The effort to define and standardize terms and categories used in morbidity statistics has produced intense discussion and a few distinct differences among biometricians. Dr. Dorn's contribution to this discussion indicates some of the issues that lie in the way of general agreement.

A Classification System for Morbidity Concepts

HAROLD F. DORN, Ph.D.

THE LOW DEATH RATE in most of the countries of Europe, North America, and Oceania, and in certain countries in other parts of the world greatly limits the usefulness of mortality statistics as a measure of the amount and characteristics of ill health of the population of these countries. The recognition of this fact has stimulated interest in the collection and analysis of a variety of morbidity statistics.

Although morbidity statistics for insured populations and for members of sick benefit associations date from the last century, corresponding statistics for the general population are of much more recent origin. General morbidity surveys of selected areas of a country or of special population groups were made as long as 40 years ago, but efforts to collect general morbidity data for the entire population of a country date from about 1940.

The publication of the findings of general morbidity surveys has made clear that there is no consensus concerning terms used to describe and measure morbidity. This is not surprising since agreement on terminology is not easy to achieve. It is a sound principle that attempts to reach agreement on the definition of standard terms should be preceded by a period of use of a variety of terms so that the adoption of a standard terminology may be based upon the demonstrated utility of the preferred terms.

The statistical description and measurement of morbidity is more complex than that of mortality. In addition, experience in the use of different terms is still rather limited so that it is doubtful if the time is ripe for an attempt to reach agreement upon a list of standard morbidity terms for use on a national basis. Nevertheless, it would be advantageous to encourage discussion of the types of terms required in the description and measurement of morbidity and also to propose the use of some terms on a trial basis in order to bring about the consensus requisite to agreement upon a standard terminology.

A large number of terms to describe the different aspects of ill health and to measure the risk of becoming ill, the amount of ill health in a population, or the amount of disability due to ill health already exist. If these existing terms are to be organized into an orderly system, it is essential first to develop a general scheme for classifying and enumerating the

Dr. Dorn is chief of the Biometrics Branch, Division of Research Services, National Institutes of Health, Public Health Service, Bethesda, Md.

unit or units of observation in morbidity studies. This paper proposes such a general scheme.

It is necessary to use certain terms in order to continue with this discussion. These should not be regarded as necessarily being preferred terms; they are used only for the purpose of facilitating this discussion. Once a general scheme is outlined, the definition of the various concepts involved can be considered.

A period of ill health is a continuous interval of time during which a person experiences a departure from a state of good health. This also has been called a spell or episode of ill health or a complaint period.

During a period of ill health, one or more separate diagnostic entities or causes of ill health may exist. These will be called illnesses or diagnoses with the understanding that illness includes conditions resulting from disease, poisoning, and injury.

The amount of ill health in a population may be measured by (*a*) the number of persons who are ill, (*b*) the number of periods of ill health, or (*c*) the number of separate illnesses or diagnoses. During a fixed interval of time one person may experience one or more periods of ill health with one or more illnesses during each period. Consequently, it is important to be clear as to which unit of measurement is being used since the definition and method of computation of morbidity rates is not the same for each unit.

For the purpose of measuring morbidity, ill health may be classified (*a*) with respect to the interval of time during which observations are made, and (*b*) from the point of view of the person affected.

If we observe a population during a specified interval of time, four categories of ill health may be observed. For convenience in exposition the term "case" will be used to denote the manifestation of ill health being observed and may refer to a person, a period of ill health, or an illness.

1. Cases existing prior to the start of the interval, continuing throughout the interval, and still existing at the end of the interval.

2. Cases existing prior to the start of the interval and terminating during the interval.

3. Cases beginning during the interval and still existing at the end of the interval.

4. Cases beginning during the interval and terminating during the interval.

Since the term "case" is here used in a general sense, category 4, for example, may be interpreted as (*a*) the number of persons becoming ill and recovering during the interval, t_1 to t_2, or (*b*) the number of periods of ill health beginning and terminating during the interval, t_1 to t_2, or (*c*) the number of separate illnesses beginning and terminating during the interval, t_1 to t_2. In general, these three numbers will not be the same.

This classification suggests three ways of counting cases in relation to time.

1. The number of cases existing at some point of time, for example at t_1. This would include categories 1 and 2 shown in diagram 1. In practice, this may be defined as the number of cases existing during a single day or as the average daily number of cases existing during the interval t_1 to t_2.

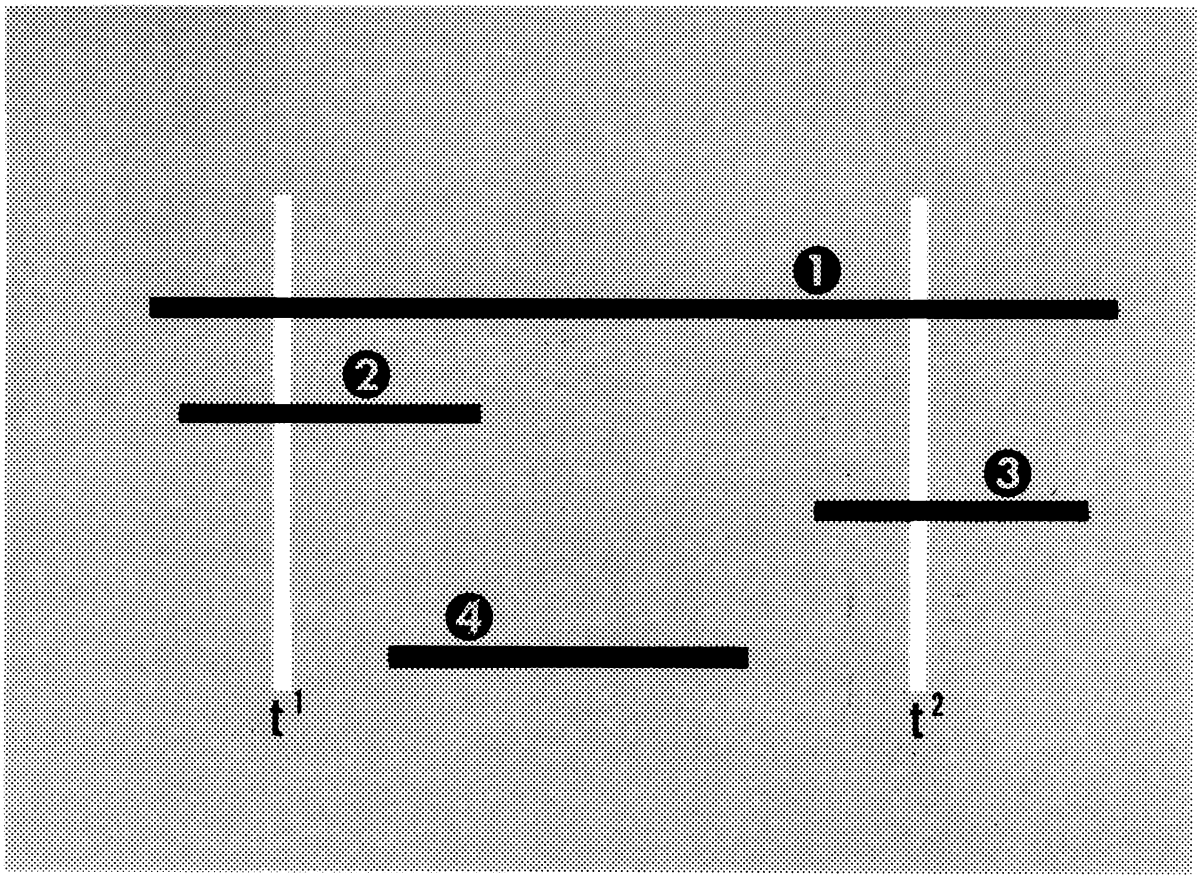
2. The number of cases existing at any time during the interval, t_1 to t_2. This is an index of the total amount of illness during the interval and would include a count of all four categories of cases shown in diagram 1.

3. The number of cases with onset during a specified interval of time, for example, between t_1 and t_2. This would include categories 3 and 4 shown in the diagram.

The relationship between the three units in which cases of ill health may be enumerated can be seen from the following classification of ill health from the point of view of the person affected (diagram 2).

During a specified interval of time a person may experience (*a*) no period of ill health, or (*b*) a single period of ill health, or (*c*) two or more periods of ill health. During each period of ill health one or more distinct illnesses or diagnoses may exist. Each illness may be (*a*) the first attack during the person's lifetime, or (*b*) the first attack during the period of ill health, or (*c*) the second or subsequent attack during the period. For the second and subsequent periods of ill health during the interval of observation, a specific illness may be classified as to whether or not it is the first attack during this interval.

Diagram 1.



First attacks may be illnesses (*a*) for which one attack gives lifelong immunity, for example, smallpox or measles; or (*b*) from which complete recovery may occur but no immunity from subsequent attacks exists, for example, the common cold or pneumonia; or (*c*) with a persistent residual pathological process characterized by alternating periods of remission of symptoms and clinical manifestation of ill health. Most chronic diseases such as bronchitis, asthma, and arthritis fall into this last category. Included also are illnesses for which even temporary remission of symptoms does not occur. This classification of first attacks applies equally to all subsequent attacks of illness except for the class of illnesses that give lifelong immunity after one attack.

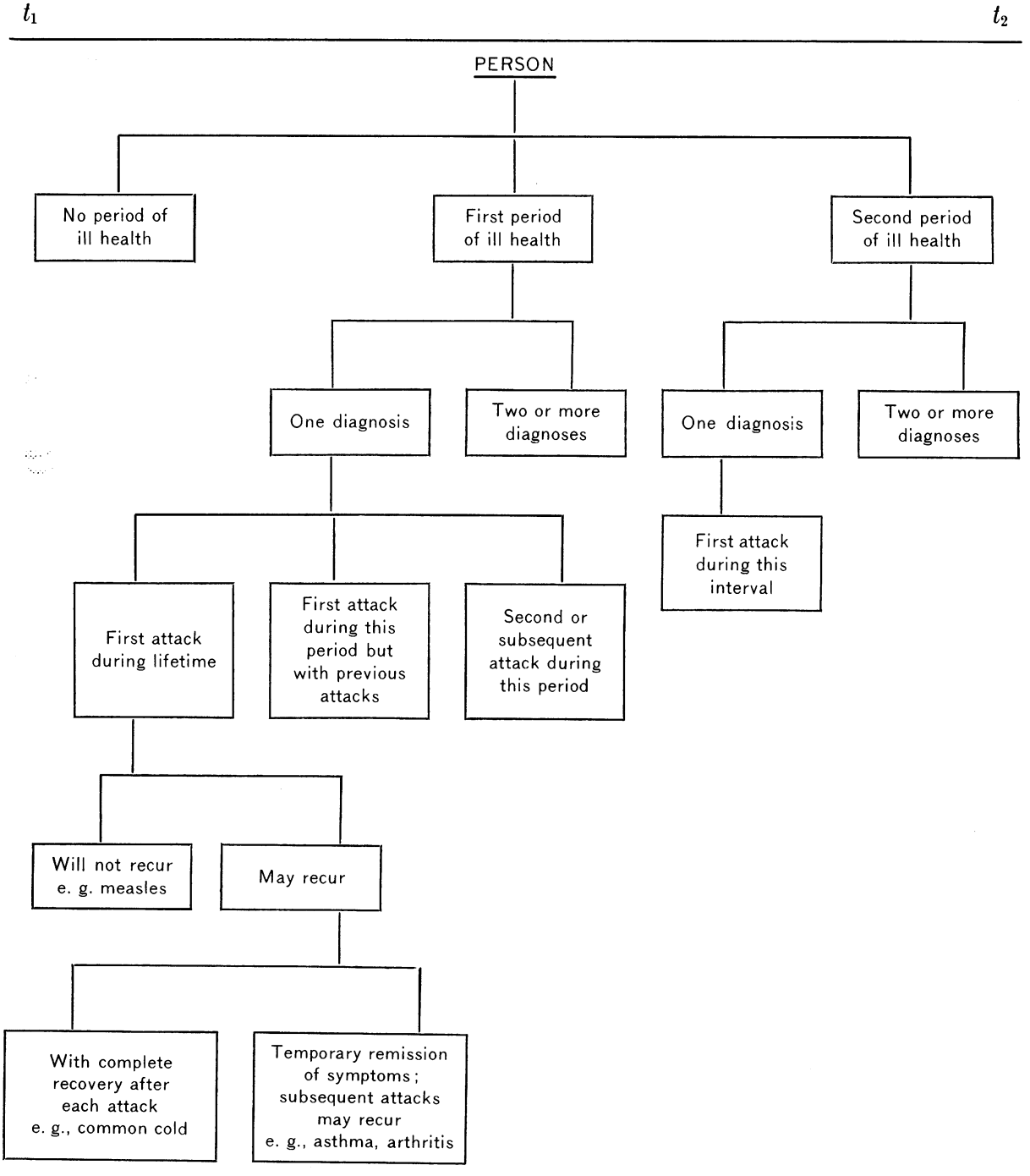
Three ways of counting cases of ill health with respect to time were mentioned above: (*a*) the number of cases existing at some point in time, (*b*) the number of cases existing at any time during an interval of time, and (*c*) the number of cases with onset during some interval of time. The first two ways result in measures of the amount of ill health in a population and the third way results in a measure of the risk of ill health or of the rate at which ill health develops in a population.

Rates computed from the first two ways of counting cases may be termed prevalence rates of ill health. They measure (*a*) the amount of ill health at a particular point of time, in practice usually a given day, that is, point prevalence; or (*b*) the amount of ill health during a specified interval of time, a month or a year—period prevalence.

Prevalence rates may be based upon a count of persons who are ill, a count of periods of ill health, or a count of illnesses. For a prevalence rate at a particular point of time, the number of ill persons and the number of periods of ill health is the same. However, the number of illnesses or diagnoses may be greater than the number of ill persons.

A prevalence rate for an interval of time, t_1 to t_2, is based on a count of cases existing at t_1 plus all cases beginning during the interval, t_1 to t_2. In this instance the prevalence rate for persons usually will be less than that for periods of ill health which in turn will be less than that for illnesses or diagnoses. A useful special form of a prevalence rate during an interval, t_1 to t_2, is the proportion of a population which is ill, computed as a daily average for the interval. In this form, it is often called the daily noneffective rate and shows the proportion of a population which is ill on an average day.

Diagram 2.



A rate computed from the third way of counting cases with respect to time, that is from a count of cases beginning during an interval, may be termed an incidence rate. This is a measure of the risk of becoming ill or of the rate at which ill health develops. It may be computed for persons, periods of ill health, or for illnesses. The method of computation and interpretation is, of course, different in each instance.

An incidence rate for persons represents the proportion of the population that is ill at least once during the interval t_1 to t_2. An incidence rate for periods of ill health represents, per unit of population, the number of separate periods of ill health developing during the interval, t_1 to t_2. The magnitude of this rate may exceed that of the unit of population on which the rate is based. For example, a rate per head may be greater than unity.

Before an incidence rate for illnesses can be computed, it is necessary to decide which of the types of illnesses shown in diagram 2 are to be used. Illnesses that give lifelong immunity after one attack present no special problem. Illnesses from which complete recovery is possible may be counted the first time they occur in each period of ill health. A person with two or more periods of ill health may have two or more common colds during an interval of observation. A second attack of the same illness during a single period of ill health creates a more difficult problem, for it is necessary to decide whether the second attack is merely a prolongation of the first or is a new attack of the same disease and hence should be counted in the computation of an incidence rate.

The greatest problem is created by some of the so-called chronic illnesses that are characterized by alternating periods of active clinical manifestation of ill health and lack of symptoms. Asthma, bronchitis, arthritis, and migraine are examples of this class of illnesses. A count of attacks of these illnesses can be based upon (*a*) the first attack during a lifetime, or (*b*) the first attack during the interval, t_1 to t_2, or (*c*) the first attack during a period of ill health, or (*d*) each separate attack during a period of ill health. Obviously the magnitude and interpretation of the incidence rates based on these methods of counting cases will differ greatly.

Incidence and prevalence rates belong to a class of rates designed to measure the frequency of ill health. Although certain incidence and prevalence rates, namely, those based on persons sick one or more times, taken as a unit or on first attacks of illness, may be interpreted as relative frequencies with a maximum value of unity and hence may be considered to be a measure of the probability of ill health, the remaining rates are in reality weighted averages and may exceed unity when expressed per head of population.

A second class of rates are those designed to measure disability. These yield the average number of days of disability (*a*) per person, (*b*) per period of ill health, or (*c*) per illness. The usual method of computation is to divide the number of days of disability during a specified interval, t_1 to t_2, by the appropriate denominator and express the quotient on a per annum basis. If the rate is for persons, the appropriate denominator would be the average number of persons in the population during the interval. There are advantages in counting only days of disability occurring within the interval, t_1 to t_2, for persons who are ill at the beginning of the interval. This rate yields the average number of days of disability per person per annum or some other unit of time.

If the rate based on persons is expressed per day, it is often called the daily noneffective rate since it is the average proportion of persons who are disabled on a given day during the interval, t_1 to t_2. This results from the fact that a day of disability is equivalent to one person disabled for 1 day. The daily disability or noneffective rate also is an average daily prevalence rate.

The computation of the average number of days of disability per illness creates knotty problems in the determination of the number of days of disability to assign to two or more illnesses occurring during the same period of ill health.

The amount of disability due to a period of ill health may be computed by dividing the total number of days of disability during the interval, t_1 to t_2, by the number of periods of

ill health during the interval. For periods that begin before the interval of observation, only the days of disability during the interval are counted. Similarly for periods not terminated at the end of the interval, only the days of disability during the interval are counted.

This definition leads to an interesting relationship between frequency and disability rates when the frequency rate is defined as the number of periods of ill health existing during the interval divided by the average number of persons in the population with the quotient expressed per head. Then, if F represents the frequency rate, D the disability rate per person, and S the disability rate per period, $F \times S = D$; or $F = D/S$.

Another special relationship exists if the frequency rate is defined as in the previous paragraph and expressed per person per day, that is, as a daily rate per person, and the disability rate is computed for persons and expressed as an average daily prevalence rate or noneffective rate per person. Then from the relationship between F, S, and D shown above, we have $F = D/S$, or

$$\text{daily morbidity rate} = \frac{\text{daily noneffective rate}}{\text{average number of days of disability per period of ill health}}$$

If the daily morbidity rate is 1 per 1,000, the daily noneffective rate, that is, the proportion of the population ill on an average day, equals the average number of days of disability per period of ill health. For example, if the daily hospital admission rate is 1 per 1,000, the proportion of the population in hospitals on an average day is equal to the average duration of stay in hospitals.

In general the above relationships between frequency and disability rates hold true only in a population with a fixed pattern of ill health. A further discussion of this point is beyond the scope of this paper.

Some persons have computed the measure of disability as the average duration of periods of ill health ending during the interval of observation, t_1 to t_2. This necessitates counting the days of disability occurring before time, t_1, for periods of ill health existing at t_1. There is no reason why this method of computation should not be used provided the result is useful. However, if this is done, the above relationships between F, D, and S no longer hold true. Furthermore, although this method of computation gives the correct duration of cases terminating during the interval, t_1 to t_2, this is not necessarily the same as the eventual duration of cases still ill at the end of the interval except for a population with a fixed pattern of ill health.

The above discussion is only an introduction to concepts useful in general morbidity statistics of a population. Terms useful for special morbidity statistics, such as hospital statistics and insurance statistics, have not been considered. The purpose of this discussion is to clarify some of the basic concepts useful in the description and measurement of morbidity in order to provide a basis for the development of widely acceptable definitions for specific terms.

BIBLIOGRAPHY

Copes, P.: The statistical measurement of morbidity frequency. Ottawa, Canada, Dominion Bureau of Statistics, 1957.

Measurement of morbidity: *In* Studies on medical and population subjects No. 8. London, Great Britain General Register Office, 1954.

Stocks, P.: Sickness in the population of England and Wales in 1944–1947. *In* Studies on medical and population subjects No. 2. London, Great Britain General Register Office, 1949.