Epidemiology is the science of studying distribution with regards to the various determinants that play a significant role in health-related issues within a specified population as well as the implications of the study results in order to control health problems. Descriptive epidemiology is intended to classify the details of a disease based on person, place and timing factors. The methods of causal reasoning, which are used to test the hypothesis are carried out mainly using inputs derived out of descriptive epidemiology. The various aspects of descriptive epidemiology are covered in the following article.

**Definition of Epidemiology**

Study of the distribution and determinants of states or events in specified populations and the application of this study to the control of health problems.

**Classification of Epidemiology**

*Experimental epidemiology* is the study of the relationships of the various factors that determine the frequency and distribution of diseases in a community. The most common type of experimental epidemiology is a randomized clinical trial. In experimental epidemiology, the deliberate manipulation of the cause is predictably followed by an alteration in the effect not due to chance.
Observational Epidemiology can be classified into two parts, namely descriptive epidemiology, and analytical epidemiology. Descriptive epidemiology seeks an answer to the descriptive questions of -who, -what, -when, -where. It is the study of occurrence and distribution of disease. Analytical epidemiology seeks an answer to the questions -why and -how. With regards to the type of studies, descriptive epidemiology deals with the frequency and the distribution of risk factors in populations and enables them to assess the extent of a disease. It can thus provide testable hypotheses (educated guesses about an association that is testable in a scientific investigation) of etiologic research. Analytical epidemiology aims to research and study risk and protector factors of diseases.

Differences between descriptive and analytical epidemiology

Descriptive epidemiology is mainly applied when knowledge about a particular disease is limited and it is intended only to illustrate a potential association. In addition, it mainly relies on data that was previously collected. It focuses on correlational studies of populations, and on individual case reports, series or cross-sectional studies.

Analytical epidemiology, on the other hand, is applied when information about a particular disease is already available. The key is in evaluating the causality associations of the disease. It uses two widely used measures: Relative Risk (or risk ratio) and Odds Ratio.

The Relative Risk (RR) is the risk of the disease in the exposed group divided by the risk of the disease in the non-exposed group:

\[
RR = \frac{a}{a+b} \frac{c}{c+d}
\]

where 

- \(a\) = # of exposed and those who have the outcome
- \(b\) = # of exposed but those who do not have the outcome
- \(c\) = # of not exposed and those who have the outcome
- \(d\) = # of not exposed and those who do not have the outcome

It follows that:

- \(a+b\) = total # of the exposed
- \(c+d\) = total # of the not exposed
- \(a+c\) = total # who have the outcome
- \(b+d\) = total # who do not have the outcome
- \(a+b+c+d\) = total study population

Odds Ratio (OR) is the tool to measure the risk of a disease when the population at risk is unknown. It is used in case-control studies.

\[
OR = \frac{a \times d}{b \times c}
\]
Both RR and OR are interpreted as follows:

- $=1$ indicates **NO association**
- $>1$ indicates **a positive association**
- $<1$ indicates **a negative association**

When is an analytic test significant? When it is unlikely that the observed association is due to chance.

The two main tests are: 95% Confidence Interval and $p$-value

**Advantage of descriptive epidemiology**

The cost incurred in conducting descriptive epidemiological studies is relatively low compared to that of analytical studies. The whole process is less time-consuming.

After conducting these studies, the **general factors** about a disease such as the highest and lowest rate of occurrence within the population and the temporal patterns of a particular disease are uncovered to a greater extent. Descriptive epidemiology data greatly assist the government in their decisions of resource allocation in order to tackle the health problem on a priority basis.

**Descriptive Epidemiology: Overall detail of – the who, the where and the when**

**Case Definition**

Case Definition is a set of Standard Diagnostic Criteria that must be fulfilled in order to identify a person as a case of a particular disease. This ensures that all persons who are counted as cases actually have the same disease. Case definition includes clinical criteria and may include restrictions on time, place and person.

The factor "**person**" consists of describing a person in terms of race, age, and sex. The factor "**place**" consists of describing a disease in terms of the geographical location within which it is found. The difference in the prevalence of the disease could be due to the same factor in two locations. The factor "**time**" involves the distribution of the disease based on the time span in which it occurs and is usually shown as a graph with the number of cases on the $-y-$ and time on the $-x-$axis.

The **time factor** represents secular (the disease might occur over a long period of time), periodic, seasonal and epidemic. **Cyclic** is an alteration that keeps occurring in a cyclical fashion. The **variation** which occurs may be due to seasonal factors that vary and recur every year or a number of years.

A disease is classified as an **epidemic** if, within a particular set location, the occurrence of the disease affects more individuals than would be expected. Other terminologies based on this occurrence include those of endemic (the occurrence is similar to what would be expected from the habitual occurrence), epidemic (the clear excess in the occurrence of a particular disease) and pandemic (the diseases affecting a worldwide population).

The diagnosis of **cholera** as a pathological agent by John Snow signifies the importance of recording the place factor in the descriptive epidemiology.
Incidence and prevalence

In a population at risk for a particular disease, the proportion of all new cases constitutes the incidence. In a population affected by a particular disease, the proportion of all present cases constitutes the prevalence. The mortality rate, on the other hand, is a marker of the proportion, taken over a particular lineated time period, of the individuals who have died.

In a defined population under study, the frequency at which a particular event occurs at a specific time interval constitutes the rate. A rate is a special ratio in which the two terms are in different units. For example, if a 12-ounce can of corn costs 69¢, the rate is 69¢ for 12 ounces or 5.75¢/ounce. A unit rate is a rate with a denominator of 1, while the ratio can be expressed as fractions or as a decimal.

Incidence

This represents the number of new cases of a disease that occurred during a specified period of time divided by the number of persons at risk of developing the disease during that same period of time.

Incidence = \( \frac{\text{number of new cases over a specific time}}{\text{number of persons at risk of disease over that specific period of time}} \)

(can also be expressed as a number of cases per 1000 persons)

Incidence can be used to help determine the causes of disease and to determine the likelihood of developing a disease.

Incidence density and cumulative incidence

Sometimes the framework within which we study a disease might keep on varying. This brought about the concept of the person-time. This is basically like a multiplication factor which represents the rate of an occurrence of a new case in a particular disease (which is nothing but the incidence) over a varied time length. This is referred to as the incidence density.

Incidence density = Number of new cases / (Population at risk * Duration of risk)

If for example you monitor 10 people and, among them, 5 get diagnosed with malaria. To get the incidence density you first add up all the varying time points and then divide 5 by the total time periods of all patients (this is represented by a patient’s years).

When we have a specific time period rather than a varying time period, the concept of cumulative incidence occurs. In this case, we are able to see the rate of all the new cases which have occurred over the time period of interest.

Cumulative incidence rate = Number of new cases / Population at risk at the beginning of the study

In a particular location of Egypt, there is a total of 100,000 people monitored over a period of 5 years and at the end around 1,000 people developed the disease under study. The cumulative incidence rate in this example is equivalent to 0.002.
Prevalence

In a particular time point, the number of people who have a particular disease in the population under consideration constitutes the prevalence. Rather than the incidence, one of the important differences is that it constitutes the ratio and takes into account both the new cases and those already present in order to obtain the final value.

Prevalence = \( \frac{\text{Number of Cases}}{\text{Number of people in the population}} \)

(Can also be expressed as a number of cases per 1,000 population)

It is useful to determine the burden of disease within a population, though it is not useful to determine the cause of the disease.

Types of Observational Studies

1. Cross-sectional study
2. Cohort study
3. Case-control study

Correlation and case report

In a correlation study, the general characteristics of the population are measured. This is essential so as to get ideas to conduct further studies. The case report is where we report the findings of either an individual case or those of a group of cases with similar features (either as a diagnosis or some other factor).

Statistical association studies can not be done using the records of a case report. A cross-section study is one where we only measure the disease prevalence along with its level of exposure simultaneously within a population. Though the prevalence of the disease can be concluded, it can be difficult to find the incidence.

Case fatality ratio and proportional mortality ratio

It should be noted that not all diseases are fatal. When we get affected by a particular disease, the probability of dying from the disease constitutes the case fatality ratio.

Secondly, though a particular disease may be fatal, the overall impact on death by the disease compared to that of other diseases within a population might be low. The proportion of death from a particular disease when compared to that of the whole population constitutes the proportional mortality ratio.

Prevalence based on the view of time

The query about the disease might naturally come as the query about the occurrence of the disease at a particular time point or over a given period. These constitute the point and period prevalence.

When the ratio of the diseased people over that of the population is taken over a particular time period, it represents the point prevalence whereas the proportion of the diseased people over that of the population over the period of time is the period prevalence.
Mortality Rate Concepts

Crude death rate

As the name suggests, we just make a crude estimate of the death rate within a given population. In a single year, the total number of deaths that occur for every 1,000 people constitute the crude death rate.

Perinatal mortality rate

The word ‘peri’ means in and around. The word natal means the event of childbirth. The perinatal mortality represents the mortality of neonates (WHO definition: within 7 days after birth) along with that of fetal death (greater than 28 weeks of the intrauterine life, other definitions have been used) for every 1,000 births.

Maternal mortality rate

For every 100,000 live births that occur, the number of maternal deaths constitutes the maternal mortality rate and it represents the gross health care system of the society. Many countries use this as a decisive factor when improving health care services.

Childbirth is a natural process and the mother is generally responsible for the survival of the child during the event. The untimely death of the mother is rare. So a considerable childbirth event is taken to obtain a comparable value.

Infant mortality rate

The number of children deaths, for every 1,000 live births, less than a year old constitute the infant mortality rate. This is also one of the sensitive indicators in the health care setup of a particular country. Every country aspires to keep the infant mortality rate as well as the maternal mortality rate as low as possible.

Standardized mortality ratio

For every population, there is a set mortality ratio based on age, gender, and other similar factors, whereby the mortality is taken as the proportion of the number of deaths between a studied population and that of the standard population. This constitutes the standardized mortality ratio.

Case fatality ratio and proportion mortality ratio

This can be better understood through a simple example. When you get a common cold, is there a need to panic? Though there might be a problem, it is unlikely in most circumstances that a common cold by itself will cause death to a patient.

But at the same time, if a person is diagnosed with fulminant hepatic failure, it is absolutely essential to get admitted to a hospital and get treated immediately. This is important as the fatality rate is higher in such a disease. This is comparable to the case fatality ratio. It represents the probability of dying from a particular disease after having contracted it.

\[
\text{Case fatality rate} = \frac{\text{Number of deaths due to the disease under consideration}}{\text{Total number of people who contracted the disease}}
\]
The second concept can be explained by the fact that though some diseases are very dangerous and may cause death, the total occurrences of the diseases might be lower compared to that of other diseases. On the other hand, though a disease may be less fatal compared to other diseases, it may be more common in terms of its occurrence. This is the proportional mortality ratio and it represents the number of people who die from a particular disease out of the total number of deaths.

**Proportional mortality rate = Number of deaths due to the disease under consideration/total number of deaths ∈ population**

**Measurements of Morbidity**

A rate is a measure of the frequency with which an event occurs in a defined population in a defined time. Its denominator must include a measure of time.

A ratio is a value obtained by dividing one quantity by another, for example, the proportion of men to women in a sample.

**Population Pyramid**

In a population, the distribution of the ages may be haphazard or it may follow a normal pattern. The difference also lies in the sex. The population pyramid gives us an opportunity to view both the age as well as the gender differences within that population represented as a single figure. This helps us to better appreciate the differences.

**Evidence-Based Medicine (EBM)**

Evidence-based medicine is the integration of the best research evidence with clinical expertise and patient values. It's a way to use the literature to help you make clinical decisions in a systemic fashion [Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn’t. Br Med J 1996;312:71-72]. EBM is the application of the critical thinking aspects of Epidemiology to the process of making clinical decisions. What is “best research evidence”? It is clinically relevant research, often from the basic sciences but typically from medical literature. What is “clinical expertise”? It is the ability to use your clinical skills and past experience to rapidly identify each patient’s unique health state and diagnosis, individual risks and benefits of potential intervention, and their personal values and expectations. 

What is “best research evidence”? It is clinically relevant research, often from the basic sciences but typically from medical literature. What is “clinical expertise”? It is the ability to use your clinical skills and past experience to rapidly identify each patient’s unique health state and diagnosis, individual risks and benefits of potential intervention, and their personal values and expectations. What are “patient values”? It is the unique preferences, concerns, and expectations each patient brings to a clinical encounter and which must be integrated into clinical decisions if they are to serve the patient.

EBM has been around since post-revolutionary Paris. But recent interest has been spurred by 4 realizations:

1. Doctors need daily information about the diagnosis, prognosis, therapy, and prevention.
2. Textbooks are often out of date; experts are often wrong; CMEs are often useless; there are too many journal articles.
Clinician’s technical knowledge declines over time.

Only a few seconds can be afforded per patient for assimilating a mountain of evidence, and only 30 minutes per week can be set aside for general reading.

Recent developments have made EBM possible, like evidence-based journals, new information system (i.e. computers), new attitudes toward lifelong learning and professional development, strategies for efficiently finding and appraising evidence, and systemic reviews and concise summaries of ongoing research.

One can actually practice EBM by converting the need for information (about prevention, diagnosis, prognosis, therapy, causation, etc.) into an answerable question; by finding the best evidence with which to answer that question; by critically appraising that evidence for its validity, impact, and applicability; by assessing the discovered evidence for its epidemiological rigor, retaining only the best; by integrating the evidence with your clinical expertise, circumstances and values, in order to make a clinical decision; by evaluating your effectiveness and efficiency in improving these steps.

Validity means the closeness to the truth or the ‘real world’. Impact means the effect size, or big of a clinical change was measured. Applicability asks how useful the result was to your specific situation?

Types of studies used to inform EBM

1. RCT: groups of patients are randomized into either experiment or control groups.
2. Cohort (prospective): following exposed and unexposed patients forward to determine the outcome.
3. Case-control (retrospective): looking at patients with the outcome of interest and looking back to see if they had the exposure in question.
4. Case series: a report on a series of patients with an outcome of interest; no control group involved.
5. Systemic review: a summary of the literature that uses explicit methods to appraise and combine studies.
6. Meta-analysis: a systemic review that uses quantitative methods to summarize the results.

What to do with this information? Use your well-phrased research question to search for evidence. Apply the Pyramid of Evidence to determine which studies should carry more weight.

Hierarchy of evidence — the evidence pyramid
References


Utilization and Application of Public Health Data in Descriptive Epidemiology via nih.gov

**Legal Note:** Unless otherwise stated, all rights reserved by Lecturio GmbH. For further legal regulations see our [legal information page](#).