fomite an inanimate object that can be the vehicle for transmission of an infectious agent (e.g., bedding, towels, or surgical instruments).
forest plot a graph that displays the point estimates and confidence intervals of individual studies included in a meta-analysis or systematic review as a series of parallel lines.
frequency the amount or number of occurrences of an attribute or health outcome among a population.
frequency distribution a complete summary of the frequencies of the values or categories of a variable, often displayed in a two-column table with the individual values or categories in the left column and the number of observations in each category in the right column.
frequency polygon a graph of a frequency distribution in which values of the variable are plotted on the horizontal axis, and the number of observations are plotted on the vertical axis. Data points are plotted at the midpoints of the intervals and are connected with straight lines.

Geometric mean see **mean, geometric**.
graph a visual display of quantitative data arranged on a system of coordinates.

Health a state of complete physical, mental, and social well-being and not merely the absence of disease or other infirmity.
health indicator any of a variety of measures (e.g., mortality rate) that indicate the state of health of a given population.
health information system a combination of health statistics from different sources. Data from these systems are used to learn about health status, health care, provision and use of services, and the impact of services and programs on health.
healthy worker effect the observation that employed persons generally have lower mortality rates than the general population, because persons with severe, disabling disease (who have higher mortality rates) tend to be excluded from the workforce.
herd immunity see **immunity, herd**.

high-risk group a group of persons whose risk for a particular disease, injury, or other health condition is greater than that of the rest of their community or population.
HIPAA the Health Insurance Portability and Accountability Act, enacted in 1996, which addresses the privacy of a person's medical information as well as postemployment insurance and other health-related concerns.
histogram a visual representation of the frequency distribution of a continuous variable. The class intervals of the variable are grouped on a linear scale on the horizontal axis, and the class frequencies are grouped on the vertical axis. Columns are drawn so that their bases equal the class intervals (i.e., so that columns of adjacent intervals touch), and their heights correspond to the class frequencies.
host a person or other living organism that is susceptible to or harbors an infectious agent under natural conditions.
host factor an intrinsic factor (e.g., age, race/ethnicity, sex, or behaviors) that influences a person's exposure, susceptibility, or response to an agent.
hyperendemic the constant presence at high incidence and prevalence of an agent or health condition within a given geographic area or population.
hypothesis a supposition, arrived at from observation or reflection, that leads to refutable predictions; any conjecture cast in a form that will allow it to be tested and refuted.
hypothesis, alternative the supposition that an exposure is associated with the health condition under study. The alternative is adopted if the null hypothesis (see **hypothesis, null**) proves implausible.
hypothesis, null the supposition that two (or more) groups do not differ in the measure of interest (e.g., incidence or proportion exposed); the supposition that an exposure is not associated with the health condition under study, so that the risk ratio or odds ratio equals 1. The null hypothesis is used in conjunction with statistical testing.

Immunity, active resistance developed in response to an antigen (i.e., an infecting agent or vaccine), usually characterized by the presence of antibody produced by the host.
immunity, herd the resistance to an infectious agent of an entire group or community (and, in particular, protection of susceptible persons) as a result of a substantial proportion of the population being immune to the agent. Herd immunity is based on having a substantial number of immune persons, thereby reducing the likelihood that an infected person will come in contact with a susceptible one among human populations, also called **community immunity**.
immunity, passive immunity conferred by an antibody produced in another host This type of immunity can be acquired naturally by an infant from its mother or artificially by administration of an antibody-containing preparation (e.g., antiserum or immune globulin).
incidence a measure of the frequency with which new cases of illness, injury, or other health condition occurs among a population during a specified period.
incidence proportion the fraction of persons with new cases of illness, injury, or other health condition during a specified period, calculated as the number of new cases divided by the size of the population at the start of the study period (see also **attack rate**).
incidence rate a measure of the frequency with which new cases of illness, injury, or other health condition occur, expressed explicitly per a time frame. Incidence rate is calculated as the number of new cases over a specified period divided either by the average population (usually mid-period) or by the cumulative person-time the population was at risk.
incubation period the time interval from exposure to an infectious agent to the onset of symptoms of an infectious disease.
independent variable see **variable, independent**.
index case see **case, index**.
indirect transmission see **transmission, indirect**.
individual data values or observations from each record (also called raw data).
infant mortality rate see **mortality rate, infant**.

infection invasion of the body tissues of a host by an infectious agent, whether or not it causes disease.
infectivity the ability of an infectious agent to cause infection, measured as the proportion of persons exposed to an infectious agent who become infected.
information bias see bias, information.
interquartile range a measure of spread representing the middle 50% of the observations, calculated as the difference between the third quartile (75th percentile) and the first quartile (25th percentile).
isolation the separation of infected persons to prevent transmission to susceptible ones. Isolation refers to separation of ill persons; **quarantine** refers to separation of potentially exposed but well persons.

Latency period the time from exposure to a causal agent to onset of symptoms of a (usually noninfectious) disease (see also **incubation period**).
life expectancy a statistical projection of the average number of years a person of a given age is expected to live, if current mortality rates continue to apply.
line graph, arithmetic-scale a graph that displays patterns or trends by plotting the frequency (e.g., number, proportion, or rate) of a characteristic or event during some variable, usually time. The y-axis, measuring frequency, uses an arithmetic scale.
line graph, semilogarithmic-scale a graph that displays patterns or trends by plotting the frequency (e.g., number, proportion, or rate) of a characteristic or event during some variable, usually time. The y-axis, measuring frequency, uses a logarithmic scale.
line listing a type of epidemiologic database, organized similar to a spreadsheet with rows and columns in which information from cases or patients are listed each column represents a variable, and each row represents an individual case or patient.
logarithmic transformation conversion of nominal or ordinal data to logarithmic data. The purpose is to examine rate of change instead of amount of change only.

Map, area (shaded, choropleth) a visual display of the geographic pattern of a health problem, in which a marker is placed on a map to indicate where each affected person lives, works, or might have been exposed.
mean (or average) commonly called the average; it is the most common measure of central tendency.
mean, arithmetic the measure of central location, commonly called the average, calculated by adding all the values in a group of measurements and dividing by the number of values in the group.
mean, geometric the mean, or average, of a set of data measured on a logarithmic scale.
measure of association a quantified relationship between exposure and a particular health problem (e.g., risk ratio, rate ratio, and odds ratio).
measure of central location a central value that best represents a distribution of data. Common measures of central location are the mean, median, and mode also called the measure of central tendency.
measure of dispersion see **measure of spread**.
measure of spread a measure of the distribution of observations out from its central value. Measures of spread used in epidemiology include the interquartile range, variance, and the standard deviation.
measurement scale the complete range of possible values for a measurement.
mechanical transmission see **transmission, mechanical**.
median the measure of central location that divides a set of data into two equal parts, above and below which lie an equal number of values (see also **measure of central location**).
medical surveillance see **surveillance, medical**.
midrange the halfway point, or midpoint, in a set of observations. For the majority of data, the midrange is calculated by adding the smallest observation and the largest observation and dividing by two. The midrange is usually calculated as an intermediate step in determining other measures.
mode the most frequently occurring value in a set of observations (see also **measure of central location**).

mode of transmission the manner in which an agent is transmitted from its reservoir to a susceptible host (see also **transmission**).
morbidity disease; any departure, subjective or objective, from a state of physiological or psychological health and well-being.
mortality death.
mortality rate a measure of the frequency of occurrence of death among a defined population during a specified time interval.
mortality rate, age-adjusted a mortality rate that has been statistically modified to eliminate the effect of different age distributions among different populations.
mortality rate, age-specific a mortality rate limited to a particular age group, calculated as the number of deaths among the age group divided by the number of persons in that age group, usually expressed per 100,000.
mortality rate, cause-specific the mortality rate from a specified cause, calculated as the number of deaths attributed to a specific cause during a specified time interval among a population divided by the size of the midinterval population.
mortality rate, crude a mortality rate from all causes of death for an entire population, without adjustment.
mortality rate, infant the mortality rate for children aged <1 year, calculated as the number of deaths reported among this age group during a given period divided by the number of live births reported during the same period, and expressed per 1,000 live births. Infant mortality rate is a universally accepted indicator of the health of a nation’s population and the adequacy of its health-care system.
mortality rate, neonatal the mortality rate for children from age birth up to, but not including, 28 days. In calculating neonatal mortality rates, the numerator is the number of deaths among this age group during a given period, and the denominator is the number of live births reported during the same period. The neonatal mortality rate is usually expressed per 1,000 live births.
mortality rate, postneonatal the mortality rate for children from age 28 days up to, but not including, 1 year. In calculating postneonatal mortality rates, the numerator is the number of deaths among this age group during a given period, and the denominator is the number of live births during the same period.. The postneonatal mortality rate is usually expressed per 1,000 live births.
mortality rate, race/ethnic-specific a mortality rate limited to a specified racial or ethnic group both numerator and denominator are limited to that group.
mortality rate, sex-specific a mortality rate among either males or females.

Natural history of disease the progression of a disease process in a person from the time it begins to the time it resolves, in the absence of treatment.
NCHS The National Center for Health Statistics, the US governmental organization responsible for national vital statistics and multiple national health surveys. Organizationally, NCHS is a component of the Centers for Disease Control and Prevention, one of the agencies of the US Department of Health and Human Services.
NHANES The National Health and Nutrition Examination Survey, a representative survey of the civilian, noninstitutionalized US population conducted by the National Center for Health Statistics, designed to (1) estimate the proportion of the US population and designated groups with selected disease and risk factors; (2) monitor trends in selected behaviors, exposures, and diseases; and (3) study the associations among diet, nutrition, and health.
necessary cause see **cause, necessary**.
neonatal mortality rate see **mortality rate, neonatal**.
nominal scale see **scale, nominal**.
normal curve the bell-shaped curve that results when a normal distribution is graphed.
normal distribution a distribution represented as a bell shape, symmetrical on both sides of the peak, which is simultaneously the mean, median, and mode, and with both tails extending to infinity.

notifiable disease a disease that, by law, must be reported to public health authorities upon diagnosis.
null hypothesis see **hypothesis, null**.
numerator the upper portion of a fraction (see also **denominator**).

Observational study see **study, observational**.
odds ratio a measure of association used in comparative studies, particularly case-control studies, that quantifies the association between an exposure and a health outcome; also called the cross-product ratio.
ordinal scale see **scale, ordinal**.
outbreak the occurrence of more cases of disease, injury, or other health condition than expected in a given area or among a specific group of persons during a specific period. Usually, the cases are presumed to have a common cause or to be related to one another in some way. Sometimes distinguished from an epidemic as more localized, or the term less likely to evoke public panic (see also **epidemic**).
outbreak, common-source an outbreak that results from persons being exposed to the same harmful influence (e.g., an infectious agent or toxin). The exposure period can be brief or can extend over days, weeks, or longer, with the exposure being either intermittent or continuous.
outbreak, point-source a common source outbreak in which the exposure period is relatively brief so that all cases occur within one incubation period.
outbreak, propagated an outbreak that spreads from person to person rather than from a common source.
outcome(s) any or all of the possible results that can stem from exposure to a causal factor or from preventive or therapeutic interventions; all identified changes in health status that result from the handling of a health problem.
outlier a value substantively or statistically different from all (or approximately all) of the other values in a distribution.

Pandemic an epidemic occurring over a widespread area (multiple countries or continents) and usually affecting a substantial proportion of the population.
***P* value** the probability of observing an association between two variables or a difference between two or more groups as large or larger than that observed, if the null hypothesis were true. Used in statistical testing to evaluate the plausibility of the null hypothesis (i.e., whether the observed association or difference plausibly might have occurred by chance).
passive immunity see **immunity, passive**.
passive surveillance see **surveillance, passive**.
pathogenicity the ability of an agent to cause disease after infection, measured as the proportion of persons infected by an agent who then experience clinical disease.
percentile a set of cut points used to divide a distribution or a set of ranked data into 100 parts of equal area with each interval between the points containing 1/100 or 1% of the observations. Forexample, the 5th percentile is a cut point with 5% of the observations below it and the remaining 95% above it.
period prevalence see **prevalence, period**.
person-time rate the incidence rate calculated as the number of new cases among a population divided by the cumulative person-time of that population, usually expressed as the number of events per persons per unit of time.
person-time the amount of time each participant in a cohort study is observed and disease-free, often summed to provide the denominator for a person-time rate.
phylogenetic tree a branching chart that indicates the evolutionary lineage or genetic relatedness of organisms.
pie chart a circular graph of a frequency distribution in which each segment of the pie is proportional in size to the frequency of corresponding category.
point prevalence see **prevalence, point**.

point-source outbreak see **outbreak, point-source**.
population the total number of inhabitants of a geographic area or the total number of persons in a particular group (e.g., the number of persons engaged in a certain occupation).
population pyramid a graphical display of the age-sex distribution of a population, constructed with a horizontal histogram of the age distribution of males pointing to the left, and the corresponding horizontal histogram of age distribution of females pointing to the right.
portal of entry a pathway into the host that gives an agent access to tissue that will allow it to multiply or act.
portal of exit a pathway by which an agent can leave its host.
postneonatal mortality rate see **mortality rate, postneonatal**.
predictive value positive the proportion of cases identified by a test, reported by a surveillance system, or classified by a case definition that are true cases, calculated as the number of true positives divided by the number of true-positives plus false-positives.
prevalence the number or proportion of cases or events or attributes among a given population.
prevalence rate the proportion of a population that has a particular disease, injury, other health condition, or attribute at a specified point in time (point prevalence) or during a specified period (period prevalence).
prevalence, period the amount of a particular disease, chronic condition, or type of injury present among a population at any time during a particular period.
prevalence, point the amount of a particular disease, chronic condition, or type of injury present among a population at a single point in time.
privacy rule a set of regulations based on the Health Insurance Portability and Accountability Act to protect the privacy of individually identifiable health information.
propagated outbreak see **outbreak, propagated**.
proportion a ratio in which the numerator is included in the denominator; the ratio of a part to the whole, expressed as a "decimal fraction" (e.g., 0 2), a fraction (1/5), or a percentage (20%).
proportion, attributable a measure of the impact of a causative factor on the public health; the proportion of a health state or event among exposed persons that can be attributed to the exposure also called attributable risk percent.
proportionate mortality the proportion of deaths among a population attributable to a particular cause during a selected period. Each cause of death is expressed as a percentage of all deaths, and the sum of the proportionate mortality for all causes must equal 100%. These proportions are not mortality rates because, in proportionate mortality, the denominator is all deaths instead of the population among whom the deaths occurred.
prospective study see **study, prospective**.

Quarantine the separation of well persons who have been exposed or are suspected to have been exposed to a communicable disease, to monitor for illness and to prevent potential transmission of infection to susceptible persons during the incubation period. Quarantine refers to separation of potentially exposed but well persons; **isolation** refers to separation of ill persons.

Race/ethnic-specific mortality rate see **mortality rate, race/ethnic-specific**.
random sample see **sample, random**.
range in statistics, the difference between the largest and smallest values in a distribution; in common use, the span of values from smallest to largest.
rate an expression of the relative frequency with which an event occurs among a defined population per unit of time, calculated as the number of new cases or deaths during a specified period divided by either person-time or the average (midinterval) population. In epidemiology, it is often used more casually to refer to proportions that are not truly rates (e.g., attack rate or case-fatality rate).

rate ratio a measure of association that quantifies the relation between an exposure and a health outcome from an epidemiologic study, calculated as the ratio of incidence rates or mortality rates of two groups.
ratio the relative size of two quantities, calculated by dividing one quantity by the other.
record in a line listing, each row is a record or observation. A record represents data related to a single case.
relative risk a general term for measures of association calculated from the data in a two-by-two table, including risk ratio, rate ratio, and odds ratio (see **risk ratio**).
representative sample see **sample, representative**.
reservoir the habitat in which an infectious agent normally lives, grows, and multiplies, which can include humans, animals, or the environment.
retrospective study see **study, retrospective**.
risk the probability that an event will occur (e.g., that a person will be affected by, or die from, an illness, injury, or other health condition within a specified time or age span).
risk factor an aspect of personal behavior or lifestyle, an environmental exposure, or a hereditary characteristic that is associated with an increase in the occurrence of a particular disease, injury, or other health condition.
risk ratio a measure of association that quantifies the association between an exposure and a health outcome from an epidemiologic study, calculated as the ratio of incidence proportions of two groups.

Sample a selected subset of a population a sample can be random or nonrandom and representative or nonrepresentative.
sample, random a sample of persons chosen in such a way that each one has the same (and known) probability of being selected.
sample, representative a sample whose characteristics correspond to those of the original or reference population.
scale, interval a measurement scale consisting of quantitative categories whose values are measured on a scale of equally spaced units, but without a true zero point (e.g., date of birth).
scale, nominal a measurement scale consisting of qualitative categories whose values have no inherent statistical order or rank (e.g., categories of race/ethnicity, religion, or country of birth).
scale, ordinal a measurement scale consisting of qualitative categories whose values have a distinct order but no numerical distance between their possible values (e.g., stage of cancer, I, II, III, or IV).
scale, ratio a measurement scale consisting of quantitative categories whose values are intervals with a true zero point (e.g., height in centimeters or duration of illness).
scatter diagram (or **scattergram**) a graphical display of the association between two variables in which a dot is plotted on the graph for each set of paired values for two continuous variables, with one variable plotted on the horizontal axis, and the other plotted on the vertical axis.
seasonality change in physiologic status or in the occurrence of a disease, chronic condition, or type of injury that conforms to a regular seasonal pattern.
secondary attack rate see **attack rate, secondary**.
secular trend see **trend, secular**.
selection bias see **bias, selection**.
semilogarithmic-scale line graph see **line graph, semilogarithmic-scale**
sensitivity the ability of a test, case definition, or surveillance system to identify true cases; the proportion of people with a health condition (or the proportion of outbreaks) that are identified by a screening test or case definition (or surveillance system).
sentinel surveillance see **surveillance, sentinel**.
sex-specific mortality rate see **mortality rate, sex-specific**.
skewed a distribution that is not symmetrical.

source (of infection) the person, animal, object, or substance from which an infectious agent is transmitted to a host.
source case see **case, source**.
specificity the ability or a test, case definition, or surveillance system to exclude persons without the health condition of interest; the proportion of persons without a health condition that are correctly identified as such by a screening test, case definition, or surveillance system.
spectrum of illness the range of manifestations a disease process can take (e.g., from asymptomatic to mild clinical illness to severe illness and death).
sporadic an event that occurs infrequently and irregularly.
spot map a visual display of the geographic pattern of a health problem, in which a marker is placed on a map to indicate where each affected person lives, works, or might have been exposed.
standard deviation a statistical summary of how dispersed the values of a variable are around its mean, calculated as the square root of the variance.
standard error (of the mean) the standard deviation of a theoretical distribution of sample means of a variable around the true population mean of that variable. Standard error is computed as the standard deviation of the variable divided by the square root of the sample size.
statistical inference generalizations developed from sample data, usually with calculated degrees of uncertainty.
statistical significance the measure of how likely it is that a set of study results could have occurred by chance alone. Statistical significance is based on an estimate of the probability of the observed or a greater degree of association between independent and dependent variables occurring under the null hypothesis (see also ***P* value**).
study, analytic a study, usually observational, in which groups are compared to identify and quantify associations, test hypotheses, and identify causes. Two common types are cohort studies and case-control studies.
study, case-control an observational analytic study that enrolls one group of persons with a certain disease, chronic condition, or type of injury (case-patients) and a group of persons without the health problem (control subjects) and compares differences in exposures, behaviors, and other characteristics to identify and quantify associations, test hypotheses, and identify causes.
study, cohort an observational analytic study in which enrollment is based on status of exposure to a certain factor or membership in a certain group. Populations are followed, and disease, death, or other health-related outcomes are documented and compared. Cohort studies can be either prospective or retrospective.
study, cross-sectional a study in which a sample of persons from a population are enrolled and their exposures and health outcomes are measured simultaneously; a survey.
study, experimental a study in which the investigator specifies the type of exposure for each person (clinical trial) or community (community trial) then follows the persons' or communities' health status to determine the effects of the exposure.
study, observational a study in which the investigator observes rather than influences exposure and disease among participants. Case-control and cohort studies are observational studies (see also **study, experimental**).
study, prospective an analytic study in which participants are enrolled before the health outcome of interest has occurred.
study, retrospective an analytic study in which participants are enrolled after the health outcome of interest has occurred. Case-control studies are inherently retrospective.
subclinical without apparent symptoms.
surveillance, active public health surveillance in which the health agency solicits reports.
surveillance, medical monitoring of a person who might have been exposed to an infectious, chemical, radiologic, or other potentially causal agent, for the purpose of detecting early symptoms.

surveillance, passive public health surveillance in which data are sent to the health agency without prompting.
surveillance, sentinel a surveillance system that uses a prearranged sample of sources (e.g., physicians, hospitals, or clinics) who have agreed to report all cases of one or more notifiable diseases.
surveillance, syndromic (1) the monitoring of the frequency of illnesses with a specified set of clinical features among a given population without regard to the specific diagnoses, if any, that are assigned to them by clinicians. (2) A system for early detection of outbreaks whereby health department staff, assisted by automated acquisition of data routinely collected for other purposes and computer generation of statistical signals, monitor disease indicators, particularly those associated with possible terrorism-related biologic and chemical agents, continually or at least daily to detect outbreaks earlier than would otherwise be possible with traditional public health methods.
survey a systematic canvassing of persons to collect information, often from a representative sample of the population.
survival curve a line graph that begins with 100% of the study population and displays the percentage of the population still surviving at successive points in time. A survival curve can also be used to depict freedom from a health problem, complication, or another endpoint.
symmetrical a type of distribution where the shapes to the right and left of the central location are the same. Normal, bell-shaped distributions are symmetrical; the mean, median, and mode are the same.
symptom any indication of disease noticed or felt by a patient.
syndrome a combination of symptoms characteristic of a disease or health condition; sometimes refers to a health condition without a clear cause (e.g., chronic fatigue syndrome).
syndromic surveillance see **surveillance, syndromic**.

Table an arrangement of data in rows and columns. In epidemiology, the data are usually summaries of the frequency of occurrence of an event or characteristic occurring among different groups.
table shell a table that is completely drawn and labeled but contains no data.
table, two-by-two a two-variable table with cross-tabulated data, in which each variable has only two categories. Usually, one variable represents a health outcome, and one represents an exposure or personal characteristic.
transmission (of infection) any mode or mechanism by which an infectious agent is spread to a susceptible host.
transmission, airborne transfer of an agent suspended in the air, considered a type of indirect transmission.
transmission, **biologic** indirect transmission by a vector in which the infectious agent undergoes biologic changes inside the vector as part of its life cycle before it is transmitted to the host (see also **transmission, mechanical**).
transmission, direct immediate transfer of an agent from a reservoir to a host by direct contact or droplet spread.
transmission, indirect transfer of an agent from a reservoir to a host either by being suspended in air particles (airborne), carried by an inanimate objects (vehicleborne), or carried by an animate intermediary (vectorborne).
transmission, mechanical indirect transmission by a vector in which the infectious agent does not undergo physiologic changes inside the vector (see also **transmission, biologic**).
transmission, vectorborne transmission of an agent by a living intermediary (e.g., tick, mosquito, or flea); considered a type of indirect transmission.
transmission, vehicleborne transmission of an agent by an inanimate object; considered a type of indirect transmission; includes foodborne and waterborne transmission.
trend movement or change in frequency over time, usually upwards or downwards**.**
trend, secular changes occurring over a substantial period, generally years or decades.

trial, clinical an experimental study that uses data from individual persons. The investigator specifies the type of exposure for each study participant and then follows each person's health status to determine the effects of the exposure.
trial, community an experimental study that uses data from communities. The investigator specifies the type of exposure for each community and then follows the communities' health status to determine the effects of the exposure.
trial, randomized clinical a clinical trial in which persons are randomly assigned to exposure or treatment groups.
two-by-two table see **table, two-by-two.**

Validity the degree to which a measurement, questionnaire, test, or study or any other datacollection tool measures what it is intended to measure.
variable any characteristic or attribute that can be measured and can have different values.
variable (or **data**), **discrete** a variable that is limited to a finite number of values; data for such a variable.
variable, continuous a variable that has the potential for having an infinite number of values along a continuum (e.g., height and weight).
variable, dependent in a statistical analysis, a variable whose values are a function of one or more other variables.
variable, independent an exposure, risk factor, or other characteristic being observed or measured that is hypothesized to influence an event or manifestation (the dependent variable).
variance a measure of the spread in a set of observations, calculated as the sum of the squares of deviations from the mean, divided by the number of observations minus 1 (see also **standard deviation**).
vector a living intermediary that carries an agent from a reservoir to a susceptible host (see also **transmission, biologic** and **transmission, mechanical**) (e.g., mosquitoes, fleas, or ticks).
vehicle an inanimate object that can carry an agent from a reservoir to a susceptible host (e.g., food, water, blood products, and bedding) (see also **transmission, indirect**).
virulence the ability of an infectious agent to cause severe disease, measured as the proportion of persons with the disease who become severely ill or die.
vital statistics systematically tabulated data about recorded births, marriages, divorces, and deaths.

X-axis the horizontal axis of a rectangular graph, usually displaying the independent variable (e.g., time).

Y-axis the vertical axis of a rectangular graph, usually displaying the dependent variable (e.g., frequency — number, proportion, or rate).
years of potential life lost (YPLL) a measure of the impact of premature death on a population, calculated as the sum of the differences between a predetermined minimally acceptable age (e.g., 65 years or current life expectancy) and the age at death for everyone who died earlier than that age.

Zoonosis an infectious disease that is transmissible from animals to humans.

References

1. WHO Definition of Health Preamble to the Constitution of the World Health Organization as adopted by the International Health Conference, 1946
2. http://www.globalhandwashing.org/Publications/Lit_review.htm
3. *The Solid Facts: Social Determinants of Health* edited by Richard Wilkinson and Michael Marmot, WHO, 2003
4. SYED M. AHMED, BARBRA BECK, CHERYL A. MAURANA & GAIL NEWTON Overcoming Barriers to Effective Community-Based Participatory Research in US Medical Schools. Education for Health, Vol. 17, No. 2, July 2004, 141 – 151.
5. Asia-Pacific Academic Consortium for Public Health. "Asia-Pacific Journal of Public Health". http://aph.sagepub.com/. Retrieved on 12 January 2009.
6. Annual Reviews. "Annual Reviews of Public Health". http://www.annualreviews.org/default.aspx. Retrieved on 29 December 2008.
7. Public Health Association of Australia. "Australian and New Zealand Journal of Public Health". http://blackwellpublishing.com/journal.asp?ref=1326-0200&site=1. Retrieved on 12 January 2009.
8. BioMed Central. "BMC Public Health". http://www.biomedcentral.com/bmcpublichealth/. Retrieved on 12 January 2009.
9. Canadian Public Health Association. "Canadian Journal of Public Health". http://www.cpha.ca/en/cjph.aspx. Retrieved on 29 December 2008.
10. Czech Republic National Institute of Public Health. "Central European Journal of Public Health". http://www1.szu.cz/svi/cejph/index_en.php. Retrieved on 12 January 2009.
11. Scottish Centre for Infection and Environmental Health. "Communicable Disease and Public Health". http://www.hpa.org.uk/cdph/index.html. Retrieved on 12 January 2008.
12. Oxford University Press. "Journal of Public Health". http://jpubhealth.oxfordjournals.org/. Retrieved on 29 December 2008.
13. SAGE Publications. "Scandinavian Journal of Public Health". http://sjp.sagepub.com/. Retrieved on 29 December 2008.
14. Harvard School of Public Health. "Harvard Public Health Review". http://www.hsph.harvard.edu/news/hphr/. Retrieved on 12 January 2008.
15. Control of Communicable Diseases Manual edited by James B. Chin, APHA, 2000
16. Encyclopedia of public health edited by Lester Breslow, Macmillan Reference 2002