**~~\* Positive predictive value~~**

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2540558/pdf/bmj00448-0038a.pdf

The positive test value is the proportion of patients (positive) who are actually within the total number of subjects (positive) as a result of the diagnostic test. (Pozitif test değeri,tanı testi sonucunda hasta (pozitif) bulunan toplam denekler içerisinde gerçekten hasta (pozitif) olanların oranıdır.) It is possible that someone who is positive for the test result is sick.

**~~\* Negatif predictive value~~**

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2540558/pdf/bmj00448-0038a.pdf

The negative test value is the ratio of those who are truly non-patients (negative) among all subjects who are non-patient (negative) as a result of the diagnostic test. (Negatif test değeri, tanı testi sonucunda non-patient (negatif) bulunan tüm denekler içerisinde gerçekten non-patient (negatif) olanların oranıdır.) The predictive value of negative test of the test result indicates the likelihood that a person known to be negative is not sick.

**~~\* Prevelance~~**

http://onlinelibrary.wiley.com/doi/10.1111/j.1651-2227.2006.00180.x/full

Prevalence is the ratio of the number of people who are truly sick to the total number of people in the study. (Prevelans, gerçekten hasta olanların çalışmadaki toplam kişi sayısına oranıdır.)

**\* ~~Sensitivity~~**

http://onlinelibrary.wiley.com/doi/10.1111/j.1651-2227.2006.00180.x/full

Sensitivity is the proportion of subjects who are actually patient (positive) who have tested the patient (positive). (Duyarlılık, gerçekte hasta (pozitif) olanlar arasında testin hasta (pozitif) sonuç verdiği deneklerin oranıdır.) That is, the sensitivity of the test shows how many of the really sick people are classified correctly (as patients). It only measures how good the test is when looking at patient subjects.

**~~\* Specificity~~**

http://onlinelibrary.wiley.com/doi/10.1111/j.1651-2227.2006.00180.x/full

Specificity is the proportion of subjects who are actually non-patient (negative) who have tested the non-patient (negative). (Özgüllük, gerçekte hasta olmayanlar arasında testin sağlam sonuç verdiği deneklerin oranıdır.) That is, the possibility of a negative test result between those who participate in the study and those who are not ill. The specificity of the test indicates that what is not sick is actually classified correctly (non-sick). It gives a measure of how good the test is when looking at only non-patient subjects.

**~~\* The Probabality of Disease Given a Positive Test (PPV with Bayes Formula)~~**

http://www.medicalbiostatistics.com/sensitivity-specificity.pdf

Sensitivity and specificity are easy to evaluate by a case-control study but predictivity requires that the subjects be followed up till such time that their disease status is confirmed as present or absent. This could be very time consuming and expensive. Thus predictivities are difficult to evaluate. Luckily there is a statistical procedure, called Bayes’ 3 rule, which helps to get one from the other, provided some ancillary information is available. The significance of Bayes’ rule: High fever, rigors, spleenomegaly, and presence of parasite in the blood are the stages that progressively confirm malaria. As the information increases, the diagnosis becomes focused, and the probability of absence or presence of the disease firms up. The probability depends on what information is already available. The chance part is the uncovered information. The probability of any event without availability of any information is called prior probability and the probability after some information is available is called posterior probability. The latter obviously depends on the kind of information available to alter the probability. The function of Bayes’ Rule is to convert one posterior probability to its directional inverse. It converts probability of A given B to probability of B given A. This is useful to convert sensitivity of a test to its predictivity. The former is P(T+|D+) and the latter is P(D+|T+). Such conversion is a great convenience when one probability is already available or can be easily obtained, and its inverse is otherwise difficult to obtain directly.

**~~\* False Positive Rate~~**

https://www.researchgate.net/profile/Arif\_Hanga/publication/275644390\_Brief\_review\_on\_Sensitivity\_Specificity\_and\_Predictivities/links/5541c40b0cf23222273171ba/Brief-review-on-Sensitivity-Specificity-and-Predictivities.pdf

False Positive (also known as false alarm) are predictions that should be false but were predicted as true. A false positive occurs when the test reports a positive result (patient) for a person who is disease free (non-patient).

**~~\* False Negative Rate~~**

https://www.researchgate.net/profile/Arif\_Hanga/publication/275644390\_Brief\_review\_on\_Sensitivity\_Specificity\_and\_Predictivities/links/5541c40b0cf23222273171ba/Brief-review-on-Sensitivity-Specificity-and-Predictivities.pdf

A false negative occurs when the test reports a negative result (non-patient) for a person who actually has the disease (patient). In real patients, the test is accidental, not sick.

**~~\* Accuracy~~**

http://www.cpdm.ufpr.br/documentos/ROC.pdf

Accuracy is the ratio of actual results (true positive or true negative) among subjects participating in the study. It measures the accuracy of the diagnostic test.

**~~\* Likelihood ratio of positive test (LR+)~~**

http://www.ifcc.org/ifccfiles/docs/190404200805.pdf

The Likelihood ratio is the ratio of the test result expected to those who have the disease to those without the disease. That is, a test result indicates how many times the subjects with the disease have the disease compared to the subjects without the disease. When the likelihood ratio is equal to 1, both probabilities (test results in subjects with the disease and test results in subjects without the disease) are equal.

Likelihood ratios for positive test (LR +) show how much more positive the test result will be in patient subjects than in non-patient subjects. This rate is usually greater than 1. The higher the LR +, the greater the test-disease rate.

**~~\* Youden Index~~**

http://www.ifcc.org/ifccfiles/docs/190404200805.pdf

This is an index that summarizes the sensitivity and authenticity of a test. Youden Index is 0 if the diagnostic test has weak accuracy and 1 if it has strong accuracy. It is used for the evaluation of overall discriminative power of a diagnostic procedure and for comparison of this test with other tests. The disadvantage of the Youden Index is that it is not sensitive to differences in sensitivity and specificity. It’s not a bad index as an approximation to the overall performance of the test, but it is not recommended to use it as a single parameter to evaluate any diagnostic test. Youden’s index is not affected by the disease prevalence, but it is affected by the spectrum of the disease, as are also sensitivity specificity, likelihood ratios and odds ratio.

**~~\* Likelihood ratio of negative test (LR-)~~**

The Likelihood ratio is the ratio of the test result expected to those who have the disease to those without the disease. That is, a test result indicates how many times the subjects with the disease have the disease compared to the subjects without the disease. When the likelihood ratio is equal to 1, both probabilities (test results in subjects with the disease and test results in subjects without the disease) are equal.

Likelihood ratio for negative test (LR-) is the ratio of the likelihood of a negative outcome in a patient to the likelihood of a negative outcome in non-ill subjects. This rate is usually less than 1.

**~~\* Odds ratio (OR)~~**

OR is the ratio of the odds ratio of non-ill subjects to the odds ratio of non-ill subjects. DOR depends on the sensitivity and specificity of the test. OR is elevated when sensitivity and specificity are high, false positive and false negative rate is low.

**\* Number needed to treat**

http://www.fpnotebook.com/prevent/Epi/NmbrNdTScrn.htm

Number needed to treat is the number of patients examined to prevent a death from occurring within a specified period of time. If the Number needed to treat is negative, the finding would be consistent with Number Needed to Treat. If the result were positive, the finding would be consistent with number needed to harm.

**~~\* Number Needed to Misdiagnose~~**

http://journals.lww.com/epidem/Fulltext/2013/01000/Number\_Needed\_to\_Misdiagnose\_\_A\_Measure\_of.27.aspx

http://www.biochemia-medica.com/2016/26/297

The number needed to misdiagnose, defined as the number of patients who need to be tested in order for one to be misdiagnosed (FP or FN) by the test. The higher the NNM of a test, the closer is the test to the gold-standard, hence, a better test.

**\* Lift**

http://michael.hahsler.net/research/recomm\_lnai2002/lnai2002.pdf

Lift is an alternative measure for confidence that takes the frequency of the consequent into account. Lift is defined as the relation of the (observed) probability of the co-occurrence of two items to the probability under the assumption that they occur independently.

**\* Precision**

Precision is the ratio of accurately estimated positives to all predicted positives. It falls in the range from 0 to 1, with 1 being the best score. 0.500 is closer to 1, so it is a very good score.

**\* Recall**

Precision is the ratio of accurately estimated positives to all predicted positives. It falls in the range from 0 to 1, with 1 being the best score. 0.500 is closer to 1, so it is a very good score.

**\* E-Measure**

http://michael.hahsler.net/research/recomm\_lnai2002/lnai2002.pdf

To find an optimal trade-off between precision and recall a single-valued measure like the E-measure can be used.

**\* Matthew's Correlation Coefficient**

https://en.0wikipedia.org/index.php?q=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnL3dpa2kvTWF0dGhld3NfY29ycmVsYXRpb25fY29lZmZpY2llbnQ

The Matthews correlation coefficient is used in [machine learning](https://en.0wikipedia.org/index.php?q=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnL3dpa2kvTWFjaGluZV9sZWFybmluZw) as a measure of the quality of binary (two-class) [classifications](https://en.0wikipedia.org/index.php?q=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnL3dpa2kvQmluYXJ5X2NsYXNzaWZpY2F0aW9u), introduced by biochemist [Brian W. Matthews](https://en.0wikipedia.org/index.php?q=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnL3dpa2kvQnJpYW5fTWF0dGhld3NfKGJpb2NoZW1pc3Qp) in 1975. It takes into account true and false positives and negatives and is generally regarded as a balanced measure which can be used even if the classes are of very different sizes. The MCC is in essence a correlation coefficient between the observed and predicted binary classifications; it returns a value between −1 and +1. A coefficient of +1 represents a perfect prediction, 0 no better than random prediction and −1 indicates total disagreement between prediction and observation.

**~~\* Gain in Certainty~~**

http://en.citizendium.org/wiki/Sensitivity\_and\_specificity#Number\_needed\_to\_diagnose

https://www.ncbi.nlm.nih.gov/pubmed/4014166

https://www.biomedcentral.com/1742-5573/content/3/1/11

http://www.psycho-oncology.info/PG\_analyse\_ajmitchell.pdf

The gain in the certainty that a condition is present is the difference between the post-test probability and the prior probability (the prevalence) when the test is positive. The gain in certainty that there is no disease is the difference between post-test probability of no disease and the prior probability of no disease (1-prevalence). Gain in certainty varies from 0 to 2 and a result of 1 indicates that the diagnostic test does not add to guessing; that is, the test provides no information; when Gain in certainty is greatest, the expected gain is maximized. Expected gain is also related to the receiver operating characteristic curve for a diagnostic test: the point on the receiver operating characteristic curve at which Gain in certainty is greatest corresponds to the point at which the distance from the major diagonal is greatest at which the slope of the receiver operating characteristic curve equals 1.

**~~\* Number needed to diagnose~~**

https://rdrr.io/cran/epiR/man/epi.tests.html

http://www.psycho-oncology.info/PG\_analyse\_ajmitchell.pdf

 The number needed to diagnose is defined as the number of paitents that need to be tested to give one correct positive test. NNT is used to calculate the number of patients that need to be examined in order to correctly detect one person with the disease.

**~~\* Misclassification rate (Error classification rate)~~**

https://www.omicsonline.org/measures-derived-from-a-2-x-2-table-for-an-accuracy-of-a-diagnostic-test-2155-6180.1000128.php?aid=3010&view=mobile

The misclassification rate is the proportion of those individuals incorrectly categorized by the test those with disease who had a negative test plus those without disease who had a positive test result). This measures the portion of all decisions that were incorrect decisions. It falls in the range from 0 to 1, with 0 being the best score.

**~~\* Rand Index (Rand Measure)~~**

https://en.0wikipedia.org/index.php?q=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnL3dpa2kvUmFuZF9pbmRleA

The Rand index or Rand measure (named after William M. Rand) in [statistics](https://en.0wikipedia.org/index.php?q=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnL3dpa2kvU3RhdGlzdGljcw), and in particular in [data clustering](https://en.0wikipedia.org/index.php?q=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnL3dpa2kvRGF0YV9jbHVzdGVyaW5n), is a measure of the similarity between two [data clusterings](https://en.0wikipedia.org/index.php?q=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnL3dpa2kvRGF0YV9jbHVzdGVyaW5n). From a mathematical standpoint, Rand index is related to the [accuracy](https://en.0wikipedia.org/index.php?q=aHR0cHM6Ly9lbi53aWtpcGVkaWEub3JnL3dpa2kvQWNjdXJhY3lfYW5kX3ByZWNpc2lvbiNJbl9iaW5hcnlfY2xhc3NpZmljYXRpb24), but is applicable even when class labels are not used. The Rand index has a value between 0 and 1, with 0 indicating that the two data groups do not agree on any pair of points and 1 indicating that the data groups are exactly the same.

**~~\* Balanced Accuracy (Average Accuracy)~~**

The balanced accuracy is the average of sensitivity and specificity can be defined also as the average accuracy obtained on either class.

**\* Discriminant power**

Discriminant power (DP) is a measure that summarizes sensitivity and specificity.

**\* F1 Score**

This measure is the combination of the harmonic averages of precision and recall. It falls in the range from 0 to 1, with 1 being the best score.

**\* Predictive Summary Index (PSI)**

The new Predictive Summary Index (PSI) reflects the true total gain in certainty obtained by performing a diagnostic test based on knowledge of disease prevalence, i.e., the overall additional certainty. We show that the overall gain in certainty can be expressed in the form of the following expression: PSI = PPV+NPV-1. PSI is a more comprehensive measure than the post-test probability or the Youden Index (J). PSI can be derived in the target (patient) population as a measure of the goodness of the predictability in a diagnostic test, using alpha and beta errors in the target population. If the test is always correct, all errors equal 0 and PSI = 1. Negative values of PSI (between -1 and 0) occur if the test is misleading, i.e., occurrence of disease is negatively associated with tests results.