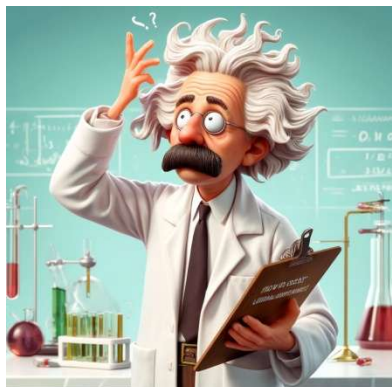


Performance Assessment of Binary Output Examinations in Medical Laboratories



1 - What is fitness for purpose?

The test meets the necessary performance criteria for a clinical binary examination **to ensure it is suitable for making accurate and reliable clinical decisions in the intended patient population.** Among others, performance criteria include clinical sensitivity and clinical specificity.

2 - How to assess clinical performance?

Assessing the clinical performance of a binary examination, stating results as 'positive' or 'negative,' involves calculating the examination's

Clinical Sensitivity and Clinical Specificity. From the physician's point of view, clinical performance is evaluated by using predictive values, i.e., how likely is it that a positive or negative result corresponds to the presence or absence of the condition, which is also influenced by the prevalence of the condition in the population being tested.

3 - What is Clinical Sensitivity?

	Clinical Diagnosis		Result Totals
	Disease, D_1	Non-Disease, D_0	
Positive, p	tp	fp	$tp + fp$
Negative, n	fn	tn	$tn + fn$
Case Totals	$tp + fn$	$tn + fp$	N

Table 1. 2x2 Contingency table for clinical diagnosis.

Clinical sensitivity (SS), also known as the True Positive Rate, measures the proportion of true positives correctly identified by the test.

Data from a group of subjects who have been confirmed to have the disease or condition through diagnosis or a gold standard test (D_1 in Table 1) are needed to calculate clinical sensitivity.

The formula for clinical sensitivity is:

$SS = tp / (tp + fn)$, with:

tp , True Positives are the cases where the test correctly identifies the presence of the disease.

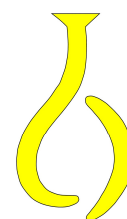
fn , False Negatives are the cases where the test incorrectly identifies the absence of the disease when it is actually present.

Table 1 represents the components of clinical sensitivity and specificity. Figure 1 illustrates an example of the overlapping of the number of components. The symbol " \cap " represents the intersection of groups. For example, $D_1 \cap n$ denotes the set of negative results obtained for known positive cases, i.e., false negatives.

4 - What is Clinical Specificity?

Clinical Specificity (SP), also known as the True Negative Rate, measures the proportion of true negatives correctly identified by the test.

Data from a group of subjects who have been confirmed not to have the disease or condition through a gold standard test (D_0 in Table 1) are needed to calculate clinical specificity.



Eurachem

A FOCUS FOR
ANALYTICAL CHEMISTRY
IN EUROPE

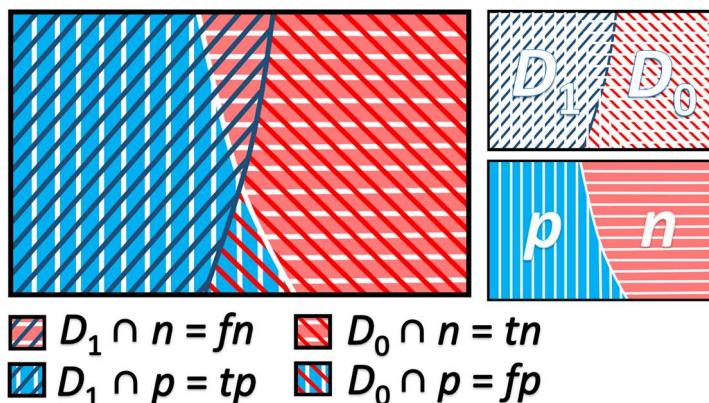


figure 1. Graphical representation of an example of $tp(D_1 \cap p)$, $tn(D_0 \cap n)$, $fn(D_1 \cap n)$, and $fp(D_0 \cap p)$.

uncertainty **involves statistical methods that estimate each parameter's confidence interval (CI)**. These intervals provide a range of values within which the true sensitivity or specificity is expected to lie with a certain level of confidence, **typically 95 %**. Please refer to the AQA:2021 for information about statistical models to compute the 95 % CI [1].

6 - How to define targets?

Target minimum values for clinical sensitivity and specificity should be identified. These values may originate from regulations or criteria based on the intended use of the binary results. A target or minimum value for the lower limit of the 95 % CI for result rate, i.e., SS or SP , ($LL_{SS,95}$ or $LL_{SP,95}$) or the upper limit of the 95 % CI for result rate ($HL_{SS,95}$ or $HL_{SP,95}$) should also be defined according to the purpose of the analysis. The target or minimum value (LL_{SS}^{tg} or LL_{SP}^{tg}) is particularly significant when the impact of false results is critical. For example, for blood components used in transfusions, the screening for infectious diseases should be performed with tests associated with an $LL_{SS,95}$ close to 100 %.

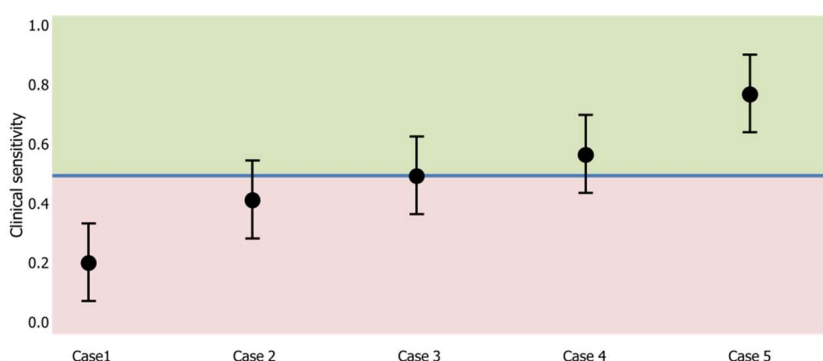


Figure 2. Compliance assessment with clinical sensitivity uncertainty using the lower 95% confidence interval limit for five performance assessment scenarios.

Figure 2 illustrates five different tests, each assessed for clinical sensitivity and its associated uncertainty. These cases reflect different levels of compliance with a target sensitivity of 0.5. Case 1 is non-compliant, Cases 2 to 4 are partially compliant, and Case 5 is compliant. From a simplified statistical power view, the laboratory should always take into account the minimum number of samples when defining its target. This number can be estimated by simulation. For example, for a clinical sensitivity of 90 % the evaluation will require a minimum of 10 samples. It should be emphasized that samples must be representative of the target population for a clinical power evaluation - for example, epidemiologically representative samples of the types and subtypes of a given virological agent are required.

Reference

[1] R. Bettencourt da Silva and S. L. R. Ellison (eds.) Eurachem/CITAC Guide: Assessment of performance and uncertainty in qualitative chemical analysis. (1st ed. 2021). Available from www.eurachem.org.