

Wie beurteile ich die Qualität diagnostischer Tests?

Arnold von Eckardstein,
Institut für Klinische Chemie, Unispital Zürich

Kriterien für die Beurteilung der klinischen Tauglichkeit von Biomarkern

- 1) Kann der Biomarker gemessen werden?**
 - a) Richtige und präzise Analyseverfahren*
 - b) bekannte und beherrschbare Präanalytik*
 - c) zugängliche Assays*
 - d) Hochdurchsatzfähig und kurze turn-around Zeit*
 - e) relativ kostengünstig*
- 2) Bringt der Biomarker neue Information?**
 - a) Starke und konsistente diagnostische oder prognostische Assoziation zwischen dem Biomarker und der Krankheit in multiplen Studien*
 - b) Information ergänzt oder verbessert Information bisheriger Tests*
 - c) Entscheidungsgrenzen in >1 Studie evaluiert*
 - d) Evaluation auch in zufälligen Populationen*
- 3) Hilft der Biomarker beim Management des Patienten?**
 - a) Überlegenheit gegenüber vorhandenen Tests, und/oder*
 - b) Assoziierte Krankheit bzw. Risiko ist therapeutisch beeinflussbar, und/oder*
 - c) Orientierung am Biomarker verbessert Patientenversorgung*

Kriterien für die Beurteilung der klinischen Tauglichkeit von Biomarkern

1) Kann der Biomarker gemessen werden?

- a) *Richtige und präzise Analyseverfahren*
- b) *bekannte und beherrschbare Präanalytik*
- c) *zugängliche Assays*
- d) *Hochdurchsatzfähig und kurze turn-around Zeit*
- e) *relativ kostengünstig*

Heutige Vorlesung
Kurs Klinische Chemie
beides

2) Bringt der Biomarker neue Information?

- a) *Starke und konsistente diagnostische oder prognostische Assoziation zwischen dem Biomarker und der Krankheit in multiplen Studien*
- b) *Information ergänzt oder verbessert Information bisheriger Tests*
- c) *Entscheidungsgrenzen in >1 Studie evaluiert*
- d) *Evaluation auch in zufälligen Populationen*

3) Hilft der Biomarker beim Management des Patienten?

- a) *Überlegenheit gegenüber vorhandenen Tests, und/oder*
- b) *Assoziierte Krankheit bzw. Risiko ist therapeutisch beeinflussbar, und/oder*
- c) *Orientierung am Biomarker verbessert Patientenversorgung*

Wie beurteile ich die Qualität diagnostischer Tests?

Lehr/Lernziele

Kennen der Beurteilungskriterien / Kenngrößen diagnostischer Tests und ihrer klinischen Implikationen:

➤ **analytische Qualität:**

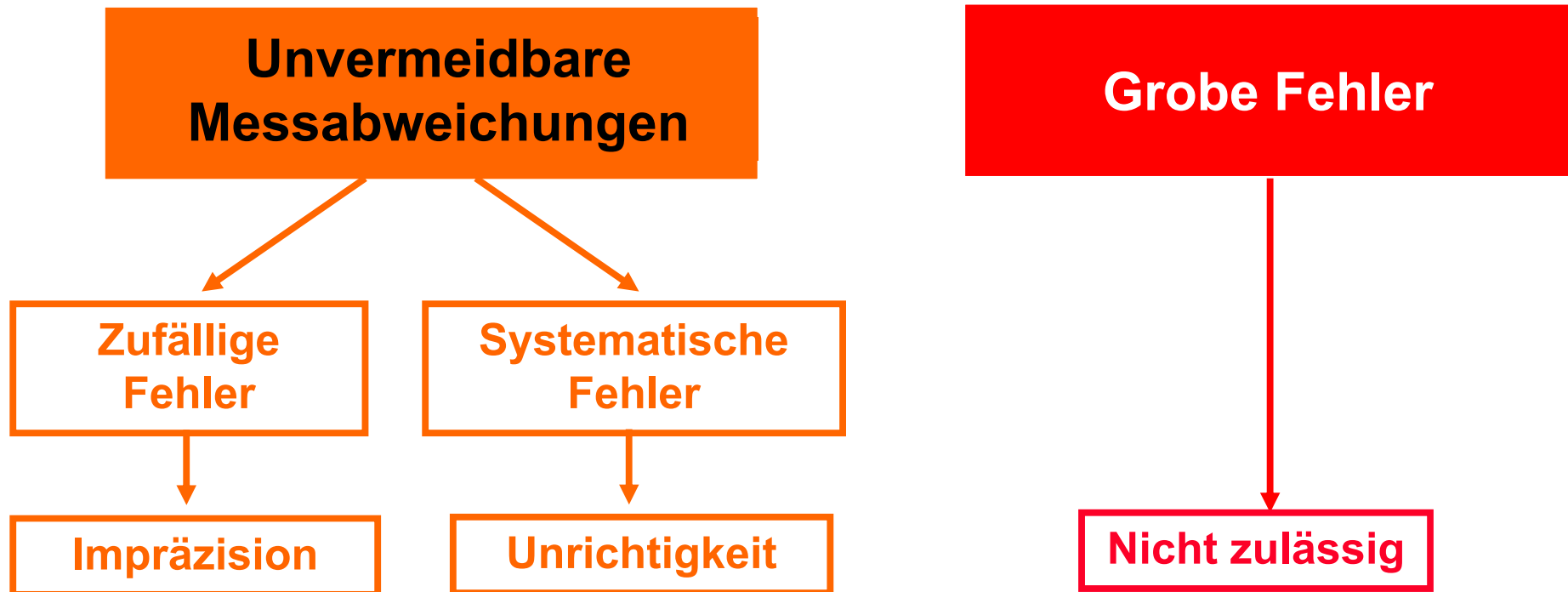
- Präzision und ihre Implikationen für Messgrenzen, minimale und kritische Differenzen
- Richtigkeit und ihre Implikationen für die Methodenabhängigkeit von Messergebnissen und Grenzwerten

➤ **diagnostische Qualität:**

- Spezifität/Sensitivität, receiver operator characteristic (ROC) Kurven
- Positive und negative Likelihood-Ratios
- Positive und negative prädiktive Werte und ihre Abhängigkeiten von Prävalenz/Prä-Testwahrscheinlichkeit einer Krankheit

Fehlerarten im Labor

Laboranalysen sind anfällig für



Verschiedene Arten von Messabweichungen (Fehler)

● Zufällige Messabweichungen

- Resultate abwechselnd leicht zu hoch und leicht zu tief
- Summe aller Fehler, die bei einer Analyse unvermeidlich auftreten
- daher: nicht zu vermeiden, aber möglichst klein zu halten
- **Mass: Unpräzision („Präzision“)**

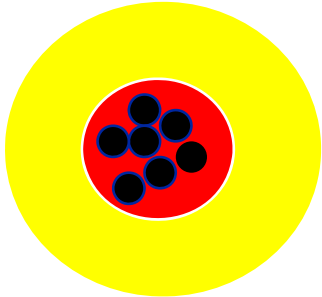
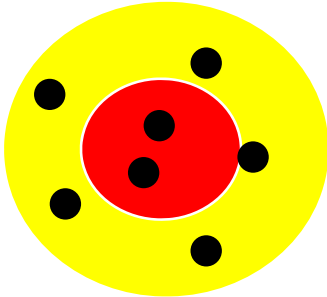
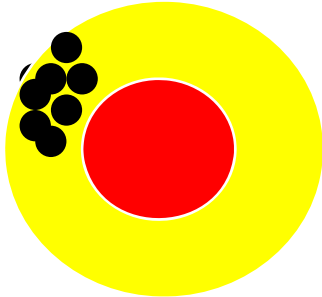
● Systematische Messabweichungen

- Resultate immer zu hoch oder zu tief: „bias“
- Fehler bei Kalibration, verfallene Reagenzien, falsch eingestellte Pipetten
- **Mass: Unrichtigkeit („Richtigkeit“)**

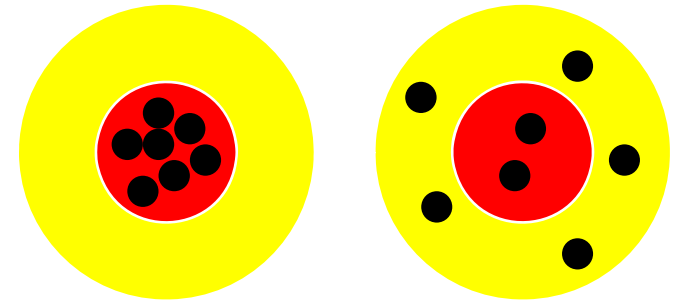
● Grobe Fehler

- häufig menschlichen Ursprungs: Verwechslungen von Proben, Abschreibe- und Umrechnungsfehler, Berechnungsfehler bei Verdünnungen

Fehlerarten

Fehlertyp	kein Fehler	zufälliger Fehler	systematischer Fehler	grober Fehler
				falsche Scheibe, falsche Methode
Präzision	optimal	schlecht	gut	
Richtigkeit	optimal	gut	schlecht	

Statistische Grundlagen für die Beurteilung der Präzision



1. Verteilung der Werte (zufällige Messabweichungen):
Normalverteilung (Gauss)

2. Mittelwert (arithmetisches Mittel): $\bar{x} = \sum(X_i) / N$

3. Streuung: Standardabweichung s

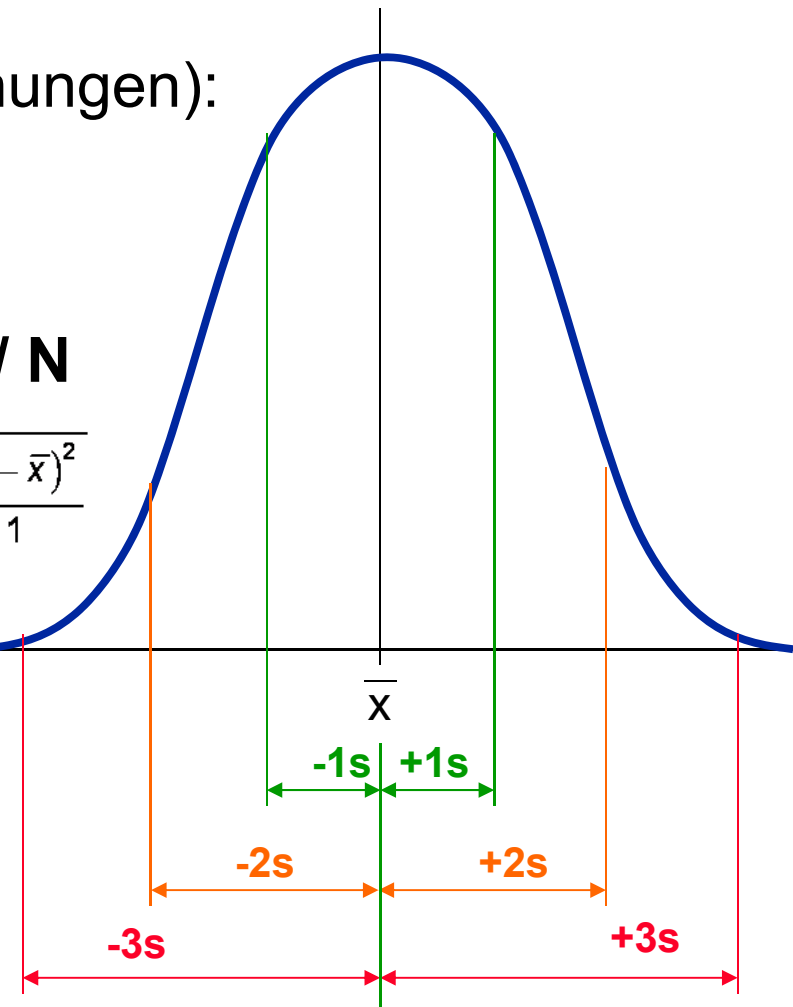
$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n - 1}}$$

$\bar{x} \pm s = 68.3\%$

$\bar{x} \pm 2s = 95.5\%$

$\bar{x} \pm 3s = 99.7\%$

4. Variationskoeffizient: $VK = s / \bar{x}$

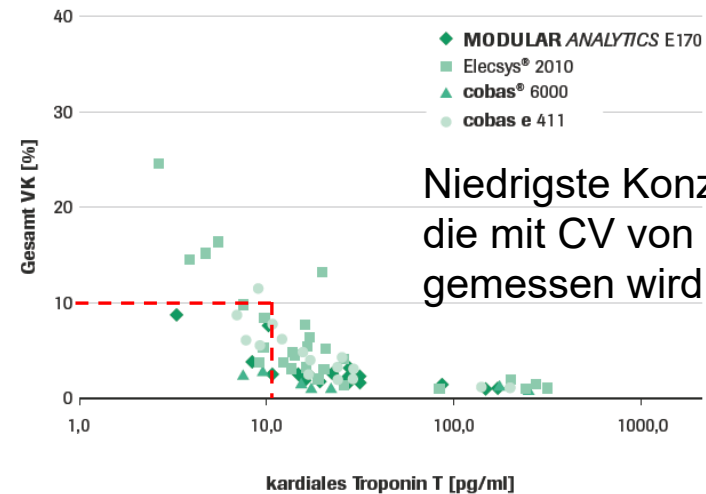


Kategorisierung von kardialen Troponin -Assays nach ihrer Impräzision (Funktionelle Assay-Sensitivität)

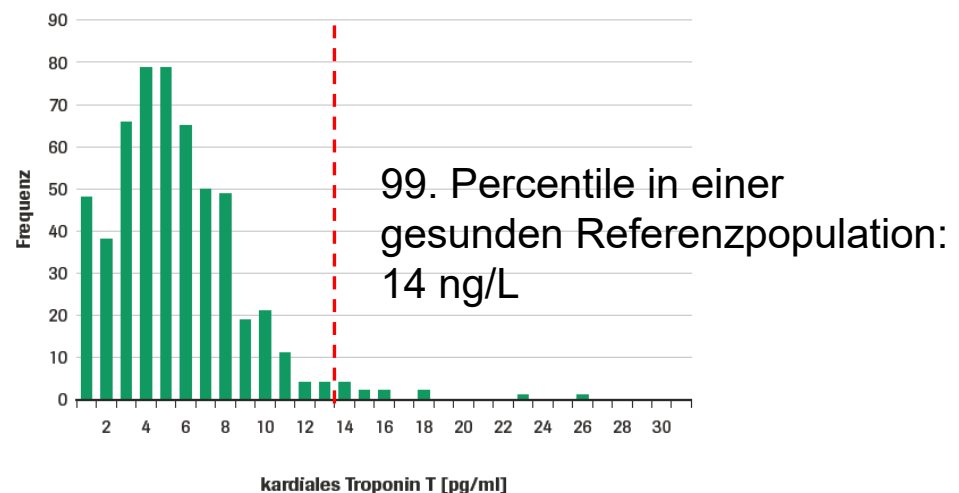
Beispiel:

Table 1. Scorecard designations of cTn assays.	
Acceptance designation	Total imprecision at the 99th percentile, CV%
Guideline acceptable	≤ 10
Clinically usable	>10 to ≤ 20
Not acceptable	>20
Assay designation	Measurable normal values below the 99th percentile, %
Level 4 (third generation, hs)	≥ 95
Level 3 (second generation, hs)	75 to <95
Level 2 (first generation, hs)	50 to <75
Level 1 (contemporary)	<50

1 Präzisionsprofil Elecsys® Troponin T hs

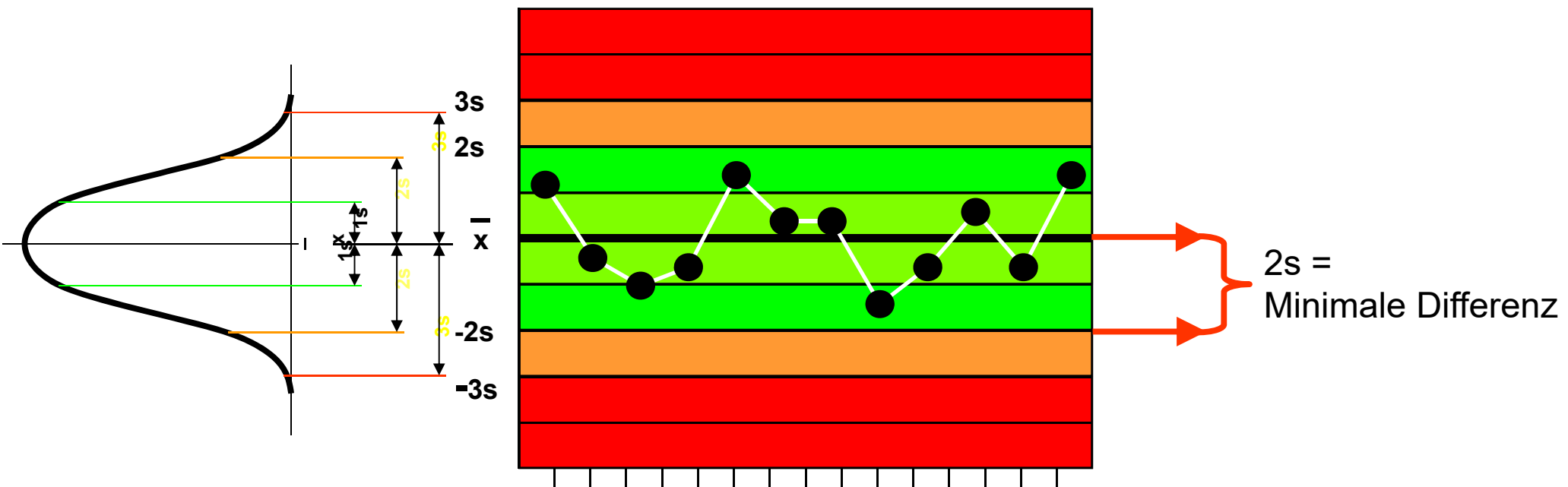


2 Referenzbereich Troponin T mit Elecsys® Troponin T hs



Über-/Unterschreiten eines Cut-offs

- Mittelwert einer Mehrfachmessung $>$ oder $<$ Cutoff
- Für eine Einfachmessung gilt: Bei $p=0.05$ (95% Sicherheit) beträgt die **minimale Differenz** $\pm 2s$, um sicher über bzw. unter dem Cutoff zu liegen.



Frage 1: Sie bestimmen die Glukosekonzentration im Nüchternplasma bei Ihrem Patienten mit einem Glukometer. Dessen Impräzision beträgt 5 %.

Ab welcher Konzentration können Sie ausreichend sicher sein, dass die wahre Glukosekonzentration >7,0 mmol/l beträgt und damit die Definition eines Diabetes erfüllt ist? (Einfachauswahl)

- a) 7,1 mmol/l
- b) 7,2 mmol/l
- c) 7,4 mmol/l
- d) 7,6 mmol/l
- e) 7,8 mmol/l

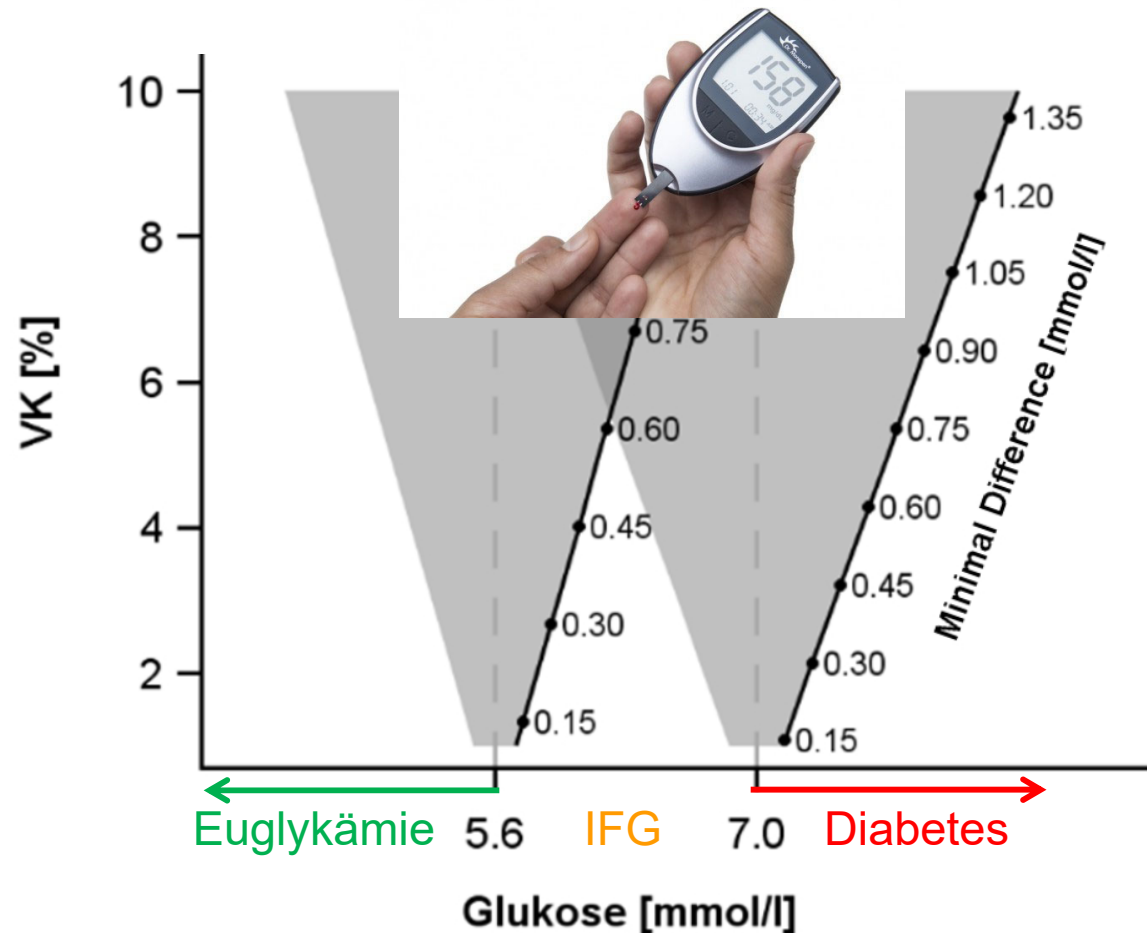
Abhängigkeit der diagnostischen Zuverlässigkeit von der Impräzision: Beispiel der Diabetes-Diagnostik

Der kleinste Abstand zwischen einem Messwert und einem Grenzwert, bei denen sie als verschieden bezeichnet werden können, berechnet sich aus der Standardabweichung (SD).

$$MD = k \times \sqrt{SD^2} = 2 \times SD$$

MD= Minimale Differenz

$k=2$ entspricht einem Konfidenzintervall von 95%



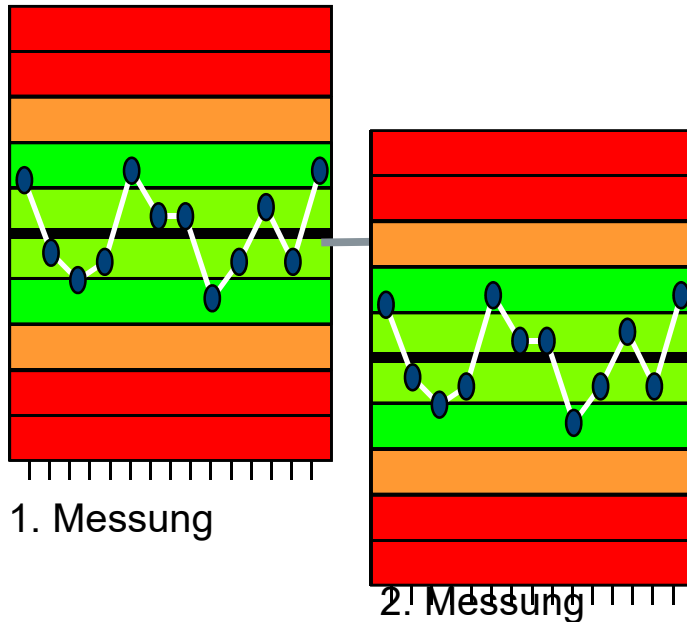
Die Ärzte eines Spitals verordnen Ihrem Patienten nach einem Myokardinfarkt und wegen eines LDL-Cholesterin-Spiegels von 4.0 mmol/l ein hochwirksames Statin und empfehlen eine Kontrolluntersuchung beim Hausarzt. Sie messen acht Wochen später den LDL-Cholesterin-Spiegel mit 3.2 mmol/l. Die Impräzision der von Ihnen und im Spitallabor verwendeten Tests für die direkte LDL- Cholesterin Bestimmung oder für die Bestimmung des HDL-Cholesterins (für die Berechnung des LDL-Cholesterins nach Friedewald) beträgt 5%.

Wie ist dieser geringer als erwartete Abfall des LDL-Cholesterins (-30 bis -50 %) erklärbar?

- a) Gute Adhärenz des Patienten (er nimmt das Statin ein)
- b) Mangelnde Adhärenz des Patienten (er nimmt das Statin nicht ein)
- c) Postaggressionsstoffwechsel beim akuten Herzinfarkt
- d) Impräzision der Messmethoden
- e) Unrichtigkeit der Methoden zur LDL- Cholesterinbestimmung



Verlaufsbeurteilung nach kritischer Differenz



$$D_{\text{crit}} = 2 * \sqrt{2} \cdot SD \approx 2.8 * SD$$
$$= D_{\text{crit}} \approx 2.8 * x_{\text{max}} * \text{VK\%/100}$$

$$D_{\text{pat}} \leq D_{\text{crit}}$$



Ab- oder Zunahme kann durch analytische Streuung allein erklärt werden

$$D_{\text{pat}} > D_{\text{crit}}$$



Ab- oder Zunahme wahrscheinlich „echt“

Verlaufsbeurteilung: Beispiel Kontrolle einer Cholesterin-senkenden Therapie

52-jähriger Mann nach Herzinfarkt: Kontrolle nach 2 Monaten

		Vorwert	aktuell
LDL Cholesterin	mmol/L	4.0	3.2

- Frage: hat die Statin Therapie wirklich eine Wirkung gezeigt oder ist die geringe Abnahme von Cholesterin durch die analytische Streuung von Tag zu Tag allein schon erklärbar ?



Verlaufsbeurteilung: Beispiel Kontrolle einer Cholesterin-senkenden Therapie

52-jähriger Mann nach Herzinfarkt: Kontrolle nach 1 Monat

		Vorwert	aktuell
LDL Cholesterin	mmol/L	4.0	3.2

- VK LDL-Cholesterin: 5%
- $D_{\text{crit}} = 2.8 \cdot 4.0 \text{ mmol/L} \cdot 5\%/100 = 0.56 \text{ mmol/L}$
- $D_{\text{pat}} = 0.8 \text{ mmol/L}$
- $D_{\text{pat}} > D_{\text{crit}}$ → kann durch analytische Streuung (Impräzision) nicht erklärt werden



Wie ermittelt man Kontroll-/ Warngrenzen?

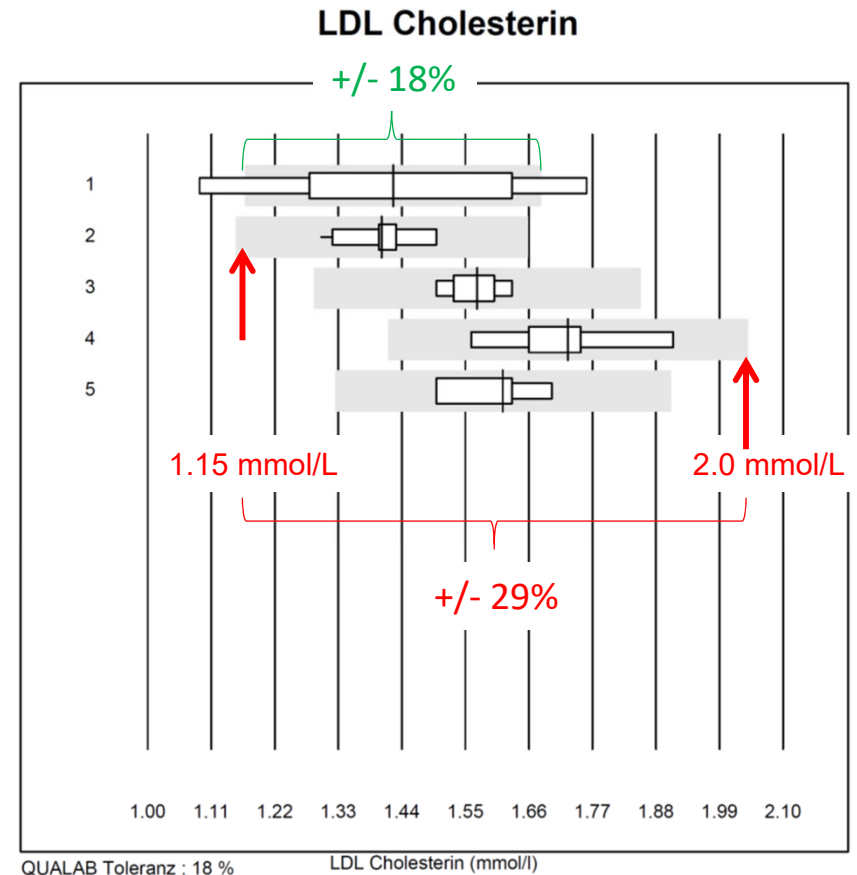


1. **Aus gesetzlichen Vorgaben (Schweiz: QUALAB)**
-> legen maximale Toleranzen für jeden Parameter fest
2. **Aus dem Kontrollbereich des Kontrollmaterialherstellers**
-> (idealerweise deklariert als 2s oder 3s-Intervalle)
3. **Aus Mehrfachmessungen der Qualitäts-Kontrolle (Vorperiode)**
-> zeitaufwendig
-> teuer

Qualität von LDL Cholesterin im Ringversuch in der Schweiz (MQ 2/2023)

erlaubte Zielbereiche
pro Methode (3*VK): +/- 18%
Also erlaubter VK: +/- 6%

De facto:
über alle Methoden: +/- 29%
Also VK: +/- 9.6 %



Nr. Methode	Total	% OK	% ungen.	% Ausr	Zielwert	VK%	Typ
1 Selectra	6	50.0	33.3	16.7	1.4	18.7	e*
2 nasschemisch	15	100.0	0.0	0.0	1.4	4.0	e
3 Roche, Cobas	15	100.0	0.0	0.0	1.6	2.8	e
4 Autolyser/DiaSys	12	75.0	0.0	25.0	1.7	6.6	e
5 Beckman	4	100.0	0.0	0.0	1.6	5.2	e*

Ein Resultat wurde abgegeben, aber nicht publiziert, da die Methodengruppe zu klein war.

Wie beurteile ich die Qualität diagnostischer Tests?

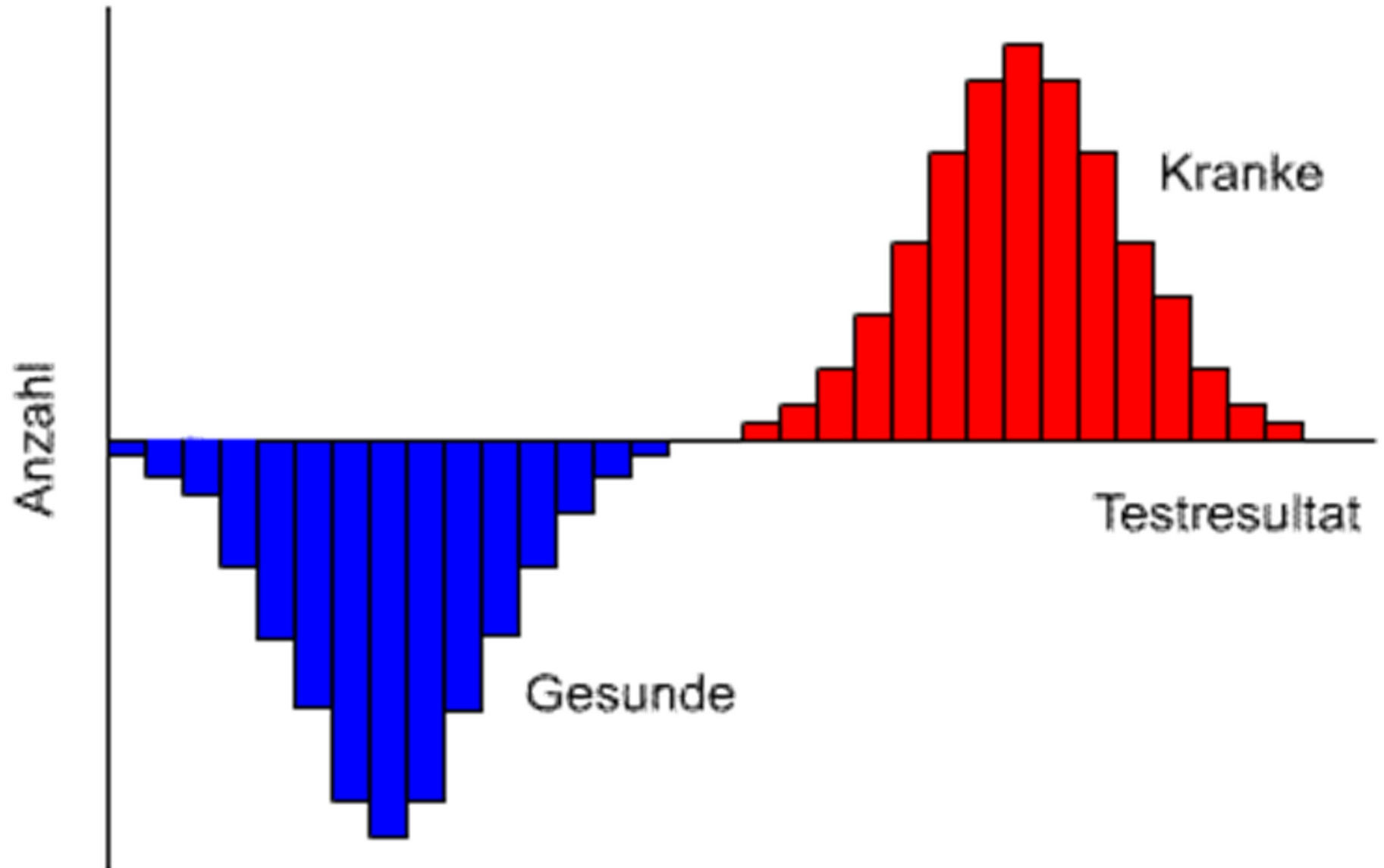
Lehr/Lernziele



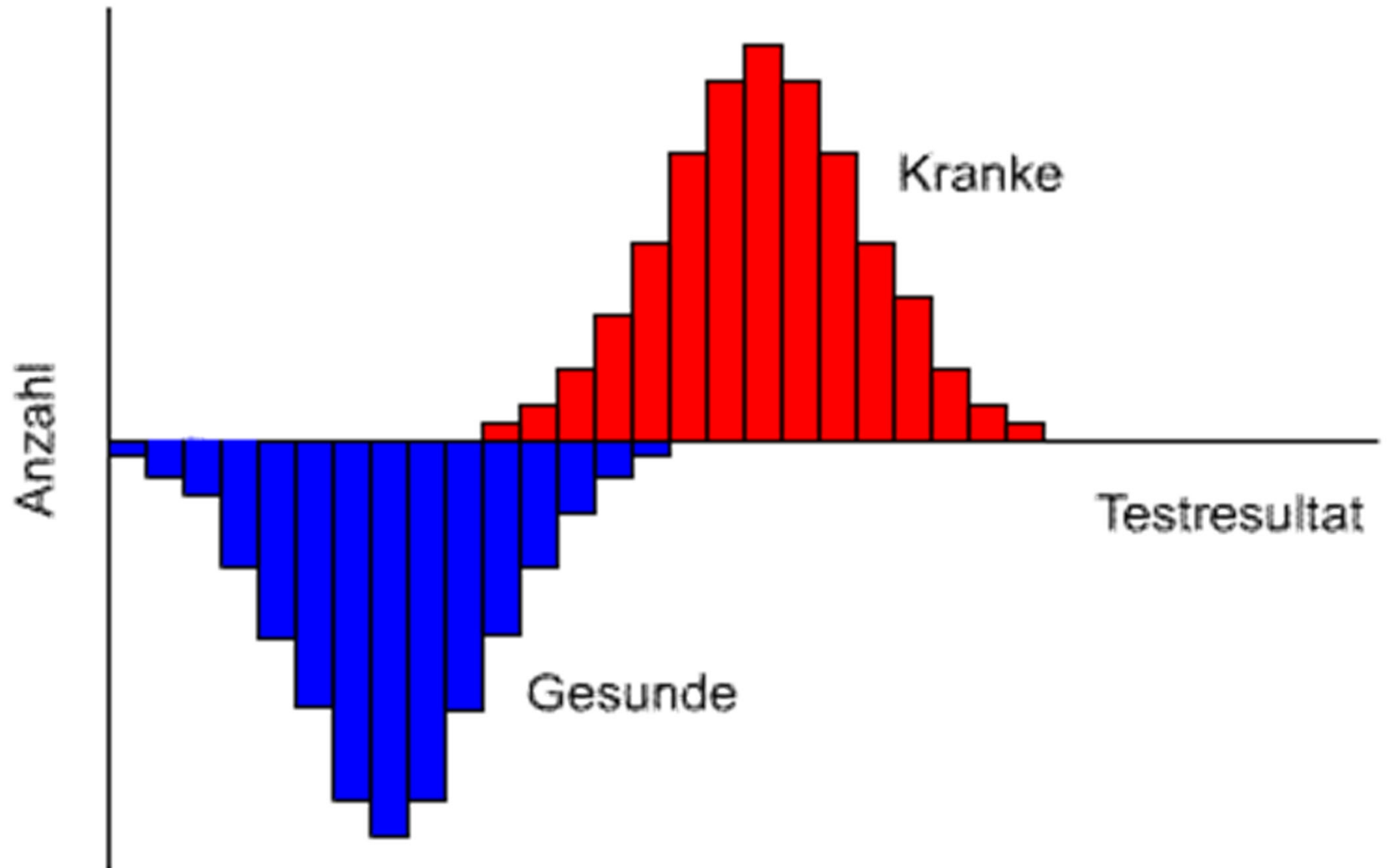
➤ **diagnostische Qualität:**

- Spezifität/Sensitivität, receiver operator characteristic (ROC) Kurven
- Positive und negative Likelihood-Ratios
- Positive und negative prädiktive Werte
und ihre Abhängigkeiten von Prävalenz/Prä-Testwahrscheinlichkeit einer Krankheit

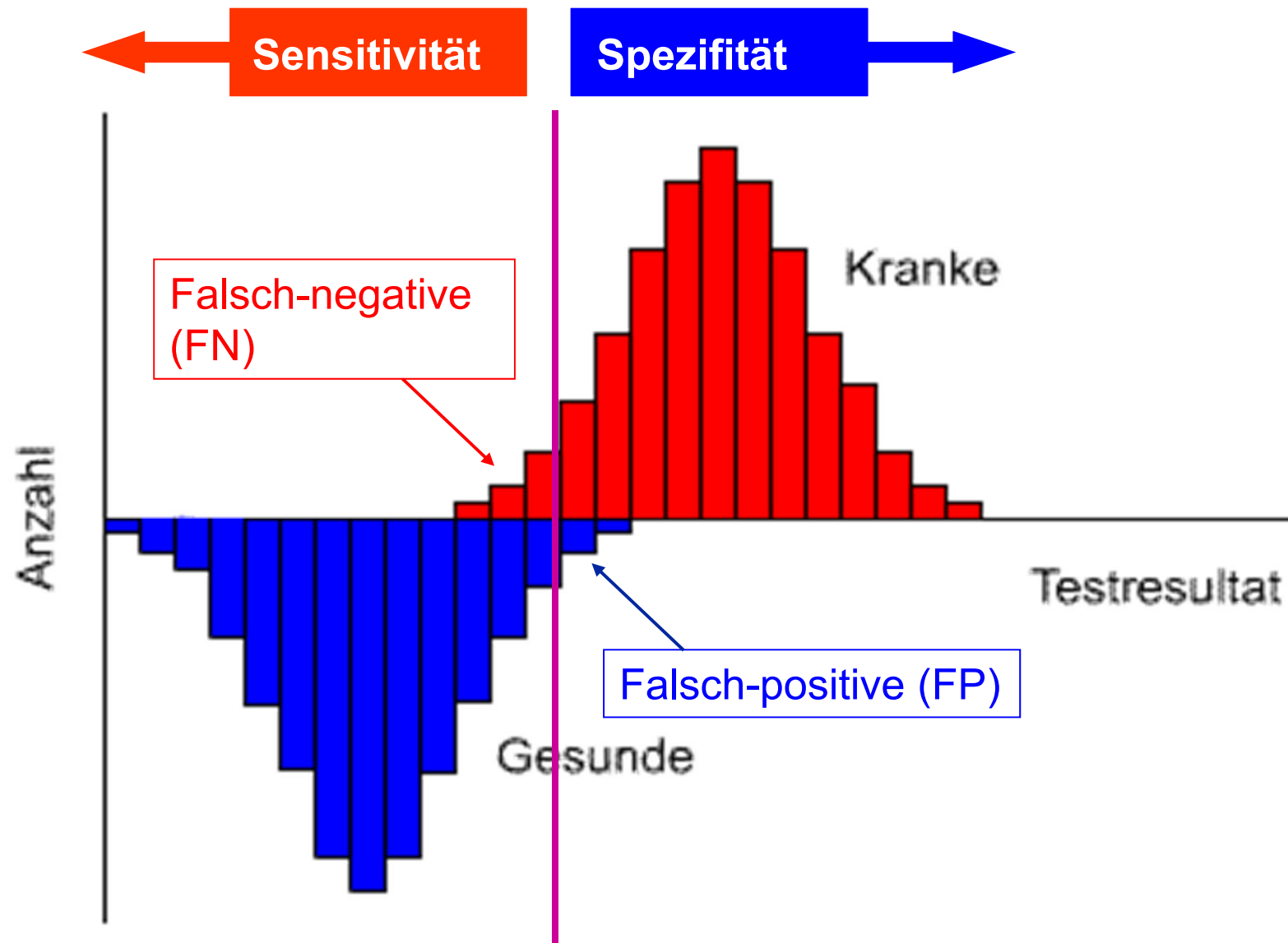
Der ideale Test



Der nicht-ideale Test



Sensitivität und Spezifität



Grundbegriffe der Beurteilung diagnostischer Tests:

Sensitivität

Diagnostische Sensitivität

Definition: Anzahl wahrer positiver (=pathologischer) Ergebnisse bei Patienten, bei denen die Krankheit mit Sicherheit besteht.

= Wahrscheinlichkeitsmass, Kranke richtig zu erfassen

$$\text{Sensitivität} = \frac{\text{Zahl der echt positiven}}{\text{Zahl der Kranken}} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

TP = true positive

FN = false negative

Grundbegriffe der Beurteilung diagnostischer Tests:

Spezifität

Diagnostische Spezifität

Definition: Anzahl wahrer negativer (=„normaler“) Ergebnisse bei Gesunden, bzw. Fehlen der bestimmten Krankheit

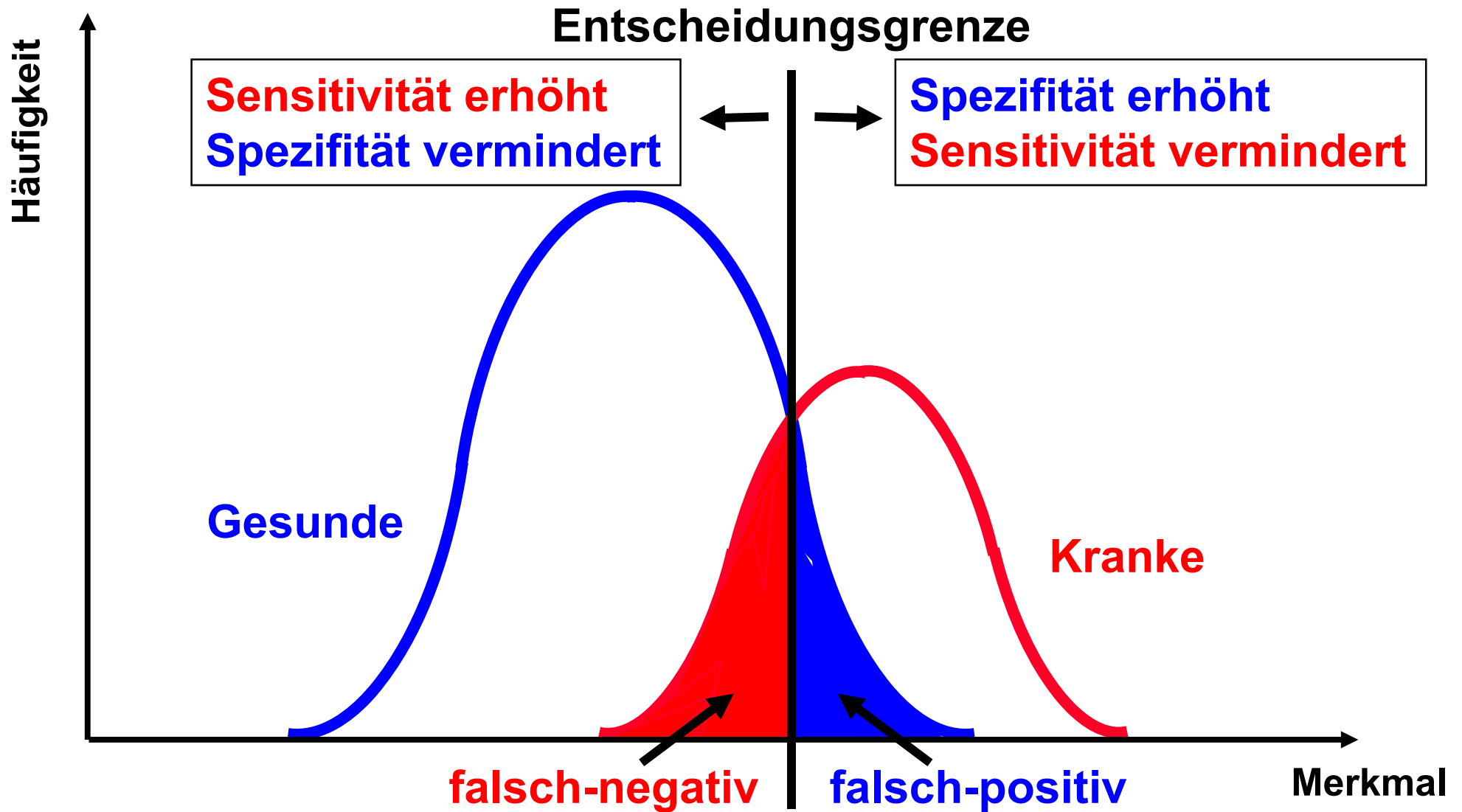
= Wahrscheinlichkeitsmass Gesunde richtig zu erfassen

$$\text{Spezifität} = \frac{\text{Zahl der echt negativen}}{\text{Zahl der Gesunden}} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

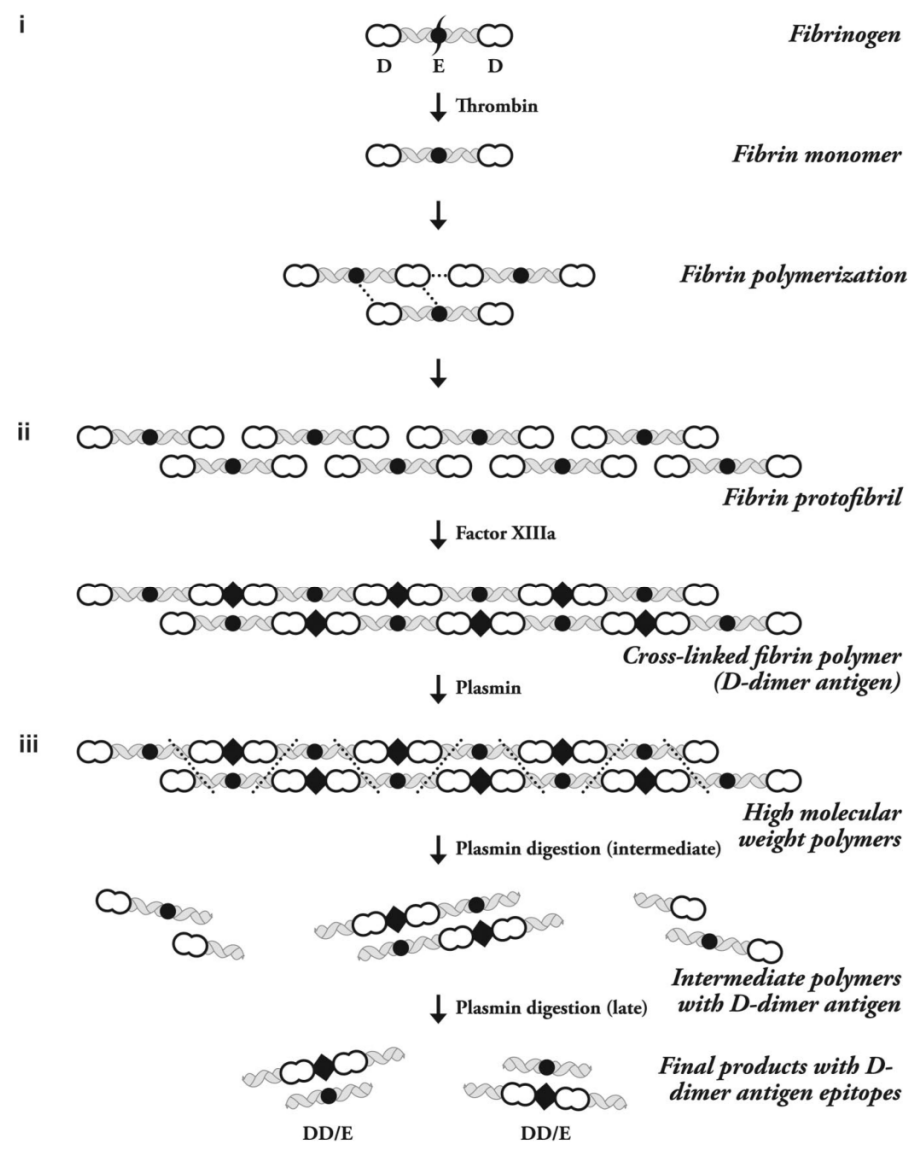
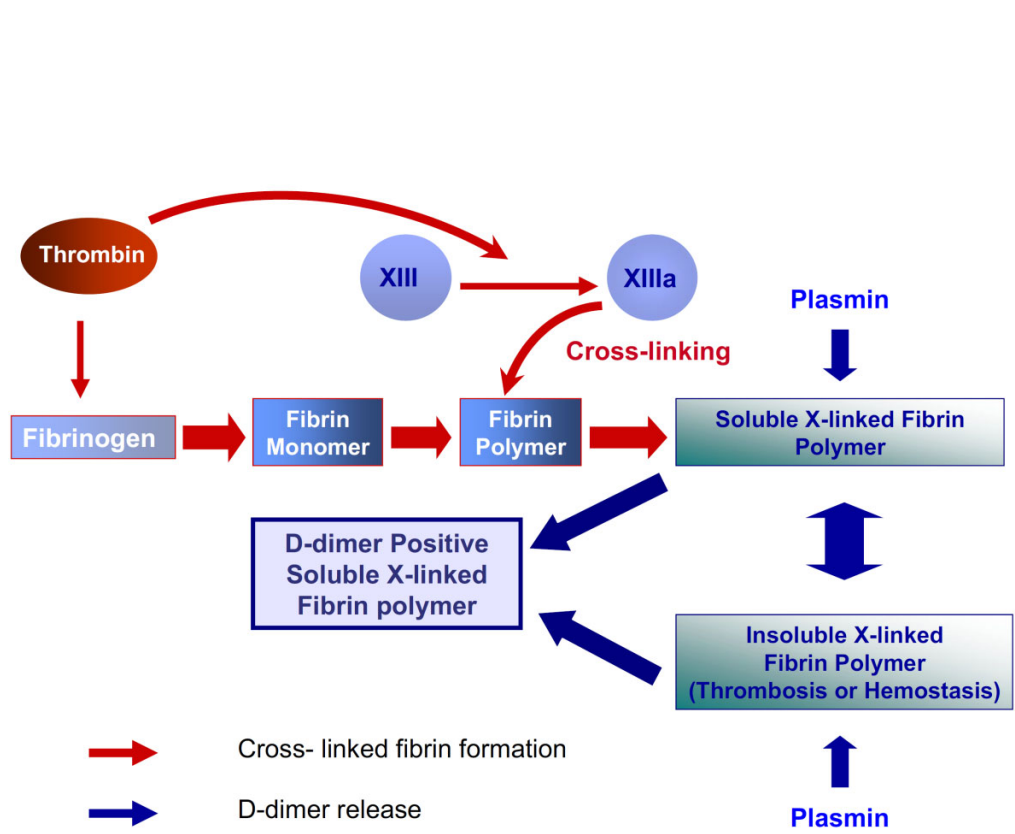
TN = true negative

FP = false positive

Sensitivität und Spezifität



Fibrinpolymerisierung und D-Dimer- Entstehung



Ein Beispiel:

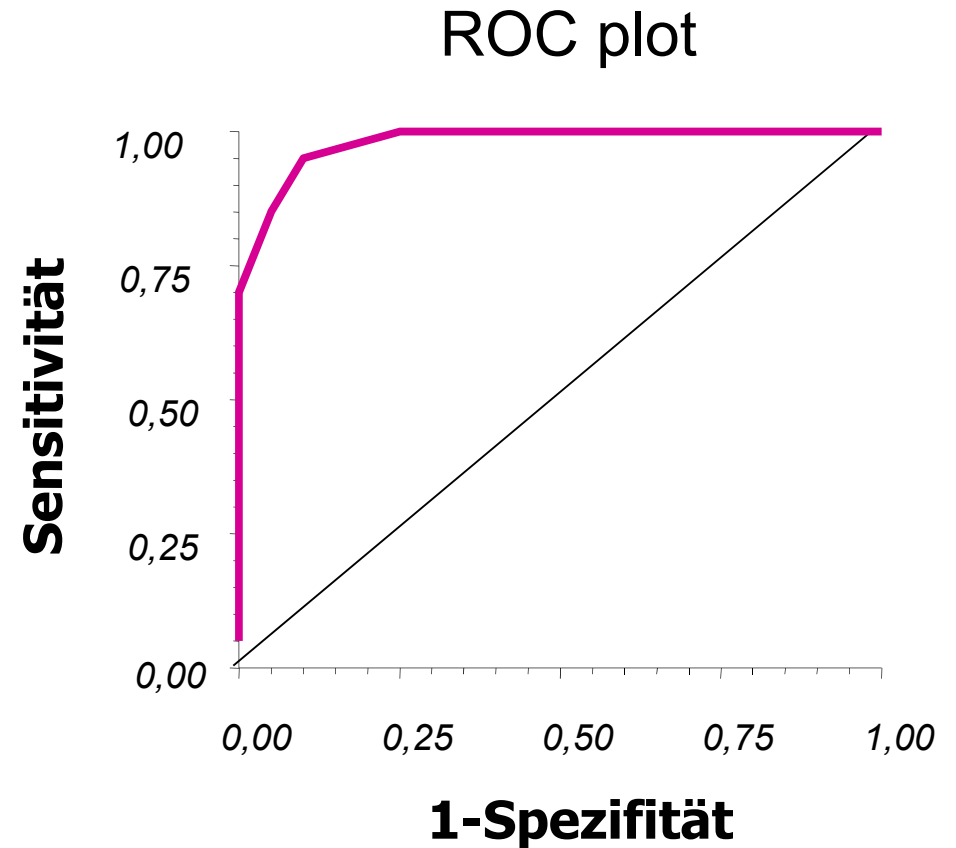
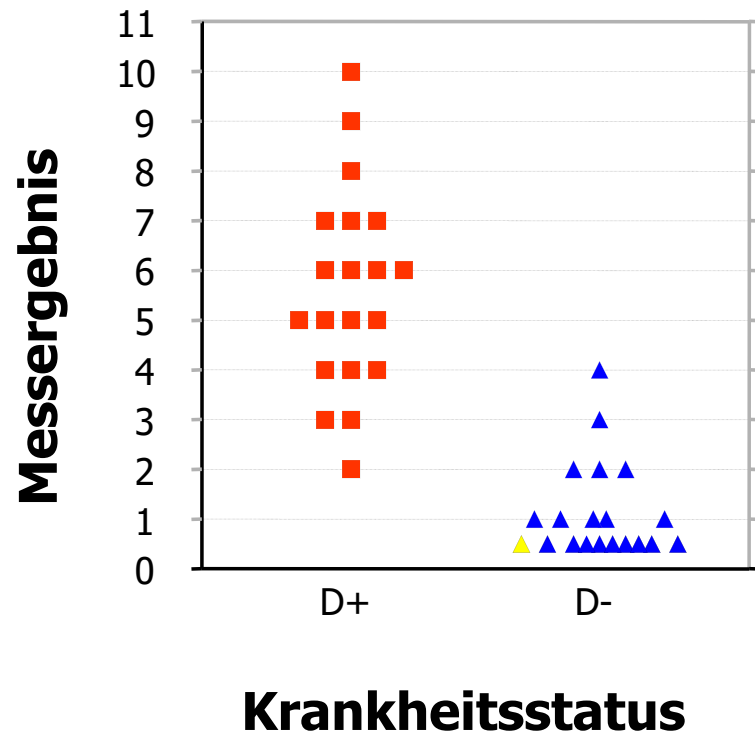
D-Dimer bei venöser Thrombose und Embolie

118 Patienten mit Verdacht auf pulmonale Embolie oder tiefer Beinvenenthrombose. Davon bei 41 verifiziert und bei 77 ausgeschlossen (Bildgebung).

Cut-off ($\mu\text{g/l}$)	Sensitivität	Spezifität
150 ng/mL	95	38
200 ng/mL	86	58
300 ng/mL	76	79
500 ng/mL	60	88
1000 ng/mL	36	96

(in %)

Receiver Operator Characteristic (ROC) -Analyse



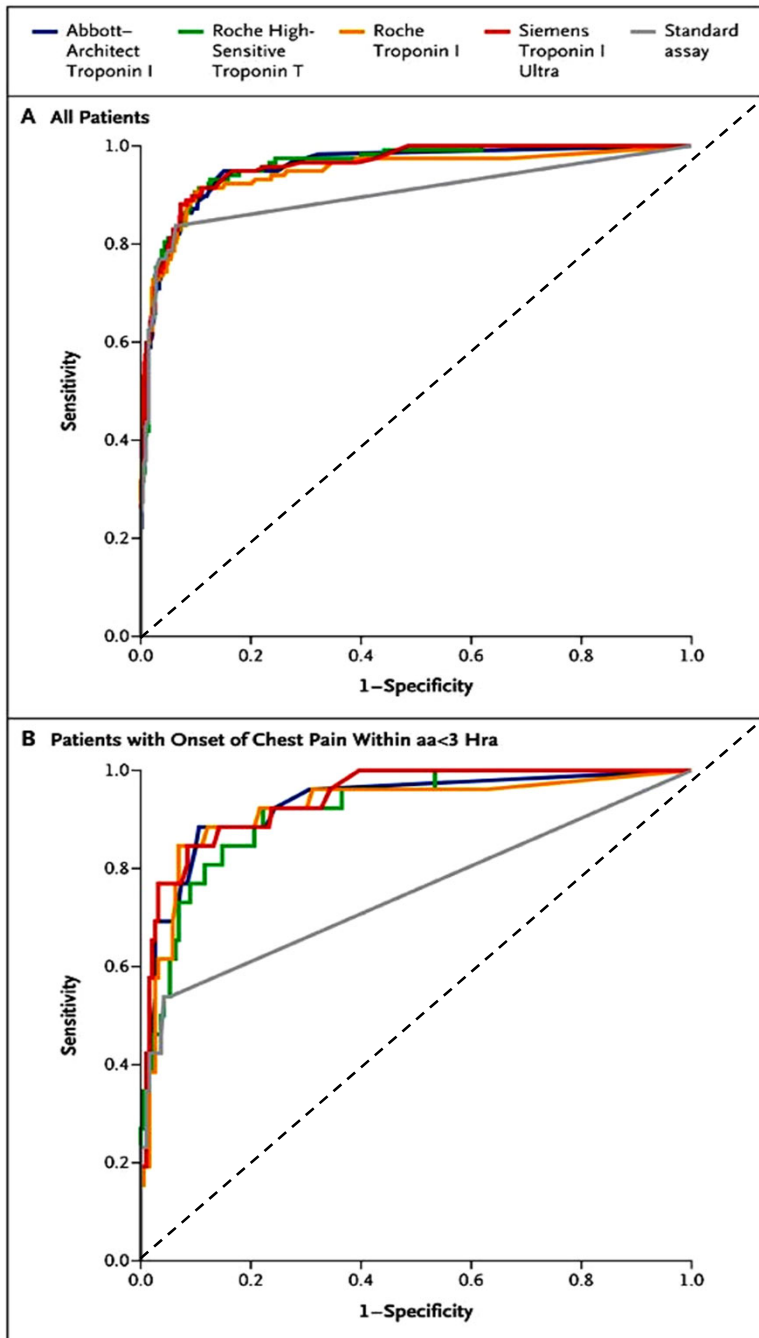
Grundbegriffe der Beurteilung diagnostischer Tests: ROC-Kurve

ROC-Kurve

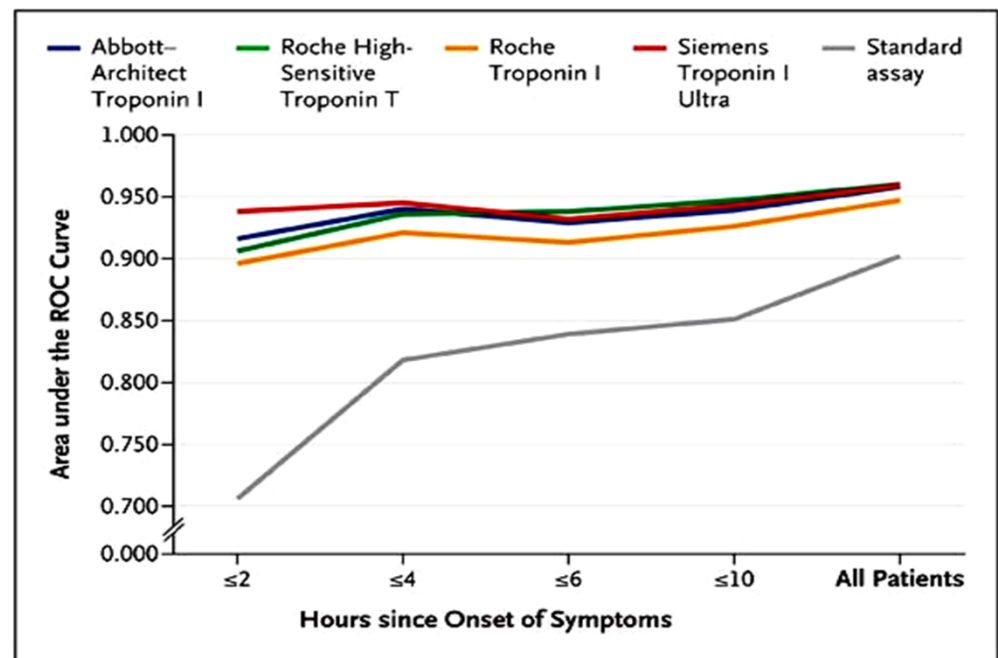
Definition: Vergleich von Sensitivität und Unspezifität
(= 100% - Spezifität)

ermöglicht

- Vergleich der diagnostischen Qualität von mehreren Tests
(Mass: Fläche unter der Kurve = AUC)
 - AUC = 1.00 idealer Test
 - AUC = 0.50 wertloser Test („würfeln“, „raten“)
 - AUC > 0.8 diagnostisch relevante Tests



Receiver-operating-characteristic Kurven Analysen für konventionelle und sensitive Troponin T Assays in Abhängigkeit von der Zeit nach Symptombeginn (718 konsekutive Patienten mit Verdacht auf AMI)



Reichlin et al.; N. Engl. J. Med. 2009; 361: 858-867

Grundbegriffe der Beurteilung diagnostischer Tests

positive Likelihood Ratio (LR+)

Positive Likelihood Ratio (LR +)

Definition: Die Wahrscheinlichkeit einer Person mit der Krankheit, positiv getestet zu werden, relativ zur Wahrscheinlichkeit einer Person ohne diese Krankheit positiv getestet zu werden.

= Verhältnis von Sensitivität zu Unspezifität (100% - Spezifität)

$$\text{Positive Likelihood Ratio (LR+)} = \frac{\text{Anteil der richtig-positiven Testresultate}}{\text{Anteil der falsch-positiven Testresultate}} = \frac{TP/(FN + TP)}{FP/(TN + FP)}$$

$$\text{Positive Likelihood Ratio (LR+)} = \frac{\text{Sensitivität}}{(100\% - \text{Spezifität})}$$

TP = true positive
FP = false positive
TN = true negative
FN = false negative

Grundbegriffe der Beurteilung diagnostischer Tests

negative Likelihood Ratio (LR-)

negative Likelihood Ratio (LR -)

Definition: Die Wahrscheinlichkeit einer Person mit der Krankheit negativ getestet zu werden relativ zur Wahrscheinlichkeit einer Person ohne Krankheit, negativ getestet zu werden

= Verhältnis von Unsensitivität (d.h. 100% - Sensitivität) zu Spezifität

$$\text{Negative Likelihood Ratio (LR-)} = \frac{\text{Anteil der falsch-negativen Testresultate}}{\text{Anteil der richtig-negativen Testresultate}} = \frac{\text{FN}/(\text{FN} + \text{TP})}{\text{TN}/(\text{TN} + \text{FP})}$$

$$\text{Negative Likelihood Ratio (LR-)} = \frac{100\% - \text{Sensitivität}}{\text{Spezifität}}$$

TP = true positive
FP = false positive
TN = true negative
FN = false negative

Interpretation von Likelihood-Ratios

Positive Likelihood Ratio (LR+)	Negative Likelihood Ratio (LR-)	Interpretation
LR+ > 10	LR- < 0.1	„überzeugende diagnostische Evidenz“
LR+ 5 – 10	LR- 0.1 – 0.2	„hohe diagnostische Evidenz“
LR+ 2 – 5	LR- 0.2 – 0.5	„schwache diagnostische Evidenz“
LR+ 1 – 2	LR- 0.5 – 1.0	„keine relevante diagnostische Evidenz“

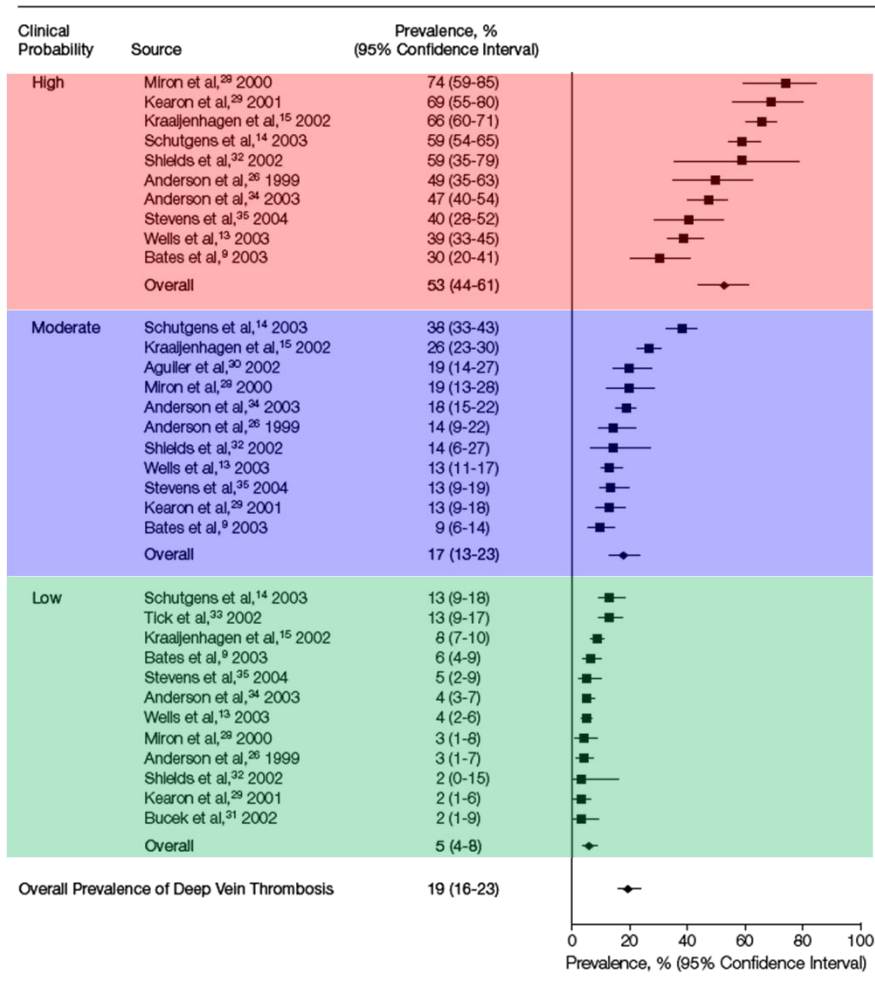
Ein Beispiel zu Likelihood Ratios: D-Dimer bei venöser Thrombose und Embolie

118 Patienten mit Verdacht auf pulmonale Embolie oder tiefer Beinvenenthrombose. Davon bei 41 verifiziert und bei 77 ausgeschlossen (Bildgebung).

Cut-off ($\mu\text{g/l}$)	Sensitivität	Spezifität	LR+	LR-
150 ng/mL	95	38	1.5	0.13
200 ng/mL	86	58	2.0	0.24
300 ng/mL	76	79	3.6	0.30
500 ng/mL	60	88	5.0	0.45
1000 ng/mL	36	96	9.0	0.67

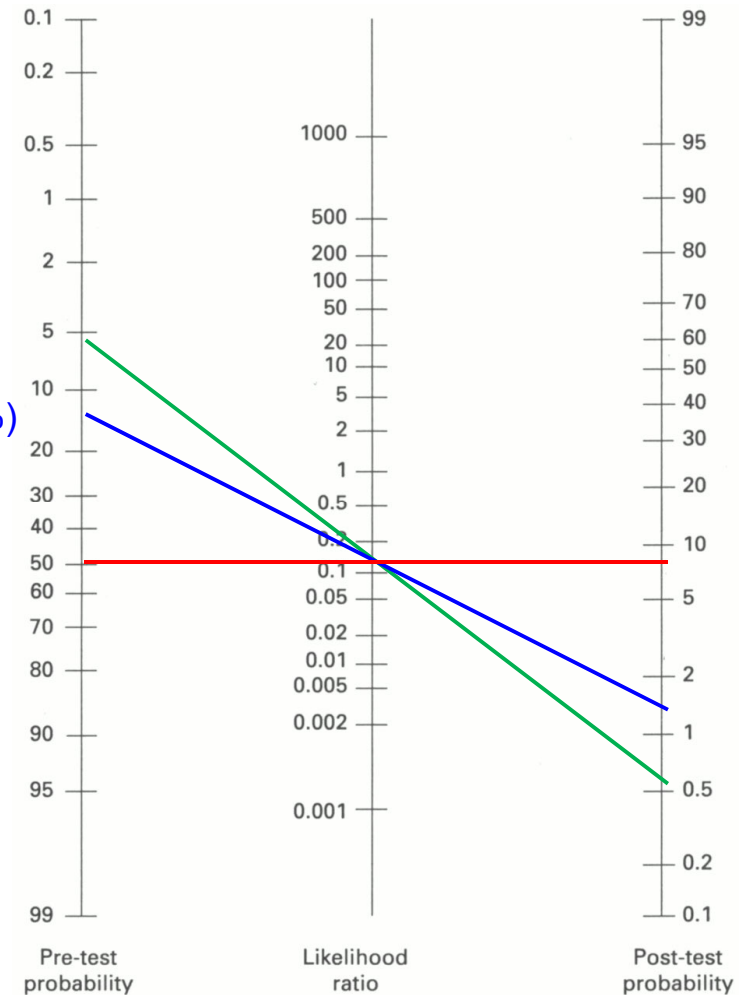
(in %)

Ausschluss venöser Thromboembolien mittels D-dimer bei unterschiedlicher Prävalenz



Sensitivität 90%, Spezifität 60%
negative Likelihood Ratio 0.17

tiefes
Risiko (5%)
mittleres
Risiko (17%)
hohes
Risiko (50%)



Grundbegriffe der Beurteilung diagnostischer Tests

Positiver Prädiktiver Wert

Positiver prädiktiver Wert:

Definition: Wahrscheinlichkeit, mit der bei Vorliegen eines positiven Tests eine bestimmte Erkrankung tatsächlich vorliegt

$$PV_{\text{pos}} = \frac{\text{wahre positive}}{\text{alle positiven}} \cdot 100 = \frac{TP}{TP + FP} \cdot 100$$

(abhängig von Prävalenz !)

TP = true positive
FP = false positive

Grundbegriffe der Beurteilung diagnostischer Tests:

negativer prädiktiver Wert

Negativer prädiktiver Wert

Definition: Wahrscheinlichkeit, mit der bei Vorliegen eines negativen Tests eine bestimmte Erkrankung ausgeschlossen werden kann.

$$PV_{\text{neg}} = \frac{\text{wahre negative}}{\text{alle negativen}} \cdot 100 = \frac{TN}{TN + FN} \cdot 100$$

(abhängig von Prävalenz !)

TN = true negative
FN = false negative

Ein Beispiel:

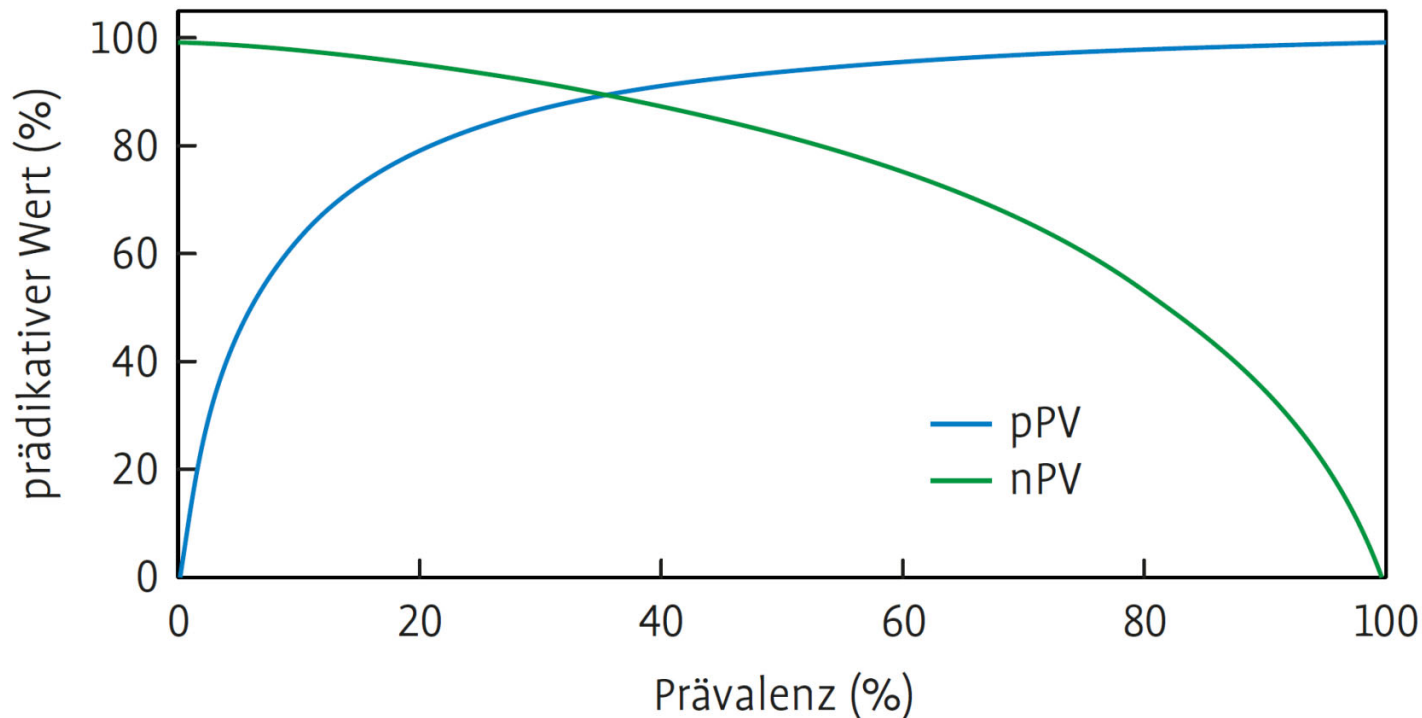
D-Dimer bei venöser Thrombose und Embolie

118 Patienten mit Verdacht auf pulmonale Embolie oder tiefer Beinvenenthrombose. Davon bei 41 verifiziert und bei 77 ausgeschlossen (Bildgebung).

Cut-off ($\mu\text{g/l}$)	Sensitivität	Spezifität	PV _{pos}	PV _{neg.}
150 ng/mL	95	38	45	94
200 ng/mL	86	58	52	88
300 ng/mL	76	79	66	86
500 ng/mL	60	88	73	83
1000 ng/mL	36	96	80	74

(in %)

Abhängigkeit der positiven (pPV) und negativen (nPV) prädiktiven Werte eines Tests von der Prävalenz der gesuchten Diagnose



Berechnung der prädiktiven Werte pPV und nPV für einen Labortest mit 95 % Spezifität und 80 % Sensitivität in Abhängigkeit von der Prävalenz. Die höchste Aussagekraft hat dieser Test, wenn die gesuchte Krankheit etwa bei jedem dritten Patienten vorliegt (Prävalenz 30 bis 40%).

Ein Beispiel: Einfluss der Vortestwahrscheinlichkeit (= Prävalenz) auf die Testqualität von D-Dimer bei venöser Thrombose und Embolie

Vortestwahrscheinlichkeit (= Prävalenz, %)	niedrig 5	mittel 15	hoch 50
Sensitivität, %	90	90	90
Spezifität, %	60	60	60
Positiver prädiktiver Wert, %	10.6	28.4	69.2
Negativer prädiktiver Wert, %	99.1	97.1	85.7

Situation typisch für

ambulante
Medizin

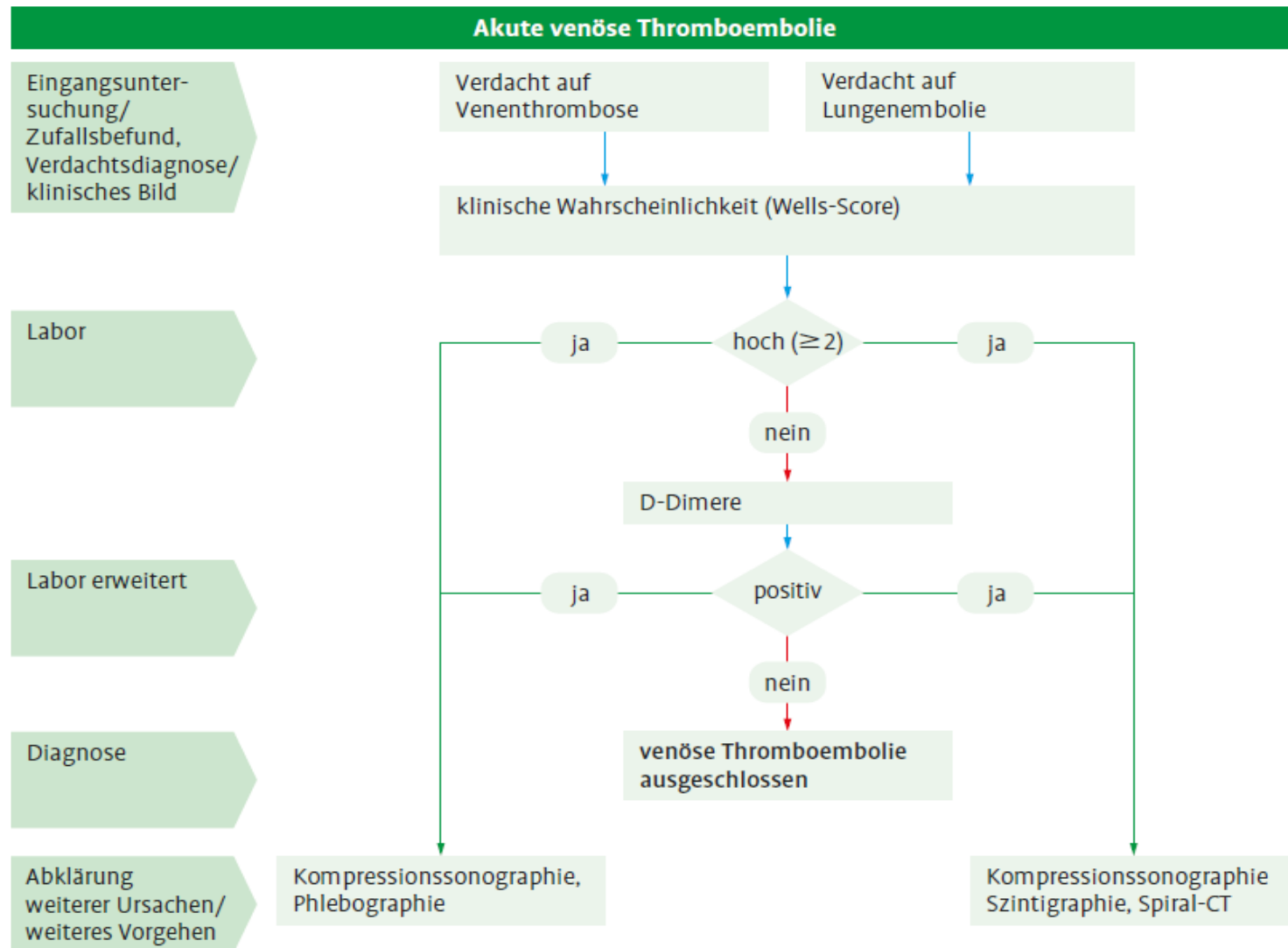
stationäre
Medizin

D-Dimer geeignet für rule-out einer
venösen Thrombose oder Embolie

ja

nein

Algorithmus zur Diagnose einer tiefen Beinvenenthrombose oder Lungenembolie



Vereinfachtes klinisches Modell zur Beurteilung einer tiefen Beinvenenthrombose (DVT)

Klinische Variable	Punkte
- Aktive Krebserkrankung (in Behandlung während der letzten 6 Monate)	1
- Paralyse, Parese oder Immobilisierung	1
- Bettlägerigkeit während der letzten 3 Tage oder grösserer chirurgischer Eingriff in den vergangenen 12 Monaten	1
- Druckschmerz entlang des Verlaufes tiefer Beinvenen	1
- Schwellung des gesamten Beines	1
- Lokale Unterschenkel-Schwellung (> 3 cm als kontralateral)	1
- Hautödeme	1
- Kollateralvenen (keine Varikose)	1
- Frühere tiefe Beinvenenthrombose	1
- Andere Diagnose als DVT mindestens so wahrscheinlich	-2

Wahrscheinlichkeit einer DVT:

hoch: 3 Punkte oder mehr; mittel: 1-2 Punkte; tief: 0 Punkte

Klinisches Modell (Well's Score) zum Ausschluss einer Lungenembolie

Variablen	Punkte
Klinische Befunde und Symptomatik einer TVT (min. Ödem und Druckdolenz im tiefen Leitvenensystem)	3,0
Alternative Diagnose weniger wahrscheinlich als LE	3,0
Herzfrequenz > 100/min	1,5
Immobilisation (> 3 Tage) oder Operation innerhalb der letzten 4 Wochen	1,5
Frühere TVT oder LE	1,5
Hämoptyse	1,0
Aktives Malignom, Chemo- oder Strahlentherapie, < 6 Monate palliativ	1,0
Klinische Vortestwahrscheinlichkeit gering	≤ 4
Klinische Vortestwahrscheinlichkeit hoch	> 4

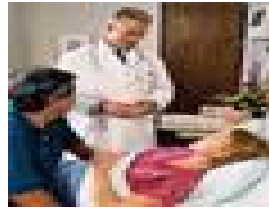
Mögliche Klinische Nutzen neuer Biomarker

Patienten Zufriedenheit



Früherkennung
Schnelle Diagnosen
Richtige Dosierungen
Bessere medizinische Ergebnisse
Bessere Lebensqualität

Klinischer Nutzen



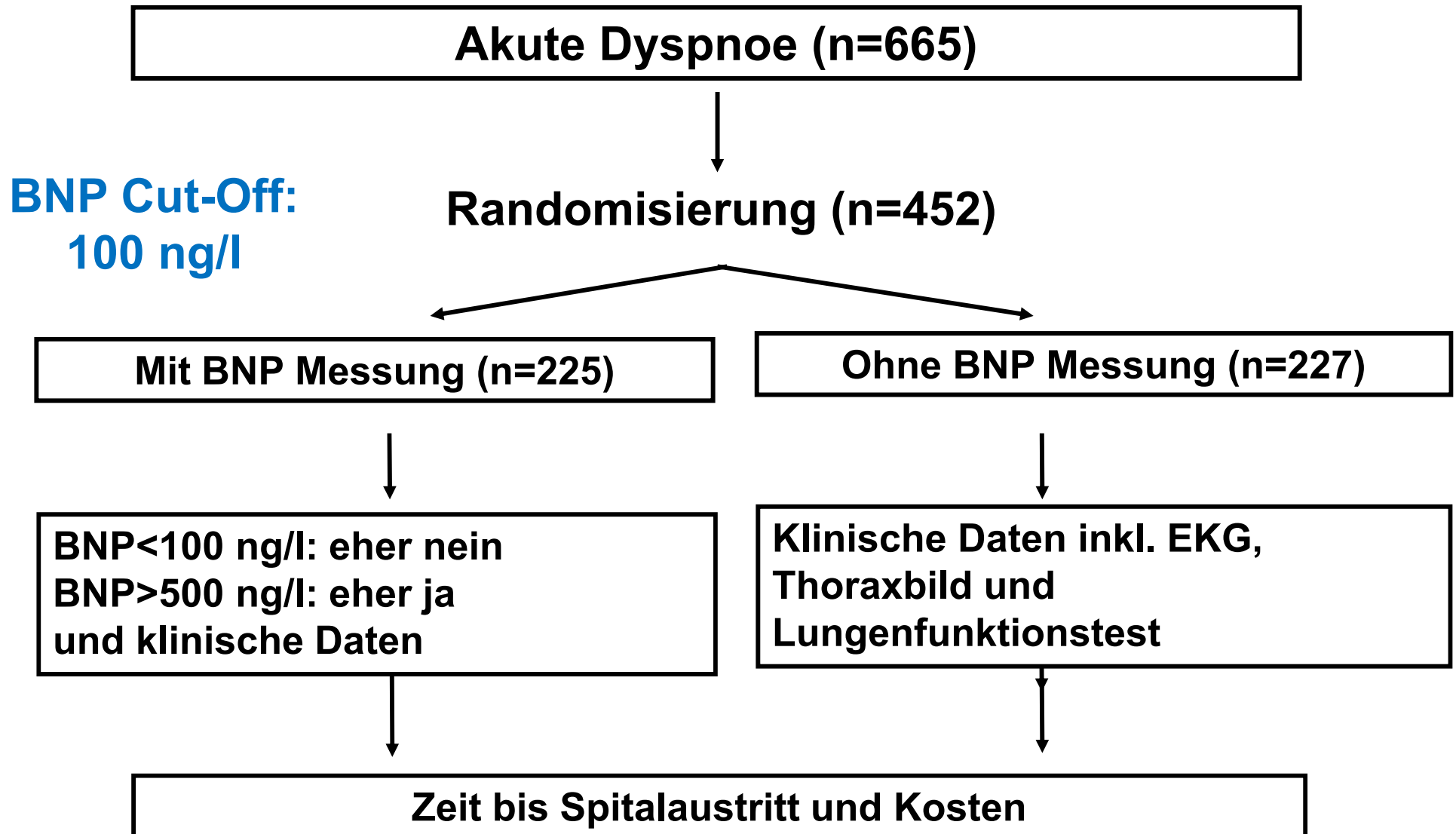
Frühe, bessere Diagnosen
Beurteilung der Effektivität von Therapien
Unterstützung des Krankheits-Managements
Besseres individuelles Gesundheitsmanagement
vermindert die Notwendigkeit späterer medizinischer Interventionen

Ökonomischer Nutzen



Vermeidung unnötiger Behandlungen
Verkürzung der Hospitalisationsdauer
Verminderung der Kosten für Behandlung & Rehabilitation
Verminderung der Gesamtkosten pro Patient

BASEL Studie zum klinischen Nutzen von BNP



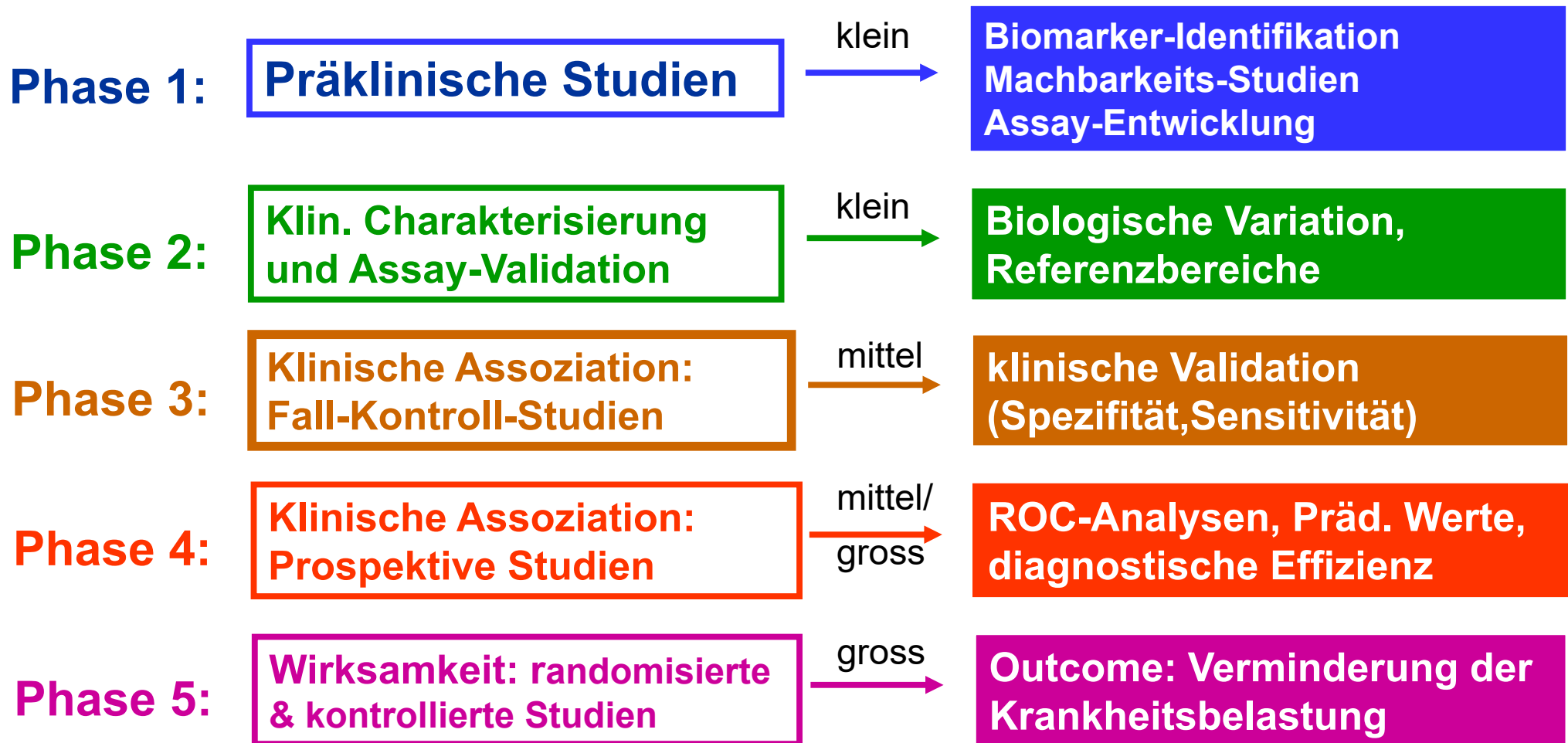
Beispiel

Basel Studie:

(N Engl J Med. 2004 Feb 12;350(7):647-54)

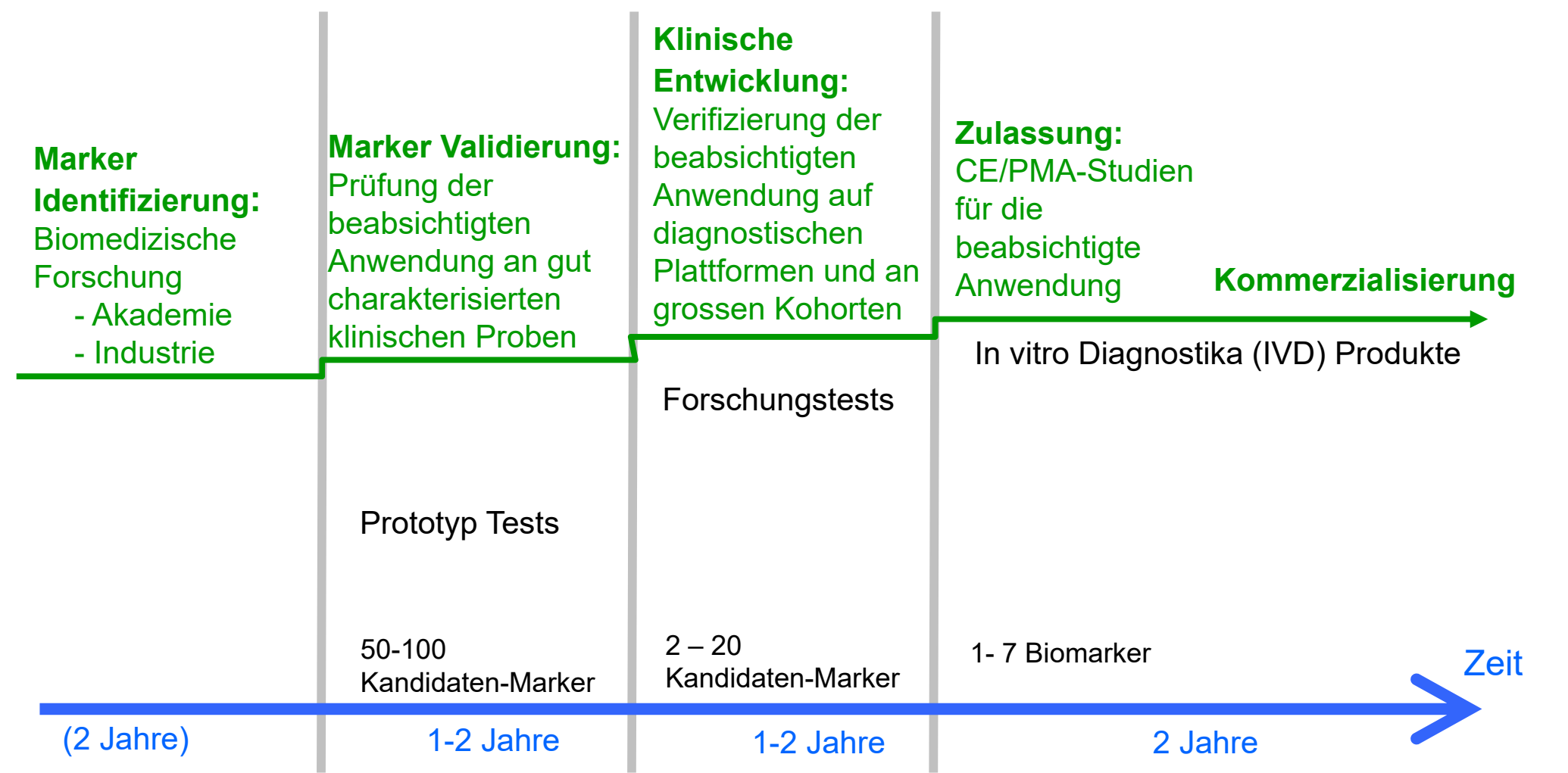
End Point	B-Type Natriuretic Peptide Group (N=225)	Control Group (N=227)	P Value
Time to treatment — min			0.03†
Median	63	90	
Interquartile range	16–153	20–205	
Time to discharge — days			0.001†
Median	8.0	11.0	
Interquartile range	1.0–16.0	5.0–18.0	
Hospitalization — no. (%)	169 (75)	193 (85)	0.008
Admission to intensive care — no. (%)	33 (15)	54 (24)	0.01
Cost of intensive care — \$			0.07
Median	874	1,516	
95% Confidence interval	423–1,324	989–2,043	
Total treatment cost — \$			0.006
Median	5,410	7,264	
95% Confidence interval	4,516–6,304	6,301–8,227	
In-hospital mortality — no. (%)	13 (6)	21 (9)	0.21‡
30-day mortality — no. (%)	22 (10)	28 (12)	0.45‡
30-day readmission rate — no. (%)	26 (12)	23 (10)	0.63

Phasen in der Entwicklung von diagnostischen Biomarkern



Grösse der untersuchten Population

Zeitablauf und (Miss)erfolgsrate in der Entwicklung diagnostischer Biomarker



Biomarker mit hohem negativ-prädiktivem Wert (PV_{neg}) und hoher Sensitivität sind gut für Ausschlussdiagnostik

$$PV_{neg} = \frac{\text{wahre negative}}{\text{alle negativen}} \cdot 100 = \frac{TN}{TN + FN} \cdot 100$$

PV_{neg} umso grösser

Je grösser der Anteil TN

Je kleiner der Anteil FN

TN = true negative
FN = false negative

$$\text{Sensitivität} = \frac{\text{Zahl der echt positiven}}{\text{Zahl der Kranken}} = \frac{TP}{TP + FN}$$

TP = true positive
FN = false negative

Grundbegriffe der Beurteilung diagnostischer Tests

$$\text{Sensitivität} = \frac{\text{Zahl der echt positiven}}{\text{Zahl der Kranken}} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{Spezifität} = \frac{\text{Zahl der echt negativen}}{\text{Zahl der Gesunden}} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

$$\text{Positive Likelihood Ratio} = \frac{\text{Sensitivität}}{(100\% - \text{Spezifität})} = \frac{\text{TP}/(\text{FN} + \text{TP})}{\text{FP}/(\text{TN} + \text{FP})}$$

$$\text{Negative Likelihood Ratio} = \frac{(100\% - \text{Sensitivität})}{(\text{Spezifität})} = \frac{\text{FN}/(\text{FN} + \text{TP})}{\text{TN}/(\text{TN} + \text{FP})}$$

$$\text{PV}_{\text{neg}} = \frac{\text{Zahl echt positive}}{\text{Zahl alle positive}} \cdot 100 = \frac{\text{TP}}{\text{TP} + \text{FP}} \cdot 100$$

$$\text{PV}_{\text{neg}} = \frac{\text{Zahl echt negative}}{\text{Zahl alle negative}} \cdot 100 = \frac{\text{TN}}{\text{TN} + \text{FN}} \cdot 100$$

TP = true positive
FP = false positive
TN = true negative
FN = false negative

Zusammenfassung: Sensitivität/Spezifität und prädiktive Werte

	krank	gesund	
positiv	TP	FP	PV_{pos}
negativ	FN	TN	PV_{neg}
	Sensitivität	Spezifität	

Rechenbeispiele: Prävalenz, Sensitivität, Spezifität, Likelihood Ratios, Vorhersagewerte

krank? \ Test	ja	nein	gesamt
positiv	33 (TP)	4 (FP)	37
negativ	37 (FN)	39 (TN)	76
gesamt	70	43	113 (alle)

Begriff	Definition	Beispiel
Prävalenz	$(TP + FN) / \text{alle}$	$(33 + 37) / 113 * 100\% = 62\%$
Sensitivität	$TP / (TP + FN)$	$33 / (33 + 37) * 100\% = 47\%$
Spezifität	$TN / (TN + FP)$	$39 / (39 + 4) * 100\% = 91\%$
pos. Likelihood Ratio	$\text{Sens.} / (100\% - \text{Spez.})$	$47\% / (100\% - 9\%) = 5.22$
neg. Likelihood Ratio	$(100\% - \text{Sens}) / \text{Spez.}$	$(100\% - 47\%) / 91\% = 0.58$
pos. Vorhersagewert	$TP / (TP + FP)$	$33 / (33 + 4) * 100\% = 89\%$
neg. Vorhersagewert	$TN / (TN + FN)$	$39 / (39 + 37) * 100\% = 51\%$