

Borrelien-Serologie

Kriterien analytischer
Leistungsfähigkeit am Beispiel
kommerziell erhältlicher
Testsysteme (ELISA)

Warum dieser Vortrag?

- Serologie kann entscheidend für Diagnose und Behandlung sein.
- Bei identischen Seren unterschiedliche Ergebnisse.
- „Die Waage fängt erst bei 100 kg an zu messen (H-P Gabel)“.
- „Es können doch nicht alle Mitarbeiter seropositiv / krank sein (Hersteller)“.

Ziel des Vortrags:

- Jeder Anwender sollte sich den Unterschied zwischen testbezogenen (analytischen) und klinischen (diagnostischen) Aussagen einer Analyse vergegenwärtigen!
- Welche technischen und methodischen Maßstäbe zur Beurteilung der Testgüte gibt es?
- Wie ist nachfolgend die diagnostische / klinische Beurteilung einzuschätzen?

Aussagebereiche und Lücken

19 verschiedener Suchteste

sollen nachfolgend

dargestellt werden:

Übersicht: Analytische Kriterien

Technisch:

- Präzision und Richtigkeit
- Linearität
- Auflösungsvermögen
- Nachweisgrenze

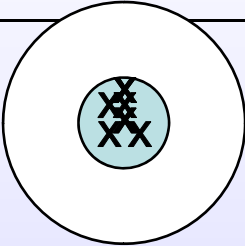
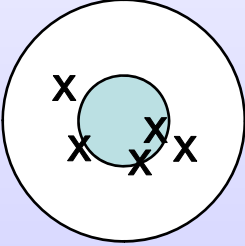
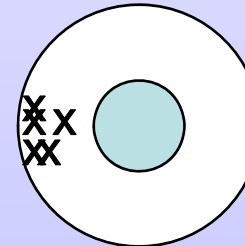
Methodisch:

- Selektivität, Interferenz, Cut
- Hierarchie der Methodenqualität

Erst danach:

Diagnostische / klinische Wertung: Sensitivität und Spezifität, Vorhersagewerte

Technische Kriterien

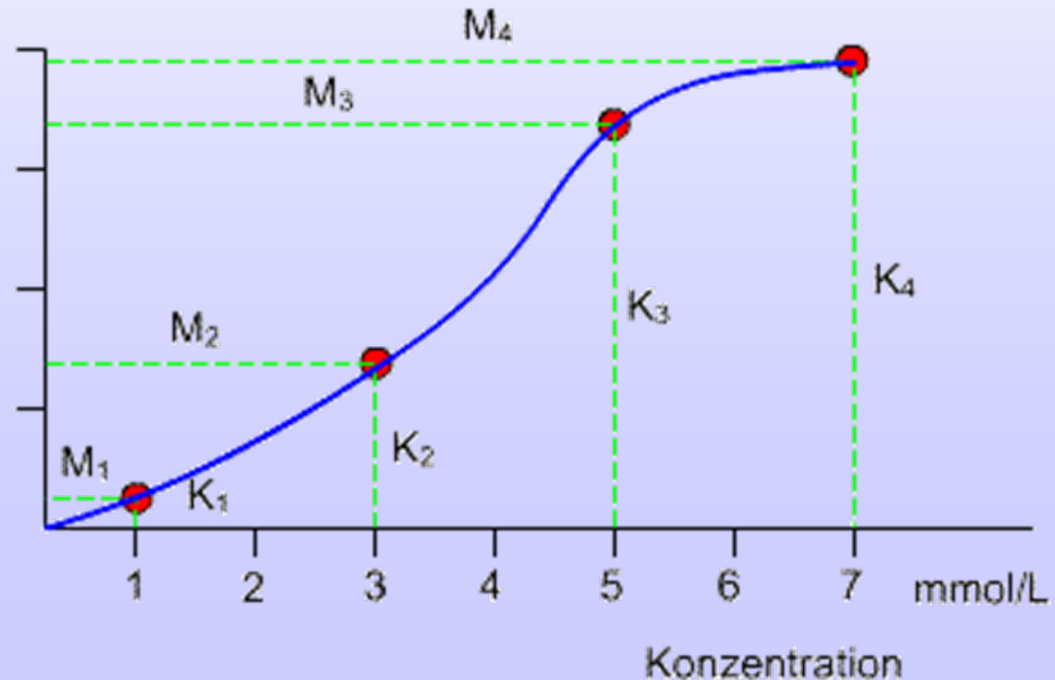
Richtigkeit	Präzision		Fehlertyp
Übereinstimmung mit dem „Zielwert“, sagt nichts über den „wahren“ Zielwert aus.	Reproduzierbarkeit		
Optimal	Optimal		-
Gut	Schlecht		Zufällig
Schlecht	Gut		Systematisch

- Angaben bei 14 von 19 Tests

Die *Linearität* erlaubt eine Schätzung von Werten.

Angaben dazu bei den Herstellern D, E, L und Q.

Mess-Signal



Die *analytische Empfindlichkeit*

Kennzeichnet die kleinste

Konzentrationsdifferenz innerhalb eines

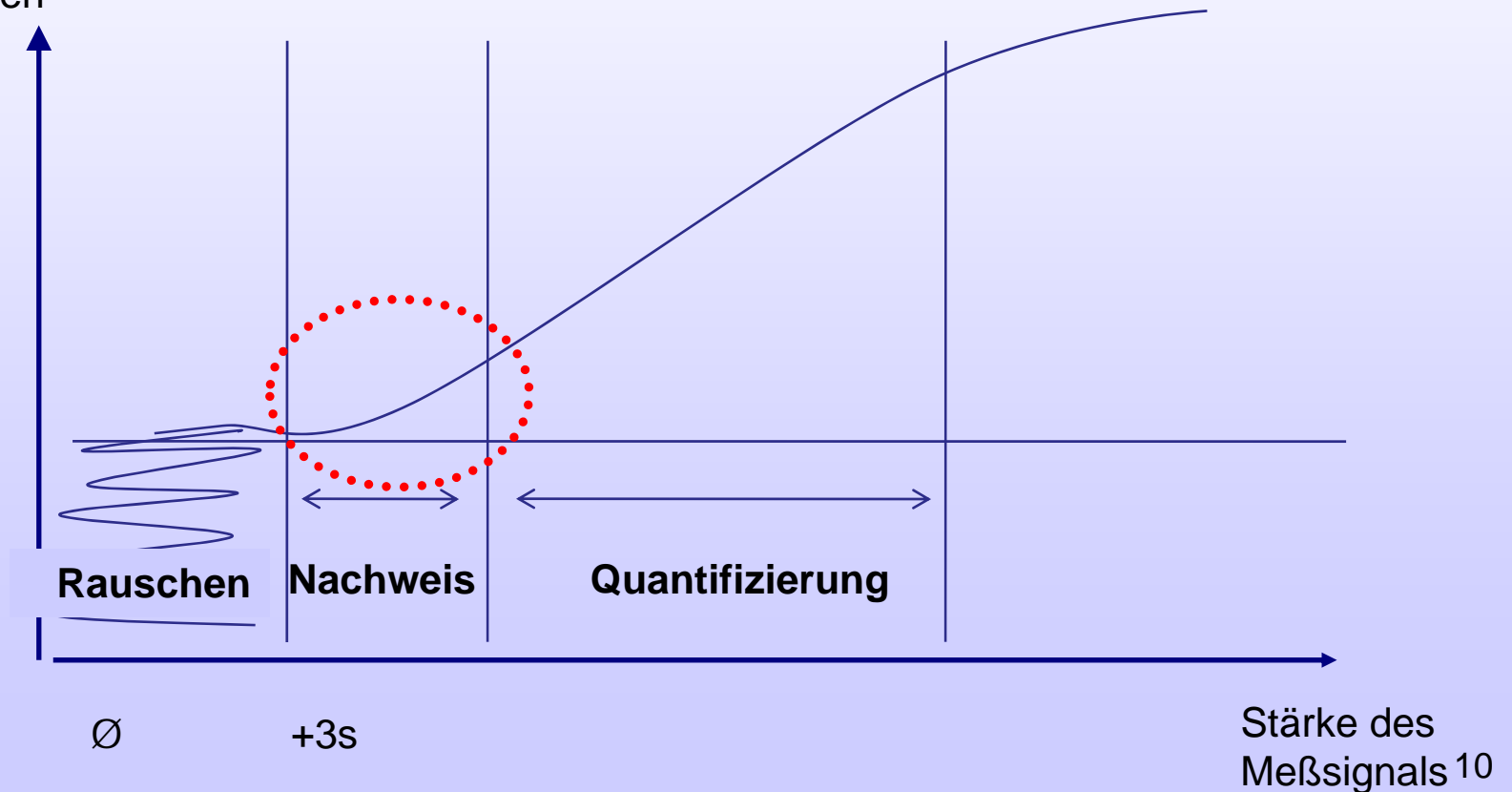
Messbereiches, die sicher unterschieden werden kann.

Keine Hersteller-Angaben

Grenzen des Meßintervalls

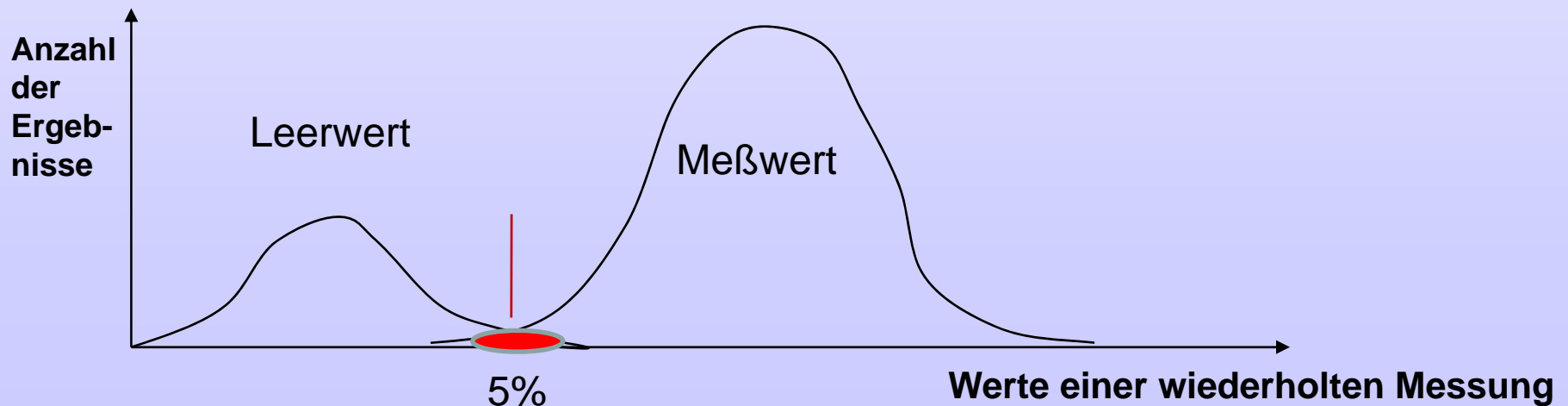
(nach Keller)

Konzentration
des Analyten



Wo kann man mit der Messung anfangen (Nachweisgrenze)?

Leerwert-Bestimmung (Matrix oder Puffer) + 3 Standardabweichungen des Leerwertes oder Toleranzgrenze für das Maß der Überlappung zwischen den Standardverteilungskurven des Leerwertes und der Standardverteilungskurve einer definierten Probe (z. B. 5%)



Herstellerangaben zur Nachweisgrenze

anhand des Substratleerwertes bei Hersteller
G, Negativ-Kontrolle plus cut-off-Faktor bei
Hersteller B.

Ansonsten werden keine Nachweisgrenzen
angegeben.

Methodische Kriterien

Die *Selektivität* (analytische Spezifität) ist die Möglichkeit, z.B. nur einen bestimmten Antikörper zu erfassen.

- Für Suchteste liegt der Schwerpunkt auf der (analytischen) Sensitivität.
- Bei rekombinanten Antigenen hoch.

- Keine Hersteller-Angaben.

Interferenz sind Wechselwirkungen, z.B. Kreuzreaktivitäten oder Matrixeffekte.

- Angaben bei 12 Herstellern.

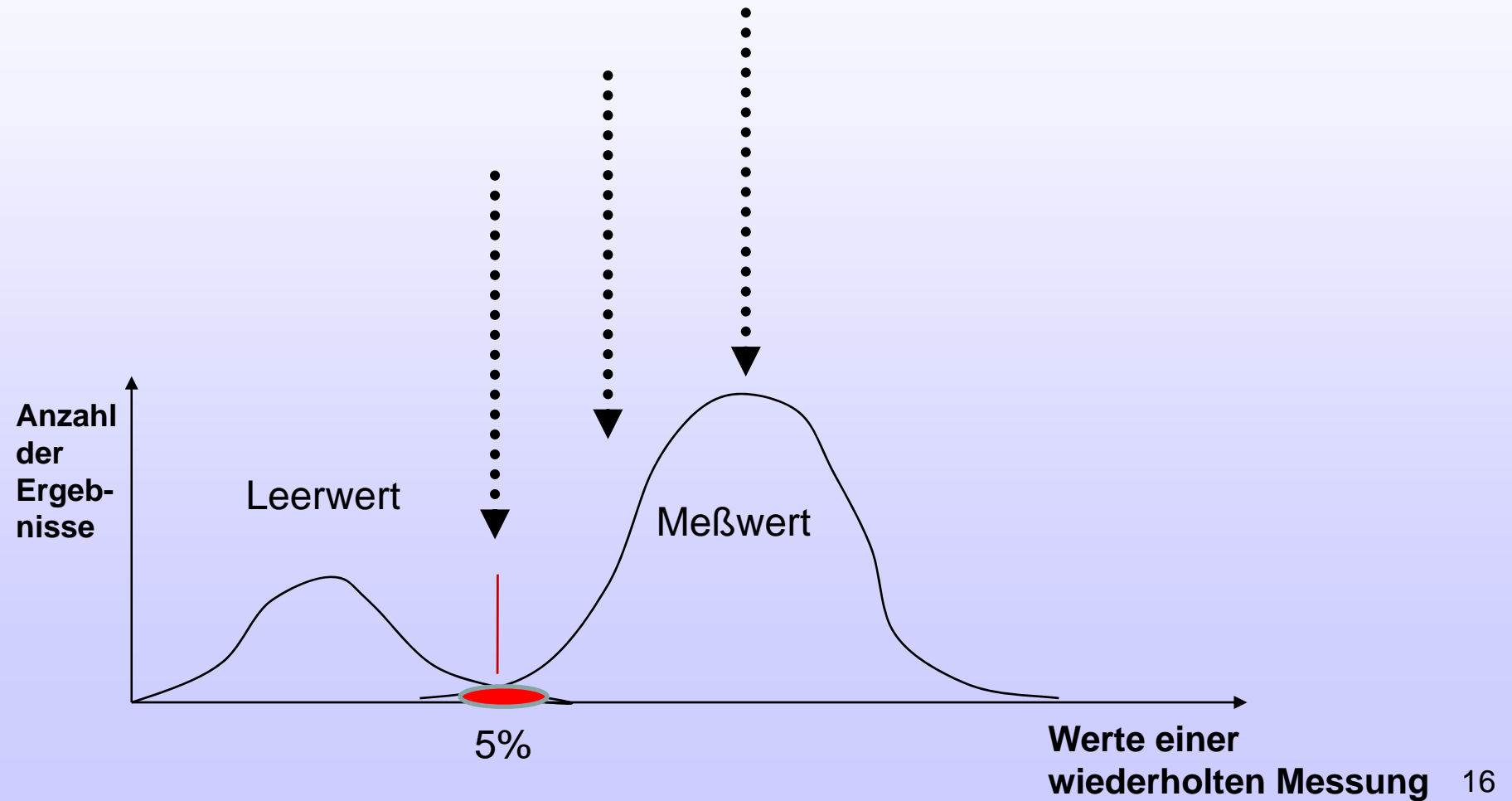
Entscheidendes analytisches Kriterium

ist der Punkt, der über die Reaktivität der Probe Auskunft gibt, der Cut.

Für die Festlegung des Cut gibt es keine festen Regeln.

Sensitivität

Spezifität / Selektivität



Werden z.B. gepoolte Blutspenderseren mit geschätzter Seroprävalenz als Referenz gewählt, dann gibt das Ergebnis der Messung nur Auskunft darüber,

ob im Probenserum mehr Antikörper vorhanden sind, als in durchschnittlichen Spenderseren.

Hersteller H ermittelt *analytische* Sensitivität und Spezifität an *klinisch definierten* Seren.

Hersteller M ermittelt den Cut mittels gepoolter Seren aus endemischen und nicht endemischen Gebieten!

Dies sind keine Angaben über analytische, sondern über diagnostische Testkriterien!

Der Hierarchie der Methodenqualität gebührt besondere Beachtung:

- I. *Definitive Methode* (Goldstandard)
- II. *Referenzmethode*
 - a. mit definitiver Methode gesichert
 - b. nicht wie a., aber **Standards** vorhanden
 - c. wie b. ohne Standards
- III. *Routinemethode*
 - a. empfohlene Methode, Störanfälligkeit definiert
 - b. empfohlene Methode, Störanfälligkeit nicht definiert.

In 16 von 19 Testen fehlen
Vergleichsmethoden oder werden nicht
genannt!

Hersteller K weist auf die fehlende Referenz-
Präparation hin.

Zusammenfassung: analytische Leistungsmerkmale

Angaben zu analytischen Leistungsmerkmalen fehlen bis auf Interferenzen, Präzision und Richtigkeit fast völlig. Dies mag Ausdruck dafür sein, dass es keine Einigkeit über den Aussagebereich der Serologie (analytisch oder diagnostisch) gibt.

Es existieren derzeit keine Referenzmethode und keine Probenstandards (Referenzpräparation)²¹!

Anwendung in der Praxis (1): Diagnostische Sensitivität und diagnostische Spezifität

Diagnostische Sensitivität:

Wahrscheinlichkeitsmaß, Kranke richtig zu erfassen.

Diagnostische Spezifität:

Wahrscheinlichkeitsmaß, Gesunde richtig auszuschließen.

Anwendung in der Praxis (2):

Positiver Vorhersagewert:

Wahrscheinlichkeit dass das Testergebnis richtig-positiv ist, ein reaktiver Wert spricht für eine Erkrankung.

Anwendung in der Praxis (3):

Negativer Vorhersagewert:

Wahrscheinlichkeit, dass das Testergebnis richtig-negativ ist; ein negativer Wert spricht gegen eine Erkrankung.

Herstellerangaben zu diagnostischen Kriterien (1):

Diagnostische Sensitivität: 6 von 19

Herstellern erproben ihre Teste an klinisch definierten Seren, 8 von 19 wählen

„Vergleichsseren“, 2 von 19 Herstellern

wählen serologische Vergleichsmethoden,

3x prozentuale Angaben, 3x keine Angaben.

Herstellerangaben zu diagnostischen Kriterien (2):

Diagnostische Spezifität: 5x „gesunde“ Probanden, 1x „nicht vorselektierte Seren“, 2x „unauffällige Seren“, 3x prozentuale Angaben, 3x „geschätzte Seroprävalenz“, 3x Blutspender, 1x serologischer Vergleich, 1x Schwangere, 4x „keine Angaben“.

Herstellerangaben zu diagnostischen Kriterien (3):

Positiver Vorhersagewert: keine
Angaben

Negativer Vorhersagewert: keine
Angaben

Derzeitige Situation zur Test-Evaluation:

Man untersucht Seren von „gesunden“ oder „kranken“ Personen und legt den Cut dorthin, wo man eine Trennung zwischen krank und gesund zu erkennen meint.

Oder man misst Blutspenderseren ein und ermittelt den Cut durch Schätzung der Seroprävalenz.

Derzeitige Situation zur Test-Evaluation (2):

Dabei wird die analytische Test-Einstellung
durch ein begrenztes „diagnostisches“
Verfahren beeinflusst!

Analytische Leistungskriterien bleiben meist
außen vor.

Schlussfolgerungen :

Man muss sich die Frage stellen, ob es vertretbar ist, einen Test vom Ergebnis her zu evaluieren:

- „Unauffälliges Serum“, Blutspender, Schwangere oder „gesunder“ Proband: das Ergebnis des Suchtestes ist „nicht-reaktiv“.

Oder muss man zuerst die analytischen Testkriterien erfüllen, bevor diagnostische Schlüsse gezogen werden können?

Die Festlegung des Cut mittels
Blutspenderseren (insbesondere der
geschätzten Seroprävalenz bei
gepoolten Proben) ist zu überdenken,
da auch „gesunde“ Blutspender Antikörper
haben können und die Menge der
Antikörper nichts über den Grad einer
möglichen Erkrankung aussagt.

Um Evidenz-basierte Aussagen über das Vorliegen von Borrelia-Antikörpern machen zu können ist eine definierte Nachweisgrenze nötig und Offenlegung der „Vergleichsmethoden“ einschließlich ihrer Einschränkungen. Einen solchen Test gibt es derzeit nicht.

Danach kann man mit Hilfe ausführlicher
Datenerhebungen und insbesondere im
Vergleich mit
Kultur- oder PCR- positiven
Patientenproben Angaben über klinische
Aspekte,
diagnostische Sensitivität, diagnostische
Spezifität und Vorhersagewerte machen.

Konsequenz für statistische Erhebungen:

- Für die Ermittlung der Seroprävalenz (Antikörperstatus: keine Krankheitsbeurteilung) werden zuverlässige Angaben zur Nachweisgrenze, zum Cut und zur analytischen Sensitivität und Selektivität benötigt.
- Aufmerksamkeit ist bei der Verwendung von Testsystemen geboten, deren Schwerpunkt auf der diagnostischen Spezifität liegt.

Konsequenz für mögliche Standardisierung:

- Die DIN 58969-44 [Spezielle Anforderungen für den Nachweis von Antikörpern gegen *Borrelia burgdorferi*] als Leit-Norm? Sie beschränkt den Aussagebereich des „Borrelien-Blots“ auf das Vorliegen von Antikörpern.
- Analytische Kriterien sollten möglichst vollzählig erfüllt sein.
- Güte des Referenztestes sollte genau bestimmt sein.

Konsequenz für die Befundung:

Die Angabe: „Antikörper vorhanden oder nicht vorhanden“ benötigt die Kenntnis der Testevaluation.

Prüfen ob es möglich ist, den Befundtext auf die nachvollziehbaren Parameter zu beschränken.

Diagnostische Schlussfolgerungen sollten weggelassen werden.

Konsequenz für die Behandlung:

Der serologische Befund gibt keinen
Aufschluss über die Behandlungs-
bedürftigkeit.

Welchen Wert hat ein serologischer Befund?

Hinweis auf frisches oder älteres Infektions-
geschehen.

Bestätigung eines Verdachtes?

Was tun, wenn Testgütekriterien nicht vollständig ermittelt werden können?

- Sprachliche Konkretisierung kann helfen.
- „Mut zur Lücke!“
- Sagen oder schreiben, wie weit man aufgrund eines Testergebnisses Aussagen treffen kann zu:
 - Antikörperstatus und
 - Krankheitsgeschehen.

„Dass gesamthaft gesehen nur so wenige Angaben über Sensitivität und Spezifität (...) von Tests vorliegen, erstaunt daher nicht mehr. Der Aufwand gemessen am Nutzen ist enorm.

Trotzdem sollte, gerade im Zeitalter der "Evidence-based Medicine", versucht werden, so viele diesbezügliche Daten zusammenzutragen wie möglich.
(<http://www.biorama.ch>)

Zusammenfassung

- Unterschied zwischen analytischen und diagnostischen Testkriterien.
- Der Cut entscheidet über das Ergebnis.
- Auswirkungen „diagnostisch-spezifischer“ Suchtests bedenken.

Uta Everth

uta@everth.de

www.everth.de

www.bzk-online.de



Literatur:

- Herbert Keller, Klinisch-chemische Labordiagnostik für die Praxis, 2. Auflage 1991
- Lothar Thomas, Labor und Diagnose, 6. Auflage