

All of the statistics are calculated based on a so-called "contingency table", which looks like this:

	Correct=Y	Correct=N	
Assigned=Y	a	b	
Assigned=N	c	d	

a, b, c, and d are counts that reflect how the assigned categories matched the correct categories. Depending on whether a macro-statistic or a micro-statistic is being calculated, these numbers will be tallied per-category or for the entire result set.

The following statistics are available:

- accuracy

This measures the portion of all decisions that were correct decisions. It is defined as $(a+d)/(a+b+c+d)$. It falls in the range from 0 to 1, with 1 being the best score.

Note that macro-accuracy and micro-accuracy will always give the same number.

- error

This measures the portion of all decisions that were incorrect decisions. It is defined as $(b+c)/(a+b+c+d)$. It falls in the range from 0 to 1, with 0 being the best score.

Note that macro-error and micro-error will always give the same number.

- precision

This measures the portion of the assigned categories that were correct. It is defined as $a/(a+b)$. It falls in the range from 0 to 1, with 1 being the best score.

- recall

This measures the portion of the correct categories that were assigned. It is defined as $a/(a+c)$. It falls in the range from 0 to 1, with 1 being the best score.

- F1

This measures an even combination of precision and recall. It is defined as $2*p*r/(p+r)$. In terms of a, b, and c, it may be expressed as $2a/(2a+b+c)$. It falls in the range from 0 to 1, with 1 being the best score.

診斷：敏感度 (sensitivity)、特異度 (specificity)、概似比 (likelihood ratio)、檢測前機率 (pre-test probability)、檢測後機率 (post-test probability)

Diagnostic test (ferritin)	Disease (Iron deficient anemia)	
	Present	Absent
Positive	731 (a)	270 (b)
Negative	78 (c)	1500 (d)

$$\text{Sensitivity} = a/a+c = 731/809 = 90\%$$

$$\text{Specificity} = d/b+d = 1500/1770 = 85\%$$

$$\text{Positive predictive value} = a/a+b = 731/1001 = 73\%$$

$$\text{Negative predictive value} = d/c+d = 1500/1578 = 95\%$$

不同的疾病盛行率 (prevalence) 會影響 diagnostic test 的 sensitivity and specificity, EBM 常用 likelihood ratio (LR), odds (一種無分母的機率表示方法) 來表達 probability, odds and probability 兩者也可以互相換算 (probability= odds / odds+1)。

$$\text{LR+ for a positive result} = \text{sens}/(1- \text{spec}) = 90\%/15\% = 6$$

$$\text{LR- for a negative result} = (1-\text{sens})/\text{spec} = 10\%/85\% = 0.12$$

$$\text{Pre-test probability (prevalence)} = a+c/a+b+c+d = 31\%$$

$$\text{Pre-test odds} = \text{prevalence}/(1-\text{prevalence}) = 31\%/69\% = 0.45$$

$$\text{Post-test odds} = \text{Pre-test odds} \times \text{Likelihood Ratio}$$

$$\text{Post-test probability} = \text{Post-test odds}/(\text{Post-test odds} + 1)$$