

## Unit I – Problem 10 – Handout (1): A Dictionary of Epidemiology



- **Attack rate:** the cumulative incidence of infection in a group observed over a period during an epidemic. This “rate” can be determined empirically (تجريبياً) by identifying clinical cases and/or by means of seroepidemiology. Because its time dimension is uncertain or arbitrarily (اعتباطياً) decided, it should probably not be described as a rate.
- **Carrier:** A person or animal that harbors a specific infectious agent in the absence of discernible (ملحوظ-قابل للإدراك) clinical disease and serves as a potential source of infection. The carrier state may occur in an individual with an infection that is inapparent throughout its course (known as healthy or asymptomatic carrier) or during the incubation period, convalescence (recovery), and post-convalescence of an individual with a clinically recognizable disease (known as incubatory carrier or convalescent carrier). The carrier state may be of short or long duration (temporary or transient or chronic carrier).
- **Confidence interval (CI):** the computed (محسوب) interval with a given probability, e.g., 95%, that the true value of a variable such as a mean, proportion, or rate is contained within the interval.
- **Endemic disease:** the constant presence of a disease or infectious agent within a given geographic area or population group; may also refer to the usual prevalence of a given disease within such area or group.
- **Epidemic:** (epi: upon – demos: people) the occurrence in a community or region of case of illness, specific health-related behavior, or other health-related events clearly in excess of normal expectancy. The community or region and the period in which the cases occur are specified precisely. The number of cases indicating the presence of an epidemic varies according to the agent, size and type of population exposed ; previous experience or lack of exposure to the disease; and time and place of occurrence. Epidemicity is thus relative to usual frequency of the disease in the same area, among the specified population, at the same season of the year. A single case of a communicable disease long absent from a population or first invasion by a disease not previously recognized in that area requires immediate reporting and full field investigation; two cases of such a disease associated in time and place may be sufficient evidence to be considered an epidemic. Recent epidemics initially identified following the occurrence of small numbers of cases include the epidemic of vaginal cancer in daughters of women who took diethylstilbestrol during pregnancy and the pandemic of AIDS that was heralded by a report of cases of pneumocystis carinii pneumonia among gay men in Los Angeles 1981. The purpose of surveillance systems such as the Epidemic Intelligence Service is to identify epidemics as early as possible so that effective control measures can be put in place. This remains the most important use of epidemiology.
- **Incidence rate:** the rate at which new events occur in a population. The numerator is the number of new events that occur in a defined period; the denominator is the population at risk of experiencing the event during this period, sometimes expressed as person-time. The incidence rate most often used in public health practice is calculated by the formula:

$$\frac{\text{Number of new events in specified period}}{\text{Number of persons exposed to risk during this period}} \times 10^n$$

In dynamic population, the denominator is the average size of the population, often the estimated population at the mid-period. If the period is year, this is the annual incidence rate. This rate is an estimate of the person-time incidence rate, i.e., the rate per  $10^n$  person-years. If the rate is low, as with many chronic diseases, it is also a good estimate of the cumulative incidence rate over a period, e.g, a year. In follow-up studies with no



censoring (رقابة), the incidence rate is calculated by dividing the number of new cases in a specified period by the initial size of the cohort of persons being followed; this is equivalent to the cumulative incidence rate during the period. If the number of new cases during a specified period is divided by the sum of the person-time units at risk for all persons during the period, the result is the person-time incidence rate.

- **Incubation period:** the time interval between invasion by an infectious agent and appearance of the first sign or symptom of the disease in question. In a vector, it is the period between entry of the infectious agent into the vector and the time at which the vector becomes infective; i.e., transmission of the infectious agent from the vector to a fresh final host is possible.
- **Pandemic:** an epidemic occurring worldwide, or over a very wide area, crossing international boundaries, and usually affecting a large number of people.
- **Rate:** a measure of the frequency of occurrence of a phenomenon. In epidemiology, demography, and vital statistics, a rate is an expression of the frequency with which an event occurs in a defined population in a specified period of time. The use of rates rather than raw numbers is essential for comparison of experience between populations at different times. Different places, or among different classes of persons. The components of a rate are the numerator, the dominator, the specified time in which events occur, and usually a multiplier, a power of 10, that converts the rate from an awkward fraction or decimal to a whole number:

In vital statistics:

$$\text{Rate} = \frac{\text{Number of events in specified period}}{\text{Average population during the period}} \times 10^n$$

In epidemiology, the denominator is usually person time.

All rates are ratios, calculated by dividing a numerator, e.g, the number of deaths, or newly occurring cases of a disease in a given period, by a denominator, e.g, the average population during that period. Some rates are proportions, i.e., the numerator is contained within the denominator. Rate has several different usages in epidemiology:

- As a synonym (مرادف) for ratio, it refers to proportions as rates, as in the terms cumulative incidence rate, prevalence rate, survival rate.
  - In other situations, rate refers only to ratios representing relative changes (actual or potential) in two quantities. This accords with OED, which gives “relative amount of variation” among its definitions for rate.
  - Sometimes rate is further restricted to refer only to ratios representing changes over time, in this usage, prevalence rate would not be a true rate because it cannot be expressed in relation to units of time but only to a point in time, in contrast, the force of mortality or force of morbidity (hazard rates) is a true rate, for it can be expressed as the number of cases developing per unit time dividing by the total size of the population at risk.
- **Ratio:** the value obtained by dividing one quantity by another: a general term of which rate, proportion, percentage, etc., are subsets. The important difference between a proportion and a ratio is that the numerator of a proportion is included in the population defined by the denominator, whereas this is not necessarily so for a ratio. A ratio is an expression of the relationship between a numerator and a denominator where the two usually are separate and distinct quantities, neither being included in the other. The



dimensionality of a ratio is obtained through algebraic cancellation, summation, etc., of the dimensionalities of its numerator and denominator terms. Both counted and measured values may be included in the numerator and the denominator. There are no general restrictions on the dimensionalities or ranges of ratios, as there are in some of its subsets (e.g., proportion, prevalence). Ratios are sometimes expressed as percentages (e.g., standardized mortality ratio). In these cases, unlike the special case of proportion, the value may exceed 100.

- **Relative risk:**

- The ratio of the risk of disease or death among exposed to the risk among the unexposed; this usage is synonymous with risk ratio.
- Alternatively, the ratio of the cumulative incidence rate in the exposed to the cumulative incidence rate in the unexposed, i.e., the rate ratio.
- The term relative risk has also been used synonymously with odds ratio, in some biostatistical articles, has been used for the ratio of forces of morbidity. The use of the term relative risk for several different quantities arises from the fact that for “rare” diseases (e.g. most cancers) all the quantities approximate one another. For common occurrences (e.g., neonatal mortality in infants under 1500-g birth weight), the approximations do not hold.

- **Secondary attack rate:** the number of cases of an infection that occur among contacts within the incubation period following exposure to a primary case in relation to the total number of exposed contacts; the denominator is restricted to susceptible contacts when these can be determined. The secondary attack rate is a measure of contagiousness (العدوى) and is useful in evaluating control measures.

- **Statistical significance:** statistical-methods allow an estimate to be made of the probability of the observed or greater degree of association between independent and dependent variables under the null hypothesis. From this estimate, in a sample of given size, the statistical significance of a result can be stated. Usually the level of statistical significance is stated by the P-value.

- **Case definition:** a set of diagnostic criteria that must be fulfilled in order to identify a person as a case of a particular disease. Case definition can be based on clinical, laboratory, or combined clinical and laboratory criteria, or a scoring system with points for each criterion that matches the features of the disease. Where the diagnosis is based on a scoring system, e.g., multiple sclerosis, it is important to abide (الالتزام بـ) by the system for surveillance purposes and when deciding whether to include or exclude cases in an epidemiologic study.

- **Case fatality rate:** the proportion of cases of a specified condition which are fatal within a specified time.

$$\text{CFR (usually expressed as a percentage)} = \frac{\text{Number of deaths from a disease (in a given period)}}{\text{Number of diagnosed cases of that disease (in the same period)}} \times 100$$

this definition can lead to paradox (مفارقة-تناقض) when more persons die of the disease than develop it during a given period. For instance, chemical poisoning that is slowly but inexorable (لا محالة) fatal may cause many persons to develop the disease over a relatively short period of time, but the deaths may not occur until some years later and may be spread over a period of years during which there are no new cases. Thus, in calculating the case fatality rate, it is necessary to acknowledge that the time dimension varies: it may be brief, e.g., covering only the period of stay in a hospital; of finite duration, e.g.,



one year; or of longer duration still. The term case fatality rate is then better replaced by a term such as survival rate or by the use of a survivorship table.

- **Surveillance:** systemic ongoing collection, collation (ترتيب), and analysis of data and the timely dissemination (النشر الفوري) of information to those who need to know so that action can be taken. Source: WHO. Surveillance is the essential feature of epidemiologic practice. Its components are fully summarized in the above brief definition. It is distinguished from monitoring by the fact that it is continuous and ongoing, whereas monitoring is intermittent or episodic (متقطع). Surveillance has been defined and described at greater length, e.g., by the Center for Disease Control and Prevention: the ongoing systemic collection, analysis, and interpretation of health data, essential to the planning, implementation (التطبيق), and evaluation of public health practice, closely integrated with the timely dissemination of these data to those who need to know. The final link in the surveillance chain is the application of these data to prevention and control. A surveillance system includes a functional capacity for data collection, analysis, and dissemination linked to public health programs. Another definition gives some details about sources of surveillance data: continuous analysis, interpretation, and feedback of systemically collected data, generally using methods distinguished by their practicality, uniformity, and rapidity rather than by accuracy or completeness. By observing trends in time, place, and persons, changes can be observed or anticipated and appropriate action, including investigative or control measures, can be taken. Sources of data may relate directly to disease or to factors influencing the disease. Thus they may include:
  - Mortality and morbidity reports based on death certificates, hospital records, general practice sentinels, or notifications.
  - Laboratory diagnosis.
  - Outbreak reports.
  - Vaccine utilization-uptake and side effects.
  - Sickness absence records.
  - Disease determinants such as biological changes in agent, vectors or reservoirs.
  - Susceptibility to disease, as by skin testing or serological surveillance e.g., serum banks.
- **Iceberg phenomenon:** that portion of disease which remain unrecorded or undetected despite physicians' diagnostic endeavors (مسابعي) and community disease surveillance procedures is referred to as the "submerged portion of the iceberg". Detected or diagnosed disease is the "tip of the iceberg". The submerged portion comprises disease not medically attended, medically attended but not accurately diagnosed, and diagnosed but not reported. Other terms have been proposed to describe this concept in parts of the world where icebergs are unknown, e.g., "ears of the hippopotamus" "crocodile's nose".