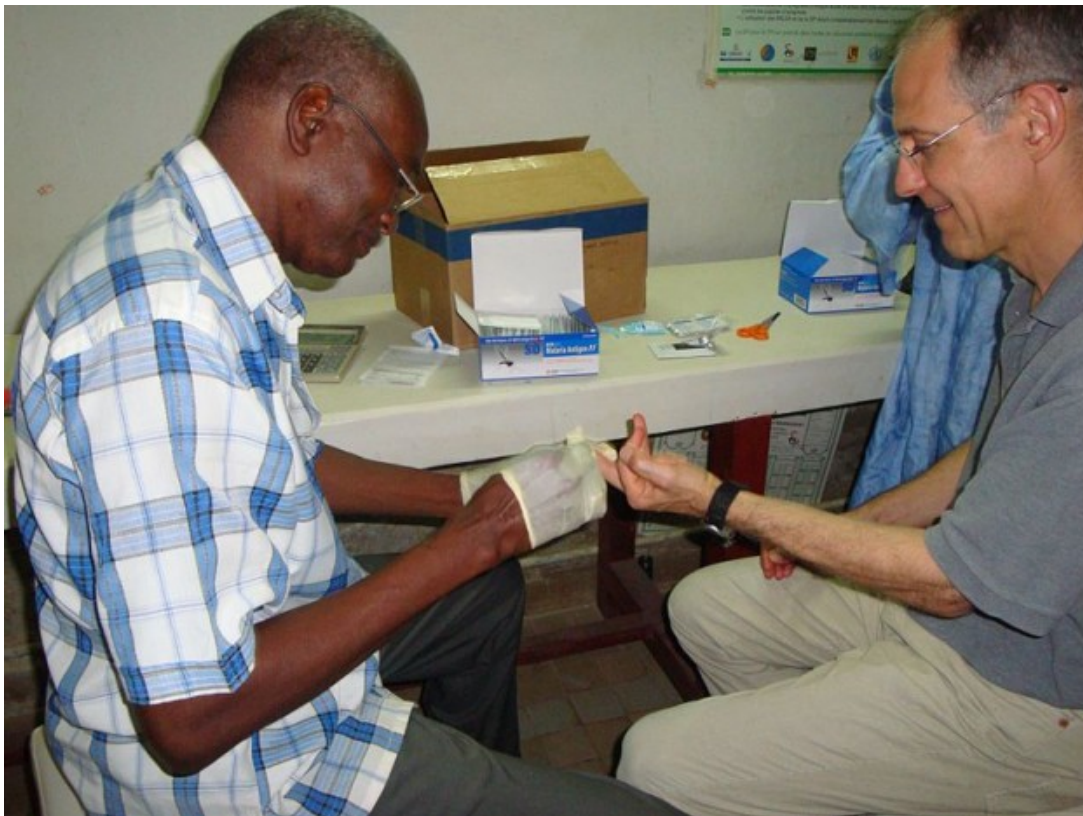


## General Medicine and Epidemiology: Statistical Principles

[See online here](#)

**The terms sensitivity, specificity, positive and negative predictive values constitute the statistical quality criteria of diagnostic testing. You should not only know their meaning and interpretation while preparing for your written exams as, if you want to make (responsible) treatment decisions as a doctor, there is no way around these parameters. Incidence rate and pretest probability are further parameters that you need to be familiar with. This article will provide you with an overview of various statistical quality criteria and their interpretation.**



### Diagnostic Testing

In medicine, a **diagnostic test** examines whether a disease or characteristic is present or not. Examples are tests for metabolic diseases during **new-born screening**, HIV testing, or even a simple pregnancy test. Broadly speaking, there is a distinction between a **screening test** and a **confirmatory test**.

**Screening tests** are designed to detect early disease or the presence of risk factors in a large population of healthy individuals. Screening tests are usually simple and non-

invasive, thus, individuals easily consent to the procedures. They are designed for high sensitivity to make sure that potential disease is not missed. This bias towards sensitivity comes with an increase in false positives.

**Confirmatory tests** are designed to establish the absence or presence of a disease and form the basis of diagnosis in symptomatic patients or patients who are positive for a screening test. They may be invasive and require some form of informed consent before undertaking them. These tests have a bias towards specificity rather than sensitivity; thus, they have high accuracy rates in the identification of true negatives.



**Image:** The blood of a two-week-old infant is collected for phenylketonuria or PKU screening. By: U.S. Air Force photo/Staff Sgt Eric T. Sheler - USAF Photographic Archives. License: [Public Domain](#).

## Statistical Quality Criteria for Diagnostic Tests

### Sensitivity and specificity

The **sensitivity** of a diagnostic procedure provides information as to what percentage of sufferers is actually recognized as sick by the test method. Sensitivity thus describes '**true positive results**'. For example, the Ebola rapid test (ReBOV test) by the company Corgenix (USA) has a sensitivity of almost 100%. In this sense, a patient suffering from Ebola is properly diagnosed in all cases with this test.

**Specificity** indicates the probability that a medical test for a given disease will recognize that a healthy person is actually healthy. With a specificity of 92%, the aforementioned Ebola rapid test will detect, on average, 92 people as healthy ('**true negative rate**') among 100 healthy patients. Eight healthy individuals would, therefore, be erroneously diagnosed as having Ebola, even though they are healthy, and may have to endure unneeded quarantine measures.

### Positive predictive value

**Positive predictive value** is the probability that people with a positive medical test result are actually ill (right = positive). The positive **predictive** value depends on the prevalence of the disease in the population. With a more common disease, the likelihood

of actually being affected is a lot higher than with very rare diseases.

## Negative predictive value

The **negative predictive value** indicates how many people, whose medical test results turn out negative, are actually healthy. It also depends on the prevalence of the disease in the population.

### Likelihood ratio

This represents the number of times individuals are more likely to have the disease. It is described by the formula: sensitivity/ (1- specificity).

## Prevalence

**Prevalence** is an epidemiological measure for **disease frequency**. It indicates how many people are suffering from disease during a particular period in time (period prevalence) or at a particular date in time (point prevalence). The value refers to a sample that is generalized and extrapolated to the population or to actual numbers.

While relatively accurate data is available for notifiable diseases, others can only be representatively 'extrapolated' through sampling. Prevalence thus describes a current state: how many people in 2015, for example, suffer from AIDS, regardless of when they contracted the disease.

### Incidence

Incidence refers to the rate of new cases of a disease, which is usually reported as the number of new cases of the disease occurring within a period of time (per month or per year).

It is also reported as the fraction of the population that is at risk of developing the disease i.e. per 100,000 or per 1,000,000 of the population.

## The Fourfold Contingency Table

The **fourfold contingency table** combines the abovementioned test quality criteria in one **model**, and it is often encountered in epidemiology.

If the meanings of the parameters are clear, calculations are straightforward. If you have been given values for the quality criteria of certain tests, you can illustrate the ratios of the parameters based on numerical examples.

	<b>Positive Test</b>	<b>Negative Test</b>	<b>Total Test Results</b>	<b>Calculation</b>
<b>Sick</b>	a (true positive)	b (false negative)	a + b (all sufferers)	Sensitivity = $a / (a + b)$
<b>Healthy</b>	c (false positive)	d (true negative)	c + d (all healthy)	Specificity = $d / (c + d)$
<b>Total Persons</b>	a + c (all positive tests)	b + d (all negative tests)		
<b>Calculation</b>	<b>Positive predictive value</b> = $a / (a + c)$	<b>Negative predictive value</b> = $d / (b + d)$		

**Example:** Suppose that 7% of the population in an affected area (this is randomly

chosen, as there are no reliable numbers regarding this) is suffering from Ebola and we are testing all the people in our small clinic. Altogether, 1,000 people are examined with the Ebola-rapid test in order to detect outbreaks early and promptly isolate affected persons.

	Positive test	Negative test	Total Test Results	Calculation
<b>Sick</b>	70	0	70 (7% prevalence)	<b>Sensitivity = 100%</b>
<b>Healthy</b>	65	865	930 (100 - 7 = 93% healthy)	<b>Specificity = 92%</b>
<b>Total Persons</b>	135	865	1.000 (tested persons)	
<b>Calculation</b>	<b>Positive predictive value = 0.52</b>	<b>Negative predictive value = 1</b>		

Out of 1,000 tested persons, 70 people (7%) are actually infected by the virus and their test is positive. Because of the lower specificity of the rapid test, 65 of the healthy 930 people also get a positive Ebola test result. Out of a total of 135 Ebola tests with a positive result, almost half, 65 people in our hospital, will have been diagnosed with a false positive and will have to be treated for safety reasons and put in isolation. Despite the apparently high values, the chance of actually having Ebola with a positive test result, in our example, is only 52%!

Such considerations become important in the medical profession in the case of well-known screening tests, such as mammography for breast cancer or the PSA level determination for prostate cancer prevention. False-positive test may result in uncertainty and invasive diagnosis for the patient, therefore, it is useful to be able to deal with test results and probabilities in order to optimally inform patients and to plan further diagnosis and treatment steps.

## Other Statistical Values

### Pretest probability

**Pretest probability** indicates how likely a test will yield a positive result; therefore, it is dependent on prevalence. For example, the pretest probability for our Ebola rapid test is much lower in East Africa than in Sierra Leone due to a much lower prevalence.

## Descriptive Studies

True 'descriptive study' is an example of 'descriptive epidemiology'.

- Who?
  - Students
- What?
  - Left-handedness
- Where?
  - Your neighborhood
- When?
  - Right now

Currently, 23% of students in your neighborhood are left-handed.

# Historic Triumphs: Smallpox

The global eradication of smallpox was certified, based on intense verification activities in countries, by a commission of eminent scientists on 9 December 1979 and subsequently endorsed by the World Health Assembly on 8 May 1980 - "[Resolution WHA 33.3](#)"

In 1975, 2-year-old Rahima Banu contracted the last known case of naturally occurring variola major Smallpox (Source: [Wikipedia](#)).

Global eradication was possible because of herd immunity.

## History

### 18th century

- 400,000 people died each year
- One-third of the survivors became blind
- Survivors also developed immunity to smallpox
- Edward Jenner took interest in cowpox to find solutions in the prevention of smallpox

### 1950

- Effort to prevent smallpox

### 1967

- Variolation
- WHO began the eradication program
- 15 million people developed smallpox annually
- 2 million people died

### 1980

- WHO and eradication of smallpox

## At a Glance: Statistical Principles

- **Sensitivity:** how many sufferers are 'true positive' and considered truly ill?
- **Specificity:** how many healthy people are considered "true negative" (healthy) by the test?
- **Positive predictive value:** positive test result but truly ill
- **Negative predictive value:** negative test result but truly healthy
- **Prevalence:** how many people are suffering from a particular disease within a specific time period?
- **Incidence:** how many people develop a particular disease during a specific period?
- **Pretest probability:** how likely is the test to be positive? Prevalence - influence!

## References

Broeck, J, & Brestoff, J. R. (2013). *Epidemiology: Principles and practical guidelines*. Dordrecht: Springer.

Checkoway, H., Pearce, N., & Crawford-Brown, D. J. (1989). *Research methods in occupational epidemiology*. New York: Oxford University Press.

Merrill, R. M. (2017). *Introduction to epidemiology*.

Sutherland, I., & Mainland, D. (1953). Elementary Medical Statistics. The Principles of Quantitative Medicine. *Journal of the Royal Statistical Society. Series A (General)*, 116(1), 90. doi:10.2307/2980957

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

Notes