Vital Statistics, Overview

Vital statistics, as a scientific discipline, is a subdomain of **demography**, the study of the characteristics of human populations. Vital statistics comprises a number of important events in human life including birth, death, fetal death, marriage, divorce, annulment, judicial separation, adoption, legitimation, and recognition. The term "vital statistics" is also applied to individual measures of these vital events. Thus, a birth rate is an example of a vital statistic and an analysis of trends in birth rates is an example of an application in the field of vital statistics. A vital statistics system is the total process of collecting by civil registration, enumeration, or indirect estimation, information on the frequency of occurrence of vital events, selected characteristics of the events and the persons concerned, and the compilation, analysis, evaluation, and dissemination of these data in summarized statistical form. Other life events of demographic importance such as change of place of residence (migration), change of citizenship (naturalization), and change of name are not included, mainly because information on these is usually derived from other statistical systems such as population registers [1].

Systems for Collecting Vital Statistics

It is generally accepted that the preferred method for individual countries to collect vital statistics is through a civil registration system. This is recognized by the United Nations (UN) and other international organizations, as well as by the many countries that have had civil registration laws and regulations in place and in operation for many years [1, 2]. Nevertheless, a number of newly emergent and developing nations, facing the difficulties and length of time it takes to create a satisfactory civil registration system, have instituted alternative procedures to acquire statistical data to describe the levels and trends for key vital events, particularly for fertility and mortality measurements. The UN recognizes the importance of a civil registration system for each country as the preferred source of vital statistics data for the long run. However, use of an alternative data collection system is recommended as an interim measure for meeting needs for essential information where a civil registration system of acceptable quality does not yet exist. Other systems include, for example, probability area samples (*see* **Probability Sampling**), purposeful area samples, records-based **surveys**, and **record linkage**. Furthermore, the UN recommends a priority order for the types of vital statistics data to be collected. The highest need is given to data on births and deaths, followed in order by marriages, divorces, fetal deaths, annulments, judicial separations, adoptions, legitimations, and recognitions [1].

Uses of Vital Records and Vital Statistics

Vital records created through a civil registration system have two classes of use. They have value individually as legal documents for the persons named thereon; they also constitute the input, when aggregated, for the various vital statistics measures that are used to study the demographics and health of populations and population subgroups.

For the individual, a birth record is a legal document establishing name, parentage, birth data, order of birth for multiple births, legitimacy, and citizenship, nationality, or geographic place of birth. A wide variety of individual rights and civil entitlements depends on these facts, including proof of age for school entrance, motor vehicle drivers' licenses, military service and other age-related activities, establishment of eligibility for family allowances, insurance benefits, tax benefits, inheritance rights, issuance of passports, etc. The death record provides documentary proof of the facts of death needed for social security and insurance purposes such as time and place of death and the medical cause of death. Proof of death and the associated facts are also used for property inheritance rights, for remarriage rights of surviving spouses, etc. Marriage and divorce records serve to document rights to special social and economic programs and benefits for the married, including tax privileges for couples, alimony, change of nationality based on marriage, and the right to remarry. Many rights of children, their parents, and their guardians are dependent on records of adoption, legitimation, and recognition.

Individual vital records may also be used administratively as the basis for initiating maternal and child health services, including child immunization programs, or for epidemiologic investigations into disease outbreaks or assessments of causes of accidents and injuries. Another important administrative use of individual vital records especially of death records (*see* **Death Certification**), is for the updating or clearing of files such as electoral rolls, social security files, **disease registers**, cohort follow-up studies, tax registers, etc.

In aggregated form, vital records become a collection of vital statistics, most often in the form of **means**, **medians**, and various ratios such as proportions and **rates**. Whether collected by civil registration or by other means, vital statistics serve as key demographic variables in the analysis of population size, growth and geographic distribution, especially when used in conjunction with periodic population **censuses**. When census data are used as a base, current intercensal estimates of population size can be made, and projections into the future can be prepared using estimates of future trends in fertility, natality, and mortality linked with estimates of net migration. In addition to the importance of vital statistics to the study of population size and growth trends, other national and subnational economic and social concerns such as health, welfare, education, occupation, housing, urbanization, family structure, and income are also affected by these measures. In the fields of public health and medicine, for example, levels and trends of **infant and perinatal mortality** are often used as surrogate measures of levels and trends in the overall health and well-being of nations. **Life expectancy** at birth is also frequently used to compare the overall effects of mortality and its determinants. **Cause of death** information provides a foundation upon which much research into diseases and disease prevention is based.

Differentials in mortality by sex, age, racial groups, and other variables are often the basis for the planning of health and medical intervention programs. In addition, the planning and provision of public and private housing, educational facilities, social security and private insurance plans, medical facilities, and consumer goods of all kinds are examples of activities dependent on vital statistics data. At the international level, vital statistics provide a basis for comparing important demographic, social, and economic differences and trends over time among countries or regions of the world.

Definitions of Selected Vital Events

Standard statistical definitions of vital events have been promulgated by international agencies [1, 5]. In some cases, legal definitions may differ from the international standards in varying degrees, but, in many cases, national vital statistics reports are either based on the standard statistical definitions or do not differ in principle. In cases where comparability among countries is compromised because of the use of nonstandard definitions, international agencies and others presenting national comparisons of tabular, graphical or descriptive vital statistics usually provide appropriate cautions to users. Nevertheless, users of vital statistics data need to ascertain the comparability of the data before drawing reliable conclusions about national differences. The **World Health Organization (WHO)** promulgates a number of vital statistics definitions as part of the **International Classification of Diseases (ICD)**. These definitions are incorporated in regulations adopted by the World Health Assembly and which each WHO member country has agreed to follow [4]. Nevertheless, it is still necessary to ensure that the standard definitions have been followed for a given data set. The international standard definitions for selected vital events are given below.

Live Birth. This is the complete expulsion or extraction from its mother of a product of conception, irrespective of the duration of the pregnancy, which, after such separation, breathes or shows any other evidence of life, such as beating of the heart, pulsation of the umbilical cord, or definite movement of voluntary muscles, whether or not the umbilical cord has been cut or the placenta is attached; each product of such a birth is considered liveborn [5].

Fetal Death. This is death prior to the complete expulsion or extraction from its mother of a product of conception, irrespective of the duration of pregnancy; the death is indicated by the fact that after such separation the fetus does not breathe or show any other evidence of life, such as beating of the heart, pulsation of the umbilical cord, or definite movement of voluntary muscles [5].

Maternal Death. This is the death of a woman 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. Maternal deaths may be subdivided into two groups: direct obstetric deaths which are the result of obstetric complications of the pregnant state (pregnancy, labor, and the puerperium), from interventions, omissions, incorrect treatment, or from a chain of events resulting from any of these; and indirect obstetric deaths which are the result of previously existing disease or disease that developed during pregnancy and which was not due to direct obstetric causes, but which was aggravated by physiologic effects of pregnancy [5].

Infant Death. This is the death of a liveborn infant who dies before completing its first year of life.

Neonatal Death. This is the death of a liveborn infant who dies during the first 28 completed days of life. These may be subdivided into early neonatal deaths, occurring during the first seven days of life, and late neonatal deaths, occurring after the completion of the seventh day but before the completion of 28 days [5].

Perinatal Death. This is the death of a fetus or newborn infant occurring after 22 completed weeks (154 days) of gestation (the time when fetal weight is normally about 500 g), but prior to the completion of seven days after birth [5].

Marriage. This is the act, ceremony or process by which the legal relationship of husband and wife is constituted. The legality of the union may be established by civil, religious, or other means recognized by the laws of each country [1].

Divorce. This is a final legal dissolution of a marriage which confers on the parties the right to remarriage under civil, religious, or other provisions, according to the laws of each country [1].

Definitions of Selected Vital Statistics Measures

Raw vital statistics most often are comprised of counts of how often a specified vital event has occurred, rather than on measurements of continuous variables such as height, weight, or blood pressure. The analysis of vital data depends mainly on the conversion of observed frequencies into indices, ratios, and probabilities. Counts of vital events often do have utility, but, for the majority of uses, absolute frequencies are not sufficient and it becomes necessary to compute relative numbers, including crude rates, various types of specific rates, percentages, probabilities, and other ratios.

Some of the more commonly encountered vital statistics relative numbers are defined and calculated as follows.

Crude Death Rate

The most common form of mortality measurement is the crude death rate. It is computed from the following formula [3]:

$$m_{\mathrm{cd}} = \left(\frac{D}{P}\right)k,$$

where m_{cd} is the crude death rate, D is the total number of deaths for a given area and time period, usually a calendar year, P is the size of population at risk of dying, usually taken as the estimated population at the midpoint of the calendar year, and k is a constant, usually taken as 1000.

The crude rate is so named to differentiate it from various specific and adjusted rates and represents the total or overall death rate without regard to the various component elements which combine to produce the total figure. The crude death rate is usually expressed as "the number of deaths per 1000 persons" for a specified place (country, city, state, etc.) for a given year.

Specific Death Rate

Detailed analyses of vital statistics frequently go beyond the overall risk of death in the population as a whole. Many studies deal with subsets of the population or with particular classes of deaths. Epidemiologists often focus on deaths from a particular disease or class of diseases. Actuaries and demographers are concerned with differences in mortality by sex and in different age groups within the population. Environmental and **occupational health** specialists are interested in the differential risks of dying in selected occupations, and in different geographic subdivisions such as urban and rural areas. To meet these kinds of needs, various specific death rates are calculated. Specific rates for different age groups are called *age-specific death rates*; rates for males and females are called *sex-specific death rates*, rates for particular causes of death are called *cause-specific death rates*. Rates may be specific for combinations of characteristics. For example, age–sex–race-specific death rates are computed separately for each age group by race and sex. Specific death rates are approximations of true probabilities. That is, the denominator of the ratio is an estimate of the total number of events of a particular type that could happen, while the numerator is a count of those that did happen.

Specific death rates are computed as follows [3]:

$$m_{\mathrm{sd}} = \left(\frac{d_i}{p_i}\right)k,$$

where m_{sd} is the specific rate for any defined ith class, d_i is the number of deaths occurring in the ith class for a given area and time, p_i is the number of persons in the ith class of the population for the same area and time, and k is a constant, usually 100 000.

For cause-specific death rates, the denominator, p_i, in the above formula is replaced by P, the total population exposed to the risk of death. Therefore, a cause-specific death rate measures the risk in the total population of dying from a specified cause of death.

Infant Mortality Rate

The infant mortality rate is considered by many as one of the important indicators of the overall level of health and social well-being of a country or other geopolitical area. This is, in part, because a large proportion of deaths in the first year of life are considered to be preventable through adequate prenatal care, good nutrition for women and infants, and improved control of the environment, including injury prevention.

The infant mortality rate is computed as follows [3]:

$$m_i = \left(\frac{d_{<1}}{B}\right)k,$$

where m_i is the infant mortality rate, $d_{<1}$ is the number of deaths to liveborn infants under one year of age during a specified time period, usually one year, B is the total number of live births during the same time period, and k is a constant, usually 1000.

The infant mortality rate is a proxy for the age-specific death rate for the "under one year of age group" and is intended to be a measure of the risk of dying during the first year of life. The numerators of the infant mortality rate and the "under one year of age" age-specific death rate are the same. For a denominator, however, a reliable estimate of the size of the population under one year of age for a given time period is hard to obtain, even in a census year. As a proxy measure, the denominator may be considered to be the number of births occurring during the period. For either of these choices of denominator, there is some mismatch with the numerator in terms of a true probability number. Not all events in the numerator arise from the events in the denominator. For example, in the infant mortality rate, some of the deaths under one year of age in a given year and counted in the numerator were actually born in the previous year and are not represented in the denominator, while some of the births represented in the denominator will die before their first birthday but the deaths will occur in the next year and are not included in the numerator. However, when the birth rate is fairly stable from one year to the next, calculation of the infant mortality rate results in a ratio that closely approximates the probability of a liveborn infant dying within the first year of life. When the birth rate is not stable from year to year, a more accurate mortality rate may be computed by following each live birth occurring during a one year period and measuring how many of them die before their first birthday.

Neonatal, Early Neonatal and Postneonatal Mortality Rates

The *neonatal mortality rate* is defined as follows [5]:

$$m_{\mathrm{n}} = \left(\frac{d_{<1\ \mathrm{mo}}}{B}\right)k,$$

where m_{n} is the neonatal mortality rate, $d_{<1\ \mathrm{mo}}$ is the number of deaths of infants under 1 month of age during a specified time period, B is the number of live births occurring during the same time period, and, k is a constant, usually 1000.

The neonatal mortality rate, like the infant mortality rate, is a proxy for an age-specific death rate. It approximates the risk of dying in the first month of life. The relative importance of an infant mortality rate compared with the corresponding neonatal mortality rate depends on the proportionate age distribution of the deaths under one year of age. Generally, when the infant mortality rate is low, a large proportion of infant deaths occur during the first month of life. The neonatal mortality rate then reflects an important measure of the mortality risk for infants. Conversely, when the infant mortality rate is high, larger proportions of deaths fall into the older age groups under a year. Often it is useful to partition the deaths of infants under one year of age into two groups: those dying before one month of age, and those dying between one month and their first birthday. The former comprise the numerator, $d_{<1\ \mathrm{mo}}$, of the neonatal mortality rate, while the latter can be used to calculate the *postneonatal mortality rate*:

$$m_{\mathrm{pn}} = \left(\frac{d_{1\ \mathrm{mo}-1\ \mathrm{yr}}}{B}\right)k,$$

where m_{pn} is the postneonatal mortality rate, $d_{1\ \mathrm{mo}-1\ \mathrm{yr}}$ is the number of deaths occurring between 1 month and 1 year of age during a specified time period, B is the number of live births occurring during the same time period, and k is the same constant used in the neonatal mortality rate, usually 1000.

In similar fashion, the neonatal deaths may be partitioned into those dying within the first week of life and the remainder that survive the first seven days but die before one month of age. The risk of dying in the first week of life is measured by the *early neonatal mortality rate*, m_{en}, as follows [5]:

$$m_{\mathrm{en}} = \left(\frac{d_{<7\ \mathrm{days}}}{B}\right)k,$$

where the components of the calculation are the same as in the neonatal and postneonatal mortality rates, except that the numerator contains only those deaths to infants occurring during the first week of life.

Perinatal Mortality Rate

The perinatal period, as defined earlier, is the period of time surrounding the event of birth. It includes the time that a fetus spends in utero after it has reached 22 weeks of gestation and continues through the birth process until the end of the first week of life after birth. The perinatal mortality rate measures mortality occurring during this period. The rate, therefore, combines deaths of fetuses of specified **gestational age** with deaths of liveborn infants who die in their first week of life. The determination of whether a fetus is born dead or whether it shows any sign of life before expiring is not always clear-cut; social, economic, and cultural factors, as well as medical and biological considerations, tend to push the fetal death rate in one direction or the other in different societies, thus making comparisons of neonatal or infant mortality among countries difficult. By using the perinatal mortality rate for comparisons, this difficulty is minimized since fetuses dying just before or during the birth process as well as those born alive but dying shortly thereafter are all included in the calculation [5]:

$$m_{\mathrm{peri}} = \left[\frac{d_{\mathrm{peri}}}{F + B}\right]k,$$

where m_{peri} is the perinatal mortality rate, d_{peri} is the number of deaths of fetuses of 22 or more weeks of gestation plus deaths of liveborn infants of less than 7 days of age during a specified period, usually a calendar year, F is the number of fetal deaths of 22 or more weeks of gestation during the same period, B is the number of live births during the same period, and, k is a constant, usually 1000.

Note that, unlike the infant and neonatal mortality rates, the denominator of the perinatal mortality rate combines both the number of live births and the number of fetal deaths of 22 or more weeks of gestation. This denominator is called "total births" and better approximates the population from which the numerator could arise than would a denominator restricted to only live births. On the other hand, it is recognized that it is easier to collect reliable counts of live births than of fetal deaths, thus introducing another source of error into the calculation of the perinatal mortality rate.

Maternal Mortality Rate

The **maternal mortality** rate is calculated as follows [5]:

$$m_{\mathrm{m}} = \left[\frac{d_{\mathrm{md}} + d_{\mathrm{mi}}}{B}\right]k,$$

where m_{m} is the maternal mortality rate, d_{md} is the number of direct maternal deaths in a specified time period, usually 1 year, d_{mi} is the number of indirect maternal deaths in the same period, B is the number of live births in the same period, and k is a constant, usually 10 000 or 100 000.

A related measure, the *direct obstetric mortality ratio*, may be calculated from the above formula but using in the numerator only the direct maternal deaths, d_{md}.

Proportionate Mortality

Proportionate mortality, sometimes known as the death ratio (*see* **Proportional Mortality Ratio (PMR)**), is defined as [3]:

$$p_{\mathrm{d}} = \left(\frac{d_{\mathrm{i}}}{D}\right) k,$$

where p_{d} is the proportionate mortality, d_{i} is the number of deaths in a specified class during a stated time period, D is the total number of deaths in the same time period, and k is a constant, usually 100 or 1000.

Proportionate mortality ratios may be calculated for any class of deaths, but their most common uses are for given causes or group of causes of death expressed as percentages of deaths from all causes, or for deaths at a specified age expressed as percentages of deaths at all ages.

Crude Birth Rate

The crude birth rate is the most frequently used overall measure of the reproduction of a population. Like its counterpart, the crude death rate, it is influenced by many factors and represents a proxy for more specific fertility measurements. It is calculated as follows [3]:

$$m_{\mathrm{cb}} = \left(\frac{B}{P}\right) k,$$

where m_{cb} is the crude birth rate, B is the total number of live births for a given area and time period, P is the total population at the midpoint of the time period, and, k is a constant, usually 1000.

Comparing Vital Statistics Data

Aggregated vital statistics data, whether in tabular or graphical form, often appear as time trends for particular variables such as causes or groups of causes of death, or for age and sex groups of the population. They also appear frequently as comparisons between countries or other geographical entities for a point in time, usually a particular year. In either case, great care must be taken to ensure that the quality of the data in the groups being compared warrants making the comparisons. In registration based systems, measures or estimates of completeness of reporting of vital events should be known. In sample based systems, the representativeness of the sample and the **nonresponse** rate is important. In the comparison of data between two or more geographic places, it is important to ascertain if common definitions and procedures were used to collect, process, analyze, and present the data; in looking at time trends, it is essential to know if the definitions of the events and the procedures for classifying the data remained constant over the entire time period being studied. This latter point is particularly important when looking at trends in causes of death since the instrument for grouping diseases into categories for study, the ICD, is revised approximately every 10 years (*see* **Morbidity and Mortality, Changing Patterns in the Twentieth Century**; **Mortality, International Comparisons**). Vital statistics data are often presented in statistical compendia published by official national and international organizations that attempt to include important notes for interpretation of the data in headnotes and footnotes to tables, appendices, etc. (*see* **Data Access, National and International**). The user is cautioned to pay careful attention to such explanatory or cautionary notes.

References

[1] United Nations (1973). Principles and Recommendations for a Vital Statistics System. *Statistical Papers, Series M, No. 19, Rev. 1*. United Nations, New York.

[2] United States Bureau of the Census (1971). *The Methods and Materials of Demography*, H. Shryock, J. Siegel et al., eds. US Government Printing Office, Washington.

[3] United States Department of Health, Education, and Welfare (1965). Techniques of Vital Statistics. Reprint of Chapters I-IV, *Vital Statistics Rates in the United States, 1900–1940*. National Center for Health Statistics, Washington.

[4] World Health Organization (1967). *WHO Nomenclature Regulations*. World Health Organization, Geneva.

[5] World Health Organization (1992). *International Statistical Classification of Diseases and Related Health Problems: 10th revision*, 3 vols. World Health Organization, Geneva.

Robert A. Israel