Chapter 1
Introduction to Health Statistics

Key Terms

Ambulatory care
Census
Descriptive statistics
Encounter
Home health (HH)
Hospice
Hospital
Inferential statistics
Inpatient
Inpatient census
Managed care organization (MCO)
Mean
Nursing facility
Outpatient
Primary data source
Research
Secondary data source
Visit
Vital statistics

Objectives

At the conclusion of this chapter, you should be able to:

- Define statistics
- Appreciate the need to study healthcare statistics
- Differentiate between descriptive and inferential statistics
- Recognize where statistics in healthcare originate
- Identify the users of healthcare statistics

Statistics

The term *statistic* refers to a number computed from a larger collection of numbers which collectively constitute a sample of data—for instance, the average value (or **mean**) of a variable belonging to a sample of data. A sample is a small part (a subset) of a larger group of data (a population). The term *statistics* is more broadly defined as a branch of mathematics concerned with collecting, organizing, summarizing, and analyzing data.

Origins of the Term

Originally, the term *statistics* referred to the collection of information about and for the "state." The word comes from the Italian word *stato,* meaning "state." One need only think of our own government and its statistics-collecting organizations, such as the Bureau of Labor Statistics, the National Center for Health Statistics (NCHS), the Administration on Aging, the Centers for Disease Control and Prevention (CDC), the Centers for Medicare and Medicaid Services (CMS), and so on.

Reasons for Studying Statistics

Statistics is really about decision making. In every area of our lives, we are expected to make decisions. To do that, we must have some information. Because information is often incomplete in healthcare settings, we must learn to estimate the characteristics of a complete population using statistics.

All organizations keep statistics in order to make decisions about their business. For example, an organization may use statistics to determine its markets, that is, to identify who is using its services and how it can increase those services. Libraries need information on what the public is reading in order to satisfy public interest and needs. Educational institutions need statistics about types of employment that students can get after graduation in order to improve the curriculum for future students.

Healthcare Operations Needs

In the healthcare industry, there are compelling reasons to collect and analyze data. For example, statistics kept on activities in the healthcare facility indicate why patients come to the facility and the costs of taking care of them. Patient care statistics and studies on performance can show the quality of care provided. Many accrediting agencies require a data analysis system as part of accreditation, and many third-party payers require facilities to collect performance data. Administrators also may use statistics for prioritizing needed services and to point to areas where efficiency and effectiveness might be increased. Additionally, healthcare facilities are interested in the types of patients they have with respect to their diagnoses in order to maintain the optimum physician specialty mix they need.

Public Health Needs

The government also needs to maintain statistics on and about the population in order to provide services. For example, the CDC is recognized as the lead agency responsible for protecting the health of the United States population by providing credible information to help individuals make the right healthcare decisions.

To obtain the knowledge they need, organizations first must have data. Data are unprocessed facts and figures that can be deliberately selected, processed, and organized to provide useful information. This leads, in turn, to facts, which are pieces of information

representing the truth. And facts lead to an improved understanding of the original, unprocessed information. Improved understanding gives individuals the power to make better decisions. The sequence, then, is as follows:

Data → Information → Facts → Improved understanding → Better decision making

Health statistics provide information about the health of people and their utilization of healthcare services. Examples of healthcare statistics include average longevity, birth rates, death rates, incidence of a particular disease in a state or the United States as a whole, and the frequency of usage of a particular type of service within a hospital.

Other common uses of statistics may include divorce rates in a country, accident rates in a state, crime statistics in a city, percent of HIV carriers in the world, and even the standings of candidates in a political race. To serve each of these and many other purposes, the figures used in the statistics must be relevant and reliable. "To be relevant" refers to the applicability of the statistics. "To be reliable" means that there is some consistency of results. For example, if your instructor graded a test and then asked another instructor to grade the same test and they both had the same results, the results could be said to be reliable.

Descriptive Statistics versus Inferential Statistics

The primary focus of descriptive statistics is to organize and describe the features of data in a study. **Descriptive statistics** describe what the data show about the characteristics of a sample; in other words, they tell us information about a particular population. For example, we may need to know the average age of our patients or which service is used most in our facility. Both these statistics are used to describe data. **Inferential statistics,** on the other hand, help us make inferences or guess about a larger group of data by drawing conclusions from a small group of data. The smaller group of data is often called a sample, which is a portion of the larger group, or population. The results obtained from the sample, if gathered carefully, are assumed to be typical of the entire population.

Sources of Healthcare Statistics

Healthcare information is derived from both primary and secondary data sources. Statistics usually come from a primary data source, or firsthand documents. In healthcare, ***primary data source*** refers to the record that was developed by healthcare professionals in the process of providing care. ***Secondary data sources*** are data derived from primary sources. Secondary data sources are facility-specific. For example, the disease and operation index is a secondary source of information. All the information in the index comes from a primary data source—the health record. Registries are also considered secondary data sources. For example, information from the medical record may be used to create a cancer or trauma registry.

The health record is one of the most important primary sources of health statistics because it contains most of the health facts about patients.

Another source of health information is the census. A **census** is defined as a count of a particular population. The US government conducts a census, that is, a count of the people residing in the United States and their location. This kind of census is called a population census. "Census" comes from the Latin word *censere,* which means "to assess or to tax." The Romans started taking a census in order to make a register of the people and

their property. One reason was to identify individuals who could serve in their armies and another was to place a value on their property so they could be taxed.

The US Constitution requires that a population census be taken every 10 years. The main reason is to determine the number of congressional representatives in the states. Over the years, Congress has authorized gathering more information about each person. The census now is used in many ways. For instance, the amount of government monies given to school districts is based partly on the number of children in a particular district. Congress also has requested that other types of censuses be taken periodically. These include a census of the types of businesses and industries in the US, for example, farms and fisheries, construction, foreign trade, manufacturing, and energy companies. All census information is available to the public.

Healthcare facilities also have a census, which is the count of patients present at a specific time and in a particular place. In **hospitals,** this is referred to as the **inpatient census.** The hospital census is a source of primary information. **Ambulatory care** facilities also may keep a census. This figure usually represents the number of **visits** or encounters during a specified period, usually one day. An **encounter** is defined in appendix B of this book as the direct personal contact between a patient and a physician or other person authorized by state licensure and, if applicable, by medical staff bylaws to order or furnish healthcare services for the diagnosis or treatment of the patient.

Hospital departments also keep statistics on the activities they perform for patients. For example, the laboratory department may keep data on the number of lab tests performed on **inpatients** and **outpatients.** The radiology department may keep track of the number of chest and hip x-rays. The physical therapy department may use statistical information such as the number of patient visits in deciding whether to hire additional physical therapists or add physical therapist assistants to their staff. These reports may be used in turn by the managers of the departments for productivity reports and combined to produce a report of activity for the entire facility. The administration of your hospital might ask you to keep data on the number of patients transferred to another hospital for a cardiac catheterization in order to determine the need for that service at your facility.

Another example of a primary source of information is **vital statistics.** The National Vital Statistics System (NVSS) is part of the National Center for Health Statistics (NCHS) of the CDC. These data are provided throughout the 50 states, Washington, DC, New York City, and five territories of the United States (Puerto Rico, the Virgin Islands, Guam, American Samoa, and the Commonwealth of the Northern Mariana Islands) to the NCHS. *Vital statistics* refers to a special group of statistics that record important events in our lives such as birth, marriage, death, divorce, and fetal death. Healthcare facilities are interested in births and deaths, fetal deaths, and induced terminations of pregnancy and generally are responsible for completing certificates for births, fetal deaths, abortions, and, occasionally, death certificates. All states have laws that require this information. The certificates are reported to the individual state registrars and maintained permanently.

The NCHS has developed standard certificates and procedures that states and territories must use to help facilitate the consistent collection of data. The standard certificates represent the minimum basic data set necessary for the collection and publication of comparable national, state, and local vital statistics data. The standard forms are revised about every 10 years, and the last revision was completed in 2003.

Data from the states and territories provide important information for use in medical research and are extremely valuable in estimating population growth in particular areas of the country and essential in planning and evaluating maternal and child health programs.

NCHS prepares and publishes national statistics based on vital statistics data because they are important in the fields of social welfare and public health. Because of their many uses, the data on these certificates must be complete and accurate.

Exercise 1.1

Identify the following as either a Primary Data Source or a Secondary Data Source:

Type of Healthcare Information	Source
1. Health insurance data pulled from national census	
2. Hospital census	
3. Productivity reports pulled from patient visit report	
4. Patient health record	
5. State vital statistics	
6. Tumor registry	
7. Hospital disease index	

Users of Health Statistics

All healthcare entities and third-party payers collect and use statistics. Following are examples of individuals and organizations that collect statistics and how they use statistics:

- Healthcare Administration: Inpatient facilities use health statistics to help address staffing issues and to determine the types of services to provide. For example, if the number of patients in the intensive care unit is increasing, the hospital administration may want to consider adding beds and staff to meet the growing need. Conversely, if a request is made to the hospital administration for new facilities and equipment that cannot be substantiated by the statistics, it is unlikely the request will be granted. Quality management departments in healthcare facilities collect data to determine how they are doing in regard to patient care and to how they can improve their patient care services.
- Healthcare Department Managers: Individual department managers in healthcare organizations use statistics to implement their department goals. For example, a manager needs to know if he or she is staying within budget. If not, the manager will need to investigate the reason.
- Cancer Registries: A cancer registry may be maintained by a separate department or may be a function of the health information department. States may also have a state cancer registry that is responsible for collecting data about cancer. A cancer registry collects data about the diagnosis, treatment, and follow-up of cancer patients. These statistics are important in tracking cancer survival rates. Facilities may choose to undergo accreditation through the American College of Surgeons. This is an evaluation by an independent team to determine whether the facility's

cancer registry meets their standards. Statistics must show the facility is providing high-quality care and follow-up to its cancer patients.

- **Nursing Facilities:** Long-term care facilities may use statistics to determine the types of payers their patients have. These statistics also are helpful in demonstrating to the public the types of patients being cared for.
- **Home Health (HH):** Home health agencies provide care to elderly, disabled, and convalescent patients in their homes. These agencies keep statistics to determine the types of services used by their patients and their outcomes. For example, a home health agency would need to know the number of nursing visits, home health aide visits, physical therapy treatments, and patients with various types of equipment such as oxygen machines or other respiratory aids. Additionally, agencies will report patient outcomes, such as the number of patients who have improved, the number of patients who were compliant with taking their medications, or the number of patients who had to be readmitted to a hospital.
- **Hospice:** Hospice programs provide care and psychosocial support to terminally ill patients and their families. These services may be given in either the home or an inpatient setting. A hospice needs to know types of illnesses in order to match the appropriate caregiver with each patient.
- Mental Health Facilities: These may be inpatient or outpatient facilities. These facilities use health statistics to determine whether they are providing the proper services for patients in the community.
- Drug and Alcohol Facilities: These programs may be inpatient, ambulatory, or a combination of the two. Statistics are important in this area to show the success rates of these facilities' clients.
- Outpatient Facilities: These include physician clinics, surgery centers, emergency centers, and the like. Outpatient facilities often use statistics to determine whether they are providing the proper level of care to the community.
- **Managed Care Organizations (MCOs):** MCOs use statistics to determine whether they are providing the correct level of care at the best cost. Additionally, MCOs contract with healthcare facilities to provide specific services to their members at a particular cost. The MCO pays the agreed upon amount each time a member uses the service. Typically, the MCO receives a discounted rate and this results in individual members of the MCO paying less out of pocket.
- Healthcare Researchers: Researchers depend on healthcare statistics to perform **research.** Some examples include research in managed care, health law and regulations, mergers and acquisitions of healthcare facilities, physician practice issues, different types of illness and risk factors, telehealth issues, pharmaceutical research, and so on.
- Accreditation Agencies: These organizations use healthcare statistics to determine the most common diagnoses and procedures and whether the resources are available to treat patients with those diagnoses.

- Federal Government: The US government collects information for public health issues. For example, the CDC reports data on births, deaths, birth defects, cancer, and HIV/AIDS, just to name a few of the categories of data. CMS uses data collected by Quality Improvement Organizations (QIOs) for its quality improvement projects.

Handy Tip: Health information management (HIM) practitioners must remember that, first and foremost, statistics must be gathered and formulated, or expressed in systematic terms or concepts, before they even exist. HIM practitioners are usually the individuals who gather and formulate this information.

Because HIM practitioners have a broad knowledge of healthcare facilities as well as immediate access to a wide range of clinical data, they are in the best position to collect, prepare, analyze, and interpret healthcare data. HIM practitioners must learn acceptable terminology, definitions, and computational methodology if they are to provide the basic and most frequently used health statistics.

Chapter 1 Test

Select the best answer to the following questions:

1. In order to be useful, the figures used in statistics must be:
   a. fair and exact.
   b. relevant and reliable.
   c. honest and justified.
   d. simple and clear.
2. To be reliable, statistical information must:
   a. have some consistency.
   b. be applicable to what is being measured.
   c. be collected from one source only.
   d. have multiple meanings.
3. Descriptive statistics makes inferences or a best guess about a larger group of data by drawing conclusions from a smaller group of data.
   a. True
   b. False
4. Which of the following is NOT a primary source of statistics?
   a. Health record
   b. Vital statistics
   c. Hospital census
   d. Disease and operation index

(continued on next page)

Chapter 1 Test (continued)

5. What is the correct sequence to go from obtaining knowledge to use of that knowledge for decision making?
   a. Knowledge → Data → Information → Facts → Improved understanding → Better decision making
   b. Data → Information → Facts → Improved understanding → Better decision making
   c. Data → Information → Facts → Statistics → Improved understanding → Better decision making
   d. Information → Facts → Data → Improved understanding → Better decision making
6. The National Center for Health Statistics keeps statistics on:
   a. the licensing information on all healthcare providers in the 50 states.
   b. cancer and other deadly diseases in the 50 states and the US-owned six territories.
   c. vital statistics, such as births, deaths, and fetal deaths in North America.
   d. vital statistics, such as births, deaths, and fetal deaths in the 50 states and the US-owned territories.
7. Vital statistics are a primary data source for information.
   a. True
   b. False
8. Which user of statistics has the primary job of supporting terminally ill patients and their families?
   a. Home health agencies
   b. Nursing facilities
   c. Hospice
   d. MCOs
9. The CDC is the lead agency that:
   a. accredits and licenses acute hospital facilities in the United States.
   b. is responsible for providing vital statistics to various agencies, such as the NCHS.
   c. develops and updates ICD-10 for the world.
   d. is responsible for protecting the health of the people of the United States.
10. Which of the following is considered to be a primary source of information?
    a. The inpatient census
    b. Vital statistics collected by the NCHS
    c. The health record
    d. All of the above
    e. b and c only