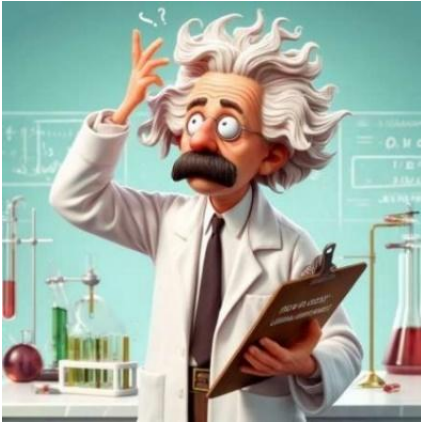


Leistungsbewertung binärer Ergebnisprüfungen in medizinischen Laboratorien



1 - Was ist Zweckmäßigkeit?

Der Test erfüllt die erforderlichen Leistungskriterien für eine klinische binäre Untersuchung, **um sicherzustellen, dass er für genaue und zuverlässige klinische Entscheidungen bei der vorgesehenen Patientengruppe geeignet ist.** Zu den Leistungskriterien gehören unter anderem die klinische Sensitivität und die klinische Spezifität.

2 - Wie lässt sich die klinische Leistung bewerten?

Die Bewertung der klinischen Leistungsfähigkeit einer binären Untersuchung, bei der die Ergebnisse als „positiv“ oder „negativ“ angegeben werden, umfasst die Berechnung der **klinischen Sensitivität und klinischen Spezifität** der Untersuchung. Aus Sicht des Arztes wird die klinische Leistungsfähigkeit anhand von Vorhersagewerten bewertet, d. h. wie wahrscheinlich es ist, dass ein positives oder negatives Ergebnis mit dem Vorliegen oder Nichtvorliegen der Erkrankung korrespondiert, was auch von der Häufigkeit der Erkrankung in der getesteten Population beeinflusst wird.

3 - Was ist klinische Sensitivität?

	Klinische Diagnose		Gesamtergebnis
	Krankheit, D_1	keine Krankheit, D_0	
Positiv, p	tp	fp	$tp+fp$
Negativ, n	fn	tn	$tn+fn$
Gesamtfälle	$tp+fn$	$tn+fp$	N

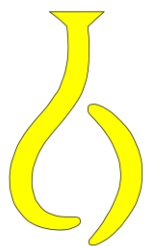
Tabelle 1. 2x2 Kontingenztafel für die klinische Diagnose.

Die klinische Sensitivität (SS), auch bekannt als Wahr-Positiv-Rate, misst den Anteil der echten positiven Ergebnisse, die vom Test korrekt identifiziert wurden. Zur Berechnung der klini-

schen Sensitivität werden Daten einer Gruppe von Probanden benötigt, bei denen die Krankheit oder der Zustand durch eine Diagnose oder einen Goldstandardtest (D_1 in Tabelle 1) bestätigt wurde.

Die Formel für die klinische Sensitivität lautet:

$$SS = tp / (tp + fn), \text{ wobei}$$



Eurachem

A FOCUS FOR
ANALYTICAL CHEMISTRY
IN EUROPE

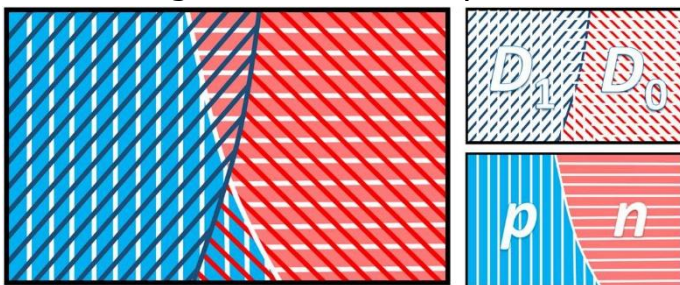
tp , wahr positiv sind Fälle, in denen der Test das Vorliegen der Krankheit korrekt identifiziert.

fn , falsch negativ sind Fälle, in denen der Test fälschlicherweise das Nichtvorhandensein der Krankheit feststellt, obwohl diese tatsächlich vorliegt.

Tabelle 1 zeigt die Komponenten der klinischen Sensitivität und Spezifität. Abbildung 1 veranschaulicht ein Beispiel für die Überschneidung der Anzahl der Komponenten. Das Symbol „ \cap “ steht für die Schnittmenge von Gruppen. Beispielsweise bezeichnet $D_1 \cap n$ die Menge der negativen Ergebnisse, die für bekannte positive Fälle erzielt wurden, d. h. falsche Negative.

4 - Was ist klinische Spezifität?

Die klinische Spezifität (SP), auch als Wahr-Negativ-Rate Negative Rate bezeichnet, misst den Anteil der durch den Test korrekt identifizierten Wahr- Negativbefunde. Zur Berechnung der klinischen Spezifität werden Daten einer Gruppe von Probanden benötigt, bei denen durch einen



$\square D_1 \cap n = fn$ $\square D_0 \cap n = tn$
 $\square D_1 \cap p = tp$ $\square D_0 \cap p = fp$

Abbildung 1. Grafische Darstellung eines Beispiels von: tp ($D_1 \cap p$), tn ($D_0 \cap n$), fn

benötigt, bei denen durch einen Goldstandard-Test (D_0 in Tabelle 1) bestätigt wurde, dass sie nicht an der Krankheit oder dem Zustand leiden

Die Formel für die klinische Spezifität lautet:

$SP = tn / (tn + fp)$, wobei:

tn , wahr negativ sind die Fälle, in denen der Test das Nichtvorhandensein der Krankheit korrekt identifiziert.

fp , falsch positiv sind die Fälle, in denen der Test fälschlicherweise das Vorhandensein der Krankheit identifiziert, obwohl diese tatsächlich nicht vorliegt.

5 - Wie berechnet man die Unsicherheit der klinischen Sensitivität und Spezifität?

Um die Unsicherheit der klinischen Sensitivität und Spezifität zu berechnen, werden **statistische Methoden zur Schätzung des Konfidenzintervalls (CI) jedes Parameters benötigt**. Diese Intervalle geben einen Wertebereich an, innerhalb dessen die tatsächliche Sensitivität oder Spezifität mit einer bestimmten Konfidenz, **in der Regel 95 %**, zu erwarten ist.



Eurachem

A FOCUS FOR
ANALYTICAL CHEMISTRY
IN EUROPE

Informationen zu statistischen Modellen zur Berechnung des 95 %-CI finden Sie in AQA:2021 [1].

6 - Wie definiert man Zielwerte?

Ziel-Mindestwerte für die klinische Sensitivität und Spezifität sollten festgelegt werden. Diese Werte können sich aus Vorschriften oder Kriterien ergeben,

abhängig von der beabsichtigten Verwendung der binären Ergebnisse. Ein Ziel- oder Mindestwert für die Untergrenze des 95 %-Konfidenzintervalls für die Ergebnisrate, d. h. SS oder SP, (LLSS.95 or LLSP.95) oder die Obergrenze des 95 %-Konfidenzintervalls für die Ergebnisrate (HLSS.95 or HLSP.95) sollte ebenfalls entsprechend dem Zweck der Analyse definiert werden. Der Ziel- oder Mindestwert (LL^{tg}_{SS} oder LL^{tg}_{SP}) ist besonders wichtig, wenn die Auswirkungen falscher Ergebnisse kritisch sind. Beispielsweise sollte bei Blutkomponenten, die für Transfusionen verwendet werden, das Screening auf Infektionskrankheiten mit Tests durchgeführt werden, die einen $LL_{SS.95}$ von nahezu 100 % aufweisen.

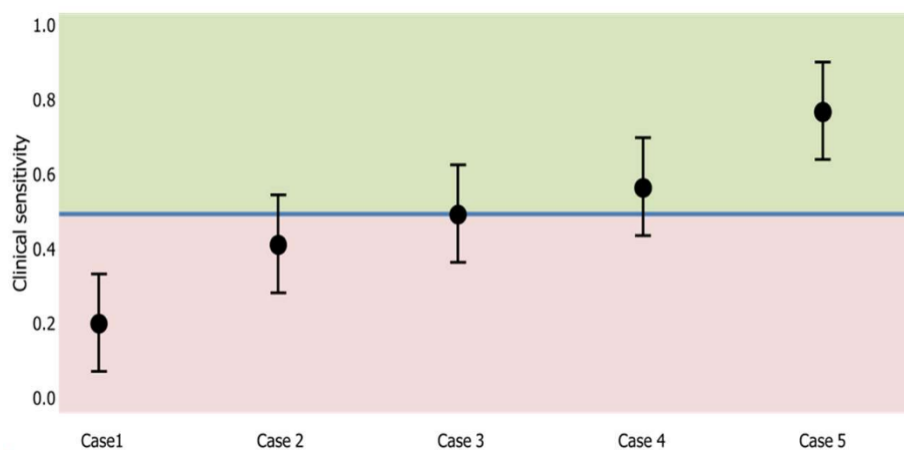


Abbildung 2. Konformitätsbewertung mit klinischer Sensitivitäts-Unsicherheit unter Verwendung der unteren 95 %-Konfidenzintervallgrenze für fünf Leistungsbewertungsszenarien.

Abbildung 2 zeigt fünf verschiedene Tests, die jeweils hinsichtlich ihrer klinischen Sensitivität und der damit verbundenen Unsicherheit bewertet wurden. Diese Fälle spiegeln unterschiedliche Grade der Einhaltung einer Zielsensitivität von 0,5 wider. Fall 1 ist nicht konform, die Fälle 2 bis 4 sind teilweise konform und Fall 5 ist konform. Aus vereinfachter statistischer Sicht sollte das Labor immer die Mindestanzahl an Proben bei der Festlegung seines Zielwertes berücksichtigen. Diese Anzahl kann durch Simulation geschätzt werden. Bei einer klinischen Sensitivität von 90 % sind beispielsweise mindestens 10 Proben für die Bewertung erforderlich. Dabei ist es wichtig, dass die Proben für eine klinische Leistungsbewertung repräsentativ für die Zielpopulation sein müssen – beispielsweise sind epidemiologisch repräsentative Proben der Typen und Subtypen eines bestimmten virologischen Erregers erforderlich.

Quellen

[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). Verfügbar unter www.eurachem.org.